

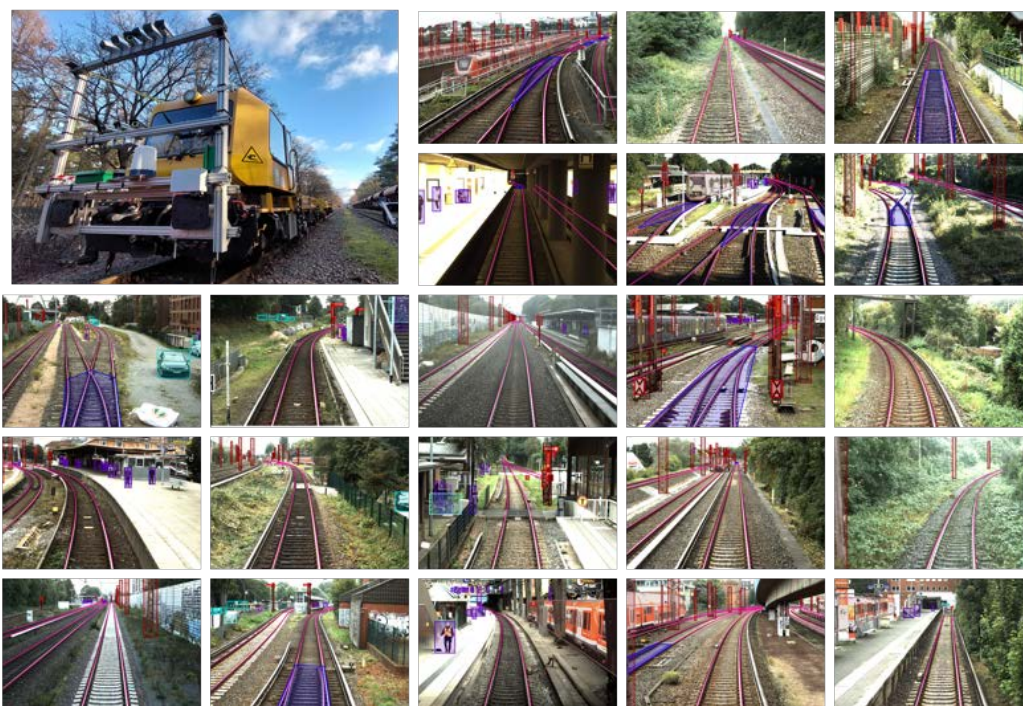


Berichte
des Deutschen Zentrums
für Schienenverkehrsforschung

Bericht 44 (2023)



Aufbereitung von Datensätzen für Anwendungen des automatisierten Fahrens im Eisenbahnbetrieb



Berichte des Deutschen Zentrums
für Schienenverkehrsforschung, Bericht 44 (2023)
Projektnummer 2021-20-D-1202

Aufbereitung von Datensätzen für Anwendungen des automatisierten Fahrens im Eisenbahnbetrieb

von

Kai-Uwe Kaden, Dr. Karsten Schwalbe, Michael Scheithauer
FusionSystems GmbH, Chemnitz

im Auftrag des Deutschen Zentrums für Schienenverkehrsforschung beim Eisenbahn-Bundesamt

Impressum

HERAUSGEBER

Deutsches Zentrum für Schienenverkehrsforschung beim Eisenbahn-Bundesamt

August-Bebel-Straße 10
01219 Dresden

www.dzsf.bund.de

DURCHFÜHRUNG DER STUDIE

FusionSystems GmbH
Annaberger Straße 240
09125 Chemnitz

ABSCHLUSS DER STUDIE

November 2022

REDAKTION

DZSF

Dr. Roman Tilly, Dr. Kai Hofmann, Forschungsbereich Querschnittsthemen, Digitalisierung, Automatisierung, Migration und Recht

BILDNACHWEIS

Eigene Darstellungen auf Grundlage von Fotos der DB Netz AG

PUBLIKATION ALS PDF

<https://www.dzsf.bund.de/Forschungsergebnisse/Forschungsberichte>

ISSN 2629-7973

[doi: 10.48755/dzsf.230012.01](https://doi.org/10.48755/dzsf.230012.01)

Dresden, Oktober 2023

Die Verantwortung für den Inhalt dieser Veröffentlichung liegt bei den Autoren.

Inhaltsverzeichnis

Abkürzungs- und Begriffsverzeichnis.....	7
Kurzbeschreibung	9
Abstract	10
1 Einleitung.....	11
2 Entwicklung der Annotationsvorschrift und des Qualitätssicherungskonzepts.....	13
2.1 Zusammenfassung Annotationsvorschrift.....	13
2.2 Anwendungsbereich (Sensoren) der Annotationsvorschrift	14
2.3 Übersicht der Annotationsgeometrien.....	14
2.3.1 Annotationsgeometrien in den Kameras und im RaDAR.....	14
2.3.2 Annotationsgeometrien im LiDAR	15
2.4 Übersicht der Objektklassen.....	16
2.4.1 Visueller Überblick	16
2.4.2 Zuordnung der Annotationsgeometrien zu den Objektklassen	21
2.4.3 Allgemeine Regeln für die Annotation	21
2.4.4 Überblick über die Attribute	23
2.5 Aufbau der Sensordaten des Referenzdatensatzes	40
2.6 Annotations- und Qualitätssicherungskonzept	41
2.6.1 Zuordnung der zu annotierenden Sequenzen	43
2.6.2 Annotation des ersten M-Frames.....	43
2.6.3 Review des ersten M-Frames.....	44
2.6.4 Annotation der Key-M-Frames	44
2.6.5 Review der Key-M-Frames.....	44
2.6.6 Ausführen der Interpolation	44
2.6.7 Korrektur der Interpolationsergebnisse.....	45
2.6.8 Automatische Konsistenzprüfungen.....	45
2.6.9 Interner Review aller M-Frames	46
2.6.10 Automatische Berechnung der SemSeg.....	46
2.6.11 Externer Review	46
2.6.12 Korrektur aller M-Frames.....	47
2.6.13 Umwandlung der Annotationsdaten	47
2.7 Beschreibung der automatischen Konsistenzprüfungen	47
2.7.1 Überprüfung von Lage und Anzahl der 2D-BBa.....	47

2.7.2	Überprüfung der Plausibilität der Abmessungen der 3D-BB	48
2.7.3	Überprüfung der Plausibilität der Tracking-IDs	49
2.7.4	Überprüfung der Plausibilität von Attributen der Kategorie „Abgeschnitten“	49
2.7.5	Überprüfung der zeitlichen und sensorübergreifenden Konsistenz der Attribute	50
2.8	Annotation einer Nulltranche	51
3	Aufbereitung des Datensatzes	52
3.1	Informationen zum Datensatz	52
3.2	Durchführung der Annotation	56
3.3	Statistik der erstellten Annotationen	60
3.4	Anonymisierung	62
4	Fazit	64
	Abbildungsverzeichnis	65
	Tabellenverzeichnis	67
	Quellenverzeichnis	68
	Anhänge	69
	Anhang 1: Sequenzübersicht	69
	Anhang 2: Grenzen der Abmessungen einzelner Objektklassen	84

Abkürzungs- und Begriffsverzeichnis

Abkürzung/Begriff	Bedeutung
2D-BBa	“two-dimensional Bounding Box (axis-aligned)” (engl.): achsenparalleles Rechteck zur Annotation von Objekten
2D-BBr	“two-dimensional Bounding Box (rotated)” (engl.): rotiertes Rechteck zur Annotation von Objekten
2D-PG	Zweidimensionales Polygon zur Annotation von Objekten
2D-PL	Zweidimensionale Polylinien zur Annotation von Objekten
3D-BB	„three-dimensional Bounding Box“ (engl.): (Rotierter) Quader zur Annotation von Objekten
3D-PL	Dreidimensionale Polylinien zur Annotation von Objekten
AKP	Annotationsprozesskoordinierende Person
AP	Annotierende Person
ATO	„Automated Train Operation“ (engl.): Automatisiertes Fahren im Eisenbahnbereich
EPL	Externe Projektleitung
EPP	Extern prüfende Person
FusionSystems Labeling Tool	Software des Auftragnehmers zur Annotation von Sensordaten
IPL	Interne Projektleitung
IPP	Intern prüfende Person
Key-M-Frame	M-Frame, der händisch annotiert wurde und dessen Annotationsgeometrien deswegen als Stützstellen für die Interpolation fungieren
KI	Künstliche Intelligenz
Labeling Guide	Annotationsvorschrift; Dokument, welches die Regeln zum Erstellen der Annotationen beinhaltet
Labeling Tool	Vom Auftragnehmer entwickelte Online-Anwendung zum Erzeugen von Annotationen
LiDAR	„Light Detection and Ranging“ (engl.), Sensor zur Ermittlung von dreidimensionalen Umgebungsinformationen mittels Lichtwellen
M-Frame	Multisensor-Frame: Zusammenschluss mehrerer zeitgleich erhobener Sensordatenframes
ML	Maschinelles Lernen, Sammelbezeichnung für Algorithmen, die anhand von Beispieldaten Muster erkennen
RaDAR	„Radio Detection and Ranging“ (engl.), Sensor zur Ortung von Objekten auf der Basis von Radiowellen

Abkürzung/Begriff	Bedeutung
Referenzdatensatz	Im Rahmen des vorliegenden Projektes annotierte und veröffentlichte Sensordaten
SemSeg	Semantische Segmentierung; Zuordnung von Punkten einer Punktwolke zu einem Objekt
Sequenz	Abfolge von im Zeitverlauf erhobenen Sensordaten
Tracking-ID	Identifikationsnummer zur Nachverfolgung

Kurzbeschreibung

Der vorliegende Forschungsbericht beschreibt das Projekt „Aufbereitung von Datensätzen für Anwendungen des automatisierten Fahrens im Eisenbahnbetrieb“. Für das automatisierte Fahren im Eisenbahnbetrieb ist es zwingend erforderlich, dass Aufgaben wie die vollständige Erfassung des Eisenbahnumfelds mittels geeigneter Sensorik und Algorithmik automatisiert erfolgen. Hierbei müssen den Sensordaten relevante Informationen, wie zum Beispiel Art, Lage und Attribute von wichtigen Objekten, akkurat zugeordnet werden. Für derartige Aufgaben haben sich in letzter Zeit Verfahren des Maschinellen Lernens als geeignetes Werkzeug herausgestellt. Diese „lernen“ entsprechende Zuordnungen anhand einer Vielzahl sogenannter Lerndaten und können nach Abschluss des Lernprozesses Objekte mit einer hohen Genauigkeit erkennen.

Jedes Verfahren kann damit nur so gut sein, wie die Lerndaten, welche ihm zugrundeliegen. Deswegen sind hochqualitative Lerndaten von immenser Wichtigkeit. Die Lerndaten bestehen dabei aus den Sensordaten (wie zum Beispiel Bilddaten und Punktwolken) und diesen zugeordneten Annotationen (wie zum Beispiel Markierungen der Objekte mittels entsprechender Geometrien und eine Auflistung relevanter Attribute). Im Rahmen des Projektes wurde ein solcher hochqualitativer Lerndatensatz erstellt.

Damit die Annotationen in Hinblick auf ihre spätere Verwendung einen maximalen Nutzen besitzen, müssen diese bestimmte Eigenschaften aufweisen, welche in einer Annotationsvorschrift festgehalten werden und deren Einhaltung im Rahmen eines Qualitätssicherungskonzeptes überprüft wird. Im vorliegenden Projekt wurde im Arbeitspaket (AP) 1 „Entwicklung der Annotationsvorgaben und des Qualitätssicherungskonzeptes“ sowohl eine solche detaillierte Annotationsvorschrift für den Bahnkontext erstellt, als auch eine ausführliche Beschreibung der Annotations- und Qualitätssicherungsprozesse erarbeitet. Die Annotationsvorschrift kann unter [1] abgerufen werden. Sie enthält detaillierte Informationen und Regeln zur Annotation der insgesamt 22 relevanten Objektklassen und deren klassenspezifischen Attribute. Dieser Schlussbericht fasst die wichtigsten Punkte der Annotationsvorschrift zusammen; er gibt einen Überblick über die Objektklassen sowie deren Attribute und legt dar, welche Annotationsgeometrien in den unterschiedlichen Sensormodalitäten (Kamera, LiDAR, RaDAR) für die jeweilige Objektklasse zu verwenden sind.

Dem im vorliegenden Projekt entwickelte Annotations- und Qualitätssicherungsprozess folgend können Annotationen erstellt werden, welche den Regeln der Annotationsvorschrift genügen. Er besteht aus einer Abfolge von 13 Einzelschritten, welche in Summe den Prozess vom manuellen Anlegen der ersten Annotationsgeometrien bis zum Export der Annotationsdaten in einem geeigneten Datenformat (im Projekt wird das ASAM OpenLABEL Format genutzt) beschreiben. Der Prozess ist dabei gezielt auf die Spezifika der im Projekt verwendeten Sensordaten (zeitlich aufeinanderfolgend und zueinander kalibriert) zugeschnitten, kann aber mittels Anpassungen auch für andere Sensordatensätze verwendet werden. Eine besondere Rolle nehmen im Qualitätssicherungsprozess die automatischen Konsistenzüberprüfungen ein, welche in diesem Schlussbericht detailliert beschrieben sind. Der Annotations- und Qualitätssicherungsprozess wurde im Rahmen einer sogenannten Nulltranche evaluiert und für geeignet empfunden.

Im AP 2 „Aufbereitung des Datensatzes“ wurden, dem Annotations- und Qualitätssicherungsprozess folgend, Annotationen für einen Referenzsensordatensatz erstellt. Der Datensatz, bestehend aus den Sensordaten mit zugehörigen Annotationen und aufgeteilt in zeitlich zusammenhängende Sequenzen, kann unter [2] abgerufen werden. Dabei aufgetretene besondere Herausforderungen sind in diesem Schlussbericht exemplarisch aufgeführt. In tabellarischen Übersichten sind wesentlichen Angaben zu den Sequenzen (Länge der Sequenz, repräsentatives Beispielbild, kurze verbale Beschreibung) und den entsprechenden Annotationen (Objektklassenverteilung) festgehalten. Weiterhin wurde im Rahmen von Arbeitspaket 2 die Richtigkeit der Anonymisierung von Personen und Kraftfahrzeugen sichergestellt.

Abstract

This research report describes the project "Development of Datasets for Applications of Automated Driving in Railway Operations". For automated driving in railway operations, it is imperative that tasks such as the complete recording of the railway environment are automated using suitable sensors and algorithms. Relevant information - such as the type, location and attributes of important objects - must be accurately assigned to the sensor data. Machine learning methods have recently emerged as a suitable tool for such tasks. These "learn" corresponding assignments based on a large number of so-called learning data and can recognize objects with a high level of accuracy after the learning process is complete.

Each method can only be as good as the learning data on which it is based. Therefore, high-quality learning data is of immense importance. The learning data consist of the sensor data (e.g. image data and point clouds) and annotations assigned to them (e.g. markings of the objects using corresponding geometries and a list of relevant attributes). Such a high-quality learning dataset was created as part of the project.

For the annotations to have maximum benefit with regard to their later use, they must have certain properties which are recorded in an annotation specification and compliance with which is checked as part of a quality assurance concept. In the present project, such a detailed annotation specification for the railway context was created in work package 1 "Development of the annotation specifications and the quality assurance concept", as well as a detailed description of the annotation and quality assurance processes. The labeling guide can be found under [1]. It contains detailed information and rules for annotating the 22 relevant object classes and their class-specific attributes. This final report summarizes the most important points of the labeling guide; it gives an overview of the object classes and their attributes and explains which annotation geometries are to be used in the different sensor modalities (camera, LiDAR, RaDAR) for the respective object class.

Following the annotation and quality assurance process developed in the present project, annotations can be created that meet the rules given in the labeling guide. It consists of a sequence of 13 individual steps, which in total describe the process from the manual creation of the first annotation geometries to the export of the annotation data in a suitable data format (the ASAM OpenLABEL format is used in the project). The process is specifically tailored to the specifics of the sensor data used in the project (chronologically consecutive and calibrated to one another) but can be used for other sensor data sets by means of adjustments. The automatic consistency checks, which are described in detail in this final report, also play a special role in the quality assurance process. The annotation and quality assurance process was evaluated as part of a so-called zero tranche and found to be suitable.

In work package 2 "Preparation of the data set", following the annotation and quality assurance process, annotations were created for a reference sensor data set. The data set, consisting of the sensor data with associated annotations and divided into chronologically related sequences, can be obtained under [2]. The special challenges that arose are listed as examples in this final report. Essential information on the sequences (length of the sequence, representative example image, short verbal description) and the corresponding annotations (distribution of object classes) are recorded in tabular overviews. Furthermore, as part of work package 2, the correctness of the anonymization of persons and vehicles was ensured.

1 Einleitung

Die Entwicklung und Zulassung des automatisierten Fahrens im Eisenbahnbetrieb (engl. „automatic train operation“, kurz: ATO) sind derzeit wichtige und zugleich technisch, organisatorisch und rechtlich anspruchsvolle Gebiete im System Bahn. Hinsichtlich der Überwachung des Zugumfelds und der Erkennung und Einordnung von Objekten in der Umgebung kommt dabei optischen Sensoren als Hardware-Komponente und künstlicher Intelligenz (KI) als Software-Komponente eine zentrale Rolle zu. Diese Komponenten werden voraussichtlich in der Lage sein müssen, das Umfeld des Zuges mindestens so zuverlässig und sicher wahrzunehmen wie derzeit der Triebfahrzeugführer bzw. die Triebfahrzeugführerin. Als optische Sensoren kommen hierbei zurzeit häufig Kombinationen aus Farbbildkameras, Infrarotkameras, Sensoren auf Basis von „Light Detection and Ranging“-Technologien (LiDAR) und Sensoren auf Basis von „Radio Detection and Ranging“-Technologien (RaDAR) zum Einsatz. Softwareseitig ist davon auszugehen, dass KI-Anwendungen mindestens die Prüfung auf etwaige Hindernisse im Bereich des zu befahrenden Gleises und dessen unmittelbarem Umfeld sowie die Erkennung und Klassifizierung definierter Arten von Objekten in diesem Bereich zuverlässig leisten müssen.

Die Algorithmen und Regeln, anhand derer die KI-Anwendungen diese Aufgaben leisten, werden zu einem signifikanten Teil mittels Verfahren des maschinellen Lernens (ML) aus Trainingsdaten automatisiert abgeleitet. Grundlage dafür sind wiederum qualitativ geeignete und ausreichend große Datensätze für Training und Validierung der KI-Anwendungen. Die Datensätze teilen sich dabei in Sensordaten und Annotationsdaten auf. Die Sensordaten werden durch die optischen Sensoren beziehungsweise Sensorsysteme erzeugt und stellen die Eingabedaten für die ML-Verfahren dar. Die Annotationsdaten beziehen sich auf einen konkreten Sensordatensatz und umfassen in der Regel räumliche Kennzeichnungen (zum Beispiel Koordinaten) relevanter erkannter Objekte sowie Klassifikationen der Objekte (zum Beispiel „Person“) und erkannte Ausprägungen bestimmter Attribute (zum Beispiel „Position: stehend“). Dies sind die Zieldaten, die KI-Anwendungen im Betrieb automatisch aus den Eingabedaten abzuleiten in der Lage sein sollen. Die hierfür notwendigen Regeln werden mittels ML und anhand von Datensätzen erlernt, bei denen Sensor-/Eingabedaten und Annotations-/Zieldaten bekannt sind. Um ML-Verfahren anwenden zu können, müssen die Annotationsdaten homogen sein, sprich die gleichen Objektarten umfassen und Objekte in gleicher Art und Weise kennzeichnen. Sofern Datensätze durch verschiedene Akteure unabhängig voneinander erhoben und annotiert werden, ist es deshalb sinnvoll, die Annotation auf Grundlage derselben Annotationsvorschrift (engl. „Labeling Guide“) durchzuführen. Die Annotationsvorschrift definiert beispielsweise die relevanten Objektklassen, mit welchen Geometrien (zum Beispiel Rechteck, Polygon, Polylinie) sie annotiert werden, welche Attribute annotiert werden, welche Attributausprägungen in welchen Fällen zu wählen sind und wie mit Grenzfällen umzugehen ist.

Im Eisenbahnbereich existieren derzeit kaum offene Datensätze (lediglich RailSem19 [3], FRSign [4] und RAWPED [5]) mit geeigneten Sensor- und Annotationsdaten [6] und auch kein Referenzrahmen zur Erstellung von Annotationsdaten für gegebene Sensordaten. Insgesamt stellen der Mangel an standardisierten Vorgehensweisen zur Datenerhebung und Datenaufbereitung von ML-geeigneten Datensätzen und die damit einhergehende mangelnde Verfügbarkeit an Datensätzen ein großes Hemmnis für die Entwicklung von KI-Anwendungen für ATO, die Entwicklung von Prüfverfahren für ATO-KI-Anwendungen und die Weiterentwicklung der Zulassungsverfahren für den Einsatz von ATO-KI-Anwendungen im Eisenbahnbetrieb dar.

Im Forschungsprojekt „Aufbereitung von Datensätzen für Anwendungen des automatisierten Fahrens im Eisenbahnbetrieb“ wurde ein Referenzrahmen und insbesondere eine Annotationsvorschrift für die Annotation von Datensätzen für die Entwicklung von KI-Anwendungen für ATO entwickelt. Diese Annotationsvorschrift wurde unter [1] veröffentlicht. Anhand dieser Annotationsvorschrift können Sensor-

daten aus einer dem Stand der Technik entsprechenden Sensorkonfiguration zu ML-geeigneten Datensätzen aufbereitet werden. Im Projekt wurde dies für einen ausgewählten Sensordatensatz, dem Referenzdatensatz, exemplarisch durchgeführt [2]. Dieser Datensatz kann als Referenz für Forschung und Entwicklung zu ATO-KI-Anwendungen verwendet werden.

2 Entwicklung der Annotationsvorschrift und des Qualitätssicherungskonzepts

Dieses Kapitel beschreibt die Projektergebnisse, die im Rahmen des Arbeitspaketes 1 „Entwicklung der Annotationsvorschrift und des Qualitätssicherungskonzepts“ entstanden sind. Nach einer kurzen Zusammenfassung des Inhaltes der Annotationsvorschrift (Abschnitt 2.1) und einer Beschreibung des Anwendungsbereichs der Annotationsvorschrift (Abschnitt 2.2) folgt eine Übersicht der verwendeten Annotationsgeometrien (Abschnitt 2.3) und Objektklassen (Abschnitt 2.4). Anschließend wird der Aufbau der Sensordaten im Referenzdatensatz erläutert (Abschnitt 2.5) und das erarbeitete Annotations- und Qualitätssicherungskonzept vorgestellt (Abschnitt 2.6). Die hierbei verwendeten automatischen Konsistenzüberprüfungen werden gesondert in Abschnitt 2.7 dargestellt. Zum Schluss wird das Vorgehen zur Herstellung eines einheitlichen Verständnisses der Annotationsvorschrift im Projekt anhand einer Nulltranche erläutert (Abschnitt 2.8).

2.1 Zusammenfassung Annotationsvorschrift

Ein zentraler Bestandteil des Projektes war die Erarbeitung einer detaillierten Annotationsvorschrift, die es erlaubt, Multi-Sensor-Datensätze (Datensätze, die aus zeitsynchronen Aufnahmen mehrerer Sensoren bestehen) für die Entwicklung von KI-Anwendungen für das autonome Fahren im Eisenbahnbetrieb aufzubereiten. Die Annotationsvorschrift beinhaltet Informationen zu den Punkten:

1. Für welche Sensordaten findet der Guide Anwendung?
2. Welche Annotationsarten (Geometrien, IDs, Attribute) existieren?
3. Welche Objektklassen werden annotiert?
4. Welche allgemeinen und welche klassenspezifischen Annotationsregeln gibt es?
5. In welchem Format werden die Annotationen abgespeichert?

