Context Aware Data Reduction for Highly Automated Driving

Vom Fachbereich Maschinenbau an der
Technischen Universität Darmstadt
zur Erlangung des Grades eines
Doktor-Ingenieurs (Dr.-Ing.)
genehmigte

Dissertation

vorgelegt von

Kai Hendrik Storms, M. Sc. (geb. Domhardt)

aus Kaiserslautern

Berichterstatter:	Prof. Dr. -Ing. Steven Peters
Mitberichterstatter:	Prof. Dr. -Ing. Markus Lienkamp

Tag der Einreichung:	04.12.2023
Tag der mündlichen Prüfung:	20.02.2024

Darmstadt 2024

D 17

Preface

> ”Quidquid agis,
> prudenter agas
> et respice finem.”
>
> *Gesta Romanorum*

This dissertation was written during my time working as a research assistant at the Institute of Automotive Engineering (FZD) of the Technical University of Darmstadt. My work was enabled in terms of scientific input and funding by the Verification & Validation Methods (VVM) project. Therefore, I would like to thank the BMWK, without whose funding the project and this work would not have been possible. Additionally, I would also like to thank all the project staff and partners of the VVM consortium who contributed through many insightful discussions.

Special thanks go to my doctoral supervisor Prof. Dr.-Ing Steven Peters for the consultations and the scientific advice. The trust placed in me, especially considering project work, greatly enabled me to learn and grow in my role.

I would also like to thank Prof. Dr.-Ing. Markus Lienkamp for taking over the role as second examiner and also for taking the time to exchange as well as discuss ideas regarding this thesis.

I would also like to thank all my colleagues who always had an open ear and time for a beer when things did not go so well or when there was something to celebrate. The feedback I received from the semi-annual doctoral seminars helped me overcome multiple methodological obstacles.

A big thank you also goes to my friends and family, who have given me the opportunity to leave work behind and ground myself anew.

Finally, but most importantly, I want to thank my wife Sarah. Who was always there for me and encouraged me to start this journey in the first place.

Darmstadt, December 2023.

Table of Contents

List of Symbols and Indices

Latin formula symbols

Symbol	Unit	Description
$\boldsymbol{A}$		Alpha (opacity) matrix
$\boldsymbol{C}$		Class matrix (segmentation map)
$\boldsymbol{I}$		Image matrix
$\boldsymbol{M}$		Mask matrix
$\boldsymbol{W}$		Noise matrix
$\boldsymbol{K}$		Kernel matrix
Y		Model output quantities
v		Threshold value
b	bit	Number of bits
c		Class
i		Iteration variable
j		Iteration variable
f		Function
p		Performance
$\boldsymbol{e}$		Unit vector
x		X coordinate
y		X coordinate
w		Width
h		Height
N		Max count
n		Count variable
MAX		Maximum possible value
a	$\frac{m}{s^2}$	Acceleration
t	s	Time
s	m	Size
v	$\frac{m}{s}$	Velocity
d	m	Distance
r	m	Position vector

Greek formula symbols

Symbol	Unit	Description
Φ		Model parameters
α		Significance level
ϵ		Error

Symbol	Unit	Description
τ	s	Duration
ϕ		Offset angle
μ		Mean
σ		Standard deviation

Calligraphic symbols and fraktur characters

Symbol	Unit	Description
$\mathbb{D}$	-	Set of Data; Dataset
$\mathbb{C}$	-	Set of object classes
$\mathcal{N}$	-	Normal distribution

Indices

Symbol	Description
$\parallel$	Parallel
$\perp$	Perpendicular
0	Initial
1	Ego
2	Object of interest
3	Static object
a	Acceleration
assurance	With respect to the assurance
b	Braking
Blackened	Blackened dataset
border	With respect to the (image) border
c	Change
C	Class
class	With the respect to the object class
corr	Corrected
d	Desired/adequate speed
dilated	Dilated
e	End
expansion	With respect to the expansion
Full	Full dataset
g	Guaranteed
G	Generator
grey	Greyscale
horz	Horizontal
irrelevant	Irrelevant
l	Lateral
lanemarking	With respect to lane markings
m	Merge
masked	Masked with respect to the relevance
max	Maximum
min	Minimum
Mixed_Synth	Mixed synth dataset
noise	White noise
O	Object

Symbol	Description
pixel	In pixel coordinates
r	Reaction
r	Radial
reduction	With respect to the reduction
relevant	Relevant
residual	Residual
road	With respect to the road
s	Start
selection	With respect to the selection
Synth	Synthetic dataset
system	With respect to the system
T	Training
test	Test split
thresh	Threshold
train	train split
Truncated	Truncated dataset
val	Val split
vert	Vertical

Accents and Operators

Symbol	Description
$\neg\square$	Negated
$\tilde{\square}$	Estimated
$\hat{\square}$	Expanded
$\check{\square}$	Reduced
$\overline{\square}$	Arithmetic mean
$\odot$	Element-wise product: $C_{i,j} = A_{i,j} \cdot B_{i,j}$
$\oslash$	Element-wise division: $C_{i,j} = A_{i,j} / B_{i,j}$
$\square'$	Rereferenced
$\square^*$	Special considerations apply
$\cup$	Union of sets
$\cap$	Intersection of sets
$\in$	Element of

Symbol	Description
$\vee$	Logical or
$\wedge$	Logical and

List of Abbreviations

A-Box	assertion box
AD	automated driving
ADAS	advanced driver assistance systems
ADE	average displacement error
ADS	automated driving system
AVC	Advanced Video Coding
BEV	birds-eye-view
CADR	Context Aware Data Reduction
DDT	dynamic driving task
DIKW	data-information-knowledge-wisdom
DNN	deep neural network
eCDF	empirical cumulative distribution function
ECU	electronic control unit
FN	false negative
FP	false positive
GAN	generative adversarial network
GRM	general relevance model
HEVC	High Efficiency Video Coding
IoU	intersection over union
IQA	image quality assessment
ITU	International Telecommunication Union
mAP	mean average precision
$\mathrm{mAP}_{.5}$	mean average precision at IoU=0.5
$\mathrm{mAP}_{.5:.95}$	mean average precision at IoU $\in$ [0.5:0.05:0.95]
mIoU	mean intersection over union
MS-SSIM	multi-scale structural similarity metric
MSE	mean square error
OD	operating domain
ODD	operating design domain
OOI	object of interest
pDI	physical driving instance
PKL	planning Kullback-Leibler divergence
PNG	portable network graphics

PSNR	peak signal-to-noise ratio
RoI	region of interest
RQ	research question
RSS	Responsibility-Sensitive Safety
SAE	Society of Automotive Engineers
SRM	specific relevance model
SSIM	structural similarity metric
SYS-REQ	system requirement
T-Box	terminology box
TP	true positive
V2X	vehicle-to-everything
vAVI	virtual automated vehicle instance

List of Figures

List of Tables

Kurzzusammenfassung

Diese Dissertation befasst sich mit der entstehenden Herausforderung der Datenverarbeitung bei der Entwicklung des automatisierten Fahrens.

Die zunehmenden Datenmengen, die bei der Entwicklung und dem Betrieb dieser Systeme anfallen, erfordern effiziente Handhabungsstrategien. Ein umfassender Überblick über den aktuellen Stand der Technik zeigt, dass die Datenreduktion eine praktikable Lösung für die Bewältigung dieser großen Datenmengen darstellt. Darüber hinaus zeigen sich, durch eine Analyse des Standes der Technik, verschiedene Herausforderungen, bei denen etablierte Methoden unzureichend sind, wie z.B. der Umgang mit dem offenen Kontext. Daraus ergibt sich die primäre Forschungsfrage dieser Dissertation:
(1) Wie kann eine Datenreduktion umgesetzt werden, um diese Herausforderungen zu bewältigen?

Die in dieser Arbeit vorgeschlagene Lösung ist ein neuartiger Ansatz zur Datenreduktion, der das Konzept der Relevanz integriert und darauf abzielt, den Informationsbedarf in den Daten genau zu definieren und zu deklarieren. Dieser Ansatz ermöglicht eine bessere Kontrolle über den Leistungsverlust in späteren Anwendungsfällen und führt zu zwei sekundären Forschungsfragen:

(2) Wie kann Relevanz formal definiert werden, um ihre Verwendung bei der Datenreduktion zu ermöglichen?

(3) Wie wirkt sich die neuentwickelte Methode der Datenreduktion auf die Leistung der nachfolgenden Anwendungsfälle aus?

Zur Beantwortung der zweiten Frage wird das Konzept der Relevanz in der Literatur intensiv untersucht. Es wird ein allgemeines Relevanzmodell für das automatisierte Fahren, sowie eine Methodik zur Ableitung und Validierung von anwendungsfallspezifischen Relevanzmodellen entwickelt. Die Anwendung dieser Methodik auf einen ausgewählten Anwendungsfall demonstriert ihre Wirksamkeit.

Zur Beantwortung der dritten Frage wird eine Architektur für relevanzgesteuerte Datenreduktion vorgeschlagen. Ein Prototyp, der diese Architektur implementiert, wird im Zusammenhang mit der Wahrnehmung und dem Training neuronaler Netze evaluiert, wobei der Schwerpunkt auf semantischen Segmentierungs- und Objekterkennungsaufgaben liegt. Die Ergebnisse zeigen, dass eine relevanzgesteuerte Datenreduktion den Leistungsverlust bei Wahrnehmungsaufgaben wirksam kontrollieren kann. Beim Training neuronaler Netze ist jedoch eine starke Aufgabenabhängigkeit zu beobachten, wodurch die Grenzen des Ansatzes und Möglichkeiten für zukünftige Forschung aufgezeigt werden.

Zusammenfassend stellt diese Arbeit einen Beitrag zu zwei Bereichen dar:
Erstens werden Methoden zur Bewältigung der Herausforderungen bei der Verarbeitung großer Mengen von Fahrzeugdaten und der Reduktion dieser Daten auf die Anteile mit relevanten Informationen entwickelt.
Zweitens werden Ansätze zur expliziten Berücksichtigung der großen Vielfalt von Relevanzkonzepten bei der Entwicklung des automatisierten Fahrens aufgezeigt.

Abstract

This research addresses the emerging challenge of data handling in the development of automated driving systems. The increasing volumes of data generated in the development and operation of these systems necessitate efficient handling strategies. A comprehensive review of the current state of the art reveals data reduction as a viable solution to manage these large data volumes. Further, an analysis of the state of the art establishes various challenges where established methodologies are insufficient, such as the open context. This leads to the primary research question of this dissertation:

(1) how to effectively implement data reduction while addressing these challenges.

The proposed solution in this work is a novel data reduction approach that integrates the concept of relevance, aiming to precisely define the informational needs within the data. This approach hypothesizes an enhanced control over performance loss in subsequent use cases, leading to two secondary research questions:

(2) How can relevance be formally defined to facilitate its use in data reduction?

(3) What is the impact of this data reduction method on the performance of subsequent use cases?

To answer the second question, the concept of relevance is extensively explored in the literature. A general relevance model for automated driving is developed, along with a methodology for deriving and validating use case-specific relevance models. Application of this methodology to a selected use case demonstrates its effectiveness.

Addressing the third question, an architecture for relevance-guided data reduction is proposed. A prototype implementing this architecture is evaluated in the contexts of perception and neural network training, focusing on semantic segmentation and object detection tasks. The findings indicate that relevance-guided data reduction can effectively control performance loss in perception tasks. However, in neural network training, a strong task dependency is observed, highlighting limitations of the approach and opportunities for future research.

In conclusion, this work represents a contribution in two areas. First, to overcoming the challenges of handling large amounts of automotive data and reducing this data to only those parts with relevant information. Second, to an explicit consideration of the wide variety of relevance concepts in the development of automated driving.

1 Introduction

This chapter will establish the basis of this thesis. First, the initial motivation for the research area is presented. From this, the need for the research topic, with its initial situation and research question (RQ)s, is derived. As a guide for the following chapters, the methodology for answering the RQs within the structure of the whole thesis is presented.

1.1 Motivation

Automated driving (AD) promises to bring about various societal benefits with its introduction.[1,2] Increased safety within the traffic environment is often cited as the primary positive effect associated with AD. The expectation of increased road safety is especially manifested in the Vision Zero[4] of the European Union and Vision for Safety[5] of the National Highway Traffic Safety Administration of the United States. Further, AD is believed to enable major shifts in how and especially where we live, changing the range of available mobility options. The magnitude and variety of benefits are expected to increase with the increasing level of AD of released vehicles.[1]

A taxonomy for levels of automation is defined by the Society of Automotive Engineers (SAE) in the standard J3016[3]. The standard gives a visual representation as shown by figure 1-1 ranging from level 0 to level 5. Level 0 indicates no automation of the driving task. Levels 1 and 2, called advanced driver assistance systems (ADAS), identify automation that assists the driver in the driving task, while the driver remains responsible for the safe operation. Level 3 is the first level of AD, where the automated driving system (ADS) can temporarily take both control and responsibility for the driving task. The driver remains a fallback layer if the system reaches the limits of its operational capabilities. This is also the distinguishing factor to level 4, as here the ADS cannot rely on the driver to take over to reach a safe state. The operation of level 4 ADS is limited to certain environments, if the operation is possible *”everywhere in all conditions”*[3], it is denoted as level 5.

With higher levels of AD, it is not only the way we drive now that will be improved. More so, new concepts for mobility are emerging and changing the way we drive in a fundamental way.[2]

1 Szimba, E.; Hartmann, M.: Assessing travel time savings and user benefits of automated driving (2020)

2 Lenz, B.; Fraedrich, E.: New Mobility Concepts and Autonomous Driving: The Potential for Change (2016)

3 Shi, E. et al.: The Principles of Operation Framework (2020)

4 Tingvall, C.; Haworth, N.: Vision Zero: An ethical approach to safety and mobility (1999).

5 U.S. Department of Transportation NHTSA: Automated Driving Systems: A Vision for Safety (2017).

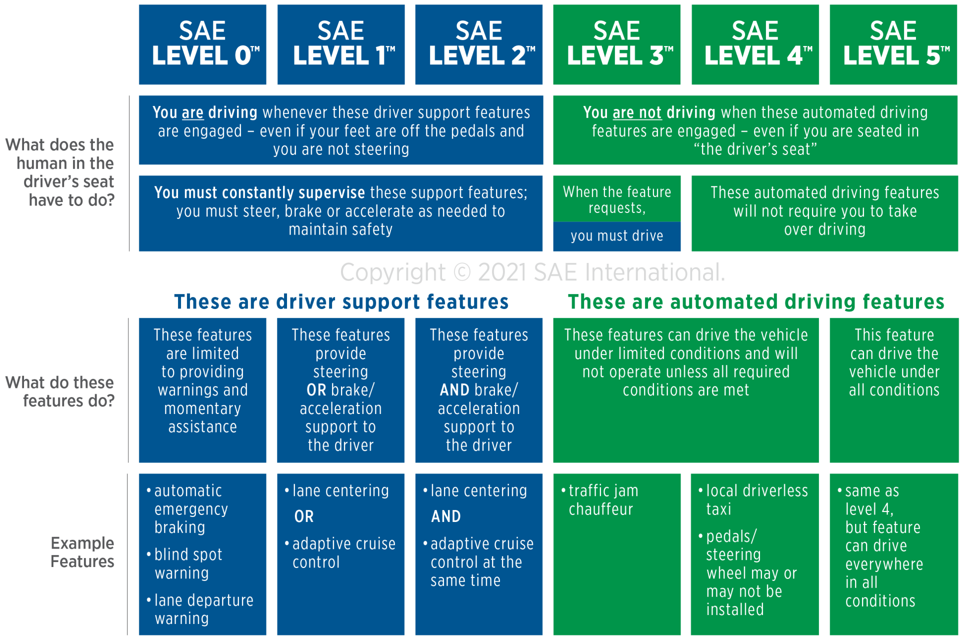

Figure 1-1: Overview of levels of automated driving according to SAE[3]

Current systems up to level 2 have increased the safety and comfort of operating a vehicle.[6] Systems of level 3 and especially level 4 allow for completely new modes of operation[2]. ADS equipped vehicles enable individual mobility for groups, that were previously reliant on other people, such as children, the elderly, and people without driver's licenses.[7] Disruption through car sharing enables a substantial change in the necessity to own and provision space for individual cars.[2] Accordingly, AD not only promises to improve our current traffic environment but also opens the way to a new understanding of mobility that can fundamentally change the way we live.

1.2 Initial Situation and Research Questions

In order to leverage these benefits, ADS realizing high levels of AD are required. Currently, ADS are only publicly available at lower levels of automation, therefore many benefits are not yet available. The development and release of the corresponding levels remains a challenge to this day. Not only the operation of ADS but also both development and release are inherently dependent on data.

Considering the operation of ADS, Kraus, Ivanov, and Leitgeb[8] state that the required bandwidth inside the car for AD is going to increase by a factor of three to four when comparing SAE level 2 to SAE level 3 and higher systems. Further, current development of AD, especially through the

[6] Horn Martin; Watzenig, D.: Automated Driving: Safer and More Efficient Future Driving (2017).

[7] Cyganski, R.: Automated Vehicles and Automated Driving from a Demand Modeling Perspective (2016), p. 240.

[8] Kraus, D. et al.: Approach for an Optical Network Design for Autonomous Vehicles (2019).

use of neural networks, necessitates large amounts of data.[9,10] It stands to reason that increasing levels of AD will also necessitate increasing amounts of data. Feng et. al[11] has shown an increase in size of publicly available datasets between 2014 and 2019 of two orders of magnitude.

Considering the aspect of creating a safety argumentation for the release of AD, there is currently no established standard method to provide such an argumentation. Wachenfeld and Winner[12] present an example stochastic safety proof for a highway pilot ADS on German highways. They assume that a reduction of fatal accidents to half the original number constitutes acceptable safety and require a 95 % confidence threshold for the reduction. Under these assumptions, using the Poisson distribution as the underlying model, Wachenfeld and Winner show that to prove acceptable safety for a highway pilot 6.62 billion kilometers driven with no fatality are necessary. To convert this distance requirement into a data requirement, a few assumptions have to be made. Using the estimated range of data rates for level 4 and 5 ADS of 4.2 TB/h to 76 TB/h[13] and assuming an average speed of 130 km/h, it yields the requirement to handle the 0.2 to 3.9 Zetabytes of data produced. The white paper[14] of the International Data Corporation and Seagate places the amount of data created in 2023 in the realm of 100 zetabytes, which means that a singular safety proof equates to up to 4% of the yearly worldwide data creation. While this example on the proof shows the necessity for more feasible safety assurance methods, and it also exemplifies the need for data reduction methods to handle the inevitably large amounts of data associated with a release of AD.

Further, it can be assumed, that with the increasing complexity of the operating environment ADS are designed for a respective increase of the data requirement will be required. A key aspect of the challenge of releasing AD is its open context. Rueß and Burton define an open context as *"an environment that cannot be fully specified in a way that desirable system behavior can be defined for each possible set of conditions "*[15]. While the open context is mostly known as a hurdle for safety argumentation, it also poses a challenge for data reduction. Through the inherent lack of specification in the open context, it is impossible to a-priori fully specify a data reduction for an ADS. Since an a-priori specification is unavailable, the working specification has to be created and adjusted while it is being used. Such a lack and imprecision of a specification hinders the application of data reduction, due to the risk of removing potentially relevant information. A data reduction method that can take into account the growing specification of the open context is needed. As it will later be shown, an explicit consideration of relevance enables this capability for data reduction.

[9] Klitzke, L. et al.: Vehicle Data Management System for Validation of ADS (2019).

[10] Houben, S. et al.: Inspect, Understand, Overcome: A Survey of Practical Methods for AI Safety (2022).

[11] Di Feng et al.: Deep Multi-modal Object Detection and Semantic Segmentation (2021).

[12] Wachenfeld, W. H. K.; Winner, H.: The Release of Autonomous Vehicles (2016).

[13] Kraus, D. et al.: Approach for an Optical Network Design for Autonomous Vehicles (2019).

[14] David Reinsel et al.: The Digitization of the World: From Edge to Core (2018).

[15] Rueß, H.; Burton, S.: Whitepaper: Safe AI. How is this possible? (2022), p.8.

Knowing the motivation behind, the initial situation and the challenges concerning the release of AD, this work has the objective of outlining the means to address the challenges of data reduction described above. This objective directly yields the first RQ:

> **Research Question 1**
> *How can the open context present in automated driving be addressed in data reduction?*

As mentioned above, the concept of relevance, which will be expaned on in the subsequent chapters, is key to solving the open context for data reduction. Due to the lack of current understanding about the concept of relevance, a second RQ emerges:

> **Research Question 2**
> *How can the concept of relevance, in the context of a given use case, be formally described?*

Given an answer to RQ1, an additional RQ presents itself. Since RQ1 only addresses if there is a method to address the open context and how it might look like, the next question towards that answer is how well does it perform? The answer for this question can only be given in relation to a given use case, for which the data reduction is designed. As such, a third RQ is formulated as follows:

> **Research Question 3**
> *What is the impact on the use case of relevance-driven data reduction?*

1.3 Methodology & Structure of the Dissertation

In order to answer the presented RQs, this thesis will be structured in the following way, as displayed in figure 1-2.

Chapter 1 presents the starting point for this work. The general motivation of why AD is a technology with societal benefit is introduced. Further, the current situation at the time of writing is laid out and a knowledge gap between the current situation and target situation is identified. To close the knowledge gap, RQs are derived.

Chapters 2 and 3 establish the connection to the current scientific corpus and lead to a starting point for the following chapters. Chapter 2 introduces some basics from previous work that are considered prerequisite knowledge for further understanding of the dissertation and which are well established in the current state of the art. Chapter 3 explores related works and delineates the current extend of understanding within the state of the art. The related work is divided into three parts, the challenges of AD itself, data reduction as the topic of this dissertation, and relevance which, as this work will show, has a crucial role in solving the challenges ahead.

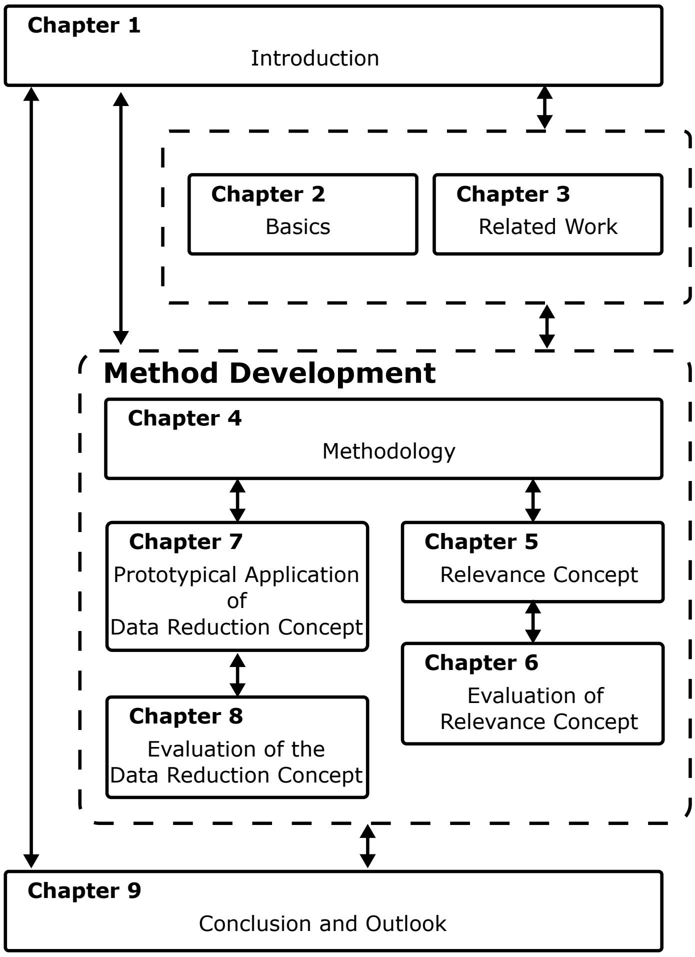

Figure 1-2: Structure of the Dissertation

Chapter 4 to 8 form the core of the dissertation, where the RQs are answered. In chapter 4 a methodology for data reduction is developed to give a possible answer to RQ. 1. As previously mentioned, the answer to RQ. 1 leads directly into the second RQ. Accordingly, chapter 5 opens a new research thread, focusing on RQ. 2. Within this chapter, an answer to RQ. 2 in the form of a relevance concept is presented, based on the findings in 3 on relevance. Subsequently, in chapter 6 an experiment for the applicability of the relevance concept is performed.

Having answered RQ. 2, chapter 7 and 8 then present experiments to explore the applicability of the overall data reduction concept. Chapter 7 implements a data reduction prototype suitable for the evaluation of the experiments. The experiments are conducted in chapter 8 while considering the development and usage for perception as the target use case. Specifically, the task of semantic segmentation and object detection are considered. Lastly, chapter 8 quantifies the effects for the aforementioned tasks and gives an answer to RQ. 3.

The final chapter, chapter 9 summarizes the progress of research performed by this dissertation and discusses the results. Finally, an outlook for future development and research directions is presented.

2 Basics

Several concepts are considered essential to this dissertation. While each is considered well established in the current scientific corpus, this chapter serves as a cursory introduction to these concepts and their termninology.

2.1 Preliminaries and Terminology

As a starting point, some basic terminology and preliminaries are introduced, which will be used throughout the dissertation.

Scene, Situation, and Scenario

In order to correctly and accurately describe the sequence of events of a driving mission, a taxonomy of the respective entities is necessary. In the following, the definitions of Ulbrich et al.[16] of the concepts *"Scene"*, *"Situation"* and *"Scenario"* will be used. In addition to the more holistic definitions below, these can be summarized as follows: A scene is a view of any measurable entities in one's surrounding environment. A situation is a scene with the added information of a task, represented by goals and values. Temporal aggregation of subsequent scenes yields a scenario.

Scene: *"A scene describes a snapshot of the environment including the scenery and dynamic elements, as well as all actors' and observers' self-representations and the relationships among those entities. Only a scene representation in a simulated world can be all-encompassing (objective scene, ground truth). In the real world, it is incomplete, incorrect, uncertain, and from one or several observers' points of view (subjective scene)"*[16a].

Situation: *"A situation is the entirety of circumstances, which are to be considered for the selection of an appropriate behavior pattern at a particular point of time. It entails all relevant conditions, options and determinants for behavior. A situation is derived from the scene by an information selection and augmentation process based on transient (e.g. mission-specific) as well as permanent goals and values. Hence, a situation is always subjective by representing an element's point of view."*[16b]

Scenario: *"A scenario describes the temporal development between several scenes in a sequence of scenes. Every scenario starts with an initial scene. Actions & events as well as goals & values*

[16] Ulbrich, S. et al.: Defining and Substantiating the Terms Scene, Situation, and Scenario (2015)[a:p.3, b:p.5, c:p.6]

may be specified to characterize this temporal development in a scenario. Other than a scene, a scenario spans a certain amount of time."[16c]

Indirection of Reference

Description is a central concept in this dissertation. In order to be able to talk about description in general, several terms surrounding the concept need to be introduced.

As the most general term, this work will follow the definition of *"entity"* as presented by Arp, Smith, and Spear, who propose that an *"entity"* is *"[a]nything that exits, including objects, processes and qualities"*.[19]

Based on thoughts from the philosopher Gottlob Frege, Ogden[17] introduced the semiotic triangle as depicted in figure 2-1. The semiotic triangle establishes three essential entities involved in describing and communicating arbitrary ideas and concepts. The first concept is the *"thing"* or tangible entity being described, the second concept is the abstract concept of the tangible entity, the final concept introduced is the entity, which evokes an abstract concept of the tangible entity. Within the terminology introduced by Ogden, the tangible entity is denoted as *"referent"*, its abstract concept is called the *"reference"*, finally, an entity evoking an abstract concept is identified by the term *"symbol"*.

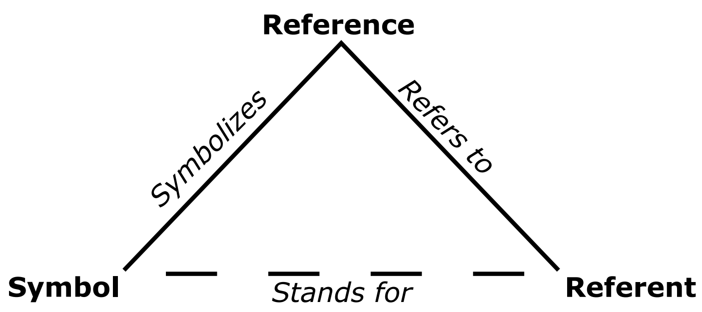

Figure 2-1: The Semiotic Triangle. (Figure recreated from Ogden [17] and Wright[18])

The following applies these abstract terms to a concrete example to demonstrate its purpose and meaning. Suppose the *"referent"* is a specific car with its physical manifestation of several material parts. The reader might use the last car they drove in or the one they own. In this example, the *"reference"* is the internal representation of said car in the mind of the reader. This is an entirely intangible entity. To finish this example, the aspects of what constitutes a *"symbol"* will be considered. Here, the specific symbol used was the sentence prompting the reader to imagine their car, or, more specifically, the string of characters *"car"*. Ogden[17] notes the fact that symbol and referent do not share a direct connection, but only an indirect connection via the

[17] Ogden, C. K.; Richards, I. A.: The meaning of meaning (1923)

[18] Wrigth, S. E.: From the Semiotic Triangle to the Semantic Web (2023)

[19] Arp, R. et al.: Building ontologies with basic formal ontology (2016), p.2.

reference. Further, between no two entities in the triangle exists a one-to-one relationship. There can be many symbols that evoke the reference and there can be multiple referents to a reference. Considering the example, a photograph of the reader's car also constitutes a *"symbol"*.

Data, Information, Knowledge

To accurately identify the quantities being processed by a data reduction process, it is important to delineate the terminology used. While there is no absolute consensus on exact definitions[24], the following will provide working definitions for the terms *"data"*, *"information"*, *"knowledge"*. The relation between these terms is generally modelled as the data-information-knowledge hierarchy, or data-information-knowledge-wisdom (DIKW) if wisdom as top layer is included.[21] Both models are displayed in figure 2-2. Ahsan and Shah[26] explicitly describe this model as an aggregation from one level of the hierarchy to the next. While the concept of *"wisdom"* is at the top of the DIKW hierarchy, it is also a not yet very well understood concept due to its high level of abstraction.[21] Since the concept of wisdom is often left out of scientific considerations and because of a lack of scientific consensus, it will also not be considered in this work.

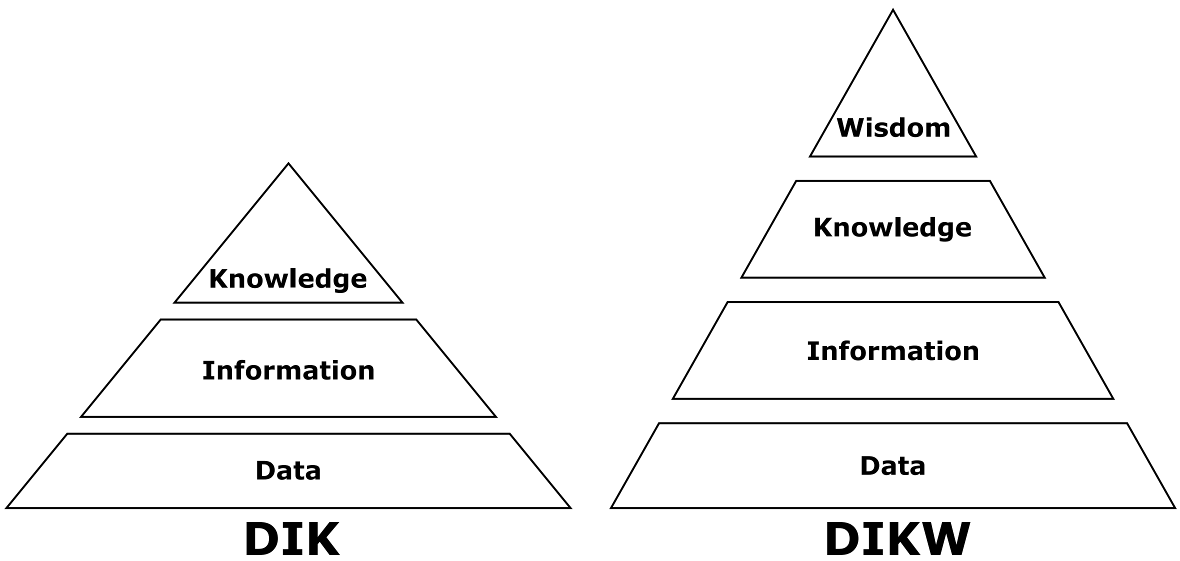

Figure 2-2: The DIK and DIKW hierarchy (figure recreated from Jifa[25])

[20] Choo, C. W.: The Knowing Organization (1996)

[21] Rowley, J.: The wisdom hierarchy: representations of the DIKW hierarchy (2007)[a:p.11, b:p.12]

[22] Ackoff, R. L.: From Data to Wisdom (1989)

[23] Sanders, J.: Defining terms: Data, information and knowledge (2016)[a:p.4]

[24] Dervos, D. A.; Coleman, A.: A Common Sense Approach to Defining Data, Information, and Metadata (2006)

[25] Jifa, G.: Data, Information, Knowledge, Wisdom and Meta-Synthesis of Wisdom (2013)[a:p.714]

[26] Ahsan, S.; Shah, A.: Data, Information, Knowledge, Wisdom: A Doubly Linked Chain? (2006)

Starting with a delineation for the term *"data"*, Dervos and Coleman[24] define data as a representation of *real world facts*. Sanders identifies these *real world facts* as a carrier of information and specifies further: while *"data can be stored"*, he offers a clear distinction to *"information [, which] [...] cannot be stored but can be represented by data"*[23a].

Using a different interpretation, Dervos and Coleman[24] state that information is a property of data that has to be revealed by successful interpretation. Similarly, Ackoff[22] determines that information is a reduced product of a data analysis process. He further identifies information as a type of functional attribution, instead of a static property of data. Sanders also comes to the conclusion, that data is always an approximation of the original information.[23] Concerning where data originates from, Ackoff[22] states that data is observed/sensed. As such, any storing of information as data, which is the only option to make information persistent, can be understood as a reduction process. This means persistent information is always reduced in one way or another. Due to this inevitability of reduction, an understanding of reduction methods is imperative.

Within this dissertation the term *"data"* is defined as a discrete, reduced representation of information as symbols. Further, the term *"information"* will be used to denote the abstract intangible reference to an entity.

Considering the term *"knowledge"* and summarizing relevant literature, Rowley notes that *"knowledge is an elusive concept which is difficult to define"*[21a]. Exploring the state of the art, different authors define *"knowledge"* with the usage of varying entities, resulting in no unified final definition of the term. Still, many authors include a relation to an action or task in their definition of knowledge. Sanders gives the definition of knowledge as *"the persistent, appropriate response to a given input"*[23a]. Jifa[25a] defines knowledge in relation to information in that knowledge adds *"how to use it"*. Using this as a baseline, the term *"knowledge"* will be used to mean the correlated/causal relation between information.

These concepts are explained with an example using three pieces of information. They are the abstract concepts of **water**, **heat**, and **steam**. These pieces of information can be stored as data, as well as evoked from the same data. For this example, a UTF-8 encoding, to differentiate it from the other written text, is chosen.
water → `\x77 \x61 \x74 \x65 \x72`
heat → `\x68 \x65 \x61 \x74`
steam → `\x73 \x74 \x65 \x61 \x6D`
An example of knowledge for these pieces of information would be the ternary relation (**water** and **heat** leads to **steam**).

Relevance

Within this dissertation, the concept of relevance plays a vital role. As such, *"relevance"* and related terms will not be used with their colloquial meaning, but with meanings later defined in chapter 5.

Data is by definition an entity that can be stored. How to formalize information and knowledge is a separate matter. For this purpose, ontologies have been proven to be the right tool to formalize these entities.[28] An ontology is an *"explicit specification of a conceptualization"*[29], where conceptualization is *"a body of formally represented knowledge"*[30]. The information represented by ontologies is divided into two groups. The first, called the terminology box (T-Box) contains *"concepts, role definitions and axioms"*[27]. The second group, called assertion box (A-Box) contains *"instances of concepts and roles"*[27]. The use of specialized algorithms and description formats, called reasoners and description logic, allows the inference of additional information and knowledge based on the specified concepts and axioms.

Within the automotive domain, ontologies have found wide application in describing the scenes and scenarios of the traffic environment.[27] Examples of applications include the description and creation of scenes for simulation-based testing[31,27], the recognition of criticality[32] and hazardous scenarios[33], as well as decision-making[34].

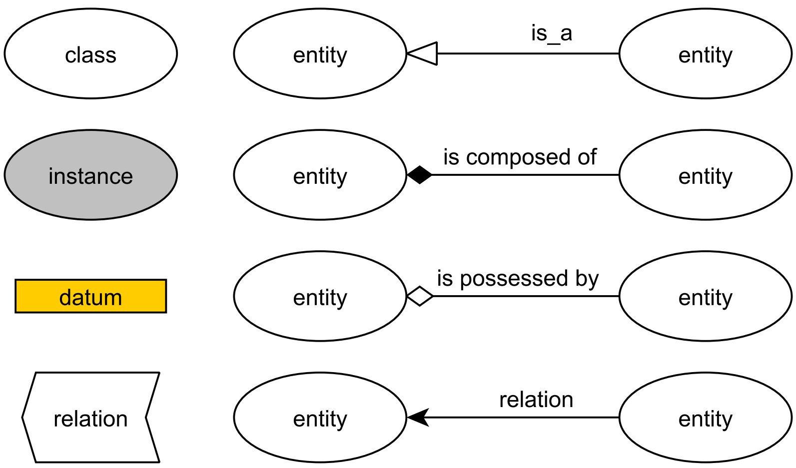

Figure 2-3: Ontological elements used for visualizing the relevance models

[27] Hülsen, M. et al.: Traffic intersection situation description ontology for advanced driver assistance (2011)p. 995

[28] Křemen, P.; Kouba, Z.: Ontology-Driven Information System Design (2012).

[29] Gruber, T. R.: A translation approach to portable ontology specifications (1993), p. 199.

[30] Guarino, N. et al.: What Is an Ontology? (2009), p. 3.

[31] Bagschik, G. et al.: Ontology based Scene Creation for the Development of Automated Vehicles (2018).

[32] Westhofen, L. et al.: Using Ontologies for the Formalization and Recognition of Criticality (2022).

[33] Kramer, B. et al.: Identification and Quantification of Hazardous Scenarios for Automated Driving (2020).

[34] Buechel, M. et al.: Ontology-based traffic scene modeling, and decision-making for automated vehicles (2017).

For the purposes of this dissertation, the concepts and notation as depicted in figure 2-3 are established. Four basic entities are used: class, instance, datum, and relation. Classes represent abstract concepts of entities, which can be further refined into subclasses. Additionally, classes can have multiple concrete instances, for example there can be multiple real cars associated with the abstract class/concept of a car. These relations are denoted as an is_a relation. Further, entities can be compositions or aggregations of other entities. The main difference between composition and aggregation relations is, that with the latter both entities exist independently of each other, while for the former the composed entity only exists through its composites. The composition is possibly best described by the example of a room being a composite of four walls. Without the existence of all the walls the room as a concept ceases to exist. These relations are denoted by is_composed_of and is_possessed_by. A datum represents an individual piece of data of arbitrary type.

Additionally, other arbitrary relations based on the use case are possible and are denoted by an arrow. While relations are usually not denoted as separate entities, this notation includes an optional dedicated symbol for relations. As relations take on a special role in relevance theory, they are treated as separate entities. This enables explicitly linking another ontological entity to a relation.

Reduction & Expansion

In a lot of literature the terms reduction, compression, and encoding are used interchangeably. For a consistent nomenclature, this dissertation will use the term reduction to reference any process that manipulates the size of data from an initial magnitude to a lesser magnitude. For mathematical notation, both reduction and expansion are modeled as functions, as described in equations 2-1 and 2-2. The input to these functions are either individual datums or a dataset. The latter being represented as the mathematical concept of a set by using the double-struck notation of the variable. Thus, datasets are denoted by $\mathbb{D}$. The reduced result of the reduction function is denoted by the check accent $\check{\mathbb{D}}$, while the expanded result of the expansion function is using the circumflex accent $\hat{\mathbb{D}}$. This notation was chosen to represent a downward and upward arrow, symbolizing the change in data size.

$$\check{\mathbb{D}} = f_{\text{reduction}}(\mathbb{D}) \tag{2-1}$$

$$\hat{\mathbb{D}} = f_{\text{expansion}}(\check{\mathbb{D}}) \tag{2-2}$$

2.2 Use Cases for Data in the Context of HAD

Software is a key part of an ADS. As such, it has many interfaces to various kinds of data and use cases for data throughout its lifecycle. To understand the importance of data for AD it needs to be understood which use cases data fulfills for AD.

In the following subsections, multiple data reliant use cases within the context of AD are identified. For each use case, the extent of the reliance on data is described, and what importance data reduction has to the respective use case.

Dynamic Driving Task

The most obvious task is the dynamic driving task (DDT) itself. The goal within this use case can, in the simplest form, be described as moving a vehicle safely to a target position. This task may be performed by a human operator, an ADS, or a shared operation between human and ADS. The components executing the DDT are not limited to the vehicle-to-be-operated itself, but may include components external to the vehicle. As such, a distinction between the intra-, and extra-vehicular components and a mix thereof can be made.

Intra-Vehicular components

Within the intra-vehicular components, the executing elements are located in close proximity of each other within the vehicle itself. Here, the DDT ranges from an ADAS function assisting the driver at SAE levels 1 and 2 to the ADS fulfilling driving missions with full autonomy at SAE level 5. Data exchange between individual components takes place with varying bandwidths and latencies depending on the need and technology used. Transferring data within a single electronic control unit (ECU) can be done at high speeds of 256 GB/s (2048 Gb/s) via internal bus systems, like PCI-e 6.0.[35] Current development trends show a doubling of bandwidth every three years, with 516 GB/s (4096 Gb/s) becoming available with PCI-e 7.0 in 2025.[35]

A more significant bottleneck exists when multiple ECUs need to transfer data between each other. A sample of existing intra-vehicle communication technologies is listed in table 2-1. Many existing technologies have bandwidths lower than 100 Mb/s stemming from their use in older SAE level 2 systems. With the focus on SAE level 3 systems and higher, Ethernet has become

[35] Group, P. C. I. S. I.: PCI-SIG® Announces PCI Express® 7.0 Specification to Reach 128 GT/s (2022)
[36] Tuohy, S. et al.: Intra-Vehicle Networks: A Review (2015)
[37] Wang, J. et al.: Networking and Communications in Autonomous Driving: A Survey (2019)[a: p.1248, b:p.1249]
[38] IEEE: Std 802.3bt-2018 - 10 Mb/s Operation and Power Delivery over a Single Pair of Conductors (2020)
[39] IEEE: Std 802.3bw-2015 - 100 Mb/s Operation over a Single Balanced Twisted Pair Cable (2016)
[40] IEEE: Std 802.3bp-2016 - 1 Gb/s Operation over a Single Twisted-Pair Copper Cable (2016)
[41] IEEE: P802.3bs - Physical Layers and Management Parameters for 200 Gb/s and 400 Gb/s Operation (2017)

Table 2-1: Different intra-vehicle communication technologies and their bandwidths

Local Network Technology	Bandwidth
LIN	19.2 Kb/s[36,37b]
CAN	1 Mb/s[36,37b]
D2B	11.2 Mb/s[37b]
FlexRay	20 Mb/s[36,37b]
MOST	150 Mb/s[36,37b]
IDB	400 Mb/s[37b]
LVDS	655 Mb/s[36,37b]
Automotive Ethernet	10 Mb/s - 1 Gb/s[38,39,40]
Ethernet	up to 400 Gb/s[41]

an established communication technology in ADS.[37b] Automotive Ethernet provides a range of bandwidths depending on the use case ranging from 10 Mb/s[38], over 100 Mb/s[39], to 1 Gb/s[40]. Recent non-automotive standards enable the use of Ethernet with a bandwidth of up to 400 Gb/s[41].

Focusing on cameras, an estimate of the raw bit rate of a current AD sensor setup can be made. For this, a setup with ten cameras using a 4k resolution at a frame rate of 15 hz and a bit depth of 16 bit per pixel is assumed. The product of these parameters yields an estimated raw data rate of 2.654 GB/s (21.232 Gb/s). Comparing this value to the available network technologies above, it becomes apparent, that even most current automotive Ethernet is not capable of handling such a raw bit rate. In contrast, only the most current non-automotive Ethernet standard[41] can support similar bandwidths. Thus, the design of an ADS needs to account for this bottleneck.

This is either done by choosing an appropriate architecture to minimize the need to transfer larger amounts of raw sensor data or by reducing the raw sensor data to a manageable size. This estimate serves to approximate the order of magnitude of raw sensor data. While different ADS have different sensor setups, current ADS show similar numbers of sensors[42], with the 2022 Waymo ADS having 14 cameras[43]. Examining the list of used sensor setups by Marti et al.[42] a general trend of growing sensor numbers in ADS can be observed. As such, the challenge of transferring sensor data is set to increase in the coming years.

Transferring data becomes even more of a challenge when data needs to be transferred out of the vehicle, generally denoted as vehicle-to-everything (V2X). This transfer is identified in more detail by the communication partners of the interaction: vehicle-to-vehicle in case of communication between two traffic participants, vehicle-to-infrastructure if the communication is between a traffic participant and traffic infrastructure. While traffic infrastructure can include traffic lights and their current state, it can also include infrastructure with sensors, relaying additional dynamic environment information to the respective traffic participant. A prominent application for this

[42] Marti, E. et al.: A Review of Sensor Technologies for Perception in Automated Driving (2019)

[43] Tangramvision: Sensing Breakdown: Waymo Jaguar I-Pace RoboTaxi (2022)

type of communication is extending the field of view in intersection scenarios in order to mitigate occlusions and collision potential.[44,45]

Currently, many different implementations and approaches to utilizing V2X communication are proposed. With the focus on data, these proposals can be assigned to groups, with varying levels of abstraction and need for bandwidth. Considering communications with lower bandwidth requirements, an example is the sharing and coordinating decision and planning data.[46] This kind of communication has been researched and shown to work in real time field tests in the project IMAGinE[47].

Other approaches also include environment description data such as a mixed object list and birds-eye-view (BEV) representation.[48] The sharing of general environment representations for further interpretation is described by the term cooperative perception. Cooperative perception is subject to the specific challenge, where data reduction is used since cooperative perception has to operate with both large data and limited bandwidth.[49]

In addition to the general environment representations that are understandable to humans, some systems opt for sharing inferred latent feature data directly extracted from deep neural network (DNN)s, as Yang et al.[50] demonstrates. This form of data sharing is less abstracted for human understanding but allows for the preservation of finer details. However, it is also more data-intensive compared to sharing high-level information. Different layers of latent features, generated by the deep learning models, can be strategically selected from one observer and fused with layers from another observer to optimize the trade-off between detail and bandwidth usage.

The approach with the highest need for bandwidth is cooperation perception with no prior extraction of information. Here, data is transferred in a raw state as it is gathered from the sensor. This raw data is then used to create a fused representation with the ego vehicle's sensor data. An example of this approach is generating a joined view, which includes otherwise occluded regions, enabling either the construction of a more complete map view or projections with the ability to effectively see through obstructions.[51]

Contrary to data transfers concerning the DDT, larger amounts of data need to be transferred when the DDT is performed from a remote, extra-vehicular situation.

[44] Zhang, C. et al.: Occlusion-Aware Planning for Autonomous Driving (2023).

[45] Buchholz, M. et al.: Handling occlusions in automated driving using a multiaccess edge computing (2021).

[46] Maag, C. et al.: Supporting cooperative driving behaviour by technology (2022).

[47] EICT GmbH: IMAGinE: Intelligent Maneuver Automation (2023).

[48] Ngo, H. et al.: Cooperative Perception With V2V Communication for Autonomous Vehicles (2023).

[49] Cui, G. et al.: Cooperative perception technology of autonomous driving (2022).

[50] Yang, Q. et al.: Machine-learning-enabled cooperative perception for connected autonomous vehicles (2021).

[51] Kim, S.-W. et al.: Multivehicle cooperative driving using cooperative perception (2014).

Extra-Vehicular components

Performing the DDT from a remote, extra-vehicular situation, is referred to as teleoperation. The primary use case for teleoperation is to provide a safe and efficient solution to the ADS reaching the limits of its operating design domain (ODD).[52]

Teleoperation can be divided into three modes of operation, remote driving, remote assistance and remote monitoring, based on the degree of involvement in the DDT:[52] The first is remote driving, where the whole DDT is being performed from a remote situation. In remote assistance, the system has to perform the DDT, while in special circumstances a remote operator can render assistance by interacting with the ADS on any layer of the planning or perception task.[53] The lowest degree of involvement presents remote monitoring, where only certain maneuvers, usually minimum risk maneuvers, can be triggered, and no other mode of interaction is possible.

All modes of operation rely on sensor data being relayed to the remote operator. Especially for the modes of remote assistance and remote monitoring, unabstracted environment representations are important, as to give the operator the ability to recognize errors in the system's perception. Evaluations have shown that more frequent sensor updates, even with a lesser bitrate, yield a better scene understanding for an operator.[54] For video streams, a bitrate of 300 kbit/s to 800 kbit/s is sufficient.[55] In addition to frame- and bitrate, latency plays a vital role. Here, a user study has shown that a latency of 300 ms should not be exceeded for teleoperation.[57] This constrains which types of data reduction can be performed for teleoperation since each data reduction algorithm further adds to the overall latency. As such, teleoperation uses a combination of well established data reduction methods, like the standardized Advanced Video Coding (AVC) H.264[58] and High Efficiency Video Coding (HEVC) H.265[59] algorithms, and simple context-specific approaches, such as region of interest (RoI) masking, where only the relevant areas of an image are transmitted.[56]

Development

Data plays a vital role in developing automated driving functions. The development of AD is currently experiencing the challenges of handling algorithmic complexity, a multitude of scenarios to be considered, growing networking and dependencies in architectures, as well as complex and

52 Majstorović, D. et al.: Survey on Teleoperation Concepts for Automated Vehicles (2022)

53 Feiler, J.; Diermeyer, F.: The Perception Modification Concept to Free the Path (2021).

54 Hofbauer, M. et al.: TELECARLA (2020)

55 Hofbauer, M.; Sc, M.: Adaptive Live Video Streaming for Teleoperated Driving (2022)

56 Hofbauer, M. et al.: Traffic-Aware Multi-View Video Stream Adaptation for Teleoperated Driving (2022)

57 Neumeier, S. et al.: Teleoperation: The Holy Grail to Solve Problems of Automated Driving? (2019).

58 International Telecommunication Union: H.264: Advanced video coding for generic audiovisual services (2019).

59 International Telecommunication Union: ITU-T Rec. H.265 (08/2021) High efficiency video coding (2021).

60 Bach, J. et al.: Data-driven development, a complementing approach for automotive systems engineering (2017)

fuzzy systems. Data driven development has the potential to mitigate these challenges and not to replace but to complement existing development strategies[60]. Data driven development is a term used to describe a closed-loop between the development of a function, its testing, and refinement of the function based on test data[61], during all stages and activities of product development.[60]

While data driven development is not exclusive to any specific type of function, it has a special status and significance for the development of artificial intelligence functions or, more specifically, the training of DNNs. Since the training of DNNs is a process to *"identify patterns and extract features from [...] data"*[62] through optimization, it requires large amounts of data to achieve a good performance for the function.[63]

This data dependence becomes especially evident when looking at the datasets, which form the basis for training DNNs. Both, the rate of novel datasets publications[64] and the upper range of available dataset size[65], are increasing.

Currently, the largest synthetic, as well as overall dataset is the SHIFT dataset with 10 TB of size[66]. For real-world data, the Argoverse2 dataset, being composed of the sensor, motion forecasting, LIDAR and map change sub datasets, has a combined size of 7 TB[67]. The SHIFT dataset was released in 2022 and the Argoverse2 dataset in 2021. Considering well established datasets like Waymo Open dataset[68] and the NuScenes[69] dataset, the former was released in 2019 with a size of 2 TB and the latter in 2020 with a size of 0,3 TB.

The role of data for development of DNNs will be highlighted in more detail in section 3.2.4 of the next chapter.

Due to the uncertain nature of the open context, ADS will operate in operating domain (OD)s where a mismatch between required, specified, and implemented behavior[70] has to be expected. To account for this circumstance, it is proposed to extend the consideration of the open context to the whole product life cylce[71]. This enables a continuous refinement and assessment of assumptions made during development, based on evidence data collected from field operation.

61 DSpace: Data Driven Software Development (11.11.2023).

62 Jan, B. et al.: Deep learning in big data Analytics: A comparative study (2019), p. 281.

63 Hestness, J. et al.: Deep Learning Scaling is Predictable, Empirically (2017), p. 6.

64 Bogdoll, D. et al.: Ad-datasets: a meta-collection of data sets for autonomous driving (2022), p. 6.

65 ad Datasets: Autonomous Driving related Collection of Datasets (28.07.2023).

66 Sun, T. et al.: SHIFT: A Synthetic Driving Dataset for Continuous Multi-Task Domain Adaptation (2022).

67 Wilson, B. et al.: Argoverse 2: Next Generation Datasets for Self-driving Perception and Forecasting (2021).

68 Sun, P. et al.: Scalability in Perception for Autonomous Driving: Waymo Open Dataset (2019).

69 Caesar, H. et al.: nuScenes (2020).

70 Stellet, J. E. et al.: Formalisation and algorithmic approach to the AD validation problem (2019), p. 47.

71 Bangarevva, P. et al.: In-Service Monitoring and Assessment of AD vehicles with AI based Algorithms (2023).

Release / Safety Assurance

While safety assurance is a key part and final step of any development process, from a data/information/knowledge side it needs to be viewed separately. This due to the *"assurance trap"*[72], where the processes and available data might be sufficient to develop a safe system, but the processes and knowledge to prove the systems safety is insufficient. As such, additional novel methods are necessary to create an argumentation for sufficient safety of an ADS.