Ziel der Annotationsvorschrift ist es zum einen, eindeutig zu dokumentieren, unter welchen Vorschriften die Annotationen im Projekt entstanden sind, und zum anderen, eine Referenz für zukünftige Annotationen im Bahnkontext zur Verfügung zu stellen. Die Annotationsvorschrift richtet sich an Forschende sowie Entwicklerinnen und Entwickler weltweit, die Sensordatensätze aufbereiten und auf der Basis der Annotationen autonome Fahrfunktionen entwickeln und testen können. Sie ist deswegen in englischer Sprache verfasst und wurde als separates Dokument veröffentlicht [1]. Dieses Kapitel enthält einen kurzen Überblick über die wesentlichen Inhalte der Annotationsvorschrift.

Die Annotationsvorschrift wurde im vorliegenden Projekt in einem iterativen Prozess ausgearbeitet. Den Ausgangspunkt bildeten initiale Vorgaben des Auftraggebers zu den Objektklassen, Annotationsgeometrien, Attributen und Annotationsregeln. Die Vorgaben basierten auf allgemeinen Überlegungen zur Nutzbarkeit von Annotationsdaten im Rahmen des automatisierten Fahrens im Eisenbahnbetrieb und wurden im Vorfeld des Annotationsprozesses mit dem Auftragnehmer diskutiert und, wo erforderlich, korrigiert bzw. spezifiziert. Der aktuelle Stand der Annotationsvorschrift kann in zukünftigen Projekten weiterentwickelt werden.

2.2 Anwendungsbereich (Sensoren) der Annotationsvorschrift

Die Ausführungen in der Annotationsvorschrift beziehen sich auf Sensoren, die an der Spitze eines Zuges befestigt sind und Sensordaten zum Zugumfeld liefern. Sie beziehen sich ausschließlich auf die folgenden Arten von Sensordaten:

- Punktwolken, die mittels LiDARen aufgenommen wurden
- Bilder, die mittels Farbbild- oder Infrarotkameras aufgenommen wurden
- Bilder, die mittels RaDARen aufgenommen wurden

2.3 Übersicht der Annotationsgeometrien

Der folgende Abschnitt gibt einen Überblick über die verschiedenen Annotationsgeometrien, die in der Annotationsvorschrift und im Datensatz verwendet werden.

2.3.1 Annotationsgeometrien in den Kameras und im RaDAR

In den Kamera- und RaDAR-Daten kommen folgende Annotationsgeometrien zum Einsatz:

- Achsenparallele Rechtecke (engl. „two-dimensional Bounding Box (axis-aligned)“, Abkürzung: 2D-BBa), siehe Abbildung 1
- Rotierte Rechtecke (engl. „two-dimensional Bounding Box (rotated)“, Abkürzung: 2D-BBr), siehe Abbildung 2
- Zweidimensionale Polylinien (Abkürzung: 2D-PL), siehe Abbildung 3
- Zweidimensionaler Polygone (Abkürzung: 2D-PG), siehe Abbildung 4



Abbildung 1: Annotation eines Objektes mit einer 2D-BBa

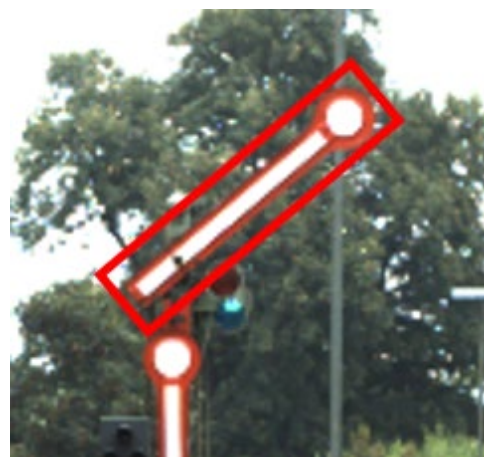


Abbildung 2: Annotation eines Objektes mit einer 2D-BBr



Abbildung 3: Annotation eines Objektes mit zwei 2D-PL



Abbildung 4: Annotation eines Objektes mit zwei 2D-PG

2.3.2 Annotationsgeometrien im LiDAR

Im LiDAR kommen folgende Annotationsgeometrien zum Einsatz:

- Dreidimensionale Quader (engl. „three-dimensional Bounding Box“, Abkürzung: 3D-BB), siehe Abbildung 5
- Dreidimensionale Polylinien (Abkürzung: 3D-PL), siehe Abbildung 6
- Semantische Segmentierungen (Abkürzung: SemSeg), siehe Abbildung 7 und Abbildung 8

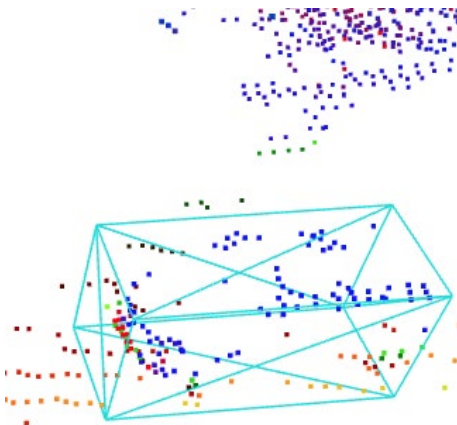


Abbildung 5: Annotation eines Objektes mit einer 3D-BB

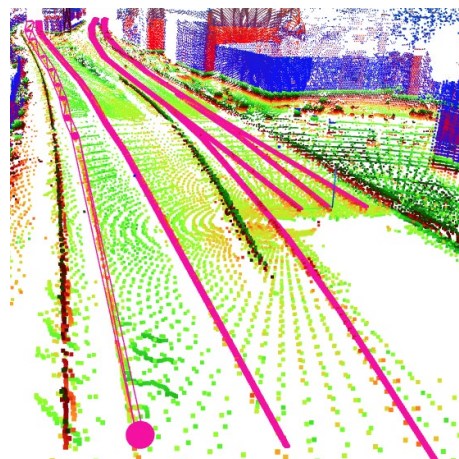


Abbildung 6: Annotation mehrerer Objekte mit 3D-PL

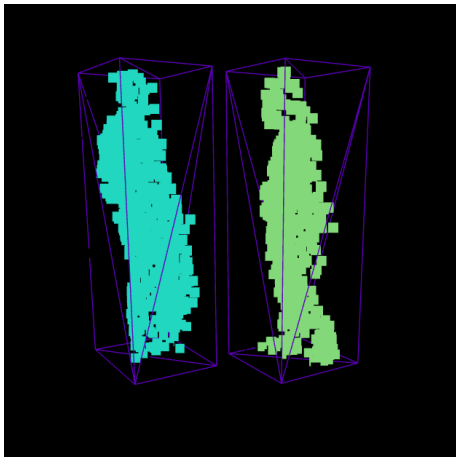


Abbildung 7: Annotation eines Objektes mit einer SemSeg

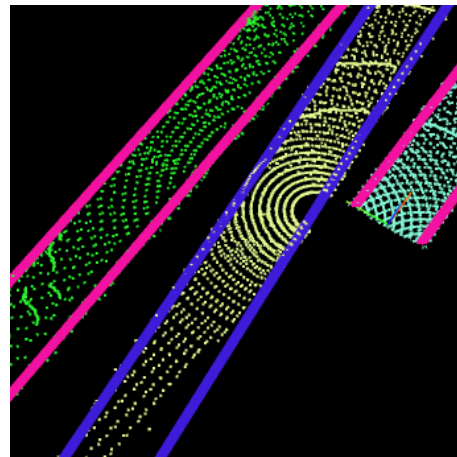


Abbildung 8: Annotation eines Objektes mit einer SemSeg

2.4 Übersicht der Objektklassen

Der folgende Abschnitt gibt einen Überblick über die verschiedenen Objektklassen, die in der Annotationsvorschrift und im Datensatz verwendet werden. Die nachfolgenden Unterabschnitte geben einen visuellen Überblick über die Objektklassen (2.4.1), eine Übersicht der jeweils verwendeten Annotationsgeometrien (2.4.2), eine Zusammenfassung der allgemeine Annotationsregeln (2.4.3) sowie einen Überblick der die Attribute der Objektklassen (2.4.4). Ausführliche Vorgaben zur Annotation finden sich in der Annotationsvorschrift.

2.4.1 Visueller Überblick

Die folgenden Abschnitte geben einen visuellen Überblick über alle im Projekt und in der Annotationsvorschrift berücksichtigten Objektklassen. Die Objektklassen teilen sich dabei in die vier Kategorien auf:

1. Dynamische Objekte
2. Schienenobjekte
3. Statische Objekte
4. Spezielle Objekte

2.4.1.1 Dynamische Objekte

In Abbildung 9 bis Abbildung 20 sind alle Klassen dynamischer Objekte dargestellt, die laut Annotationsvorschrift zu annotieren sind. In den Klammern werden die jeweiligen englischen Bezeichnungen in der Annotationsvorschrift angegeben.



Abbildung 9: Die Objektklasse „Person“ („Person“) umfasst alle Menschen und menschenähnliche Attrappen (engl. „Dummy“) inklusive am Körper getragener Objekte

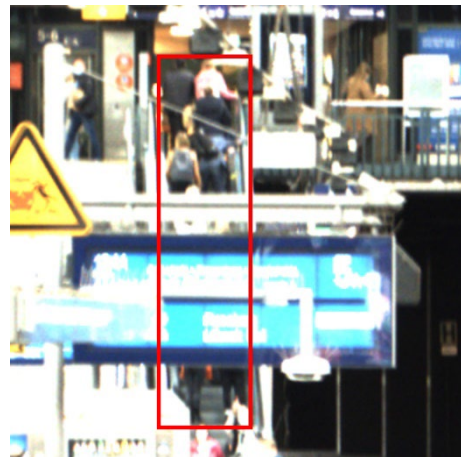


Abbildung 10: Die Objektklasse „Menschenmenge“ („Crowd“) umfasst Gruppen von Menschen, die nur schwer voneinander unterscheidbar sind



Abbildung 11: Die Objektklasse „Zug“ („Train“) umfasst alle Arten von Schienenfahrzeugen.



Abbildung 12: Die Objektklasse „Waggons“ („Wagons“) umfasst alle Arten von gezogenen Objekten im Schienentransport



Abbildung 13: Die Objektklasse „Fahrrad“ („Bicycle“) umfasst alle Arten von unmotorisierten Ein- und Zweirädern



Abbildung 14: Die Objektklasse „Gruppe von Fahrrädern“ („Group of Bicycles“) umfasst Ansammlungen von Fahrrädern, die sich nur schwer getrennt voneinander annotieren lassen

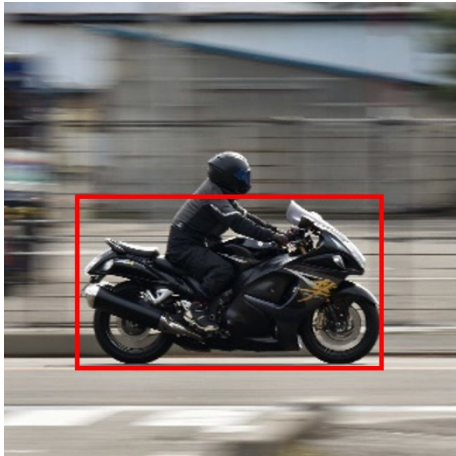


Abbildung 15: Die Objektklasse „Motorrad“ („Motorcycle“) umfasst alle Arten von Motorrädern, Mopeds und deren elektrische Varianten



Abbildung 16: Die Objektklasse „Straßenfahrzeug“ („Road Vehicle“) umfasst alle Arten von motorisierten Automobilen und deren elektrische Varianten



Abbildung 17: Die Objektklasse „Tier“ („Animal“) umfasst alle Tierarten.



Abbildung 18: Die Objektklasse „Gruppe von Tieren“ („Group of Animals“) umfasst Ansammlungen von Tieren, die sich nur schwer getrennt voneinander annotieren lassen



Abbildung 19: Die Objektklasse „Rollstuhl“ („Wheelchair“) umfasst alle Arten von handgetriebenen und elektrischen Rollstühlen (Quelle: pxby666/Pixabay)



Abbildung 20: Die Objektklasse „Hemmschuh“ („Drag Shoe“) umfasst Objekte, die Zugräder an Ort und Stelle halten

2.4.1.2 Schienenobjekte

In Abbildung 21 bis Abbildung 23 sind alle Klassen von Schienenobjekten dargestellt, die laut Annotationsvorschrift zu annotieren sind.



Abbildung 21: Die Objektklasse „Gleis“ („*Track*“) umfasst Schienenpaare, die ein Gleis bilden, sowie das Gleisbett dazwischen

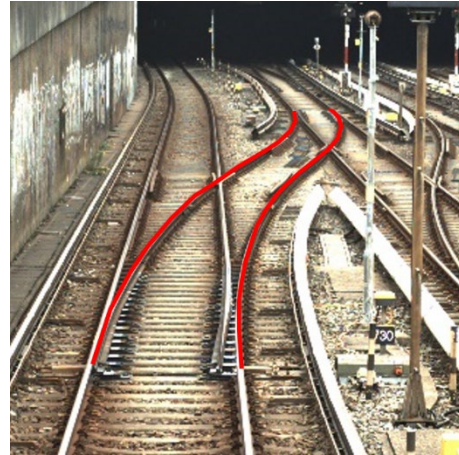


Abbildung 22: Die Objektklasse „Übergang“ („*Transition*“) umfasst kurze Schienenpaare, die zwei Hauptgleise verbinden, sowie das Gleisbett dazwischen



Abbildung 23: Die Objektklasse „Weiche“ („*Switch*“) umfasst Teile der Schienen, die es Zügen ermöglichen, das Gleis zu wechseln

2.4.1.3 Statische Objekte

In Abbildung 24 bis Abbildung 28 sind alle Klassen statischer Objekte dargestellt, die laut Annotationsvorschrift zu annotieren sind.

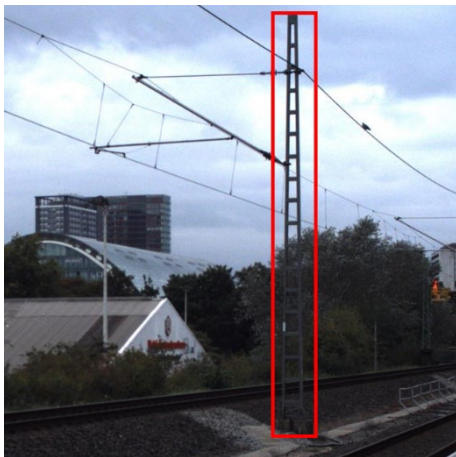


Abbildung 24: Die Objektklasse „Oberleitungsmast“ („*Catenary Pole*“) umfasst alle Masten, die Oberleitung beziehungsweise den Fahrdrat für die Stromversorgung der elektrifizierten Züge tragen



Abbildung 25: Die Objektklasse „Signalmast“ („*Signal Pole*“) umfasst alle Masten, an denen vollbahnrelevante Signale angebracht sind



Abbildung 26: Die Objektklasse „Signal“ („*Signal*“) umfasst alle Arten von vollbahnrelevanten Form- und Lichtsignalen

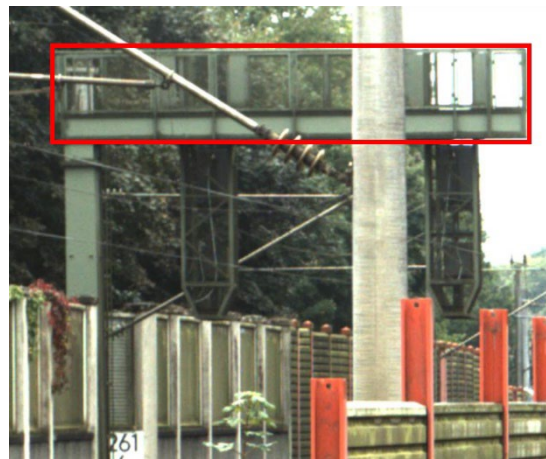


Abbildung 27: Die Objektklasse „Signalbrücke“ („*Signal Bridge*“) umfasst horizontale Strukturen, an denen bahrelevante Signale angebracht sind

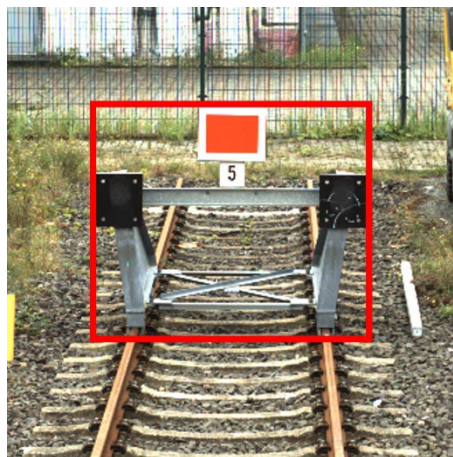


Abbildung 28: Die Objektklasse „Prellbock“ („*Buffer Stop*“) umfasst Objekte, die das Ende eines Gleises markieren

2.4.1.4 Spezielle Objekte

In Abbildung 29 und Abbildung 30 sind alle Objektklassen dargestellt, die laut Annotationsvorschrift zu annotieren sind und nicht in eine der anderen Kategorien passen.

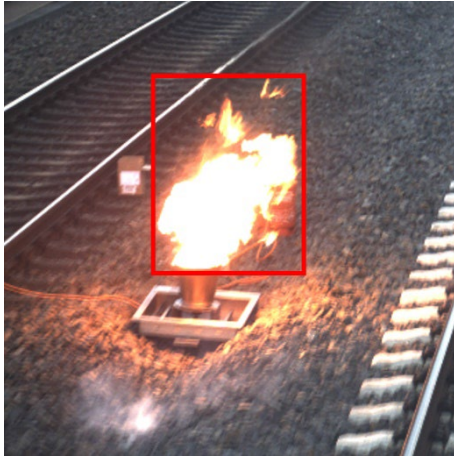


Abbildung 29: Die Objektklasse „Flamme“ („Flame“) umfasst alle Arten von Feuer



Abbildung 30: Die Objektklasse „Rauch“ („Smoke“) umfasst alle Rauchschwaden, die die Sicht verdecken

2.4.2 Zuordnung der Annotationsgeometrien zu den Objektklassen

Jeder Objektklasse können, je nachdem in welcher Sensormodalität das Objekt auftritt, eindeutig eine oder mehrere Annotationsgeometrien (Übersicht siehe Abschnitt 2.3) zugeordnet werden. Diese Zuordnung ist in Tabelle 1 dargestellt. Beispielsweise werden Objekte der Klasse „Person“ in den Sensordaten der Kameras (Farbbild und Infrarot) mit 2D-BBa, im LiDAR mit 2D-BB und SemSeg sowie im RaDAR mit 2D-BBr annotiert. Objekte der Klasse Rauch werden lediglich in den Kameradaten mit 2D-PG annotiert.

2.4.3 Allgemeine Regeln für die Annotation

Die Annotationsvorschrift enthält verschiedene Regeln, die bei der Annotation alle Objektklassen zu beachten sind. An dieser Stelle werden die wichtigsten kurz zusammengefasst:

1. Für jede Annotationsgeometrie existiert eine eindeutige Identifikationsnummer.
2. Jeder Annotationsgeometrie muss eine weitere Identifikationsnummer zur Nachverfolgung (Tracking-ID) zugeordnet sein. Die Tracking-ID wird verwendet, um verschiedene reale Objekte zu unterscheiden. Annotationen, die zu demselben realen Objekt gehören, müssen dieselbe Tracking-ID haben (in den verschiedenen Sensoren und über die gesamte Zeit).
3. Geschlossene Annotationsgeometrien (alle außer 2D-PL und 3D-PL) sind so zu wählen, dass alle Punkte bzw. Pixel, die zu einem Objekt gehören, im Inneren der Geometrie liegen und die Geometrie gleichzeitig so enganliegend wie möglich ist.
4. Annotationsgeometrien sind auf den entsprechenden Sensorbildbereich begrenzt, auch wenn Objekte aus dem Erfassungsbereich des Sensors „herausragen“.
5. Sind Objekte verdeckt, dann wird deren wahre Größe geschätzt und die Annotationsgeometrie so erstellt, dass sie die geschätzte wahre Größe einschließt.

TABELLE 1: ZUORDNUNG DER ANNOTATIONSGEOMETRIEN ZU DEN OBJEKTKLASSEN

Objektklasse	Sensormodalität und Annotationsgeometrie							
	Kamera (Farbbild und Infrarot)				LiDAR			RaDAR
	2D-BBa	2D-BBr	2D-PL	2D-PG	3D-BB	3D-PL	Sem Seg	2D-BBr
Person	●				●		●	●
Menschenmenge		●			●		●	●
Zug				● ¹	●		●	●
Waggons				●	●		●	●
Fahrrad	●				●		●	●
Gruppe von Fahrrädern		●			●		●	●
Motorrad	●				●		●	●
Straßenfahrzeug	●				●		●	●
Tier	●				●		●	●
Gruppe von Tieren		●			●		●	●
Rollstuhl	●				●		●	●
Hemmschuh	●				●		●	
Gleis			● ²			● ²	●	
Übergang			● ²			● ²	●	
Weiche				●	●		●	
Oberleitungsmast	●				●		●	●
Signalmast	●				●		●	●
Signal		●			●		●	
Signalbrücke	●				●		●	●
Prellbock	●				●		●	●
Flamme	●							
Rauch				●				

¹ Ein 2D-PG für die Zugfront und ein 2D-PG für den gesamten Zug² Eine 2D-PL bzw. 3D-PL für jede Schiene

6. In den Kamerabildern müssen Objekte nur annotiert werden, wenn sie mindestens 25 Pixel breit und/oder mindestens 25 Pixel hoch sind. In den LiDAR-Punktwolken müssen die Objekte nur annotiert werden, wenn sie aus mindestens drei Punkten bestehen.
7. In den Kamerabildern müssen alle Annotationsgeometrien eine Genauigkeit von drei Pixeln aufweisen.

2.4.4 Überblick über die Attribute

Im Folgenden werden die Attribute (deutsche Bezeichnung und englische Bezeichnung laut Annotationsvorschrift) inklusive der Attributart und der möglichen Attributausprägungen (deutsche Bezeichnung und englische Bezeichnung laut Annotationsvorschrift), jeweils pro Objektklasse, tabellarisch aufgeführt. In den Fußnoten werden einzelne Attributausprägungen näher erläutert. Attributarten können Einfachauswahl (ein Wert aus der Menge der Attributausprägungen), Mehrfachauswahl (ein oder mehrere Werte aus der Menge der Attributausprägungen) oder Boolean (Wert aus der Menge „wahr“ oder „falsch“) sein.

2.4.4.1 Person

TABELLE 2: ATTRIBUTE DER OBJEKTKLASSE „PERSON“

Attributname	Attributart	Attributausprägungen	
Funktion <i>function</i>	Einfachauswahl	Passagier (Standard)	<i>passenger (default)</i>
		Arbeiterkraft	<i>worker</i>
		Sicherheitsmitarbeiter	<i>securityStaff</i>
		Weitere Mitarbeiter	<i>otherStaff</i>
		Uniformierte Person	<i>uniformedPerson</i>
		Sonstiges	<i>other</i>
		Unbekannt	<i>unknown</i>
Aufsteller <i>isDummy</i>	Boolean	wahr	<i>true</i>
		falsch (Standard)	<i>false (default)</i>
Alter <i>age</i>	Einfachauswahl	Kind ³	<i>child</i>
		Erwachsener (Standard)	<i>adult (default)</i>
Hilfsmittel <i>aid</i>	Einfachauswahl	keins (Standard)	<i>none (default)</i>
		Blindenstock	<i>whiteCane</i>
		Gehhilfe	<i>walkingAid</i>
		Sonstiges	<i>other</i>

³ Personen zählen als Kind, wenn sie jünger als 14 sind oder, wenn das Alter nicht erkennbar ist, kleiner als 1,60 m groß.