While there is currently no single accepted method for providing a safety argumentation, a much researched approach is scenario based testing.[74] The essence of scenario based testing is compiling all scenarios relevant to describing the ODD of an ADS into a catalog and verifying the safety of the ADS within each scenario. This catalog is compiled from various sources including crash data analysis, naturalistic driving data, and expert knowledge.[75,76]

Scenario descriptions are available at different levels of detail. Concrete scenarios are fully specified and can be simulated. Compared to concrete scenarios, logical scenarios are less specific. Logical scenarios are machine-readable descriptions that specify only ranges or distributions, instead of concrete values, for the parameters of the scenario. Functional scenarios are the most abstract. They provide only human-readable textual descriptions of the scenario. While the level of description varies, each serves a specific purpose. Safety testing requires concrete scenarios. However, concrete scenarios are not suitable for fully describing the ODD, since the ODD contains an infinite number of concrete scenarios. Therefore, more abstract scenario descriptions such as functional or logical scenarios are used for this task.

The scenario-based testing approach thus requires scenario data in two key steps. The first step is to analyze and extract those functional and logical scenarios that are most suitable to describe the ODD. The second step contains the deriving and execution of concrete scenarios as test cases. The results of the test cases are then used as evidence in a safety argumentation. A prominent method employing scenario based testing for generating a safety argumentation was established by the PEGASUS project[73]. This method is showcased in figure 2-4. Starting from the lower left with an initial set of data, knowledge, and use cases, it proceeds through the analysis and extraction step to the top right. The second step of deriving, executing, and analyzing test cases can be seen from the top right to the top left. Finally, the PEGASUS method displays a closed-loop, refining the initial data, knowledge, and use case, in the case that the evidence produced is insufficient to provide a safety argumentation.

[72] Wachenfeld, W. H. K.; Winner, H.: The Release of Autonomous Vehicles (2016), p. 439.

[73] PEGASUS Project: PEGASUS Method: An Overview (2019)

[74] Neurohr, C. et al.: Fundamental Considerations around Scenario-Based Testing (2020), p. 121.

[75] Fremont, D. J. et al.: Formal Scenario-Based Testing of Autonomous Vehicles (2020), p. 2.

[76] Consortium, P.: PEGASUS METHOD (2019).

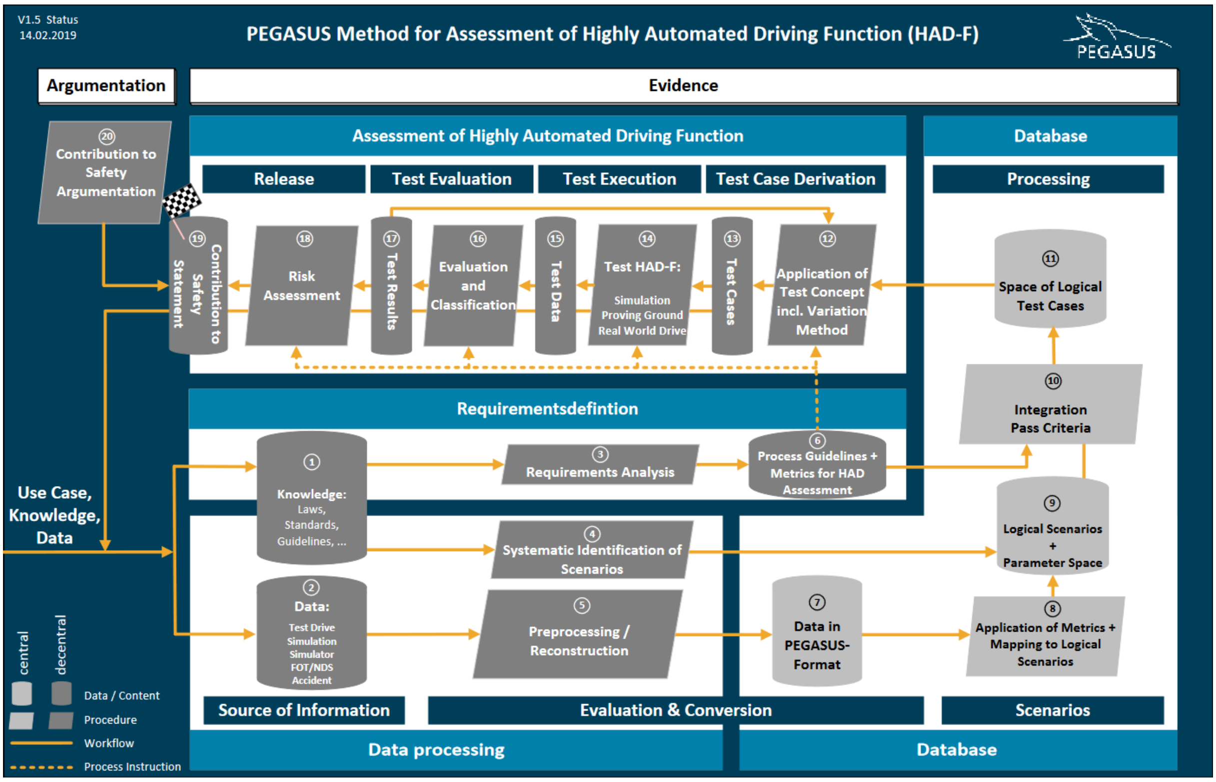

Figure 2-4: The safety assurance method, with data flow as proposed by the PEGASUS project. (Image take from [73])

Besides just generating the knowledge to describe the ODD and performing tests in concrete scenarios, maintaining traceability is important[77]. Traceability is key to any safety argument, since it assures compliance, drives efficiency, enables handling complexity, and acts as a safeguard for quality.[78] Considering the importance of data within the use case of releasing and assuring the safe of an ADS, it can be summarized that the role of data is an essential building block of creating a safety argumentation. Large amounts of data are necessary to generate knowledge about the ODD. Similarly, large amounts of data need to be generated as evidence for the safety argumentation. Throughout the whole process, this data needs to be stored to ensure traceability.

Monitoring

The data use case of monitoring is concerned with observing and ensuring the proper operation of an ADS within its specified bounds. Similar to the DDT this can be done from an in-vehicle or a remote location. Contrary to remote operation monitoring does not include a closed-loop design and cannot influence the actions of the ADS. The only purpose is to gather and report information about the current state and performance of an ADS within the traffic environment.

[77] Johanson, M. et al.: Big Automotive Data (2014), p. 738.

[78] Steinkirchner, K.; Bühler, C.: Usage of Baselining and Traceability Demonstrator Developed in VVM (2022).

[79] Bundesministeriums für Digitales und Verkehr: Autonome-Fahrzeuge-Genehmigungs- und-Betriebs-Verordnung – AFGBV) (86/22)[a: p. 47, b: p.53]

The level of detail in both data gathering and reporting can be varied and change with the goal pursued. Full traces of DDT data can be recorded and evaluated for development to improve on the current version of the ADS. As previously stated in the subsection above concerning the DDT continuous recording and transfer of full sensor and system data is impossible with today's technology. To this end event-based recording strategies are employed, identifying relevant events during the DDT, recording and transferring only small intervals of event related data. The most prominent type of event-based recording is accident reporting.[80] The German jurisdiction mandates since 2022 that any ADS has to record data once an accident is detected or a risk minimal maneuver is performed with subsequent remote operation.[79a] No full trace is mandated but rather high level data about the vehicle, the ADS itself, and the environment are required.[79b]

Another goal besides the improvement is the statistical monitoring of the real and expected performance of the ADS. Here, only the general event is of interest, which is then compiled into reports. An example of this type of report is the safety report of Waymo[81].

Johanson et al. estimate data amounts for a vehicle fleet of 1000 non-automated vehicles. State-of-health is given as the least data intensive monitoring only requiring 36 GB/year. With medium requirements, remote diagnostic read-outs necessitates 3.6 TB/year. A full monitoring of the controller area network bus with a data rate of 250 kb/s (31.25 kB/s) is estimated to come to 1.6 PB/year.[82]

Using the same assumptions for the sensor data of an ADS the total data of a 1000 vehicle fleet would yield several exabytes per year. While these numbers are significantly less than the data rate produced by the sensors of an ADS, they signify the magnitudes of data that can be accumulated over time when monitoring an entire fleet. The conclusion to be drawn is that it is essential to identify the events and pieces of information most relevant to the targeted goal in mind.

2.3 Metrics for Data Reduction Evaluation

In order to evaluate data reduction methods, two aspects need to be assessed. The first aspect concerns how data reduction affects the effort needed to handle data. The second aspect is focused on the effect that data reduction has on the performance on the data for a given use case.

The relationship between effort reduction and residual performance resembles a trade-off, where optimizing one aspect may come at the expense of the other. This trade-off can be visualized as a

[80] Parmar, P.; Sapkal, A. M.: Real time detection and reporting of vehicle collision (2017).

[81] Victor, T. et al.: Safety performance of the Waymo rider-only automated driving system (2023).

[82] Johanson, M. et al.: Big Automotive Data (2014), p. 738.

Pareto front, representing the spectrum of possible variations in data reduction methods, thereby delineating the boundaries of what can be achieved.

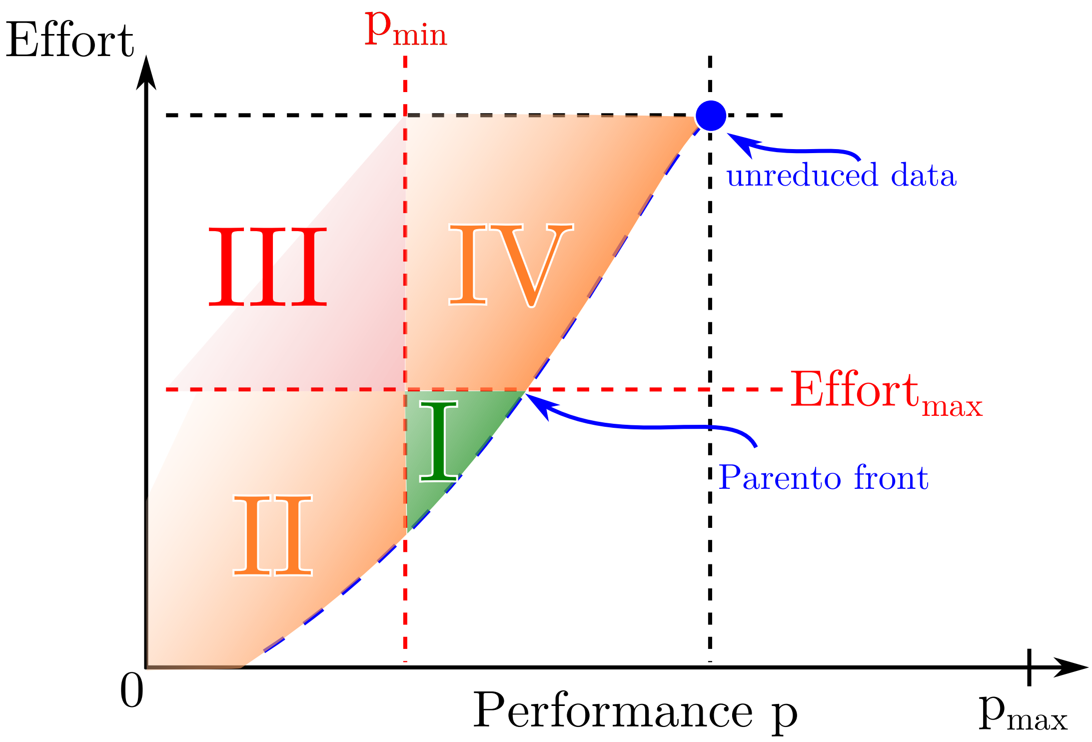

Figure 2-5: Change of effort over performance for Data Reduction

Both of these aspects are governed by two constraining requirements: The minimum required performance: This sets the threshold below which the data performance cannot fall without rendering it unusable for the intended purpose. The maximum available effort: This defines the upper limit of effort that can be put forth. By considering these constraints, the application of data reduction can be classified into four distinct quadrants, as depicted in figure 2-5:

I. Possible application: In this quadrant, data reduction methods can be applied successfully, meeting both the minimum required performance and the maximum available effort.

II. Performance requirement is violated: Here, data reduction techniques may lead to a performance level that falls below the minimum required, rendering them unsuitable for the given purpose.

III. Violation of both constraints: In this scenario, data reduction methods fail to meet both the minimum required performance and the maximum available performance, making them infeasible.

IV. Effort requirement is violated: Data reduction methods in this quadrant exceed the maximum available effort while still meeting the minimum required performance.

It is important to note that the Pareto front may intersect with the constraints in a manner that renders one or more quadrants unreachable. If quadrant I is not reachable, a variation that places the data reduction effect in quadrant IV can still be considered, since the residual performance is sufficient for the use case. If only quadrants II and III are possible, the available variations of data reduction are insufficient to be applied, necessitating further development need.

2.3.1 Effect of the Data Reduction towards Data Handling Effort

Data handling requires effort to be put into the process. In the context of this thesis, handling means applying any type of process, such as transforming, storing, analyzing, etc., to data. Additionally, in this thesis, the term effort is defined as the allocation of resources, which can include various factors such as financial expenditure, human or machine work performed, provision of technical resources, and more. Applying data reduction to data affects the effort required to handle the data. This effect can be divided into two parts, a positive and a negative effect. The former reduces, and the latter increases the effort to handle data.

For the application of data reduction to be feasible, the positive effect on effort must outweigh the negative effect, resulting in an overall effect of reduced effort. The nature of the positive and negative effects on effort is briefly described below. Determining the exact magnitude of the effects requires sufficient effort models for the use case of interest and accurate knowledge of the values for the model parameters. Such models are considered out of scope for this thesis due to their business centric focus.

Positive Effect

The positive effect on effort is caused by the decreased size of data after applying a data reduction to it. The decrease in data size is expressed through a comparison of the reduced data against the unreduced data. To account for different sizes of the unreduced data, the decrease is generally given as a difference in information density. For this density, the ratio between semantic symbols, necessary to extract information, and syntactic symbols, employed to describe the former, is employed. A widely used example is bits per pixel given in equation 2-3 describing the potential level of detail for image data. The positive effect irrespective of a specific use case can then be expressed by the reduction factor. The positive effect for a given use case is then given by the data reduction factor, as described in equation 2-4, in conjunction with the data amount and use case specific parameters. Considering storing data as an example these parameters are storage duration and storage cost per amount.

$$bpp_{\boldsymbol{I}} = \frac{b_{\mathrm{I}}}{\mathrm{N}_{\mathrm{pixel},\boldsymbol{I}}} \tag{2-3}$$

$$reduction factor = \frac{bpp_{\check{\boldsymbol{I}}}}{bpp_{\boldsymbol{I}}} \tag{2-4}$$

To fully describe the use case specific parameters, the path of the data flow being provisioned for the use case has to be considered. The provisioning of the data to the use case may involve multiple data sources as well as multiple transmission steps, where the data can be stored in between. These steps can conceptualized as a graph, as the example in figure 2-6 shows.

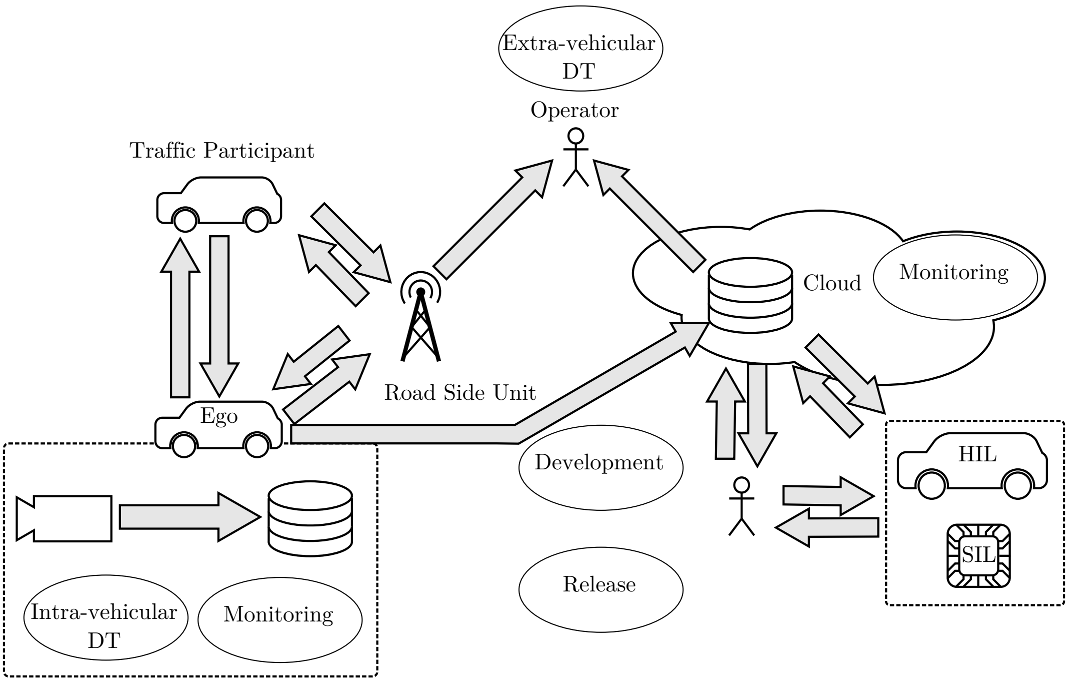

Figure 2-6: Example of a data flow graph for the use cases within the automated driving domain

Each of these steps has individual parameters associated, like transmission, storage costs or transmission bandwidth. Considering the steps of provisioning the data for a use case thus enables the application of data reduction, as shown in figure 2-7, in various instances of the data flow. While this representation requires extensive knowledge and modeling of the infrastructure involved in provisioning data, this view of data procurement enables a holistic estimate of the positive effect of data reduction.

Negative Effect

The negative effects on effort associated with data handling when applying data reduction methods are multifaceted. These negative effects can be divided into two categories: those that depend on the size of the subject data and those that remain invariant.

One aspect of size-dependent negative effects is the computational costs incurred during the reduction and subsequent expansion of data. The process of data reduction demands computational resources, and the subsequent expansion to its original form can also be resource-intensive.

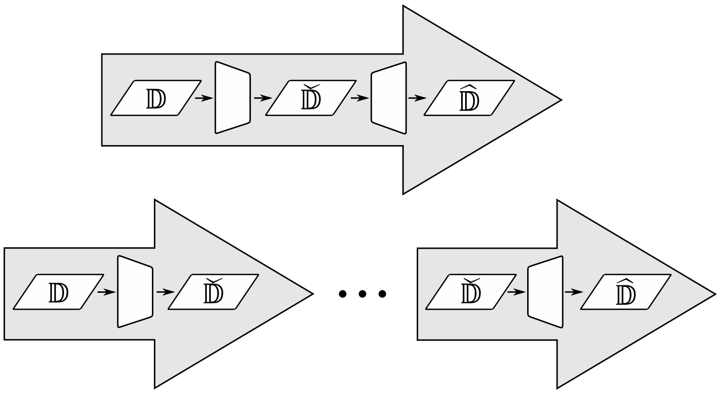

Figure 2-7: Realization of data reduction within a dataflow, either in a single transmission (top) or over multiple transimission steps (bottom)

Additionally, as the size of the data increases, the time required for these operations may escalate, potentially causing delays in data processing. Conversely, some negative effects are independent of the size of the data being reduced. The financial cost associated with the development or procurement of the data reduction method can be considered as an example.

Furthermore, an additional negative effect arises from maintaining and monitoring the data reduction method. Ongoing development in the target use case can change the resulting relationship between residual performance and reduced effort necessitating continued development and adjustment of the data reduction method. This entails dedicating resources to ensure that the data reduction method remains effective and aligned with the desired effort and performance requirements. The effort invested in monitoring and maintenance is separate from the initial implementation of data reduction and can persist over time.

2.3.2 Residual Performance on the Reduced Data

Concerning the second aspect of assessing a data reduction method, the performance p of the target use case on the reduced data versus the unreduced data has to be considered. Similar to the data reduction factor the residual performance can be described as a performance reduction factor, as given by equation 2-5.

Regarding the second aspect of assessing a data reduction method, it is necessary to examine the performance p in the target use case for both the reduced data and the unreduced data. Analogous to the data reduction factor, the residual performance can be described as a performance reduction factor, defined by the equation 2-5:

$$p_{\text{res}} = \frac{\check{p}}{p} \tag{2-5}$$

While the possible use cases can cover a wide range of applications, this thesis focuses primarily on perception-related tasks. Consequently, the following sections will delineate the current measures used to quantify performance in perception-related tasks. Three main areas are considered to represent perception in the automotive domain, human perception, and machine perception. The latter category is further subdivided into semantic segmentation and object detection tasks.

Human Perception

Zhai et al.[83] provides an overview of available means to assessing the quality of an image for human perception. Zhai firstly established that image quality assessment (IQA) can in a first step be split into two groups, objective and subjective. While the *"subjective assessment is the most reliable way to evaluate the quality of images, because human eyes are usually the ultimate receiver of the images"*[83a], the objective assessment, performed by mathematical metrics has the advantage of repeatability and increased efficiency per assessment.[86]

Within the domain of IQA exist four categories of assessment with three being objective algorithms and one being assessment via subjective study.

1. Reduced-reference IQA
2. No-reference IQA
3. Full-reference IQA
4. IQA through study

No-reference IQA

No-reference IQA is used when only the subject image is given and the respective original image and its features are unknown. While No-reference IQA is quite challenging, due to the lack of a direct comparison, the application No-reference IQA methods has the advantage, that no prior knowledge of the original image is necessary. In place of an original image, prior knowledge about the kind of the expected distortion can still be leveraged.[83b] A common example for No-reference IQA is noise detection.

[83] Zhai, G.; Min, X.: Perceptual image quality assessment: a survey (2020)[a:p.3, b:p.14]

[84] Wang, Z.; Bovik, A. C.: Mean squared error: Love it or leave it? A new look at Signal Fidelity Measures (2009)

[85] Wang, Z. et al.: Image quality assessment: from error visibility to structural similarity (2004)

[86] Wang, Z. et al.: Why is image quality assessment so difficult? (2002)

[87] Z. Wang et al.: Multiscale structural similarity for image quality assessment (2003)

Reduced-reference IQA

Reduced-reference IQA denotes IQA where the subject image and select features of the original image are given. Reduced-reference IQA can be used to assess the image quality after a transmission process, where the features of the original image are transmitted separately as a kind of checksum.

Full-reference IQA

In Full-reference IQA, both subject and original image are given. Since the original image is known the concept of *quality* can be equated to *similarity* in this context. The main use case of Full-reference IQA is output assessment of processes that manipulate images. While the topic of Full-reference IQA can be further disambiguated into spatial domain and transform domain methods, the spatial domain methods appear to the most commonly used methods in the state of the art.[83]

In this scope, the mean square error (MSE) metric (as defined by 2-6) and its extension the peak signal-to-noise ratio (PSNR) metric (as defined by 2-7) were the prime choice for evaluating image similarity for more than 50 years.[84] Their dominance can be explained by their ease of use and understanding, as well as their mathematical properties, since they define valid distance metrics.[84] MSE only quantifies the distance between the values of each individual datum pair of subject and original. PSNR further contextualizes this values by referencing it to the maximum possible value MAX for a datum. Considering an image with 8-bit per channel MAX equals $2^8 - 1 = 255$.

$$MSE = \frac{1}{N} \sum_{i=1}^{N} [x_i - y_i]^2 \tag{2-6}$$

$$PSNR = 10 \log_{10} \left(\frac{MAX_{\boldsymbol{I}}^2}{MSE} \right) \tag{2-7}$$

Despite the above-mentioned points Zhou and Bovik note that these positives do not translate when MSE or PSNR are used to evaluate human perception of image quality.[84] This significant shortcoming led to the development of the structural similarity metric (SSIM)[85] which compares image quality on three components that influence human perception: luminance, contrast and structure. multi-scale structural similarity metric (MS-SSIM)[87] extends SSIM by further incorporating different influencing factors of human perception. Viewing distance to the image, as well as perception capability of the viewer are approximated by evaluating the SSIM score of the image at different scale and distortion levels.

In conclusion, MS-SSIM can be viewed as *"one of the most precise FR-IQA measures"*[88] for state of the art similarity analysis between a subject and a reference image.

IQA through study

As previously mentioned the most robust, but also most expensive method to assert image quality is through subjective evaluation by potential users. The International Telecommunication Union (ITU) outlines several study approaches that can be grouped into two categories, direct assessment and indirect assessment via task.

For the first option, study participants are directly asked to rate the quality of an image on a given scale. For a subjective assessment, the mean opinion score[89] is a well established quality measure. Study participants express their opinion of the presented image quality on a scale from 1 (Bad) to 5 (Excellent). This can be done by presenting both the original and subject image or only the subject image to the participant.

For the second option, participants of the study are asked to solve a specific task given a sample image.[90] Based on the results, a statistic for the correctness of the given answers to the task is constructed. Given the descriptive statistic on performance for the task, a check for validity against an acceptance threshold of $\alpha = 0.05$ is performed. Lastly, concerning timed tasks, the ITU recommends to evaluate both time per task and correctness of task.

Machine Perception

Perception in AD is inherently reliant on machines and only in special use cases like teleoperation performed by humans. It must be considered, that the working mechanism of machine perception is inherently different from human perception. Thus, other metrics to evaluate the quality of reduced images are needed. While it is possible to apply the generalized IQA metrics from human perception for machine perception, it is not necessary nor advisable. As mentioned above, evaluating human perception has the disadvantage, that evaluating the concrete task at hand is quite expensive and time-consuming. For machine perception, this disadvantage disappears, due to faster inference speeds and low computational costs. Therefore, the state of the art in this case is to directly evaluate the results of the target task. Following, the state of the art for the perception tasks of semantic segmentation and object detection will be presented.

[88] Bakurov, I. et al.: Structural similarity index (SSIM) revisited: A data-driven approach (2022), p. 4.

[89] ITU-T: ITU-T Rec. P.800.1 (07/2016) Mean opinion score (MOS) terminology (2016).

[90] ITU-T: ITU-T Rec. P.912 (03/2016) Subjective video quality assessment methods for recognition tasks (2016).

Semantic Segmentation

The task of semantic segmentation refers to assigning a spatial feature, such as an image's pixel, a semantic value, such as "car". Considering a single pixel $\boldsymbol{I}_{\mathrm{i,j}}$ of an image and a set of possible semantic classes $\mathbb{C}$ to assign, the whole task can be thought of as a classification task. When $\mathbb{C}$ is of size two, it is called a binary classification. The result of a binary classification can be described by the terms true positive (TP), true negative, false positive (FP) or false negative (FN) as shown in figure 2-8.

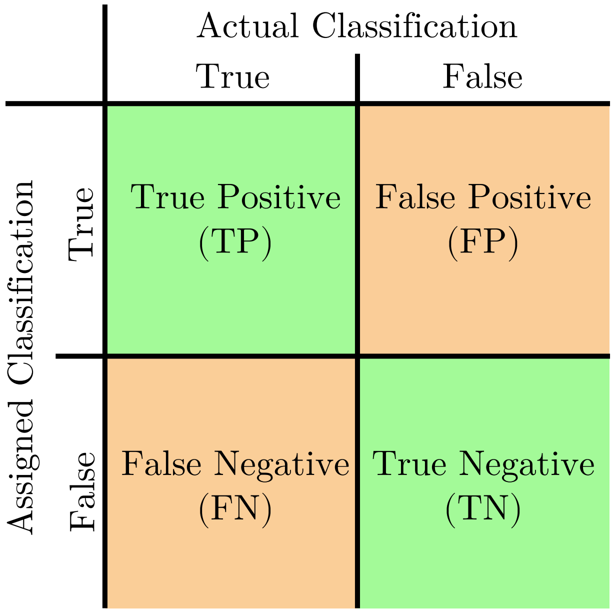

Figure 2-8: Basic classification results of a binary classificator

In order to ascertain the quality of a classification task, the above definitions are usually used to express two describing values, precision and recall. Both describe important characteristics of a binary classification. Precision as defined in equation 2-8 quantifies how likely a true classification is going to be correct. Recall as defined in equation 2-9 on the other hand quantifies how complete the set of true classifications is.

$$\text{Precision} = \frac{TP}{TP + FP} \tag{2-8}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{2-9}$$

While these two values give a good picture of the quality of a classification, in some instances other measures are needed. Considering the domain of machine learning, the necessity to quantify the quality of a classification with a singular number arises. To that end, measures like the Jaccard index J, also called intersection over union (IoU) or the F-1 score are used. Both try to combine precision and recall into one value, striking a compromise in considering both recall and precision [91]. The F-1 score is defined as given in equation 2-10.

$$\text{F-1} = \frac{2}{\text{Precision}^{-1} + \text{Recall}^{-1}} = \frac{2TP}{2TP + FP + FN} \tag{2-10}$$

The definition for IoU measure is given in equation 2-11. Here, two types of notation are used. First, the equation is expressed with the results depicted in figure 2-8. The second notation uses the comparison of two classification matrices. A classification matrix are multiple assignment of a class value c from a set of available classes $\mathbb{C}$. With one classification matrix being the actual classification $\boldsymbol{I}_c$ and the other, denoted by the tilde accent, the assigned classification$\widetilde{\boldsymbol{I}}_c$. Considering that N samples are used to describe classification quality the mean IoU as given by equation 2-12 is used. This measure is also sufficient for non-binary classifications, since they can be re-expressed as multiple binary classifications.

$$IoU = \frac{TP}{TP + FP + FN} = J(\boldsymbol{I}, \widetilde{\boldsymbol{I}}) = \frac{\left|\boldsymbol{I}_c \cap \widetilde{\boldsymbol{I}}_c\right|}{\left|\boldsymbol{I}_c \cup \widetilde{\boldsymbol{I}}_c\right|}, \text{with } c \in \mathbb{C} \tag{2-11}$$

$$mIoU = \frac{\sum_{n=1}^{N} J(\boldsymbol{I}_n, \widetilde{\boldsymbol{I}}_n)}{N} \tag{2-12}$$

Object Detection

Besides the task of classification, the task of object detection is also to be considered. This task is more complex than semantic segmentation, which is often used as one of multiple steps in object detection. The aim of object detection is to identify the existence and location of objects that are of a class within $\mathbb{C}$. Here, the presented measure of IoU will not suffice itself, since the individual pixels are not the main interest, but the higher concept of objects. Instead, the measure of precision is used and calculated for each class c to be detected. Averaging the precision over a set of data samples $\mathbb{D}$ yields the measure average precision. In order to reduce the $|\mathbb{C}|$ measures to a single number, it is further averaged over all classes to what is known as mean average precision (mAP). The definitions of both measures are given in equation 2-13 and 2-14 respectively.

[91] Zhang, E.; Zhang, Y.: F-Measure (2018).

[92] Padilla, R. et al.: A Survey on Performance Metrics for Object-Detection Algorithms (2020)

$$\mathrm{AP}_c = \frac{1}{|\mathbb{D}|} \sum_{I \in \mathbb{D}} \frac{TP_{c,n}}{TP_{c,n} + FP_{c,n}}, \text{with } c \in \mathbb{C} \tag{2-13}$$

$$\mathrm{mAP} = \frac{1}{|\mathbb{C}|} \sum_{c \in \mathbb{C}} \mathrm{AP}_c \tag{2-14}$$

Commonly an arbitrary IoU threshold of 0.5 is used to count a detection as TP. In this case the measure is denoted by mean average precision at IoU=0.5 ($\mathrm{mAP}_{.5}$). Since there is no consensus on which non-arbitrary IoU threshold constitutes a TP, multiple thresholds are usually evaluated, to account for these uncertainties. The mean average precision at IoU $\in$ [0.5:0.05:0.95] ($\mathrm{mAP}_{.5:.95}$) is a generally used example, using IoU threshold between 0.5 and 0.95 in 0.05 increments.[92]

3 Related Works

This chapter delineates the current state of the art with relation to the two core topics of this thesis, data reduction and relevance.

3.1 The Release of Automated Driving

Within AD there are currently several challenges that need to be addressed. The most prominent is the safety assurance for AD. While there are several promising developments toward implementing self-driving cars, substantiated methods to argue their safety are still lacking. This mismatch of capabilities results in the *"approval trap"*[94], which is considered one of the leading challenges for AD. Complexity and general uncertainty have been named as the key contributors to the challenge of developing and releasing AD.[93]

These challenges manifest themselves in three specific types of uncertainty: aleatory, epistemic, and ontological. Aleatory uncertainty describes the uncertainty, which can be observed as the result of probabilistic models. A common example of aleatory uncertainties in AD is the unpredictability of human behavior. Epistemic uncertainty represents a lack of knowledge about a model. The complexity of ADS and of the OD can be seen as a cause of epistemic uncertainty. The key difference between these uncertainties is that epistemic uncertainty can be reduced, by gathering more data, but aleatory uncertainty cannot be reduced.[98] Finally, ontological uncertainty *"can be defined as a condition of complete ignorance in the model of a relevant aspect of the system"*[97a]. Gansch and Adee[97] assign the first two types of uncertainty to what is known as the *known-unknowns* and the ontological to the *unknown-unknowns*. An ontological uncertainty within the OD of AD is commonly referred to as *"open context"*[93,95]. An open context entails several conclusions. First, such an OD cannot be structured, due to the unknown elements within it.[96] Second, complex systems operating in an open context cannot be without any risk.[96]

[93] Rueß, H.; Burton, S.: Whitepaper: Safe AI. How is this possible? (2022)

[94] Wachenfeld, W. H. K.; Winner, H.: The Release of Autonomous Vehicles (2016), p. 439.

[95] Poddey, A. et al.: On the validation of complex systems operating in open contexts (2019)

[96] Stellet, J. E. et al.: Formalisation and algorithmic approach to the AD validation problem (2019)

[97] Gansch, R.; Adee, A.: System Theoretic View on Uncertainties (2020)[a: p.4]

[98] Kiureghian, A. D.; Ditlevsen, O.: Aleatory or epistemic? Does it matter? (2009), p. 105.

3.2 Data Reduction Methods

Methods to reduce the amount of data needed to be stored or transmitted are both quite ubiquitous and old. In this section, the current state of the art of data reduction methods is briefly presented. In general, two groups of data reduction methods can be distinguished: lossless and lossy methods. While lossless methods retain all information in reduced form and are therefore reversible, lossy methods only retain parts of the original information and cannot be reversed without adding distortion to the original data.[99]

In addition to these two groups of data reduction, special attention will be given to data reduction in the context of deep learning. In this case, data reduction is both the final goal and a means to an end. The latter is manifested through the use of deep learning techniques being employed to create more efficient data reduction methods and means to an end, where data reduction serves as a method to enable better data provisioning for the training of DNNs.

3.2.1 Redundancy Reduction / Lossless Methods

Yang and Bourbakis[100] present an overview of state of the art lossless reduction methods. These methods can be grouped into two sequential stages. The first being decorrelation, where structural redundancies are removed by prediction algorithms, the second being entropy coding, where the information per bit is maximized. Yang and Bourbakis[100] state that the reduction effect of entropy coding is approaching its theoretical maximum. Therefore, the focus of research for lossless data reduction is on the decorrelation stage. Current standard formats for lossless image reduction have an average reduction factor of around 0.5 to 0.2.[101]

3.2.2 Information Reduction / Lossy Methods

Lossy data reduction methods, aim to reduce data by only shifting the usage of data towards relevant information and using as little data as possible to represent irrelevant data. For further differentiation, a distinction between the necessary abstraction of the reduction concept can be made. Reduction on high-level entities requires a semantic understanding of the use case,

[99] Li, Z.-N. et al.: Fundamentals of Multimedia (2021)

[100] Yang, M.; Bourbakis, N.: An overview of lossless digital image compression techniques (2005)

[101] Barina, D.: Comparison of Lossless Image Formats (2021).

[102] Wang, C.: Silent Testing for Safety Validation of Automated Driving in Field Operation (2021)[a: p.69]

[103] Shaham, T. R.; Michaeli, T.: Deformation Aware Image Compression (2018)a:p.4, b:p.3

[104] Fu, C. et al.: Texture Segmentation Based Video Compression Using Convolutional Neural Networks (2018)

[105] Di Chen et al.: Pixel-Level Texture Segmentation Based AV1 Video Compression (2018)

[106] Daniele Mari et al.: Content-Aware Compression and Transmission Using Semantics (2023)

[107] Furman, V. et al.: Image and Video Compresssion for Remote Vehicle Assistance (2016)

like object classes and situations, while reduction on low-level entities, like colors and spatial frequencies, only necessitates a semantic understanding of the data itself. To reflect the abstract targeted aspect of reduction on high-level entities, the term selection is used. For the reduction of low-level entities, the term compression is chosen.

Compression can generally be applied to all dimensions present in the data. A general example of compression is subsampling, where every n-th sample is included in the data.

Many compression methods for image and video data are primarily designed with human perception in mind. A widely used standard for lossy image compression is the JPEG format[108]. The key working principle of JPEG is limiting the data used for sudden color value changes between neighboring pixels, with differing thresholds for each color channel. State of the art video compression methods like AVC H.264[109] and HEVC H.265[110] utilize a variety of different techniques to achieve high data reduction factors like quantization, grouping similar structures and self optimizing encodings.[111]

To further optimize existing reduction methods, Shaham and Michaeli[103] propose a novel method and similarity score which takes into account that *"human observers are indifferent to slight local translations"*[103a]. Based on this information they reduce the overall error in the reconstructed image by introducing small deformations to the original image, thus making it *"more compressible"*[103b].

The most simple approach to lossy data reduction can be described by selection or omission. Relevant data is selected, and irrelevant data is omitted from further processing. In the domain of AD, selection is a common technique to reduce large amounts of measurement data. Selection methods can generally be differentiated between scene selection and intra-scene selection methods. Scene selection is concerned with the question, which scenes should be selected and which scenes should be omitted. Intra-scene selection on the other hand considers which components within a single scene should be selected or omitted.

On scene selection, depending on the use case for the data, only a few scenes might be relevant, such as close encounters with other traffic participants or those within a certain ODD. Conversely, many samples might also be a priori irrelevant, such as active recordings while parked. These samples are identified by use of metrics and then respectively selected or omitted. The application of the metrics can either be online as an event data trigger or offline as post recording filtering.

A special case of the event data trigger is the concept of silent testing[102]. In silent testing, one or more virtual automated vehicle instance (vAVI), which are open-loop instances of the system

[108] International Standardization Organization: Information technology: Digital compression and coding of continuous-tone still images: JPEG File Interchange Format (JFIF) (2013)

[109] International Telecommunication Union: H.264: Advanced video coding for generic audiovisual services (2019).

[110] International Telecommunication Union: ITU-T Rec. H.265 (08/2021) High efficiency video coding (2021).

[111] Amer, Hossam: Image/Video Compression: Human and Computer Vision Perspectives (2020), p. 2.

under test, are run virtually in parallel with the physical driving instance (pDI) on the same perception input. This parallelism enables new triggering conditions. The vAVI are evaluated based on criticality metrics and above a set threshold, data is being saved from a ring buffer for further evaluation.[102a]

Additionally, behavioral differences between vAVI and pDI can be used as markers for deeper investigation. Here, the behavior of the pDI is assumed to be safe if no explicit wrong behavior, like an accident, is reported, and thus deviations from the pDI behavior are taken as an indicator of unsafe behavior and as a trigger for recording data. In the event that criticality metrics report a significantly poor performance of the pDI, such as an accident, it is considered an indicator of a perception error and also a trigger for data recording.[112,113]

On intra-scene selection, a distinction is made within a single scene, which entities are relevant and which are irrelevant[127,106,107]. The approaches to determine which entities are relevant and irrelevant differ. Closely related to traditional data reduction approaches, DNNs can be employed to detect similar repeating patterns like textures and resynthesize them after removal. [104,105]

A 2016 patent[107] by Google describes an approach to data reduction, by selecting spatial features within a single scene. The patent performs notes that the selection is performed by matching image regions to a map. Based on relevance criteria not specified in the patent, a RoI is created for the image. The irrelevant pixels outside the RoI are then reduced, while the pixels inside the RoI are left unaltered. The patent names options for handling the irrelevant pixels, by either replacing them with a uniform color or reducing the data of these pixels by removing high discrete cosine transform coefficients.

A similar approach is presented by Wang et al.[114,115] where they propose a video compression technique. Here, a RoI for a given video frame is selected. The selection is based on the relevance assigned to the classes of a prior semantic segmentation of the frame. Following, two new frames are created, one with the pixels inside the RoI on a black background and one with the pixels outside the RoI. Both frames are then reduced using the H.264[116] or H.265[117] standard with a higher reduction factor for the non-RoI frame. Wang et al. note an improvement for a later semantic segmentation task for videos reduced using the proposed technique against the baseline of using a single reduction factor for the whole frame.

---

[112] Wang, C. et al.: Online Safety Assessment of Automated Vehicles Using Silent Testing (2022), p. 13072.

[113] Wang, C. et al.: Reduction of Uncertainties for Safety Assessment Under Parallel Simulations (2021), p. 110.

[114] Wang, Y. et al.: Semantic-Aware Video Compression for Automotive Cameras (2023).

[115] Wang, Y. et al.: A Two-stage H.264 based Video Compression Method for Automotive Cameras (2022).

[116] International Telecommunication Union: H.264: Advanced video coding for generic audiovisual services (2019).

[117] International Telecommunication Union: ITU-T Rec. H.265 (08/2021) High efficiency video coding (2021).

3.2.3 Data Reduction by Deep Learning

A novel approach to data reduction is the usage of DNNs like autoencoders. Here, a neural network, called an encoder, is used to transform the input data into an intermittent representation. The intermittent representation, termed the latent space, serves as an interface between the encoder and a second neural network, the decoder. The decoder is used to approximate the input data from the corresponding point in the latent space. Training of autoencoders utilizes the supervised training technique, optimizing both encoder and decoder on pairs of corresponding input and output data[118]. This working principle is shown in figure 3-1.

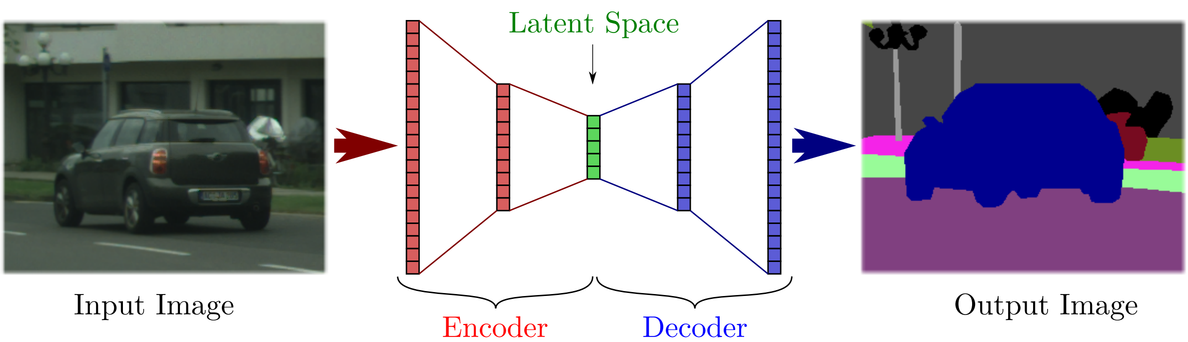

Figure 3-1: Example of the structure of an autoencoder using semantic segmentation of an input image.

If the dimensionality of the latent space is smaller than that of the input data, this approach can be used as a form of reduction.[119] The reduced data can either be stored in a latent space or in a preset format as decided by the user. In the case of the latent space, the semantic meaning of the storage format is defined by the neural network during the training process.

The semantic meaning of the latent space is in most cases not known to a human user. Extracting the semantic meaning poses a non-trivial task. A contributing factor is entanglement, meaning there is no one to one correspondence between a semantic property as understood by a human and a machine.[120]

[118] Janiesch, C. et al.: Machine learning and deep learning (2021), p. 687.

[119] Luo, C.: Understanding Diffusion Models: A Unified Perspective (2022), p. 2.

[120] Pham, C.-H. et al.: Principal Component Analysis Autoencoder for Organising the Latent Space (2022).

[121] Chamain, L. D. et al.: End-to-End optimized image compression for machines, a study (2021)

[122] Toderici, G. et al.: Full Resolution Image Compression with Recurrent Neural Networks (2016)

[123] Schiopu, I.; Munteanu, A.: Deep-Learning based Lossless Image Coding (2020)

[124] Li, M. et al.: Learning Convolutional Networks for Content-weighted Image Compression (2017)

[125] Patwa, N. et al.: Semantic-Preserving Image Compression (2020)[a: p.69]

[126] Luo, S. et al.: DeepSIC: Deep Semantic Image Compression (2018)

[127] Sun, X. et al.: A Task-Driven Scene-Aware LiDAR Point Cloud Coding Framework (2023)

[128] Varischio, A. et al.: Hybrid Point Cloud Semantic Compression for Automotive Sensors (2021)

For image compression, a key motivation to using neural networks over conventional methods is specialization towards machine tasks, since conventional methods are designed with human perception in mind[121]. Different autoencoder architectures have been proposed[129,121,122,124] in recent years. These approaches are already able to outperform the widely used JPEG standard.[122] Chamain et al.[121] present an approach, where they propose a strong coupling between encoder, decoder, and task algorithm. They suggest not to train the autoencoder and the task algorithm separately. Instead, a single loss function should be applied to update the parameters of the encoder, decoder, and task algorithm in an end-to-end fashion to achieve an optimal result. Considering the effect that data reduction can have on a subsequent task, like semantic analysis, Luo et al.[126] proposed a method, where semantic tags are included in the reduced image for later use.

While the focus of the research is primarily on image compression, machine learning approaches to reduce data from other sensor modalities such as LIDAR point clouds[130,127,106,128,131,132] or sound recordings[133] are also emerging.

Due to the nature of the approximate reconstruction of encoder/decoder networks, the process is inherently lossy. This loss can be mitigated by also encoding the reconstruction error[123]. Other methods propose reconstruction error mitigation, by predicting and considering relevant pixels in the encoding process[124]. It has also been proposed, to predict the reconstruction error using DNN and perform a correction step using the predicted reconstruction error.[123]

3.2.4 Data Reduction for Deep Learning

Deep learning and the resulting models as analytic tasks are inherently reliant on the amount and quality of data used. Any manipulation, such as data reduction, on data has the potential to negatively influence the result.[125] In particular, training DNNs is a complex task that has three rather straightforward levers for yielding better results: improved model architectures, more computational time, and larger datasets.[134]

Focusing on the third option, the limiting factor is not technological, as current of training data set sizes are orders of magnitude smaller than any theoretical storage limit. Rather, it is a technical and economical limit, where the burden of provisioning and processing the training data outweighs the benefits of the trained network. For this reason, having effective reduction strategies is key to being able to utilize all three options to create better DNNs. Currently, there

[129] Ballé, J. et al.: Variational image compression with a scale hyperprior (2018).

[130] Till Beemelmanns et al.: 3D Point Cloud Compression with Recurrent Neural Network (2022).

[131] International Standardization Organization: Visual volumetric coding and point cloud compression (2021).

[132] International Standardization Organization: Geometry-based point cloud compression (2023).

[133] Anastasia Natsiou, Sean O'Leary, and Luca Longo: An Exploration of the Latent Space of a CVAE (2023).

[134] Hestness, J. et al.: Deep Learning Scaling is Predictable, Empirically (2017)[p. 6]

is no trend indicating that there is a technological saturation point in the near future for the return rate of employing more data in training as long as the model size is scaled accordingly.[134]

Contrary to the above stated paradigm, that more data is better, the most natural approach to data reduction of a training dataset is a reduction in training samples. As mentioned above, this affects the quality of the results, depending on the information quality of the training samples. While useful, a priori asserting the information quality of individual training samples is a difficult task. Other approaches suggest using conventionally reduced data, such as by using the JPEG or similar formats. The usage of such compression formats is prevalent in large-scale public datasets. While having a high data reduction effect, introducing these types of compression artifacts leads to performance loss.[135]

Another method of reducing the amount of stored data is the inclusion of synthetic data that can be generated at the time of training, thus extending the training data. A simple example is data augmentation, where the sample size is increased by creating deviations or augmentations from an original sample.[136] Common augmentations are any geometric transformations like rotation, flipping, cropping, or introducing shifts in color space. Another type of augmentation is generating a mixed deviation from two or multiple samples.[137] Further approaches have been proposed, where the effective sample size for generative adversarial network (GAN)s can be further increased by also augmenting the output of the generator which then serves as input to a discriminator.[138]

Intuition might suggest an approach to data reduction in deep learning, where samples can be discarded once they have been learned during training. Due to the effect of *"catastrophic forgetting"* training samples have to be retained in order to also retain their learning effect.[139] This effect is used to a positive end in fine-tuning generalist models to a specific task, but poses a substantial problem in incremental learning, where training processes for the same task might be separated by a longer period of time. Rios and Itti[140] therefore suggest the training of a generator alongside the task function, in order to later remove already learned samples from the training set and to replace them with a reduced representation from which a substitute can be generated for later training processes. This approach is most interesting when training is suspended for a longer period of time.

[135] Ehrlich, M. et al.: Analyzing and Mitigating JPEG Compression Defects in Deep Learning (2021).

[136] Taylor, L.; Nitschke, G.: Improving Deep Learning with Generic Data Augmentation (2018).

[137] Zhang, Z. et al.: Bag of Freebies for Training Object Detection Neural Networks (2019).

[138] Tran, N.-T. et al.: On Data Augmentation for GAN Training (2021).

[139] French, R.: Catastrophic forgetting in connectionist networks (1999).

[140] Rios, A.; Itti, L.: Closed-Loop Memory GAN for Continual Learning (2018).

3.3 Relevance

In this section, the concept of relevance will be elaborated and placed in the current state of science. This includes definitions of the concept of relevance as well as derived and related terms, relations to neighboring concepts, and current methods to transform this abstract concept into a tangible form. The context in which this concept is discussed must be considered. First, there is the context of Relevance Theory, which considers the above question its primary focus. Second the context of application, in this case AD, needs to be included as well.

3.3.1 The Concept of Relevance in the Automotive Domain

In the context of automated driving, relevance is a concept that, in the scientific community is most of the time implicitly considered. As this section will show, in most works the question ”What is relevant?” takes precedence over the question of what defines the concept behind this term.

Ulbrich et al.[141] delineates several terms in the automotive domain and defines their conceptional composition by their *”function-relevant”* components. While no exact criteria for function relevance are given, they link it to an influencing interaction between an object and the planning task.

The most common consideration of relevance in the automotive domain is through safety/criticality metrics. While a distinction between relevance and the current criticality term might seem negligible at first sight, Neurohr et al. define that *”[c]riticality (of a traffic situation) is the combined risk of the involved actors when the traffic situation is continued”*[143]. Westhofen et al.[142] present an analysis of the state of the art for criticality metrics in use for the automotive domain.

In general criticality metrics share the following properties:

Criticality metrics consider individual scenes or whole scenarios. Contrary to the above definition of criticality the majority of criticality metrics only consider pairwise relations between two objects. Aggregation of the criticality for all objects is mostly done by extracting the extremes of a set of pairwise relations. Similarly, for scene level metrics the extension from scene to scenario is performed similarly by aggregating over the individual scenes in a scenario.[142a]

Criticality metrics present their results on an ordinal to ratio output scale. Most commonly the unit of the result represents a physical relation such as time, acceleration, jerk, energy, or

[141] Ulbrich, S. et al.: Defining and Substantiating the Terms Scene, Situation, and Scenario (2015).

[142] Westhofen, L. et al.: Criticality Metrics for Automated Driving (2023)[a: p. 13]

[143] Neurohr, C. et al.: Criticality Analysis for the Verification and Validation of Automated Vehicles (2021).

probability, with a few metrics giving their result in the form of a unit-less index.[144]

A linking to relevance is in many cases only implicitly present. A stronger linking is usually achieved through set target values, which are highly dependent on the context they are defined in. Transfer and application outside the defined context remain insufficiently substantiated. As a result, many critical values for these metrics exist, as can be seen for the time-to-collision metric, where Westhofen et al. cite critical values in the range of 1 s[145,146] to 3 s[147] have been presented.

There are also examples of explicit consideration of the concept of relevance in automated driving. Lyssenko et al.[149] argue for relevance metrics that consider the underlying task to bridge the gap between perception and safety metrics. In the paper, the perception performance of pedestrians is evaluated and linked to the distance as a connection to the driving task. In a subset of works, relevance is also explicitly judged. Philion et al.[150] describe a relevance estimation that is linked to the influence of a concrete function implementation on a task. Shalev-Shwartz, Shammah, and Shashua[151] introduce Responsibility-Sensitive Safety (RSS) which aims to define rules for safe behavior. Further, the criteria presented also delineate whether an actor has a *duty of care*. The duty of care is defined as *"the obligation to avoid negligence, particularly to take reasonable care not to cause physical, economic, or emotional loss or harm to others"*[148]. An assertion of *duty of care* can be equated to binary relevance judgment.

3.3.2 The Concept of Relevance in the Relevance Theory Domain

What do we mean when we talk about 'relevance'?

This question might initially appear to be trivial, but on further examination reveals itself to be a complex subject matter in itself. While the intuitive meaning of relevance might appear quite clear as *"pertaining to the matter at hand"*[153], its meaning depending on the use case differs. Within information science *"[r]elevance is the measure of retrieval performance"*[154] and describes how well presented information fits the user retrieving it. While in communication science relevance of information is defined over its *"positive cognitive effect"*[152], meaning its ability to influence the recipient of the information towards *"to the fulfillment of cognitive functions or goals."*[152].

---

[144] Westhofen, L.: Time To Collision (TTC) - Criticality Metrics (2023).

[145] Hayward, J. C.: NEAR-MISS DETERMINATION THROUGH USE OF A SCALE OF DANGER (1972).

[146] Huber, B. et al.: Evaluation of Virtual Traffic Situations based on Multidimensional Criticality Analysis (2020).

[147] Autey, J. et al.: Safety evaluation of right-turn smart channels using automated traffic conflict analysis. (2012).

[148] Allaby, M.; Park, C.: A dictionary of environment and conservation (2013)

[149] Lyssenko, M. et al.: Towards Task-oriented Relevance Metrics for Pedestrian Detection (2021).

[150] Philion, J. et al.: Learning to Evaluate Perception Models Using Planner-Centric Metrics (2020).

[151] Shalev-Shwartz, S. et al.: On a Formal Model of Safe and Scalable Self-driving Cars (2017).

[152] Sperber, D.; Wilson, D.: Relevance: communication and cognition (2001)

[153] Saracevic, T.: RELEVANCE: A review of and a framework (1975).

[154] Schamber, L. et al.: A re-examination of relevance: toward a dynamic, situational definition∗ (1990).

Due to this ambiguity, further application necessitates more elaboration of the detailed meaning of relevance.

For this purpose, this section will explore the following questions: How is the concept relevance modeled? What influences a given relevance? How can relevance be asserted? Due to the nature of different disparate views on relevance over time, a complete overview of the state of the art would exceed the scope of this thesis. As such, a focus on the most prevalent streams of thought is presented.

Assessing the state of the art, modeling relevance appears to be an ongoing task, with multiple models coexisting and competing. Relevance is generally modeled and understood as a relation between a set number of entities.[155,156] While the same entities are present in several of the models they might serve a different purpose in the respective model.

In a meta-paper[155] Mizzaro summarizes various approaches to model relevance. Building on the findings in the meta-paper, Mizzaro proposes an ontological model for relevance as presented in figure 3-2.[157] Here, relevance is the relation between two generalized entities, the information item and the problem of a user. Both sides of the relation are subject to different levels of abstraction, where the *real information need* within a problem might differ from what the users *perceived information need*. This is then further differentiated by how the user expresses his information need, either in human language or machine language, which is called *request* and *query* respectively.

Considering the *information item*, three subtypes are distinguished: *physical entity* or *document*[158], *information*, and *surrogate*[157]. The first type of *information item* is the *physical entity*, as it is the initially existing entity, from which all other types are derived. The *surrogate* is a representation of a *physical entity*.

From the *physical entity information* can be derived, which is the main type when considering knowledge management and is usually the entity of interest when considering relevance.

Mizzaro states that each combination of *information item* and *problem* constitutes a different relevance.[155] The relevance also changes depending on which component of the information item or problem is considered. These components are decomposed into topic, context, and task. He further includes the concept that relevance is subject to change over time.

[155] Mizzaro, S.: Relevance: The whole history (1997)

[156] Saracevic, T.: RELEVANCE: Part II: nature and manifestations of relevance (2007).

[157] Mizzaro, S.: How many relevances in information retrieval? (1998)

[158] As the focus of Mizzaro's work is mostly retrieval of text documents, he refers to the *physical entity* mostly as *document*. For clarity, the term *physical entity* will be used, in favor of *document*.

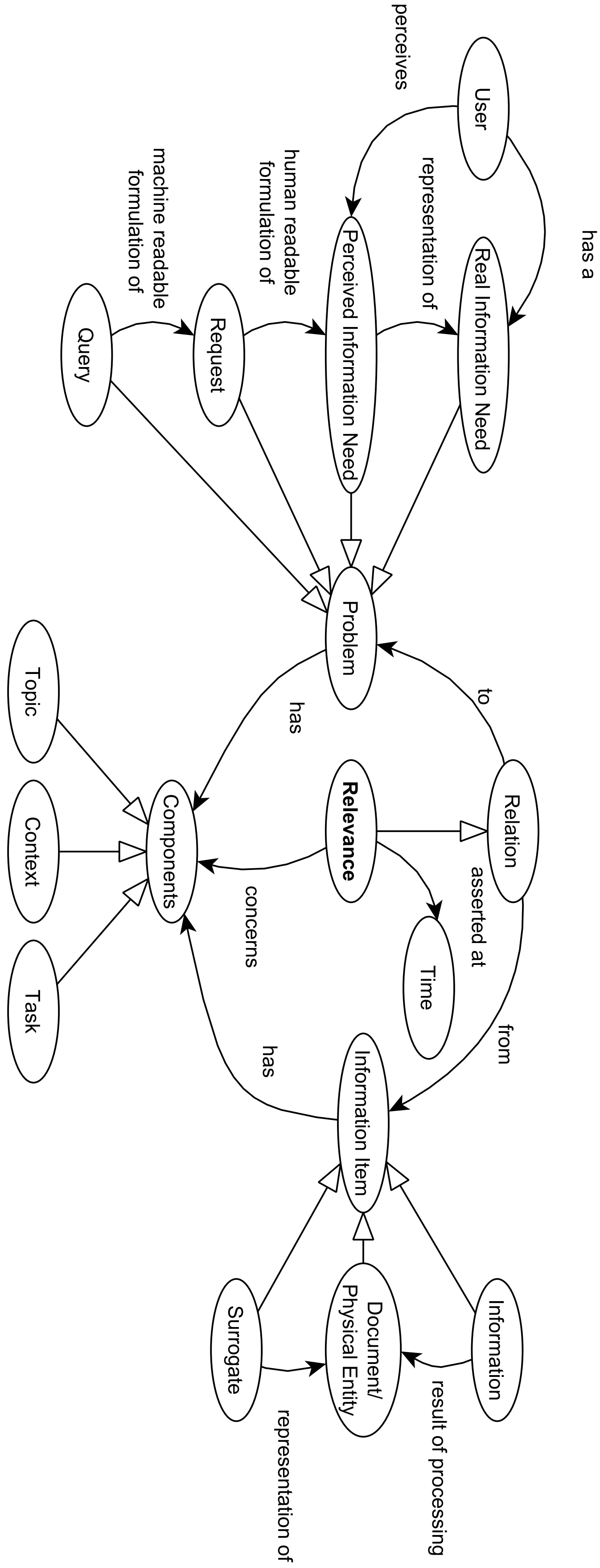

Figure 3-2: Illustration of relevance based on description from Mizzaro [157]

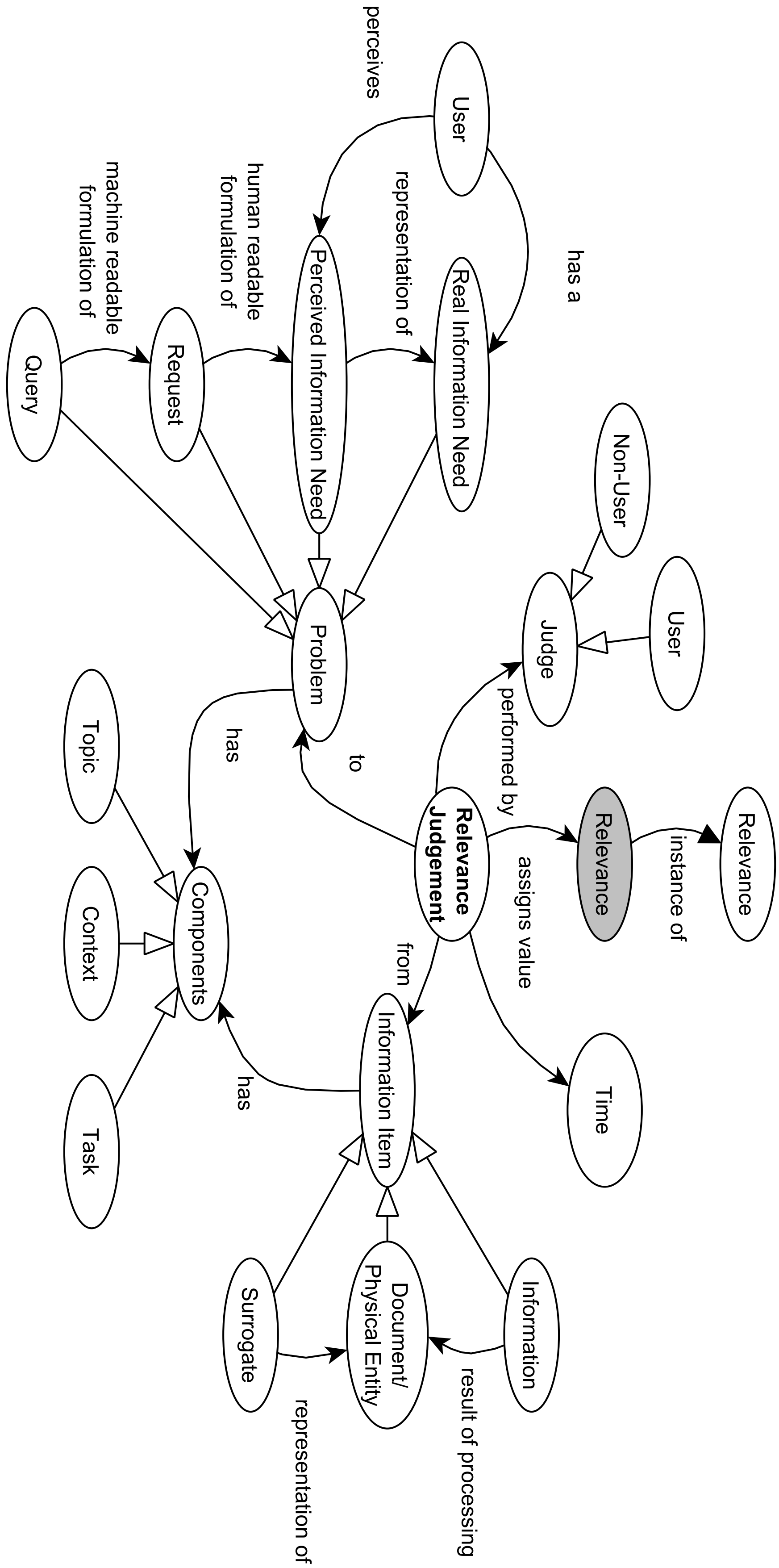

Figure 3-3: Illustration of relevance judgments based on description from Mizzaro [157]

In addition to providing a model for relevance, Mizzaro also summarizes a model for judging relevance, as depicted in figure 3-3. He distinguishes several attributes that influence a relevance judgment. First, the type of relevance being judged. Second, whether the judge is a user or a non-user. Third, what the judge can use to evaluate relevance, both for the problem and the information item. This accounts for the possibility that the relevance of a physical entity to a user's problem is to be judged, while only the query and a surrogate are available. Finally, the time of judgment as opposed to the time of relevance. This accounts for the fact that a judgment may consider past, present, or predicted relevance.

Saracevic's seminal paper[159], which has been widely cited and recognized throughout the broader scientific discourse, presents five key attributes for relevance: Relation, Intention, Context, Inference, and Interaction. Like Mizzaro's connection between the information item and problem, Saracevic's *Relation* denotes the same concept. The other attributes, while not identical to those outlined by Mizzaro, correspond to similar concepts. For instance, *intention* aligns with the task component associated with both the information item and the problem. While Saracevic identifies *context* as a key attribute of relevance, Mizzaro's summary places the concept of *context* together with *task* and the *topic* as a component of information item and problem. Similarly, *Inference*, as conceptualized by Saracevic, refers to the assessment of relevance, often visualized on a spectrum rather than a mere binary evaluation. Lastly, *Interaction* parallels the time component, emphasizing the dynamic relationship between the evolving information item and the user-defined problem.

In general language usage, reference is often made to the concept of relevance. However, a closer look at the current state of the art reveals that relevance is an umbrella term that encompasses a variety of differentiated concepts. The following will highlight several of these concepts to understand the complexity of the term and its implications in different meanings.

A key distinction can be made whether a relevance reflects the user view of the information item. A relevance not considering the user, is denoted in literature by the terms *objective relevance*, *topicality* or *topical relevance*, *logical relevance*, *subject relevance* and *system relevance* or *algorithmic relevance*.[160,161,162,163,164] For brevity, the term *objective relevance* will be used to describe this group. As the name *topical relevance* suggests, this objective relevance considers the topic component. In more detailed definitions it is identified as the relevance of an information item to a request or query for what concerns the topical component or subject.[165] Descriptions

[159] Saracevic, T.: RELEVANCE reconsidered (1996).

[160] Spink, A. et al.: From Highly Relevant to Not Relevant (1998)

[161] Eisenberg, M. B.: Measuring relevance judgments (1988)

[162] Cooper, W. S.: On selecting a measure of retrieval effectiveness (1973)

[163] Cooper, W. S.: On selecting a measure of retrieval effectiveness part II. Implementation of the philosophy (1973)

[164] O'Brien, A.: Relevance as an aid to evaluation in OPACs (1990)

[165] Saracevic, T.: RELEVANCE: Part II: nature and manifestations of relevance (2007)

of system or algorithmic relevance, indicate an understanding that this relevance is rather an implementation, via a system or algorithm, of other relevances.[166]

Similarly, the state of the art presents several terms for relevances, that take aspects involving the user into consideration. While the group is termed *subjective relevance* multiple subgroups of terms, having the same meaning, emerge:[168,160,162,163,164,167].

1. *pertinence* and *cognitive relevance*
2. *usefulness*, *situational relevance*, and *utility*
3. *motivational relevance*, *affective relevance*

Sperber and Wilson identified two key properties for *"subjective relevance"*. First, an effect on the subject's view of the world, and second, how much processing is necessary to obtain the effect based on the information.[152]

The term *pertinence* is the most directly related subjective relevance terms to an objective relevance. It can be understood as the subjective extension of *topical relevance*, additionally taking into account the second of Sperber and Wilson's properties. While an information item might be *topical relevant*, as concerning the topic of interest, it has to be interpretable or usable for the user in order to be considered *pertinent*[169]. The relevance term *usefulness* further adds to the concept of pertinence a consideration of Sperber and Wilson's first property. *"An entity has utility if it is pertinent and makes a useful contribution beyond what the user knew already"*[169]. As such, information has utility if it has a relation to the performance of a task and contributes to the way the task is performed. Examples are *"[u]sefulness in decision making, appropriateness of information in resolution of a problem, reduction of uncertainty"* [165]. While *"affective relevance"* is similar to *utility*, the key difference lies in that it does not concern the *result* of a task, but the motivation of whether a task should be executed [165].

Carbonell and Goldstein introduce the concept of *marginal relevance*[170]. This relevance is defined as a combination of the relevance of an information item and its novelty to the user in a given context. The degree of novelty is defined via the dissimilarity to other information items present in the users' context.

While the nomenclature in use for relevance is more extensive as presented here, this overview is

[166] Saracevic, T.: RELEVANCE reconsidered (1996).

[167] Regazzi, J. J.: Performance measures for information retrieval systems—an experimental approach (1988)

[168] Cosijn, E.; Ingwersen, P.: Dimensions of relevance (2000).

[169] Soergel, D.: Indexing and retrieval performance: The logical evidence (1994) p.590

[170] Carbonell, J.; Goldstein, J.: The Use of MMR, Diversity-Based Reranking for Reordering Documents (1998).

considered sufficient, since it covers the most widely spread terms for relevance.

In addition to general terminology and definitions, literature further identifies several attributes, which influence the magnitude of the assigned relevance. First, considering the judge of the relevance, a judge may be subject to a *bias*, or be restricted to his own *knowledge* and *experience*. The *intelligence* of the judge is further noted as an attribute.[171] When considering an algorithm as the judge, this can be understood as the sophistication of the algorithm.[172]

In addition, the circumstances under which the assessment is made can also influence the degree of relevance attributed. Factors such as the *amount of information* available for assessment, the *order of presentation* of the information, the *specification of the task* at hand, and the *time of the judgment* attribute to determining relevance.[171]

Regarding the *information item*, the quality of the information contents like its *accuracy*, *clarity*, *recency*, and absence of *ambiguity* are noted. Further, the information's *completeness* to the *information need* is named. From a utility perspective, the *usefulness* of the information, especially in task resolution, stands out. For instance, in a document search, the *suggestiveness* of the document to other documents also influences relevance. In addition to the content and effect of the information, the process of extracting information from a physical entity is also considered. Here, *ease of use*, *difficulty*, *time*, *cost*, *efficiency* as well as *accessibility* are given as possible attributes.[167,171,172,173,174]

Considering the *information item* the following influence to relevance attributed to it. First, the *accuracy* or lack of *ambiguity* of the presented information has to be considered. Additionally, whether the information has *completeness* and *recency* for the information need, is identified as an influencing factor. Evaluating the utility side of the information object, the *usefulness* in solving a task, is a key factor. Considering a document search, the *suggestiveness* of the document to other documents also influences relevance.[171,152,172,173]

As previously noted, the relevance of individual *information items* can also be assessed in conjunction with other present *information items*. Here, the concept of *novelty* of the *information* in contrast to already known *information* is considered. Regarding the above-mentioned circumstance that information can be inaccurate or ambiguous, the ability to alleviate these situations is also conceptualized as *noise reduction*.[172,173,174]

[171] Schamber, L.: Relevance and Information Behavior. (1994)

[172] Taylor, A. et al.: Relationships between categories of relevance criteria and stage in task completion (2007)

[173] Vakkari, P.; Hakala, N.: Changes in Relevance Criteria and Problem Stages in Task Performance (2000)

[174] Taylor, R. S.: Value-added processes in information systems (1986)

3.3.3 Validation of Relevance Models

The contents of this subsection have been previously published as part of a conference paper[175]. The state of the art shows various methods on how to conceptualize and assert object relevance. The validation of these relevance concepts however has not been the focus of research and is therefore underexplored.

Most datasets do not consider the validity of inclusion or exclusion of objects, with respect to safety in planning or for evaluation of perception metrics.[176,177,178] In other datasets, considerations can be found that the validity of inclusion may be less for some objects than other. As an example, the ONCE dataset[179] considers objects at different distances separately. A common approach is to show plausibility through application and visualization of examples.[180,181,182] Other works propose the usage of planners for evaluation in closed-loop simulations [183] and also apply it to safety aware prediction metrics[184]. For the planning Kullback-Leibler divergence (PKL) metric [183], validity is argued in two steps. First, an argument for plausibility on the basis of showing sensitivity to distance and velocity as commonly accepted features of relevance for the automotive domain. Second, the results are subjected to human comparative evaluation via Amazon Mechanical Turk. This evaluation showed a human preference towards the PKL metric over the NuScenes Detection Score by 80%. However, the use of a planner introduces a tight coupling to the specific implementation and also to the errors within the implementation. As such, the results generated by this type of method are questionable in their validity.[185]

3.4 Interim Conclusion

Having reviewed the works related to this dissertation an intermediate conclusion on the current state of the art can be made. Here, two key points emerge. First, the amount of data is an essential factor in increasing the performance of DNNs and larger amounts of data are needed. Second, one of the main challenges for the release of automated driving is controlling the open context.

[175] Storms, K. et al.: SURE-Val: Safe Urban Relevance Extension and Validation (2023).
[176] Geiger, A. et al.: Are we ready for Autonomous Driving? (2012).
[177] Sun, P. et al.: Scalability in Perception (2019).
[178] Wilson, B. et al.: Argoverse 2 (2021).
[179] Mao, J. et al.: One Million Scenes for Autonomous Driving: ONCE Dataset (2021).
[180] Topan, S. et al.: Interaction-Dynamics-Aware Perception Zones for Obstacle Detection Safety Evaluation (2022).
[181] Volk, G. et al.: A Comprehensive Safety Metric (2020).
[182] Philipp, R. et al.: Systematization of Relevant Road Users (2022).
[183] Topan, S. et al.: Interaction-Dynamics-Aware Perception Zones for Obstacle Detection Safety Evaluation (2022)
[184] Jha, S. et al.: Watch Out for the Safety-Threatening Actors: Proactively Mitigating Safety Hazards (2022).
[185] Mao, J. et al.: 3D Object Detection for Autonomous Driving (2023).

As such, there is a need for additional options in data reduction, to further reduce the burden of provisioning large amounts of data. The state of the art showed many diverse methodologies for data reduction, while a lack of methods for partial, intra-scene data reduction methods is apparent. Furthermore, based on current data centric methods, it is necessary to not only consider the open context in safety argumentation, but also at the provisioning stage and hence also in data reduction.

It can be concluded, that a novel intra-scene data reduction methodology, that explicitly considers the open context is needed for an efficient development and release of automated driving.

As stated before, a relevance centric approach has the potential to mitigate the challenges posed by the open context. Reviewing the state of the art for relevance, many concepts emerge. Despite the abundance of theoretical frameworks, a unified model to conceptualize, assess, or assign relevance within the domain has not yet been presented. It can be assumed that a relevance centric way of thinking must not just be transferred to the automotive domain, but rather rebuilt based on the building blocks provided by that state of the art from relevance theory and information retrieval.

Summarizing the state of the art of relevance validation a distinct lack in methods and substantiation thereof must be noted. Proposed methods argue their validity by comparing chosen examples and performance relative to baselines. As such, no method exists which provides acceptance criteria and for which validity can be assured.

4 Methodology

The previous chapter concluded a need for a novel data reduction methodology for an efficient development and release of AD. Two aspects of this methodology were deduced from the state of the art. First, the need for novel intra-scene data reduction methodology, and second, an explicit consideration of the challenge posed by the open context for AD. On the one hand, while the intra-scene aspect is a novelty in the state of the art, it is only another data reduction possibility among others. The consideration of the open context on the other hand is something that has no corresponding state of the art. As such, the chapter will focus on the second aspect, which is solidified with the following research question.

Research Question 1
How can the open context present in automated driving be addressed in data reduction?

To provide an answer to this research question, a concept for such a data reduction methodology will be derived. Since the consideration of the open context is the key aspect of the data reduction, the name of Context Aware Data Reduction (CADR) is assigned to it. In the initial step, functional requirements will be expanded from the above-mentioned aspects on a functional level. Subsequently, based on these functional requirements a system design for CADR will be defined.

4.1 Functional Specification for Context-Aware Data Reduction

Data reduction methods rely fundamentally on relevance models. Such a model serves as the backbone for decision-making within a data reduction method, delineating which pieces of information are relevant and must be retained versus those that are irrelevant and can be discarded. Additionally, the relevance model dictates the distribution of data, determining how much data storage should be allocated to each piece of information based on its relevance. Therefore, the CADR approach requires the usage of a relevance model to address the specific needs and considerations intrinsic to automated vehicles.

Traceability is defined by ISO/IEC/IEEE 24765 as the *"discernible association among two or more logical entities, such as requirements, system elements, verifications, or tasks"*[186a]. ISO/DIS

[186] ISO: Systems and software engineering - Vocabulary (2017)[a: p.378, b: p.236]

21448, ISO 26262, and ISO/TR 4804 rely on the concept of traceability.[187,188,189] ISO 26262 places traceability as a key component of a safety case. Specifically ISO 26262-2 states that work products referenced in the safety case must be traceable from one to another.[190] Data especially from field operation provides *"objective evidence"*[186b] for a given safety case. As such, data and manipulation thereof also need to be traceable. This is formalized in the following requirement (REQ):

REQ: CADR shall enable traceability.

Many data reduction approaches only use implicit relevance models. An implicit relevance model is specified only by the algorithmic implementation of the approach, but is not otherwise documented. This means that the relevance model can only be derived by analyzing the implementation and evaluating the results of the data reduction. Such relevance models, due to their implicit and unstated nature, cannot be traced and linked to previous or subsequent processes. Thus, they prevent a holistic view of the system development process and complicate the creation of a safety case. It is then argued that any use of relevance in processing or reasoning must make the usage of the applied relevance model available for use by other processes. These features are only possible with an explicit description of a relevance concept. Thus, the following functional requirement applies:

REQ: CADR shall use an explicit relevance model.

A principal component of the system is its capacity to handle an open context, an ODD that cannot be entirely structured or exhaustively defined by existing ontologies. An open context is characterized by uncertainties and incomplete structures that defy full ontological categorization. These uncertainties and incomplete structures have the potential to cause a false assertion of relevance and subsequent data reduction. To account for the open context, this potential needs to be handled by the CADR approach, giving the functional requirement that:

REQ: CADR shall include an assurance concept.

The system's ability to function within an open context is supported by its adaptability to the evolving nature of the ODD. The structure of the ODD, at any given moment, is recognized as being incomplete, with the possibility of expansion as new facets are described and uncertainties are clarified. This inherent flexibility is essential, as the act of data reduction, predicated on the relevance model, could inadvertently lead to the loss of information that might later prove to be significant.

To mitigate the risks associated with information degradation through successive applications of data reduction, the system shall incorporate a feature termed *continuous extendability*. This

[187] International Standardization Organization: Road vehicles - Functional safety (2018).

[188] ISO/DIS: Road vehicles - Safety of the intended functionality (2021).

[189] ISO: Road vehicles — Safety and cybersecurity for automated driving systems (2020).

[190] International Standardization Organization: Road vehicles - Functional safety - Part 2 (2018), p. 39.

property ensures that information, which has not yet been deemed irrelevant, is not permanently discarded. Rather, the original, unreduced datasets are always accessible for re-evaluation. Continuous extendability allows for the ongoing refinement of the ODD, the reduction of uncertainty, and the iterative enhancement of the relevance model, as shown in figure 4-1.

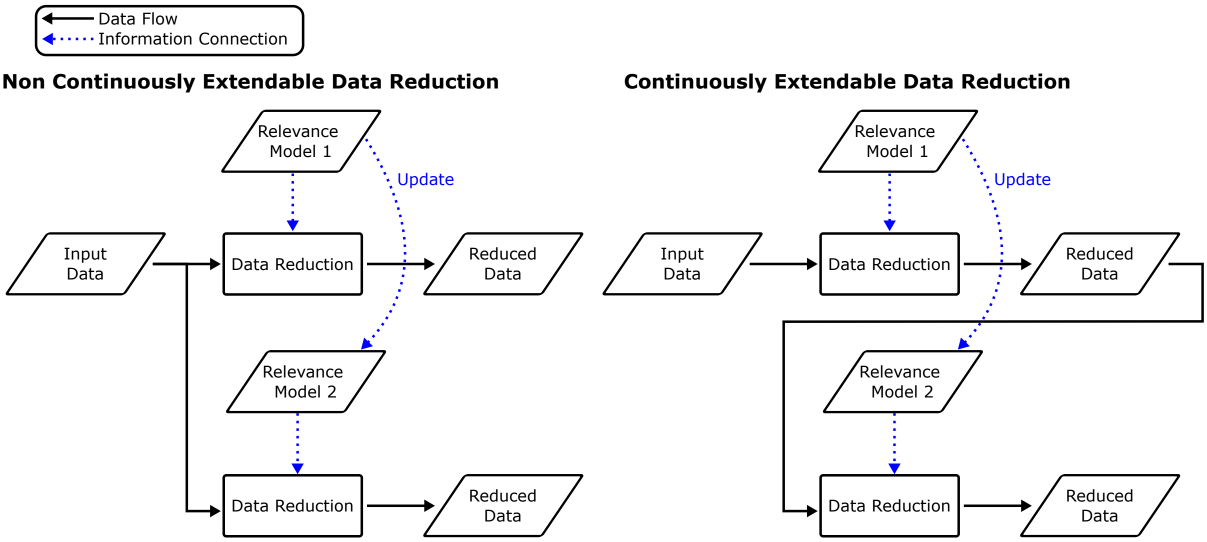

Figure 4-1: The property of continuous extendability of a data reduction upon updating the relevance model.

REQ: CADR shall provide continuous extendability.

In conclusion, the functional specification for CADR sets forth a structured methodology for developing a data reduction system that is not only aware of the complexities of automated driving, but is also robust and flexible enough to evolve with the expanding knowledge and understanding of the ODD. This approach is conceptualized to ensure that ADS remain both efficient and effective in the face of the dynamic and uncertain nature of real-world driving environments.

4.2 System Design for Context-Aware Data Reduction

Based on the previously defined functional requirements a framework system is designed for further implementations of the CADR approach. The proposed system is depicted in figure 4-2. Following, the individual modules comprised in the system will be described in more detail.

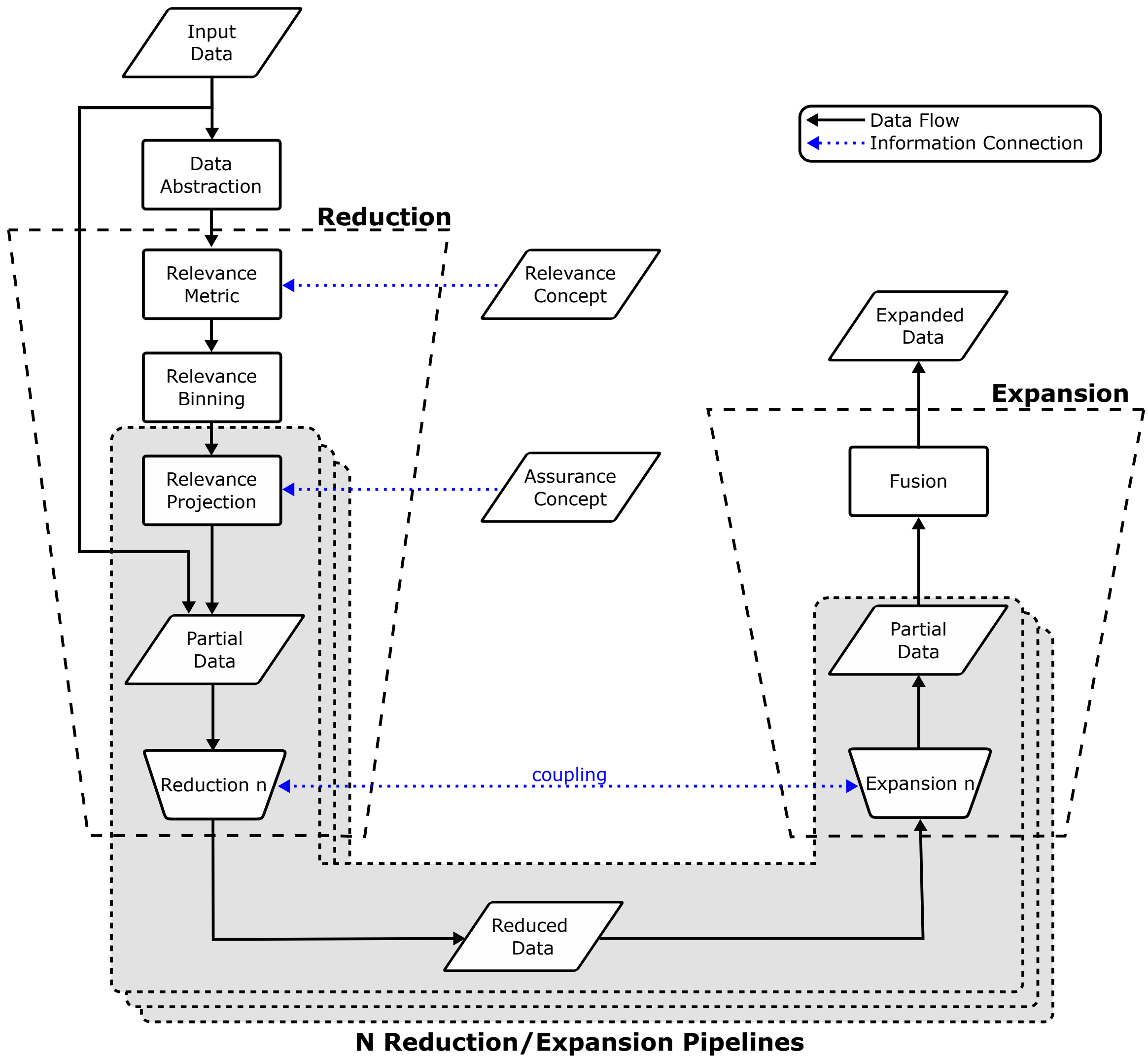

Figure 4-2: Flowchart for the CADR architecture

Data Abstraction

The data abstraction module, as depicted in figure 4-3, transfers the input data to the abstraction level of the used relevance model. While the representation of a relevance model is an ontology in nature, the representation is not confined to ontological representation formats. Contrary the abstracted form can be any arbitrary representation, with the only constraint being, that the represented entities and properties are semantically equivalent to those used in the relevance model. This semantic equivalence then allows relevance judgments to be made. In ADS, perception modules provide a parametric description of the environment that concisely represents what has been deemed important. These modules can be seen as the first part of the data abstraction process.

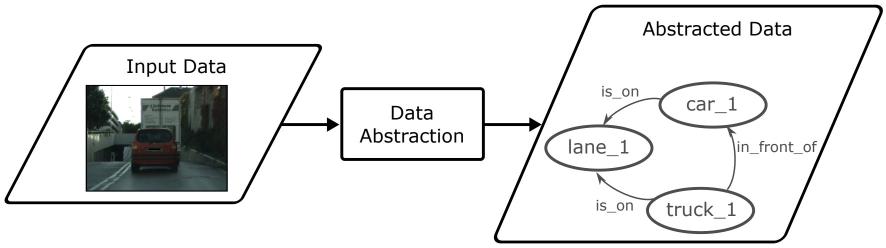

Figure 4-3: Data Abstraction

Relevance Attribution

The relevance attribution module, shown in figure 4-4 encompasses one or multiple relevance metrics. Based on the relevance metrics each ontological element represented by the relevance model is attributed a value on the used relevance scale.

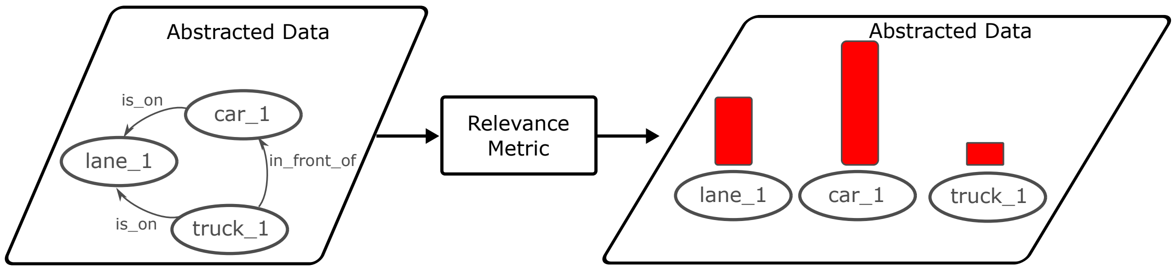

Figure 4-4: Relevance Metric

Relevance Binning

The task of the relevance binning component is to divide the relevance measure of the used metric into N intervals of relevance. These are then assigned to the corresponding N Reduction/Expansion pipelines.

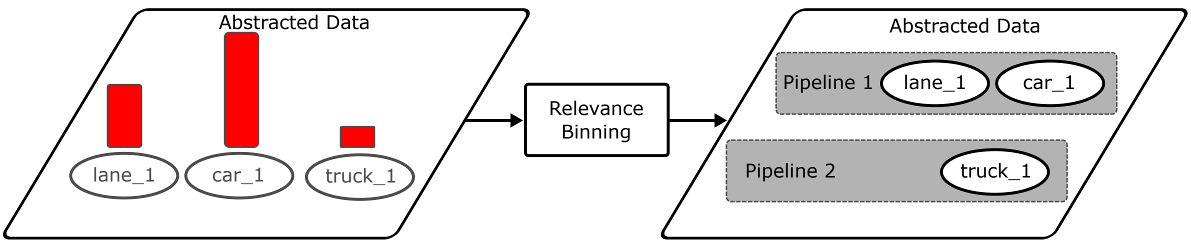

Figure 4-5: Relevance Binning

Relevance Projection

Within the relevance projection module, the relevance judgments of each bin are projected to the representation space of the input data. In the case of image data, this results in a pixel mask I_{mask}, denoting which pixel corresponds to the respective relevance bin.

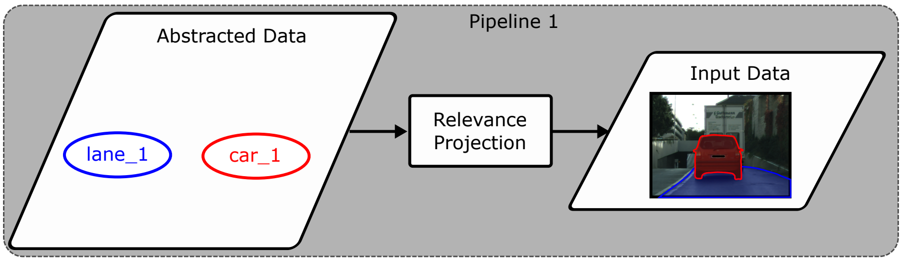

Figure 4-6: Relevance Projection

Reduction/Expansion Pipeline

The framework comprises N reduction/expansion pipelines, with each pipeline integrating a reduction module and an expansion module. These modules are responsible for implementing an arbitrary data reduction algorithm. The variation among pipelines can be subtle, such as a change in parameters, or more pronounced with different reduction/expansion algorithms being employed.

The design of each pipeline shows a fundamental coupling between the reduction and expansion functions. This is primarily because the expansion module is tasked with performing the inverse operation of the reduction algorithm. For instances of lossless reduction, the expansion function $f_{\text{expansion},n}$ is the mathematical inverse of the reduction function $f_{\text{reduction},n}$, as defined by:

$$f_{\text{expansion},n} = f^{-1}_{\text{reduction},n} \tag{4-1}$$

In scenarios involving lossy reduction, the expansion function seeks to approximate the inverse of the reduction function, introducing an error ϵ, described as:

$$f_{\text{expansion},n} + \epsilon = f^{-1}_{\text{reduction},n} \tag{4-2}$$

Previously, relevance was solely associated with an arbitrary external task. However, the approach is extended to incorporate relevance as an integral aspect of the data reduction and expansion task itself. The magnitude of the expansion error ϵ resulting from lossy reductions is contingent upon the expansion algorithm's implementation, its parameters, and the input data characteristics.

It is observed that the magnitude of the expansion error can differ based on the input. To address this sensitivity, the concept of relevance to the data expansion task is introduced. This concept enables the methodology to identify which input data elements disproportionately affect the expansion error. With this knowledge, the reduction process can be refined to preemptively preserve data that would otherwise contribute excessively to the error upon expansion, thus enhancing the accuracy and reliability of the data processing pipeline.

Fusion

The fusion module's task is to take in the expanded outputs from each relevance pipeline and combine them into the final expanded data. This can be as easy as a simple merge operation, as shown in figure 4-7. If multiple lossy data reduction pipelines are applied, other considerations might apply making this step more complex.

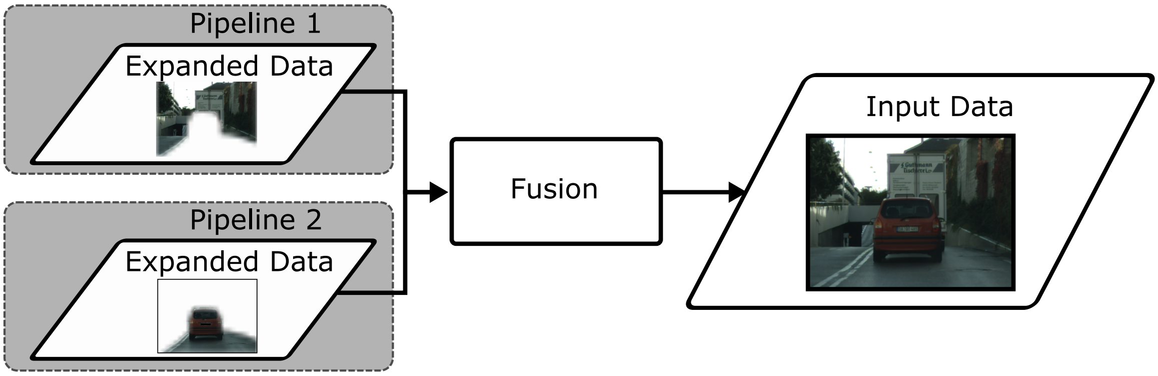

Figure 4-7: Fusion

4.3 Assurance Concept

As stated in section 4.1 the CADR approach has to incorporate an assurance concept to ensure that important information is not lost. This concept is especially vital when a relevance concept may be incorrectly applied, which can lead to a break in the traceability for the safety argumentation through data loss. To address and mitigate this risk, the relevance concept used should be conservative. This means that in situations of uncertainty, where data could be seen as both relevant and irrelevant, the conservative approach would default to considering the entity as relevant. Such a stance ensures that any data that might be pertinent is not discarded.

Moreover, a relevance concept might fail to consider all relevant aspects of the subject system. To prevent this, instead of creating a single, holistic relevance concept, a more fine grained approach can be used. For this case, a functional decomposition[191] can be applied. The subject system is broken down into its individual sub functions, and a conservative relevance concept is applied to each. These individual concepts are then united using a logical disjunction. Figure 4-8 visualizes this concept.

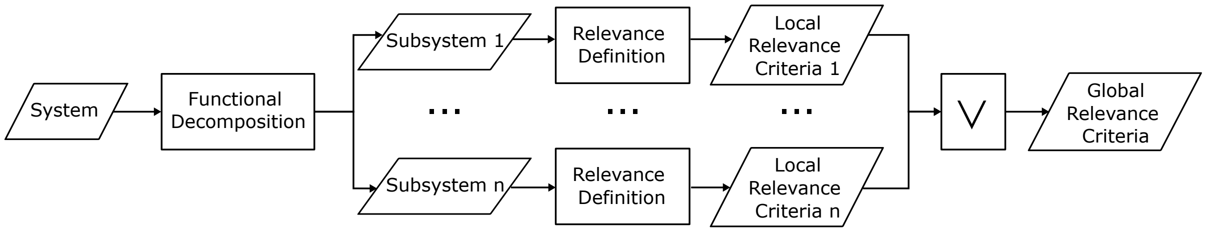

Figure 4-8: Functional Decomposition

In practice, this approach dictates that if any one sub function determines a piece of information to be relevant, then it is then deemed relevant to the entire system. This layered approach to relevance helps to maintain the integrity of the system's functionality by ensuring comprehensive data consideration. The utilization of functional decomposition in conjunction with a conservative approach is expected to yield sufficient assurance for the application of the CADR approach.

[191] Amersbach, C.: Functional Decomposition Approach (2019).

5 Relevance Concept

Chapter 4 defined the need for an explicit relevance model, as to be able to link it to other engineering processes. Thus, RQ. 2 is posed to answer how this need can be satisfied.

> **Research Question 2**
> *How can the concept of relevance, in the context of a given use case, be formally described?*

In chapter 2, ontologies were already identified as a suitable tool to structure and formalize information and knowledge. Ontologies are widely used to describe information in the automotive domain to describe and reason about information from the traffic environment. Since relevance is primarily thought of as a relation, describing relevance information involves entities that describe and link to these environment information. In order to reflect this characteristic, a nested ontological architecture as depicted in figure 5-1 is proposed. As described in section 2.1, an ontology contains two types of information. The first is the terminology box (T-Box), which states the concepts and axioms that serve as ground truth for how the described world is structured. The second, called assertion box (A-Box), states or asserts the observations made in that world. Here, the entire environment ontology is treated as an assertion described within the relevance ontology. Thus, the relevance ontology contains the used abstract concepts about the environment as well as concrete instances of a current situation. This allows the relevance ontology to reason about the implications and correctness of the environment ontologies A-Box as well as T-Box.

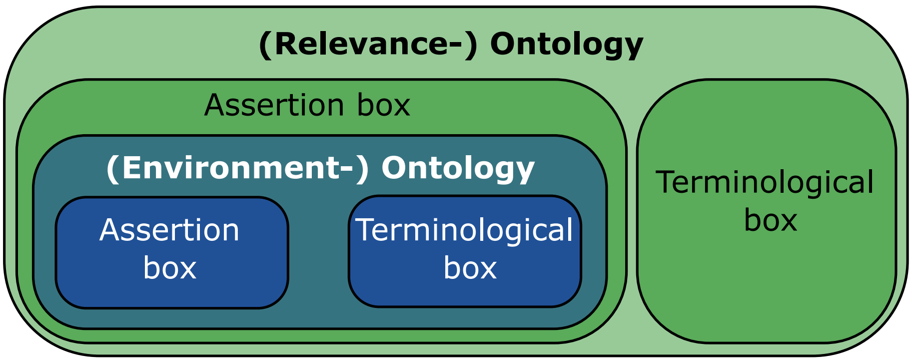

Figure 5-1: Nested ontological architecture of relevance and environment ontology

In the following, two types of models will be introduced. First, an abstract description of classes and relations within the relevance ontologies T-Box is modeled. This model is designated as general relevance model (GRM). The GRM is intended to serve as a system agnostic template which can then be applied and further specified for different target systems. These specifications are designated as specific relevance model (SRM). For a given system, the SRM defines concrete instances for the classes defined in the GRM. While the SRM can be as detailed as possible a full specification is not necessarily beneficial or necessary. Looking at scenario description as a potential application for an environment ontology, multiple levels of abstraction/concretization have been established over time, ranging from concrete scenarios to functional scenarios.[192,193] While the former gives a machine-readable specification of a single scenario with fixed parameters, the latter only gives a human-readable behavior based description for a range of scenarios. As to not be limited to a single set of fixed parameters, or a single implementation of a system, allowing for partial specifications allows for a more generalized applicability of a SRM.

5.1 Definition of a GRM

This section will consider which types of relevance are useful for the current challenges of AD development. The identified relevances will be represented in a model. This model may not reflect further uses for relevance not considered or not yet known. Thus, the presented model may require extension or adaptation to be applicable to these novel uses. Using the notation given in section 2.1 the proposed GRM is visualized in figure 5-2.

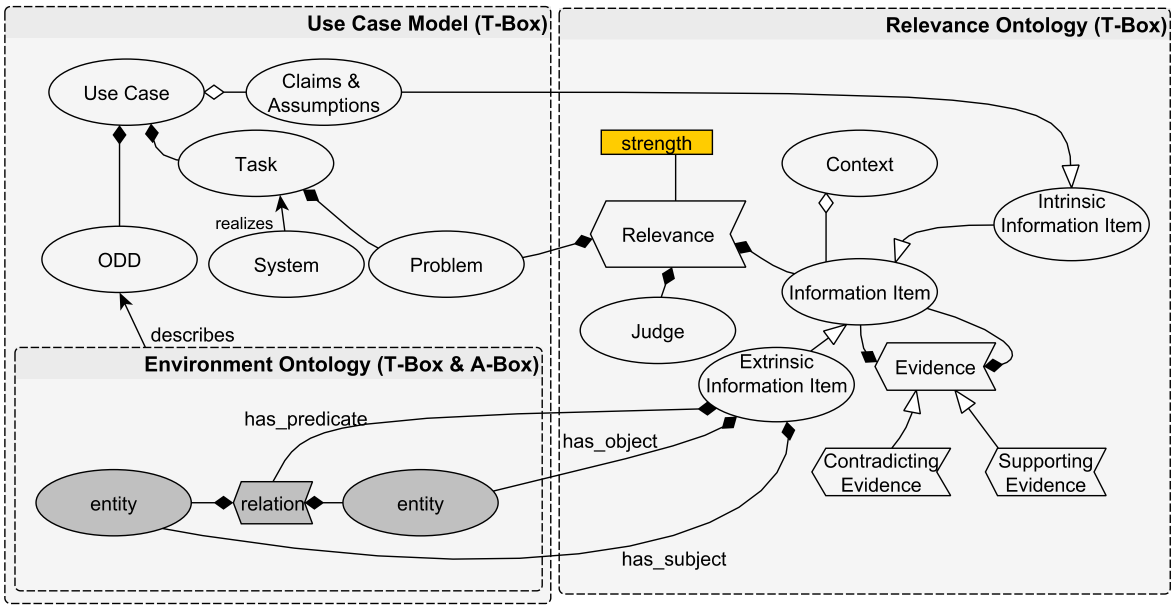

Figure 5-2: Ontology of the concepts that make up the GRM

192 Neurohr, C. et al.: Criticality Analysis for the Verification and Validation of Automated Vehicles (2021).

193 Menzel, T. et al.: Scenarios for Development, Test and Validation of Automated Vehicles (2018).

To ensure the ability to reason about the relevance aspect of relations in ontological structures, the GRM needs to reflect the above nested ontological architecture. The environment ontologies A-Box only contains instances of the environment ontologies T-Box. Thus, it is necessary to first define the entities of the environment ontologies T-Box. Subsequently, the entities within the environments A-Box can be instantiated by concrete observations from the environment scene. Two classes within the environment T-Box are predefined as the interface between the environment ontology and the relevance ontology, these are entity and relation.

The rest of the environment T-Box can only be defined once the subject use case is known. The use case entails two aspects: what needs to be done and in which environment does it need to be done? To cover the first aspect, a description of the subject task is required. In addition, the GRM must represent a system that implements a solution to the task. The second aspect requires a description of the ODD. Describing the ODD and creating a system to solve a task usually requires models as simplifications of reality. The claims and assumptions of the models are represented in the GRM as a type of *information item*. Since these *information items* are not sensed or derived from the external world during operation, they are denoted as *intrinsic information items*. Conversely, *information items* observed within the environment are denoted as *extrinsic information items*. *Extrinsic information items* are conceptualized as a relation between two entities within the environment ontology. From the point of view of relevance ontologies, this relation is described as a composition of subject, predicate, and object. This allows the representation of both information elements from the environments T-Box, as in table 5-1, and A-Box, as in table 5-2.

Table 5-1: Examples of extrinsic information items from the environments T-Box. Here, the terminology is introduced, that both road and car are subclasses of thing is stated, as well as that car can have an is_on relation to road.

Information Item	Entity	Relation	Entity
T1	car	is_on	road
T2	road	is_a	thing
T3	car	is_a	thing

Table 5-2: Examples of extrinsic information items from the environments A-Box. In this example, it is asserted that there is an entity ”car_1” that is of type car and another ”road_1” that is of type road. Further, the relation that car_1 is_on road_1 is asserted.

Information Item	Entity	Relation	Entity
A1	car_1	is_instance_of	car
A2	road_1	is_instance_of	road
A3	car_1	is_on	road_1

To further visualize this example with figure 5-3, information items T3 and A3 from tables 5-1 and 5-2 are depicted in the context of the nested ontological architecture established in figure 5-1. As can be seen, the purpose of the environment ontology is to describe the observed environment through assertions, as well as the concepts used to describe the environment. In contrast, the relevance ontology is not directly concerned with the environment. Its purpose is to describe the informational workings of the environment ontology. Using this example, a result of the relevance ontology might be to reason about, the is_on relation. Assuming that the environment ontology requires every car to be on a road, observing cars that are not on any road can be used to refine the environment ontology.

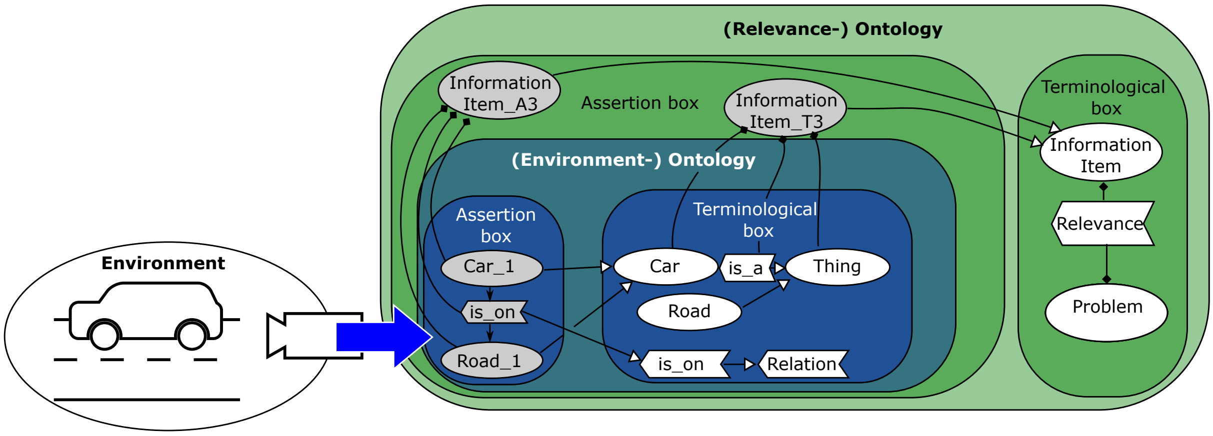

Figure 5-3: Example for the application and use of the nested ontological architecture of relevance and environment ontology

Regarding the task, it can be considered as the requirement to perform an action within the environment as defined in the ODD. In the scope of AD, the task is considered to require a closed-loop interaction, such that it can also react to changes within the environment. As such, the task requires a sense step before a reaction can be rendered. In terms of relevance theory, this sense step represents an *information need* for a *problem*.

As established in the review of the state of the art, relevance within the GRM is also conceptualized as the relation between a *problem* and an *information item*. Similar to criticality metrics the concept that relevance can be expressed on different scales is also included.

However, it is crucial to include the role of a *judge* in determining relevance, since relevance can also be judged for an object that is not the ego object. This aspect is not adequately captured in traditional criticality metrics. Furthermore, two additional concepts are integrated to enhance the application of relevance in automated driving. The significance of an information item may be diminished or even nullified if it is redundant to another information item, a concept termed *marginal relevance*. To address this, the GRM incorporates the notion of *redundancy* by compiling all current information items into what is termed as the *(information) context*. Within the scope of the GRM, the context remains an abstract concept to be later filled by concrete *information items*.