Attributname		Attributart	Attributausprägungen	
abgelenkt <i>isDistracted</i>		Boolean	wahr	<i>true</i>
			falsch (Standard)	<i>false (default)</i>
getragene Objekte <i>carrying</i>		Mehrfachauswahl	keins (Standard)	<i>none (default)</i>
			Koffer	<i>suitcase</i>
			Rucksack	<i>backpack</i>
			Regenschirm	<i>umbrella</i>
			Baby	<i>baby</i>
			Sonstiges	<i>other</i>
verbunden mit <i>connectedTo</i>		Mehrfachauswahl	Tracking-IDs der zu einer der aufgeführten Klassen gehörenden Objekte	<i>Tracking-IDs of objects belonging to one of the mentioned classes</i>
Pose <i>pose</i>		Einfachauswahl	stehend (Standard)	<i>upright (default)</i>
			sitzend	<i>sitting</i>
			liegend	<i>lying</i>
			Sonstiges	<i>other</i>
Verdeckung <i>occlusion</i>		Einfachauswahl	0-25 % (Standard)	<i>0-25 % (default)</i>
			25-50 %	<i>25-50 %</i>
			50-75 %	<i>50-75 %</i>
			75-99 %	<i>75-99 %</i>
			100 %	<i>100 %</i>
			undefiniert ⁴	<i>undefined</i>

⁴ Die Verdeckung kann im RaDAR nicht ermittelt werden und ist dort deswegen „undefiniert“

2.4.4.2 Menschenmenge

TABELLE 3: ATTRIBUTE DER OBJEKTKLASSE „MENSCHENMENGE“

Attributname	Attributart	Attributausprägungen	
Größe <i>size</i>	Einfachauswahl	< 25 (Standard)	< 25 (default)
		25-50	25-50
		50-75	50-75
		> 75	> 75
Verdeckung <i>occlusion</i>	Einfachauswahl		0-25 % (default)
		25-50 %	25-50 %
		50-75 %	50-75 %
		75-99 %	75-99 %
		100 %	100 %
		undefiniert	undefined

2.4.4.3 Zug

TABELLE 4: ATTRIBUTE DER OBJEKTKLASSE „ZUG“

Attributname	Attributart	Attributausprägungen	
Typ <i>type</i>	Einfachauswahl	Lokomotive	<i>locomotive</i>
		Intercity	<i>intercity</i>
		Regionalverkehr (Standard)	regional (default)
		Nahverkehr	<i>commuter</i>
		Bauwesen	<i>construction</i>
		Sonstiges ⁵	<i>other</i>
Vorderseite abgeschnitten <i>isTruncated-Front</i>	Boolean	wahr	<i>true</i>
		falsch⁶ (Standard)	false (default)

⁵ „Sonstiges“ bedeutet, dass eine Zuordnung nicht möglich ist.

⁶ Abschneiden kann nur in den Kameras erfolgen, bei den anderen Sensoren ist dieses Attribut immer „falsch“. Das gilt auch für „Hinterseite abgeschnitten“.

Attributname		Attributart	Attributausprägungen	
Hinterseite abgeschnitten	<i>isTruncated-Back</i>	Boolean	wahr	<i>true</i>
			falsch (Standard)	<i>false (default)</i>
Vorderseite	<i>isFront</i>	Boolean	wahr	<i>true</i>
			falsch (Standard)	<i>false (default)</i>
verbunden mit	<i>connectedTo</i>	Mehrfachauswahl	Tracking-IDs der zu einer der genannten Klassen gehörenden Objekte	<i>Tracking-IDs of objects belonging to one of the mentioned classes</i>
Verdeckung	<i>occlusion</i>	Einfachauswahl	0-25 % (Standard)	<i>0-25 % (default)</i>
			25-50 %	<i>25-50 %</i>
			50-75 %	<i>50-75 %</i>
			75-99 %	<i>75-99 %</i>
			100 %	<i>100 %</i>
			undefiniert	<i>undefined</i>

2.4.4.4 Waggon

TABELLE 5: ATTRIBUTE DER OBJEKTKLASSE „WAGGONS“

Attributname		Attributart	Attributausprägungen	
Typ	<i>type</i>	Einfachauswahl	Intercity	<i>intercity</i>
			Regionalverkehr	<i>regional</i>
			Fracht (Standard)	<i>freight (default)</i>
			Bauwesen	<i>construction</i>
			Sonstiges	<i>other</i>
Vorderseite abgeschnitten	<i>isTruncated-Front</i>	Boolean	wahr	<i>true</i>
			falsch (Standard)	<i>false (default)</i>
Hinterseite abgeschnitten	<i>isTruncated-Back</i>	Boolean	wahr	<i>true</i>
			falsch (Standard)	<i>false (default)</i>
verbunden mit	<i>connectedTo</i>	Mehrfachauswahl	Tracking-IDs der zu einer der genannten Klassen gehörenden Objekte	<i>Tracking-IDs of objects belonging to one of the mentioned classes</i>

Attributname		Attributart	Attributausprägungen	
Verdeckung	<i>occlusion</i>	Einfachauswahl	0-25 % (Standard)	0-25 % (default)
			25-50 %	25-50 %
			50-75 %	50-75 %
			75-99 %	75-99 %
			100 %	100 %
			undefiniert	<i>undefined</i>

2.4.4.5 Fahrrad

TABELLE 6: ATTRIBUTE DER OBJEKTKLASSE „FAHRRAD“

Attributname		Attributart	Attributausprägungen	
Zustand	<i>state</i>	Einfachauswahl	geschoben (Standard)	<i>pushed (default)</i>
			gefahren	<i>ridden</i>
			getragen	<i>carried</i>
			abgestellt	<i>steady</i>
verbunden mit	<i>connectedTo</i>	Mehrfachauswahl	Tracking-IDs der zu einer der genannten Klassen gehörenden Objekte	<i>Tracking-IDs of objects belonging to one of the mentioned classes</i>
Verdeckung	<i>occlusion</i>	Einfachauswahl	0-25 % (Standard)	0-25 % (default)
			25-50 %	25-50 %
			50-75 %	50-75 %
			75-99 %	75-99 %
			100 %	100 %
			undefiniert	<i>undefined</i>

2.4.4.6 Gruppe von Fahrrädern

TABELLE 7: ATTRIBUTE DER OBJEKTKLASSE „GRUPPE VON FAHRRÄDERN“

Attributname		Attributart	Attributausprägungen	
Verdeckung	<i>occlusion</i>	Einfachauswahl	0-25 % (Standard)	0-25 % (default)
			25-50 %	25-50 %
			50-75 %	50-75 %
			75-99 %	75-99 %
			100 %	100 %
			undefiniert	<i>undefined</i>

2.4.4.7 Motorrad

TABELLE 8: ATTRIBUTE DER OBJEKTKLASSE „MOTORRAD“

Attributname		Attributart	Attributausprägungen	
verbunden mit	<i>connectedTo</i>	Mehrfachauswahl	Tracking-IDs der zu einer der genannten Klassen gehörenden Objekte ⁷	<i>Tracking-IDs of objects belonging to one of the mentioned classes</i>
Verdeckung	<i>occlusion</i>	Einfachauswahl	0-25 % (Standard)	0-25 % (default)
			25-50 %	25-50 %
			50-75 %	50-75 %
			75-99 %	75-99 %
			100 %	100 %
			undefiniert	<i>undefined</i>

⁷ Wenn das Motorrad mit einer Person oder mehreren Personen verbunden ist, wird dieses Attribut entsprechend gesetzt.

2.4.4.8 Straßenfahrzeug

TABELLE 9: ATTRIBUTE DER OBJEKTKLASSE „STRAßENFAHRZEUG“

Attributname	Attributart	Attributausprägungen	
Typ <i>type</i>	Einfachauswahl	Auto (Standard)	<i>car (default)</i>
		Kleinbus	<i>van</i>
		Lastkraftwagen	<i>truck</i>
		Bus	<i>bus</i>
		Baufahrzeug	<i>constructionVehicle</i>
		Anhänger	<i>trailer</i>
		Sonstiges	<i>other</i>
verbunden mit <i>connectedTo</i>	Mehrfachauswahl	Tracking-IDs der zu einer der genannten Klassen gehörenden Objekte ⁸	<i>Tracking-IDs of objects belonging to one of the mentioned classes</i>
Verdeckung <i>occlusion</i>	Einfachauswahl	0-25 % (Standard)	<i>0-25 % (default)</i>
		25-50 %	<i>25-50 %</i>
		50-75 %	<i>50-75 %</i>
		75-99 %	<i>75-99 %</i>
		100 %	<i>100 %</i>
		undefiniert	<i>undefined</i>

⁸ Wenn ein Straßenfahrzeug mit einem anderen Straßenfahrzeug verbunden ist, wird dieses Attribut entsprechend gesetzt.

2.4.4.9 Tier

TABELLE 10: ATTRIBUTE DER OBJEKTKLASSE „TIER“

Attributname	Attributart	Attributausprägungen	
Größe⁹ <i>size</i>	Einfachauswahl	klein	<i>small</i>
		mittel (Standard)	<i>medium (default)</i>
		groß	<i>large</i>
Spezies <i>species</i>	Einfachauswahl	Hund (Standard)	<i>dog (default)</i>
		Rotwild	<i>deer</i>
		Fuchs	<i>fox</i>
		Kaninchen	<i>rabbit</i>
		Wildschwein	<i>wildboar</i>
		Kuh	<i>cow</i>
		Sonstiges	<i>other animals</i>
Pose <i>pose</i>	Einfachauswahl	stehend (Standard)	<i>upright (default)</i>
		sitzend	<i>sitting</i>
		liegend	<i>lying</i>
		Sonstiges	<i>other</i>
Verdeckung <i>occlusion</i>	Einfachauswahl	0-25 % (Standard)	<i>0-25 % (default)</i>
		25-50 %	<i>25-50 %</i>
		50-75 %	<i>50-75 %</i>
		75-99 %	<i>75-99 %</i>
		100 %	<i>100 %</i>
		undefiniert	<i>undefined</i>

⁹ Um eine Vorstellung von diesem Attribut zu bekommen, beachten Sie bitte die folgenden Beispiele: Hund (klein bis mittel), Hirsch (groß), Fuchs (klein), Kaninchen (klein), Wildschwein (mittel), usw.

2.4.4.10 Gruppe von Tieren

TABELLE 11: ATTRIBUTE DER OBJEKTKLASSE „GRUPPE VON TIEREN“

Attributname	Attributart	Attributausprägungen	
Verdeckung <i>occlusion</i>	Einfachauswahl	0-25 % (Standard)	0-25 % (default)
		25-50 %	25-50 %
		50-75 %	50-75 %
		75-99 %	75-99 %
		100 %	100 %
		undefiniert	<i>undefined</i>

2.4.4.11 Rollstuhl

TABELLE 12: ATTRIBUTE DER OBJEKTKLASSE „ROLLSTUHL“

Attributname	Attributart	Attributausprägungen	
verbunden mit <i>connectedTo</i>	Mehrfachauswahl	Tracking-IDs der zu einer der genannten Klassen gehörenden Objekte ¹⁰	<i>Tracking-IDs of objects belonging to one of the mentioned classes</i>
Verdeckung <i>occlusion</i>	Einfachauswahl	0-25 % (Standard)	0-25 % (default)
		25-50 %	25-50 %
		50-75 %	50-75 %
		75-99 %	75-99 %
		100 %	100 %
		undefiniert	<i>undefined</i>

¹⁰ Wenn dieses Objekt mit einem anderen Objekt verbunden ist, wird dieses Attribut entsprechend gesetzt. Eine Person kann im Rollstuhl sitzen oder eine Person kann den Rollstuhl schieben.

2.4.4.12 Hemmschuh

TABELLE 13: ATTRIBUTE DER OBJEKTKLASSE „HEMMSCHUH“

Attributname		Attributart	Attributausprägungen	
auf Gleis	<i>onTrack</i>	Einfachauswahl	-2	-2
			-1	-1
			0 (Standard)	0 (default)
			1	1
			2	2
			Sonstiges ¹¹	<i>other</i>
			keins ¹²	<i>none</i>
Gleisseite	<i>railSide</i>	Einfachauswahl	Schiene links (Standard)	<i>leftRail (default)</i>
			Schiene rechts	<i>rightRail</i>
			Keins	<i>none</i>
Verdeckung	<i>occlusion</i>	Einfachauswahl	0-25% (Standard)	<i>0-25 % (default)</i>
			25-50 %	<i>25-50 %</i>
			50-75 %	<i>50-75 %</i>
			75-99 %	<i>75-99 %</i>
			100 %	<i>100 %</i>
			undefiniert	<i>undefined</i>

¹¹ Liegt der Hemmschuh außerhalb des Tracking-ID-Bereichs, sollte er als „Sonstiges“ bezeichnet werden.

¹² Befindet sich der Hemmschuh nicht auf einer Spur, so ist er mit „Keins“ zu kennzeichnen.

2.4.4.13 Gleis

TABELLE 14: ATTRIBUTE DER OBJEKTKLASSE „GLEIS“

Attributname	Attributart	Attributausprägungen	
Start <i>start</i>	Einfachauswahl	0¹³ (Standard)	0 (default)
		+1 ¹⁴	+1
		+2 ¹⁵	+2
		-1 ¹⁶	-1
		-2 ¹⁷	-2
Gleisseite¹⁸ <i>railSide</i>	Einfachauswahl	Schiene links (Standard)	leftRail (default)
		Schiene rechts	rightRail
Verdeckung <i>occlusion</i>	Einfachauswahl	0-25 % (Standard)	0-25 % (default)
		25-50 %	25-50 %
		50-75 %	50-75 %
		75-99 %	75-99 %
		100 % ¹⁹	100 %
		undefiniert	undefined

¹³ Ego-Gleis: Das Gleis, auf dem der Testzug fährt.¹⁴ Gleis auf der rechten Seite neben Ego-Gleis.¹⁵ Gleis auf der rechten Seite neben dem Gleis +1.¹⁶ Gleis auf der linken Seite neben dem Ego-Gleis.¹⁷ Gleis auf der linken Seite neben dem Gleis -1.¹⁸ Bestimmt, ob die Polylinie zur linken oder rechten Schiene des jeweiligen Gleisbetts gehört. Es soll nur die äußere Oberkante der Schiene gekennzeichnet werden.¹⁹ Beide Schienen zusammenzählen als 100 %. Wenn eine Schiene vollständig verdeckt ist und die andere vollständig sichtbar ist, werden 50 % als Verdeckung ausgewählt.

2.4.4.14 Übergang

TABELLE 15: ATTRIBUTE DER OBJEKTKLASSE „ÜBERGANG“

Attributsname		Art der Auswahl	Attributausprägungen	
Gleisseite	<i>railSide</i>	Einfachauswahl	Schiene links (Standard)	<i>leftRail (default)</i>
			Schiene rechts	<i>rightRail</i>
Startspur	<i>startTrack</i>	Einfachauswahl	-2	-2
			-1	-1
			0 (Standard)	<i>0 (default)</i>
			+1	+1
			+2	+2
Endspur	<i>endTrack</i>	Einfachauswahl	-2	-2
			-1	-1
			0 (Standard)	<i>0 (default)</i>
			+1	+1
			+2	+2
Verdeckung	<i>occlusion</i>	Einfachauswahl	0-25 % (Standard)	<i>0-25 % (default)</i>
			25-50 %	25-50 %
			50-75 %	50-75%
			75-99 %	75-99 %
			100 %	100 %
			undefiniert	<i>undefined</i>

2.4.4.15 Weiche

TABELLE 16: ATTRIBUTE DER OBJEKTKLASSE „WEICHE“

Attributname		Attributart	Attributausprägungen	
auf Gleis	<i>onTrack</i>	Einfachauswahl	-2	-2
			-1	-1
			0 (Standard)	0 (default)
			+1	+1
			+2	+2
Zustand	<i>state</i>	Einfachauswahl	Links	<i>left</i>
			Mitte (Standard)	center (default)
			Rechts	<i>right</i>
			Unbekannt	<i>unknown</i>
Verdeckung	<i>occlusion</i>	Einfachauswahl	0-25 % (Standard)	0-25 % (default)
			25-50 %	25-50 %
			50-75 %	50-75 %
			75-99 %	75-99 %
			100 %	100 %
			undefiniert	<i>undefined</i>

2.4.4.16 Oberleitungsmast

TABELLE 17: ATTRIBUTE DER OBJEKTKLASSE „OBERLEITUNGSMAST“

Attributname		Attributart	Attributausprägungen	
Typ	<i>type</i>	Einfachauswahl	massiv (Standard)	<i>solid (default)</i>
			strukturiert	<i>structured</i>
oben abgeschnitten	<i>isTruncatedTop</i>	Boolean	wahr	<i>true</i>
			falsch²⁰ (Standard)	<i>false (default)</i>
unten abgeschnitten	<i>isTruncatedBottom</i>	Boolean	wahr	<i>true</i>
			falsch (Standard)	<i>false (default)</i>
Verdeckung	<i>occlusion</i>	Einfachauswahl	0-25 % (Standard)	<i>0-25 % (default)</i>
			25-50 %	<i>25-50 %</i>
			50-75 %	<i>50-75 %</i>
			75-99 %	<i>75-99 %</i>
			100 %	<i>100 %</i>
			undefiniert	<i>undefined</i>

2.4.4.17 Signalmast

TABELLE 18: ATTRIBUTE DER OBJEKTKLASSE „SIGNALMAST“

Attributname		Attributart	Attributausprägungen	
Typ	<i>type</i>	Einfachauswahl	massiv	<i>solid</i>
			strukturiert (Standard)	<i>structured (default)</i>
oben abgeschnitten	<i>isTruncatedTop</i>	Einfachauswahl	wahr	<i>true</i>
			falsch (Standard)	<i>false (default)</i>
unten abgeschnitten	<i>isTruncatedBottom</i>	Einfachauswahl	wahr	<i>true</i>
			falsch (Standard)	<i>false (default)</i>

²⁰ Abschneiden kann nur in den Kameras erfolgen, bei den anderen Sensoren ist dieses Attribut immer „falsch“. Das gilt auch für „unten abgeschnitten“.

Attributname		Attributart	Attributausprägungen	
verbunden mit	<i>connectedTo</i>	Mehrfachauswahl	Tracking-IDs der zu einer der genannten Klassen gehörenden Objekte	<i>Tracking-IDs of objects belonging to one of the mentioned classes</i>
Verdeckung	<i>occlusion</i>	Einfachauswahl	0-25 % (Standard)	0-25 % (default)
			25-50 %	25-50 %
			50-75 %	50-75 %
			75-99 %	75-99 %
			100 %	100 %
			undefiniert	<i>undefined</i>

2.4.4.18 Signal

TABELLE 19: ATTRIBUTE DER OBJEKTKLASSE „SIGNAL“

Attributname		Attributart	Attributausprägungen	
Typ	<i>type</i>	Einfachauswahl	Licht (Standard)	<i>light (default)</i>
			Form	<i>shape</i>
Signalausrichtung	<i>signalFace</i>	Einfachauswahl	Vorderseite (Standard)	<i>front (default)</i>
			Hinterseite	<i>back</i>
			unbekannt ²¹	<i>unknown</i>
verbunden mit	<i>connectedTo</i>	Mehrfachauswahl	Tracking-IDs der zu einer der genannten Klassen gehörenden Objekte	<i>Tracking-IDs of objects belonging to one of the mentioned classes</i>
Verdeckung	<i>occlusion</i>	Einfachauswahl	0-25 % (Standard)	0-25 % (default)
			25-50 %	25-50 %
			50-75 %	50-75 %
			75-99 %	75-99 %
			100 %	100 %
			undefiniert	<i>undefined</i>

²¹ „Unbekannt“ ist zu verwenden, wenn eine Entscheidung nicht möglich ist.

2.4.4.19 Signalbrücke

TABELLE 20: ATTRIBUTE DER OBJEKTKLASSE „SIGNALBRÜCKE“

Attributname		Attributart	Attributausprägungen	
Typ	<i>type</i>	Einfachauswahl	fest	<i>solid</i>
			strukturiert (Standard)	<i>structured (default)</i>
links abgeschnitten	<i>isTruncatedLeft</i>	Boolean	wahr	<i>True</i>
			falsch²² (Standard)	<i>false (default)</i>
rechts abgeschnitten	<i>isTruncatedRight</i>	Boolean	wahr	<i>true</i>
			falsch (Standard)	<i>false (default)</i>
verbunden mit	<i>connectedTo</i>	Mehrfachauswahl	Tracking-IDs der zu einer der genannten Klassen gehörenden Objekte	<i>Tracking-IDs of objects belonging to one of the mentioned classes</i>
Verdeckung	<i>occlusion</i>	Einfachauswahl	0-25 % (Standard)	<i>0-25 % (default)</i>
			25-50 %	<i>25-50 %</i>
			50-75 %	<i>50-75 %</i>
			75-99 %	<i>75-99 %</i>
			100 %	<i>100 %</i>
			undefiniert	<i>undefined</i>

²² Abschnitten kann nur in den Kameras erfolgen, bei den anderen Sensoren ist dieses Attribut immer „falsch“. Das gilt auch für „Rechts abgeschnitten?“.

2.4.4.20 Prellbock

TABELLE 21: ATTRIBUTE DER OBJEKTKLASSE „PRELLBOCK“

Attributname	Attributart	Attributausprägungen	
auf Gleis <i>onTrack</i>	Einfachauswahl	-2	-2
		-1	-1
		0 (Standard)	0 (default)
		+1	+1
		+2	+2
		Sonstiges ²³	<i>other</i>
		Keins ²⁴	<i>none</i>
Verdeckung <i>occlusion</i>	Einfachauswahl	0-25 % (Standard)	0-25 % (default)
		25-50 %	25-50 %
		50-75 %	50-75 %
		75-99 %	75-99 %
		100 %	100 %
		undefiniert	<i>undefined</i>

2.4.4.21 Flamme

TABELLE 22: ATTRIBUTE DER OBJEKTKLASSE „FLAMME“

Attributname	Attributart	Attributausprägungen	
Größe ²⁵ <i>size</i>	Einfachauswahl	groß	<i>big</i>
		klein (Standard)	<i>small (default)</i>
Verdeckung <i>occlusion</i>	Einfachauswahl	0-25 % (Standard)	0-25 % (default)
		25-50 %	25-50 %
		50-75 %	50-75 %
		75-99 %	75-99 %
		100 %	100 %
		undefiniert	<i>undefined</i>

²³ Liegt der Prellbock außerhalb des Tracking-ID-Bereichs, sollte er als „Sonstiges“ bezeichnet werden.²⁴ Befindet sich der Prellbock nicht auf einer Spur, so ist er mit „Keins“ zu kennzeichnen.²⁵ Brände mit einer Höhe von weniger als 4 m gelten als klein, Brände mit einer Höhe von mehr als 4 m gelten als groß.

2.4.4.22 Rauch

TABELLE 23: ATTRIBUTE DER OBJEKTKLASSE „RAUCH“

Attributname	Attributart	Attributausprägungen	
Größe²⁶ <i>size</i>	Einfachauswahl	groß (Standard)	<i>big (default)</i>
		klein	<i>small</i>
Farbe <i>color</i>	Einfachauswahl	weiß	<i>white</i>
		grau (Standard)	<i>gray (default)</i>
		Schwarz	<i>black</i>
		Sonstiges	<i>other</i>
Verdeckung <i>occlusion</i>	Einfachauswahl	0-25 % (Standard)	<i>0-25 % (default)</i>
		25-50 %	<i>25-50 %</i>
		50-75 %	<i>50-75 %</i>
		75-99 %	<i>75-99 %</i>
		100 %	<i>100 %</i>
		undefiniert	<i>undefined</i>

2.5 Aufbau der Sensordaten des Referenzdatensatzes

Die Sensordaten, die im Rahmen des Projektes annotiert wurden, weisen über die in Abschnitt 2.2 definierten Eigenschaften hinaus weitere Merkmale auf: Sie wurden als Videos mit einer Frequenz von 10 Hz zeitsynchron aufgezeichnet und die entsprechenden Sensoren sind zueinander kalibriert. Details zu der eingesetzten Sensorik finden sich in Abschnitt 3.1. Dieser Sachverhalt findet bei der Beschreibung des Annotations- und Qualitätssicherungskonzept, welches im folgenden Abschnitt beschrieben wird, Beachtung. Für Datensätze, die diese Merkmale nicht erfüllen, muss dieses entsprechend angepasst werden: Es können keine Projektionen der Annotationsgeometrien zwischen Sensoren vorgenommen werden (siehe Unterabschnitt 2.6.2), Annotationen können nicht interpoliert werden (siehe Unterabschnitte 2.6.4 bis 2.6.7) und die automatischen Konsistenzüberprüfungen „Überprüfung der Plausibilität der Tracking-ID“ (Unterabschnitt 2.7.3) sowie „Überprüfung der zeitlichen und sensorübergreifenden Konsistenz der Attribute“ (Unterabschnitt 2.7.5) entfallen.