Section 3.1 highlights that the *open context* is a significant challenge for AD, as identified in the state of the art. An *open context* is marked by uncertainty and an incomplete description of the ODD. Therefore, these challenges must be reflected in a relevance concept tailored for AD. This is achieved by introducing an *evidence* relation into the relevance ontology. Evidence can manifest as either a contradiction or support between two information items. This approach allows for the representation of uncertainty in the ODD by identifying contradictions in the observed environment's A-Box. For instance, a potential class mismatch in late sensor fusion, where different sensors may classify the same object differently, exemplifies this. Additionally, a discrepancy between reality and the specified ODD can be detected by observing contradictions between the environment's T-Box and A-Box. For example, if the T-Box assumes that every car is on a lane, but in the hypothetical A-Box, a car is observed without an associated lane, this contradiction indicates that the model assumption in the T-Box is flawed and needs refinement. These thus state a relevance to the development use case and, by extension, to the use case of safety assurance.

5.2 Definition of a Method to Derive a SRM

The contents of this section have previously been published in separate papers[194,195] co-authored by the author of this work together with Ken Mori, and are presented here with modifications towards the application within this dissertation. An analogous approach can be found in the dissertation of Ken Mori[196].

Having defined the GRM which encapsulates the abstract notion of what influences relevance, a next step is necessary. The GRM only considers abstract entities without any concretization, as such it does not provide a relevance concept for a given system that can then be used in a data-driven process. Therefore, a method is needed to perform this transfer from abstract to concrete, which this section will present. Figure 5-4 shows the proposed method.

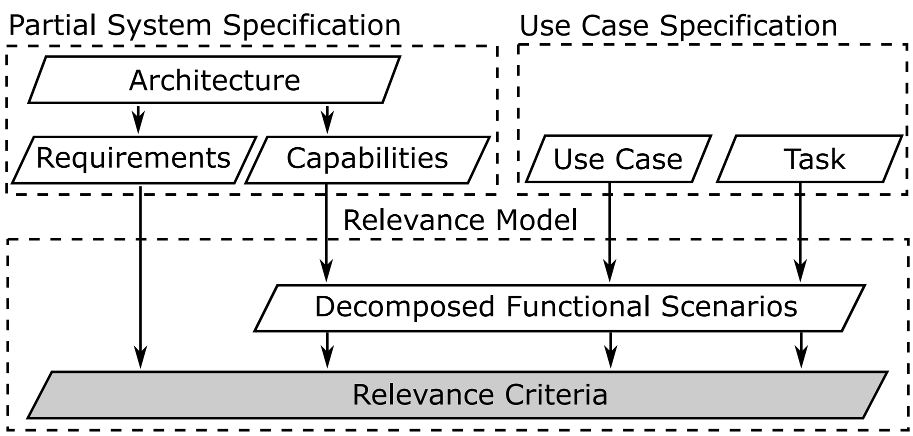

Figure 5-4: Method for deriving a SRM

[194] Mori, K. et al.: Conservative Estimation of Perception Relevance of Dynamic Objects (2023).

[195] Storms, K. et al.: SURE-Val: Safe Urban Relevance Extension and Validation (2023).

[196] Mori, K.: Defining Object Detection Requirements for Safe Automated Driving (2023).

The method consists of the following three major steps. Definition of the problem through the use case definition. Specification of the partial system, defining the action space for solving the problem. The final step is relevance modeling, where the use case is decomposed, behavioral requirements are derived and described through criteria.

5.2.1 Use Case Specification

The first step is the specification of the use case. Here, the task to be performed is defined as well as the ODD. These definitions are on a functional level, as only a human-readable textual definition is required at this point. Additionally, claims and assumptions are specified as intrinsic information items. The claims and assumptions include normative statements and axioms. These can be legal texts or representations of the laws of physics. Together with the description of the ODD, these statements form the T-Box of the environment ontology. The task and the ODD are sufficient to infer the topicality of an information item.

5.2.2 Partial System Specification

As the second step, a partial system specification is performed, yielding a system architecture that fulfills the aforementioned. The partial aspect of the system specification refers to the circumstance that only the interfaces are required to be specified, as they are the elements linking the system to the ODD, as well as the claims and assumptions. The system architecture includes functional requirements as well as a description of the capabilities of the system with respect to the task. While the input interface of the system must match the information contained within the ODD, the output interface is defined by the capabilities. The capabilities enable the system to perform actions that interact with the environment. Similar to a logical scenario that defines the ranges of possible parameters, the capabilities define the ranges of possible actions. Combining all capabilities yields the action space of the system. The action space is the key component for inferring the utility of an information item.

5.2.3 Relevance Model

For a given use case, multiple different contexts within the use case can exist. In each of these contexts, different behavioral requirements apply to the use case. To derive these behavioral requirements, first, the use case needs to be decomposed into functional scenarios with respect to the subject task. This approach has previously been used in various different methods[198,199,200,201,197].

[197] Schönemann, V. et al.: Maneuver-based Adaptive Safety Zone for Valet Parking (2019)

[198] Shalev-Shwartz, S. et al.: On a Formal Model of Safe and Scalable Self-driving Cars (2017).

[199] PEGASUS Project: PEGASUS Method: An Overview (2019).

[200] Philipp, R. et al.: Systematization of Relevant Road Users (2022).

[201] Vater, L. et al.: Test Case Selection Method for the Verification of Automated Driving Systems (2021).

For each identified functional scenario, specific behavioral requirements can then be derived from the high level requirements. This enables further specification of relevance criteria which are formalized as analytic equations. These criteria are solely dependent on a set of parameters describing the current situation.

Uncertainty of the exact value of a parameter causes ambiguity of the context. Therefore, possible high uncertainty of parameters needs to be explicitly considered. To this end, three principles are proposed:

1. Consideration of the worst-case.
2. Superposition of potential values/scenarios, when no single worst-case cannot be a priori established.
3. Argumentation for a minimum acceptable value, if a worst-case exists but yields not feasible solutions.

The first principle of only evaluating the worst-case is quite straightforward and thus does not need further elaboration. The second principle excludes the explicit parameter from the relevance criteria, if not single, but only multiple local worst-cases are known. Instead, the parameter is implicitly considered by superposition of multiple functional scenarios which cover all potential local worst-case values of the parameter. This approach was previously used by Schönemann et. al[197]. A common example in the domain of AD is the road. Considering only a small arbitrary subset consisting of only two elements, a roundabout, and an intersection, a complex interaction within these road environments has to be considered. Thus, depending on the task, it is not obvious whether a roundabout or an intersection is the worst-case. Both the roundabout and the intersection are considered separately and the results are then merged into a single result.

Application of the first and second principles is not feasible if a global or local worst-case assumption leads to a collapse of the relevance criteria. A collapse of the relevance criteria describes the state where the relevance judgment always defaults to either relevant or irrelevant, irrespective of all other influencing factors. In the automotive context, a road friction coefficient of zero or an infinite reaction time are examples of this phenomenon. Assuming these values is equivalent to reducing the action space of the object to zero. In such a case, the third principle leaves the parameter as a variable to be set not by measurement or estimation, but to be defined as the result of an argumentation. Here, the parameter is assumed to have the lowest acceptable value within a confidence threshold. This confidence threshold must then be considered in all subsequent applications of the resulting relevance judgments. The parameter is thus considered to be guaranteed to the relevance model by outside considerations. Such an outside consideration might be limiting the ODD to where no more extreme values can occur or constraining the system in order to always keep a sufficient reserve to guarantee the compliance of the value.

Given the previously defined functional scenarios, behavioral requirements need to be derived. These behavioral requirements, add further detail to the high level requirements from the use case. To derive behavioral requirements normative intrinsic information items are evaluated for normative statements. If applicable to the functional scenario, the normative statement of the intrinsic information item is translated into a behavioral requirement within the scope of the functional scenario. Criteria are then defined which determine the adherence to these behavioral requirements within the functional scenario.

5.3 Definition of a Validation Concept

The last section presented a method to define a SRM. This section will consider how the validity of a given SRM can be substantiated. Two main approaches for ensuring the validity of relevance models exist. The first and most basic approach is empirical validation post usage in the use case. This sets a baseline for future reference. The second approach is to demonstrate validity from a previously established baseline.

In this thesis, the validation of the relevance model is conceptualized as a comparison against a known valid baseline. As previously established, validation of relevance models is currently underexplored in literature. Thus, the only baseline for the validity of relevance judgments is considered to be human perception. While no human relevance judgment models exist that are applicable to the used environment representation, human behavior models exist. Since the behavior is directly influenced by what is perceived, it can be used to examine, which objects a human perception judges as relevant. In order to extract this human relevance judgment for an OOI, the human behavior needs to be examined once with the OOI being perceived for behavior planning and once with the OOI absent.

Under ideal conditions the influence of the absence of the OOI is studied in driving simulators with human test subjects. To avoid the large costs associated with establishing a human baseline via driving simulator studies, an approximation of the human baseline is proposed. The approximation of the human behavior can be done by using a motion prediction algorithm trained on human behavior. A similar approach has been previously used by Philion et. al[202], where a planning algorithm was used instead of a motion prediction algorithm. The usage of a motion prediction of a planning algorithm has several advantages. While the objective of a planning algorithm does not have an unambiguous single target, the motion prediction task has only one target state. That is, there are multiple trajectories that can be planned to fulfill a use case, but only one trajectory that can be predicted that matches the ground truth trajectory. Further, a prediction task can be performed open-loop on real world data without the need for a simulation environment.

[202] Philion, J. et al.: Learning to Evaluate Perception Models Using Planner-Centric Metrics (2020).

The following validation method is proposed. Using an evaluation dataset, the prediction network is run on different inputs. One input is the original data without any object manipulation. The other input is the data that has only those objects present that were previously judged as relevant. Both predictions are then compared to the actual ground truth and the discrepancy between prediction and ground truth is quantified. This resultant discrepancy is displayed in figure 5-5 for a relevant OOI. In order to validate a whole relevance model instead of individual relevance judgments, a further step is necessary. For the whole relevance model to be considered valid, the overall introduced change in the discrepancy between prediction and ground truth needs to be summarized and not exceed a predetermined acceptance threshold.

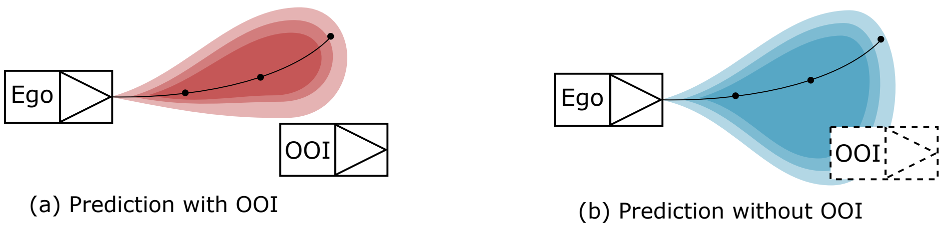

Figure 5-5: Relevance confirmation for ideal prediction

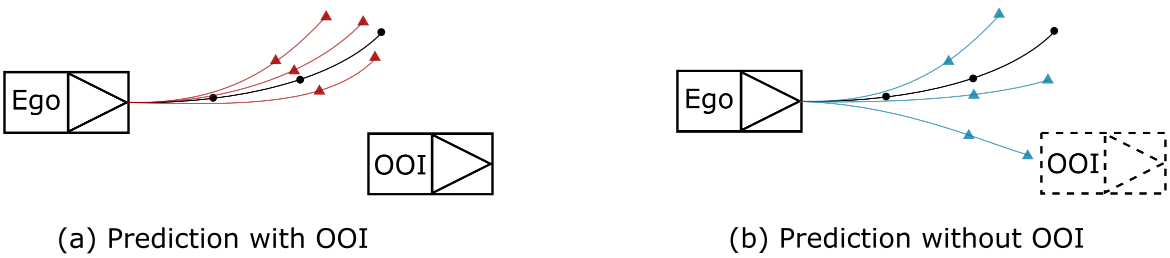

Figure 5-6: Relevance confirmation for discrete prediction

To account for the circumstance of not having an ideal continuous probability distribution predicted, the prediction algorithm is run multiple times with a different random seed to approximate the ideal prediction algorithm. This approximation is shown in figure 5-6 and stands in contrast to figure 5-5. For the discrete prediction used, the individual discrepancy is described as average displacement error (ADE). To describe the discrepancy for the whole model, the empirical cumulative distribution function (eCDF) of the ADE of both the prediction with the OOI present and the prediction with the OOI omitted is considered. This is to avoid local performance issues and the effects of non-deterministic predictions. As an acceptance threshold for a given relevance criterion, the hypothesis that the sampled error distributions remain unchanged between both inputs is used. Whether the error distribution has remained unchanged is evaluated using a statistical test. For this purpose, the two sample Cramer-von Mises test is used to test if both sampled error distributions are from the same underlying continuous distribution.[203] The test yields, based on distribution similarity and sample size, a confidence value to which a confidence threshold as an acceptance criterion can be applied.

[203] Anderson, T. W.: On the Distribution of the Two-Sample Cramer-von Mises Criterion (1962).

6 Evaluation of the Relevance Concept

The contents of this chapter have previously been published in separate papers[204,205] co-authored by the author of this work together with Ken Mori, and are presented here with modification towards the application within this dissertation. An analogous approach can be found in the dissertation of Ken Mori[206].

A proof of concept for the presented abstract methodology is applied to a chosen use case. Before the methodology is applied, the use case and the system need to be specified.

6.1 Use Case Specification

For evaluation purposes, of the proposed methodology, the collision free traversal of an urban environment is considered as a use case. As part of the DDT, the task of collision avoidance is considered as the subject task. To showcase the end to end applicability of the methodology, the perception component is chosen as a subject, since it is the initial step in any solution finding. The extent of the use case is limited to a minimal scope so as not to distract from the methodology being evaluated. Consequently, any fulfillment of the use case is considered a success without regard for any performance indicators, such as speed, comfort, or similar quantities.

Since the overarching aspect of this work is data reduction, special considerations apply to the error types in relevance judgment. The impact of a FP, that is falsely judging an object as relevant when it is irrelevant, is only a worsened data reduction factor. Conversely, a FN that judges an object as irrelevant when it is relevant, can have a significant impact on the downstream task. To meet these considerations, the relevance of each object must be judged conservatively, where a FP is acceptable, while a FN is not.

The question to be answered by the application of the abstract methodology in this experiment can be summarized as:

> *Which objects are guaranteed to be irrelevant to a perception module at a single point in time, in the meaning that they do not influence a collision-free performance of the DDT in an urban environment?*

[204] Mori, K. et al.: Conservative Estimation of Perception Relevance of Dynamic Objects (2023).

[205] Storms, K. et al.: SURE-Val: Safe Urban Relevance Extension and Validation (2023).

[206] Mori, K.: Defining Object Detection Requirements for Safe Automated Driving (2023).

6.2 Partial System Specification

As previously mentioned, a definition of relevance is inherently tied to a subject and a task that needs to be performed. Since the task is realized within a system, a partial system specification is required. To this end, a high-level architecture is defined. This architecture yields, in addition to the task requirements, requirements internal to the system, as well as capabilities that enable the system to fulfill the requirements.

In the following, first, the system architecture will be described to the extent that is sufficient for the experiment at hand. Second, a basic set of requirements is derived from the use case. For a further detailed relevance concept that yields criteria for a SRM, concrete requirements for the downstream task are necessary. In order to define more detailed requirements, a similarly more detailed description of the subject system has to be provided. Thus, within section 6.2.3, a set of minimal system capabilities for the use case is defined.

6.2.1 System Architecture

A modular Sense-Plan-Act architecture is assumed as minimal architecture, as it is a common example for architectures in AD and used as a basis for more complex architectures. The subject task of perception is included within the sense module.[207] Within this architecture, only the external interfaces for the perception module and its downstream tasks are specified. The specific internal mechanisms of each module are limited to a black box model.

As the interface between the sense and plan modules, a general object-centric representation is chosen. The choice for an object list was made since it is the common representation between perception and planning.[208]

The downstream task of the plan module is considered to yield a series of momentary actions describing a trajectory. The trajectories given by the plan module are then processed and executed downstream by the act module.

6.2.2 System Requirements

In order to specify the task component, system requirements (SYS-REQs) are needed. These system requirements determine constraints for the set of actions possible. As the first requirement, the downstream act module must be able to execute a planned trajectory to be valid:

SYS-REQ 1: *The actions must adhere to physical limitations.*

[207] Amersbach, C.: Functional Decomposition Approach (2019).

[208] Hoss, M. et al.: A Review of Testing Object-Based Environment Perception for Safe Automated Driving (2022).

As a second requirement, the actions needed to execute the planned trajectory must adhere to applicable traffic rules:

SYS-REQ 2: *The actions must adhere to applicable legal restrictions.*

6.2.3 System Capabilities

Complementary to the defined system requirements, the partial system specification needs to include a set of capabilities for the system, which enable it to fulfill the aforementioned requirements. The used system architecture is considered to be subject to latency, as depicted in figure 6-1. Assuming an event, like an action or initial perception, of another object, this point in time is defined as t_0. The system requires a specific time to sense the event, plan a corresponding trajectory, and utilize its actuators to realize the trajectory. The latency τ_r of the system is defined as the time elapsed from the first sensing of the event to the first externally measurable state change of the system, i.e. the first acceleration change in response to the event.

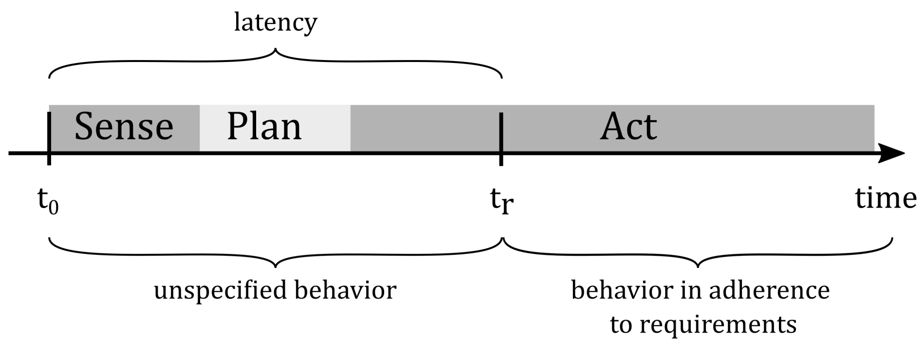

Figure 6-1: Following the occurrence of an event, an actor can only react after a certain amount of time t_r. During that time a behavior in response to that event cannot be assumed. After the latency has elapsed it is assumed that the actor has to behave in adherence to the behavioral requirements derived from the event.

The previously mentioned capabilities might be limited before their theoretical maximum. This may be due to external circumstances narrowing the physical limits, or system specifications restricting the acceptable uncertainties or risks associated with the capability. Considering the limited informational scope of the relevance concept, the system is simplified to have only two key capabilities providing acceleration and braking deceleration. Acceleration can be provided in either a longitudinal or lateral direction. The latter is provided by steering. The action space of the subject system is therefore described by three capability parameters.

- minimum guaranteed braking deceleration a_b
- minimum guaranteed acceleration a_g
- maximum guaranteed latency τ_r [209]

[209] It should be noted, that while the system latency is denoted as the duration τ_r, the time when the system latency has elapsed is denoted as t_r

6.3 Application of Method

Using the previously established use case and partial system specification, this section concerns the application of the method to derive the concrete relevance criteria.

6.3.1 Use Case Decomposition

As stated in section 6.2.1, the environment is assumed to be represented by a generic object list.

An object is defined at a given point in time by a set of parameters defining the object's state and its relation to other objects. While certain parameters can be determined with high accuracy, other parameters are subject to uncertainty. In order to limit the influence of parameters with higher uncertainty on the relevance criteria, only a minimal set of parameters is used. The states and relation of an object pair in this experiment are completely described by their position vector $\vec{r}$ and velocity vector $\vec{v}$ on a two-dimensional plane.

Since prediction of future object behavior and interpretation of road geometry is a difficult task that is subject to uncertainty, it is excluded from further explicit consideration as described in section 5.2.2. Instead, future behavior is implicitly considered by choosing the worst-case OOI behavior for the best-case ego behavior. Meaning, that while the ego object is trying to avoid a collision, the assumption for the OOI's behavior is such, that it is trying to force a collision. In a head-on scenario, this assumption equates to the ego object braking and steering away from the OOI, while the OOI accelerates and steers to intercept the ego object. While uncertainty in the road geometry is considered implicitly as well, it is done by superposition and evaluation of multiple applicable hypothetical scenarios. Each hypothetical scenario can be understood, that a road geometry exists, which enables the assumed behaviors.

For a given point in time, any object list can contain many objects. As previously used by Topan et al.[210], this experiment only considers pairwise interactions between the ego object and an OOI. Including additional objects in the interaction only introduces restrictions on their behavior. For ego, these restrictions are already covered, since all behavioral requirements of the functional scenarios are considered superposed, where any positive relevance judgment leads to the OOI being considered as relevant. For the OOI, any behavioral restriction can lead to a worst-case being excluded. As such, not including any additional object beyond the pairwise interaction overestimates the behaviors and leads to a conservative estimate while also limiting the complexity of relevance criteria.

Scenarios in literature are commonly described by lane based coordinates.[211,212,213] Since the

[210] Topan, S. et al.: Interaction-Dynamics-Aware Perception Zones for Obstacle Detection Safety Evaluation (2022).

[211] Shalev-Shwartz, S. et al.: On a Formal Model of Safe and Scalable Self-driving Cars (2017).

[212] Philipp, R. et al.: Systematization of Relevant Road Users (2022).

[213] Schönemann, V. et al.: Maneuver-based Adaptive Safety Zone for Valet Parking (2019).

worst-case road geometry is only considered after disambiguation of the applicable scenario, road centric coordinates cannot be used to describe the considered functional scenarios. Thus, a different coordinate system is necessary for the representation of the functional scenarios.

Since cartesian vehicle centric coordinates inherently imply a straight road in traveling direction, they are considered unsuitable. In contrast, a polar coordinate system is considered more suitable, as the radial distance to decrease to zero for a collision to occur. The following will decompose the use case into functional scenarios based on the relative positioning of the objects and their radial and tangential velocities with respect to each other. The resulting decomposition is structured by the primary direction of the interaction in the coordinate system. As such, a distinction is made between radial and tangential scenarios.

As a shorthand, each of the functional scenarios is abbreviated by a three letter combination. The first letter represents the primary axis of interaction, either radial (R) or tangential (T). After a separating period (.), two letters denote the relative movement direction of the ego and the OOI. The first letter refers to the ego vehicle and the second to the OOI. Possible relative movement directions are towards (T), and away (A); if the relative movement direction is inconsequential to the considerations, (X) is used to signify that (T) and (A) are possible.

Radial Scenarios

This section considers all functional scenarios in which the primary interaction is in the radial direction. The basic constellation and transfer from the perceived situation (A) to the evaluated hypothetical scenario is depicted in figure 6-2. Here, the hypothetical road does not correspond with the direction of movement of the objects. (A) shows the extent of the observed world, consisting only of the position, bounding box, and velocity of the objects. Part (B) of the figure describes the simplification process of projecting velocities, the assumption of a circular shape for the objects as well as a 1-D simplification of the world. Finally, (C) yields the hypothetical world that is evaluated, depicting the simplifications as well as the assumed action spaces for the objects, as denoted by the acceleration vectors.

The following notation is used for the equations distinguishing the functional scenarios. Location and velocity vectors are denoted with the numerical index of 1 or 2, where 1 is used to refer to the ego object and 2 to refer to the OOI. For equations that are applicable to either the ego object or the OOI, the variable i is used instead. When multiple indices are used, the index in the first position denotes the object being referred to. The Euclidean distance vector between both objects is given by:

$$\vec{d} = \vec{r}_2 - \vec{r}_1 \tag{6-1}$$

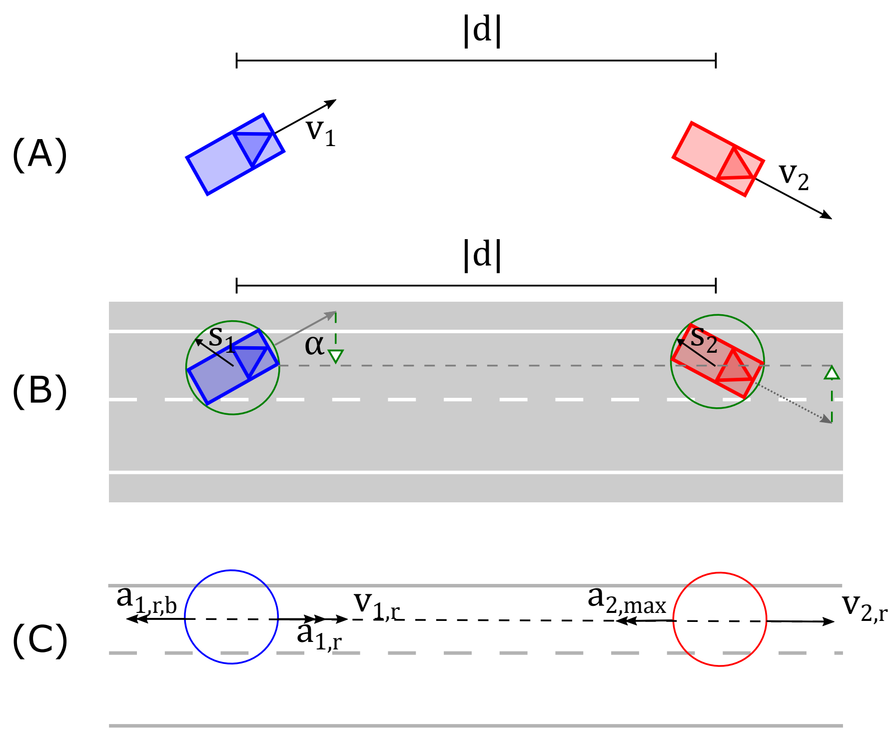

Figure 6-2: Development of a simplified environment model for a radial scenario. Shown for scenario R.TA. (A) Perceived real world (B) simplifications through hypothetical road model (C) simplified model

Overall, considering the full factorial combination of relative radial movement directions, four functional scenarios are possible. In the following equations, the second index of 0 indicates the initial state to distinguish whether a functional scenario is applicable.

- ego moving towards OOI, OOI moving away from ego (R.TA):

$$\vec{d}_0 \cdot \vec{v}_{1,0} > 0 \;\wedge\; \vec{d}_0 \cdot \vec{v}_{2,0} > 0 \tag{6-2}$$

- ego moving away from OOI, OOI moving towards ego (R.AT):

$$\vec{d}_0 \cdot \vec{v}_{1,0} < 0 \;\wedge\; \vec{d}_0 \cdot \vec{v}_{2,0} < 0 \tag{6-3}$$

- both vehicles moving away from each other (R.AA):

$$\vec{d}_0 \cdot \vec{v}_{1,0} < 0 \;\wedge\; \vec{d}_0 \cdot \vec{v}_{2,0} > 0 \tag{6-4}$$

- both vehicles moving towards each other (R.TT):

$$\vec{d}_0 \cdot \vec{v}_{1,0} > 0 \;\wedge\; \vec{d}_0 \cdot \vec{v}_{2,0} < 0 \tag{6-5}$$

To fully operate in an urban domain, one special functional scenario must be considered. During the performance of the DDT in an urban domain, it is possible to encounter a permanent bottleneck that can only be passed by traversing an area where the ego has lower priority than other OOIs. Since the bottleneck is permanent, a non-traversal of the lower priority would lead to a deadlock and non-fulfillment of the use case for the ego object. In many cases, this bottleneck is caused by a third stationary object, denoted by the index 3. The following functional scenario describes the applicability of an overtaking behavior. While the scenario is modeled with a static object in mind, it is also applicable to other bottlenecks, for which the traversal of the low priority area can be described by a virtual stationary object.

This functional scenario is considered as a subclass of R.TT and thus denoted as R.TT'. The conditions for the applicability of the scenario are given by:

- both vehicles moving towards each other, with a static object in between (R.TT'):
    - The static object is relevant for ego according to R.TA
    - Ego moving towards static object: $(\vec{r}_{3,0} - \vec{r}_{1,0}) \cdot \vec{v}_{1,0} > 0$
    - OOI moving towards static object: $(\vec{r}_{3,0} - \vec{r}_{2,0}) \cdot \vec{v}_{2,0} > 0$
    - OOI located behind static object: $(\vec{r}_{3,0} - \vec{r}_{1,0}) \cdot (\vec{r}_{3,0} - \vec{r}_{2,0}) < 0$

Additional conditions, like a known ego intention for passing, can be appended to the shown list of conditions. In this application, such conditions are omitted, since data on ego intention is generally not available in open datasets.

For the radial scenarios, a simplified one dimensional model is used. All quantities are projected on a new coordinate system defined by the connecting vector between ego object and OOI. For every projection, the conservativeness of the operation is considered.

The radial distance between ego object and OOI is given by:

$$d = |\vec{d}| \tag{6-6}$$

The Euclidean distance is the shortest path between two points, while the actual traversable path between both objects might be longer. This is considered a conservative simplification. Any tangential component of distance and velocity is ignored, resulting in:

$$v_{\mathrm{i,r}} = |\vec{v}_{\mathrm{i,r}}| = \left| \vec{v_{\mathrm{i}}} \cdot \frac{\vec{d}}{|\vec{d}|} \right| \tag{6-7}$$

Here, the index r in the second position is introduced to denote the radial axis. The reduction of the scenario to a one dimensional coordinate system is considered conservative, since any tangential motion may remove the potential for a collision by evasion. The available acceleration is conservatively assumed to be equal to the maximum possible acceleration when considering worst-case behavior. The available braking acceleration is close to or equal to zero if considering the worst-case, like a low friction surface or cornering at high speeds. As such, it is left as a variable to be filled by later argumentation with respect to a concrete system implementation. Further, the guaranteed braking acceleration is reduced with respect to the direction of ego movement, since part of the available acceleration potential may be used by the friction of tangential motion, for example by following a winding road. Therefore, the available potential for braking acceleration in radial direction $a_{\mathrm{i,r,b}}$ is limited to:

$$a_{\mathrm{i,r,b}} = cos(\alpha) \cdot a_{\mathrm{i,b}} = \frac{v_{\mathrm{i,r}}}{v_{\mathrm{i}}} \cdot a_{\mathrm{i,b}} \tag{6-8}$$

Here, the index in the third position b refers to the braking procedure.

While acceleration and steering are limited due to the friction coefficient of the road and the vehicle contact force on the road surface, additional limits for each are given separately by the vehicle[214] and preference of human passengers[215]. Additionally, the worst-case for acceleration and steering is that no residual potential is available. Thus, the available radial acceleration is treated as a quantity that needs to be guaranteed. As a consequence, it is not subject to an additional reduction like the braking acceleration, and $a_{\mathrm{i,r,g}} = a_{\mathrm{i,g}}$ applies.

Tangential Scenarios

For interactions along the tangential axis, two functional scenarios are applicable:

- OOI moving towards ego (T.XT):

$$\vec{d}_0 \cdot \vec{v}_{2,0} < 0 \tag{6-9}$$

- OOI moving away from ego (T.XA):

$$\vec{d}_0 \cdot \vec{v}_{2,0} \geq 0 \tag{6-10}$$

For both functional scenarios, the ego motion is irrelevant to the applicability. If the motion of the OOI is towards the ego object, a potential merging scenario is assumed, where the ego moves in front of the OOI and continues to travel in the same direction as the OOI. This is depicted in figure 6-3, alongside with the initial states of the objects, as well as the state at the of merging onto the lane of the OOI. Similar to the R.TT' functional scenario, additional conditions like the ego intention to merge in front of another vehicle can be considered for a more precise applicability.

[214] Bokare, P. S.; Maurya, A. K.: Acceleration-Deceleration Behaviour of Various Vehicle Types (2017).

[215] Bertolazzi, E. et al.: Supporting Drivers in Keeping Safe Speed and Safe Distance (2010).

In the case where the motion of OOI is away from the ego object, the tangential dynamics are irrelevant. The ego object can momentarily only merge behind the OOI, that is already considered in the radial R.TA scenario. For the ego object to merge in front of the OOI requires multiple intermediary actions, meaning the OOI will potentially become relevant once those actions have been performed.

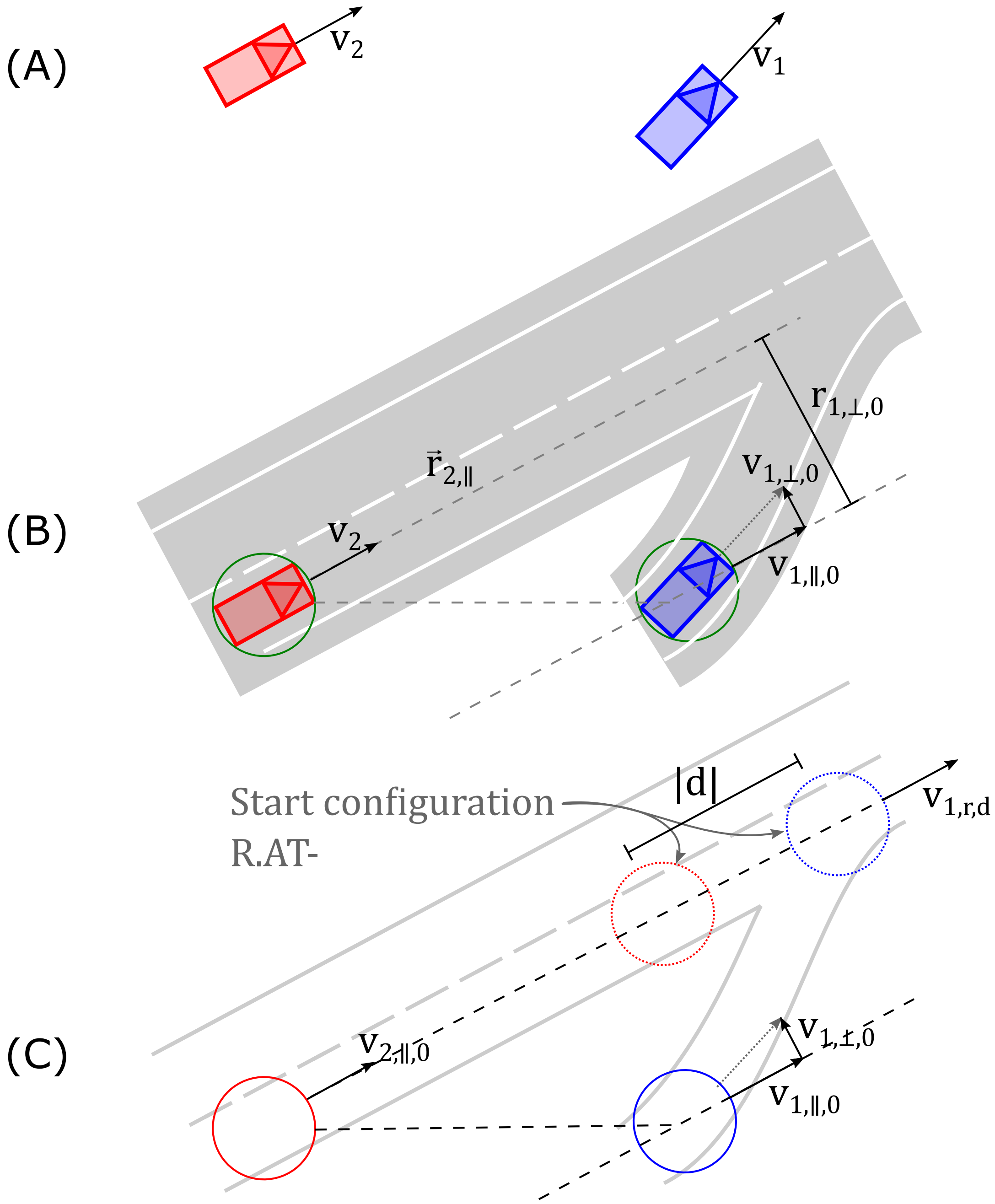

Figure 6-3: Development of a simplified environment model for a tangential scenario. Scenario T.XT is shown. (A) Perceived real world (B) simplifications through hypothetical road model (C) simplified model, with start configuration for R.AT- scenario once merging is complete.

6.3.2 Relevance Criteria for Functional Scenarios

In this subsection, the relevance criteria for the previously established functional scenarios are defined. Each criterion is expressed in the form of an analytic equation. As previously mentioned, each functional scenario is considered superposed to the others. As given by equation 6-11, an object is considered relevant if any criterion results in a relevance judgment. Conversely, a single criterion not yielding a relevant judgment, does not imply that the object is irrelevant. Irrelevance can only be established if all criteria yield an irrelevant judgment. For simplicity, irrelevance is simply noted as the negation of relevance as given by equation 6-12.

$$R.TA \vee R.AT \vee R.TT \vee R.AA \vee T.XT \Rightarrow relevant \tag{6-11}$$

$$\neg relevant \Rightarrow irrelevant \tag{6-12}$$

Initially, behavioral requirements are extracted from legal requirements defined by German traffic law[216]. Within the following relevance criteria, the ego object's adherence to the individual legal requirements is assumed after the system's latency. Before the system's latency has elapsed, the ego object's behavior is assumed to be according to the worst-case.

The trajectories representing the corresponding behaviors are presented in the form of parameterized equations. To define these equations, multiple simplifying assumptions are made, similarly to the worst-time-to-collision metric[217]. All objects are assumed to be spheres with the radius being equal to the half diagonal of the object. Further, Kamm's circle is utilized as an over approximation of the worst-case action space for both objects. Within Kamm's circle, the maximum acceleration $a_{\max}$ is assumed to be isotropic, meaning it is independent of the direction of acceleration[218]. While kinematic constraints can impose additional limitations on worst-case actions, the usage of the assumptions above is preestablished in literature[217,219] and is considered conservative.

These assumptions enable a specification of the behavior of the OOI and the action space for the ego object. Thus, the resulting relevance judgment can be formulated as:

> *All objects whose perception can change the set of viable momentary actions that allow for a successful compliance with the acceptable behavior.*

[216] Bundesministerium für Justiz und Verbraucherschutz: Straßenverkehrs-Ordnung (2013).

[217] Wachenfeld, W. H. K.; Winner, H.: The Release of Autonomous Vehicles (2016)

[218] Schmidt, C.: Unfallvermeidung im Straßenverkehr für Einzel- und Mehrobjektszenarien (2013).

[219] Althoff, M.; Magdici, S.: Set-Based Prediction of Traffic Participants on Arbitrary Road Networks (2016).

R.TA: Ego moving towards OOI, OOI moving away from ego

The R.TA functional scenario is considered to be equal to the ego object following the OOI. The German traffic regulation (StVO) applies the following behavioral requirement in such cases[220]:

> SYS-REQ2.1: *The ego vehicle shall be able to brake to halt behind a vehicle in front to avoid a collision in the event that the front car suddenly brakes.*

A similar requirement, as to not cause a rear end collision, has been leveraged by the RSS model[221]. This requirement for collision avoidance can be stated by use of the minimum distance d_{min} between ego object and OOI over the time of the scenario:

$$d_{\mathrm{min}} > 0 \tag{6-13}$$

Within the one dimensional environment model, the action space of the objects is limited to accelerating or braking in radial direction. During the system latency, the worst-case behavior for the ego object is considered accelerating towards OOI with maximum potential a_{max}. The worst-case behavior for the OOI does not change throughout the scenario and is constantly braking at maximum potential a_{max}. After the system latency has elapsed and the behavior of the OOI has been perceived by the ego, the ego is assumed to perform the best-case valid reaction, that is braking itself. This braking acceleration is restricted to the guaranteed potential $a_{1,\mathrm{b}}$. Given a constant acceleration the resulting object position r_{i} and velocity v_{i} are calculated by equation (6-14) and equation (6-15).

$$r_{\mathrm{i,r}} = r_{\mathrm{i,r,0}} + v_{\mathrm{i,r,0}}\ \tau + \frac{1}{2} a_{\mathrm{i,r,0}} \tau^2 \tag{6-14}$$

$$v_{\mathrm{i,r}} = v_{\mathrm{i,r,0}} + a_{\mathrm{i,r,0}}\ \tau \tag{6-15}$$

For both vehicles, the required braking distance to come to a standstill $r_{\mathrm{i,r,b}}$ is:

$$r_{\mathrm{i,r,b}} = \frac{v_{\mathrm{i,r}}^2}{2 a_{\mathrm{i,r,b}}} \tag{6-16}$$

The ego vehicle is assumed to initially accelerate during the system latency and then brake to a standstill. Thus, its position once coming to a full stop is calculated by adding the distances for both accelerating and braking, while also considering the gained speed:

[220] Bundesministerium für Justiz und Verbraucherschutz: Straßenverkehrs-Ordnung (2013), p.3.
[221] Shalev-Shwartz, S. et al.: On a Formal Model of Safe and Scalable Self-driving Cars (2017), p.6-7.

$$
\begin{aligned}
r_{1,\mathrm{r,s}} = r_{1,\mathrm{r},0} + v_{1,\mathrm{r},0}\ \tau_{1,\mathrm{r}} + \frac{1}{2} a_{1,\mathrm{r},0} \tau_{1,\mathrm{r}}^2 \\
+ \frac{(v_{1,\mathrm{r},0} + \tau_{1,\mathrm{r}} a_{1,\mathrm{r},0})^2}{2 a_{1,\mathrm{r,b}}}
\end{aligned}
\tag{6-17}
$$

Considering that a one dimensional coordinate system is used, the minimal distance $d_{\min}$ can be approximated by using the object positions once both have come to full stop. Additionally, the initial distance between both objects d_0 and their respective sizes s_i are considered. If a collision has occurred, the difference between equation (6-16) for the ego object and equation (6-14) for the OOI will be larger than the sum of initial distance and object sizes. Thus, $d_{\min}$ will be negative. Composing all requirements, considerations, and equations into a single form, the result is:

$$
\begin{aligned}
0 < d_{\min} = d_0 - s_1 - s_2 + \frac{v_{2,\mathrm{r},0}^2}{2 a_{\max}} - v_{1,\mathrm{r},0}\ \tau_{1,\mathrm{r}} \\
- \frac{1}{2} a_{\max} \tau_{1,\mathrm{r}}^2 - \frac{(v_{1,\mathrm{r},0} + \tau_{1,\mathrm{r}} a_{\max})^2}{2 a_{1,\mathrm{r,b}}}
\end{aligned}
\tag{6-18}
$$

If the condition is evaluated to be true, no collision occurs even given the worst-case momentary behavior available from the action space. If the condition is evaluated to be false, at least one possible behavior from the available action space can set the ego on a trajectory where a collision can be unavoidable. Adherence to SYS-REQ2.1 is then not guaranteed and the OOI is considered relevant to restricting the action space of the ego object.

R.AT: Ego moving away from OOI, OOI moving towards ego

In this scenario, the ego object is considered to be followed by another object. The behavioral requirement from German law[222] is, that

SYS-REQ2.2: *other vehicles should not be unnecessarily impeded.*

In this relevance model, the meaning of "impeded" is interpreted as not imposing restrictions to the action space regarding collision avoidance on the OOI. This functional scenario can be understood as the R.TA scenario with switched roles between ego and OOI. Hence, the requirement of not causing a rear end collision from R.TA applies to the OOI. This functional scenario can further be distinguished into two sub scenarios. In the first sub scenario, denoted as R.AT+, the ego vehicle is moving at a speed that is considered adequate. For the second sub scenario (R.AT-), the speed of the ego vehicle is considered below adequate. The meaning of adequate is here taken as a speed that is generally acceptable for constant traveling, both by

[222] Bundesministerium für Justiz und Verbraucherschutz: Straßenverkehrs-Ordnung (2013).

societal and legal standards. The exact value of an adequate speed is left to be determined within the individual sub scenarios.

R.AT+: Ego with desired speed moving away from OOI, OOI moving towards ego

Within the scope of the R.AT+ scenario, the initial speed driven by the OOI is considered adequate. In this context, SYS-REQ2.2 can be refined to the following requirement:

> SYS-REQ2.3: *The ego vehicle may not restrict the actions of the following vehicle by unnecessarily braking.*

By swapping the roles in the R.TA scenario, the results of the previous section can be applied again:

$$
\begin{aligned}
0 < d_{\min} = d_0 - s_1 - s_2 + \frac{v_{1,\mathrm{r},0}^2}{2a_{\max}} - v_{2,\mathrm{r},0}\,\tau_{2,\mathrm{r}} \\
- \frac{1}{2}a_{\max}\tau_{2,\mathrm{r}}^2 - \frac{(v_{2,\mathrm{r},0} + \tau_{2,\mathrm{r}}a_{\max})^2}{2a_{2,\mathrm{r},\mathrm{b}}}
\end{aligned}
\tag{6-19}
$$

It needs to be noted that $\tau_{2,\mathrm{r}}$ refers to the reaction of the OOI, which is conceptional equal to the ego vehicle's system latency. Since the worst-case for reaction time is infinite, it is again necessary to argue for a guaranteed value. Similarly, the guaranteed braking acceleration $a_{2,\mathrm{r},\mathrm{b}}$ refers to the OOI and must be conservatively overestimated for a guaranteed value. Equation 6-19 describes all situations where the ego vehicle braking at full potential does not cause a restriction of the OOIs action space. If the equation yields a false statement, the ego object restricts the action space and thus impedes the OOI. Violation of the equation thus leads to the OOI being relevant.

R.AT-: Ego with less than desired speed moving away from OOI, OOI moving towards ego

Contrary to the R.AT+ scenario, within the R.AT- scenario, the ego vehicle's speed is initially below an adequate speed. Here, adequate speed is considered to be tied to the ego vehicle's desired target speed $v_{1,\mathrm{r},\mathrm{d}}$. This value can change depending on location and appropriate subdomain. To have a conservative overestimation, the upper legal limit can be assumed. This functional scenario can be thought of as the ego vehicle changing lanes in front of a faster vehicle.

In this case, the ego vehicle has to initially accelerate to the desired target speed. Once the target speed has been reached, further considerations and equations of the R.AT+ scenario apply. While the ego vehicle's speed is not adequate, SYS-REQ2.2 can be refined as:

SYS-REQ2.4: *The ego vehicle may not restrict the actions of the following vehicle by having insufficient speed.*

Here, the system latency is initially not considered, since acceleration to the desired target speed is not in reaction to the OOI. The acceleration towards the adequate speed is considered to be constant throughout the scenario. Equally, the OOI constantly accelerates towards the ego vehicle.

Applying these considerations, the distance between both objects can be calculated by using equation (6-14) as:

$$d = d_0 - s_1 - s_2 + (v_{1,\mathrm{r},0} - v_{2,\mathrm{r},0})\tau + \frac{1}{2}(a_{1,\mathrm{r,g}} - a_{\max})\tau^2 \tag{6-20}$$

Given the assumption that the acceleration potential for OOI $a_{\max}$ is greater than the guaranteed acceleration potential for the ego object $a_{1,\mathrm{g}}$, the distance is minimal once the ego object has accelerated to the adequate speed. The state once the adequate speed has been reached is denoted by the index d.

$$\tau_\mathrm{d} = \frac{v_{1,\mathrm{r,d}} - v_{1,\mathrm{r},0}}{a_{1,\mathrm{r,g}}} \tag{6-21}$$

Substituting the time needed to reach an adequate speed, as given by equation (6-21), into equation (6-20) and (6-15) results in equation (6-22) defining the distance between ego object and OOI at the time of reaching adequate speed. At this point in time, the speed of the OOI is:

$$v_{2,\mathrm{r,d}} = v_{2,\mathrm{r},0} + a_{\max}\tau_\mathrm{d} \tag{6-22}$$

For further considerations, the equation (6-19) from the R.AT+ scenario is reused with additional modifications. First, the initial distance in R.AT+ given as $d_0 - s_1 - s_2$ is replaced with the residual distance given by substituting equation (6-21) into equation (6-20). Similarly, $v_{1,\mathrm{r},0}$ and $v_{2,\mathrm{r},0}$ are replaced with $v_{1,\mathrm{r,d}}$ and $v_{2,\mathrm{r,d}}$ respectively. This yields the final criterion for relevance in the R.AT- scenario:

$$
\begin{aligned}
0 < d_{\min} = d(t = t_{\mathrm{d}}) &+ \frac{v_{1,\mathrm{r,d}}^2}{2a_{\max}} - v_{2,\mathrm{r,d}}\,\tau_{2,\mathrm{r}} \\
&- \frac{1}{2}a_{\max}\tau_{2,\mathrm{r}}^2 - \frac{(v_{2,\mathrm{r,d}} + \tau_{2,\mathrm{r}}a_{\max})^2}{2a_{2,\mathrm{r,b}}}
\end{aligned}
\tag{6-23}
$$

R.TT: Both vehicles moving towards each other

This functional scenario corresponds to two objects either driving on opposite lanes towards each other or having a lateral movement towards each other on neighboring lanes. In the case of two vehicles moving towards each other, the RSS model makes the assumption of correct response for both vehicles.[223] This assumption is also found in the StVo. Here, the legal requirement to be able to stop within half of the visible distance is made.[224] This legal requirement does not derive from worst-case assumptions, which are only constrained by the dynamic limits of the objects.[225]

The worst-case assumptions in this scenario are as follows. During the system's latency, both the ego vehicle and the OOI accelerate toward each other. After the latency has elapsed, the ego vehicle starts to brake, while the OOI continues to accelerate.

In this scenario, only the general legal requirement, that "*A person using the road shall act in such a way as not to harm or endanger or, more than is unavoidable in the circumstances, to hinder or harass any other person*"[226] can be applied. As this is a very broad requirement, it needs further interpretation to yield concrete behavioral requirements.

If an accident cannot be avoided, mitigating the consequences of the accident is reasonable. Since the ego object's action space is limited to the ego object itself, the only mitigation option considered in this scenario is the minimization of the ego object's speed. Thus, the legal requirement above is refined to the following behavioral requirement:

> SYS-REQ2.5: *the ego vehicle shall brake to a standstill before the other vehicle collides with it.*

This results in the requirement that $d_{\min}$ has to be larger than zero at the time of the ego vehicle reaching a speed of zero, which is equal to the previously used condition in equation (6-13). This requirement needs to be fulfilled given a worst-case braking reaction by the ego vehicle as described in equation (6-16). The speed during the system latency and after are defined by equations (6-24) and (6-25).

[223] Shalev-Shwartz, S. et al.: On a Formal Model of Safe and Scalable Self-driving Cars (2017).

[224] Bundesministerium für Justiz und Verbraucherschutz: Straßenverkehrs-Ordnung (2013).

[225] Wachenfeld, W. H. K.; Winner, H.: The Release of Autonomous Vehicles (2016).

[226] Bundesministerium für Justiz und Verbraucherschutz: Straßenverkehrs-Ordnung (2013), I. Section 1 (2).

$$v_{1,\mathrm{r}} = v_{1,0} + a_{\max}\tau \quad \text{for} \quad \tau \leq \tau_\mathrm{r} \tag{6-24}$$

$$v_{1,\mathrm{r}} = v_{1,\mathrm{r,b}} - a_{1,\mathrm{r,b}}(\tau - \tau_\mathrm{r}) \quad \text{for} \quad \tau \geq \tau_\mathrm{r} \tag{6-25}$$

$v_{1,\mathrm{r,b}}$ is the ego object's speed after the system latency has elapsed and given by inserting τ_r into equations (6-24). The time the ego vehicle needs to come to a standstill $\tau_{1,\mathrm{b}}$ is determined by substituting $v_{1,\mathrm{r,b}}$ with equation (6-24) and $v_{1,\mathrm{r}}$ with 0. Solving for $\tau_{1,\mathrm{b}}$ yields:

$$\tau_{1,\mathrm{b}} = \tau_{1,\mathrm{r}} + \frac{(v_{1,\mathrm{r},0} + \tau_{1,\mathrm{r}}\ a_{\max})}{a_{1,\mathrm{r,b}}} \tag{6-26}$$

The position of the OOI can be described by equation (6-14), where the speed and acceleration of the OOI are negative. For the final distance measure, the position of the ego object after stopping, described in equation (6-17) and the sizes of the objects are subtracted. With the worst-case assumption and the condition (6-13) the final criterion is:

$$\begin{aligned} 0 < d_{\min} = d_0 &- s_1 - s_2 - v_{1,\mathrm{r},0}\ \tau_{1,\mathrm{r}} \\ &- \frac{1}{2} a_{\max}\tau_{1,\mathrm{r}}^2 - \frac{(v_{1,\mathrm{r},0} + \tau_{1,\mathrm{r}} a_{\max})^2}{2a_{1,\mathrm{r,b}}} \\ &- v_{2,\mathrm{r},0}\tau_{1,\mathrm{b}} - \frac{1}{2} a_{\max}\tau_{1,\mathrm{b}} \end{aligned} \tag{6-27}$$

As before, any OOI violating this criterion is considered relevant, since a head-on collision cannot be excluded from worst-case considerations.

R.TT': Both vehicles moving towards each other, with a static object in between

The R.TT' functional scenario is considered a special case of the previous R.TT functional scenario. As such, the behavioral requirement SYS-REQ2.5 from the R.TT scenario equally applies. A prerequisite for the R.TT' scenario is, that the simpler R.TT scenario has not led to any relevant judgment for the OOI. Given this prerequisite and the preconditions in section 6.3.1, that a static object in between the ego object and OOI exists, the R.TT' scenario is evaluated. The R.TT' scenario depicted in figure 6-4 considers that an assumed permanently static object blocks further fulfillment of the use case if the static object is not passed. Since this scenario introduces a third object, it is accordingly denoted by the usage of index 3 for object identification.

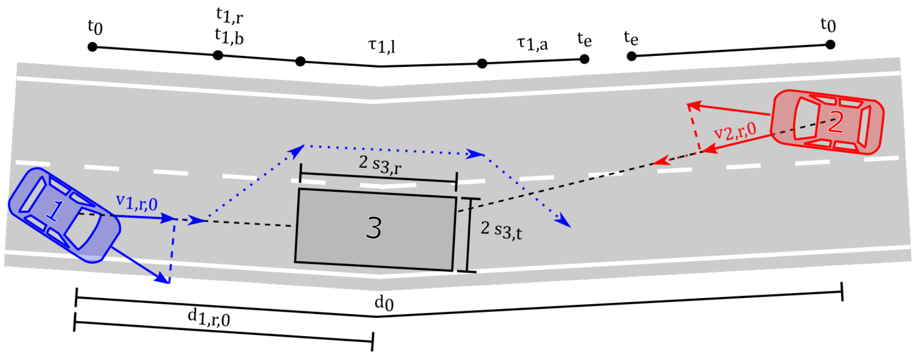

Figure 6-4: Sequence and variables for the R.TT' scenario.

The initial step in evaluating the R.TT' scenario only considers the static object as OOI. Due to the static object not having any speed, the R.TA and R.TT are equivalent. Both functional scenarios evaluate whether the static object is yet relevant for a future passing maneuver. In terms of the R.TA and R.TT scenarios, it is required that the initial distance between the ego and static object is less than the combined reaction and braking distance:

$$\begin{aligned} 0 < d_{1,\mathrm{r,min}} = d_{1,\mathrm{r},0} - s_1 - s_{3,\mathrm{r}} \\ - v_{1,\mathrm{r},0}\ \tau_{1,\mathrm{r}} - \frac{1}{2} a_{\max} \tau_{1,\mathrm{r}}^2 \\ - \frac{(v_{1,\mathrm{r},0} + \tau_{1,\mathrm{r}} a_{\max})^2}{2 a_{1,\mathrm{r,b}}} \end{aligned} \tag{6-28}$$

For further consideration, the following worst-case assumptions apply. During the system's latency, a braking action with maximum acceleration of the ego vehicle is assumed. Equation (6-29) describes the distance covered by the ego during the braking action, while equation (6-30) describes the speed of the ego object at the end of the system's latency. The duration of the braking maneuver $\tau_{1,\mathrm{b}}$ is defined as the minimal value of either the system latency or the time to standstill as defined in equation (6-31).

$$d_{1,\mathrm{b}} = v_{1,\mathrm{r},0} \tau_{1,\mathrm{b}} - \frac{1}{2} a_{\max} \tau_{1,\mathrm{b}}^2 \tag{6-29}$$

$$v_{1,\mathrm{r,b}} = v_{1,\mathrm{r},0} - a_{\max} \tau_{1,\mathrm{b}} \tag{6-30}$$

$$\tau_{1,\mathrm{b}} = min \left\{ \tau_{1,\mathrm{r}}\ ,\ \frac{v_{1,\mathrm{r},0}}{a_{\max}} \right\} \tag{6-31}$$

Once the system latency has elapsed, the ego object is assumed to react to the situation, requiring a passing maneuver. Thus, a lateral movement towards the opposite lane is initiated to start the passing procedure. The duration of the lateral movement $\tau_{1,\mathrm{l}}$ is defined in equation (6-32). When the opposite lane is reached, the ego object performs an acceleration within the limits of the guaranteed acceleration. For the acceleration phase, the duration $\tau_{1,\mathrm{a}}$ is given in equation (6-34) and the final velocity $v_{1,\mathrm{r,a}}$ is defined by equation (6-33). After traversing the passing distance of $2(s_{3,\mathrm{r}} + s_1)$, the ego object is assumed to return to its original position with a lateral movement. After the ego object has concluded the lateral movement and has returned to its original lane, the passing maneuver is considered to be completed.

$$\tau_{1,\mathrm{l}} = 2\sqrt{\frac{s_{3,\mathrm{t}} + s_1}{a_{1,\mathrm{g}}}} \tag{6-32}$$

$$v_{1,\mathrm{r,a}} = v_{1,\mathrm{r,b}} + a_{1,\mathrm{g}}\tau_{1,\mathrm{a}} \tag{6-33}$$

$$\tau_{1,\mathrm{a}} = \frac{-v_{1,\mathrm{r,b}} + \sqrt{2a_{1,\mathrm{g}}(2s_1 + 2s_{3,\mathrm{r}}) + v_{1,\mathrm{r,b}}^2}}{a_{1,\mathrm{g}}} \tag{6-34}$$

For the ego vehicle, the final distance, defined as the sum of distances covered by the maneuvers, is given by equation (6-35). The duration of the whole scenario is calculated as described in equation (6-38). During this time, the OOI is assumed to perform a worst-case acceleration at maximum potential towards the ego object. The distance covered by the OOI is defined by equation (6-36), resulting in an OOI speed given in equation (6-37).

$$d_{1,\mathrm{e}} = d_{1,\mathrm{b}} + v_{1,\mathrm{r,b}}\tau_{1,\mathrm{l}} + d_{1,\mathrm{a}} + v_{1,\mathrm{r,a}}\tau_{1,\mathrm{l}} \tag{6-35}$$

$$d_{2,\mathrm{e}} = v_{2,\mathrm{r},0}t_\mathrm{e} + \frac{1}{2}a_{\max}t_\mathrm{e}^2 \tag{6-36}$$

$$v_{2,\mathrm{r,e}} = v_{2,\mathrm{r},0} + a_{\max}t_\mathrm{e} \tag{6-37}$$

$$t_\mathrm{e} = \tau_{1,\mathrm{r}} + 2\tau_{1,\mathrm{l}} + \tau_{1,\mathrm{a}} \tag{6-38}$$

Using these components, the final criterion can be assembled by using them as initial conditions for the R.TT criterion:

$$0 < d_{\min} = (d_0 - d_{1,\mathrm{e}} - d_{2,\mathrm{e}}) - s_1 - s_2 - v_{1,\mathrm{r,a}}\,\tau_{1,\mathrm{r}} \\ - \frac{1}{2}a_{\max}\tau_{1,\mathrm{r}}^2 - \frac{(v_{1,\mathrm{r,a}} + \tau_{1,\mathrm{r}}a_{\max})^2}{2a_{1,\mathrm{r,b}}} - v_{2,\mathrm{r,e}}\tau_{1,\mathrm{b}} - \frac{1}{2}a_{\max}\tau_{1,\mathrm{b}} \tag{6-39}$$

R.AA: Both vehicles moving away from each other

Similar to the R.TT functional scenario, the R.AA can be considered to correspond to two vehicles traveling longitudinal on opposing lanes or laterally on neighboring lanes. In the latter case, also depicted in figure 6-5, it is reasonable to assume that a reaction by the ego object is required if the OOI reverses the direction of motion. In this scenario, the same requirement from R.TT SYS-REQ2.5 applies. Similarly, the same fundamental equation (6-27) applies, while negative speeds are inserted in this functional scenario. If equation (6-27) is evaluated to be false, avoidance of a collision requires an appropriate reaction within the system's latency time, thus making the OOI considered relevant.

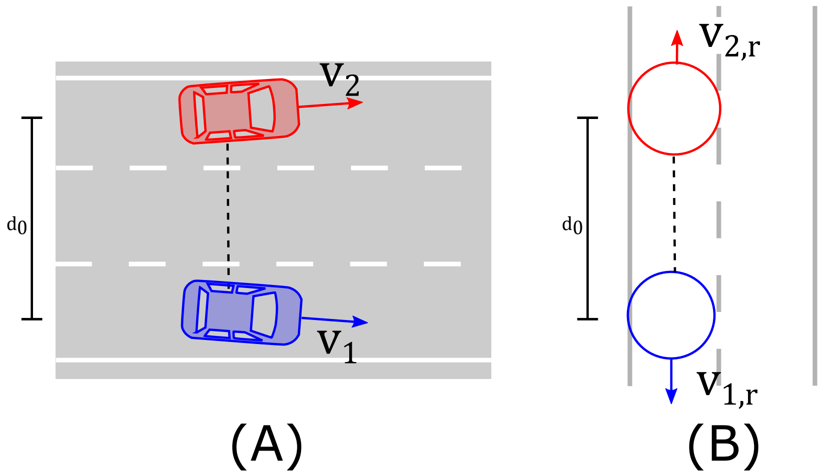

Figure 6-5: Example for the R.AA scenario (A) Possible applicable state for the R.AA scenario. (B) Simplified model with virtual road

T.XT: OOI moving towards ego

The T.XT functional scenario applies when the ego object is considered to be in front of the OOI due to its tangential motion. Based on this constellation, the relevance of the OOI is examined, given the assumption, that the ego object is merging in front of the OOI. In this context, merging is considered to mean changing one's position and direction so that the trajectories of both objects align. This is understood to correspond to either two vehicles on neighboring lanes or two vehicles on separate roads meeting at an intersection. In either case, the behavioral requirement to not impede others from SYS-REQ2.2 is considered, since not having the right of way or any other form of priority is a conservative assumption for this model.

In the context of an intersection, two basic interactions are possible as shown in figure 6-6, crossing trajectories (A) and aligning trajectories (B). Since the ego object needs to merely exit the intersection in interaction (A), while in interaction (B) it also needs to accelerate to an adequate speed so as to not impede the OOI, interaction is considered to be conservative. Objects identified as relevant in interaction (A) would also be identified as relevant when examining interaction (B). As such, only interaction (B) will be further used.

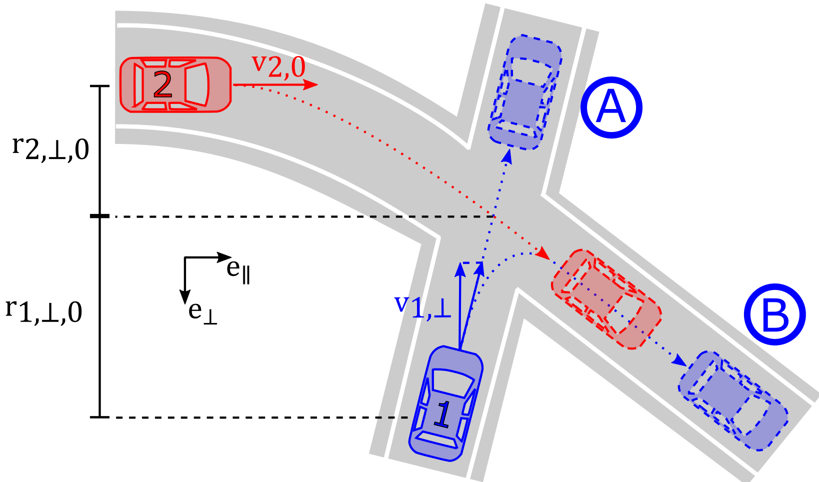

Figure 6-6: Simplified depiction of the worst-case intersection and variables for consideration if the intersection in the T.XT scenario is considered.[227]

Having excluded interaction (A) from the intersection scenario, assuming neighboring lanes or intersecting roads will lead to the same equations. For simplicity, the functional scenario is described in terms of an intersection.

Since the road layout is not known, a worst-case for the location of the merging point, i.e. the intersection, needs to be determined. The coordinate system for the assumed worst-case lane is defined by the unit vectors parallel $\vec{e}_{||}$ and orthogonal $\vec{e}_{\perp}$ to the OOIs direction of movement. The velocity components of the ego object within this coordinate system are given as:

$$v_{1,\perp} = \vec{v}_1 \cdot \vec{e}_{\perp} \quad \text{and} \quad v_{1,||} = \vec{v}_1 \cdot \vec{e}_{||} \tag{6-40}$$

Given sufficient distance, the ego object has the option not to enter the intersection by braking and thus to avoid merging in front of the OOI. The critical distance to standstill is given by equation (6-41).

[227] This image was created together with Ken Mori and also appears in the dissertation[228] created concurrently.

[228] Mori, K.: Defining Object Detection Requirements for Safe Automated Driving (2023)

$$d_{1,\perp,\mathrm{b}} = v_{1,\perp,0}\tau_{1,\mathrm{r}} + \frac{v_{1,\perp,0}^2}{2a_{1,\perp,\mathrm{b}}} \tag{6-41}$$

Since the intersection might not be located in the initial direction of travel for the OOI, it is possible for the intersection to be co-linear with the OOIs heading and closer to the ego object. The OOI is assumed to accelerate towards the intersection at maximum potential. The lateral distance traveled by the OOI to reach the intersection is defined by equation (6-42), using the ego vehicle's system latency and the ego vehilces's time to standstill from equation (6-43).

$$d_{2,\perp} = \frac{1}{2}a_{\max}(\tau_{1,\mathrm{r}} + \tau_{1,\mathrm{b}})^2 \tag{6-42}$$

$$\tau_{1,\mathrm{b}} = \frac{v_{1,\perp,0} + \tau_{1,\mathrm{r}}a_{\max}}{a_{1,\mathrm{b}}} \tag{6-43}$$

For the OOI to be considered relevant, a merging interaction has to be non-excludable for the ego object under worst-case consideration. This condition is true if the projected location of the intersection is reachable by the OOI and within the braking distance of the ego object :

$$r_{1,\perp,0} + r_{2,\perp,0} < v_{1,\perp,0}\tau_{1,\mathrm{r}} + \frac{v_{1,\perp,0}^2}{2a_{1,\perp,\mathrm{b}}} + \frac{1}{2}a_{\max}\left[\tau_{1,\mathrm{r}} + \frac{v_{1,\perp,0} + \tau_{1,\mathrm{r}}a_{\max}}{a_{1,\mathrm{b}}}\right]^2 \tag{6-44}$$

Given this precondition, the actual relevance of the OOI with regard to the merging interaction can be evaluated. While the previously assumed acceleration of the OOI towards the ego object constituted the worst-case to whether an intersection interaction has to be considered, for the actual interaction a different assumption applies. Here, a lane in the direction of movement of the OOI as depicted in figure 6-7 is assumed. Any curvature in the OOI trajectory would yield lower speed for the OOI at the time when the ego object enters the intersection. It is further assumed that an adequate speed for driving on this lane exists. After merging onto the lane, the ego object has the requirement to not impede the OOI, as previously specified in the R.AT- and R.AT+ scenario. Applicability of the R.AT- and R.AT+ scenario is dependent on whether the ego object initial velocity component parallel to the lane $v_{1,\perp,0}$ is above or below the adequate speed.

In this functional scenario, the following worst-case assumptions are used. First, the ego object accelerates towards the intersection during the system latency. Second, the OOI accelerates continuously throughout the scenario towards the intersection. Given the latter assumption, the ego velocity component orthogonal to the lane $v_{1,\perp}$ has to be zero, when the ego object is at

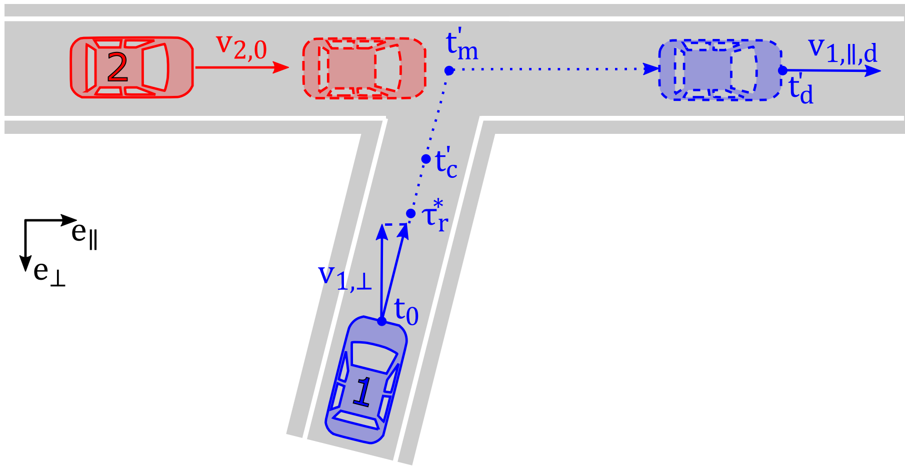

Figure 6-7: Simplified depiction of the worst-case intersection and temporal positions for relevance evaluation in the T.XT scenario.[229]

the intersection point. The ego object's speed and position towards the lane during the system latency is given by equation (6-45) and (6-46).

$$v_{1,\perp} = v_{1,\perp,0} + a_{\max} \cdot \tau_{1,\mathrm{r}^*} \tag{6-45}$$

$$r_{1,\perp} = r_{1,\perp,0} + v_{1,\perp,0} \cdot \tau_{1,\mathrm{r}^*} + \frac{1}{2} a_{\max} \tau^2 \tag{6-46}$$

Due to the orientation of the coordinate system, $v_{1,\perp}$ is negative before reaching the intersection. For the ego object not to reach the intersection point during the latency of the system, the following assumption has to be made, signifying the possibility that the ego object can come to a full stop before the reaction time has elapsed.:

$$\tau_{1,\mathrm{r}^*} = min\left\{ \tau_{1,\mathrm{r}} \ , \ -\frac{v_{1,\perp,0}}{a_{\max}} \right\} \tag{6-47}$$

After the system latency has elapsed, the ego object is assumed to perform deliberate actions in order to merge in front of the OOI. The start of these deliberate actions is denoted by the index s. The deliberate actions are constrained to the limits of guaranteed acceleration $a_{1,\mathrm{g}}$. Depending on residual distance to the intersection and ego speed, the actions can contain an initial orthogonal acceleration phase and a concluding orthogonal deceleration phase. The change from the acceleration phase and the beginning of the orthogonal deceleration phase is denoted by

[229] This image was created together with Ken Mori and also appears in the dissertation[237] created concurrently.

[230] Mori, K.: Defining Object Detection Requirements for Safe Automated Driving (2023)

the index c. While the acceleration might not be necessary, deceleration is always required to achieve an orthogonal speed of zero at the intersection point. The point in time when the ego object merges at the intersection point is denoted by the index m. For the following equations, the substitution $t' = t - \tau_{1,\mathrm{r}^*}$ will be used for readability. The ego object's orthogonal velocities up to the intersection point are given by:

$$v_{1,\perp} = v_{1,\perp,\mathrm{s}} - a_{1,\mathrm{g}} \cdot t' \quad \text{for} \quad t' \leq t'_{\mathrm{c}} \tag{6-48}$$

$$v_{1,\perp} = v_{1,\perp,\mathrm{c}} + a_{1,\mathrm{g}} \cdot (t' - t'_{\mathrm{c}}) \quad \text{for} \quad t' > t'_{\mathrm{c}} \tag{6-49}$$

Conversely, the ego vehicle's orthogonal position to the lane is:

$$r_{1,\perp} = r_{1,\perp,\mathrm{s}} + v_{1,\perp,\mathrm{s}} \cdot t' - \frac{1}{2} a_{1,\mathrm{g}} t'^2 \quad \text{for} \quad t' \leq t'_{\mathrm{c}} \tag{6-50}$$

$$r_{1,\perp} = r_{1,\perp,\mathrm{c}} + v_{1,\perp,\mathrm{c}} \cdot (t' - t'_{\mathrm{c}}) + \frac{1}{2} a_{1,\mathrm{g}} (t' - t'_{\mathrm{c}})^2 \quad \text{for} \quad t' > t'_{\mathrm{c}} \tag{6-51}$$

Once the ego vehicle reaches the target lane, no further orthogonal movement is assumed. These conditions at the merging time t'_{m} are:

$$v_{1,\perp,\mathrm{m}} = v_{1,\perp}(t = t'_{\mathrm{m}}) = 0 \tag{6-52}$$

$$r_{1,\perp,\mathrm{m}} = r_{1,\perp}(t = t'_{\mathrm{m}}) = 0 \tag{6-53}$$

For the time required to reach the target lane, the system of previous equations needs to be solved. This is done by initially substituting (6-48) and (6-49) into (6-52), yielding:

$$0 = v_{1,\perp,\mathrm{m}} = v_{1,\perp,\mathrm{s}} + a_{1,\mathrm{g}} \cdot (t'_{\mathrm{m}} - 2t'_{\mathrm{c}}) \tag{6-54}$$

Further, (6-45), (6-46) and (6-51) are inserted into (6-53), resulting in the following equation:

$$\begin{aligned} 0 = {} & r_{1,\perp,\mathrm{s}} + v_{1,\perp,\mathrm{s}} \cdot t'_{\mathrm{c}} - \frac{1}{2} a_{1,\mathrm{g}} {t'_{\mathrm{c}}}^2 \\ & + (v_{1,\perp,\mathrm{s}} - a_{1,\mathrm{g}} \cdot t_{\mathrm{c}}) \cdot (t'_{\mathrm{m}} - t'_{\mathrm{c}}) \\ & + \frac{1}{2} a_{1,\mathrm{g}} (t'_{\mathrm{m}} - t'_{\mathrm{c}})^2 \end{aligned} \tag{6-55}$$

Solving for t_{m} yields the temporal components t_{m} and t_{c} of the movement to the target lane:

$$t'_{\mathrm{c}} = \frac{v_{1,\perp,\mathrm{s}}}{a_{\max}} + \sqrt{\frac{r_{1,\perp,\mathrm{s}}}{a_{1,\perp,\mathrm{g}}}} \tag{6-56}$$

$$t'_{\mathrm{m}} = t'_{\mathrm{c}} - \frac{v_{1,\perp,\mathrm{s}}}{a_{1,\perp,\mathrm{g}}} \tag{6-57}$$

During the orthogonal movement to the target lane, it is assumed that the ego object maintains a constant velocity parallel to the target lane. Once the target lane has been reached, the resulting constellation between ego object and OOI is considered as the initial condition for the R.AT-functional scenario. As such, the ego object accelerates towards a desired speed that is considered adequate. The radial coordinate in the R.AT- scenario is equivalent to the coordinate parallel to the lane in this scenario. Reusing equation (6-21) from R.AT- for time to adequate speed, and considering the time for system latency and reaching the target lane, yields the time to adequate speed for the T.XT scenario:

$$t'_{\mathrm{d}} = \tau_{1,\mathrm{r}} + t'_{\mathrm{m}} + \frac{v_{1,\parallel,\mathrm{d}} - v_{1,\parallel,0}}{a_{1,\parallel,\mathrm{g}}} \tag{6-58}$$

Like in the R.AT- scenario, the threshold for an adequate speed is not known. As a conservative approximation, the speed of OOI can be used, defining $v_{1,\parallel,\mathrm{d}} = v_{2,0}$. Inserting this assumption into (6-22) and (6-20) results in:

$$v'_{2,\mathrm{d}} = v_{2,0} + a_{\max} t'_{\mathrm{d}} \tag{6-59}$$

$$
\begin{aligned}
d(t = t'_{\mathrm{d}}) = r_{1,\parallel,0} - s_1 - s_2 + (v_{1,\parallel,0} - v_{2,0})t'_{\mathrm{d}} \\
+ \frac{1}{2}(a_{1,\parallel,\mathrm{g}} - a_{\max}){t'_{\mathrm{d}}}^2
\end{aligned}
\tag{6-60}
$$

Finally, substituting in (6-23) for t_{d} and $v_{2,\mathrm{r,d}}$ yields the following criterion:

$$
\begin{aligned}
0 < d_{\min} = d(t = t'_{\mathrm{d}}) + \frac{v_{1,\parallel,\mathrm{d}}^2}{2a_{\max}} - v'_{2,\mathrm{d}}\,\tau_{\mathrm{r}} \\
- \frac{1}{2}a_{\max}\tau_{\mathrm{r}}^2 - \frac{(v'_{2,\mathrm{d}} + \tau_{\mathrm{r}} a_{\max})^2}{2a_{1,\parallel,\mathrm{g}}}
\end{aligned}
\tag{6-61}
$$

Any OOI violating the criterion is considered relevant to the ego object, as the actions during system latency have a direct influence on whether the OOI is being impeded.

6.4 Results

To study the usage of the proposed relevance concept, the equations above are implemented and applied to the NuScenes dataset[231]. For the application of the criteria, concrete parameter values are needed. Since no concrete ADS or later argumentation for safety is present in the context of this experiment, the values cannot be derived from them. Instead, plausible values are chosen that are plausible under typical conditions encountered in the urban domain. Thus, the transferability of the implementation for use in a concrete ADS and safety argumentation is not given. For such a transfer, the values must be redetermined.

Since the maximum acceleration is limited by friction, it is assumed that $a_{\max} = 10\ \frac{m}{s^2}$. While this ignores existing techniques to increase acceleration potential, like increasing the down force by using aerodynamic enhancements, the existence of these features is very limited in the usual urban domain. For braking accelerations when an emergency stop is required, the lower bound of viable values is chosen. Here, $a_{\mathrm{b}} = 7\ \frac{m}{s^2}$ is chosen, while also considering wet road surfaces.[233] The guaranteed acceleration potential is influenced by the current speed driven and which type of object, especially which type of vehicle is considered. While the possible lateral acceleration does not decrease with increasing speed, drivers are less likely to utilize the lateral acceleration potential with increasing speed[232]. Thus, the value of $a_{\mathrm{g}} = 0.5\ \frac{m}{s^2}$ is chosen based on literature[234,232]. Considering the reaction time $\tau_{1,\mathrm{r}}$ to an unforeseen event, human performance is

[231] Caesar, H. et al.: nuScenes (2020).

[232] Ali, G. et al.: Quantifying the effect on the frequency of longitudinal and lateral accelerations (2021)

[233] Poul Greibe: Braking distance, friction and behaviour: Findings, analyses and recommendations (2007).

[234] Bokare, P. S.; Maurya, A. K.: Acceleration-Deceleration Behaviour of Various Vehicle Types (2017).

indicated to be around 1.5 s.[235] In the implementation, these values are applied to both objects for simplicity.

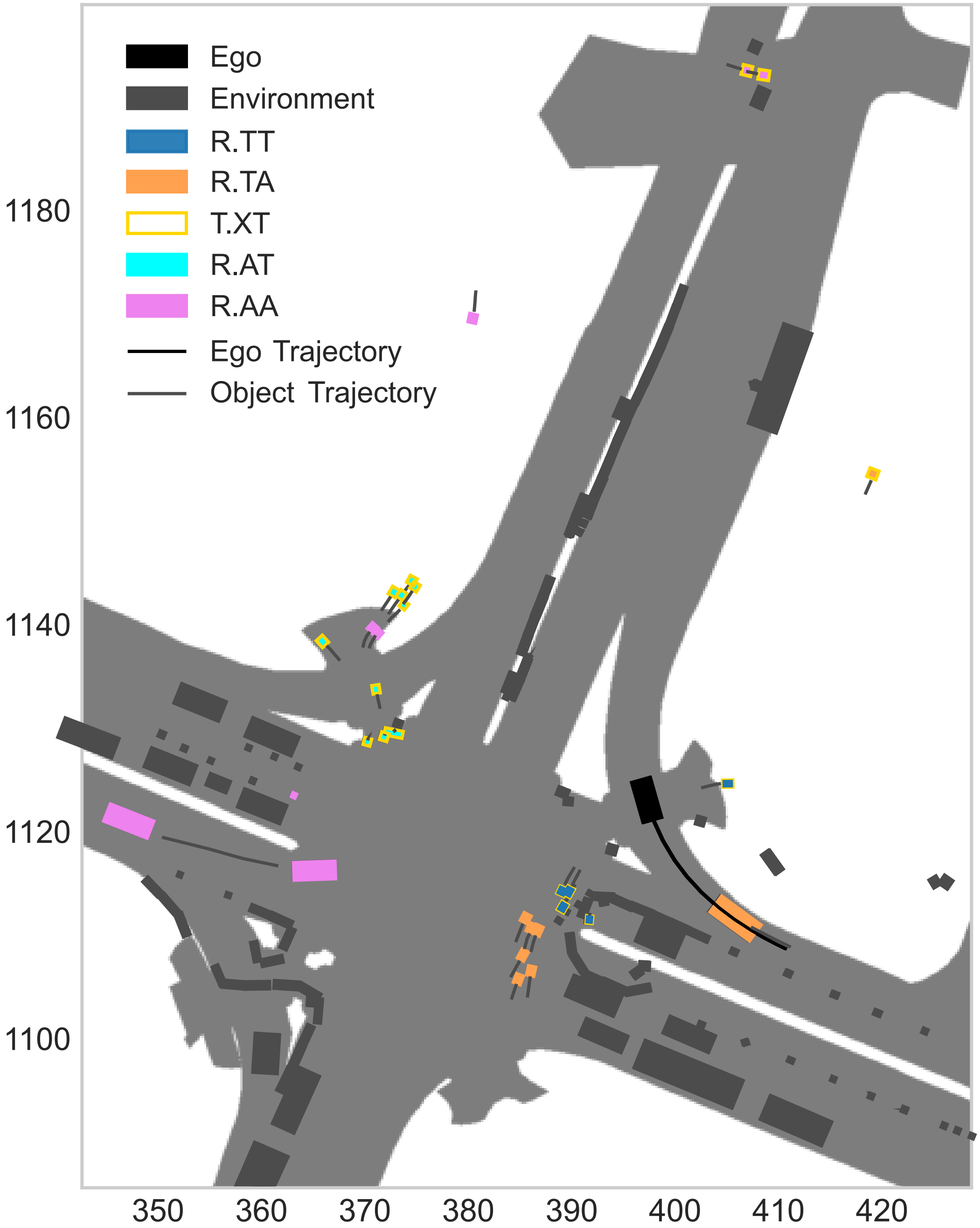

Figure 6-8: Applicable relevance criteria in an example scene of the NuScenes dataset. The displayed scene is located in Singapore with left hand traffic.[236]

[235] Green, M.: How Long Does It Take to Stop? Methodological Analysis of Driver Perception-Brake Times (2000).

[236] This image was created together with Ken Mori and also appears in the dissertation[237] created concurrently.

[237] Mori, K.: Defining Object Detection Requirements for Safe Automated Driving (2023)

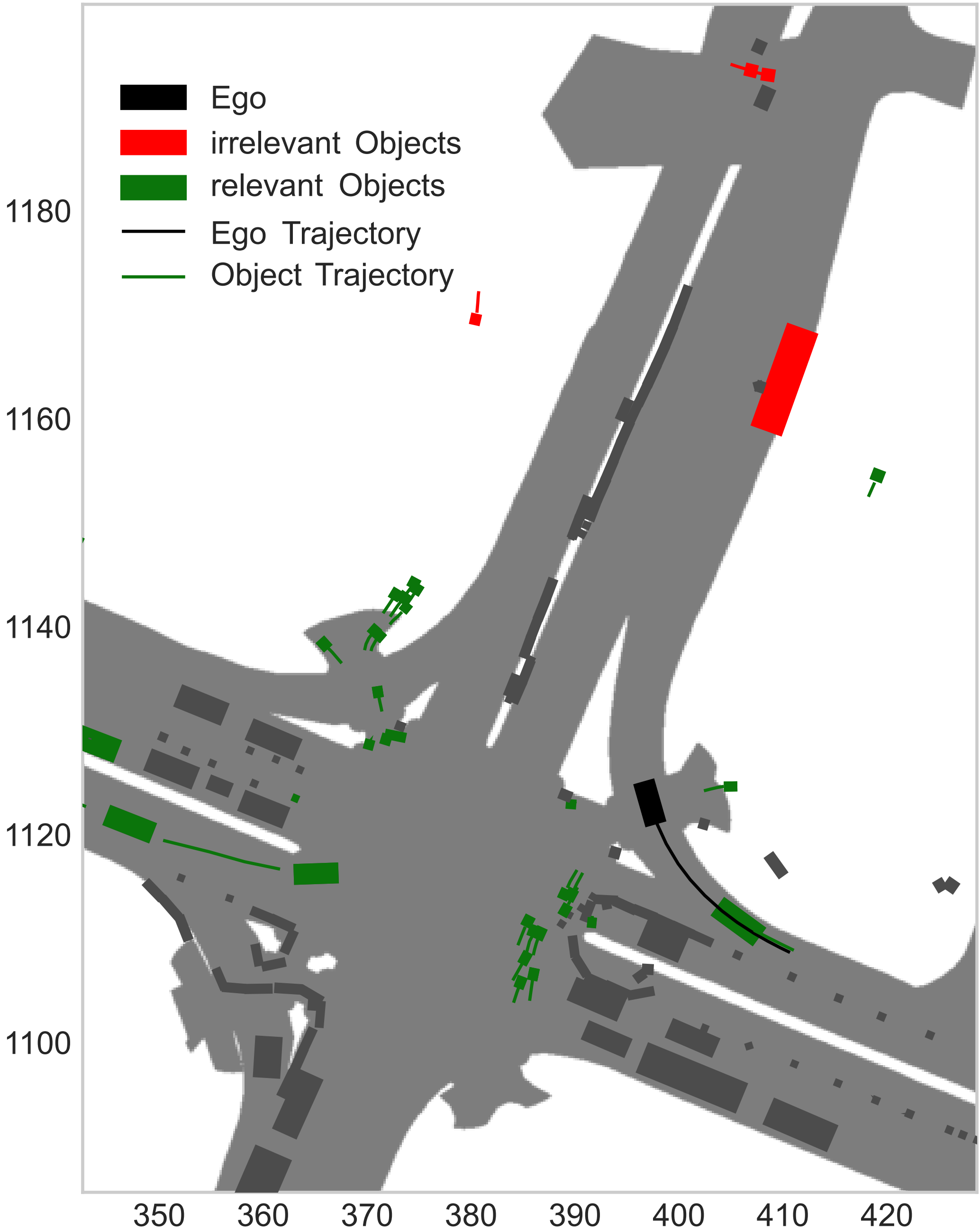

Figure 6-9: Resulting relevance judgments of the applied relevance criteria shown in figure 6-8. The displayed scene is located in Singapore with left hand traffic.[236]

Following, the results of the application are visualized. In figure 6-8 and figure 6-9 an intersection scene from the NuScenes dataset is shown. Figure 6-8 shows which criteria were applicable to which object. Figure 6-9 then displays which objects were judged relevant. Relative distance to the ego object can be observed as a distinguishing attribute for relevance, with the nearest irrelevant object being the stationary bus on the northern road. All relevant objects in this scene are in closer proximity to the ego object. Figure 6-10 visualizes the distribution of relative distances to the ego object at which objects were judged relevant or irrelevant as an eCDF.

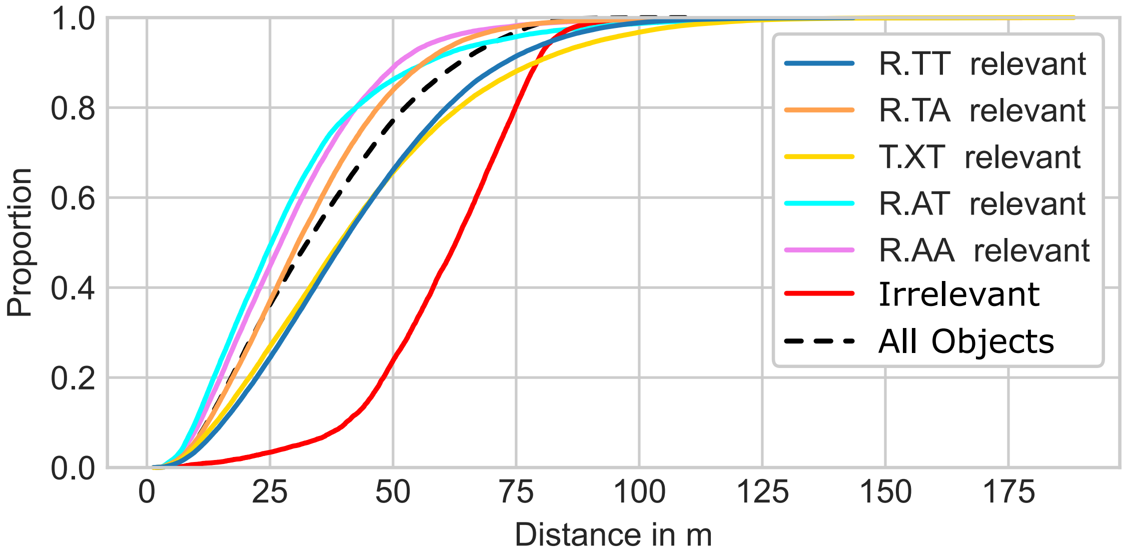

Figure 6-10: eCDF of distances for relevant judgments of individual relevance criteria on NuScenes.[236]

The observed irrelevant objects range from close to zero meters, to roughly 100 meters in distance. A rather sharp increase can be observed at around 40 meters, with less than 10 % of irrelevant objects being closer than that to the ego object. Considering the urban domain of the used dataset, the irrelevant objects are close to the edge of the observable area. Comparing the irrelevant objects to all the observed objects, it must be noted that 50 % of irrelevant objects are further away than more than 90 % of the overall observed objects.

A key property that can be noted is that objects are judged as relevant at distances of 175 meters and more. Considering that those distances correspond to the furthest annotation distances in the dataset, many objects were judged as relevant either in the first or last time step they were annotated. Thus, it cannot be ruled out that, given a hypothetical additional annotation, beyond the capabilities of the ego vehicle's sensors, the object would still have been judged as relevant. Therefore, it has to be concluded that the extent of the NuScenes dataset is insufficient to encompass all relevant objects.

Concluding, figure 6-11 represents the proposed SRM as a concretization of the GRM defined in the previous chapter. While this depiction does not visualize all components within the SRM, especially the information items concerning claim & assumptions, it serves as an example for the explicit statement of a relevance model.

6.5 Validation

Within this section, the validation procedure described in section 5.3 is applied to the criteria defined in section 6.3. In addition to the actual validation, the validation procedure is applied

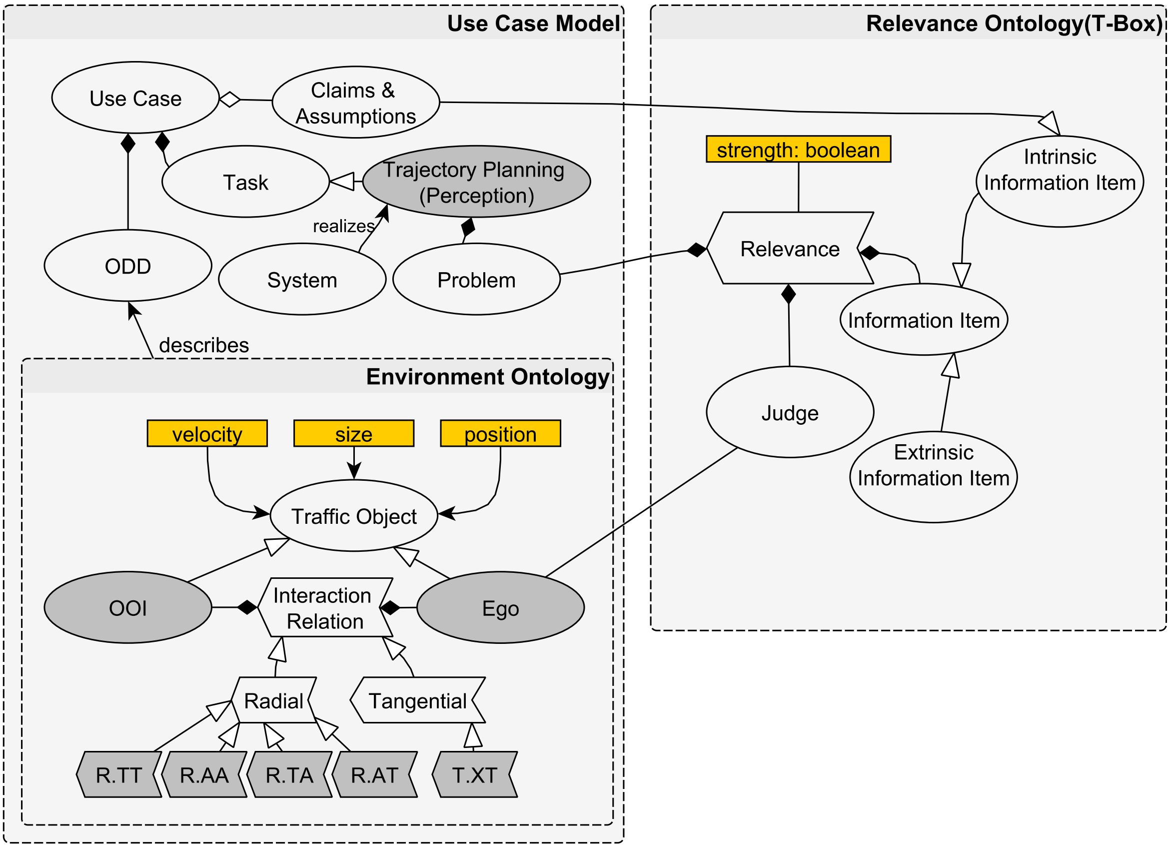

Figure 6-11: Ontological depiction of the proposed SRM

three more times. Twice with known invalid relevance criteria and once with no relevance criteria applied to test and verify the procedure itself. The following inputs to the prediction network in the validation method are used:

- **A**: All objects are used as input without omission. This is the original input and equal to no relevance criteria being applied
- **R**: Only relevant objects as identified by the criteria from section 6.3 used as input.
- **RV1**: No objects are used as input, as the known invalid relevance criterion ”All objects are irrelevant” is applied.
- **RV2**: Only objects with a lateral distance of greater than 2 m with respect to the ego vehicle used as input.

Since a probabilistic predictor is used, the noise introduced by stochastic effects must also be considered, as the prediction noise also influences the prediction error. While the prediction error is defined as the average Euclidean distance between the predicted trajectory and the ground truth trajectory, the prediction noise is established by comparing multiple runs for the original input. As such the prediction noise is defined as the average euclidean distance between two predicted trajectories, based on the same input. For more than two predicted trajectories, each

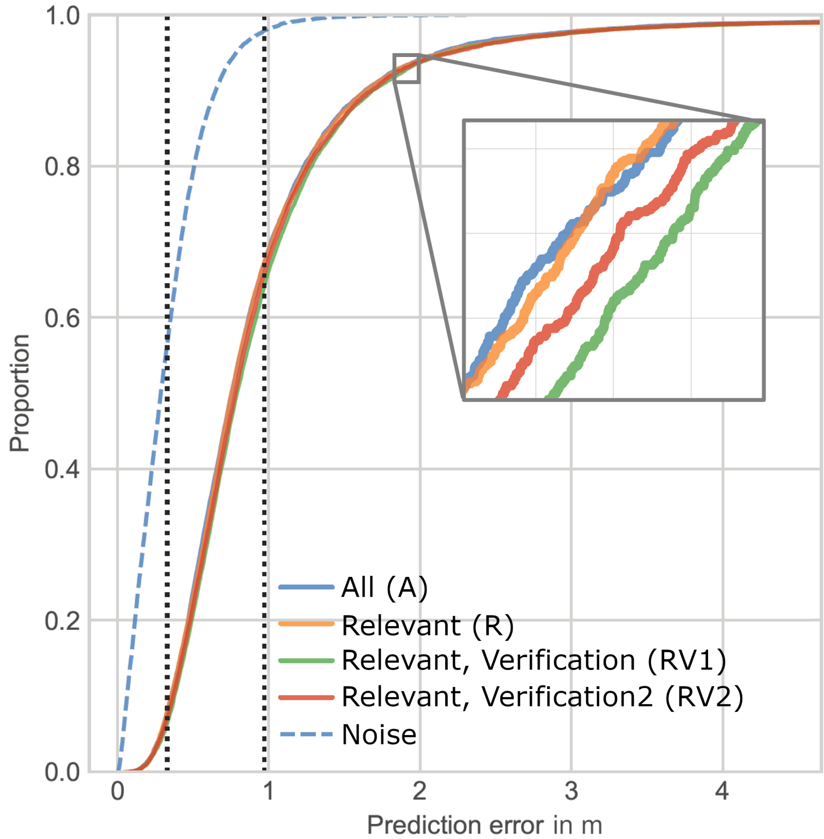

Figure 6-12: Error and noise eCDF distributions for the used relevance models

combination of pairs of predicted trajectories is considered. The prediction error for each input and the prediction noise are summarized by an eCDF in figure 6-12. While the error distribution averages at around 0.96 m, a significant average prediction noise at 0.33 m can be observed. The eCDFs of the error distributions are visually very similar, with the averages varying only in the range of single-digit centimeters. Observing the zoomed in section of the plot, it can be noted that the eCDF curve for A-A and A-R are close together and constantly crossing, while both errors for the verification inputs are continually larger. At the same time, all error distributions are significantly larger than the prediction noise eCDF which marks the theoretical optimum.

Since the procedure takes and compares two inputs, each validation run is identified by the notation of Input1-Input2. As such, A-R denotes the validation run for the proposed relevance concept. A-A is the validation run to verify the influence of no relevance concept applied. The application of no relevance concept is considered as an always valid baseline. The evaluation of A-A further quantifies the influence of prediction noise, due to the stochastic nature of the used prediction, on the validation method. A-RV1, A-RV2 denotes the two validation runs to verify the detectability of known invalid criteria. While in A-RV1 all objects are removed, in A-RV2 only 5% of the original objects are omitted. Thus, the rate of omission in A-RV2 is less than in A-R.

Performing the proposed Cramer-von Mises test on each combination of resulting predictions from both inputs results in several p-values for each validation run. These p-values are summarized in the box plot within figure 6-13. As an acceptance criterion, the generally used threshold of 5% or 0.05 is used and marked by a red horizontal line. It can be observed, that even with the influence of noise, the p-values of A-A are high, setting a plausible benchmark for a known valid procedure. Considering the validation run A-R of the proposed relevance model, some outliers have lower values than A-A, while overall equally high p-values are exhibited with no single p-value failing the test. Contrary, the verification runs of known invalid criteria A-RV1 and A-RV2 both resulted in significantly lower p-values. While all p-values for the A-RV1 run are below the acceptance threshold, the p-values for the A-RV2 run are scattered both above and below the threshold. Since individual p-values for A-RV2 fail the test, the whole validation run is considered failed and the A-RV2 is thus detected as invalid as expected.

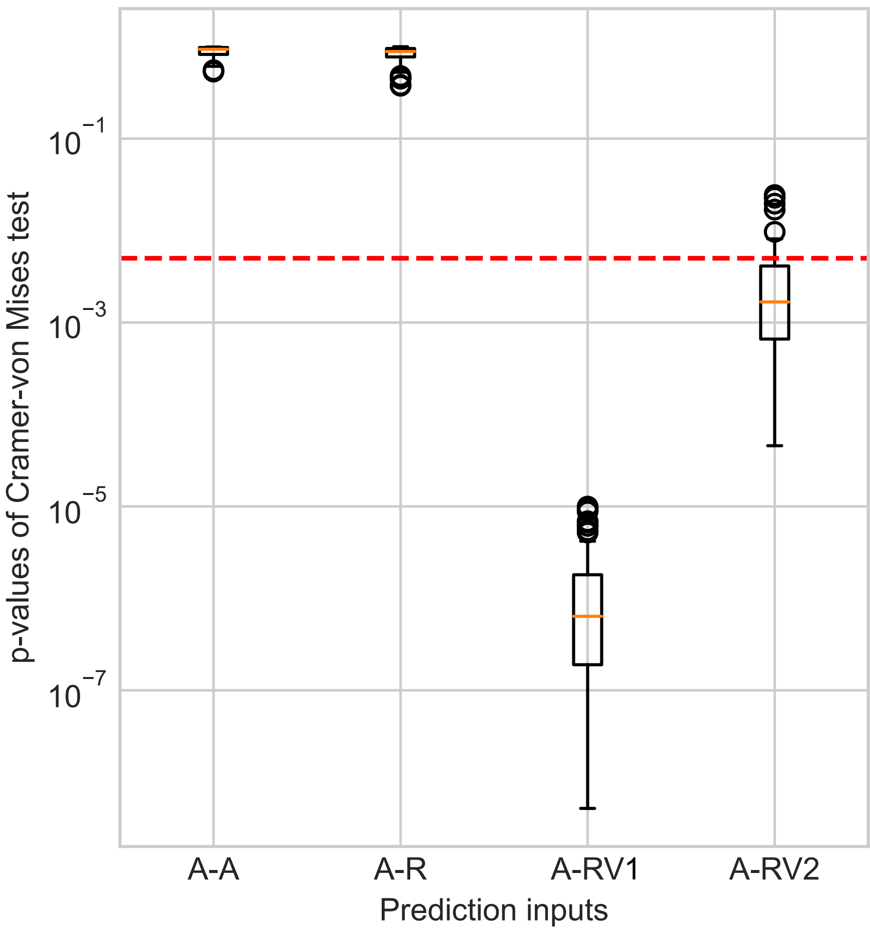

Figure 6-13: Results of the performed validation runs, with A-A being the reference, A-R being the model under test, and A-RV1 and A-RV2 being invalid verification models

6.6 Interim Conclusion

The previous chapter 5 established the research question of how relevance can formally be described. To contribute to the resolution of this question, a methodology was presented that describes relevance in a two-staged fashion. First, the previous review of literature was compiled into a GRM reflecting the information requirements of AD and the present solutions from the state of the art. Further, both a method to derive relevance criteria and a method to validate the previously defined criteria were presented.

In this chapter, both methods were applied and evaluated in an experiment on the NuScenes dataset. For this purpose, a hypothetical subject system was defined, consisting of a high level ODD and a partial specification of the system. The experiment demonstrated how behavioral requirements are derived from normative documents using the example of the German StVo and how the relevance criteria are created.

Following the application to the Nuscenes dataset and discussion thereof, the validation method was evaluated. The execution of the validation method both validated the previously defined relevance model and verified the proposed validation method itself.

Concluding, an answer to the initially posed research question was presented, applied, and verified.

7 Prototypical Application of Data Reduction Concept

Having established possible answers to RQ. 1 and RQ. 2, two key portions have been considered. A data reduction methodology has been proposed to consider the challenges that emerge with the development of AD, and a relevance concept has been proposed to control the loss of information to target functions. As such, a final research question remains, concerning the applicability of the data reduction methodology:

> **Research Question 3**
> *What is the impact on the use case of relevance-driven data reduction?*

For the application of CADR, a prototypical implementation is presented. The prototype is then evaluated with respect to the risk of information loss and the potential benefit of data reduction. The effects of information loss are evaluated for two separate use cases. The first being the usage of reduced data in the inference of DNNs and the second being the usage of reduced data in the training of DNNs.

7.1 Use Case Definition & Partial System Specification

For the subsequent experiments, a use case and target task for the prototypical CADR implementation has to be defined. Here, the use case of the development and usage of perception functions for AD will be considered. A focus is placed on the development of the perception tasks of semantic segmentation and object detection. The resulting perception functions are thus considered the subject system, as specified later by the external task specification. While the relevance concept in chapter 6 was evaluated on the NuScenes dataset, it is insufficient due to its lack of ground truth semantic segmentation. As such, the following experiment must be performed on a different dataset. For this purpose, the Cityscapes[238] dataset is selected.

[238] Cordts, M. et al.: The Cityscapes Dataset for Semantic Urban Scene Understanding (2016).

7.2 Application of the proposed SRM

In order to provide an answer to RQ. 3, both the CADR approach in conjunction with a relevance model must be evaluated. This inherent coupling between the data reduction approach and the relevance model means that the impact can only be evaluated for both components, but not isolated for the data reduction approach. The impact resulting from the data reduction approach can only be approximated by comparing it to the potential impact of the relevance model. Using a highly conservative relevance model will inevitably lead to a low impact on the target task, while an aggressive relevance model has the potential to yield a higher impact.

Observing the relevance projection of the proposed relevance model on the cityscapes dataset, it can be seen that the conservative relevance judgments translate to a conservative influence within the data representation. Figures 7-1 and 7-2 show a representative scene from the dataset, both from a BEV perspective and from the camera sensor.

Within the cityscapes dataset, a minority of relevance judgments yield an irrelevant judgment as can be seen in figure 7-3. Only 45 % of scenes include any objects that are judged as irrelevant. Figure 7-4 shows the eCDF of the distances for each type of relevance judgment. Since the relevance model is strongly influenced by the distance of the respective object, irrelevant objects are on average further from the ego object and therefore the camera sensor, than relevant objects. This leads to only 0.1 % of the overall pixels of the training split being deemed as irrelevant. Excluding the scenes with no irrelevant pixels the mean irrelevant pixel ratio increases to 0.2 % with a maximum of 2 %. Therefore, the proposed relevance model is considered to be conservative for the evaluation of the data reduction approach.