Sensordaten verschiedener Sensoren, die im Rahmen der Zeitsynchronisierung demselben Zeitpunkt zugeordnet werden können, bilden zusammen einen sogenannten Multisensor-Frame (M-Frame), wie in Abbildung 31 dargestellt. Zwischen den Sensordaten innerhalb eines M-Frames können über die Kalibrierparameter räumliche Beziehungen hergestellt werden, insbesondere können Annotationen aus dem LiDAR in die Bildsensoren (Kameras und RaDAR) projiziert werden. Mehrere in unmittelbarer zeitlicher Abfolge aufgenommene M-Frames bilden eine Sequenz.

²⁶ Rauchwolken mit einer Höhe von weniger als 4 m gelten als klein, Rauchwolken mit einer Höhe von mehr als 4 m gelten als groß.

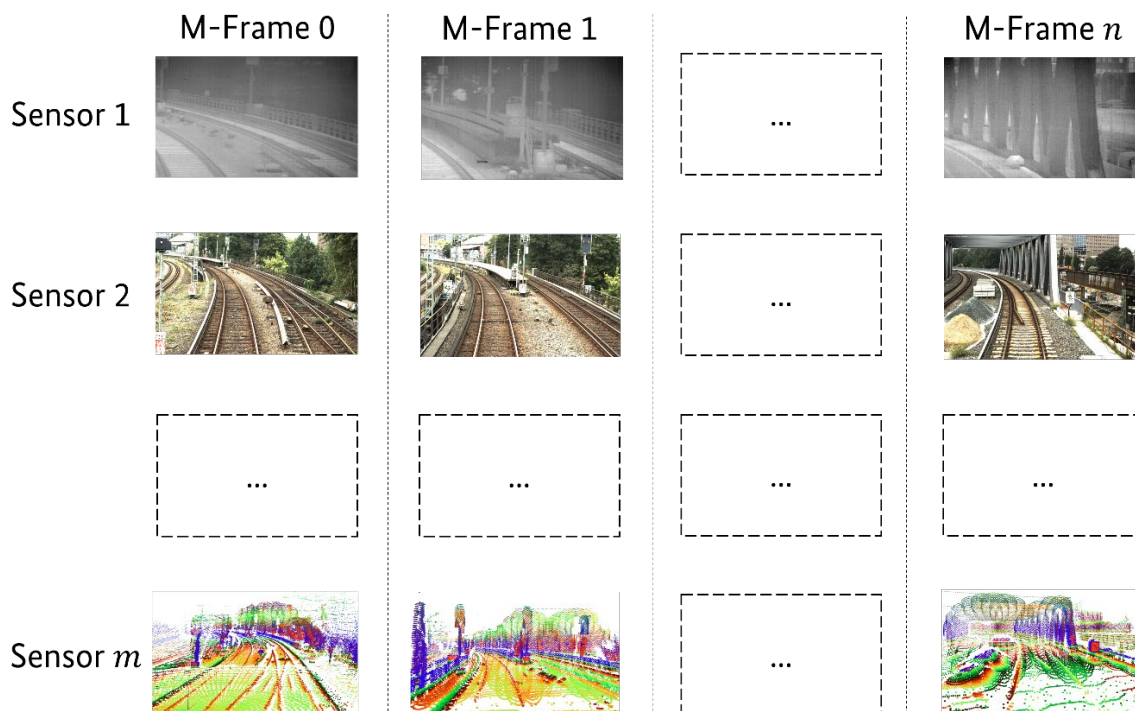


Abbildung 31: Aufbau einer Sequenz

2.6 Annotations- und Qualitätssicherungskonzept

Bei der Erstellung und der Qualitätssicherung von Annotationen sind verschiedene Personen in unterschiedlichen Rollen beteiligt. Die Rollen werden in Tabelle 24 erläutert.

TABELLE 24: ROLLENBESCHREIBUNG

Rolle	Beschreibung der Tätigkeiten
Annotierende Person (Ap)	Die Aufgabe einer AP besteht im korrekten Setzen der Annotationen (Geometrien und Attribute).
Annotationsprozess- Koordinierende Person (Akp)	Die AKP koordiniert den gesamten Annotationsprozess von der Zuordnung einzelner Sequenzen zu den APs bis zum Abschluss des Annotationsvorgangs nach der finalen Korrektur.
Intern prüfende Person (Ipp)	Eine IPP überprüft/reviewt auf Auftragnehmerseite die Fehlerfreiheit der erstellten Annotationen.
Extern prüfende Person (Epp)	Eine EPP überprüft/reviewt auf Auftraggeberseite stichprobenartig die Korrektheit der erstellten Annotationen. Der Schwerpunkt liegt darauf, ob die Annotationsvorschrift aus Auftraggebersicht grundsätzlich richtig interpretiert und angewendet wird.
Interne Projektleitung (Ipl)	Die IPL ist auf Auftragnehmerseite für das Gesamtprojekt verantwortlich.
Externe Projektleitung (Epl)	Die EPL ist auf Auftraggeberseite für das Gesamtprojekt verantwortlich.

Ziel des Annotations- und Qualitätssicherungsprozesses ist die Erstellung von Annotationen entsprechend den Vorgaben der Annotationsvorschrift. Die präzise Einhaltung der Vorgaben und damit eine hohe Qualität der Annotationen ist von hoher Bedeutung für die spätere Nutzung der Annotationen im Rahmen des Trainings von ML-Verfahren. Im Folgenden wird beschrieben, wie dies im Rahmen des Projektes erreicht wurde. Die Schritte werden jeweils für alle im Rahmen des Projektes zu annotierenden Sequenzen durchgeführt. Die Beschreibung der Prozessschritte gliedert sich dabei jeweils in eine allgemeine Beschreibung und, falls erforderlich, Hinweise zur konkreten Umsetzung im Rahmen des Projektes.

Abbildung 32 enthält eine Übersicht über den Annotations- und Qualitätssicherungsprozess. Die einzelnen Arbeitsschritte sind in den folgenden Unterabschnitten erläutert.

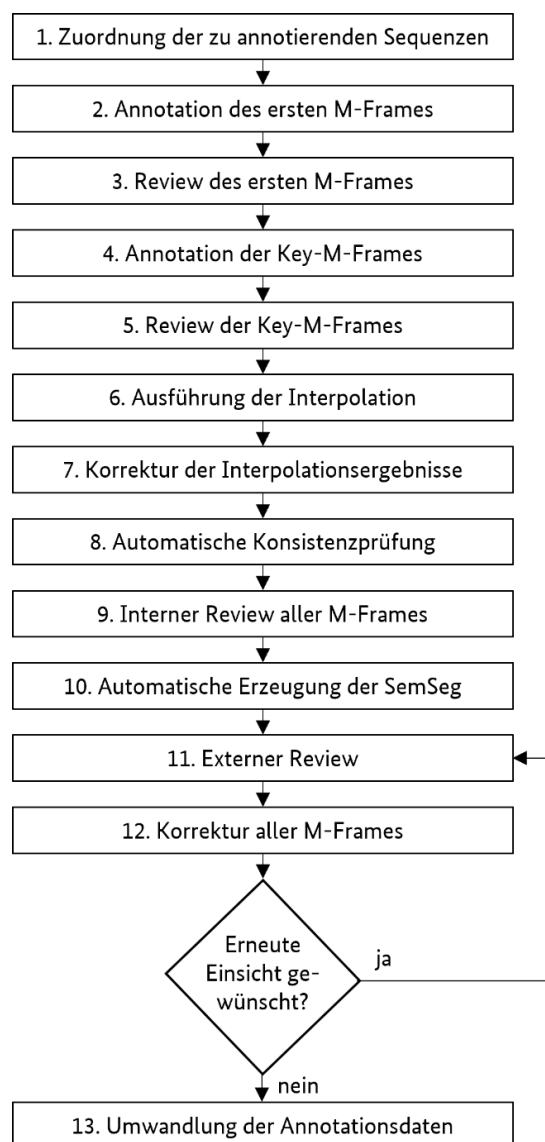


Abbildung 32: Übersicht des Annotations- und Qualitätssicherungsprozesses

2.6.1 Zuordnung der zu annotierenden Sequenzen

Basierend auf einer im Vorfeld zu erstellenden Priorisierungsliste weist die AKP die zu annotierenden Sequenzen einer AP und einer IPP zu. Die AP wird informiert, dass sie mit dem Annotieren der ihr zugeordneten Sequenz beginnen kann.

Im Projekt wurde dem Auftragnehmer vom Auftraggeber eine solche Priorisierungsliste übergeben. Um Fehl- bzw. Doppelarbeiten zu vermeiden, wird ein rollenbasiertes Rechtesystem eingesetzt, welches den Zugriff auf Daten reguliert. Damit wird beispielsweise ausgeschlossen, dass eine AP an einer Sequenz arbeitet, die ihr nicht zugewiesen wurde. Auch eine Abweichung von der Priorisierungsliste oder eine doppelte Bearbeitung durch eine zweite AP werden dadurch verhindert.

2.6.2 Annotation des ersten M-Frames

Die AP beginnt mit der Annotation des ersten M-Frames. Im Referenzdatensatz sind die Sensordaten zueinander kalibriert, so dass folgende Annotationsreihenfolge angewendet werden konnte:

1. Die AP annotierte alle in den LiDAR-Daten sichtbaren Objekte mit Geometrien und Attributen entsprechend der Annotationsvorschrift.
2. Die AP löste die 3D-zu-2D-Projektion aus, welche automatisch Annotationen in den Kamera-/RaDAR-Daten (auf der Basis der vorhandenen Kalibrierungsparameter) mit der jeweils passenden Tracking-ID erzeugte.
3. Die Geometrien und sensorspezifischen Attribute (siehe Abschnitt 2.4) der projizierten Annotationen wurden korrigiert, sodass diese den Anforderungen der Annotationsvorschrift genügten.
4. Objekte, die nicht im LiDAR sichtbar waren, wurden in einem Kamera-Bild annotiert und anschließend wurden per 2D-zu-3D-Projektion Annotationen mit passender Tracking-ID im LiDAR angelegt, die durch die AP entsprechend ihres Szenenverständnis im LiDAR positioniert wurden.
5. Für diese im LiDAR eigentlich nicht sichtbaren Objekte wurde der Wert des Attributs „Verdeckung“ auf 100 % gesetzt. Ein erneutes Auslösen der 3D-zu-2D-Projektion erzeugte nun in allen Kamera-Bildern entsprechende Annotationen mit passender Tracking-ID. Für diese erzeugten Annotationen werden die Geometrien sowie die sensorspezifischen Attribute korrigiert.

Alle Annotationen werden durch die AP überprüft und ggf. korrigiert. Ist dies abgeschlossen, informiert die AP die AKP, dass der Annotationsvorgang des ersten M-Frames beendet ist.

Es können während des Annotationsprozesses Fälle auftreten, bei denen unklar ist, wie eine Geometrie oder ein Attribut zu setzen sind. Tritt ein solcher Grenzfall auf, wird die IPL informiert, welcher die Behandlung des Grenzfalls mit der EPL abklärt. Das Ergebnis dieser Klärung wird von der IPL über die AKP an die APs und IPPs weitergegeben, die EPL informiert die EPPs. Die beschlossene Behandlung von Grenzfällen wird in die Annotationsvorschrift aufgenommen.

Im Rahmen des Projektes wurde ein vom Auftragnehmer selbst entwickeltes Programm, das „Fusion-Systems Labeling Tool“, verwendet. Dieses erlaubt eine effektive und effiziente Durchführung der beschriebenen Annotationsschritte. Für die Besprechung der Grenzfälle wurde das wöchentliche Treffen zwischen Auftragnehmer und Auftraggeber genutzt.

2.6.3 Review des ersten M-Frames

Die AKP ordnet die Sequenzen, in denen der erste M-Frame fertig annotiert wurde, einer IPP zu. Die IPP prüft sämtliche Annotationsobjekte in allen Sensoren und berichtigt, falls erforderlich, Geometrien und Attribute. Anschließend informiert sie die AKP und die AP wird angewiesen, mit dem Annotieren der ihr zugewiesenen Sequenz fortfahren kann. Zusätzlich erhält die AP ein individuelles Feedback von der IPP über die vorgenommenen Anpassungen, um sich kontinuierlich verbessern zu können.

2.6.4 Annotation der Key-M-Frames

Die starke zeitliche Korrelation aufeinanderfolgender M-Frames erlaubt es, dass nicht alle M-Frames manuell annotiert werden müssen. Es reicht aus, nur jeden n -ten M-Frame manuell zu annotieren und die restlichen Annotationen dann mittels Interpolation zu erzeugen und gegebenenfalls manuell zu korrigieren. Manuell annotierte M-Frames werden deswegen als Key-M-Frames bezeichnet.

Bei der Annotation der Key-M-Frames wird iterativ vorgegangen, wobei der bereits annotierte erste M-Frame gleichzeitig der erste Key-M-Frame ist. Die Annotationen des ersten Key-M-Frames werden von der AP in den zweiten Key-M-Frame kopiert und bezüglich der Geometrie und Attribute angepasst. Objekte, die im neuen Key-M-Frame nicht mehr auftauchen, werden gelöscht (wenn sie in der gesamten Sequenz nicht mehr auftauchen) bzw. auf „Verdeckung“ = 100 % gesetzt (wenn sie nochmal in der Sequenz auftauchen). Objekte, die erstmalig auftauchen, werden analog zu Schritt 2.6.2 neu annotiert.

Nachdem die AP sich vergewissert hat, dass alle Annotationen fehlerfrei sind, kopiert sie die Annotationen des zweiten Key-M-Frames in den dritten und wiederholt die im vorigen Absatz beschriebenen Annotationsschritte. Dies wiederholt sich nun, bis alle Key-M-Frames annotiert sind. Ist die AP mit dem Annotieren des letzten Key-M-Frames fertig, informiert sie die AKP, dass die Annotation der Key-M-Frames beendet ist.

Im Projekt erwies es sich als sinnvoll, jeden fünften M-Frame zu annotieren. Somit wurde jeder fünfte M-Frame zu einem Key-M-Frame. Der Kopiervorgang wurde mittels des FusionSystems Labeling Tools durchgeführt.

2.6.5 Review der Key-M-Frames

Die AKP informiert die IPP, welche auch bereits den ersten Key-M-Frame geprüft hat, dass sie nun die Key-M-Frames der entsprechenden Sequenz reviewen kann. Die IPP betrachtet für sämtliche Key-M-Frames sämtliche Annotationsobjekte in allen Sensoren und berichtigt, falls erforderlich, Geometrien und Attribute. Weiterhin erstellt sie, falls nötig, fehlende Annotationen oder löscht Annotationen, die nicht der Annotationsvorschrift entsprechen. Sobald dieser Prozess abgeschlossen ist, informiert die IPP die AKP. Zusätzlich erhält die AP ein individuelles Feedback von der IPP. Sollte sich aus den Feedbackgesprächen ergeben, dass eine Spezifizierung der Annotationsvorschrift erforderlich ist, so wird diese durchgeführt.

2.6.6 Ausführen der Interpolation

Wie bereits in Unterabschnitt 2.6.4 ausgeführt, können auf Basis der annotierten Key-M-Frames Annotationen in den M-Frames, die keine Key-M-Frames sind, erzeugt werden. Über die Tracking-ID können dabei eindeutig Korrespondenzen zwischen zwei Annotationsobjekten in zwei aufeinanderfolgenden Key-M-Frames hergestellt werden. Die Positionen der Ankerpunkte, welche die jeweilige Annotations-

geometrie definieren (Eckpunkte für 2D-BBa/2D-BBr/3D-BB, Punkte der 2D-PL/2D-PG/3D-PL), können für Zeitpunkte zwischen diesen beiden aufeinanderfolgenden Key-M-Frames interpoliert werden. Die Interpolation erfolgt für alle Ankerpunkte eines Objektes jeweils getrennt für die unterschiedlichen Koordinaten.

Neben den Annotationsgeometrien müssen auch die Attribute der durch die Interpolation entstandenen Annotationen gesetzt werden. Hierfür wird im Zuge der Interpolation überprüft, ob es zeitliche Änderungen der Attributausprägungen von Objekten gibt. Falls ja, wird die Attributausprägung, die sich auf die kleinere M-Frame-Nummer bezieht, übernommen und eine entsprechende objektbezogene Warnung erzeugt, auf die die IPP im nächsten Schritt eingeht. Falls nicht, werden die (identischen) Attributausprägungen für die interpolierten Objekte übernommen.

Die Interpolation wird von der AP ausgelöst, der die Sequenz zugeordnet ist. Anschließend fährt diese mit dem nächsten Schritt, der Korrektur der Interpolationsergebnisse, fort.

Im Projekt wurden für die Interpolation kubische Splines²⁷ verwendet. Die Interpolation wurde im FusionSystems Labeling Tool integriert.

2.6.7 Korrektur der Interpolationsergebnisse

Die AP sichtet die Ergebnisse der Interpolation und korrigiert diese falls erforderlich, insbesondere mit Hinblick auf die geltenden Präzisionsanforderungen an die Annotationen. Sollte sich die Ausprägung eines Attributs geändert haben, wird die AP durch die objektbezogene Warnung darauf hingewiesen und korrigiert die Ausprägung, falls erforderlich.

2.6.8 Automatische Konsistenzprüfungen

Zur Unterstützung des Reviews können verschiedene Verfahren zur automatischen Konsistenzprüfung angewendet werden, die nach der Korrektur der Interpolationsergebnisse von der AP ausgelöst werden. Im Projekt wurden folgende Verfahren angewendet:

1. Überprüfung von Lage und Anzahl der 2D-BBa
2. Überprüfung der Plausibilität der Abmessungen der 3D-BB
3. Überprüfung der Plausibilität der Tracking-IDs
4. Überprüfung der Plausibilität von Attributen der Kategorie „Abgeschnitten“
5. Überprüfung der zeitlichen und sensorübergreifenden Konsistenz der Attribute

Im Abschnitt 2.7 sind diese Verfahren detailliert beschrieben. Die Ergebnisse der Konsistenzüberprüfungen werden als objektbezogene Warnungen in geeigneter Form übergeben. Nach der Durchführung dieser Konsistenzüberprüfungen wird die AKP über die erfolgten Konsistenzüberprüfungen informiert.

Die oben genannten automatischen Konsistenzüberprüfungen sind im FusionSystems Labeling Tool integriert. Die erzeugten objektbezogenen Warnungen werden in Form von kurzen Texten in den Dateien, welche auch die Annotationsdaten enthalten, hinterlegt.

²⁷ Kubische Splines sind über folgende Eigenschaften definiert: Zwischen zwei Stützstellen wird eine ganzrationale Funktion dritten Grades verwendet, um den Funktionswert (hier Koordinaten der Ankerpunkte) in Abhängigkeit der Zeit zu beschreiben. An den Stützstellen müssen benachbarte Funktionen sowohl in Funktionswert als auch in der ersten und zweiten Ableitung übereinstimmen. Damit die Funktionen eindeutig bestimmt werden können, muss noch eine Randbedingung definiert werden. Hier wurde die sogenannte „freie Randbedingung“ verwendet, dass die zweite Ableitung an der ersten und an der letzten Stützstelle Null ist. Aus den geforderten Bedingungen lassen sich die Funktionsterme eindeutig berechnen.

2.6.9 Interner Review aller M-Frames

Die AKP weist nach Abschluss der automatischen Konsistenzprüfung die der Sequenz zugeordnete IPP an, alle M-Frames erneut zu reviewen. Dazu werden zuerst die objektbezogenen Warnungen der Konsistenzüberprüfung ausgewertet und daraus resultierende Anpassungen der Annotationsgeometrien und Attribute vorgenommen. Im FusionSystems Labeling Tool wird die IPP automatisch durch die Ergebnisse der automatischen Konsistenzprüfung geleitet. Danach sieht sie alle Annotationen erneut durch und korrigiert diese, falls erforderlich. Nach Abschluss der Korrekturen wird die AKP informiert, dass der interne Review der M-Frames abgeschlossen ist. Die AP erhält erneut ein individuelles Feedback von der IPP.

2.6.10 Automatische Berechnung der SemSeg

Die AKP löst die Berechnung der SemSeg aus, welche in das FusionSystems Labeling Tool integriert wurde. Für 3D-BB werden alle (dreidimensionalen) Punkte berechnet, die innerhalb der 3D-BB liegen und ihre zugehörigen Indizes als Liste ausgegeben. Für Gleise und Übergänge erfolgt die Berechnung anhand der Lage der 3D-PL, welche die linke und die rechte Schiene beschreiben, und eines vordefinierten Querschnittes, welcher in Abbildung 33 dargestellt ist.

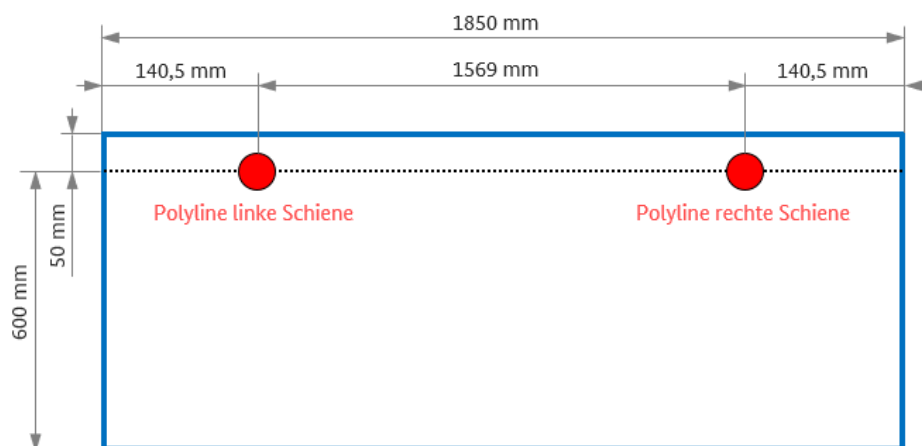


Abbildung 33: Schienenquerschnitt für die SemSeg

2.6.11 Externer Review

Nach Abschluss der automatischen Berechnung der SemSeg informiert die IPL die EPL, dass die Sequenz extern geprüft werden kann. Die AKP stellt die annotierte Sequenz den EPPs zur Verfügung. Die EPPs prüfen stichprobenartig die Richtigkeit der Annotationen. Sollten sie Mängel feststellen, vermerken sie diese in geeigneter Form. Die EPP vermerkt weiterhin, wenn er erneute Einsicht in die dann korrigierten Annotationen wünscht. Nach Abschluss des externen Reviews werden IPL und EPL entsprechend informiert.

Im FusionSystems Labeling Tool können Annotationsmängel direkt im Programm festgehalten werden, indem die EPP für die betreffende Annotation einen Text-Kommentar erstellt und/oder die Geometrien sowie Attribute direkt ändert. Zusätzlich wird ein tabellarisches Reviewprotokoll erstellt.

2.6.12 Korrektur aller M-Frames

Die IPL informiert die AKP, sobald der externe Review einer Sequenz abgeschlossen ist. Daraufhin weist die AKP die der Sequenz zugeordnete IPP an, die Ergebnisse des externen Reviews in die Annotationen einzuarbeiten. Sollte die EPP eine erneute Einsicht wünschen, wird Schritt 2.6.11 wiederholt. Ergeben sich aus dem externen Review systematische Fehler, werden diese über die AKP an die APs und die IPPs weitergegeben. Falls erforderlich, wird die Annotationsvorschrift entsprechend ergänzt.

2.6.13 Umwandlung der Annotationsdaten

Sobald alle Annotationen finalisiert sind, erfolgt eine Konvertierung der Annotationsdaten in das vom Auftraggeber geforderte „ASAM OpenLABEL“-Format [7]. Dieses stellt einen offenen Standard für Annotationsdaten dar, welches von Organisationen aus Industrie und Forschung gemeinsam entwickelt wurde. In der Annotationsvorschrift wird das „ASAM OpenLABEL“-Format detaillierter beschrieben. Mit der Konvertierung ist der Annotations- und Qualitätssicherungsprozess abgeschlossen.