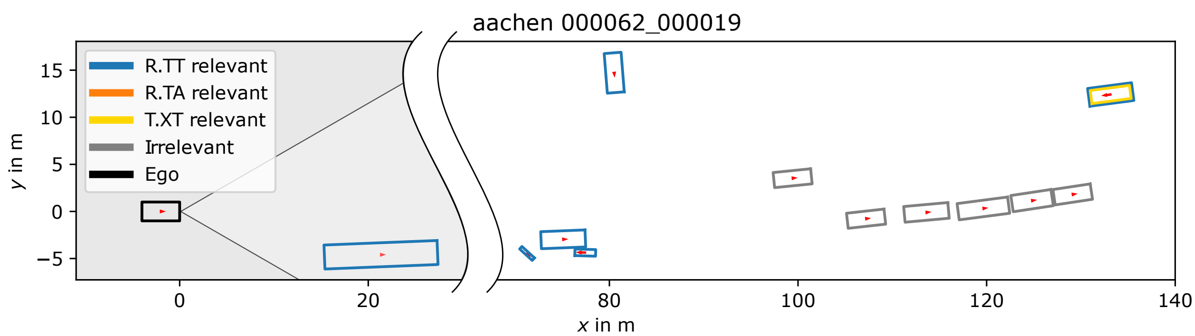

Figure 7-1: Birdseye-view of relevance judgments for the ”aachen_000062_000019” scene. Here, the parked vehicles at a distance of more than 100 meters were judged as irrelevant, as none of the relevance criteria applied to them. The oncoming vehicle exhibits a positive evaluation for both the R.TT and the T.XT criteria

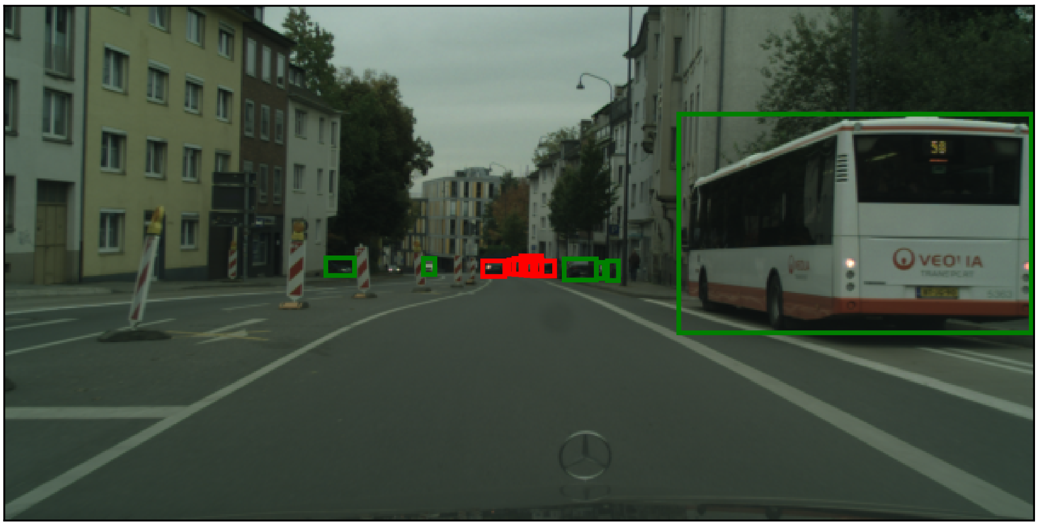

Figure 7-2: Ego view of relevance judgments in Cityscapes for the ”aachen_000062_000019” scene.

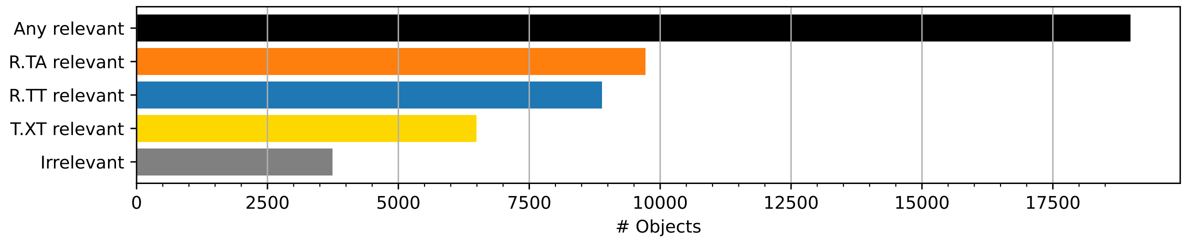

Figure 7-3: Absolute of number of occurrences for relevance judgments in the Cityscapes dataset.

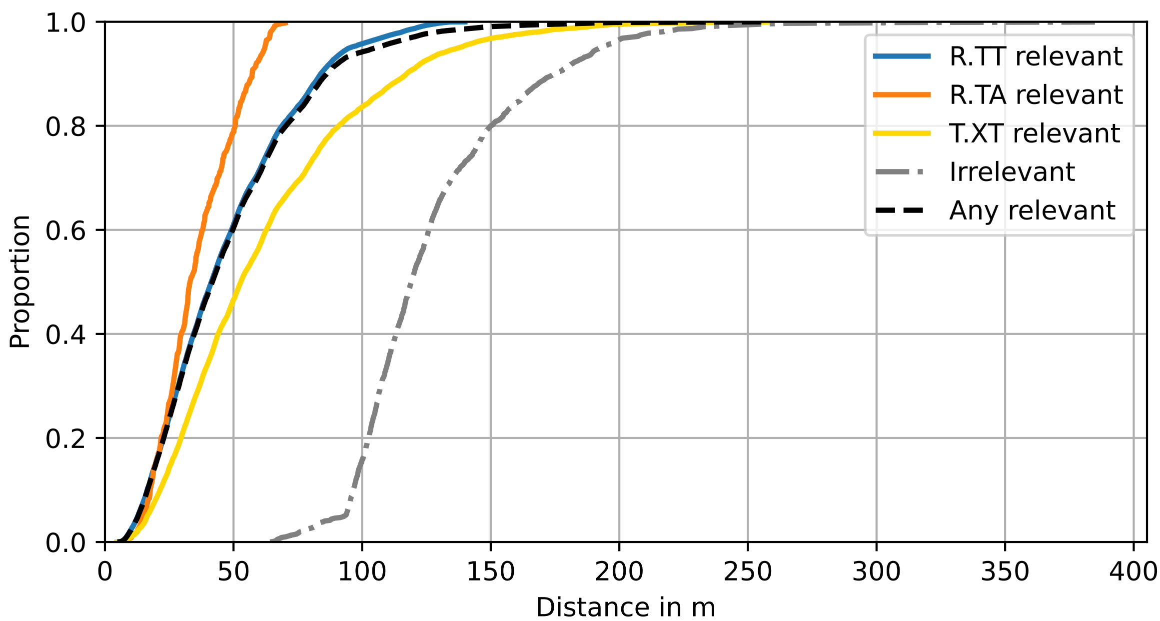

Figure 7-4: Distribution of distances of objects categorized by their assigned relevance judgment in the Cityscapes dataset.

7.3 Definition of a SRM for the Prototypical Application

As stated in the previous section, for the purpose of evaluating the proposed data reduction approach, a new SRM must be defined. Since the SRM is not the focus of this evaluation, only two requirements apply. First, it needs to be sufficiently aggressive to enable a strong impact on the target task. Second, the judgments must be feasible so that they do not cause an impact predominantly on the grounds of false judgments.

Using these two requirements a simple SRM was adopted from literature[239,240]. Here, using a topical relevance definition, certain classes are deemed relevant, based on their assumed utility in performing a DDT. Relevant are all traffic participants, as well as all entities which convey normative behavioral information for the traffic environment, such as traffic lights and signs, and lane markings. In addition to the classes already annotated by the ground truth provided by the Cityscapes benchmark for the pixel-level semantic labeling task[241], lane markings are added to include all plausible traffic infrastructure. The resultant relevance concept is thus described by the set of relevant split of the whole class taxonomy as given by equations 7-1 to 7-3.

$$\mathbb{C} = \mathbb{C}_{\text{relevant}} \cup \mathbb{C}_{\text{irrelevant}} \tag{7-1}$$

$$\mathbb{C}_{\text{relevant}} = \left\{ \begin{matrix} \text{bicycle, bus, car, motorcycle, person, rider,} \\ \text{traffic light, traffic sign, train, truck, lane marking} \end{matrix} \right\} \tag{7-2}$$

$$\mathbb{C}_{\text{irrelevant}} = \{\text{building, fence, pole, sidewalk, sky, terrain, vegetation, wall, road surface}\} \tag{7-3}$$

7.4 Implementation

For the whole experiment, the process as shown in figure 7-5 will be implemented, consisting of the three sub processes. The data reduction process is the implemented CADR data reduction, as depicted in figure 7-6. The implementation includes both the reduction and expansion steps. Thus, the original dataset $\mathbb{D}$ containing the images I is transformed to the reduced form denoted by the check accent $\check{\mathbb{D}}$ and $\check{I}$ and finally to the expanded form $\hat{\mathbb{D}}$ and $\hat{I}$, denoted by a circumflex.

[239] Wang, Y. et al.: A Two-stage H.264 based Video Compression Method for Automotive Cameras (2022).

[240] Wang, Y. et al.: Semantic-Aware Video Compression for Automotive Cameras (2023), p.2.

[241] Cityscapes Dataset: Benchmark Suite (2023).

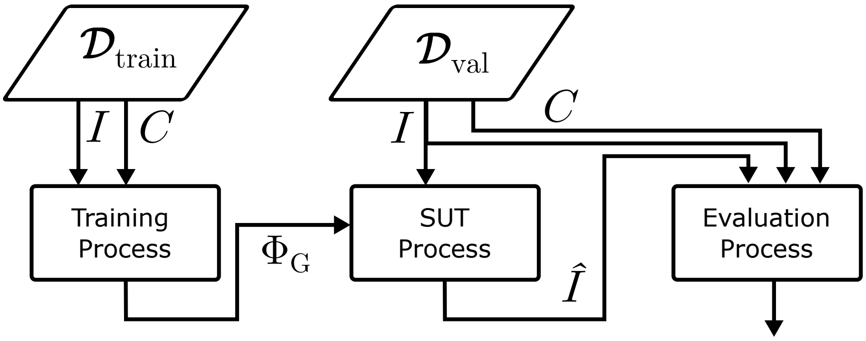

Figure 7-5: Data flowchart for the application and evaluation of CADR. The process is divided into a training, data reduction, and evaluation subprocess.

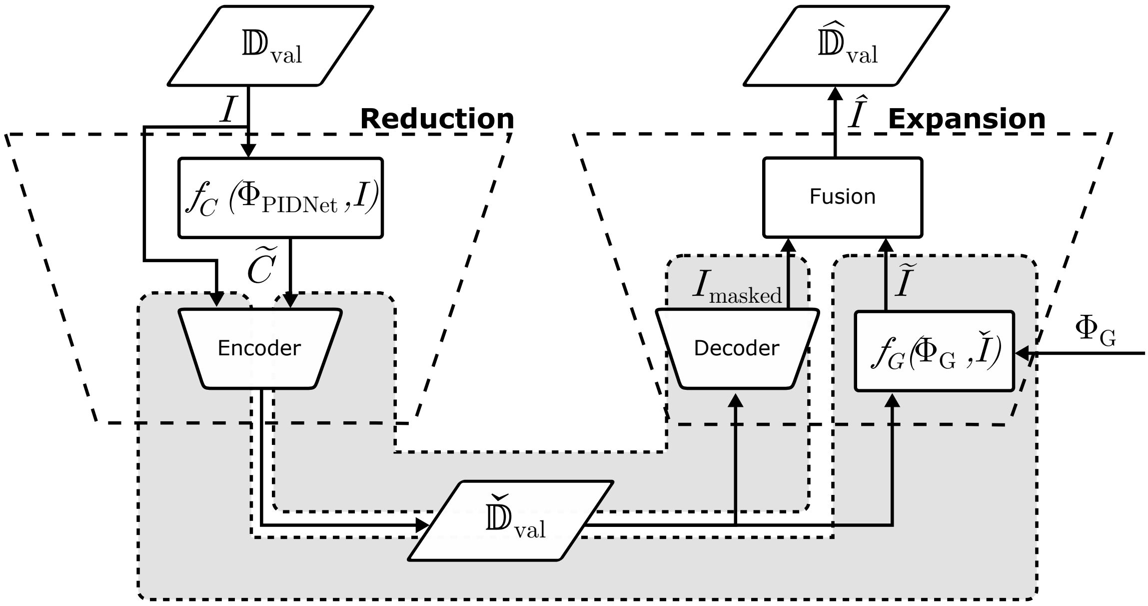

Figure 7-6: Data flowchart for the CADR data reduction subprocess.

Two data reduction pipelines are utilized, one for irrelevant judgments and one for relevant judgments. For the data reduction pipeline of the irrelevant data, a GAN f_{G} was selected to utilize the class information and relevant image data as a reduced storage format. The model parameters Φ_{G} are the result of the training process shown in figure 7-7. While the dataset already includes class labels C for the ground truth evaluation, only inferred class labels $\tilde{C}$ are utilized in the data reduction pipelines. In comparison, usage of inferred class labels $\tilde{C}$ is considered more plausible in a real application setting as provisioning the manually annotated class labels C is a resource intensive and expensive task. As such, C is only used in the training process, while $\tilde{C}$ is used in the data reduction process. The class labels are inferred by using a pretrained semantic segmentation function f_C with the model parameters Φ_{PIDNet}.

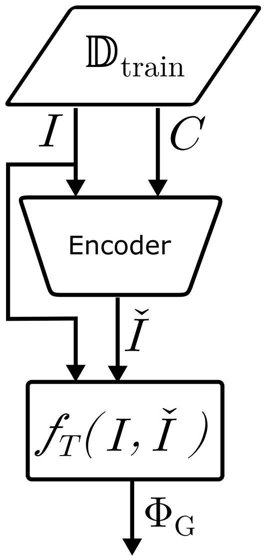

Figure 7-7: Data flowchart for the CADR training subprocess.

7.4.1 Data Reduction Module

The following describes the implementation of the data reduction module, by following the architecture defined in section 4.2.

Data Abstraction

The data abstraction module is realized by the semantic segmentation function f_C, yielding the inferred semantic map $\widetilde{C}$. As such, the semantic map is the representation space for the application of the relevance metric.

Relevance Metric & Relevance Concept

Due to choosing the semantic map as the representation space for the abstracted data, the relevance judgments can be expressed as a binary mask. Considering the mask relevant to the subject system, it is composed of those pixels that include semantic information defined as relevant in section 7.3. In this implementation, these are effectively given by selecting the relevant classes $\mathbb{C}_{\text{relevant}}$ from the semantic map $\widetilde{\boldsymbol{C}}$.

$$\boldsymbol{M}_{\text{class},i,j} = \begin{cases} 1, if\ \widetilde{\boldsymbol{C}}_{i,j} \in \mathbb{C}_{\text{relevant}} \\ 0, otherwise \end{cases} \tag{7-4}$$

[242] Stockman, George and Shapiro, Linda G.: Computer Vision (2000). a: p. 97; b: pp. 119; c: p. 81.

Since lane markings are also defined as relevant, but are not part of the predetermined set of classes of the cityscapes dataset, their detection must be handled separately. Algorithm 1 `DetectLaneMarking` details how an additional mask for lane markings $\boldsymbol{M}_{\text{lanemarking}}$ is determined. The algorithm detects lane markings based on the assumption, that the pixels of a lane marking have a higher brightness value than the pixels of the surrounding road. As such, a threshold determination[242a] is used to distinguish lane markings. Since due to changing brightness and road conditions the optimal threshold value v_{thresh} does not remain constant, the k-means algorithm[242b] is employed. Further, an opening operation[242c] is employed by the usage of algorithm 2 `ResizeMask`. Here, the mask is first eroded and then dilated by a given minimum size in order to enforce this minimum size for each detected lane marking.

Algorithm 1 : DetectLaneMarking$(\boldsymbol{I}, \boldsymbol{C})$

`Set minimum acceptable size of of pixel cluster in final mask.`

$s_{\min} \leftarrow 2$

foreach *i,j* **do**

`Create a mask for pixels of class road or ground`

$$\boldsymbol{M}_{\text{road},i,j} \leftarrow \begin{cases} 1, when\ \boldsymbol{C}_{i,j} \in \{c_{\text{road}}, c_{\text{ground}}\} \\ 0, otherwise \end{cases}$$

end

`Create image only containing road by applying road mask.`

$\boldsymbol{I}_{\text{road}} \leftarrow \boldsymbol{I} \odot \boldsymbol{M}_{\text{road}}$

`Conversion of RGB image to greyscale image by averaging all the k=3 color channels.`

$\boldsymbol{I}_{\text{road,grey},i,j} \leftarrow \frac{\sum_{k=1}^{3} \boldsymbol{I}_{\text{road}i,j,k}}{3}$

`Determines two centroids of the brightness distribution`

$\mu_1, \mu_2 \leftarrow \text{KMeans}(\boldsymbol{I}_{\text{road,grey}}, 2)$

`Set threshold half way between both distribution means.`

$v_{\text{thresh}} \leftarrow \frac{\mu_1, + \mu_2}{2}$

`Conversion of RGB image to greyscale image by averaging all the k=3 color channels.`

$\boldsymbol{M}_{\text{lanemarking}} \leftarrow \boldsymbol{I}_{\text{road,grey}} > v_{\text{thresh}}$

`Filter detections below minimum size`

$\boldsymbol{M}_{\text{lanemarking}} \leftarrow \text{ResizeMask}(\text{ResizeMask}(\boldsymbol{M}_{\text{lanemarking}}, -(s_{\min} - 1)), (s_{\min} - 1))$

return $\boldsymbol{M}_{\text{lanemarking}}$

Algorithm 2 : $\mathrm{ResizeMask}(\boldsymbol{M}, n)$

$\boldsymbol{K}_{\text{resize}} \leftarrow \begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}$; Convolution Kernel

if $n < 0$ **then**

  Perform erosion operation as an opening operation on the inverse mask.

  $\boldsymbol{M} \leftarrow \neg\mathrm{Resize}(\neg\boldsymbol{M}, -n)$

else

  Perform dilation operation n times.

  $i \leftarrow 0$

  repeat

    The dilation operation is performed through a convolution operation with the kernel. Through normalization the matrix is again transformed to a mask, where 0 indicates irrelevant entries and 1 indicates relevant entries.

    $\boldsymbol{M} \leftarrow |\boldsymbol{M} * \boldsymbol{K}_{\text{resize}}| \oslash (\boldsymbol{M} * \boldsymbol{K}_{\text{resize}})$

    $i \leftarrow i + 1$

  until $i = n$;

end

return M

Considering the imperfect accuracy of the inferred semantic classes $\tilde{C}$, an assurance concept is needed to account for this inaccuracy. While the semantic segmentation function f_C is trained to neither over- nor underestimate the semantic classes within an image, in terms of relevance an overestimation of relevant classes is acceptable. As such an overestimation of relevant classes can be assured by dilating the class mask, yielding $\boldsymbol{M}_{\text{assurance}}$ as defined by equation 7-5.

$$\boldsymbol{M}_{\text{assurance}} = \mathrm{ResizeMask}(\boldsymbol{M}_{\text{class}}, 1) \tag{7-5}$$

Considering the mask relevant to the expansion process, it is composed of pixels that include information which helps improve the performance of the expansion as described in section 7.4.3. In this implementation, two kinds of information of the original data are extracted and persisted through the reduction for this purpose. The first being the border areas between two semantic classes, the second being single pixels scattered throughout the irrelevant data.

Expanding the reduced data to accurately represent the area where two semantic classes border each other is important for later usage of the expanded data, so that the extent of each class instance is preserved. As such, the borders between irrelevant classes are selected as relevant to the expansion process, represented by the mask $\boldsymbol{M}_{\text{border,irrelevant,+}}$:

$$\boldsymbol{M}_{\text{border,irrelevant},+} = \text{ResizeMask}(\boldsymbol{M}_{\text{border,irrelevant}}, 1) \tag{7-6}$$

$\boldsymbol{M}_{\text{border,irrelevant},+}$ is dilated in a similar fashion like $\boldsymbol{M}_{\text{assurance}}$. The undilated mask $\boldsymbol{M}_{\text{border,irrelevant}}$ is derived by detecting the borders using the negated class mask for irrelevant semantic classes. This is done by an edge detection convolution given by equations 7-6 and 7-8.

$$\boldsymbol{M}_{\text{border,irrelevant}} = |\neg \boldsymbol{M}_{\text{class}} * \boldsymbol{K}_{\text{border}}| \oslash (\neg \boldsymbol{M}_{\text{class}} * \boldsymbol{K}_{\text{border}}) \tag{7-7}$$

with

$$\boldsymbol{K}_{\text{border}} = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} = \boldsymbol{K}_{\text{border,vert}} * \boldsymbol{K}_{\text{border,horz}} = \begin{bmatrix} -1 \\ 1 \end{bmatrix} * \begin{bmatrix} -1 & 1 \end{bmatrix} \tag{7-8}$$

With the current information being preserved, the data exhibits the symptom that strongly influences the resulting quality of the expanded data. Relevant and irrelevant data points are not distributed in such a way that there is always a preserved relevant data point close to an irrelevant data point that is to be expanded. As such, when expanding such a distant data point, it has to be inferred solely based on the semantic data still present. Based on the architecture used, this leads to strong artifacts or missing structural information in the synthetic data.

To counter this effect single data points can be added to the selection mask in order to act as seed points for later generative expansion. For this purpose, $\boldsymbol{M}_{\text{noise}}$ is defined based on the threshold determination of a white noise matrix $\boldsymbol{W}_{i,j}$. In this implementation, the threshold υ_{noise} is arbitrarily set as 0.7.

$$\boldsymbol{M}_{\text{noise},i,j} = \boldsymbol{W}_{i,j} > \upsilon_{\text{noise}} \tag{7-9}$$

Relevance Binning

Due to the fact that a binary relevance concept has been chosen, the relevance judgments have already yielded a binning effect, assigning each piece of data to either the data reduction pipeline for relevant or irrelevant data.

Relevance Projection

Since the relevance judgments are already performed within the representation of the input data, no sophisticated projection step has to be applied. Instead, merely the relevance mask $\boldsymbol{M}_{\text{selection}}$ is composed of the several sub-masks, that can be grouped by their relevance either to the subject system or to the expansion process. This composition is performed by a union operation [243] as given by equations 7-10 to 7-12. In this implementation the explicit declaration of an assurance concept, which would be represented by an additional sub-mask $\boldsymbol{M}_{\text{assurance}}$ is omitted.

$$\boldsymbol{M}_{\text{selection}} = \boldsymbol{M}_{\text{relevant,system}} \cup \boldsymbol{M}_{\text{assurance}} \cup \boldsymbol{M}_{\text{relevant,expansion}} \tag{7-10}$$

$$\boldsymbol{M}_{\text{relevant,system}} = \boldsymbol{M}_{\text{class}} \cup \boldsymbol{M}_{\text{lanemarking}} \tag{7-11}$$

$$\boldsymbol{M}_{\text{relevant,expansion}} = \boldsymbol{M}_{\text{border,irrelevant,+}} \cup \boldsymbol{M}_{\text{noise}} \tag{7-12}$$

Finally, $\boldsymbol{M}_{\text{selection}}$ is to be applied to the input data $\boldsymbol{I}$, yielding the image data of only the pixels to be preserved $\boldsymbol{I}_{\text{masked}}$.

$$\boldsymbol{I}_{\text{masked}} = \boldsymbol{I} \odot \boldsymbol{M}_{\text{selection}} \tag{7-13}$$

7.4.2 Data Storage Format

In order to fully instantiate a prototypical data reduction pipeline, a storage format for the data needs to be defined. As the input, storage, and output data are image data, a wide array of commonly established formats is available. In the evaluated pipeline, both the input and output data are stored in the portable network graphics (PNG)[244] format, which has established itself as a de-facto standard for lossless image data storage. In order to make use of the state of the art encoding algorithms for further data reduction and to limit the evaluation to the context aware aspect of CADR, the PNG format is also chosen as the storage format.

Contrary to the input and output data, the storage data includes additional data to the red, green, and blue color channel data of the camera. This additional data, concerning the semantic class and its masking status, needs to be stored as well. The PNG format allows for up to four separate channels of data to be stored. These channels include the red, green, and blue data, which are also used in the input and output data, as well as the alpha channel A, which is used to convey opacity for each image pixel. Thus, when combined semantic class and masking status can be stored in the alpha channel. Figure 7-8 show the bit structure for a single pixel of the reduced data. As shown the alpha channel stores the bits of the semantic class with the masking status bit appended. This effectively results in duplicating the class hierarchy as seen in figure 7-9, with

[243] Luce, R. D.: A Note on Boolean Matrix Theory (1952), p. 382.

[244] International Standardization Organization: PNG: Functional specification (2023).

each class being split into a relevant and irrelevant counterpart. Further, since one bit is used up by the masking status data, the maximum size of the class hierarchy is halved. For a standard 8-bit PNG format, this approach is limited to a maximum of 128 classes. When using the 16-bit support of the standard, up to 32768 classes can be supported.

$$\boldsymbol{A} = \widetilde{\boldsymbol{C}} + \boldsymbol{M}_{\text{selection}} \cdot 2^{b_{\text{class}}} \tag{7-14}$$

$$\check{\boldsymbol{I}} = concat(\boldsymbol{I}_{\text{masked}}, \boldsymbol{A}) = \begin{bmatrix} \boldsymbol{I}_{\text{masked}} \\ \boldsymbol{A} \end{bmatrix} \tag{7-15}$$

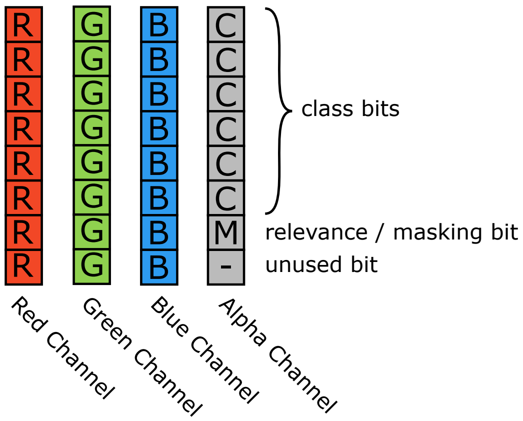

Figure 7-8: Structure of the compressed and encoded data, for a single pixel of $\check{\boldsymbol{I}}$. Each box represents a single bit and each column represents an 8-bit channel.

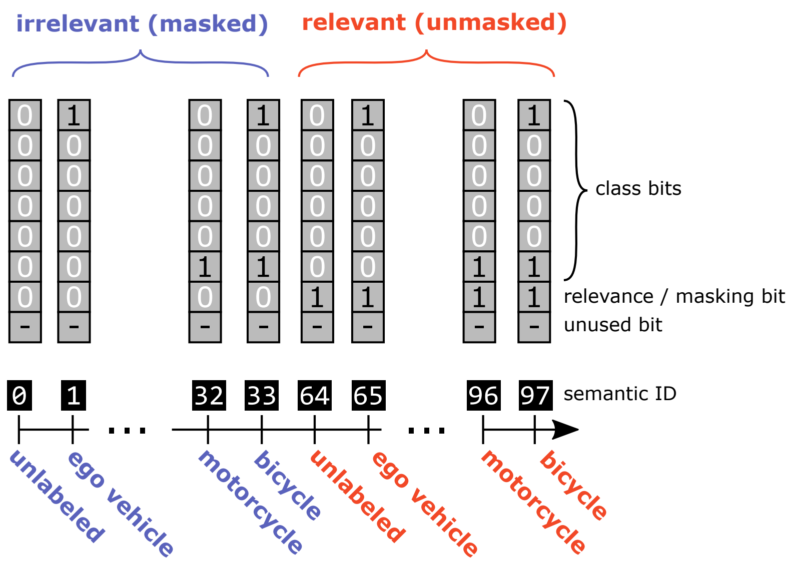

Figure 7-9: Bit values of the Alpha channel containing class and relevance information for the used class structure.

7.4.3 Data Decompression Module

From the unified storage format, the data for both expansion pipelines is extracted. This is done by first regenerating the relevance mask $\boldsymbol{M}_{\text{relevance}}$ from the PNGs alpha layer as defined in equation 7-16, which will be used to extract the unmodified relevant data in the fusion step. Equation 7-17 describes the data synthesis for the irrelevant data, here the whole stored information is used. The fusion is performed as described by equation 7-18, yielding the final expanded image $\hat{\hat{\boldsymbol{I}}}$.

$$\boldsymbol{M}_{\text{relevance},i,j} = \boldsymbol{A}_{i,j} \geq 2^{b_{\text{class}}} \tag{7-16}$$

$$\widetilde{\boldsymbol{I}} = f_G(\Phi_{\text{G}}, \check{\boldsymbol{I}}) \tag{7-17}$$

$$\hat{\boldsymbol{I}} = \widetilde{\boldsymbol{I}} \odot \neg \boldsymbol{M}_{\text{relevance}} + \boldsymbol{I}_{\text{masked}} \odot \boldsymbol{M}_{\text{relevance}} \tag{7-18}$$

In order to obtain the generative model $\boldsymbol{\Phi}_{\mathbb{D}_{\text{train}}}$, the selected GAN was trained on the original and reduced training dataset, $\mathbb{D}_{\text{Full,train}}$ and $\mathbb{D}_{\text{synth,train}}$ for 130 epochs. Contrary to other DNNs, GANs do not converge towards zero in their loss function, but instead approach an unstable equilibrium.[245] Knowing this, the training was stopped once the generator output was subjectively deemed sufficient for the task at hand.

[245] Google LLC: Machine Learning Common Problems (2023).

8 Evaluation of the Data Reduction Concept

In chapter 4, a novel concept for relevance-driven data reduction was presented. To fully understand this concept, its impact on the quality of a given data oriented process needs to be determined. As stated in section 2.3, the quality of a data oriented process is defined through its effectiveness and its efficiency. With these characteristics, the aim of this chapter is twofold, centered around the research question 3.

The first aim is to provide the means to answer the research question. Which methods are necessary to achieve a fully qualified evaluation of relevance-driven data reduction? The second aim is to provide an answer to the research question for the prototypical implementation for CADR as a relevance-driven data reduction, as outlined in chapter 7. The presented prototype will be evaluated with respect to the impact on the effectiveness and efficiency of the data. For this evaluation, the impact on effectiveness will be quantified through the risk of information loss. The impact on the efficiency of the data is defined as the data reduction factor.

Since the specifics of the use case are chosen and therefore the tasks and SRM are artificially derived specifically for this dissertation, no acceptance criteria exist. For any substantiated acceptance criteria, a more complex application environment, like in an actual industrial context would be required. As such, the answer to the research question will be in the form of a qualitative discussion at the end of this chapter.

8.1 Information Loss / Applicability for Use Case

Loss of information and a worsened usability for a given use case can generally be described by the loss of performance of a task within the considered use case. So, in order to evaluate the loss of information introduced by the data reduction, the kind of task considered must first be defined. Since the prototypical implementation is concerned with image data, the task of perception is usually the first task in any kind of data processing. Further, a perception task can either be performed by a machine or by a human.

Therefore, in the next sections, first the information loss in machine perception will be evaluated. Second, the information loss for machine perception training will be evaluated. Finally, a method to evaluate the information loss for human perception is presented. For the purpose of these evaluations, the following datasets are created, as shown in figure 8-1:

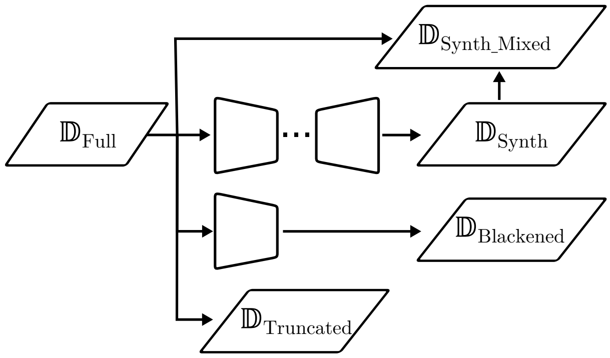

Figure 8-1: Data flowchart delineating the creation process for the used datasets

- $\mathbb{D}_{\text{Full}}$: The unaltered Cityscapes dataset. This dataset serves as a benchmark, since using the whole unaltered dataset is the default option of no data reduction being applied.
- $\mathbb{D}_{\text{Synth}}$: The processed dataset as proposed in chapter 7
- $\mathbb{D}_{\text{Mixed\_Synth}}$: This dataset is composed to equal parts of images from $\mathbb{D}_{\text{Full}}$ and $\mathbb{D}_{\text{Synth}}$. This dataset provides a basic understanding of how the approach proposed in chapter 7 behaves when applied to only parts of the data.
- $\mathbb{D}_{\text{Blackened}}$: The distinction between relevant and irrelevant parts of the data is a key aspect of CADR. While some data is irrelevant, a core principle of CADR is also that everything irrelevant must still be present as a ”plausible lie”. This dataset is created to also consider the distinction of relevance, while at the same time completely omitting the irrelevant data. As such, it serves as a baseline alternative to $\mathbb{D}_{\text{Synth}}$.
- $\mathbb{D}_{\text{Truncated}}$: A fraction of the Full dataset. This dataset is roughly equal in data size to the Synth dataset. To mimic the process of truncation in reality, whole recordings of cities were removed from the dataset until it was only slightly larger in data size when compared $\mathbb{D}_{\text{Synth}}$. This mirrors a recording campaign being shorter and performing certain recording drives to provision extra data. This dataset provides an additional baseline, since reducing the overall samples is the most simplistic alternative to $\mathbb{D}_{\text{Synth}}$.

In order to place these datasets into a data reduction context, their respective data sizes are listed in table 8-1. Each dataset can be further distinguished into training $\mathbb{D}_{-,\text{train}}$, validation $\mathbb{D}_{-,\text{val}}$ and testing $\mathbb{D}_{-,\text{test}}$ sub datasets. Since the public data does not include semantic labels for the testing sub dataset it will not be further considered in this evaluation. In figure 8-2 samples are presented from selected datasets.

Table 8-1: Dataset sizes in GB for experiment setup

	$\mathbb{D}_{\text{Full}}$	$\mathbb{D}_{\text{Synth}}$	$\mathbb{D}_{\text{Truncated}}$	$\mathbb{D}_{\text{Mixed}_\text{Synth}}$	$\mathbb{D}_{\text{Blackened}}$
RGB	6.46	2.84	3.01	4.65	1.42
Label	0.62	0.27	0.36	0.45	-
Sum	**7.08**	**3.11**	**3.37**	**5.10**	**1.42**

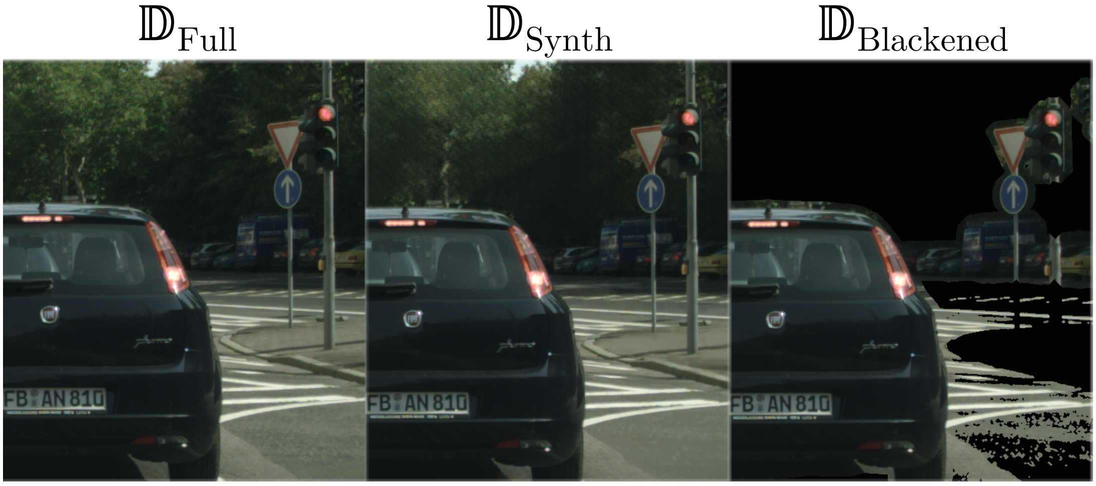

Figure 8-2: Sample images of the considered datasets for evaluation

8.1.1 Information Loss for Machine Perception

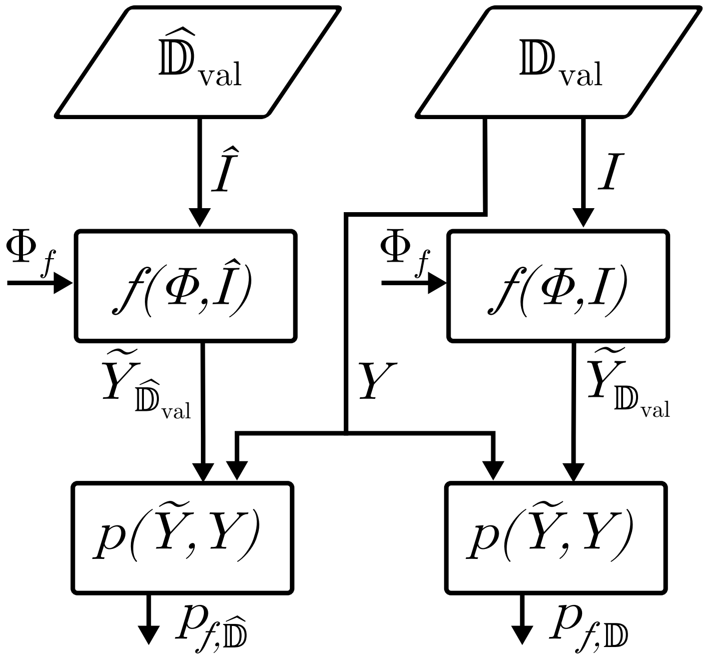

Figure 8-3: Data flowchart for the CADR inference evaluation subprocess. In this evaluation f is realized by either f_C or f_O, depending on which task is being evaluated.

In order to evaluate the loss of information introduced by the data reduction, the process as depicted in figure 8-3 is proposed. For the evaluation, two subject tasks will be considered, semantic segmentation and object detection. Functions implementing solutions to the respective tasks will be denoted as f_C and f_O. For both functions publicly available pretrained models Φ are used. The loss of information can be estimated by measuring the change in performance for each subject task. $\mathbb{D}_{\text{Full,val}}$ will be used as a baseline reference from which the change is measured. The loss of performance will be determined for both the proposed $\mathbb{D}_{\text{Synth,val}}$ dataset and the $\mathbb{D}_{\text{Black,val}}$ dataset as an alternative.

Using the two subject tasks, the experiments have two focuses. First, determining if the proposed approach has the property of loss separability, and if so, to what degree. Second, is the usage of the proposed approach feasible against baseline methods such as those represented by $\mathbb{D}_{\text{Black}}$.

Semantic Segmentation

First, the separability of performance loss between relevant and irrelevant entities will be evaluated. For this purpose, the task of semantic segmentation will be considered. Semantic segmentation is suitable for this evaluation since it requires a class assignment to all pixels, including those deemed irrelevant.

The changes in performance are determined by calculating the mean intersection over union (mIoU) score. As a subject function for the semantic segmentation task PIDNet[246] is selected as one of the top performing[247] semantic segmentation DNN for the used Cityscapes dataset. Evaluating the quality of semantic segmentation the resulting mIoU performance measures as listed in table 8-2 can be observed.

Table 8-2: Quality of a Semantic Segmentation based on synthetic images as compared to that of original images

	Class	$p_{f_{\mathrm{C}},\mathbb{D}_{\text{Full,val}}}$ mIoU	$p_{f_{\mathrm{C}},\mathbb{D}_{\text{Synth,val}}}$ mIoU	$p_{f_{\mathrm{C}},\mathbb{D}_{\text{Blackened,val}}}$ mIoU
	Ø	**0,809**	**0,777 (-3.96%)**	**-**
$\mathbb{C}_{\text{relevant}}$	**bicycle**	0,789	0,788 (±0%)	0,599 (-24%)
	bus	0,905	0,899 (-1%)	0,351 (-61%)
	car	0,955	0,949 (-1%)	0,008 (-16%)
	motorcycle	0,672	0,661 (-2%)	0,084 (-88%)
	person	0,842	0,834 (-1%)	0,224 (-73%)
	rider	0,687	0,682 (-1%)	0,177 (-74%)
	road	0,983	0,977 (-1%)	0,687 (-30%)
	traffic light	0,739	0,729 (-1%)	0,284 (-62%)
	traffic sign	0,814	0,814 (±0%)	0,427 (-48%)
	train	0,858	0,854 (±0%)	0,231 (-73%)
	truck	0,811	0,819 (+1%)	0,159 (-80%)
	Ø	**0,823**	**0,819 (-0.49%)**	**0,366 (-56%)**
$\mathbb{C}_{\text{irrelevant}}$	**building**	0,932	0,901 (-3%)	-
	fence	0,664	0,551 (-17%)	-
	pole	0,666	0,632 (-5%)	-
	sidewalk	0,865	0,822 (-5%)	-
	sky	0,944	0,927 (-2%)	-
	terrain	0,666	0,581 (-13%)	-
	vegetation	0,926	0,898 (-3%)	-
	wall	0,651	0,448 (-31%)	-
	Ø	**0,789**	**0,720 (-8.75%)**	**-**

[246] Xu, J. et al.: PIDNet: A Real-time Semantic Segmentation Network Inspired by PID Controllers (2022).

[247] paperswithcode.com: Cityscapes test Benchmark (Real-Time Semantic Segmentation) (2023).

The observed performance for $\mathbb{D}_{\text{Full,val}}$ is consistent with the stated mIoU value of 0.809[248]. While on the partially synthetic counterpart, the performance is *0.777*. As the prototype implementation is a lossy type of data reduction, the overall loss of information is in line with expectations. Considering this relative difference in performance would result in a reranking of the approach from first to a new sixth place, as depicted in figure 8-4. In order to further examine and interpret the information loss, a distinction between the relevant and irrelevant data within the validation must be made.

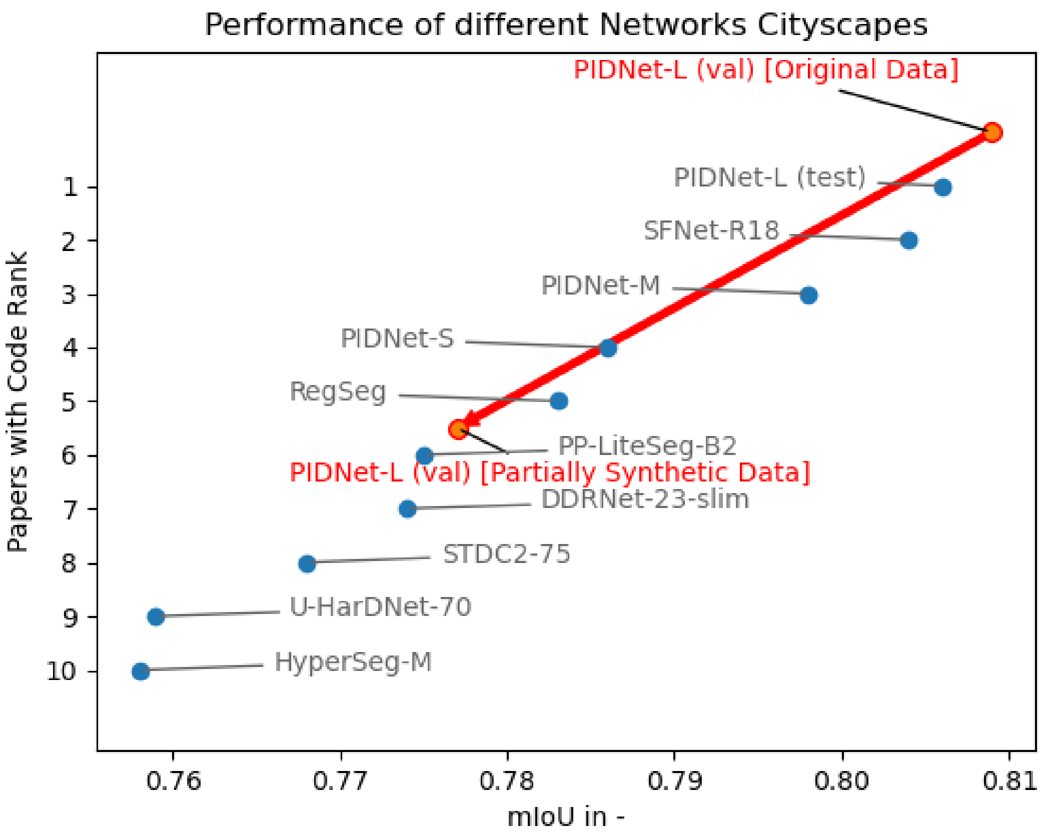

Figure 8-4: Ranking of the mIoU on the original data and the mIoU on the reduced data compared to the Papers with Code top 10 leaderboard.

Within the context of this evaluation it is the goal to observe the influence of the used data on the performance of a subject system. Hence only the data relevant to the subject function will be considered as relevant. The data relevant to the reduction process will be considered equally irrelevant as the rest of the data.

Having made this distinction, the performance values for only the relevant and irrelevant parts of data can be evaluated. Using $\mathbb{D}_{\text{Full,val}}$ exhibits a performance of *0.823* on the relevant and *0.789* for the irrelevant data. Substituting $\mathbb{D}_{\text{Full,val}}$ with $\mathbb{D}_{\text{Synth,val}}$ the performance observed is now *0.819* on the relevant and *0.720* for the irrelevant data.

[248] github.com/XuJiacong/PIDNet: This is the official repository for our recent work: PIDNet (2023).

This shows that the initially observed *3.98%* degradation of performance does not equally impact the relevant and irrelevant data. The resulting performance for the relevant data is only lessened by *0.49%* while for the irrelevant data the reduction in performance is *8.75%*. As a result, the loss of performance for the relevant data is *18* times less prevalent than it is within the irrelevant data. Table 8-2 further shows the performance results for object classes, grouped by the class relevance. While there is a difference in performance among the individual object classes, overall there are no significant outliers that contradict the observed deviation between relevant and irrelevant data. Further, within the performance values of the irrelevant classes only a small subset is responsible for most of the performance loss. For the classes, fence, terrain, and wall, it has to be noted that they exhibit a loss larger than ten percent, while all other classes experience a loss of less than or equal to five percent.

Considering the second evaluation focus of feasibility compared to baseline methods. Using $\mathbb{D}_{\text{Blackened,val}}$ as input only the performance of segmenting the relevant classes $\mathbb{C}_{\text{relevant}}$ can be considered, since the majority of data pertaining $\mathbb{C}_{\text{irrelevant}}$ is within blackened areas. On average and throughout the relevant classes, a significant drop in performance can be noted. The magnitude of the performance drop allows for the assumption, that the input data $\mathbb{D}_{\text{Blackened,val}}$ is out of distribution for the used neural network. Compared to $\mathbb{D}_{\text{Synth,val}}$, $\mathbb{D}_{\text{Blackened,val}}$ exhibits a 114 times larger performance loss on the relevant classes. While this observation in itself is not surprising, it shows that the *"plausible lie"* provided by $\mathbb{D}_{\text{Synth}}$ is suitable enough to not cause any out of distribution events and is feasible when compared to a baseline method without a *"plausible lie"*.

Object Detection

To fully evaluate the inference of the subject function, the task of object detection is considered next. Here, an object detection task is derived from the Cityscapes benchmark for the Instance-Level Semantic Labeling Task[249]. For all annotated instances of relevant classes, the extent of each instance was rewritten as a bounding box with centroid $(x_{\text{center}}, y_{\text{center}})$, width w and height h, as shown in figure 8-5, further including the respective class ID.

In order to evaluate the object detection task, a state of the art YOLOv8 nano neural network is used. For the object detection task, the mAP metric is used to measure the performance. In table 8-3 both the $\text{mAP}_{.5}$ and $\text{mAP}_{.5:.95}$ metrics, as described in section 2.3.2, are observed.

Averaged over all classes the performance change from $\mathbb{D}_{\text{Full,val}}$ to $\mathbb{D}_{\text{Synth,val}}$ is minimal with -1 % $\text{mAP}_{.5}$ and ± 0 % $\text{mAP}_{.5:.95}$. Looking at the individual object classes a fluctuation from improved and worsened performance can be seen. An outlier of note is the trailer class, which worsened by 39 % and 48 % respectively. A plausible explanation for this can be found in the number of instances to detect. While the whole validation dataset includes 16115 instances, only 13 or 0.1 % of the instances are of the trailer class. In the training set the same percentage are of the trailer class, resulting in only 61 samples for training. As such, it is plausible that this outlier is due to an

[249] Cityscapes Dataset: Benchmark Suite (2023).

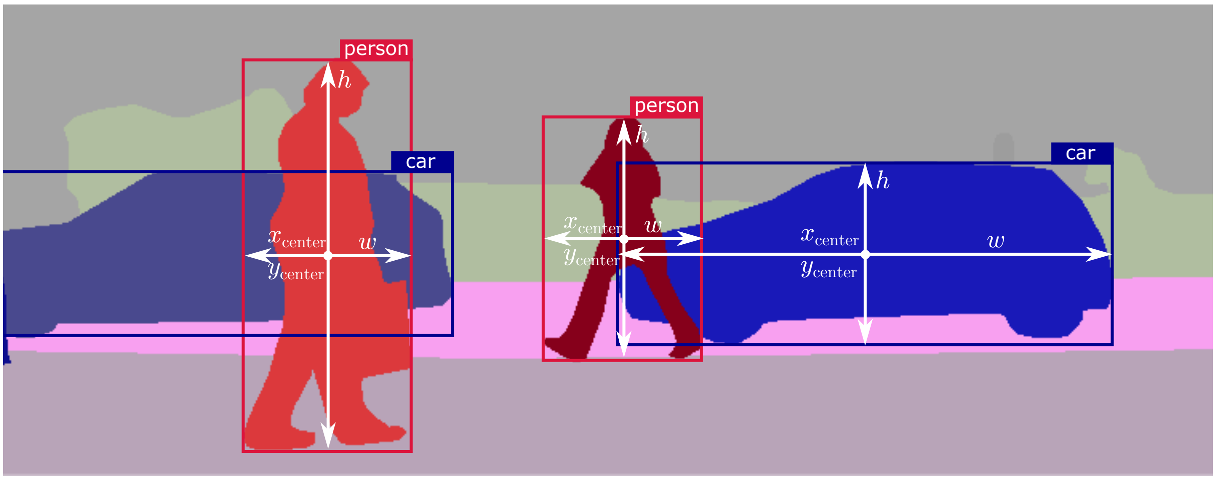

Figure 8-5: Generation from Panoptic Semantic Segmentation Task to an Object Detection Task.

Table 8-3: Performance values, produced by YOLOv8 and measured in mAP

Class	$p_{f_O,\mathbb{D}_{\text{Full,val}}}$ mAP$_{50}$	mAP$_{50\text{-}95}$	$p_{f_O,\mathbb{D}_{\text{Synth,val}}}$ mAP$_{50}$	mAP$_{50\text{-}95}$	$p_{f_O,\mathbb{D}_{\text{Blackened,val}}}$ mAP$_{50}$	mAP$_{50\text{-}95}$
Ø	**0,373**	**0,213**	**0,369 (-1%)**	**0,213 (±0%)**	**0,143 (-62%)**	**0,067 (-69%)**
traffic light	0,303	0,127	0,287 (-5%)	0,119 (-6%)	0,039 (-87%)	0,008 (-94%)
traffic sign	0,342	0,186	0,348 (+2%)	0,187 (+1%)	0,092 (-73%)	0,028 (-85%)
person	0,241	0,114	0,220 (-9%)	0,103 (-10%)	0,171 (-29%)	0,057 (-50%)
rider	0,487	0,269	0,487 (±0%)	0,271 (+1%)	0,121 (-75%)	0,042 (-84%)
car	0,710	0,476	0,708 (±0%)	0,473 (-1%)	0,434 (-39%)	0,221 (-54%)
truck	0,365	0,244	0,368 (+1%)	0,255 (+5%)	0,250 (-32%)	0,135 (-45%)
bus	0,566	0,386	0,589 (+4%)	0,409 (+6%)	0,321 (-43%)	0,166 (-57%)
caravan	0,334	0,205	0,310 (-7%)	0,220 (+7%)	0,046 (-86%)	0,025 (-88%)
trailer	0,052	0,039	0,032 (-39%)	0,020 (-48%)	0,008 (-84%)	0,005 (-87%)
train	0,387	0,192	0,377 (-3%)	0,178 (-7%)	0,092 (-76%)	0,053 (-73%)
motorcycle	0,280	0,115	0,296 (+6%)	0,127 (+10%)	0,024 (-91%)	0,013 (-88%)
bicycle	0,405	0,201	0,405 (±0%)	0,199 (-1%)	0,120 (-70%)	0,046 (-77%)

insufficient sample size, both in the training phase as well as in the evaluation phase. While the overall model yields results with low accuracy, the accuracy for the trailer class is especially low with 0.052 mAP$_{.5}$ and 0.039 mAP$_{.5:.95}$ as compared to the average of 0.373 and 0.213, indicating insufficient training samples. Additionally, for this class, the high relative change can also be understood as a numerical effect due to the low accuracy. In absolute numbers, the observed change is 0.02 and 0.19. This absolute magnitude is also observed for other classes.

In contrast to the evaluation of $\mathbb{D}_{\text{Synth,val}}$, using $\mathbb{D}_{\text{Blackened,val}}$ instead of $\mathbb{D}_{\text{Full,val}}$ causes a significant loss of performance of over 50 % on all scores and most object classes. On average a performance loss of -62 % mAP$_{.5}$ and -69 % mAP$_{.5:.95}$ is observed. Similarly to the performance in the semantic segmentation task, $\mathbb{D}_{\text{Blackened,val}}$ is not a feasible data reduction approach for perception.

8.1.2 Information Loss for Training Data

In subsection 8.1.1, the possibility to perceive information from reduced and unreduced data was evaluated. For the purpose of perceiving information from novel data, neural networks leverage learned information from data to which they were previously exposed to in a training process. In this subsection, the efficacy of reduced data compared to unreduced data for the usage as training data will be evaluated as described by the process in figure 8-6.

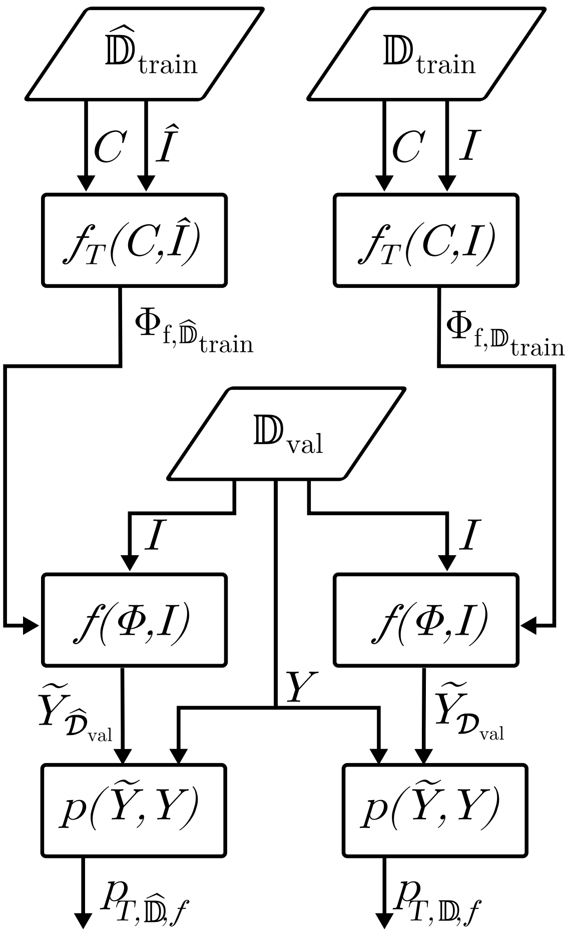

Figure 8-6: Data flowchart for the CADR training data evaluation subprocess.

To this point, the previously used neural networks PIDNet and YOLOv8 realizing the subject functions of semantic segmentation and object detection respectively will be reused. Contrary to subsection 8.1.1 not only the final performance of the trained neural net needs to be evaluated, but the development of the performance over the training process. During the training the performance development will be observed based on the metric used in subsection 8.1.1. Each network will be trained by using $\mathbb{D}_{\text{Full,train}}$, $\mathbb{D}_{\text{Synth,train}}$ and $\mathbb{D}_{\text{Truncated,train}}$.

$\mathbb{D}_{\text{Blackened,train}}$ is also used in training for the object detection task. Using this dataset to train a semantic segmentation function is omitted, since learning to segment labels of irrelevant classes from no available data does not result in a sensible task. While the segmentation task can be modified to also use $\mathbb{D}_{\text{Blackened,train}}$ as training data, this would yield a completely different task.

Semantic Segmentation

The training results for the semantic segmentation task are visualized in figure 8-7. As expected, training with the unmodified dataset $\mathbb{D}_{\text{Full,train}}$ clearly yields the best result, with the highest mIoU values in both the initial start phase and the beginning saturation phase.

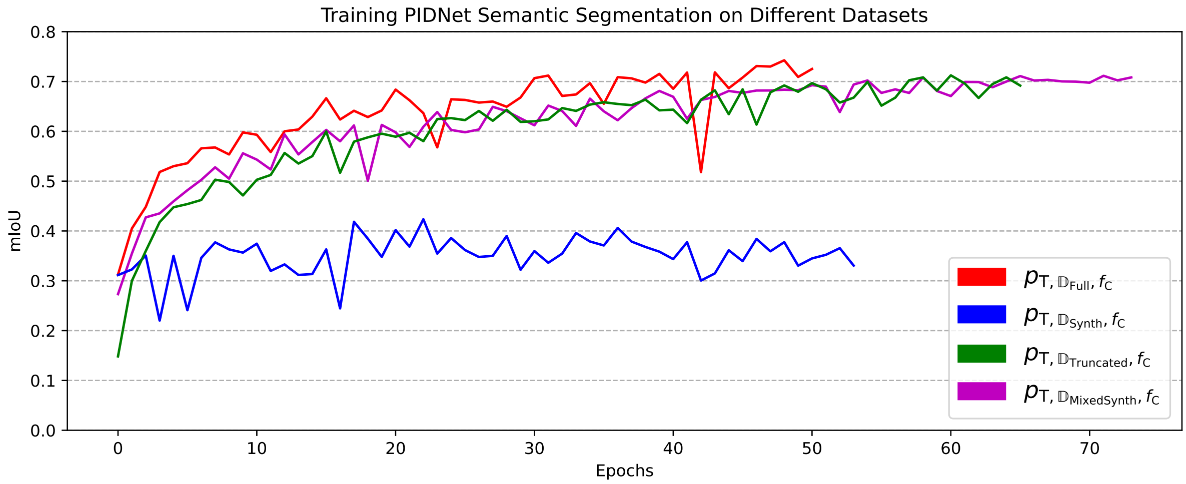

Figure 8-7: Quality of Semantic Segmentation based on training on the datasets in table 8-1

$\mathbb{D}_{\text{Truncated,train}}$ shows a similar training behavior as $\mathbb{D}_{\text{Full,train}}$ but with a seemingly constant absolute offset. On the other hand it $\mathbb{D}_{\text{Synth,train}}$ is observed to result in a different behavior where the performance remains constant with negligible progress during training. Here, $\mathbb{D}_{\text{Mixed,Synth,train}}$ is introduced as a further dataset to evaluate the effect on information extraction from synthetic data. The training progress of $\mathbb{D}_{\text{Mixed,Synth,train}}$ closely follows that of $\mathbb{D}_{\text{Truncated,train}}$. $\mathbb{D}_{\text{Truncated,train}}$ is composed of 50 % of the samples of $\mathbb{D}_{\text{Full,train}}$. $\mathbb{D}_{\text{Mixed,Synth,train}}$ includes the same amount of samples from $\mathbb{D}_{\text{Full,train}}$, plus the missing 50 % in the form of samples from $\mathbb{D}_{\text{Synth,train}}$. This information in conjunction with the near exact same training performance leads to the conclusion that in this experiment $\mathbb{D}_{\text{Synth,train}}$ does not contribute significant information to the training.

Object Detection

The cityscapes dataset was created for a semantic segmentation task. In order to train an object detection on cityscapes, a ground truth for this task had to be derived. For this purpose, the semantic segmentation label ground truth was used as a starting point. All classes with instance level annotation were considered, thus removing the road class from the relevant classes, due to its incompatibility with an object detection task. For each instance annotation, a 2-D bounding box was defined around the semantic instance segmentation, preserving the class information.

As a result, the subject neural network can be trained on this newly created ground truth data. Due to its derived origin, multiple small, hard to detect, objects are included in the dataset, leading to seemingly low performance scores. Since all trained neural networks are equally exposed to this effect, it can be ignored in the following evaluation.

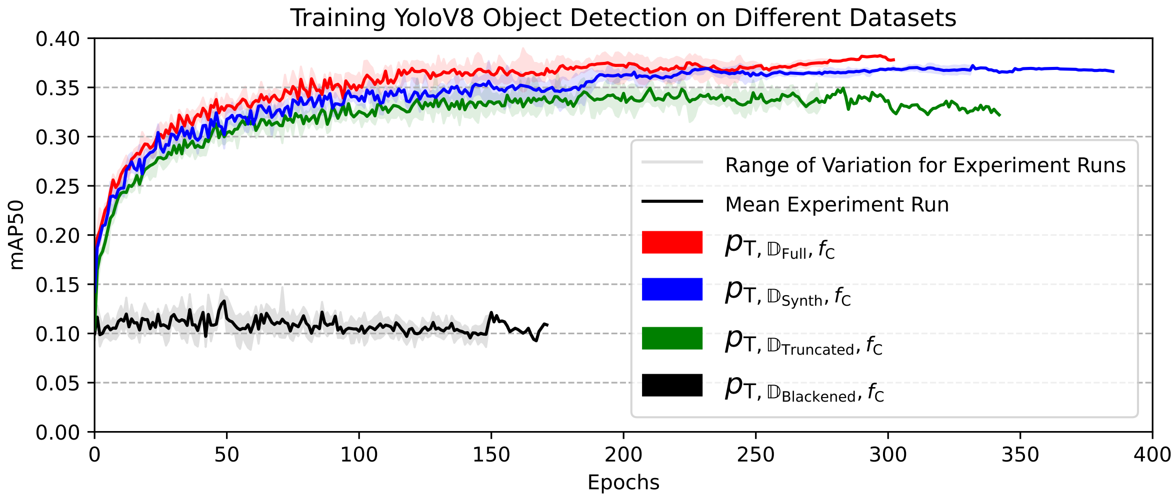

Figure 8-8: Quality of Object Detection based on training on the datasets in table 8-1

Figure 8-8 depicts the training process for a YOLOv8 neural network given the considered datasets. Contrary to the previous section 8.1.2 where the shortcomings of using synthetic data as training data were apparent, these results show a different result. As to not rely on a single stochastic process for a close evaluation, every training process was repeated five times. Displayed are the upper and lower limits as well as the mean value for each set of training processes.

Upon initial observation, $\mathbb{D}_{\mathrm{Full,train}}$, $\mathbb{D}_{\mathrm{Synth,train}}$ and $\mathbb{D}_{\mathrm{Truncated,train}}$ perform comparably, while $\mathbb{D}_{\mathrm{Blackened,train}}$ is clearly visible as not yielding any improvements. This lack of training effect is also visible in the triggering of the early stopping criterion of the YOLOv8 framework, where the training process is stopped if no improvement has been registered in the last 100 epochs.

Having a more detailed look at $\mathbb{D}_{\mathrm{Full,train}}$, $\mathbb{D}_{\mathrm{Synth,train}}$ and $\mathbb{D}_{\mathrm{Truncated,train}}$, all appear to have a similar, albeit offset, training performance. Again, as expected, $\mathbb{D}_{\mathrm{Full,train}}$ exhibits the best performance, with $\mathbb{D}_{\mathrm{Synth,train}}$ and $\mathbb{D}_{\mathrm{Truncated,train}}$ performing worse by an offset of roughly 0.02 and 0.04 mAP respectively. A fact to highlight is the circumstance, that all datasets reached saturation by triggering the early stopping criterion. Here, it is intuitive to try to offset the worsened performance of $\mathbb{D}_{\mathrm{Truncated,train}}$ by extending the number of training epochs, since due to the decreased number of samples in the dataset, each training epoch uses less computational resources. However, this approach does not seem promising, since the early stopping is indicative of a complete utilization of the available information.

8.1.3 Information Loss for Human Perception

In order to evaluate information loss for human perception, similarity between the partially synthetic image and the original image for a human viewer needs to be considered. To get a quick overview of the general quality a metric built to approximate human perception can be used. For this purpose, the MS-SSIM[250] metric discussed in section 2.3.2 will be applied. Figure 8-9 shows the MS-SSIM value eCDF when comparing $\mathbb{D}_{\mathrm{Synth}}$ to $\mathbb{D}_{\mathrm{Full}}$. The Values range from 0.75 to 0.95 with a median value of 0.89.

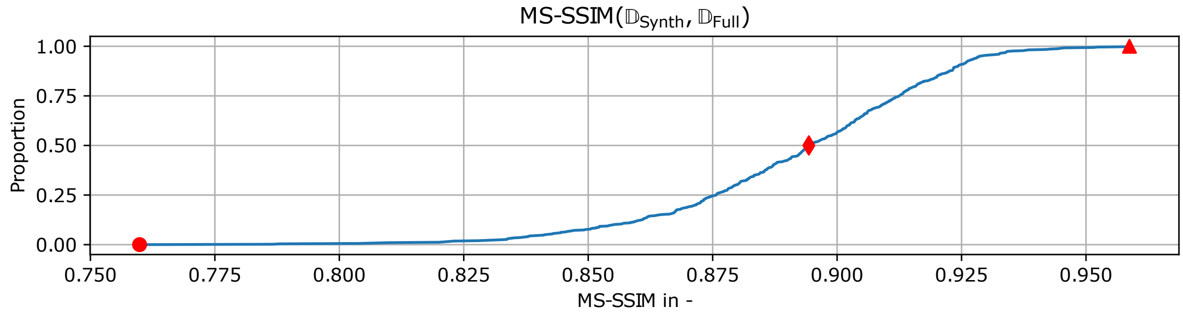

Figure 8-9: eCDF of the MS-SSIM Scores on the validation set. The red markers denote the lowest (circle), median (diamond) and highest MS-SSIM scores

While MS-SSIM values provide a relative reference for comparing two different image qualities, they are difficult to relate to absolute image quality. Considering the sample images in figure 8-10 for the lowest, median, and highest MS-SSIM scores of the dataset, a few observations can be made.

Examining the upper image pair with the lowest score, the effects of data reduction become apparent in vegetation and shadows. While the shadows on the street preserve their general structure and only appear to have been blurred, the vegetation exhibits stronger artifact patterns. Looking at the image pairs with median and highest MS-SSIM scores, it is evident that the partially synthetic image seems to be an appropriate substitute for the original image. At the same time, even the partially synthetic image with the highest score displays a strong artifact in the vegetation above the bus.

The subjective examination seems sufficient to conclude that the synthesis process is adequate for most of the image regions and inadequate for large areas of vegetation but lacks any robust criteria. In order to evaluate any substantiated acceptance criteria, the conduct of a study is necessary.

[250] Z. Wang et al.: Multiscale structural similarity for image quality assessment (2003).

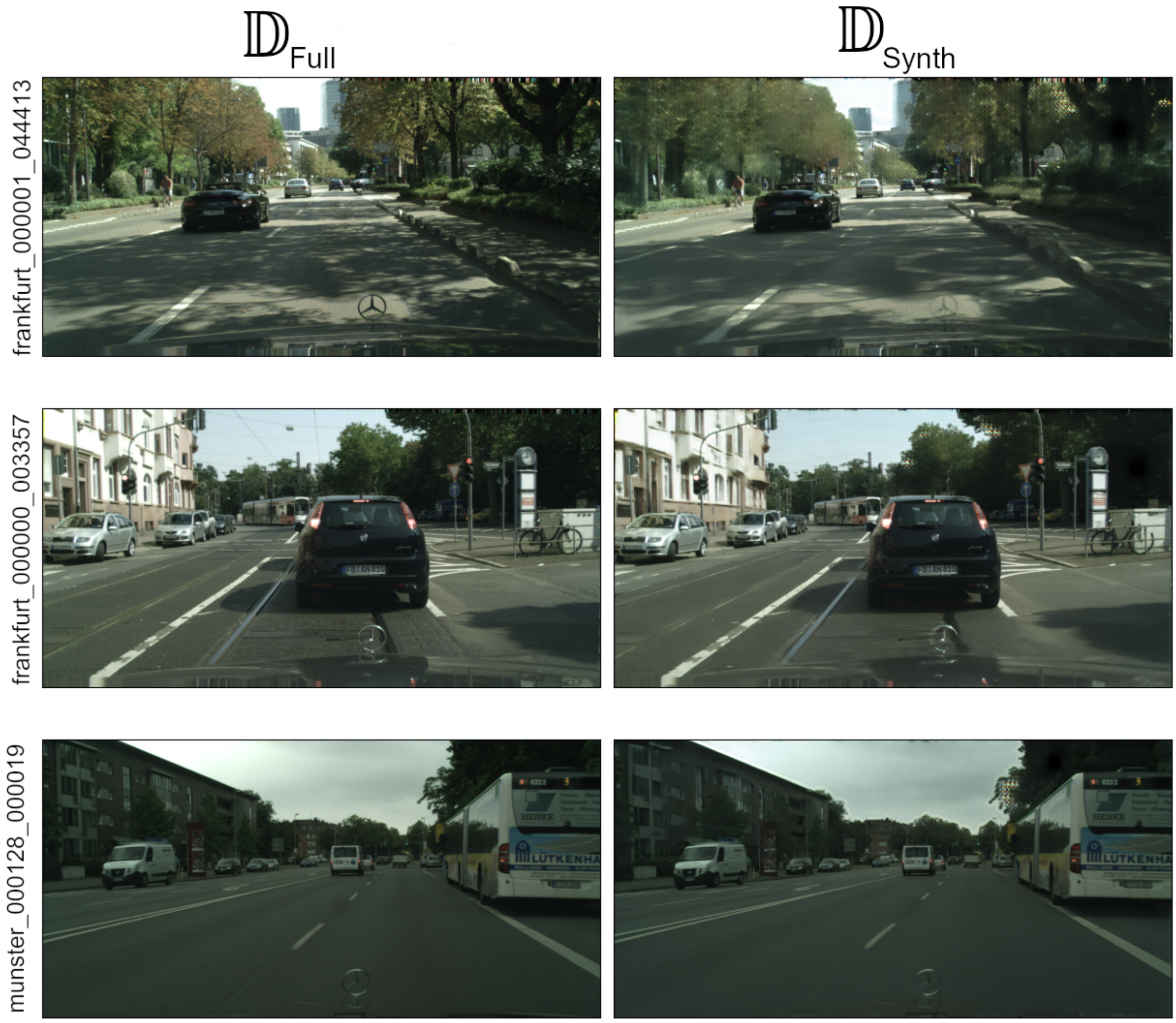

Figure 8-10: Samples for the lowest, median and highest MS-SSIM scores, as marked in figure 8-9.

The following delineates the process of conducting and evaluating a suitable study, shown in figure 8-11. The design of the study and evaluation is based on the ITU Recommendation P.912 ”Subjective video quality assessment methods for recognition tasks” as briefly presented in section 2.3.2. Initially, the study needs to be prepared. Here, the recognition task for the study is defined and data is provisioned. For this task, a dataset is created comprised of scenes from both the original dataset and the dataset after the subject data reduction has been applied. Each scene is present in both datasets. For each study participant, the scenes are then randomly allocated to equal parts to either the original dataset or the subject dataset. For experimental evaluation purposes, a 2D object detection task for traffic signs was chosen, and 21 scenes from the $\mathbb{D}_{\text{Full}}$ and $\hat{\mathbb{D}}_{\text{Synth}}$ dataset were selected. For demonstrative purposes, the study was performed with 10 participants. In actual application the number of participants has to be chosen with respect to the acceptance thresholds, to ensure sufficient statistical power for the hypothesis test.

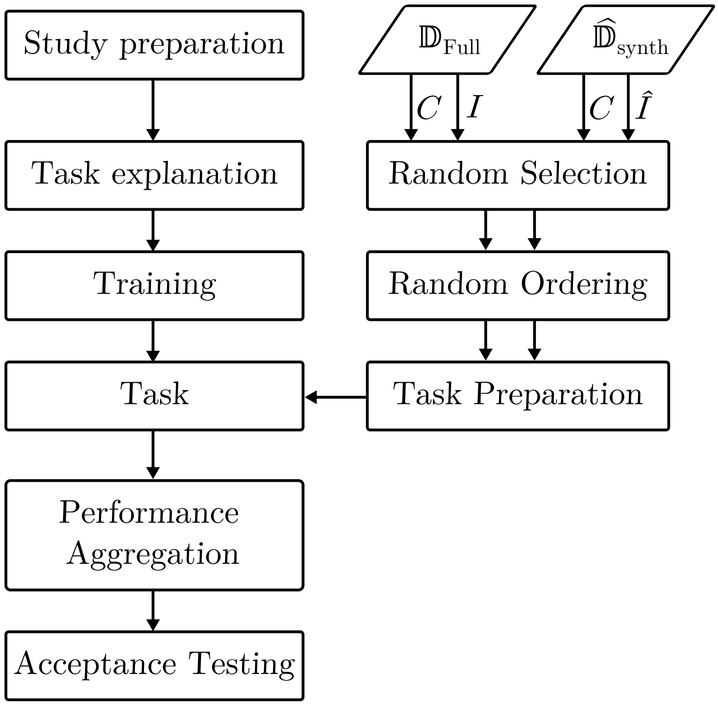

Figure 8-11: Flowchart, detailing the steps required to assert the effects of data reduction on human perception

To perform the study, each participant is given an explanation to give an understanding of the task. Following, separate scenes for training are presented as to accustom the participants to the usage. The explanation used in the experiment is shown in figure 8-12. Here, the target annotations are explained and a 60 s time limit is introduced to force the participants to annotate all signs in an equal time frame.

Because each participant may have a slightly different understanding of the task and a different tendency to perform than other participants, the results cannot be aggregated across participants. For this reason, each participants' performance is aggregated individually on the difference in performance between the scenes shown from the original dataset and the subject dataset. These performance aggregates provide the final population of samples for testing whether the impact on performance is acceptable. In the performed experiment, the participants are evaluated using the mAP$_{.5}$ measure. For each scene, a relative score is determined as the mAP$_{.5}$ value for the current scene in relation to the average mAP$_{.5}$ score of all scenes.

Due to the removal of information in the subject dataset the impact on the performance on the original data $p_{\mathbb{D}}$ is expected to be negative. The distributions of the populations of both performances are expected and assumed to be normal distributions. The performance on the subject data $p_{\hat{\mathbb{D}}}$ is modelled as the original performance worsened by both a systematic error ϵ_{sys} and a random error ϵ_{rand}. The former shifting the mean and the latter changing the variance of the normal distribution. Correcting for both errors yields the corrected distribution $p_{\hat{\mathbb{D}},corr}$ approximating the original distribution. This model is described by equations 8-1 to 8-3.

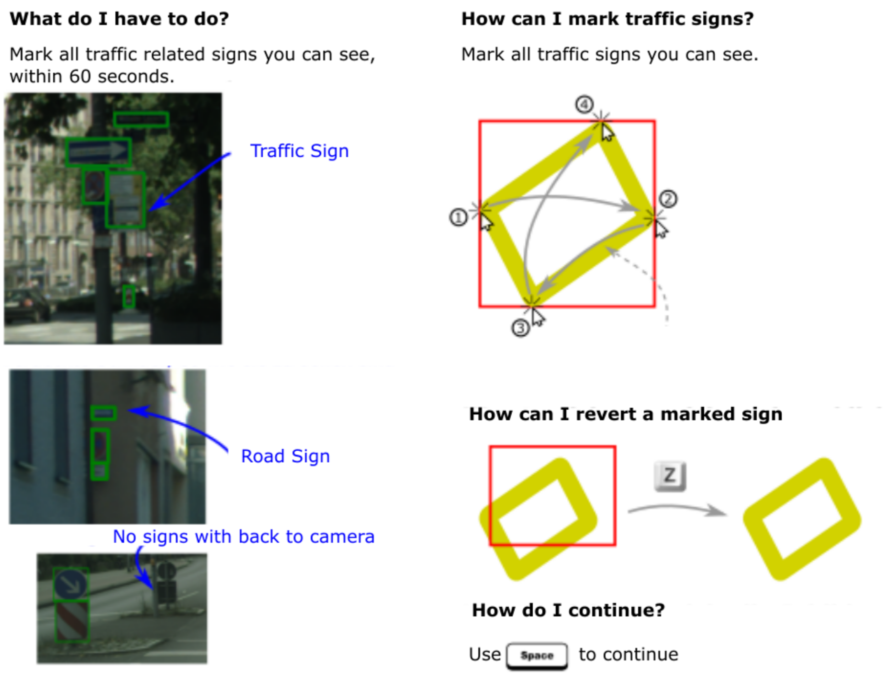

Figure 8-12: Tutorial screen as used in the experiments demonstrative study on the effects of data reduction on human perception

$$p_{\mathbb{D}} \approx \mathcal{N}(\mu_{\mathbb{D}}, \sigma_{\mathbb{D}}) \tag{8-1}$$

$$p_{\hat{\mathbb{D}}} \approx \mathcal{N}(\mu_{\hat{\mathbb{D}}}, \sigma_{\hat{\mathbb{D}}}) \tag{8-2}$$

$$p_{\hat{\mathbb{D}},corr} = \mathcal{N}(\mu_{\hat{\mathbb{D}}} - \epsilon_{\text{sys}}, \sigma_{\hat{\mathbb{D}}}/\epsilon_{\text{rand}}) \approx p_{\mathbb{D}} \tag{8-3}$$

Following, the bounds of the possible error values are determined. To this purpose a plausible range of error values is varied in a fully factorial manner. A statistical test is used to determine whether the hypotheses in equation 8-4 and 8-6 have to be rejected based on a confidence threshold. For the test of equal means, a one-sided T-test is applied. Welch's T-test is chosen instead of Student's T-test, since Welch's T-test does not require equal variances, which are assumed here to be unequal[251]. Conversely, to test for equal variances Levene's Test is applied[252]. In both cases, the implementation of the Python package Scipy is used.[253,254]

[251] Ruxton, G. D.: The unequal variance t-test is an underused alternative to Student's t-test and the Mann–Whitney U test (2006).

[252] Standarts, N. I. of; Technology: 1.3.5.10. Levene Test for Equality of Variances (2022).

[253] The SciPy community: scipy.stats.levene — SciPy v1.11.3 Manual (2023).

[254] The SciPy community: scipy.stats.ttest_ind — SciPy v1.11.3 Manual (2023).

$$H_{0,sys} : \mu_{\hat{\mathbb{D}},corr} = \mu_{\mathbb{D}} \tag{8-4}$$

$$H_{1,sys} : \mu_{\hat{\mathbb{D}},corr} \neq \mu_{\mathbb{D}} \tag{8-5}$$

$$H_{0,var} : \sigma^2_{\hat{\mathbb{D}},corr} = \sigma^2_{\mathbb{D}} \tag{8-6}$$

$$H_{1,var} : \sigma^2_{\hat{\mathbb{D}},corr} \neq \sigma^2_{\mathbb{D}} \tag{8-7}$$

Having performed the prior testing reveals the value sets for performance errors that cannot be ruled out. Performing this procedure for the experiment population, with a probability of error of 5%, reveals the possible value sets for the error model of the mAP$_{.5}$ score, as shown in figure 8-13. For these possible value sets for the error model, the null hypotheses cannot be rejected. Given acceptance thresholds for performance, the data reduction approach can now be accepted or discarded for human perception, based on whether there exists any unrejected null hypotheses for error value sets that violate the acceptance thresholds. If multiple independent performance measures are used, such as time spent on the task, the evaluation process can equally be executed. Multiple performance measures then yield multiple error models, for which individual acceptance criteria have to be applied.

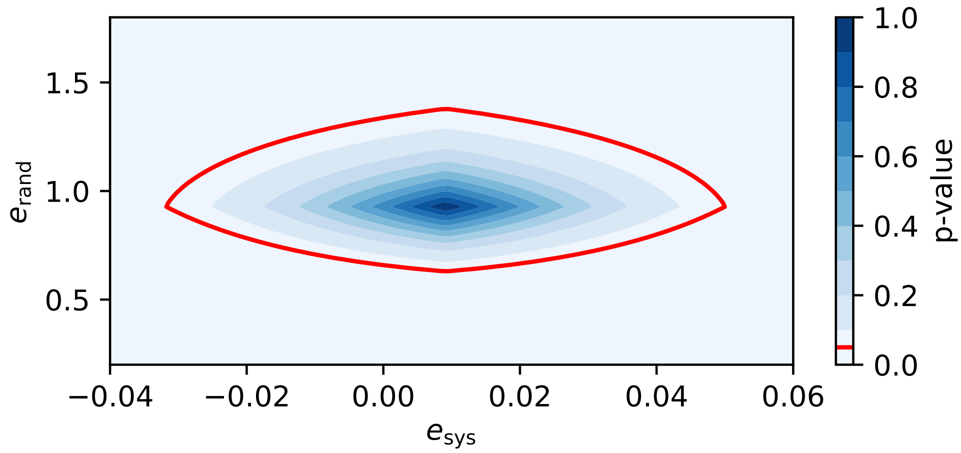

Figure 8-13: Bounds of the possible pairs of system and random error as based on the statistical tests within the human perception study. The red line marks the significance level $\alpha = 0.05$. The blue contour plot signifies the resultant p-value for the respective correction by ϵ_{sys}) and ϵ_{rand}).

8.2 Data Reduction Potential

The following section will evaluate the efficiency of the data as quantified by the data reduction factor. Reconsidering the observations from section 7.2 on the impact of the proposed SRM, it is evident that the data reduction effect is currently minimal. With only 0.1 % of pixels being deemed as irrelevant, the data reduction effect is equally or possibly even lower. As such, the SRM proposed in chapter 6 shows most of its effect in information reduction for dynamic objects.

While the proposed SRM represents a first step towards relevance-driven data reduction, the relevance model used in the CADR prototype is considered to be more akin to a mature relevance model in terms of the affected data parts. Observing the data reduction effect as produced by the SRM implemented in the CADR prototype, as described in chapter 7, enables an estimate for the possible potential for data reduction. Figure 8-14 shows a significant reduction in file size, when compared to the full dataset. Comparison with the file size of the blackened dataset gives insight into the size requirements of the additionally encoded data, as it is the differentiating factor between $\mathbb{D}_{\text{Synth}}$ and $\mathbb{D}_{\text{Blackened}}$. The additionally encoded data takes about 0.5 MB or on average roughly half of the storage amount. Overall the data reduction factor ranges between 0.2 and 0.8 with an average of just below 0.5. These numbers exemplify the sensitivity to the conditions in the subject scene. Thus, the decision to employ the data reduction approach can change based on the scene at hand.

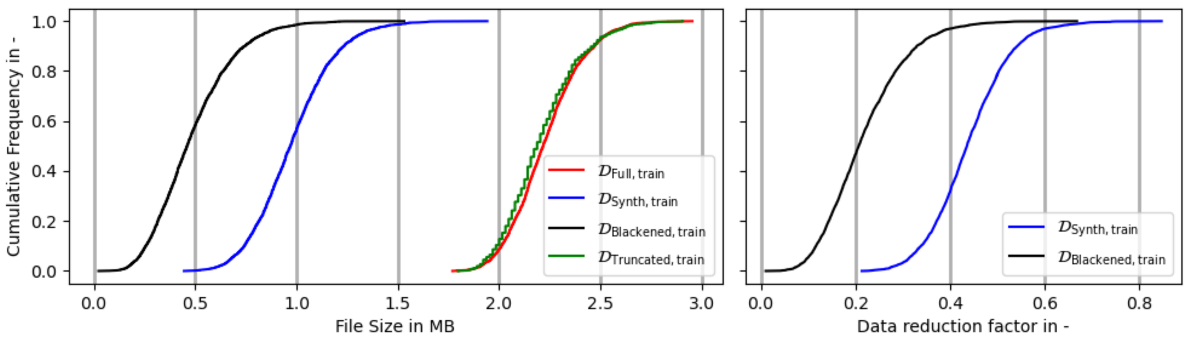

Figure 8-14: eCDF of the file size (left) and reduction factors (right) for the used datasets. The data reduction factor for $\mathbb{D}_{\text{Full}}$ and $\mathbb{D}_{\text{Truncated}}$, were omitted since they are 1 by definition

8.3 Interim Conclusion

In the previous sections, experiments were conducted and evaluated. The aim was to answer RQ. 3 and estimate the impact of a data oriented process when a relevance-driven data reduction is applied. The impact on quality is considered as loss of information and data reduction potential.

This impact is visualized in figure 8-15, showing the relations described in figure 2-5. Here, the loss of information is described by the ratio between performance on the unreduced data and performance on the reduced data. The data reduction potential is similarly expressed by the data reduction factor of the reduced data.

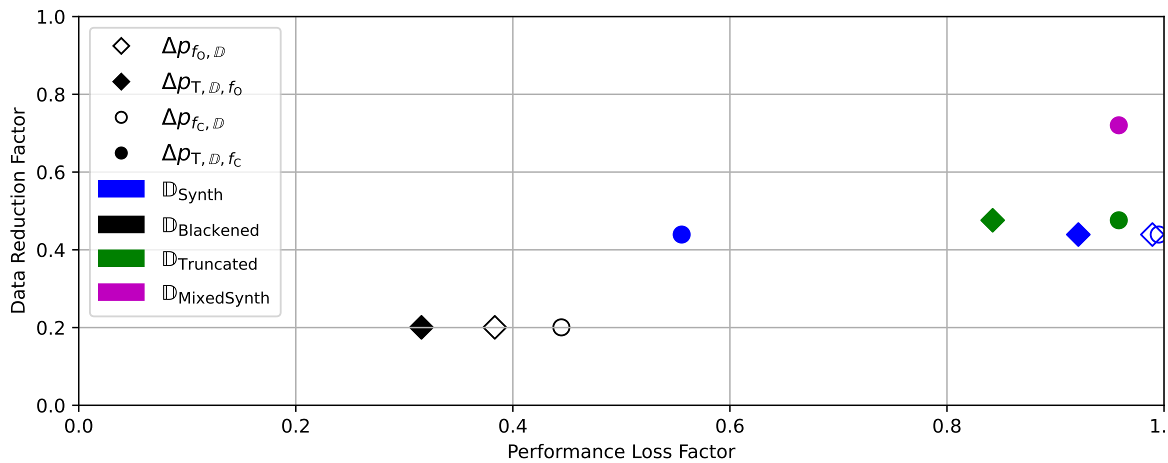

Figure 8-15: Data reduction factor compared to performance loss factor.
Symbols denote the different tasks of semantic segmentation (○) and object detection (◇). Inference symbolized by an unfilled symbol, while training has a filled symbol. The used dataset are disambiguated by color. Note that some datasets are only applicable to a subset of tasks.
Top right corner is the point of no change, while bottom right corner is the theoretical maximum.

Subsection 8.1.1 examined the impact when used for machine perception. The experiment conducted shows the separability between relevant and irrelevant information with relevance-driven data reduction. While overall some information is lost, most of the loss is isolated to the irrelevant information, with only a marginal amount affecting the relevant information. Here, a separability factor of 18 was observed. When compared to a baseline method for data reduction, the experiments conducted showed that this baseline method is not feasible, yielding a worsened performance loss by a factor of 100.

Subsection 8.1.2 explored the usability of reduced data for use as training data. The two experiments conducted exhibited different results indicating a strong dependence on the concrete task. While the training of an object detection showed feasibility compared to the baseline methods, the training of a semantic segmentation does not lead to this conclusion. The direct application of CADR yields worsened results, which can only be mitigated upon further modification. Since the modified results only approach the performance of the baseline method, the added effort of the CADR approach does not appear feasible in this case.

In the subsection 8.1.3, the impact of the proposed CADR approach on human perception was investigated. Literature and data evaluation revealed that established metrics such as MS-SSIM, while providing a quantitative measure of similarity, do not fully capture the subjective quality perceived by humans. The experimental data suggest that while some image regions are satisfactorily synthesized, complex areas such as vegetation are prone to noticeable artifacts. A methodological study is proposed for a nuanced analysis of the impact of data reduction on human task performance. The proposed study is demonstrated through an example experiment involving traffic sign detection annotation. By comparing participant performance on object detection tasks between original and synthetically altered datasets, and by applying statistical models to account for systematic and random errors, the study provides a scientifically sound approach to determining acceptable thresholds of image quality for human perception.

Section 8.2 showed that in order to deploy a relevance driven data reduction, it is first necessary to sufficiently mature the respective relevance model in order to yield an adequate data reduction effect. Early conservative relevance models lack the needed data reduction factor to justify the effort and accompanying negative effects of data reduction. More aggressive models on the other hand exhibit sufficient potential.

Concluding the two aims set at the beginning of the chapter have been achieved. First, methods to describe the impact of relevance driven data reductions have been presented. Second, the impact of the proposed CADR method was qualified. A quantified answer was given for the subject implementation. Depending on individual acceptance criteria the results for inference appear in a plausible range. The results for training show a wide spread in applicability depending on the subject task. In the experiment, usage of reduced data for training of an object detection task showed good performance, while the performance for training a semantic segmentation task was severely lacking. Following, it is concluded that it is necessary to evaluate the usability for a concrete training task before a data reduction is deployed.

9 Conclusion and Outlook

This chapter summarizes the conclusions of the present dissertation, highlighting key contributions and findings in relation to the research questions posed. It concludes with an outlook on potential future research directions.

9.1 Conclusion

This research addresses the emerging challenge of data handling in the development of ADS. The overarching goal is to advance the state of the art in data reduction by specifically considering the unique challenges of automated driving. A significant challenge identified is the release of ADS, which is hindered by the *open context* problem.

Research Question 1

"How can the open context present in automated driving be addressed in data reduction?"

Addressing the first research question, a novel approach to data reduction was presented, specifically tailored for the challenges of automated driving, called Context Aware Data Reduction (CADR). The findings indicate that an explicit consideration of relevance enables the identification of information and information needs to control the *open context*. A relevance-driven data reduction approach like CADR allows for effective management of performance loss within the data's information content.

These findings were facilitated by considering two further research questions, concerning the key functions of the proposed CADR approach.

Research Question 2

"How can the concept of relevance, in the context of a given use case, be formally described?"

The second research question led to significant contributions. Based on the state of the art in relevance theory and knowledge representation, an ontological representation of the concept of relevance in the domain of automated driving was derived. This representation includes an abstract template for relevance and a method for deriving and validating a concrete relevance model. As a demonstration, this method was applied to the sample use case of a collision-free dynamic driving task (DDT) in an urban setting. The findings show that ontologies are suitable for representing knowledge about the operating design domain (ODD) and relevance. However, the current state of the art does not provide a standardized consensus on a unified relevance concept. The proposed method successfully derives a valid relevance model within a given use case, but it also highlights that current datasets might not be extensive enough in their field of view to include all relevant objects.

Research Question 3

"What is the impact on the use case of relevance-driven data reduction?"

The third research question is focused on the contributions and findings related to the impact of this approach. Methods to estimate the impact on the use case were presented, covering both, machine and human perception. In the case of perception tasks like semantic segmentation and object detection, a minimal impact on performance was observed when focusing on relevant information. In contrast, other reference approaches to data reduction showed a significantly larger impact on performance. The impact on training neural networks was more ambiguous. Training for object detection exhibited a low impact, while training for semantic segmentation showed a larger impact.

In conclusion, a novel approach to data reduction, specifically designed for the challenges of automated driving, was presented and its applicability was demonstrated. The exact application and effectiveness of this approach are highly dependent on the specific use case and context.

9.2 Outlook

Considering future research directions, several aspects emerge that can be grouped into three main areas. The first aspect focuses on the abstract concept of relevance-driven data reduction. The second aspect delves into the concrete application of an implemented data reduction strategy. The third aspect addresses the general relevance modeling within the use cases of AD.

First Aspect: Relevance-Driven Data Reduction

Within this aspect, additional research and development efforts are required to enhance the concept of relevance-driven data reduction. A key area of focus is the architecture designed to be agnostic to sensor modalities, supporting multiple fused sensor representations. There is a need for further research into the specific implementation details of how this can be realized. Also, as identified in the context of research question three, a strong use case dependence was observed when using reduced data as the basis for a training process. Here, further research is needed to understand the root causes of this dependence. Additionally, the development of methods to determine the suitability of the use case, as to a priori determine whether the data reduction approach can be applied to the use case without significant performance loss.

Second Aspect: Application

Due to the lack of matured infrastructure and actual realized use cases of automated driving, the data reduction approach was only evaluated in a simplified academic environment for experiments. As such, it is proposed that significant future research direction should involve the details of applying data reduction in complex industrial processes. As broadly described in section 2.2, the positive and negative effects of data reduction are more complex than can be addressed within the scope of this dissertation. A characterization of possible data flows and data provisioning, along with an applicability study for data reduction, are required before the proposed approach can be utilized in any industrial environment.

Third Aspect: Relevance Modeling

This work represents only an initial exploration into the world of relevance theory from the automotive domain. There is a further need to refine, modify, and expand the concepts presented in this dissertation. As the nomenclature of relevances within the field of relevance theory is not yet unified, this work refrained from proposing its own novel nomenclature for relevance in the domain of automated driving. This omission is not a statement of the lack of importance of such a standardized nomenclature. On the contrary, it underscores that an exact taxonomy of defined relevance is too important to be encumbered with a premature definition. Future efforts should aim to contribute to a unified and standardized nomenclature that can effectively support the evolving needs of relevance theory in the context of automated driving.

Finally, the author hopes that this work will stimulate an explicit consideration of relevance and enable further development to overcome data-centric challenges in the field of automated driving.

Bibliography

Ackoff, R. L.: From Data to Wisdom (1989)
Ackoff, R. L.: From Data to Wisdom, in: Journal of applied systems analysis, pp. 3–9, 1989

ad Datasets: Autonomous Driving related Collection of Datasets (28.07.2023)
ad Datasets: Autonomous Driving related Collection of Datasets, URL: https://ad-datasets.com/, 28.07.2023

Ahsan, S. et al.: Data, Information, Knowledge, Wisdom: A Doubly Linked Chain? (2006)
Ahsan, Syed; Shah, Abad: Data, Information, Knowledge, Wisdom: A Doubly Linked Chain?, in: International Conference on Information and Knowledge Engineering. 2006

Ali, G. et al.: Quantifying the effect on the frequency of longitudinal and lateral accelerations (2021)
Ali, Gibran; McLaughlin, Shane; Ahmadian, Mehdi: Quantifying the effect of roadway, driver, vehicle, and location characteristics on the frequency of longitudinal and lateral accelerations, in: Accident Analysis & Prevention, Vol. 161, p. 106356, 2021

Allaby, M.; Park, C.: A dictionary of environment and conservation (2013)
Allaby, Michael; Park, Chris: A dictionary of environment and conservation, OUP Oxford, 2013

Althoff, M. et al.: Set-Based Prediction of Traffic Participants on Arbitrary Road Networks (2016)
Althoff, Matthias; Magdici, Silvia: Set-Based Prediction of Traffic Participants on Arbitrary Road Networks, in: IEEE Transactions on Intelligent Vehicles, Vol. 1, pp. 187–202, 2016

Amer, Hossam: Image/Video Compression: Human and Computer Vision Perspectives (2020)
Amer, Hossam: Image/Video Compression: Human and Computer Vision Perspectives, 2020

Amersbach, C.: Functional Decomposition Approach (2019)
Amersbach, Christian: Functional Decomposition Approach: Reducing the Safety Validation Effort for Highly Automated Driving, Dissertation, Technische Universität Darmstadt, 2019

Anastasia Natsiou, Sean O'Leary, and Luca Longo: An Exploration of the Latent Space of a CVAE (2023)
Anastasia Natsiou, Sean O'Leary, and Luca Longo: An Exploration of the Latent Space of a Convolutional Variational Autoencoder for the Generation of Musical Instrument Tones, 2023

Anderson, T. W.: On the Distribution of the Two-Sample Cramer-von Mises Criterion (1962)
Anderson, T. W.: On the Distribution of the Two-Sample Cramer-von Mises Criterion, in: The Annals of Mathematical Statistics, Vol. 33, pp. 1148–1159, 1962

Arp, R.; Smith, B.; Spear, A. D.: Building ontologies with basic formal ontology (2016)
Arp, Robert; Smith, Barry; Spear, Andrew D.: Building ontologies with basic formal ontology, The MIT Press, 2016

Autey, J. et al.: Safety evaluation of right-turn smart channels using automated traffic conflict analysis. (2012)
Autey, Jarvis; Sayed, Tarek; Zaki, Mohamed H.: Safety evaluation of right-turn smart channels using automated traffic conflict analysis. In: Accident; analysis and prevention, Vol. 45, pp. 120–30, 2012

Bach, J. et al.: Data-driven development, a complementing approach for automotive systems engineering (2017)
Bach, Johannes; Langner, Jacob; Otten, Stefan; Holzäpfel, Marc; Sax, Eric: Data-driven development, a complementing approach for automotive systems engineering, in: 2017 IEEE International Systems Engineering Symposium (ISSE), pp. 1–6, 2017

Bagschik, G. et al.: Ontology based Scene Creation for the Development of Automated Vehicles (2018)
Bagschik, Gerrit; Menzel, Till; Maurer, Markus: Ontology based Scene Creation for the Development of Automated Vehicles, in: 2018 IEEE Intelligent Vehicles Symposium (IV), pp. 1813–1820, 2018

Bakurov, I. et al.: Structural similarity index (SSIM) revisited: A data-driven approach (2022)
Bakurov, Illya; Buzzelli, Marco; Schettini, Raimondo; Castelli, Mauro; Vanneschi, Leonardo: Structural similarity index (SSIM) revisited: A data-driven approach, in: Expert Systems with Applications, Vol. 189, p. 116087, 2022

Ballé, J. et al.: Variational image compression with a scale hyperprior (2018)
Ballé, Johannes; Minnen, David; Singh, Saurabh; Hwang, Sung Jin; Johnston, Nick: Variational image compression with a scale hyperprior, URL: http://arxiv.org/pdf/1802.01436v2, 2018

Bangarevva, P. et al.: In-Service Monitoring and Assessment of AD vehicles with AI based Algorithms (2023)
Bangarevva, Patil; Corell, Thomas; Hota, Rudra: In-Service Monitoring and Assessment of AD vehicles with AI based Algorithms, in: VALID 2023: The Fifteenth International Conference on Advances in System Testing and Validation Lifecycle, 2023

Barina, D.: Comparison of Lossless Image Formats (2021)
Barina, David: Comparison of Lossless Image Formats, URL: http://arxiv.org/pdf/2108.02557v1, 2021

Bertolazzi, E. et al.: Supporting Drivers in Keeping Safe Speed and Safe Distance (2010)
Bertolazzi, Enrico; Biral, Francesco; Da Lio, Mauro; Saroldi, Andrea; Tango, Fabio: Supporting Drivers in Keeping Safe Speed and Safe Distance: The SASPENCE Subproject Within the European Framework Programme 6 Integrating Project PReVENT, in: IEEE Transactions on Intelligent Transportation Systems, Vol. 11, pp. 525–538, 2010

Bogdoll, D. et al.: Ad-datasets: a meta-collection of data sets for autonomous driving (2022)
Bogdoll, Daniel; Schreyer, Felix; Zöllner, J Marius: Ad-datasets: a meta-collection of data sets for autonomous driving, in: arXiv preprint arXiv:2202.01909, 2022

Bokare, P. S. et al.: Acceleration-Deceleration Behaviour of Various Vehicle Types (2017)
Bokare, P. S.; Maurya, A. K.: Acceleration-Deceleration Behaviour of Various Vehicle Types, in: Transportation Research Procedia, Vol. 25, pp. 4733–4749, 2017

Buchholz, M. et al.: Handling occlusions in automated driving using a multiaccess edge computing (2021)
Buchholz, Michael; Müller, Johannes; Herrmann, Martin; Strohbeck, Jan; Völz, Benjamin; Maier, Matthias; Paczia, Jonas; Stein, Oliver; Rehborn, Hubert; Henn, Rüdiger-Walter: Handling occlusions in automated driving using a multiaccess edge computing server-based environment model from infrastructure sensors, in: IEEE Intelligent Transportation Systems Magazine, Vol. 14, pp. 106–120, 2021

Buechel, M. et al.: Ontology-based traffic scene modeling, and decision-making for automated vehicles (2017)
Buechel, Martin; Hinz, Gereon; Ruehl, Frederik; Schroth, Hans; Gyoeri, Csaba; Knoll, Alois: Ontology-based traffic scene modeling, traffic regulations dependent situational awareness and decision-making for automated vehicles, in: 2017 IEEE Intelligent Vehicles Symposium (IV), pp. 1471–1476, 2017

Bundesministerium für Justiz und Verbraucherschutz: Straßenverkehrs-Ordnung (2013)
Bundesministerium für Justiz und Verbraucherschutz: Straßenverkehrs-Ordnung, URL: https://www.gesetze-im-internet.de/stvo_2013/StVO.pdf, 2013

Bundesministeriums für Digitales und Verkehr: Autonome-Fahrzeuge-Genehmigungs- und-Betriebs-Verordnung – AFGBV) (86/22)
Bundesministeriums für Digitales und Verkehr: Verordnung zur Genehmigung und zum Betrieb von Kraftfahrzeugen mit autonomer Fahrfunktion in festgelegten Betriebsbereichen: Autonome-Fahrzeuge-Genehmigungs- und-Betriebs-Verordnung – AFGBV), 86/22

Caesar, H. et al.: nuScenes (2020)
Caesar, Holger; Bankiti, Varun; Lang, Alex H.; Vora, Sourabh; Liong, Venice Erin; Xu, Qiang; Krishnan, Anush; Pan, Yu; Baldan, Giancarlo; Beijbom, Oscar: nuScenes: A multimodal dataset for autonomous driving, in: 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020

Carbonell, J. et al.: The Use of MMR, Diversity-Based Reranking for Reordering Documents (1998)
Carbonell, Jaime; Goldstein, Jade: The Use of MMR, Diversity-Based Reranking for Reordering Documents and Producing Summaries, in: Proceedings of the 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 335–336, 1998

Chamain, L. D. et al.: End-to-End optimized image compression for machines, a study (2021)
Chamain, Lahiru D.; Racape, Fabien; Begaint, Jean; Pushparaja, Akshay; Feltman, Simon: End-to-End optimized image compression for machines, a study, in: 2021 Data Compression Conference (DCC), pp. 163–172, 2021

Choo, C. W.: The Knowing Organization (1996)
Choo, C. W.: The Knowing Organization: How Organizations Use Information to Construct Meaning, Create Knowledge and Make Decisions, in: International Journal of Information Management, Vol. 16, pp. 329–340, 1996

Cityscapes Dataset: Benchmark Suite (2023)
Cityscapes Dataset: Benchmark Suite, URL: https://www.cityscapes-dataset.com/benchmarks/#scene-labeling-task, 2023

Consortium, P.: PEGASUS METHOD (2019)
Consortium, PEGASUS: PEGASUS METHOD, 2019

Cooper, W. S.: On selecting a measure of retrieval effectiveness (1973)
Cooper, William S.: On selecting a measure of retrieval effectiveness, in: J. Am. Soc. Inf. Sci. Vol. 24, pp. 87–100, 1973

Cooper, W. S.: On selecting a measure of retrieval effectiveness part II. Implementation of the philosophy (1973)
Cooper, William S.: On selecting a measure of retrieval effectiveness part II. Implementation of the philosophy, in: J. Am. Soc. Inf. Sci. Vol. 24, pp. 413–424, 1973

Cordts, M. et al.: The Cityscapes Dataset for Semantic Urban Scene Understanding (2016)
Cordts, Marius; Omran, Mohamed; Ramos, Sebastian; Rehfeld, Timo; Enzweiler, Markus; Benenson, Rodrigo; Franke, Uwe; Roth, Stefan; Schiele, Bernt: The Cityscapes Dataset for Semantic Urban Scene Understanding, in: Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016

Cosijn, E. et al.: Dimensions of relevance (2000)
Cosijn, Erica; Ingwersen, Peter: Dimensions of relevance, in: Information Processing & Management, Vol. 36, pp. 533–550, 2000

Cui, G. et al.: Cooperative perception technology of autonomous driving (2022)
Cui, Guangzhen; Zhang, Weili; Xiao, Yanqiu; Yao, Lei; Fang, Zhanpeng: Cooperative perception technology of autonomous driving in the internet of vehicles environment: A review, in: Sensors, Vol. 22, p. 5535, 2022

Cyganski, R.: Automated Vehicles and Automated Driving from a Demand Modeling Perspective (2016)
Cyganski, Rita: Automated Vehicles and Automated Driving from a Demand Modeling Perspective, in: Gerdes, J. Christian; Lenz, Barbara; Maurer, Markus; Winner, Hermann (Hrsg.): Autonomous Driving, Springer Berlin Heidelberg and Imprint: Springer, 2016

Daniele Mari et al.: Content-Aware Compression and Transmission Using Semantics (2023)
Daniele Mari; Elena Camuffo; Simone Milani: CACTUS: Content-Aware Compression and Transmission Using Semantics for Automotive LiDAR Data, 2023

David Reinsel et al.: The Digitization of the World: From Edge to Core (2018)
David Reinsel; John Gantz; John Rydning: The Digitization of the World: From Edge to Core, 2018

Dervos, D. A. et al.: A Common Sense Approach to Defining Data, Information, and Metadata (2006)
Dervos, Dimitris A.; Coleman, Anita: A Common Sense Approach to Defining Data, Information, and Metadata, in: Proceedings of the Ninth Interantional Society for Knowledge Organization 2006 Conference, Vol. 9, 2006

Di Chen et al.: Pixel-Level Texture Segmentation Based AV1 Video Compression (2018)
Di Chen; Chen, Qingshuang; Zhu, Fengqing: Pixel-Level Texture Segmentation Based AV1 Video Compression, in: (Hrsg.): 2018 IEEE International Conference on Acoustics, Speech, and Signal Processing, IEEE, 2018

Di Feng et al.: Deep Multi-modal Object Detection and Semantic Segmentation (2021)
Di Feng; Haase-Schütz, Christian; Rosenbaum, Lars; Hertlein, Heinz; Glaeser, Claudius; Timm, Fabian; Wiesbeck, Werner; Dietmayer, Klaus: Deep Multi-modal Object Detection and Semantic Segmentation for Autonomous Driving: Datasets, Methods, and Challenges, in: IEEE Transactions on Intelligent Transportation Systems, Vol. 22, pp. 1341–1360, 2021

DSpace: Data Driven Software Development (11.11.2023)
DSpace: Data Driven Software Development, URL: https://www.dspace.com/en/pub/home/applicationfields/comp/data_driven_development.cfm, 11.11.2023

Ehrlich, M. et al.: Analyzing and Mitigating JPEG Compression Defects in Deep Learning (2021)
Ehrlich, Max; Davis, Larry; Lim, Ser-Nam; Shrivastava, Abhinav: Analyzing and Mitigating JPEG Compression Defects in Deep Learning, in: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) Workshops, pp. 2357–2367, 2021

EICT GmbH: IMAGinE: Intelligent Maneuver Automation (2023)
EICT GmbH: IMAGinE: Intelligent Maneuver Automation – cooperative hazard avoidance in realtime, URL: https://www.imagine-online.de/en/home.html, 2023

Eisenberg, M. B.: Measuring relevance judgments (1988)
Eisenberg, Michael B.: Measuring relevance judgments, in: Information Processing & Management, Vol. 24, pp. 373–389, 1988

Feiler, J. et al.: The Perception Modification Concept to Free the Path (2021)
Feiler, Johannes; Diermeyer, Frank: The Perception Modification Concept to Free the Path of An Automated Vehicle Remotely, in: 7th International Conference on Vehicle Technology and Intelligent Transport Systems, pp. 405–412, 2021

Fremont, D. J. et al.: Formal Scenario-Based Testing of Autonomous Vehicles (2020)
Fremont, Daniel J.; Kim, Edward; Pant, Yash Vardhan; Seshia, Sanjit A.; Acharya, Atul; Bruso, Xantha; Wells, Paul; Lemke, Steve; Lu, Qiang; Mehta, Shalin: Formal Scenario-Based Testing of Autonomous Vehicles: From Simulation to the Real World, in: 2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC), pp. 1–8, 2020

French, R.: Catastrophic forgetting in connectionist networks (1999)
French, Robert: Catastrophic forgetting in connectionist networks, in: Trends in Cognitive Sciences, Vol. 3, 1999

Fu, C. et al.: Texture Segmentation Based Video Compression Using Convolutional Neural Networks (2018)
Fu, Chichen; Di Chen; Delp, Edward J.; Liu, Zoe; Zhu, Fengqing: Texture Segmentation Based Video Compression Using Convolutional Neural Networks, URL: http://arxiv.org/pdf/1802.02992v1, 2018

Furman, Vadim; Chatham, Andrew Hughes; Ogale, Abhijit; Dolgov, Dmitri. “Image and Video Compresssion for Remote Vehicle Assistance”. US 2016/0283804 A1. 2016.