Im Projekt ergab sich die Notwendigkeit der Verwendung eines internen Datenformats, das nicht identisch mit dem „ASAM OpenLABEL“-Format ist. Zwischen internen Datenformat und dem (finalen) „ASAM OpenLABEL“-Format existieren folgende Unterschiede:

- Im internen Datenformat existieren neben den Annotationsgeometrien und Attributen auch Angaben, welche für die Projektionen (2D-zu-3D bzw. 3D-zu-2D), die Interpolation, die automatische Konsistenzüberprüfung und den Reviewprozess wichtig sind. Derartige Angaben sind im „ASAM OpenLABEL“-Format nicht enthalten.
- Im internen Datenformat existiert eine Annotationsdatei pro M-Frame, im „ASAM OpenLABEL“-Format eine pro Sequenz.

Im Zuge der Konvertierung erfolgte aus Datenschutzgründen auch eine Verfremdung des Zeitstempels der Sensordaten.

2.7 Beschreibung der automatischen Konsistenzprüfungen

2.7.1 Überprüfung von Lage und Anzahl der 2D-BBa

Um die Konsistenz der erstellten 2D-BBa zu prüfen, werden über einen Objektdetektor Vergleichswerte erzeugt. Diese werden genutzt, um zu überprüfen, ob (1) die Anzahl der manuell erstellten 2D-BBa der Anzahl der vom Objektdetektor erkannten 2D-BBa entspricht und ob (2) erstellte und detektierte 2D-BBa nah beieinander liegen. Sollte eines der beiden Kriterien nicht zutreffen, ist dies ein Indiz dafür, dass die Annotationen fehlerbehaftet sind. Das Konzept ist in Abbildung 34 schematisch dargestellt. Die manuell erstellten Annotationen (2D-BBa in Rot, Abbildung 34a) für die Objektklasse „Person“ werden mit den Ergebnissen des Objektdetektors (2D-BBa in Blau, Abbildung 34b) verglichen (Abbildung 34c). Der Objektdetektor wurde unter Verwendung von manuell erstellten Annotationen der Key-Frames (2D-BBa der 6 Farbbildkameras) mittels Maschinellen Lernens antrainiert. Die Anwendung dieses Detektors auf alle M-Frames der Sequenz liefert klassenspezifische 2D-BBa. Als Objektdetektor wurde YOLOv4 [8] verwendet.

Für den Vergleich der manuell erstellten Annotationen mit den Ergebnissen des Objektdetektors wird über die ungarische Methode [9] eine Assoziation der 2D-BBa erstellt, sodass nahegelegene 2D-BBa einander zugeordnet werden. Die ungarische Methode beschreibt ein Lösungsverfahren für Zuordnungsprobleme, bei denen jedem Element einer Menge (hier manuell erstellte 2D-BBa) ein Element einer anderen Menge (hier durch Objektdetektor erkannte 2D-BBa) zugeordnet werden soll. Dabei kann jeder Zuordnung ein Skalar zugeordnet werden (hier das Verhältnis von Schnittmenge der beiden 2D-BBa zu ihrer Vereinigungsmenge). Das Ergebnis der ungarischen Methode ist eine Zuordnung, die die Summe der resultierenden Skalare minimiert.

Sollte eine Assoziation einer vom Detektor erstellten 2D-BBa zu einer händischen Annotation oder umgekehrt nicht möglich sein oder zwei assoziierte 2D-BBa zu unterschiedlich²⁸ sein, wird eine entsprechende Warnung erzeugt (im Projekt umgesetzt durch eine separate Annotationsdatei mit entsprechenden Informationen zum Beispiel mit Vorschlägen zur Lage von 2D-BBa). In Abbildung 34c ist zum Beispiel zu erkennen, dass sowohl die manuell erstellte 2D-BBa b_1 einem Detektorergebnis zugeordnet werden kann (nämlich b_3) als auch b_2 (nämlich b_5). Für das Detektorergebnis b_4 existiert keine solche zuordenbare manuell erstellte 2D-BBa, was ein (in diesem Fall zutreffendes) Indiz dafür ist, dass die AP beim Annotieren ein Objekt übersehen hat.

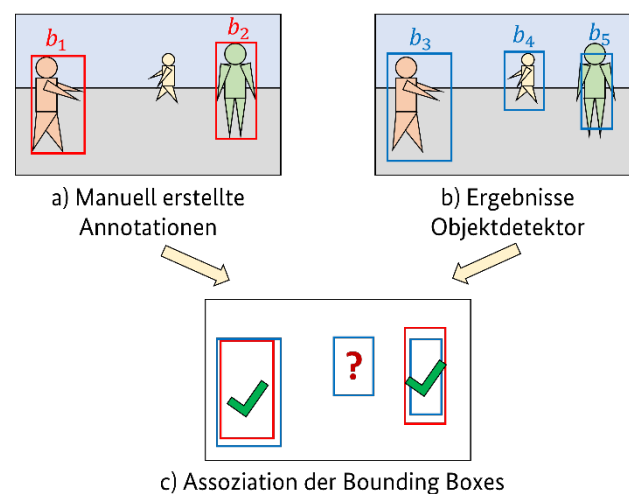


Abbildung 34: Assoziation von manuell erstellten Annotationen und den Ergebnissen des Objektdetektors

2.7.2 Überprüfung der Plausibilität der Abmessungen der 3D-BB

Bei dieser Konsistenzüberprüfung wird die Plausibilität der Abmessungen der 3D-BB überprüft. Beispielsweise sollten Objekte der Klasse Person eine Größe von 2,20 m nicht überschreiten. Je nach Objektklasse können neben einer maximalen und minimalen Höhe auch Bedingungen für die Breite und/oder die Tiefe der Objekte definiert und geprüft werden. Die zu überprüfenden Minimal- und Maximalwerte für Höhe, Breite und Tiefe sind in Abschnitt 8.2 aufgeführt.

²⁸ Als Maß für den Unterschied zwischen zwei 2D-BBa diene hier die Summe aus den Beträgen der Differenz von (1) den x-Werten der beiden linken oberen Eckpunkte, (2) den y-Werten der beiden linken oberen Eckpunkte, (3) den x-Werten der beiden rechten unteren Eckpunkte und (4) den y-Werten der beiden rechten unteren Eckpunkte. Überschreitet dieser Unterschied einen Schwellwert, hier 50 Pixel, dann gelten die beiden 2D-BBa als „zu unterschiedlich“.

2.7.3 Überprüfung der Plausibilität der Tracking-IDs

Um sicherzustellen, dass die vergebenen Tracking-IDs korrekt sind, wird deren Konsistenz im zeitlichen Verlauf überprüft. Das dahinterstehende Konzept ist in Abbildung 35 dargestellt. In Abbildung 35a) befinden sich die 2D-BBa-Annotationen mit der entsprechenden Tracking-ID (in Rot) für einen Frame, in Abbildung 35b) die 2D-BBa-Annotationen mit Tracking-ID für den Nachfolger-Frame (in Blau). Eine Assoziation über die ungarische Methode (siehe Unterabschnitt 2.7.1) erlaubt es, 2D-BBas einander zuzuordnen, die sich in räumlicher Nähe zueinander befinden, wie in Abbildung 35c) dargestellt. In der Abbildung können 2D-BBa b_1 und b_4 , b_2 und b_5 sowie b_3 und b_6 einander zugeordnet werden. Anschließend kann überprüft werden, ob zugeordnete 2D-BBas über die gleiche Tracking-ID verfügen. Sollte dieses nicht der Fall sein, wie bei b_2 und b_5 , ist dies ein Indiz dafür, dass eine Tracking-ID falsch vergeben wurde.

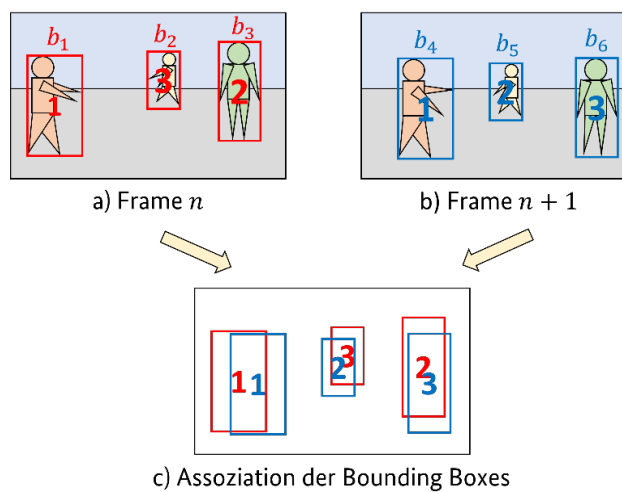


Abbildung 35: Assoziation von Annotationen eines Frames mit den Annotationen des Folgeframes

2.7.4 Überprüfung der Plausibilität von Attributen der Kategorie „Abgeschnitten“

Für einige Objektklassen existieren Attribute wie „oben abgeschnitten“ („isTruncatedTop“), „unten abgeschnitten“ („isTruncatedBottom“), „links abgeschnitten“ („isTruncatedLeft“) und „rechts abgeschnitten“ („isTruncatedRight“). Diese Attribute geben an, ob ein Objekt vom Bildbereich „abgeschnitten“ wird. Anhand der Parameter der 2D-BBa (Koordinaten des Mittelpunkts, Abmessungen) kann überprüft werden, ob diese Attribute korrekt gesetzt wurden. Das Konzept dahinter ist in Abbildung 36 dargestellt. Das Koordinatensystem des Bildes ist wie folgt definiert: Der Koordinatenursprung liegt in der oberen linken Ecke, die positive x -Achse verläuft von links nach rechts, die positive y -Achse von oben nach unten, der Pixel in der oberen linken Ecke hat die Koordinaten $(0; 0)$. Die Abmessungen des Bildes in Pixel werden mit L_x und L_y bezeichnet. 2D-BBa werden durch die Koordinaten des Mittelpunktes $M(x_M; y_M)$ und ihre Breite Δx sowie Höhe Δy beschrieben. Dann müssen folgende Kriterien erfüllt sein:

- Linker Bildrand: Sollte die 2D-BBa an den linken Rand stoßen ($x_M - \frac{\Delta x}{2} \leq 0$), dann muss das Attribut „isTruncatedLeft“ den Wert „wahr“ haben. Andernfalls ($x_M - \frac{\Delta x}{2} > 0$) muss dieses Attribut entsprechend den Wert „falsch“ haben.

- Rechter Bildrand: Sollte die 2D-BBa an den rechten Rand stoßen ($x_M + \frac{\Delta x}{2} \geq L_x$), dann muss das Attribut „isTruncatedRight“ den Wert „wahr“ haben. Andernfalls ($x_M + \frac{\Delta x}{2} < L_x$), dann muss dieses Attribut entsprechend den Wert „falsch“ haben.
- Oberer Bildrand: Sollte die 2D-BBa an den oberen Rand stoßen ($y_M - \frac{\Delta y}{2} \leq 0$), dann muss das Attribut „isTruncatedTop“ den Wert „wahr“ haben. Andernfalls ($y_M - \frac{\Delta y}{2} > 0$) muss dieses Attribut entsprechend den Wert „falsch“ haben.
- Unterer Bildrand: Sollte die 2D-BBa an den unteren Rand stoßen ($y_M + \frac{\Delta y}{2} \geq L_y$), dann muss das Attribut „isTruncatedBottom“ den Wert „wahr“ haben. Andernfalls ($y_M + \frac{\Delta y}{2} < L_y$) muss dieses Attribut entsprechend den Wert „falsch“ haben.

Ist ein Kriterium nicht erfüllt, dann ist das ein Indiz dafür, dass ein „truncated“-Attribut falsch gesetzt wurde.

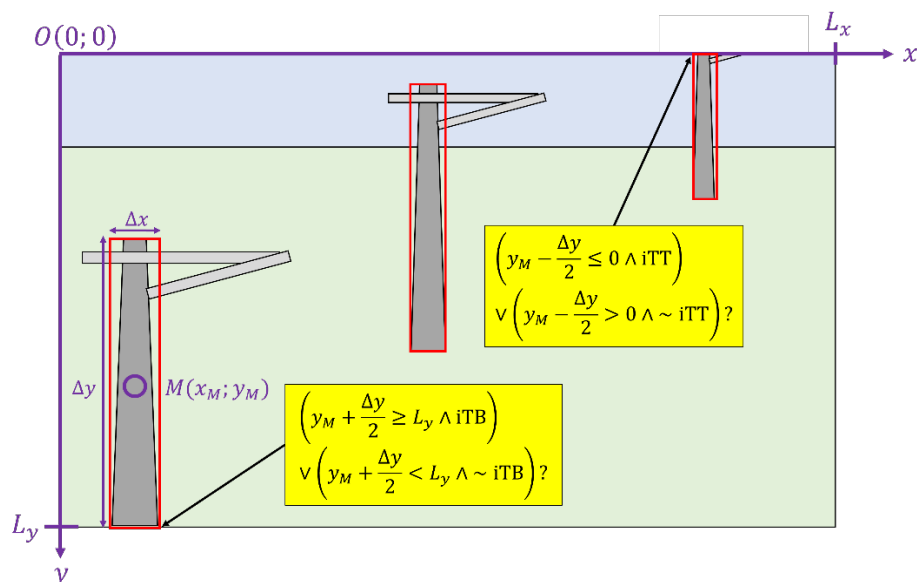


Abbildung 36: Untersuchung der „Abgeschnitten“-Attribute auf Plausibilität (iTT = „isTruncatedTop“, iTB = „isTruncatedBottom“)

2.7.5 Überprüfung der zeitlichen und sensorübergreifenden Konsistenz der Attribute

Nicht alle Attribute weisen die gleichen Charakteristika auf, wenn man ihren zeitlichen Verlauf oder ihre Ausprägungen über die verschiedenen Sensoren betrachtet. So kann sich zum Beispiel die Spezies eines Tiers nicht von einem Frame zum anderen oder von Sensor zu Sensor ändern. Die Größe des Feuers kann sich zwar mit der Zeit ändern, allerdings ist diese zu einem bestimmten Zeitpunkt für alle Sensoren gleich. Die Verdeckung eines Objektes schließlich kann sich sowohl mit der Zeit verändern als auch für verschiedene Sensoren unterschiedlich sein.

In Bezug auf die Zeit gibt es folglich zwei verschiedene Kategorien von Attributen: Solche, die sich mit der Zeit nicht ändern können (Frame-übergreifend) und solche, die sich von Frame zu Frame ändern können (Frame-spezifisch). Gleiches gilt für die verschiedenen Sensoren: Es gibt Attribute, die zu einem Zeitpunkt in verschiedenen Sensoren unterschiedlich ausgeprägt sind (Sensor-spezifisch) und solche, die in allen Sensoren gleich sein müssen (Sensor-übergreifend). Betrachtet man beide Kategorien zugleich,

ergeben sich vier unterschiedliche Kategorien: Frame-übergreifend und Sensor-übergreifend, Frame-übergreifend und Sensor-spezifisch, Frame-spezifisch und Sensor-übergreifend, Frame-spezifisch und Sensor-spezifisch. Jedes Attribut kann eindeutig einer dieser Kategorien zugeordnet werden. Die Zuordnung ist in der Konsistenzmatrix für Attribute (Tabelle 25) abgebildet.

TABELLE 25: KONSISTENZMATRIX FÜR ATTRIBUTE

	Frame-Spezifisch	Frame-Übergreifend
Sensor-spezifisch	Verdeckung (alle Klassen) Vorderseite abgeschnitten (Zug, Wagons) Hinterseite abgeschnitten (Zug, Wagons) oben abgeschnitten (Oberleitungsmast, Signalmast) unten abgeschnitten (Oberleitungsmast, Signalmast) links abgeschnitten (Signalbrücke) rechts abgeschnitten (Signalbrücke)	Existiert hier nicht
Sensor-übergreifend	abgelenkt (Person) Hilfsmittel (Person) Farbe (Rauch) Funktion (Person) Pose (Person, Tier) getragene Objekte (Person) verbunden mit (Person, Zug, Wagons, Fahrrad, Motorrad, Straßenfahrzeug, Rollstuhl, Signalmast, Signal, Signalbrücke) Zustand (Fahrrad, Weiche) Startspur (Übergang) Endspur (Übergang) auf Gleis (Hemmschuh, Weiche, Prellbock) Signalausrichtung (Signal) Größe (Menschenmenge, Flamme, Rauch)	Alter (Person) Aufsteller (Person) Vorderseite (Zug) Typ (Zug, Wagons, Straßenfahrzeug, Oberleitungsmast, Signalmast, Signal, Signalbrücke) Spezies (Tier) Größe (Tier) Gleisseite (Hemmschuh, Gleis, Übergang)

2.8 Annotation einer Nulltranche

Um frühzeitig ein einheitliches Verständnis hinsichtlich der ausgearbeiteten Annotationsvorschrift herzustellen sowie etwaige offene oder zu unspezifisch formulierte Anforderungen in dieser Vorschrift zu finden, wurde zunächst eine Nulltranche annotiert. Diese bestand aus 35 Sequenzen mit jeweils zehn M-Frames, die aus verschiedenen Teilen des Gesamtdatensatzes ausgewählt wurden. Anhand der Nulltranche wurde auch der Annotations- und Qualitätssicherungsprozess getestet.

3 Aufbereitung des Datensatzes

Dieses Kapitel beschreibt die Projektergebnisse, die im Rahmen des Arbeitspaketes 2 „Aufbereitung des Datensatzes“ entstanden sind. Nach einer kurzen Darstellung der wesentlichen Informationen zum annotierten Datensatz (Abschnitt 3.1) wird die Annotation der Daten im Allgemeinen und in einigen Spezialfällen erläutert (Abschnitt 3.2). Abschnitt 3.3 gibt einen Überblick über die im Rahmen des Projektes erstellten Annotationen. Abschnitt 3.4 beschreibt die Vorgehensweise zur Anonymisierung der Farbbildkameradaten.

Der annotierte Datensatz kann unter [2] abgerufen werden. Eine detaillierte Beschreibung des Exportformats und der Datenstruktur der Annotationen findet sich in der Annotationsvorschrift, welche unter diesem [Link](#) abgerufen werden kann.

3.1 Informationen zum Datensatz

Der im Rahmen des Projektes annotierte Sensordatensatz wurde während mehrerer Datenerhebungsfahrten durch die DB Netz AG aufgezeichnet. Die Details der eingesetzten Sensoren können Tabelle 26 und Tabelle 27 entnommen werden. Die LiDAR-Daten wurden bereits vor Projektbeginn zu einer Gesamtpunktwolke fusioniert. Zusätzlich zu den optischen Sensordaten wurden Positionsdaten mittels entsprechender Sensorik aufgezeichnet (Tabelle 28).

TABELLE 26: TECHNISCHE DATEN DER KAMERAS

Farbbildkameras	Infrarotkameras
<ul style="list-style-type: none">- Daten: Farbbilder (8 bit)- Aufnahmefrequenz: 10 Hz- Blickfeld: vorn geradeaus, vorn links, vorn rechts	<ul style="list-style-type: none">- Sensor: Teledyne Calibir DXM640- Daten: Graustufenbilder (8 bit)- Aufnahmefrequenz: 10 Hz- Auflösung: 640 x 480 Pixel- Anzahl: 3- Blickfeld: vorn geradeaus, vorn links, vorn rechts
<u>5 MP Kameras</u> <ul style="list-style-type: none">- Sensor: Teledyne GenieNano C2420- Auflösung: 2464 x 1600 Pixel- Anzahl: 3	
<u>12 MP Kameras</u> <ul style="list-style-type: none">- Sensor: Teledyne GenieNano 5GigE C4040- Auflösung: 4000 x 2500 Pixel- Anzahl: 3	

TABELLE 27: TECHNISCHE DATEN DER LIDARE UND DES RADARS

LiDARe	RaDAR
<ul style="list-style-type: none"> - Daten: 3D Punktwolken - Aufnahmefrequenz: 10 Hz 	<ul style="list-style-type: none"> - Sensor: Navtech CIR204/H - Daten: Graustufenbild (8 bit) - kartesische Vogelperspektive - Auflösung: ca. 1,2 MP - Aufnahmefrequenz: 4 Hz - Anzahl: 1
<u>Long-Range LiDAR:</u> <ul style="list-style-type: none"> - Sensor: Livox Tele-15 - Abtaste: 50.000 – 84.000 Punkte pro Frame - Anzahl: 3 	
<u>Mid-Range LiDAR:</u> <ul style="list-style-type: none"> - Sensor: HesaiTech Pandar64 - Abtaste: 60.000 – 115.000 Punkte pro Frame - Anzahl: 1 	
<u>Short-Range LiDAR:</u> <ul style="list-style-type: none"> - Sensor: Waymo Honeycomb - Abtaste: 20.000 – 40.000 Punkte pro Frame - Anzahl: 2 	

TABELLE 28: TECHNISCHE DATEN DER POSITIONS- UND BESCHLEUNIGUNGSSENSOREN

Positions- und Bewegungssensoren
<ul style="list-style-type: none"> - Sensor: NovAtel PwrPAk7D-E1 - Daten: <ul style="list-style-type: none"> • lineare und rotatorische Beschleunigung • Länge und Breite in WGS84 - Aufnahmefrequenz: 100/10 Hz

Der Datensatz gliedert sich in verschiedene Sequenzen, die bei Befahrungen verschiedener Gleisabschnitte in Deutschland unter verschiedenen Umweltbedingungen erhoben wurden. Die Sequenzen gliedern sich wiederum in Teilsequenzen, die zwischen einer und zehn Sekunden lang sind. Die Sensoren wurden zueinander kalibriert. Die Daten dieser einzelnen Sensordatenframes wurden parallel aufgenommen und auf eine einheitliche Erhebungsfrequenz von 10 Hz (bis auf das RaDAR mit 4 Hz) synchronisiert. Der Zusammenschluss dieser dreizehn Sensordatenframes wird als ein M-Frame bezeichnet. Durch die niedrigere Abtastfrequenz des RaDARs von 4 Hz enthält der M-Frame im Mittel für jedes zweite RaDAR-Bild ein Platzhalterbild. Ein exemplarischer M-Frame ist in Abbildung 37 dargestellt.

Die im Datensatz enthaltenen Sequenzen sind überblicksweise in Tabelle 28 dargestellt. Eine detailliertere Beschreibung der Sequenzen findet sich in Anhang 8.1.