Gansch, R. et al.: System Theoretic View on Uncertainties (2020)
Gansch, Roman; Adee, Ahmad: System Theoretic View on Uncertainties, in: 2020 Design, Automation & Test in Europe Conference & Exhibition (DATE), pp. 1345–1350, 2020

Geiger, A. et al.: Are we ready for Autonomous Driving? (2012)
Geiger, Andreas; Lenz, Philip; Urtasun, Raquel: Are we ready for Autonomous Driving? The KITTI Vision Benchmark Suite, in: Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3354–3361, 2012

github.com/XuJiacong/PIDNet: This is the official repository for our recent work: PIDNet (2023)
github.com/XuJiacong/PIDNet: This is the official repository for our recent work: PIDNet, URL: https://github.com/XuJiacong/PIDNet, 2023

Google LLC: Machine Learning Common Problems (2023)
Google LLC: Machine Learning Common Problems, 2023

Green, M.: How Long Does It Take to Stop? Methodological Analysis of Driver Perception-Brake Times (2000)
Green, Marc: How Long Does It Take to Stop? Methodological Analysis of Driver Perception-Brake Times, in: Transportation Human Factors, Vol. 2, pp. 195–216, 2000

Group, P. C. I. S. I.: PCI-SIG® Announces PCI Express® 7.0 Specification to Reach 128 GT/s (2022)
Group, Peripheral Component Interconnect Special Interest: PCI-SIG® Announces PCI Express® 7.0 Specification to Reach 128 GT/s, URL: https://www.businesswire.com/news/home/20220621005137/en, 2022

Gruber, T. R.: A translation approach to portable ontology specifications (1993)
Gruber, Thomas R.: A translation approach to portable ontology specifications, in: Knowledge Acquisition, Vol. 5, pp. 199–220, 1993

Guarino, Nicola; Oberle, Daniel; Staab, Steffen. "What Is an Ontology?" In: *Handbook on Ontologies*. Ed. by Steffen Staab; Rudi Studer. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 1–17. ISBN: 978-3-540-92673-3. DOI: 10.1007/978-3-540-92673-3_0. URL: https://doi.org/10.1007/978-3-540-92673-3_0.

Hayward, J. C.: NEAR-MISS DETERMINATION THROUGH USE OF A SCALE OF DANGER (1972)
Hayward, John C.: NEAR-MISS DETERMINATION THROUGH USE OF A SCALE OF DANGER, in: Highway Research Record, 1972

Hestness, J. et al.: Deep Learning Scaling is Predictable, Empirically (2017)
Hestness, Joel; Narang, Sharan; Ardalani, Newsha; Diamos, Gregory; Jun, Heewoo; Kianinejad, Hassan; Patwary, Md Mostofa Ali; Yang, Yang; Zhou, Yanqi: Deep Learning Scaling is Predictable, Empirically, URL: http://arxiv.org/pdf/1712.00409v1, 2017

Hofbauer, M. et al.: Traffic-Aware Multi-View Video Stream Adaptation for Teleoperated Driving (2022)
Hofbauer, Markus; Kuhn, Christopher B.; Khlifi, Mariem; Petrovic, Goran; Steinbach, Eckehard: Traffic-Aware Multi-View Video Stream Adaptation for Teleoperated Driving, in: 2022 IEEE 95th Vehicular Technology Conference: (VTC2022-Spring), pp. 1–7, 2022

Hofbauer, M. et al.: TELECARLA (2020)
Hofbauer, Markus; Kuhn, Christopher B.; Petrovic, Goran; Steinbach, Eckehard: TELECARLA: An Open Source Extension of the CARLA Simulator for Teleoperated Driving Research Using Off-the-Shelf Components, in: 2020 IEEE Intelligent Vehicles Symposium (IV), pp. 335–340, 2020

Hofbauer, M. et al.: Adaptive Live Video Streaming for Teleoperated Driving (2022)
Hofbauer, Markus; Sc, M.: Adaptive Live Video Streaming for Teleoperated Driving, Dissertation, Technischen Universität München e, 2022

Horn Martin; Watzenig, D.: Automated Driving: Safer and More Efficient Future Driving (2017)
Horn Martin; Watzenig, Daniel: Automated Driving: Safer and More Efficient Future Driving, 1st ed. 2017. Edition, Springer International Publishing and Imprint: Springer, 2017

Hoss, M. et al.: A Review of Testing Object-Based Environment Perception for Safe Automated Driving (2022)
Hoss, Michael; Scholtes, Maike; Eckstein, Lutz: A Review of Testing Object-Based Environment Perception for Safe Automated Driving, in: Automotive Innovation, Vol. 5, pp. 223–250, 2022

Houben, S. et al.: Inspect, Understand, Overcome: A Survey of Practical Methods for AI Safety (2022)
Houben, Sebastian; Abrecht, Stephanie; Akila, Maram; Bär, Andreas; Brockherde, Felix; Feifel, Patrick; Fingscheidt, Tim; Gannamaneni, Sujan Sai; Ghobadi, Seyed Eghbal; Hammam, Ahmed; Haselhoff, Anselm; Hauser, Felix; Heinzemann, Christian; Hoffmann, Marco; Kapoor, Nikhil; Kappel, Falk; Klingner, Marvin; Kronenberger, Jan; Küppers, Fabian; Löhdefink, Jonas; Mlynarski, Michael; Mock, Michael; Mualla, Firas; Pavlitskaya, Svetlana; Poretschkin, Maximilian; Pohl, Alexander; Ravi-Kumar, Varun; Rosenzweig, Julia; Rottmann, Matthias; Rüping, Stefan; Sämann, Timo; Schneider, Jan David; Schulz, Elena; Schwalbe, Gesina; Sicking, Joachim; Srivastava, Toshika; Varghese, Serin; Weber, Michael; Wirkert, Sebastian; Wirtz, Tim; Woehrle, Matthias: Inspect, Understand, Overcome: A Survey of Practical Methods for AI Safety, Vol. 35, pp. 3–78, 2022

Huber, B. et al.: Evaluation of Virtual Traffic Situations based on Multidimensional Criticality Analysis (2020)
Huber, Bernd; Herzog, Steffen; Sippl, Christoph; German, Reinhard; Djanatliev, Anatoli: Evaluation of Virtual Traffic Situations for Testing Automated Driving Functions based on Multidimensional Criticality Analysis, in: 2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC), pp. 1–7, 2020

Hülsen, M. et al.: Traffic intersection situation description ontology for advanced driver assistance (2011)
Hülsen, Michael; Zöllner, J. Marius; Weiss, Christian: Traffic intersection situation description ontology for advanced driver assistance, in: 2011 IEEE Intelligent Vehicles Symposium (IV), pp. 993–999, 2011

IEEE: P802.3bs - Physical Layers and Management Parameters for 200 Gb/s and 400 Gb/s Operation (2017)
IEEE: IEEE Draft Standard for Ethernet Amendment 10: Media Access Control Parameters, Physical Layers and Management Parameters for 200 Gb/s and 400 Gb/s Operation, in: IEEE P802.3bs(TM)/D3.23.3, July 2017, pp. 1–393, 2017

IEEE: Std 802.3bt-2018 - 10 Mb/s Operation and Power Delivery over a Single Pair of Conductors (2020)
IEEE: IEEE Standard for Ethernet - Amendment 5: Physical Layer Specifications and Management Parameters for 10 Mb/s Operation and Associated Power Delivery over a Single Balanced Pair of Conductors, in: IEEE Std 802.3cg-2019 (Amendment to IEEE Std 802.3-2018 as amended by IEEE Std 802.3cb-2018, IEEE Std 802.3bt-2018, IEEE Std 802.3cd-2018, and IEEE Std 802.3cn-2019), pp. 1–256, 2020

IEEE: Std 802.3bw-2015 - 100 Mb/s Operation over a Single Balanced Twisted Pair Cable (2016)
IEEE: IEEE Standard for Ethernet Amendment 1: Physical Layer Specifications and Management Parameters for 100 Mb/s Operation over a Single Balanced Twisted Pair Cable (100BASE-T1), in: IEEE Std 802.3bw-2015 (Amendment to IEEE Std 802.3-2015, pp. 1–88, 2016

IEEE: Std 802.3bp-2016 - 1 Gb/s Operation over a Single Twisted-Pair Copper Cable (2016)
IEEE: IEEE Standard for Ethernet Amendment 4: Physical Layer Specifications and Management Parameters for 1 Gb/s Operation over a Single Twisted-Pair Copper Cable, in: IEEE Std 802.3bp-2016 (Amendment to IEEE Std 802.3-2015 as amended by IEEE Std 802.3bw-2015, IEEE Std 802.3by-2016, and IEEE Std 802.3bq-2016), pp. 1–211, 2016

International Standardization Organization: Visual volumetric coding and point cloud compression (2021)
International Standardization Organization: Coded representation of immersive media: Part 5: Visual volumetric video-based coding (V3C) and video-based point cloud compression (V-PCC), URL: https://www.iso.org/standard/73025.html, 2021

International Standardization Organization: Geometry-based point cloud compression (2023)
International Standardization Organization: Coded representation of immersive media: Part 9: Geometry-based point cloud compression, URL: https://www.iso.org/standard/73025.html, 2023

International Standardization Organization: Information technology: Digital compression and coding of continuous-tone still images: JPEG File Interchange Format (JFIF) (2013)
International Standardization Organization: Information technology: Digital compression and coding of continuous-tone still images: JPEG File Interchange Format (JFIF), 2013

International Standardization Organization: PNG: Functional specification (2023)
International Standardization Organization: Portable Network Graphics (PNG): Functional specification, URL: https://www.iso.org/standard/29581.html, 2023

International Standardization Organization: Road vehicles - Functional safety (2018)
International Standardization Organization: Road vehicles - Functional safety, 2018

International Standardization Organization: Road vehicles - Functional safety - Part 2 (2018)
International Standardization Organization: Road vehicles - Functional safety - Part 2: Management of functional safety, 2018

International Telecommunication Union: H.264: Advanced video coding for generic audiovisual services (2019)
International Telecommunication Union: H.264: Advanced video coding for generic audiovisual services, 2019

International Telecommunication Union: ITU-T Rec. H.265 (08/2021) High efficiency video coding (2021)
International Telecommunication Union: ITU-T Rec. H.265 (08/2021) High efficiency video coding, 2021

ISO: Road vehicles — Safety and cybersecurity for automated driving systems (2020)
ISO: Road vehicles — Safety and cybersecurity for automated driving systems — Design, verification and validation, 2020

ISO: Systems and software engineering - Vocabulary (2017)
ISO: Systems and software engineering - Vocabulary, 2017

ISO/DIS: Road vehicles - Safety of the intended functionality (2021)
ISO/DIS: Road vehicles - Safety of the intended functionality, 2021

ITU-T: ITU-T Rec. P.800.1 (07/2016) Mean opinion score (MOS) terminology (2016)
ITU-T: ITU-T Rec. P.800.1 (07/2016) Mean opinion score (MOS) terminology, 2016

ITU-T: ITU-T Rec. P.912 (03/2016) Subjective video quality assessment methods for recognition tasks (2016)
ITU-T: ITU-T Rec. P.912 (03/2016) Subjective video quality assessment methods for recognition tasks, 2016

Jan, B. et al.: Deep learning in big data Analytics: A comparative study (2019)
Jan, Bilal; Farman, Haleem; Khan, Murad; Imran, Muhammad; Islam, Ihtesham Ul; Ahmad, Awais; Ali, Shaukat; Jeon, Gwanggil: Deep learning in big data Analytics: A comparative study, in: Computers & Electrical Engineering, Vol. 75, pp. 275–287, 2019

Janiesch, C. et al.: Machine learning and deep learning (2021)
Janiesch, Christian; Zschech, Patrick; Heinrich, Kai: Machine learning and deep learning, in: Electronic Markets, Vol. 31, pp. 685–695, 2021

Jha, S. et al.: Watch Out for the Safety-Threatening Actors: Proactively Mitigating Safety Hazards (2022)
Jha, Saurabh; Cui, Shengkun; Kalbarczyk, Zbigniew; Iyer, Ravishankar K.: Watch Out for the Safety-Threatening Actors: Proactively Mitigating Safety Hazards, in: ArXiv, 2022

Jifa, G.: Data, Information, Knowledge, Wisdom and Meta-Synthesis of Wisdom (2013)
Jifa, Gu: Data, Information, Knowledge, Wisdom and Meta-Synthesis of Wisdom - Comment on Wisdom Global and Wisdom Cities, in: Procedia Computer Science, Vol. 17, pp. 713–719, 2013

Johanson, M. et al.: Big Automotive Data (2014)
Johanson, Mathias; Belenki, Stanislav; Jalminger, Jonas; Fant, Magnus; Gjertz, Mats: Big Automotive Data: Leveraging large volumes of data for knowledge-driven product development, in: 2014 IEEE International Conference on Big Data (Big Data), pp. 736–741, 2014

Kim, S.-W. et al.: Multivehicle cooperative driving using cooperative perception (2014)
Kim, Seong-Woo; Qin, Baoxing; Chong, Zhuang Jie; Shen, Xiaotong; Liu, Wei; Ang, Marcelo H; Frazzoli, Emilio; Rus, Daniela: Multivehicle cooperative driving using cooperative perception: Design and experimental validation, in: IEEE Transactions on Intelligent Transportation Systems, Vol. 16, pp. 663–680, 2014

Kiureghian, A. D. et al.: Aleatory or epistemic? Does it matter? (2009)
Kiureghian, Armen Der; Ditlevsen, Ove: Aleatory or epistemic? Does it matter?, in: Structural Safety, Vol. 31, pp. 105–112, 2009

Klitzke, L. et al.: Vehicle Data Management System for Validation of ADS (2019)
Klitzke, Lars; Koch, Carsten; Haja, Andreas; Köster, Frank: Real-world Test Drive Vehicle Data Management System for Validation of Automated Driving Systems, in: Proceedings of the 5th International Conference on Vehicle Technology and Intelligent Transport Systems, pp. 171–180, 2019

Kramer, B. et al.: Identification and Quantification of Hazardous Scenarios for Automated Driving (2020)
Kramer, Birte; Neurohr, Christian; Büker, Matthias; Böde, Eckard; Fränzle, Martin; Damm, Werner: Identification and Quantification of Hazardous Scenarios for Automated Driving, in: Model-Based Safety and Assessment, pp. 163–178, 2020

Kraus, D. et al.: Approach for an Optical Network Design for Autonomous Vehicles (2019)
Kraus, Daniel; Ivanov, Hristo; Leitgeb, Erich: Approach for an Optical Network Design for Autonomous Vehicles, in: 2019 21st International Conference on Transparent Optical Networks (ICTON), pp. 1–6, 2019

Křemen, P. et al.: Ontology-Driven Information System Design (2012)
Křemen, Petr; Kouba, Zdeněk: Ontology-Driven Information System Design, in: IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), Vol. 42, pp. 334–344, 2012

Lenz, B. et al.: New Mobility Concepts and Autonomous Driving: The Potential for Change (2016)
Lenz, Barbara; Fraedrich, Eva: New Mobility Concepts and Autonomous Driving: The Potential for Change, in: Gerdes, J. Christian; Lenz, Barbara; Maurer, Markus; Winner, Hermann (Hrsg.): Autonomous Driving, Springer Berlin Heidelberg and Imprint: Springer, 2016

Li, M. et al.: Learning Convolutional Networks for Content-weighted Image Compression (2017)
Li, Mu; Zuo, Wangmeng; Gu, Shuhang; Zhao, Debin; Zhang, David: Learning Convolutional Networks for Content-weighted Image Compression, URL: http://arxiv.org/pdf/1703.10553v2, 2017

Li, Z.-N.; Drew, M. S.; Liu, J.: Fundamentals of Multimedia (2021)
Li, Ze-Nian; Drew, Mark S.; Liu, Jiangchuan: Fundamentals of Multimedia, Springer eBook Collection, 3rd ed. 2021. Edition, Springer International Publishing and Imprint Springer, 2021

Luce, R. D.: A Note on Boolean Matrix Theory (1952)
Luce, R. Duncan: A Note on Boolean Matrix Theory, in: Proceedings of the American Mathematical Society, Vol. 3, pp. 382–388, 1952

Luo, C.: Understanding Diffusion Models: A Unified Perspective (2022)
Luo, Calvin: Understanding Diffusion Models: A Unified Perspective, URL: http://arxiv.org/pdf/2208.11970v1, 2022

Luo, S. et al.: DeepSIC: Deep Semantic Image Compression (2018)
Luo, Sihui; Yang, Yezhou; Song, Mingli: DeepSIC: Deep Semantic Image Compression, URL: http://arxiv.org/pdf/1801.09468v1, 2018

Lyssenko, M. et al.: Towards Task-oriented Relevance Metrics for Pedestrian Detection (2021)
Lyssenko, Maria; Gladisch, Christoph; Heinzemann, Christian; Woehrle, Matthias; Triebel, Rudolph: From Evaluation to Verification: Towards Task-oriented Relevance Metrics for Pedestrian Detection in Safety-critical Domains, in: 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 38–45, 2021

Maag, C. et al.: Supporting cooperative driving behaviour by technology (2022)
Maag, Christian; Kraft, Ann-Kathrin; Neukum, Alexandra; Baumann, Martin: Supporting cooperative driving behaviour by technology–HMI solution, acceptance by drivers and effects on workload and driving behaviour, in: Transportation research part F: traffic psychology and behaviour, Vol. 84, pp. 139–154, 2022

Majstorović, D. et al.: Survey on Teleoperation Concepts for Automated Vehicles (2022)
Majstorović, Domagoj; Hoffmann, Simon; Pfab, Florian; Schimpe, Andreas; Wolf, Maria-Magdalena; Diermeyer, Frank: Survey on Teleoperation Concepts for Automated Vehicles, in: 2022 IEEE International Conference on Systems, Man, and Cybernetics (SMC), pp. 1290–1296, 2022

Mao, J. et al.: One Million Scenes for Autonomous Driving: ONCE Dataset (2021)
Mao, Jiageng; Niu, Minzhe; Jiang, Chenhan; Liang, Hanxue; Chen, Jingheng; Liang, Xiaodan; Li, Yamin; Ye, Chaoqiang; Zhang, Wei; Li, Zhenguo; Yu, Jie; Xu, Hang; Xu, Chunjing: One Million Scenes for Autonomous Driving: ONCE Dataset, in: ArXiv, 2021

Mao, J. et al.: 3D Object Detection for Autonomous Driving (2023)
Mao, Jiageng; Shi, Shaoshuai; Wang, Xiaogang; Li, Hongsheng: 3D Object Detection for Autonomous Driving: A Comprehensive Survey, in: International Journal of Computer Vision, 2023

Marti, E. et al.: A Review of Sensor Technologies for Perception in Automated Driving (2019)
Marti, Enrique; Miguel, Miguel Angel de; Garcia, Fernando; Perez, Joshue: A Review of Sensor Technologies for Perception in Automated Driving, in: IEEE Intelligent Transportation Systems Magazine, Vol. 11, pp. 94–108, 2019

Menzel, T. et al.: Scenarios for Development, Test and Validation of Automated Vehicles (2018)
Menzel, Till; Bagschik, Gerrit; Maurer, Markus: Scenarios for Development, Test and Validation of Automated Vehicles, URL: http://arxiv.org/pdf/1801.08598v3, 2018

Mizzaro, S.: How many relevances in information retrieval? (1998)
Mizzaro, Stefano: How many relevances in information retrieval?, in: Interact. Comput. Vol. 10, pp. 303–320, 1998

Mizzaro, S.: Relevance: The whole history (1997)
Mizzaro, Stefano: Relevance: The whole history, in: Journal of the American Society for Information Science, Vol. 48, pp. 810–832, 1997

Mori, K.: Defining Object Detection Requirements for Safe Automated Driving (2023)
Mori, Ken: Defining Object Detection Requirements for Automated Driving, Dissertation, Technische Universität Darmstadt, 2023

Mori, K. et al.: Conservative Estimation of Perception Relevance of Dynamic Objects (2023)
Mori, Ken; Storms, Kai; Peters, Steven: Conservative Estimation of Perception Relevance of Dynamic Objects for Safe Trajectories in Automotive Scenarios, in: 2023 IEEE International Conference on Mobility, Operations, Services and Technologies (MOST), pp. 83–95, 2023

Neumeier, S. et al.: Teleoperation: The Holy Grail to Solve Problems of Automated Driving? (2019)
Neumeier, Stefan; Wintersberger, Philipp; Frison, Anna-Katharina; Becher, Armin; Facchi, Christian; Riener, Andreas: Teleoperation: The Holy Grail to Solve Problems of Automated Driving? Sure, but Latency Matters, in: Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, pp. 186–197, 2019

Neurohr, C. et al.: Criticality Analysis for the Verification and Validation of Automated Vehicles (2021)
Neurohr, Christian; Westhofen, Lukas; Butz, Martin; Bollmann, Martin Herbert; Eberle, Ulrich; Galbas, Roland: Criticality Analysis for the Verification and Validation of Automated Vehicles, in: IEEE Access, Vol. 9, pp. 18016–18041, 2021

Neurohr, C. et al.: Fundamental Considerations around Scenario-Based Testing (2020)
Neurohr, Christian; Westhofen, Lukas; Henning, Tabea; Graaff, Thies de; Möhlmann, Eike; Böde, Eckard: Fundamental Considerations around Scenario-Based Testing for Automated Driving, in: 2020 IEEE Intelligent Vehicles Symposium (IV), pp. 121–127, 2020

Ngo, H. et al.: Cooperative Perception With V2V Communication for Autonomous Vehicles (2023)
Ngo, Hieu; Fang, Hua; Wang, Honggang: Cooperative Perception With V2V Communication for Autonomous Vehicles, in: IEEE Transactions on Vehicular Technology, 2023

O'Brien, A.: Relevance as an aid to evaluation in OPACs (1990)
O'Brien, A.: Relevance as an aid to evaluation in OPACs, in: Journal of Information Science, Vol. 16, pp. 265–271, 1990

Ogden, C. K. et al.: The meaning of meaning (1923)
Ogden, Charles Kay; Richards, Ivor Armstrong: The meaning of meaning: A study of the influence of thought and of the science of symbolism, 1923

Padilla, R. et al.: A Survey on Performance Metrics for Object-Detection Algorithms (2020)
Padilla, Rafael; Netto, Sergio; Da Silva, Eduardo: A Survey on Performance Metrics for Object-Detection Algorithms, in: Proceedings of the 2020 International Conference on Systems, Signals and Image Processing (IWSSIP), 2020

paperswithcode.com: Cityscapes test Benchmark (Real-Time Semantic Segmentation) (2023)
paperswithcode.com: Cityscapes test Benchmark (Real-Time Semantic Segmentation), URL: https://paperswithcode.com/sota/real-time-semantic-segmentation-on-cityscapes?p=pidnet-a-real-time-semantic-segmentation, 2023

Parmar, P. et al.: Real time detection and reporting of vehicle collision (2017)
Parmar, Parag; Sapkal, Ashok M.: Real time detection and reporting of vehicle collision, in: 2017 International Conference on Trends in Electronics and Informatics (ICEI), pp. 1029–1034, 2017

Patwa, N. et al.: Semantic-Preserving Image Compression (2020)
Patwa, Neel; Ahuja, Nilesh; Somayazulu, Srinivasa; Tickoo, Omesh; Varadarajan, Srenivas; Koolagudi, Shashidhar: Semantic-Preserving Image Compression, in: 2020 IEEE International Conference on Image Processing (ICIP), pp. 1281–1285, 2020

PEGASUS Project: PEGASUS Method: An Overview (2019)
PEGASUS Project: PEGASUS Method: An Overview, URL: https://www.pegasusprojekt.de/files/tmpl/Pegasus-Abschlussveranstaltung/PEGASUS-Gesamtmethode.pdf, 2019

Pham, C.-H. et al.: Principal Component Analysis Autoencoder for Organising the Latent Space (2022)
Pham, Chi-Hieu; Ladjal, Saïd; Newson, Alasdair: PCA-AE: Principal Component Analysis Autoencoder for Organising the Latent Space of Generative Networks, in: Journal of Mathematical Imaging and Vision, Vol. 64, pp. 569–585, 2022

Philion, J. et al.: Learning to Evaluate Perception Models Using Planner-Centric Metrics (2020)
Philion, Jonah; Kar, Amlan; Fidler, Sanja: Learning to Evaluate Perception Models Using Planner-Centric Metrics, in: arXiv:2004.08745 [cs], 2020

Philipp, R. et al.: Systematization of Relevant Road Users (2022)
Philipp, Robin; Rehbein, Jana; Grun, Felix; Hartjen, Lukas; Zhu, Zhijing; Schuldt, Fabian; Howar, Falk: Systematization of Relevant Road Users for the Evaluation of Autonomous Vehicle Perception, in: 2022 IEEE International Systems Conference (SysCon), pp. 1–8, 2022

Poddey, A. et al.: On the validation of complex systems operating in open contexts (2019)
Poddey, Alexander; Brade, Tino; Stellet, Jan Erik; Branz, Wolfgang: On the validation of complex systems operating in open contexts, in: ArXiv, Vol. abs/1902.10517, 2019

Poul Greibe: Braking distance, friction and behaviour: Findings, analyses and recommendations (2007)
Poul Greibe: Braking distance, friction and behaviour: Findings, analyses and recommendations based on braking trials, URL: https://www.trafitec.dk/sites/default/files/publications/braking%20distance%20-%20friction%20and%20driver%20behaviour.pdf, 2007

Regazzi, J. J.: Performance measures for information retrieval systems—an experimental approach (1988)
Regazzi, John J.: Performance measures for information retrieval systems—an experimental approach, in: Journal of the American Society for Information Science, Vol. 39, pp. 235–251, 1988

Rios, A. et al.: Closed-Loop Memory GAN for Continual Learning (2018)
Rios, Amanda; Itti, Laurent: Closed-Loop Memory GAN for Continual Learning, URL: http://arxiv.org/pdf/1811.01146v3, 2018

Rowley, J.: The wisdom hierarchy: representations of the DIKW hierarchy (2007)
Rowley, Jennifer: The wisdom hierarchy: representations of the DIKW hierarchy, in: Journal of Information Science, Vol. 33, pp. 163–180, 2007

Rueß, H. et al.: Whitepaper: Safe AI. How is this possible? (2022)
Rueß, Harald; Burton, Simon: Whitepaper: Safe AI. How is this possible?, 2022

Ruxton, G. D.: The unequal variance t-test is an underused alternative to Student's t-test and the Mann–Whitney U test (2006)
Ruxton, Graeme D.: The unequal variance t-test is an underused alternative to Student's t-test and the Mann–Whitney U test, in: Behavioral Ecology, Vol. 17, pp. 688–690, 2006

Sanders, J.: Defining terms: Data, information and knowledge (2016)
Sanders, John: Defining terms: Data, information and knowledge, in: 2016 SAI Computing Conference (SAI), pp. 223–228, 2016

Saracevic, T.: RELEVANCE reconsidered (1996)
Saracevic, Tefko: RELEVANCE reconsidered, 1996

Saracevic, T.: RELEVANCE: A review of and a framework (1975)
Saracevic, Tefko: RELEVANCE: A review of and a framework for the thinking on the notion in information science, in: Journal of the American Society for Information Science, Vol. 26, pp. 321–343, 1975

Saracevic, T.: RELEVANCE: Part II: nature and manifestations of relevance (2007)
Saracevic, Tefko: Relevance: A review of the literature and a framework for thinking on the notion in information science. Part II: nature and manifestations of relevance, in: Journal of the American Society for Information Science and Technology, Vol. 58, 2007

Schamber, L.: Relevance and Information Behavior. (1994)
Schamber, Linda: Relevance and Information Behavior., in: 1994

Schamber, L. et al.: A re-examination of relevance: toward a dynamic, situational definition∗ (1990)
Schamber, Linda; Eisenberg, Michael B.; Nilan, Michael S.: A re-examination of relevance: toward a dynamic, situational definition∗, in: Information Processing & Management, Vol. 26, pp. 755–776, 1990

Schiopu, I. et al.: Deep-Learning based Lossless Image Coding (2020)
Schiopu, Ionut; Munteanu, Adrian: Deep-Learning based Lossless Image Coding, in: IEEE Transactions on Circuits and Systems for Video Technology, p. 1, 2020

Schmidt, C.: Unfallvermeidung im Straßenverkehr für Einzel- und Mehrobjektszenarien (2013)
Schmidt, Christian: Fahrstrategien zur Unfallvermeidung im Straßenverkehr für Einzel- und Mehrobjektszenarien, 2013

Schönemann, V. et al.: Maneuver-based Adaptive Safety Zone for Valet Parking (2019)
Schönemann, Valerij; Duschek, Mara; Winner, Hermann: Maneuver-based Adaptive Safety Zone for Infrastructure-Supported Automated Valet Parking, in: 2184-495X, 2019

Shaham, T. R. et al.: Deformation Aware Image Compression (2018)
Shaham, Tamar Rott; Michaeli, Tomer: Deformation Aware Image Compression, URL: http://arxiv.org/pdf/1804.04593v1, 2018

Shalev-Shwartz, S. et al.: On a Formal Model of Safe and Scalable Self-driving Cars (2017)
Shalev-Shwartz, Shai; Shammah, Shaked; Shashua, Amnon: On a Formal Model of Safe and Scalable Self-driving Cars, URL: http://arxiv.org/pdf/1708.06374v6, 2017

Shi, E. et al.: The Principles of Operation Framework (2020)
Shi, Elisabeth; Gasser, Tom Michael; Seeck, Andre; Auerswald, Rico: The Principles of Operation Framework: A Comprehensive Classification Concept for Automated Driving Functions, in: SAE International Journal of Connected and Automated Vehicles, Vol. 3, 2020

Soergel, D.: Indexing and retrieval performance: The logical evidence (1994)
Soergel, Dagobert: Indexing and retrieval performance: The logical evidence, in: Journal of the American Society for Information Science, Vol. 45, pp. 589–599, 1994

Sperber, D.; Wilson, D.: Relevance: communication and cognition (2001)
Sperber, Dan; Wilson, Deirdre: Relevance: communication and cognition, 2nd ed.. Edition, Blackwell Publishers, 2001

Spink, A. et al.: From Highly Relevant to Not Relevant (1998)
Spink, Amanda; Greisdorf, Howard; Bateman, Judy: From Highly Relevant to Not Relevant: Examining Different Regions of Relevance, in: Information Processing & Management, pp. 599–621, 1998

Standarts, N. I. of et al.: 1.3.5.10. Levene Test for Equality of Variances (2022)
Standarts, National Institute of; Technology: 1.3.5.10. Levene Test for Equality of Variances, URL: https://www.itl.nist.gov/div898/handbook/eda/section3/eda35a.htm, 2022

Steinkirchner, K. et al.: Usage of Baselining and Traceability Demonstrator Developed in VVM (2022)
Steinkirchner, Kim; Bühler, Christian: Usage of Baselining and Traceability Demonstrator Developed in VVM, VVM Mid-term presentation, 2022

Stellet, J. E. et al.: Formalisation and algorithmic approach to the AD validation problem (2019)
Stellet, Jan Erik; Brade, Tino; Poddey, Alexander; Jesenski, Stefan; Branz, Wolfgang: Formalisation and algorithmic approach to the automated driving validation problem, in: 2019 IEEE Intelligent Vehicles Symposium (IV), pp. 45–51, 2019

Stockman, George and Shapiro, Linda G.: Computer Vision (2000)
Stockman, George and Shapiro, Linda G.: Computer Vision, 1. Edition, Prentice Hall PTR, 2000

Storms, K. et al.: SURE-Val: Safe Urban Relevance Extension and Validation (2023)
Storms, Kai; Mori, Ken; Peters, Steven: SURE-Val: Safe Urban Relevance Extension and Validation, 2023

Sun, P. et al.: Scalability in Perception (2019)
Sun, Pei; Kretzschmar, Henrik; Dotiwalla, Xerxes; Chouard, Aurelien; Patnaik, Vijaysai; Tsui, Paul; Guo, James; Zhou, Yin; Chai, Yuning; Caine, Benjamin; Vasudevan, Vijay; Han, Wei; Ngiam, Jiquan; Zhao, Hang; Timofeev, Aleksei; Ettinger, Scott; Krivokon, Maxim; Gao, Amy; Joshi, Aditya; Zhao, Sheng; Cheng, Shuyang; Zhang, Yu; Shlens, Jonathon; Chen, Zhifeng; Anguelov, Dragomir: Scalability in Perception for Autonomous Driving: Waymo Open Dataset, in: ArXiv, 2019

Sun, P. et al.: Scalability in Perception for Autonomous Driving: Waymo Open Dataset (2019)
Sun, Pei; Kretzschmar, Henrik; Dotiwalla, Xerxes; Chouard, Aurelien; Patnaik, Vijaysai; Tsui, Paul; Guo, James; Zhou, Yin; Chai, Yuning; Caine, Benjamin; Vasudevan, Vijay; Han, Wei; Ngiam, Jiquan; Zhao, Hang; Timofeev, Aleksei; Ettinger, Scott M.; Krivokon, Maxim; Gao, Amy; Joshi, Aditya; Zhang, Yu; Shlens, Jonathon; Chen, Zhifeng; Anguelov, Dragomir: Scalability in Perception for Autonomous Driving: Waymo Open Dataset, in: 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2443–2451, 2019

Sun, T. et al.: SHIFT: A Synthetic Driving Dataset for Continuous Multi-Task Domain Adaptation (2022)
Sun, Tao; Segu, Mattia; Postels, Janis; Wang, Yuxuan; Van Gool, Luc; Schiele, Bernt; Tombari, Federico; Yu, Fisher: SHIFT: A Synthetic Driving Dataset for Continuous Multi-Task Domain Adaptation, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 21371–21382, 2022

Sun, X. et al.: A Task-Driven Scene-Aware LiDAR Point Cloud Coding Framework (2023)
Sun, Xuebin; Wang, Miaohui; Du, Jingxin; Sun, Yuxiang; Cheng, Shing Shin; Xie, Wuyuan: A Task-Driven Scene-Aware LiDAR Point Cloud Coding Framework for Autonomous Vehicles, in: IEEE Transactions on Industrial Informatics, Vol. 19, pp. 8731–8742, 2023

Szimba, E. et al.: Assessing travel time savings and user benefits of automated driving (2020)
Szimba, Eckhard; Hartmann, Martin: Assessing travel time savings and user benefits of automated driving - A case study for a commuting relation, 2020

Tangramvision: Sensing Breakdown: Waymo Jaguar I-Pace RoboTaxi (2022)
Tangramvision: Sensing Breakdown: Waymo Jaguar I-Pace RoboTaxi. LiDAR, radar, cameras, ultrasonic, oh my!, URL: https://www.tangramvision.com/blog/sensing-breakdown-waymo-jaguar-i-pace-robotaxi, 2022

Taylor, A. et al.: Relationships between categories of relevance criteria and stage in task completion (2007)
Taylor, Arthur; Cool, Colleen; Belkin, Nicholas; Amadio, William J.: Relationships between categories of relevance criteria and stage in task completion, in: Information Processing & Management, Vol. 43, pp. 1071–1084, 2007

Taylor, L. et al.: Improving Deep Learning with Generic Data Augmentation (2018)
Taylor, Luke; Nitschke, Geoff: Improving Deep Learning with Generic Data Augmentation, in: Sundaram, Sureh (Hrsg.): Proceedings of the 2018 IEEE Symposium Series on Computational Intelligence (SSCI 2018), IEEE, 2018

Taylor, R. S.: Value-added processes in information systems (1986)
Taylor, Robert Saxton: Value-added processes in information systems, Greenwood Publishing Group, 1986

The SciPy community: scipy.stats.levene — SciPy v1.11.3 Manual (2023)
The SciPy community: scipy.stats.levene — SciPy v1.11.3 Manual, URL: https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.levene.html#scipy.stats.levene, 2023

The SciPy community: scipy.stats.ttest_ind — SciPy v1.11.3 Manual (2023)
The SciPy community: scipy.stats.ttest_ind — SciPy v1.11.3 Manual, URL: https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.ttest_ind.html#scipy.stats.ttest_ind, 2023

Till Beemelmanns et al.: 3D Point Cloud Compression with Recurrent Neural Network (2022)
Till Beemelmanns; Yuchen Tao; Bastian Lampe; Lennart Reiher; Raphael Van Kempen; Timo Woopen; Lutz Eckstein: 3D Point Cloud Compression with Recurrent Neural Network and Image Compression Methods, 2022

Tingvall, C. et al.: Vision Zero: An ethical approach to safety and mobility (1999)
Tingvall, Claes; Haworth, Narelle: Vision Zero: An ethical approach to safety and mobility, in: ITE International Conference Road Safety & Traffic Enforcement, Vol. 6, 1999

Toderici, G. et al.: Full Resolution Image Compression with Recurrent Neural Networks (2016)
Toderici, George; Vincent, Damien; Johnston, Nick; Hwang, Sung Jin; Minnen, David; Shor, Joel; Covell, Michele: Full Resolution Image Compression with Recurrent Neural Networks, URL: http://arxiv.org/pdf/1608.05148v2, 2016

Topan, S. et al.: Interaction-Dynamics-Aware Perception Zones for Obstacle Detection Safety Evaluation (2022)
Topan, Sever; Leung, Karen; Chen, Yuxiao; Tupekar, Pritish; Schmerling, Edward; Nilsson, Jonas; Cox, Michael; Pavone, Marco: Interaction-Dynamics-Aware Perception Zones for Obstacle Detection Safety Evaluation, URL: http://arxiv.org/pdf/2206.12471v1, 2022

Tran, N.-T. et al.: On Data Augmentation for GAN Training (2021)
Tran, Ngoc-Trung; Tran, Viet-Hung; Nguyen, Ngoc-Bao; Nguyen, Trung-Kien; Cheung, Ngai-Man: On Data Augmentation for GAN Training, in: IEEE transactions on image processing : a publication of the IEEE Signal Processing Society, Vol. 30, pp. 1882–1897, 2021

Tuohy, S. et al.: Intra-Vehicle Networks: A Review (2015)
Tuohy, Shane; Glavin, Martin; Hughes, Ciaran; Jones, Edward; Trivedi, Mohan; Kilmartin, Liam: Intra-Vehicle Networks: A Review, in: IEEE Transactions on Intelligent Transportation Systems, Vol. 16, pp. 534–545, 2015

U.S. Department of Transportation NHTSA: Automated Driving Systems: A Vision for Safety (2017)
U.S. Department of Transportation NHTSA: Automated Driving Systems: A Vision for Safety, 2017

Ulbrich, S. et al.: Defining and Substantiating the Terms Scene, Situation, and Scenario (2015)
Ulbrich, Simon; Menzel, Till; Reschka, Andreas; Schuldt, Fabian; Maurer, Markus: Defining and Substantiating the Terms Scene, Situation, and Scenario for Automated Driving, in: 2015 IEEE 18th International Conference on Intelligent Transportation Systems, pp. 982–988, 2015

Vakkari, P. et al.: Changes in Relevance Criteria and Problem Stages in Task Performance (2000)
Vakkari, Pertti; Hakala, Nanna: Changes in Relevance Criteria and Problem Stages in Task Performance, in: Journal of Documentation, Vol. 56, pp. 540–562, 2000

Varischio, A. et al.: Hybrid Point Cloud Semantic Compression for Automotive Sensors (2021)
Varischio, Andrea; Mandruzzato, Francesco; Bullo, Marcello; Giordani, Marco; Testolina, Paolo; Zorzi, Michele: Hybrid Point Cloud Semantic Compression for Automotive Sensors: A Performance Evaluation, URL: http://arxiv.org/pdf/2103.03819v1, 2021

Vater, L. et al.: Test Case Selection Method for the Verification of Automated Driving Systems (2021)
Vater, Lennart; Pütz, Andreas; Tellis, Levasseur; Eckstein, Lutz: Test Case Selection Method for the Verification of Automated Driving Systems, in: ATZelectronics worldwide, Vol. 16, pp. 40–45, 2021

Victor, T. et al.: Safety performance of the Waymo rider-only automated driving system (2023)
Victor, T; Kusano, K; Gode, T; Chen, R; Schwall, M: Safety performance of the Waymo rider-only automated driving system at one million miles, 2023

Volk, G. et al.: A Comprehensive Safety Metric (2020)
Volk, Georg; Gamerdinger, Jörg; Betnuth, Alexander von; Bringmann, O.: A Comprehensive Safety Metric to Evaluate Perception in Autonomous Systems, in: 2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC), pp. 1–8, 2020

Wachenfeld, W. H. K. et al.: The Release of Autonomous Vehicles (2016)
Wachenfeld, Walther H. K.; Winner, Hermann: The Release of Autonomous Vehicles, in: Gerdes, J. Christian; Lenz, Barbara; Maurer, Markus; Winner, Hermann (Hrsg.): Autonomous Driving, Springer Berlin Heidelberg and Imprint: Springer, 2016

Wang, C.: Silent Testing for Safety Validation of Automated Driving in Field Operation (2021)
Wang, Cheng: Silent Testing for Safety Validation of Automated Driving in Field Operation, Dissertation, Technische Universität Darmstadt, 2021

Wang, C. et al.: Online Safety Assessment of Automated Vehicles Using Silent Testing (2022)
Wang, Cheng; Storms, Kai; Winner, Hermann: Online Safety Assessment of Automated Vehicles Using Silent Testing, in: IEEE Transactions on Intelligent Transportation Systems, Vol. 23, pp. 13069–13083, 2022

Wang, C. et al.: Reduction of Uncertainties for Safety Assessment Under Parallel Simulations (2021)
Wang, Cheng; Xiong, Fanglei; Winner, Hermann: Reduction of Uncertainties for Safety Assessment of Automated Driving Under Parallel Simulations, in: IEEE Transactions on Intelligent Vehicles, Vol. 6, pp. 110–120, 2021

Wang, J. et al.: Networking and Communications in Autonomous Driving: A Survey (2019)
Wang, Jiadai; Liu, Jiajia; Kato, Nei: Networking and Communications in Autonomous Driving: A Survey, in: IEEE Communications Surveys & Tutorials, Vol. 21, pp. 1243–1274, 2019

Wang, Y. et al.: A Two-stage H.264 based Video Compression Method for Automotive Cameras (2022)
Wang, Yiting; Chan, Pak Hung; Donzella, Valentina: A Two-stage H.264 based Video Compression Method for Automotive Cameras, in: 2022 IEEE 5th International Conference on Industrial Cyber-Physical Systems (ICPS), pp. 01–06, 2022

Wang, Y. et al.: Semantic-Aware Video Compression for Automotive Cameras (2023)
Wang, Yiting; Chan, Pak Hung; Donzella, Valentina: Semantic-Aware Video Compression for Automotive Cameras, in: IEEE Transactions on Intelligent Vehicles, Vol. 8, pp. 3712–3722, 2023

Wang, Z. et al.: Why is image quality assessment so difficult? (2002)
Wang, Zhou; Bovik, Alan C.; Lu, Ligang: Why is image quality assessment so difficult?, in: IEEE International Conference on Acoustics Speech and Signal Processing, pp. IV-3313-IV–3316, 2002

Wang, Z. et al.: Mean squared error: Love it or leave it? A new look at Signal Fidelity Measures (2009)
Wang, Zhou; Bovik, Alan Conrad: Mean squared error: Love it or leave it? A new look at Signal Fidelity Measures, in: (Hrsg.): IEEE Signal Processing Magazine, 2009

Wang, Z. et al.: Image quality assessment: from error visibility to structural similarity (2004)
Wang, Zhou; Bovik, Alan Conrad; Sheikh, Hamid Rahim; Simoncelli, Eero P.: Image quality assessment: from error visibility to structural similarity, in: IEEE transactions on image processing: a publication of the IEEE Signal Processing Society, Vol. 13, pp. 600–612, 2004

Westhofen, L.: Time To Collision (TTC) - Criticality Metrics (2023)
Westhofen, Lukas: Time To Collision (TTC) - Criticality Metrics, URL: https://criticality-metrics.readthedocs.io/en/latest/time-scale/TTC.html, 2023

Westhofen, L. et al.: Using Ontologies for the Formalization and Recognition of Criticality (2022)
Westhofen, Lukas; Neurohr, Christian; Butz, Martin; Scholtes, Maike; Schuldes, Michael: Using Ontologies for the Formalization and Recognition of Criticality for Automated Driving, in: IEEE Open Journal of Intelligent Transportation Systems, Vol. 3, pp. 519–538, 2022

Westhofen, L. et al.: Criticality Metrics for Automated Driving (2023)
Westhofen, Lukas; Neurohr, Christian; Koopmann, Tjark; Butz, Martin; Schütt, Barbara; Utesch, Fabian; Neurohr, Birte; Gutenkunst, Christian; Böde, Eckard: Criticality Metrics for Automated Driving: A Review and Suitability Analysis of the State of the Art, in: Archives of Computational Methods in Engineering, Vol. 30, pp. 1–35, 2023

Wilson, B. et al.: Argoverse 2 (2021)
Wilson, Benjamin; Qi, William; Agarwal, Tanmay; Lambert, John; Singh, Jagjeet; Khandelwal, Siddhesh; Pan, Bowen; Kumar, Ratnesh; Hartnett, Andrew; Kaesemodel Pontes, Jhony; Ramanan, Deva; Carr, Peter; Hays, James: Argoverse 2: Next Generation Datasets for Self-Driving Perception and Forecasting, in: Proceedings of the Neural Information Processing Systems Track on Datasets and Benchmarks, Vol. 1, 2021

Wilson, B. et al.: Argoverse 2: Next Generation Datasets for Self-driving Perception and Forecasting (2021)
Wilson, Benjamin; Qi, William; Agarwal, Tanmay; Lambert, John; Singh, Jagjeet; Khandelwal, Siddhesh; Pan, Bowen; Kumar, Ratnesh; Hartnett, Andrew; Pontes, Jhony Kaesemodel; Ramanan, Deva; Carr, Peter; Hays, James: Argoverse 2: Next Generation Datasets for Self-driving Perception and Forecasting, in: Proceedings of the Neural Information Processing Systems Track on Datasets and Benchmarks (NeurIPS Datasets and Benchmarks 2021), 2021

Wrigth, S. E.: From the Semiotic Triangle to the Semantic Web (2023)
Wrigth, Sue Ellen: From the Semiotic Triangle to the Semantic Web, in: Terminology Science & Research / Terminologie : Science et Recherche, pp. 111–135, 2023

Xu, J. et al.: PIDNet: A Real-time Semantic Segmentation Network Inspired by PID Controllers (2022)
Xu, Jiacong; Xiong, Zixiang; Bhattacharyya, Shankar P.: PIDNet: A Real-time Semantic Segmentation Network Inspired by PID Controllers, URL: https://arxiv.org/pdf/2206.02066, 2022

Yang, M. et al.: An overview of lossless digital image compression techniques (2005)
Yang, Ming; Bourbakis, N.: An overview of lossless digital image compression techniques, in: 48th Midwest Symposium on Circuits and Systems, 2005, 1099–1102 Vol. 2, 2005

Yang, Q. et al.: Machine-learning-enabled cooperative perception for connected autonomous vehicles (2021)
Yang, Qing; Fu, Song; Wang, Honggang; Fang, Hua: Machine-learning-enabled cooperative perception for connected autonomous vehicles: Challenges and opportunities, in: IEEE Network, Vol. 35, pp. 96–101, 2021

Z. Wang et al.: Multiscale structural similarity for image quality assessment (2003)
Z. Wang; E. P. Simoncelli; A. C. Bovik: Multiscale structural similarity for image quality assessment, in: The Thrity-Seventh Asilomar Conference on Signals, Systems & Computers, 2003, 1398–1402 Vol.2, 2003

Zhai, G. et al.: Perceptual image quality assessment: a survey (2020)
Zhai, Guangtao; Min, Xiongkuo: Perceptual image quality assessment: a survey, in: Science China Information Sciences, Vol. 63, pp. 1–52, 2020

Zhang, C. et al.: Occlusion-Aware Planning for Autonomous Driving (2023)
Zhang, Chi; Steinhauser, Florian; Hinz, Gereon; Knoll, Alois: Occlusion-Aware Planning for Autonomous Driving With Vehicle-to-Everything Communication, in: IEEE Transactions on Intelligent Vehicles, 2023

Zhang, Ethan; Zhang, Yi. “F-Measure”. In: *Encyclopedia of Database Systems*. Ed. by Ling Liu; M. Tamer Özsu. New York, NY: Springer New York, 2018, pp. 1492–1493. ISBN: 978-1-4614-8265-9. DOI: 10.1007/978-1-4614-8265-9_483. URL: https://doi.org/10.1007/978-1-4614-8265-9_483.

Zhang, Z. et al.: Bag of Freebies for Training Object Detection Neural Networks (2019)
Zhang, Zhi; He, Tong; Zhang, Hang; Zhang, Zhongyue; Xie, Junyuan; Li, Mu: Bag of Freebies for Training Object Detection Neural Networks, URL: http://arxiv.org/pdf/1902.04103v3, 2019

Own Publications

Storms, Kai; Mori, Ken; Peters, Steven: SURE-Val: Safe Urban Relevance Extension and Validation, FAS Workshop 2023, 2023-10, Berkheim, 2023.

Mori, Ken; **Storms, Kai**; Peters, Steven: Conservative Estimation of Perception Relevance of Dynamic Objects for Safe Trajectories in Automotive Scenarios, IEEE International Conference on Mobility: Operations, Services, and Technologies, 2023-05-17, Detroit, MI, USA, Volume 1, pp. 83-95, 2023.

Wang, Cheng; **Storms, Kai**; Winner, Hermann: Online Safety Assessment of Automated Vehicles Using Silent Testing, IEEE Transactions on Intelligent Transportation Systems, 2021-10, volume 23, pp 13069-13083, 2021.

Storms, Kai; Winner, Hermann: Context Aware Data Reduction: Selectively Lossless Data Reduction through Partially Synthetic Representations for Highly Automated Driving, WKM Symposium, 2021-06-15, Stuttgart (virtual), 2021. (Oral Presentation)

Gutenkunst, Christian; **Storms, Kai**: Enabling Analysis of Perception Phenomena for Highly Automated Driving by Using Redundant Sensor Setups in Automotive Scenarios, European Microwave Week 2021, 2022-02-16, London (virtual), 2022. (Oral Presentation)

Junietz, Philipp; Wachenfeld, Walter; Schönemann, Valerij; **Storms, Kai**; Tribelhorn, Wadim; Winner, Hermann: Gaining Knowledge on Automated Driving's Safety—The Risk-Free VAAFO Tool, Control Strategies for Advanced Driver Assistance Systems and Autonomous Driving Functions, 01-2019, some place, volume 476 , pp 47–65, 2019.

Supervised Theses

Yi Cui; Ruidi He; Zhihao Liaotian; Yuzhen Zhang; Yanhua Zhang; Yifei Wang: Development and Implementation of a Planer for Intersection Scenario for Automated Vehicle.
Advanced Design Project Nr. 151/20

Xu Li: Development and Implementation of an Algorithm for Predicting the State of Tracked Objects.
Master-Thesis Nr. 785/20

Laurenz Müller: Integration and Evaluation of Sensitivity Analysis Methods of Scenarios for highly Automated Driving.
Bachelor-Thesis Nr. 1386/21

Patrick Stahl: Development of a Selection Method of Safety Relevant Image Regions in Scenarios for highly Automated Driving.
Master-Thesis Nr. 821/21

Yiggit Ali Akaya: Learning Behavior Models for Actors in Urban Traffic.
Master-Thesis Nr. 830/21

Junyi Guo: Development of an Augmented Reality System for the Generation of Test Cases for Teleoperated Driving.
Master-Thesis Nr. 840/21

Tobias Hoyer: Development of a methodology to identify and map relevant entity-phenomenon interactions for highly automated driving.
Master-Thesis Nr. 841/21

Yifan Liu;Yumao Liu;Qiliang Jian;Adrienne Wachtel;Leon Nauerz: Design of a Test Dummy for Automated Driving.
Advanced Design Project Nr. 162/22

Zhuolin Zhao; Kexuan, Xia; Haoran Ding; Yutong Wu; Yibo Lyu: Implementation of a methodology to derive relevance information from driving data.
Advanced Design Project Nr. 168/22

Franziska Parthey: Systematic analysis, design, and combination of elements of safety argumentation chains for AEB systems.
Master-Thesis Nr. 846/22

Jinlong Wu: Investigation of Generative Methods for Selective Reduction of Image Data.
Master-Thesis Nr. 857/22

Faris Sa‘di: Development of a methodology to determine equivalence of scenarios for automated driving..
Master-Thesis Nr. 862/22

Yi Zhan: Evaluation of possible applications and influencing factors for the use of data reduction for highly automated driving..
Master-Thesis Nr. 866/22