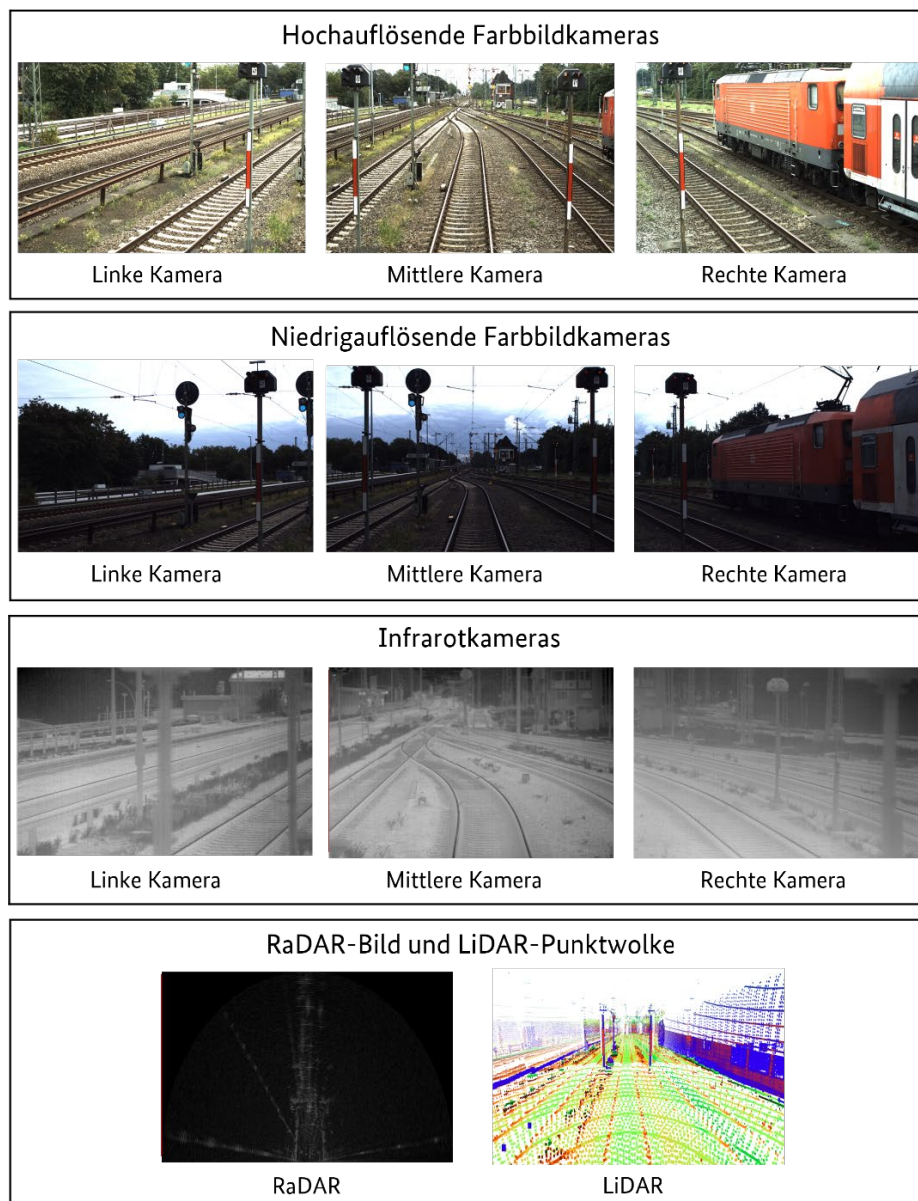


Abbildung 37: Verschiedene Sensorframes bilden einen M-Frame

TABELLE 29: SEQUENZÜBERSICHT

Nr.	Sequenzname	Teilsequenzname	Anzahl M-Frames
1	1_calibration	1_calibration_1.1	10
		1_calibration_1.2	100
2	2_station_berliner_tor	2_station_berliner_tor_2.1	10
3	3_fire_site	3_fire_site_3.1	100
		3_fire_site_3.2	10
		3_fire_site_3.3	10
		3_fire_site_3.4	10

Nr.	Sequenzname	Teilsequenzname	Anzahl M-Frames
4	4_station_pedestrian_bridge	4_station_pedestrian_bridge_4.1	10
		4_station_pedestrian_bridge_4.2	10
		4_station_pedestrian_bridge_4.3	100
		4_station_pedestrian_bridge_4.4	100
		4_station_pedestrian_bridge_4.5	76
5	5_station_bergedorf	5_station_bergedorf_5.1	100
		5_station_bergedorf_5.2	10
6	6_station_klein_flottbek	6_station_klein_flottbek_6.1	10
		6_station_klein_flottbek_6.2	100
7	7_approach_underground_station	7_approach_underground_station_7.1	10
		7_approach_underground_station_7.2	10
		7_approach_underground_station_7.3	10
8	8_station_altona	8_station_altona_8.1	100
		8_station_altona_8.2	100
		8_station_altona_8.3	10
9	9_station_alte_woehr	9_station_alte_woehr_9.1	10
		9_station_alte_woehr_9.2	10
		9_station_alte_woehr_9.3	10
		9_station_alte_woehr_9.4	10
		9_station_alte_woehr_9.5	10
		9_station_alte_woehr_9.6	10
		9_station_alte_woehr_9.7	10
10	10_station_suelldorf	10_station_suelldorf_10.1	10
11	11_main_station	11_main_station_11.1	10
12	12_vegetation_steady	12_vegetation_steady_12.1	98
13	13_station_ohlsdorf	13_station_ohlsdorf_13.1	10
14	14_signals_station	14_signals_station_14.1	10
		14_signals_station_14.2	10
		14_signals_station_14.3	10
15	15_construction_vehicle	15_construction_vehicle_15.1	10
16	16_under_bridge	16_under_bridge_16.1	10
17	17_signal_bridge	17_signal_bridge_17.1	10

Nr.	Sequenzname	Teilsequenzname	Anzahl M-Frames
18	18_vegetation_switch	18_vegetation_switch_18.1	10
19	19_vegetation_curve	19_vegetation_curve_19.1	10
20	20_vegetation_squirrel	20_vegetation_squirrel_20.1	10
21	21_station_wedel	21_station_wedel_21.1	100
		21_station_wedel_21.2	100
		21_station_wedel_21.3	40

3.2 Durchführung der Annotation

Die Annotationen wurden entsprechend den Anforderungen der Annotationsvorschrift und gemäß dem in Abschnitt 2.5 beschriebenen Prozess erstellt. Wie in Abschnitt 2.8 ausgeführt, wurde zuerst eine Nulltranche annotiert und vom Auftraggeber abgenommen. Danach erfolgte die Annotation des restlichen Datensatzes. In Abbildung 38 sind die erstellten Annotationen für den in Abbildung 37 dargestellten M-Frame abgebildet.

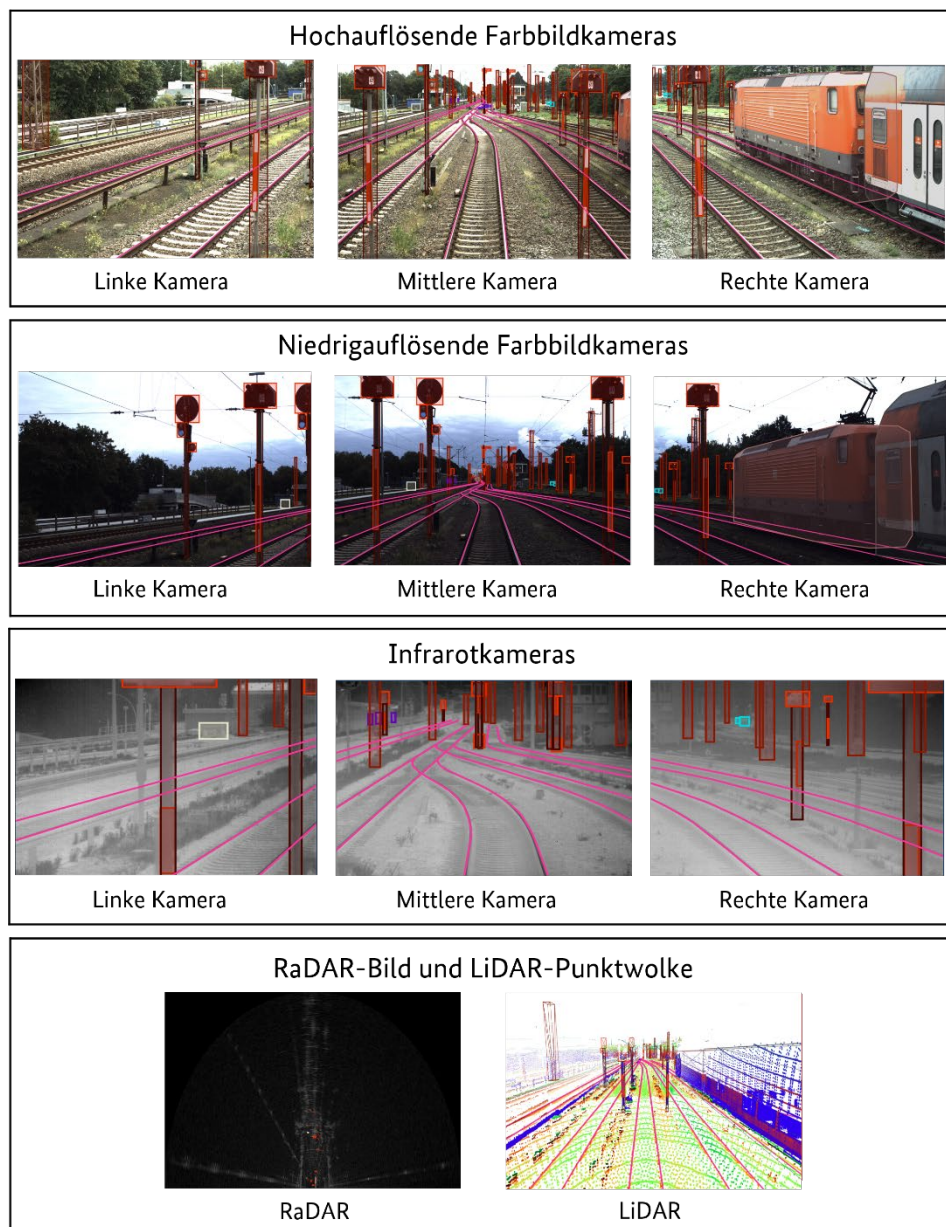


Abbildung 38: Annotation eines M-Frames

Einige Objekt- und Erhebungskonstellationen stellten besondere Herausforderungen bei der Annotation dar, die nachfolgend beschrieben werden.

Besonders **komplexe Objekte** erfordern sehr viel Annotationszeit und eine hohe Konzentration der AP. Beispielsweise weisen Doppelkreuzweichen (siehe Abbildung 39) einen komplexen Aufbau aus zwei Gleisen, mehreren Übergängen und vier Weichen auf. Wenn Doppelkreuzweichen, wie in diesem Beispiel, nur seitlich zu sehen sind, ist die Übersicht über die einzelnen Schienenverläufe und Bestandteile zusätzlich erschwert.



Abbildung 39: Doppelkreuzweiche aus seitlicher Ansicht mit begonnener Annotation

Durch **Überbelichtung** (siehe Abbildung 40) kann das Schätzen der Größe eines Objektes und/oder der Ausprägung der objektspezifischen Attribute erschwert werden. In diesem Beispiel ist mehr als die Hälfte eines Oberleitungsmastes (links im Bild) durch Überbelichtung nicht erkennbar. Die Schätzung der Größe muss in diesem Fall auf Basis umliegender ähnlicher Objekte geschehen oder, wie in diesem Beispiel, anhand kleiner, noch erkennbarer Teile des Mastkorpus.



Abbildung 40: Überbelichteter Oberleitungsmast (ganz links im Bild)

Rauch (siehe Abbildung 41) erschwert die Sicht auf dahinterliegende Objekte, besonders in den niedrigauflösenden Farbbildern. Die Herausforderung besteht darin, die genaue Größe und Position der Objekte zu annotieren.

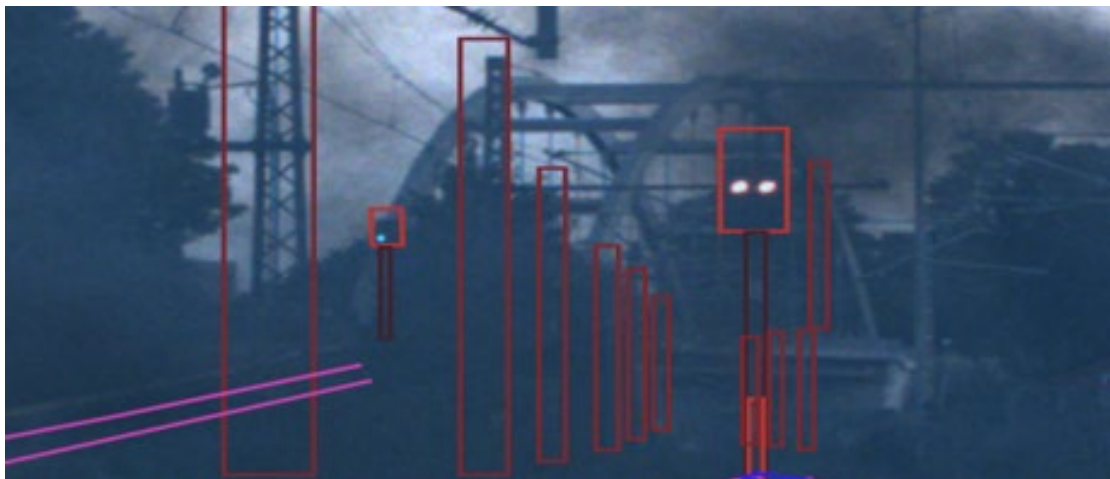


Abbildung 41: Annotationen von Objekten hinter Rauch im niedrigauflösenden Farbbild

Befinden sich **Regentropfen** (siehe Abbildung 42) auf der Infrarotkamera, kommt es zu einer starken Minderung der Qualität des Infrarotbildes. Dadurch sind Objekte nur noch schwer erkennbar, was die genaue Annotation erschwert und den Zeitaufwand pro Annotation erhöht. Im Beispiel sieht man, dass die Oberleitungsmasten und Schienen nur noch sehr schwer erkennbar sind.

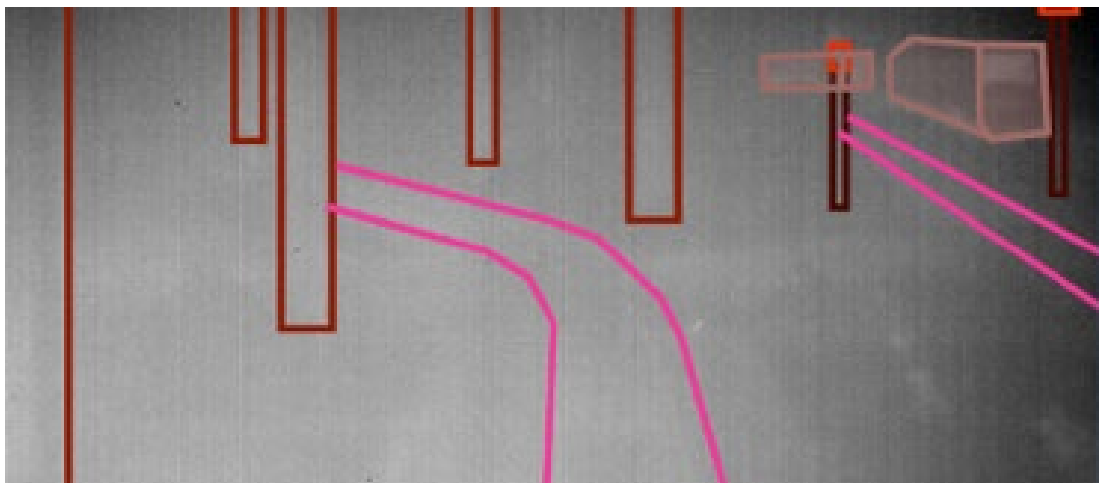


Abbildung 42: Schwer erkennbare Objekte im Infrarotbild bei Regen

Weist eine Sequenz pro M-Frame eine **überdurchschnittlich hohe Menge an zu annotierenden Objekten** auf (siehe Abbildung 43), muss die AP viel Konzentration aufbringen, um alle Objekte zu annotieren und bei allen Objekten die nötigen Anpassungen vorzunehmen. Die Wahrscheinlichkeit, etwas zu übersehen, ist in einer solchen Sequenz sehr hoch, besonders, wie in diesem Beispiel, bei vielen kleineren Objekten im Hintergrund.

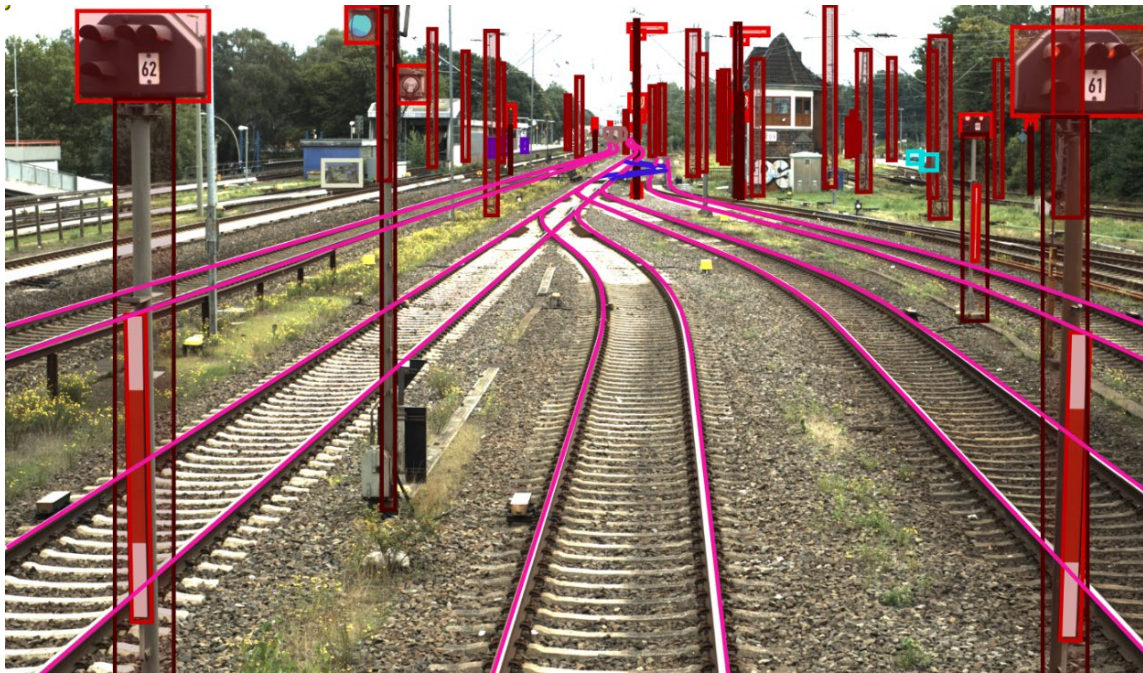


Abbildung 43: Eine Sequenz mit überdurchschnittlich vielen Objekten pro M-Frame

3.3 Statistik der erstellten Annotationen

Die Verteilung der Annotationen auf die verschiedenen Objektklassen ist in Abbildung 44 dargestellt, die Aufteilung auf die Sensoren in Abbildung 45, die Verteilung auf die verschiedenen Sequenzen in Abbildung 46. Der veröffentlichte Referenzsensordatensatz mit Annotationen besteht aus 21 Sequenzen untergliedert in 45 Teilsequenzen mit insgesamt 204.091 Annotationen, welche sich auf 1.534 M-Frames aufteilen. Besonders häufig treten dabei die Objektklassen Personen, Signale und Oberleitungsmaste auf. Er ist der erst frei verfügbare Datensatz, der eine so umfangreiche Auswahl an Objekten aus dem Bahnkontext und sowohl Farbbild-, Infrarot-, LiDAR- als auch RaDAR-Daten enthält.

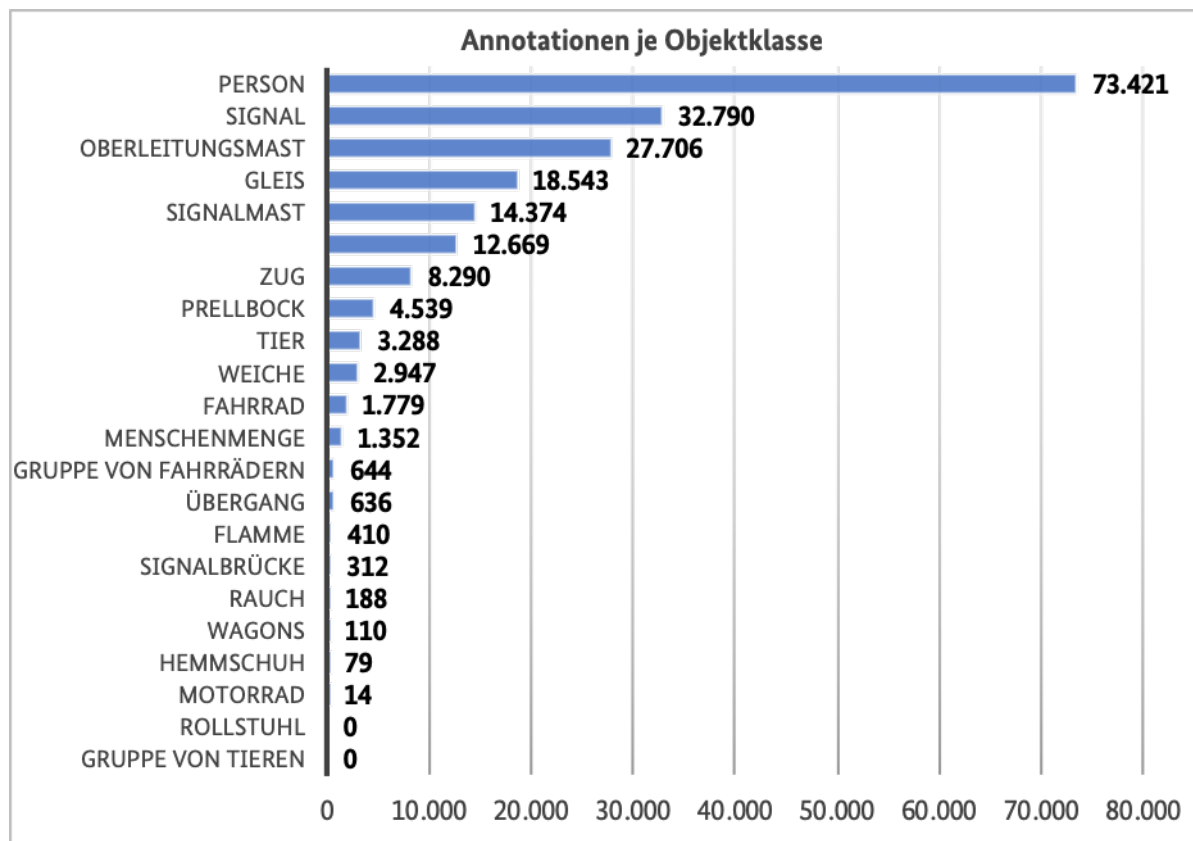


Abbildung 44: Anzahl der Annotationen je Objektklasse

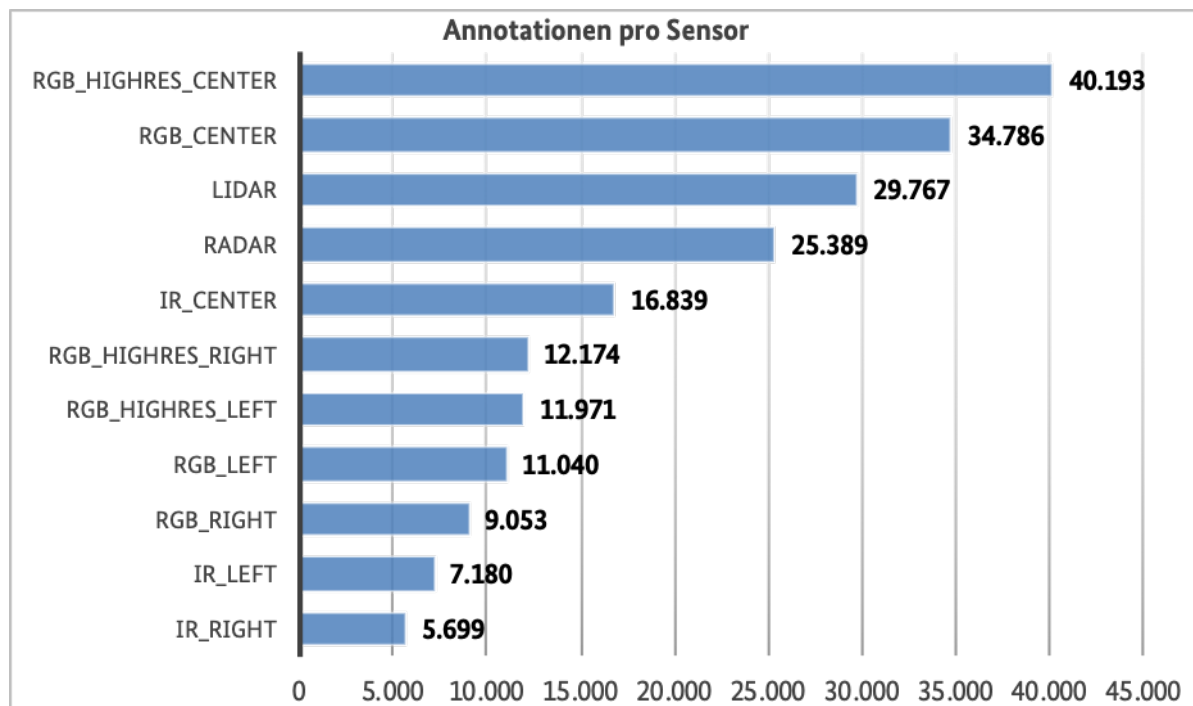


Abbildung 45: Anzahl der Annotationen je Sensor

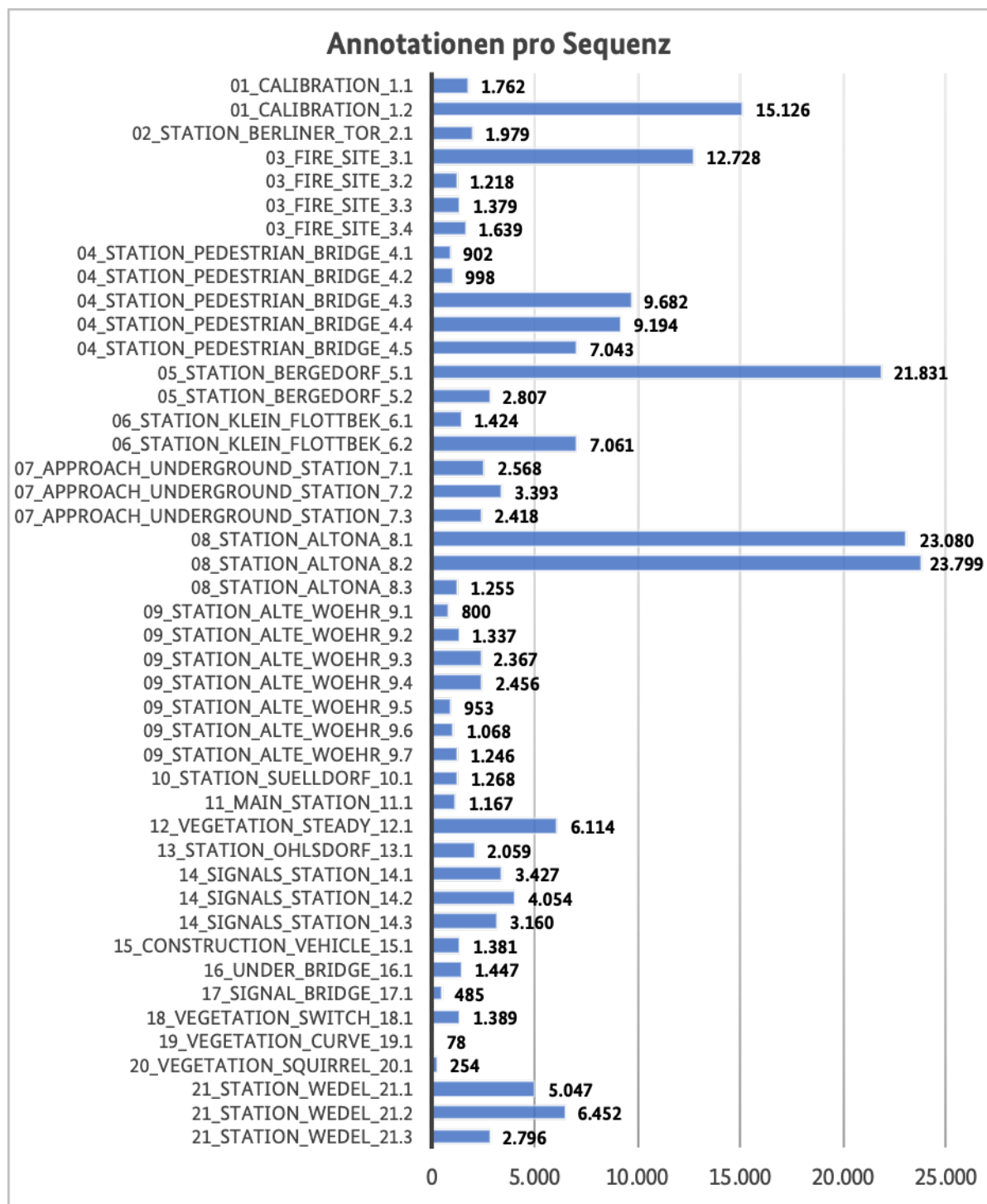


Abbildung 46: Anzahl der Annotationen je Sequenz

3.4 Anonymisierung

Die zu annotierenden Daten wurden vor Projektbeginn durch ein automatisiertes Verfahren voranonymisiert. Dabei wurden Gesichter und Nummernschilder von Fahrzeugen unkenntlich gemacht. Die Anonymisierung erfolgte mittels eines Weichzeichners, wobei Nummernschilder als Rechtecke und Gesichter als Ellipsen segmentiert werden (siehe Abbildung 47).

Im Projekt wurden die Aufnahmen aller sechs Farbbildsensoren manuell auf Vollständigkeit der Voranonymisierung geprüft. Bei fehlender Anonymisierung (Falschnegativ) wurde der entsprechende Bereich manuell als 2D-BBa markiert. Für die Anonymisierung von Personen wurde die Regel definiert, dass ein Personenbezug dann anzunehmen ist, wenn mindestens eines der drei Gesichtsmerkmale Augen, Nase oder Mund erkennbar ist. In diesem Fall erfolgt die Verfremdung des gesamten Gesichtes. Umgekehrt bedeutet ein Hinterkopf, dass an dieser Stelle keine Verfremdung erforderlich ist.



Abbildung 47: Gesichter werden ellipsenförmig durch einen Blur-Effekt unkenntlich gemacht

Die segmentierten Bildbereiche wurden durch eine Software anonymisiert, indem an den entsprechenden Stellen ein Gaußscher Weichzeichner angewendet wurde. Dabei wurde die openCV-Implementierung Gaussian Blur mit einer „kernel size“ als Funktion der Auflösung des Bildes angewendet:

$$\text{kernel size} = f(\text{width}, \text{height}) = a \cdot \sqrt{\text{width}^2 + \text{height}^2}$$

Der Faktor a steuert die Stärke des Weichzeichners. Die Anonymisierung wurde mit $a = 0,35$ durchgeführt, da dieser Wert ein optisch vergleichbares Ergebnis mit der automatisierten Anonymisierung ergab.

Im Ergebnis der Anonymisierung wurden 110.022 Bilder begutachtet und in Summe 6.483 Objekte nachanonymisiert, siehe Tabelle 30.

TABELLE 30: ERGEBNISSE DER ANONYMISIERUNG

	Nachanonymisiert	Anteil
Gesichter	6.173	95,2 %
Nummernschilder	310	5,8 %
Insgesamt	6.483	100,0 %

Wurden Bildbereiche gefunden, die fälschlicherweise anonymisiert wurden (Falschpositiv), wurden diese durch den Bildinhalt der Originalaufnahmen ersetzt. Insgesamt wurden so 2.299 falschpositive Anonymisierungen korrigiert.

4 Fazit

Für die Umsetzung der Überwachung des Zugumfelds und der Erkennung und Einordnung von Objekten in der Umgebung des Zuges bei ATO sind verschiedene Detektionsfunktionen wie Erkennung, Lokalisierung und Nachverfolgung von Objekten eine wichtige Voraussetzung. Diese basieren auf Verfahren des maschinellen Lernens, erfordern also hochqualitative Lerndaten (annotierte Sensordaten) für ihre Umsetzung. Speziell für den Bahnbereich existieren hierbei nur wenige geeignete Datensätze.

Der im Projekt erstellte annotierte, multisensorielle Datensatz wird den Anforderungen für Detektionsfunktionen im Bahnkontext gerecht, da er viele bahnspezifische Objektklassen und Klassenattribute enthält sowie mittels der Tracking-ID eine Zuordnung der Annotationen zu Realweltobjekten über Sensoren und Zeit hinweg erlaubt. Als Ergebnis des Projektes sind entstanden:

1. eine detaillierte Annotationsvorschrift zum Markieren und Beschreiben von Objekten im Bahnkontext,
2. eine Beschreibung des Annotations- und Qualitätssicherungsprozesses zur Erzeugung hochqualitativer Annotationen
3. ein vollständig anonymisierter Referenzsensordatensatz mit präzisen Annotationen, der öffentlich zugänglich ist

Die Annotationsvorschrift umfasst dabei sowohl allgemeine als auch klassenspezifische Annotationsregeln. Für jede Objektklasse sind die erforderlichen Annotationsgeometrien sowie die klassenspezifischen Attribute und deren mögliche Ausprägungen gegeben. Des Weiteren ist das Exportformat der Annotationen auf Basis des ASAM OpenLABEL [7] Formats detailliert beschrieben. Die Annotationsvorschrift bildet eine wesentliche Grundlage für die Erstellung weiterer homogener, annotierter Datensätze.

Der Annotations- und Qualitätssicherungsprozess umfasst dreizehn Einzelschritte zur Erstellung hochqualitativer Annotationen. Dabei wechseln sich manuelle Tätigkeiten (Erstellen und Anpassen von Annotationen, Review von Annotationen) mit automatischen Prozessschritten (Projektion und Interpolation von Annotationen, automatische Konsistenzüberprüfungen, Datenkonvertierung) ab.

Der veröffentlichte Referenzsensordatensatz mit Annotationen besteht aus 21 Sequenzen untergliedert in 45 Teilsequenzen mit insgesamt 204.091 Annotationen, welche sich auf 1.534 M-Frames aufteilen. Besonders häufig treten dabei die Objektklassen Personen, Signale und Oberleitungsmaste auf. Er ist der erste frei verfügbare Datensatz, der eine so umfangreiche Auswahl an Objekten aus dem Bahnkontext und sowohl Farbbild-, Infrarot-, LiDAR- als auch RaDAR-Daten enthält.

In Zukunft sollte betrachtet werden, welche Anpassungen der Annotationsvorschrift und des Annotations- und Qualitätssicherungsprozesses für weitere Sensorkonfigurationen notwendig sind.

Abbildungsverzeichnis

Abbildung 1: Annotation eines Objektes mit einer 2D-BBa.....	14
Abbildung 2: Annotation eines Objektes mit einer 2D-BBr	14
Abbildung 3: Annotation eines Objektes mit zwei 2D-PL.....	15
Abbildung 4: Annotation eines Objektes mit zwei 2D-PG	15
Abbildung 5: Annotation eines Objektes mit einer 3D-BB.....	15
Abbildung 6: Annotation mehrerer Objekte mit 3D-PL.....	15
Abbildung 7: Annotation eines Objektes mit einer SemSeg	16
Abbildung 8: Annotation eines Objektes mit einer SemSeg	16
Abbildung 9: Die Objektklasse „Person“ („ <i>Person</i> “) umfasst alle Menschen und menschenähnliche Attrappen (engl. „Dummy“) inklusive am Körper getragener Objekte.....	17
Abbildung 10: Die Objektklasse „Menschenmenge“ („ <i>Crowd</i> “) umfasst Gruppen von Menschen, die nur schwer voneinander unterscheidbar sind	17
Abbildung 11: Die Objektklasse „Zug“ („ <i>Train</i> “) umfasst alle Arten von Schienenfahrzeugen.....	17
Abbildung 12: Die Objektklasse „Waggons“ („ <i>Wagons</i> “) umfasst alle Arten von gezogenen Objek- ten im Schienentransport.....	17
Abbildung 13: Die Objektklasse „Fahrrad“ („ <i>Bicycle</i> “) umfasst alle Arten von unmotorisierten Ein- und Zweirädern.....	17
Abbildung 14: Die Objektklasse „Gruppe von Fahrrädern“ („ <i>Group of Bicycles</i> “) umfasst Ansamm- lungen von Fahrrädern, die sich nur schwer getrennt voneinander annotieren lassen ..	17
Abbildung 15: Die Objektklasse „Motorrad“ („ <i>Motorcycle</i> “) umfasst alle Arten von Motorrädern, Mopeds und deren elektrische Varianten.....	18
Abbildung 16: Die Objektklasse „Straßenfahrzeug“ („ <i>Road Vehicle</i> “) umfasst alle Arten von motori- sierten Automobilen und deren elektrische Varianten	18
Abbildung 17: Die Objektklasse „Tier“ („ <i>Animal</i> “) umfasst alle Tierarten.....	18
Abbildung 18: Die Objektklasse „Gruppe von Tieren“ („ <i>Group of Animals</i> “) umfasst Ansammlungen von Tieren, die sich nur schwer getrennt voneinander annotieren lassen	18
Abbildung 19: Die Objektklasse „Rollstuhl“ („ <i>Wheelchair</i> “) umfasst alle Arten von handgetriebenen und elektrischen Rollstühlen (Quelle: pxby666/Pixabay)	18
Abbildung 20: Die Objektklasse „Hemmschuh“ („ <i>Drag Shoe</i> “) umfasst Objekte, die Zugräder an Ort und Stelle halten.....	18
Abbildung 21: Die Objektklasse „Gleis“ („ <i>Track</i> “) umfasst Schienenpaare, die ein Gleis bilden, sowie das Gleisbett dazwischen	19
Abbildung 22: Die Objektklasse „Übergang“ („ <i>Transition</i> “) umfasst kurze Schienenpaare, die zwei Hauptgleise verbinden, sowie das Gleisbett dazwischen	19
Abbildung 23: Die Objektklasse „Weiche“ („ <i>Switch</i> “) umfasst Teile der Schienen, die es Zügen ermöglichen, das Gleis zu wechseln	19
Abbildung 24: Die Objektklasse „Oberleitungsmast“ („ <i>Catenary Pole</i> “) umfasst alle Masten, die Oberleitung beziehungsweise den Fahrdrabt für die Stromversorgung der elektrifi- zierten Züge tragen.....	20

Abbildung 25: Die Objektklasse „Signalmast“ („ <i>Signal Pole</i> “) umfasst alle Masten, an denen vollbahnrelevante Signale angebracht sind.....	20
Abbildung 26: Die Objektklasse „Signal“ („ <i>Signal</i> “) umfasst alle Arten von vollbahnrelevanten Form- und Lichtsignalen.....	20
Abbildung 27: Die Objektklasse „Signalbrücke“ („ <i>Signal Bridge</i> “) umfasst horizontale Strukturen, an denen bahnrelevante Signale angebracht sind.....	20
Abbildung 28: Die Objektklasse „Prellbock“ („ <i>Buffer Stop</i> “) umfasst Objekte, die das Ende eines Gleises markieren.....	20
Abbildung 29: Die Objektklasse „Flamme“ („ <i>Flame</i> “) umfasst alle Arten von Feuer.....	21
Abbildung 30: Die Objektklasse „Rauch“ („ <i>Smoke</i> “) umfasst alle Rauchschwaden, die die Sicht verdecken.....	21
Abbildung 31: Aufbau einer Sequenz.....	41
Abbildung 32: Übersicht des Annotations- und Qualitätssicherungsprozesses.....	42
Abbildung 33: Schienenquerschnitt für die SemSeg.....	46
Abbildung 34: Assoziation von manuell erstellten Annotationen und den Ergebnissen des Objektdetektors.....	48
Abbildung 35: Assoziation von Annotationen eines Frames mit den Annotationen des Folgeframes.....	49
Abbildung 36: Untersuchung der „Abgeschnitten“-Attribute auf Plausibilität (ITT = „isTruncatedTop“, iTB = „isTruncatedBottom“).....	50
Abbildung 37: Verschiedene Sensorframes bilden einen M-Frame.....	54
Abbildung 38: Annotation eines M-Frames.....	57
Abbildung 39: Doppelkreuzweiche aus seitlicher Ansicht mit begonnener Annotation.....	58
Abbildung 40: Überbelichteter Oberleitungsmast (ganz links im Bild).....	58
Abbildung 41: Annotationen von Objekten hinter Rauch im niedrigauflösenden Farbbild.....	59
Abbildung 42: Schwer erkennbare Objekte im Infrarotbild bei Regen.....	59
Abbildung 43: Eine Sequenz mit überdurchschnittlich vielen Objekten pro M-Frame.....	60
Abbildung 44: Anzahl der Annotationen je Objektklasse.....	61
Abbildung 45: Anzahl der Annotationen je Sensor.....	61
Abbildung 46: Anzahl der Annotationen je Sequenz.....	62
Abbildung 47: Gesichter werden ellipsenförmig durch einen Blur-Effekt unkenntlich gemacht.....	63

Tabellenverzeichnis

Tabelle 1: Zuordnung der Annotationsgeometrien zu den Objektklassen.....	22
Tabelle 2: Attribute der Objektklasse „Person“	23
Tabelle 3: Attribute der Objektklasse „Menschenmenge“	25
Tabelle 4: Attribute der Objektklasse „Zug“	25
Tabelle 5: Attribute der Objektklasse „Waggons“	26
Tabelle 6: Attribute der Objektklasse „Fahrrad“	27
Tabelle 7: Attribute der Objektklasse „Gruppe von Fahrrädern“	28
Tabelle 8: Attribute der Objektklasse „Motorrad“	28
Tabelle 9: Attribute der Objektklasse „Straßenfahrzeug“	29
Tabelle 10: Attribute der Objektklasse „Tier“	30
Tabelle 11: Attribute der Objektklasse „Gruppe von Tieren“	31
Tabelle 12: Attribute der Objektklasse „Rollstuhl“	31
Tabelle 13: Attribute der Objektklasse „Hemmschuh“	32
Tabelle 14: Attribute der Objektklasse „Gleis“	33
Tabelle 15: Attribute der Objektklasse „Übergang“	34
Tabelle 16: Attribute der Objektklasse „Weiche“	35
Tabelle 17: Attribute der Objektklasse „Oberleitungsmast“	36
Tabelle 18: Attribute der Objektklasse „Signalmast“	36
Tabelle 19: Attribute der Objektklasse „Signal“	37
Tabelle 20: Attribute der Objektklasse „Signalbrücke“	38
Tabelle 21: Attribute der Objektklasse „Prellbock“	39
Tabelle 22: Attribute der Objektklasse „Flamme“	39
Tabelle 23: Attribute der Objektklasse „Rauch“	40
Tabelle 24: Rollenbeschreibung.....	41
Tabelle 25: Konsistenzmatrix für Attribute	51
Tabelle 26: Technische Daten der Kameras	52
Tabelle 27: Technische Daten der LiDARs und des RaDARs	53
Tabelle 28: Technische Daten der Positions- und Beschleunigungssensoren	53
Tabelle 29: Sequenzübersicht.....	54
Tabelle 30: Ergebnisse der Anonymisierung.....	63
Tabelle 31: Detaillierte Sequenzübersicht.....	69
Tabelle 32: Grenzen der Abmessungen einzelner Objektklassen (Angaben in Metern)	85

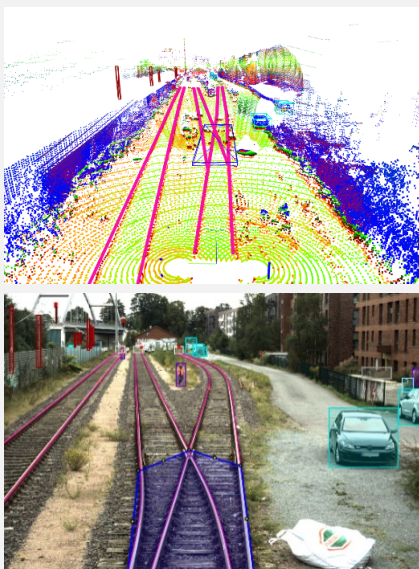
Quellenverzeichnis


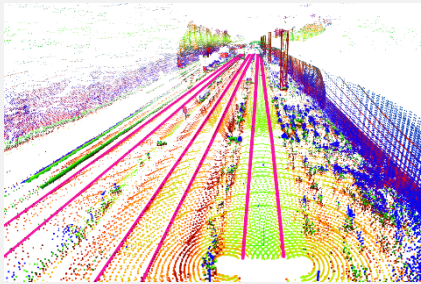
- [1] K. Schwalbe, K. Förster und M. Scheithauer, „Labeling Guide for Multi-sensor Data for Object Recognition in the Railway Domain, „Deutsches Zentrum für Schienenverkehrsforschung beim Eisenbahn-Bundesamt“, Dresden, 2023.
- [2] R. Tilly, P. Neumaier, K. Schwalbe, P. Klasek, R. Tagiew, P. Denzler, T. Klockau, M. Boekhoff und M. Köppel, „Open Sensor Data for Rail 2023 [Datensatz]“, 2023. [Online]. Available: 10.57806/9mv146r0.
- [3] O. Zendel, M. Murschitz, M. Zeilinger, D. Steininger, S. Abbasi und C. Beleznaï, „RailSem19: A Dataset for Semantic Rail Scene *Understanding*, „*IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*“, p. 1221–1229, 2019.
- [4] J. Harb, N. Rebenä, R. Chosidow, G. Roblin, R. Potarusov und H. Hajri, „FRSign: A large-scale traffic light dataset for autonomous trains“, 2020. [Online]. Available: abs/2002.05665. [Zugriff am 6 März 2023].
- [5] T. Toprak, B. Belenlioglu, B. Aydin, C. Guzelis und M. A. Selver, „Conditional Weighted Ensemble of Transferred Models for Camera Based Onboard Pedestrian Detection in Railway Driver Support Systems“, *IEEE Transactions on Vehicular Technology*, Bd. 69, Nr. 5, pp. 5041–5054, 2020.
- [6] R. Tagiew, T. Buder, R. Tilly, K. Hofmann und C. Klotz, „Datensätze für das autonome Fahren als Grundlage für GoA3+“, *ETR – Eisenbahntechnische Rundschau*, Bd. 09, pp. 10–14, 09 2021.
- [7] ASAM e.V., „ASAM OpenLABEL“, 07 10 2022. [Online]. Available: <https://www.asam.net/standards/detail/openlabel/>.
- [8] A. Bochkovsky, C.-Y. Wang und H.-Y. M. Liao, „Optimal Speed and Accuracy of Object Detection“, 2022. [Online]. Available: <https://arxiv.org/abs/2004.10934>. [Zugriff am 6 März 2023].
- [9] H. W. Kuhn, „The Hungarian method for the assignment problem“, *Naval Research Logistics Quarterly* 2, pp. 83–97, 1955.

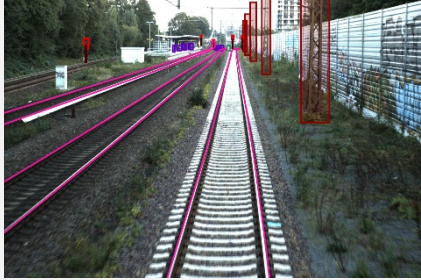
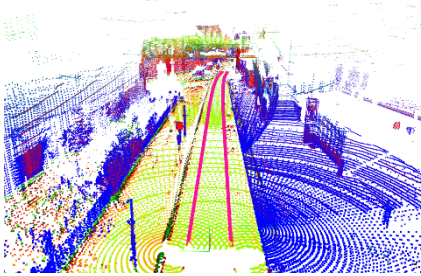

Anhänge

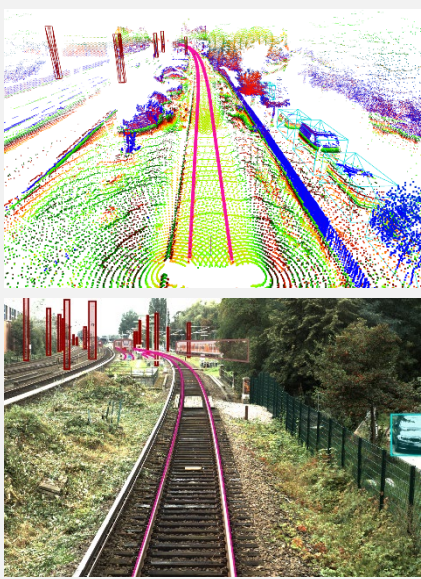
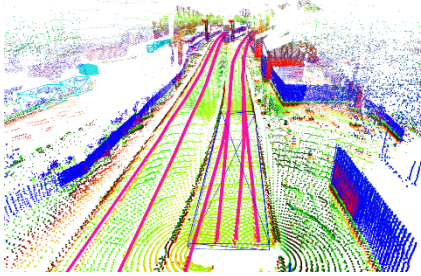
Anhang 1: Sequenzübersicht


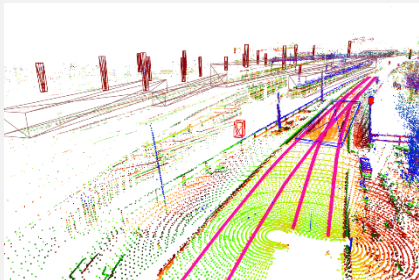

TABELLE 31: DETAILLIERTE SEQUENZÜBERSICHT

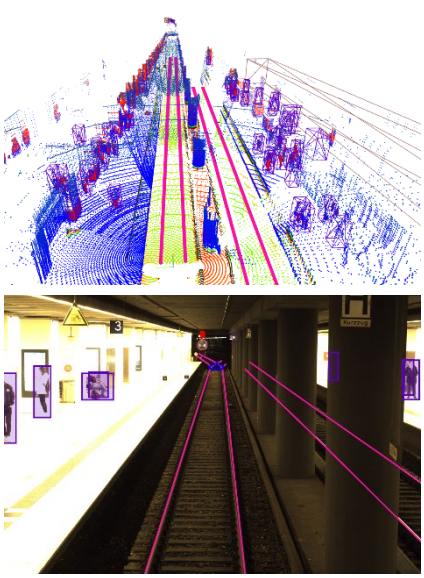
Nr.	Beispielbilder	Sequenzname	Teilsequenzname	Anzahl M-Frames	Beschreibung
1		1_calibration	1_calibration_1.1	10	<p>Der Ego-Zug fährt auf eine Abstellgleisanlage, bestehend aus drei Tracks mit Prellböcken am Ende.</p> <p>In Teilsequenz 1.1 befindet sich der Ego-Zug auf dem rechten Gleis. In 1.2 fährt er auf dem linken Gleis. Der Abstand zu den Prellböcken ist derselbe.</p> <p>Häufige Objektklassen sind Personen (Schienenarbeiter), Fahrzeuge, Oberleitungsmasten und Schienen.</p>
			1_calibration_1.2	100	

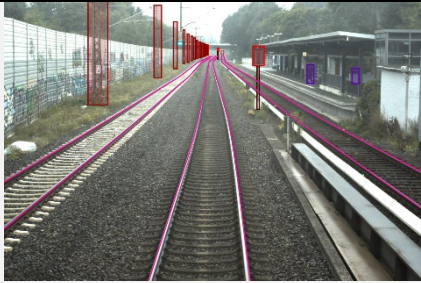
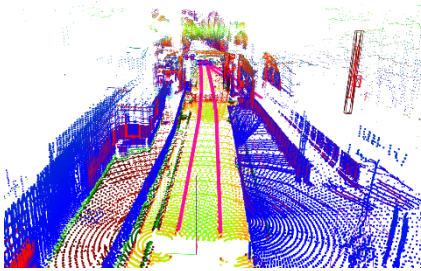

Nr.	Beispielbilder	Sequenzname	Teilsequenzname	Anzahl M-Frames	Beschreibung
2		2_station_berliner_tor	2_station_berliner_tor_2.1	10	<p>Der Ego-Zug fährt an einem Bahnsteig innerhalb einer Stadt vorbei.</p> <p>Häufige Objektklassen sind Schienen, Personen (Passagiere und Schienenarbeiter), Signale und Oberleitungsmasten.</p>
3		3_fire_site	3_fire_site_3.1	100	<p>Der Ego-Zug fährt durch ein Gebiet mit Lärmschutzmauer rechts und Vegetation links. Weiter vorn sieht man ein Feuer.</p> <p>Der Ego Zug fährt je nach Teilsequenz auf verschiedenen Gleisen und die Entfernung zum Feuer ist unterschiedlich. In 3.1 und 3.2 fährt er auf dem rechten Gleis, zwischen den beiden Teilsequenzen existiert ein zeitlicher Abstand von 24 Sekunden. In 3.3 befindet sich der Zug ebenso auf dem rechten Gleis in unmittelbarer Nähe zum Feuer. In 3.4 fährt er auf</p>
			3_fire_site_3.2	10	
			3_fire_site_3.3	10	
			3_fire_site_3.4	10	

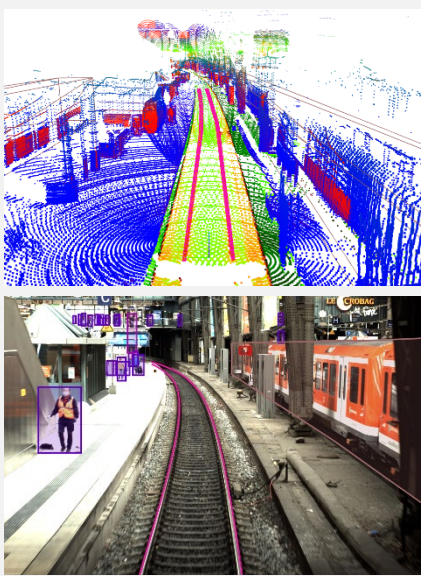
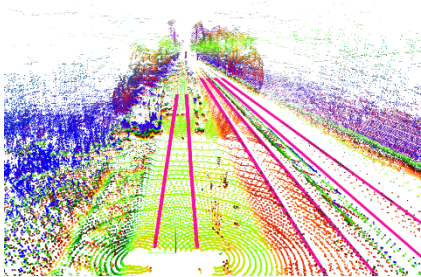
Nr.	Beispielbilder	Sequenzname	Teilsequenzname	Anzahl M-Frames	Beschreibung
					dem mittleren Gleis, links neben ihm sieht man einen weiteren Zug. Häufige Objektklassen sind Schienen, Personen (Passagiere, Schienenarbeiter sowie Feuerwehkräfte), Signale und Oberleitungsmasten.
4	 	4_station_pedestrian_bridge	4_station_pedestrian_bridge_4.1	10	Der Ego-Zug steht an einem Bahnsteig mit Blick auf dessen Ende.
			4_station_pedestrian_bridge_4.2	10	Häufige Objektklassen sind Schienen, Personen (Passagiere sowie Schienenarbeiter), Signale und Fahrzeuge.
			4_station_pedestrian_bridge_4.3	100	Die Teilsequenzen unterscheiden sich lediglich hinsichtlich der sichtbaren Personen.
			4_station_pedestrian_bridge_4.4	100	
			4_station_pedestrian_bridge_4.5	76	

Nr.	Beispielbilder	Sequenzname	Teilsequenzname	Anzahl M-Frames	Beschreibung
5		5_station_bergedorf	5_station_bergedorf_5.1	100	<p>Der Ego-Zug fährt auf einen Bahnsteig zu. Zwischen 5.1 und 5.2 existiert ein zeitlicher Abstand von 57 Sekunden, wodurch der Ego-Zug im zweiten Framebereich bereits am Bahnsteig angekommen ist.</p> <p>Häufige Objektklassen im ersten Framebereich sind Schienen, Züge sowie Oberleitungsmasten. Im zweiten Framebereich sind es vor allem Personen (Passagiere), Schienen und Signale.</p>
			5_station_bergedorf_5.2	10	
6		6_station_klein_flottbek	6_station_klein_flottbek_6.1	10	<p>In 6.1 fährt der Ego-Zug auf einen Bahnsteig mit wartenden Passagieren zu. In 6.2 ist er bereits dort angekommen, ein zeitlicher Abstand zwischen den beiden Teilsequenzen lässt sich nicht ermitteln.</p> <p>Häufige Objektklassen sind Schienen, Personen (Passagiere) und Signale.</p>
			6_station_klein_flottbek_6.2	100	

Nr.	Beispielbilder	Sequenzname	Teilsequenzname	Anzahl M-Frames	Beschreibung
					
7	 	7_approach_underground_station	7_approach_underground_station_7.1	10	Der Ego-Zug fährt auf einen großen Bahnhof zu. In 7.3 ist er dort angekommen und steht vor einem mehrspurigen Tunnelsystem. Ein zeitlicher Abstand zwischen den Teilsequenzen lässt sich nicht ermitteln.
			7_approach_underground_station_7.2	10	
			7_approach_underground_station_7.3	10	In den ersten beiden Teilsequenzen gehören Züge, Schienen, Oberleitungsmasten und Signale zu den häufigsten Objektklassen. In 7.3 sind es Schienen, Weichen und Signale.
8		8_station_altona	8_station_altona_8.1	100	
			8_station_altona_8.2	100	

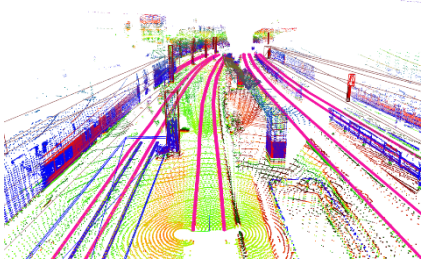

Nr.	Beispielbilder	Sequenzname	Teilsequenzname	Anzahl M-Frames	Beschreibung
			8_station_altona_8.3	10	<p>Der Ego-Zug steht an einem Bahnsteig in einer Untergrundstation. Circa 50 Sekunden nach Beendigung der zweiten Teilsequenz ist in 8.3 ein entgegenkommender Zug ebenso eingefahren und verdeckt die Sicht auf den rechten Bahnsteig.</p> <p>Häufige Objektklassen sind Schienen und Personen (Passagiere).</p>
9		9_station_alte_woehr	9_station_alte_woehr_9.1 9_station_alte_woehr_9.2 9_station_alte_woehr_9.3 9_station_alte_woehr_9.4 9_station_alte_woehr_9.5 9_station_alte_woehr_9.6 9_station_alte_woehr_9.7	10 10 10 10 10 10 10	<p>Der Ego-Zug befindet sich an einem Bahnsteig. Auf der linken Seite sieht man eine Lärmschutzmauer. Je nach Teilsequenz befindet sich der Zug kurz vor dem, neben dem oder hinter dem Bahnsteig. In 9.1 und 9.2 fährt der Ego-Zug auf dem mittleren Schienenpaar, in den anderen Teilsequenzen ist es das linke Schienenpaar. Des Weiteren herrscht in einigen Teilsequenzen regnerisches Wetter, wobei sich auch die Anzahl der Passagiere unterscheidet.</p>

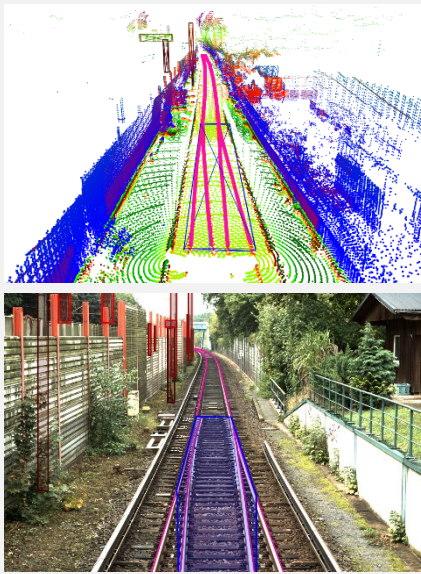
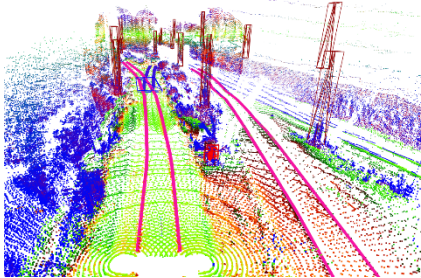
Nr.	Beispielbilder	Sequenzname	Teilsequenzname	Anzahl M-Frames	Beschreibung
					Häufige Objektklassen sind Schienen, Oberleitungsmasten und in einigen Teilsequenzen Personen (Passagiere).
10	 	10_station_suelldorf	10_station_suelldorf_10.1	10	<p>Der Ego-Zug fährt zunächst durch einen Waldabschnitt und kommt im Anschluss an einem Bahnsteig vorbei. Die Kamera schaut in dieser Sequenz nach hinten.</p> <p>Häufige Objektklassen sind Schienen, Signale und Personen (Passagiere).</p>

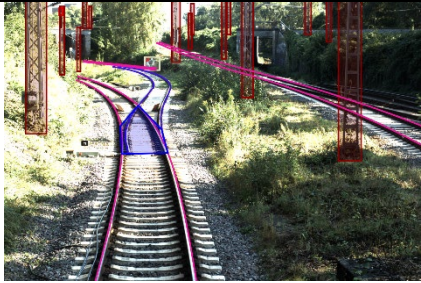
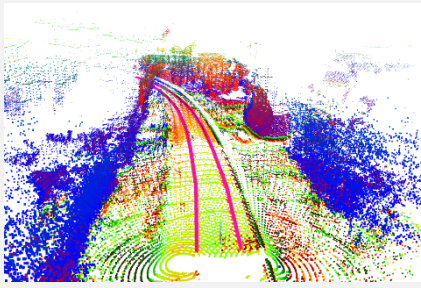

Nr.	Beispielbilder	Sequenzname	Teilsequenzname	Anzahl M-Frames	Beschreibung
11		11_main_station	11_main_station_11.1	10	<p>Der Ego-Zug fährt an einem Bahnsteig in einem Hauptbahnhof ein. Es befinden sich weitere stehende Züge im Bahnhof. Weiter entfernt sieht man einen Tunnel.</p> <p>Eine häufige Objektklasse sind Personen (Passagiere).</p>
12		12_vegetation_steady	12_vegetation_steady_12.1	98	<p>Der Ego-Zug steht auf einem Streckenabschnitt, der beidseitig von Bäumen umgeben ist.</p> <p>Häufige Objektklassen sind Schienen und Signale.</p>

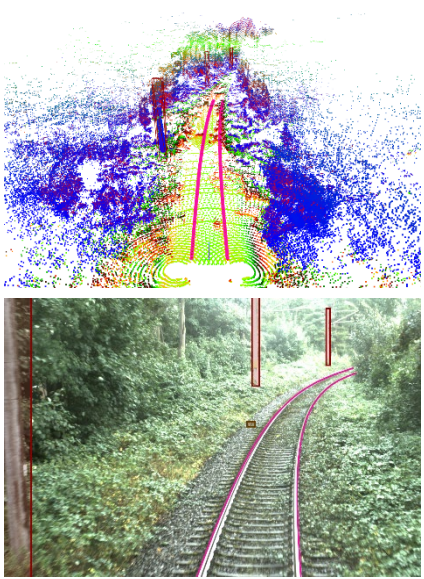
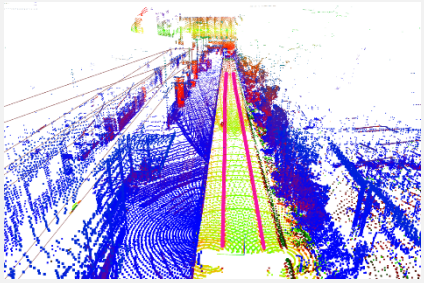
Nr.	Beispielbilder	Sequenzname	Teilsequenzname	Anzahl M-Frames	Beschreibung
					
13	 	13_station_ohlsdorf	13_station_ohlsdorf_13.1	10	<p>Der Ego-Zug fährt auf einen Bahnsteig zu. Es sind weitere stehende Züge zu sehen.</p> <p>Häufige Objektklassen sind Schienen, Signale und Personen (Passagiere).</p>


Nr.	Beispielbilder	Sequenzname	Teilsequenzname	Anzahl M-Frames	Beschreibung
14		14_signals_station	14_signals_station_14.1	10	<p>Der Ego-Zug fährt an einem Bahnsteig vorbei. Da die Kamera in dieser Sequenz nach hinten schaut, entfernt sich der Ego-Zug mit jeder Teilsequenz weiter von dem Bahnsteig. Die Teilsequenzen schließen unmittelbar aneinander an.</p> <p>Häufige Objektklassen sind Schienen, Signale und Personen (Passagiere).</p>
			14_signals_station_14.2	10	
			14_signals_station_14.3	10	
15		15_construction_vehicle	15_construction_vehicle_15.1	10	<p>Der Ego-Zug fährt durch einen Streckenabschnitt mit viel Vegetation an zwei anderen Zügen vorbei. Bei einem der beiden Züge handelt es sich um ein Arbeitsfahrzeug.</p> <p>Häufige Objektklassen sind Schienen, Signale und Oberleitungsmasten.</p>

Nr.	Beispielbilder	Sequenzname	Teilsequenzname	Anzahl M-Frames	Beschreibung
					
16	 	16_under_bridge	16_under_bridge_16.1	10	<p>Der Ego-Zug steht neben einer Brücke in einem städtischen Streckenabschnitt. Es sind zwei entgegengerichtete Züge zu sehen.</p> <p>Häufige Objektklassen sind Schienen und Oberleitungsmasten.</p>

Nr.	Beispielbilder	Sequenzname	Teilsequenzname	Anzahl M-Frames	Beschreibung
17		17_signal_bridge	17_signal_bridge_17.1	10	<p>Der Ego-Zug fährt durch einen schmalen Streckenabschnitt mit einer Lärmschutzmauer auf der linken und Vegetation auf der rechten Seite auf eine Brücke zu.</p> <p>Häufige Objektklassen sind Schienen und Signale.</p>
18		18_vegetation_switch	18_vegetation_switch_18.1	10	<p>Der Ego-Zug fährt durch einen Streckenabschnitt mit viel Vegetation auf einen Tunnel zu.</p> <p>Häufige Objektklassen sind Schienen und Oberleitungsmasten.</p>

Nr.	Beispielbilder	Sequenzname	Teilsequenzname	Anzahl M-Frames	Beschreibung
					
19	 	19_vegetation_curve	19_vegetation_curve_19.1	10	<p>Der Ego-Zug fährt durch eine Kurve in einem Streckenabschnitt mit viel Vegetation.</p> <p>Außer den Schienen sind keine weiteren Objekte zu sehen.</p>

Nr.	Beispielbilder	Sequenzname	Teilsequenzname	Anzahl M-Frames	Beschreibung
20		20_vegetation_squirrel	20_vegetation_squirrel_20.1	10	Der Ego-Zug fährt auf eine Kurve in einem Streckenabschnitt mit viel Vegetation zu. Ein Eichhörnchen befindet sich auf den Gleisen des Ego-Zuges. Häufige Objektklassen sind Schienen und Oberleitungsmasten.
21		21_station_wedel	21_station_wedel_21.1	100	Der Ego-Zug steht an einem Bahnsteig am Ende eines Gleises. Links neben ihm steht ein weiterer Zug. Häufige Objektklassen sind Schienen und Personen.
			21_station_wedel_21.2	100	

Nr.	Beispielbilder	Sequenzname	Teilsequenzname	Anzahl M-Frames	Beschreibung
			21_station_wedel_21.3	40	

Anhang 2: Grenzen der Abmessungen einzelner Objektklassen

Für die einzelnen Objektklassen gibt es in vielen Fällen (physikalische) Unter- und Obergrenzen die Höhe bzw. Breite/Tiefe betreffend. Diese Grenzen werden zur Plausibilitätsüberprüfung der erstellten 3D-BB genutzt und sind in Tabelle 31 dargestellt. Bei manchen Objektklassen hängen die Grenzen von bestimmten Attributausprägungen ab, welche in der zweiten Spalte spezifiziert sind. Die Beschreibung der entsprechenden Attribute kann der Annotationsvorschrift entnommen werden. Die Höhe eines Objekts ist die Kantenlänge (der 3D-BB) entlang der Unten-Oben-Richtung (z-Koordinate im LiDAR). Die Breite ist die Kantenlänge entlang der Links-Rechts-Richtung (y-Koordinate im LiDAR) und die Tiefe ist die Kantenlänge entlang der Vorne-Hinten-Richtung (x-Koordinate im LiDAR). Da sich die meisten Objekte um die z-Achse drehen können, können Breiten- und Tiefenbeschränkungen austauschbar verstanden werden, d.h. dass eine der Kantenlängen (abseits der Höhe) die Breitengrenzen und die andere die Tiefengrenzen erfüllen muss.

TABELLE 32: GRENZEN DER ABMESSUNGEN EINZELNER OBJEKTKLASSEN (ANGABEN IN METERN)

Objektklasse	Attribute	Minimale Höhe	Maximale höhe	Minimale Breite	Maximale Breite	Minimale Tiefe	Maximale Tiefe	Kommentar
Person	„Alter“ = „Kind“ und „Pose“ = „aufrecht“	0,60	1,50	0,15	0,50	0,07	0,30	
	„Alter“ = „Kind“ und „Pose“ = „sitzend“	0,25	0,90	0,15	0,50	0,15	1,10	Ausgestreckte Beine
	„Alter“ = „Kind“ und „Pose“ = „liegend“	0,07	0,30	0,60	1,50	0,15	0,50	
	„Alter“ = „Erwachsener“ und „Pose“ = „aufrecht“	1,50	2,20	0,30	0,80	0,15	0,60	
	„Alter“ = „Erwachsener“ und „Pose“ = „sitzend“	0,50	1,30	0,30	0,80	0,40	1,60	Ausgestreckte Beine
	„Alter“ = „Erwachsener“ und „Pose“ = „liegend“	0,15	0,60	1,50	2,50	0,30	0,80	

Objektklasse	Attribute	Minimale Höhe	Maximale höhe	Minimale Breite	Maximale Breite	Minimale Tiefe	Maximale Tiefe	Kommentar
Menschenmenge		0,60	2,20	1,00	-	0,20	-	
Zug	„Typ“ = „Lokomotive“	2,70	5,00	2,00	3,50	3,00	40,00	Kleinsten Zug: Köf ²⁹ Maximale Höhe in Deutschland: 4,57 Maximale Breite in Deutschland: 3,15
	„Typ“ = „Inter-city“	2,40	4,80	2,40	3,50	150,00	520,00	
	„Typ“ = „Regionalverkehr“	2,80	5,00	2,40	3,50	150,00	500,00	
	„Typ“ = „Nahverkehr“	2,80	5,00	2,80	3,50	35,00	100,00	
Waggons	„Typ“ = „Inter-city“	2,40	4,80	2,40	3,50	150,0	520,00	
	„Typ“ = „Regionalverkehr“	2,80	5,00	2,40	3,50	150,0	500,00	
	„Typ“ = „Fracht“	1,40	5,00	2,80	3,50	12,5	-	
	„Typ“ = „Bauwesen“	2,00	5,00	2,00	3,50	6,00	-	

²⁹ Abkürzung für: Kleinlokomotive mit Ölmotor (Dieselmotor) und Flüssigkeitsgetriebe

Objektklasse	Attribute	Minimale Höhe	Maximale höhe	Minimale Breite	Maximale Breite	Minimale Tiefe	Maximale Tiefe	Kommentar
Fahrrad		0,60	1,20	0,20	0,80	0,90	4,00	
Gruppe von Fahrrädern		0,60	1,20	0,20	-	0,90	-	
Motorrad		0,80	1,20	0,60	1,20	1,30	2,40	
Straßenfahrzeug	„Typ“ = „Auto“	1,40	1,60	2,30	6,00	1,10	2,20	Beispiel für ein kleines Auto: Renault twizy
	„Typ“ = „Kleinbus“	1,10	2,00	1,80	8,00	1,30	2,40	
	„Typ“ = „Lastkraftwagen“	1,80	4,20	4,00	20,00	1,80	2,60	
	„Typ“ = „Bus“	2,60	4,20	5,00	30,00	2,40	2,55	Weltrekord Breite: 30
	„Typ“ = „Bau-fahrzeug“	1,40	4,20	3,00	-	1,80	2,55	
	„Typ“ = „Anhänger“	-	4,00	-	8,00	-	2,40	
Tier	„Spezies“ = „Hund“	0,10	1,40	0,20	1,50	0,06	0,70	
	„Spezies“ = „Rotwild“	0,15	2,50	0,20	2,40	0,10	6,00	Maximale Geweihbreite: 6
	„Spezies“ = „Fuchs“	0,10	0,60	0,20	1,60	0,10	0,30	Maximale Schwanzlänge: 0,4
	„Spezies“ = „Kaninchen“	0,10	0,50	0,10	0,80	0,05	0,30	

Objektklasse	Attribute	Minimale Höhe	Maximale höhe	Minimale Breite	Maximale Breite	Minimale Tiefe	Maximale Tiefe	Kommentar
	„Spezies“ = „Wildschwein“	0,15	1,20	0,20	2,20	0,10	1,00	
	„Spezies“ = „Kuh“	0,30	1,70	0,40	2,20	0,15	1,00	
Rollstuhl		0,50	1,00	1,00	1,40	0,50	1,00	
Hemmschuh		0,14	0,20	0,38	0,60	0,10	0,20	Typische Größe 1: 0,387 x 0,16 x 0,193 Typische Größe 2: 0,43 x 0,1 x 0,145
Oberleitungs- mast		7,00	20,00	-	-	-	-	
Signalmast		5,00	15,00	-	-	-	-	
Prellbock		0,50	2,50	0,80	2,00	0,30	3,00	