

# Schlussbericht

## zum Projekt OPTmicro

Verbundpartner

 <p>Förderkennzeichen: 01FV2051A</p>	<p>Hochschule Merseburg Fachbereich INW / Mechatronische Systeme Ansprechpartner: Prof. Dr.-Ing. Stephan Schmidt Eberhard-Leibnitz-Straße 2 06217 Merseburg</p> <p>Tel: +49-3461-46-2974 E-mail: <a href="mailto:stephan.schmidt@hs-merseburg.de">stephan.schmidt@hs-merseburg.de</a></p>
 <p>Förderkennzeichen: 01FV2051B</p>	<p>FusionSystems GmbH Annaberger Straße 240 09125 Chemnitz</p> <p>Tel: +49 371 8365250 E-mail: <a href="mailto:karsten.schwalbe@fusionsystems.de">karsten.schwalbe@fusionsystems.de</a></p>

# 1 Inhaltsverzeichnis

1	Inhaltsverzeichnis.....	2
2	Projektkontext (kurzer Bericht) .....	3
2.1	Aufgabenstellung.....	3
2.2	Voraussetzungen des Vorhabens .....	3
2.3	Planung und Ablauf des Vorhabens .....	4
2.4	Wissenschaftlicher und technischer Ausgangspunkt .....	4
2.5	Zusammenarbeit mit anderen Stellen.....	5
3	Projektdurchführung und -ergebnisse (ausführlicher Bericht).....	6
3.1	Verwendung der Zuwendung und erzielte Ergebnisse .....	6
3.1.1	AP 1 Spezifikation und Konzeption .....	6
3.1.2	AP 2 Aufbau Sensorik .....	11
3.1.3	AP 3 Labeling- Tool.....	15
3.1.4	AP 4 Datenerhebung .....	18
3.1.5	AP 5 Datenaufbereitung.....	21
3.1.6	AP 6 Baseline, Toolchain und Datenveröffentlichung.....	26
3.2	Wichtigste Positionen des zahlenmäßigen Nachweises.....	28
3.3	Notwendigkeit und Angemessenheit der geleisteten Arbeit.....	28
3.4	Verwertbarkeit der Ergebnisse.....	28
3.5	Fortschritt auf dem Gebiet .....	29
3.6	Veröffentlichung der Ergebnisse .....	29

## 2 Projektkontext (kurzer Bericht)

### 2.1 Aufgabenstellung

Weltweit stehen Städte, insbesondere in Europa, unter enormen Wandlungsdruck ihren Verkehr mittelfristig CO<sub>2</sub>-neutral und nachhaltig zu gestalten. Neue Paradigmen wie die 15-Minuten Stadt verändern grundlegend die Art und Weise von Mobilität und die Aufteilung von öffentlichem Raum. Straßen werden von spielenden Kindern geprägt sein, werden zum Treffpunkt für Menschen und auf die Bedürfnisse der „schwächeren“ Verkehrsteilnehmer hin gestaltet. Beispiele wie Begegnungszonen in Deutschland, Superblocks in Barcelona, die autofreie Innenstadt in Madrid sowie der 70% Radanteil in Groningen zeigen eindeutig: Die Entwicklung zur Straße für Menschen findet statt. E-Scooter, Hoverboards, boomender e-Bike-Absatz und sich jährlich verdoppelnde Verkaufszahlen von Lastenrädern machen deutlich, dass Verkehr in Städten zukünftig deutlich digitaler und mikromobiler sein wird und diese Fahrzeuge zunehmend in sogenannten unstrukturierten Verkehrsräume operieren müssen bzw. gefahren werden. Durch den Einsatz von Mikromobilen kann es zudem gelingen die CO<sub>2</sub>-Emissionen des Verkehrs um 44% zu senken. Dies wird durch den einfachen Effekt erzielt, dass diese Fahrzeuge deutlich kleiner, leichter und teilweise muskelangetrieben sind und so signifikant energieeffizienter gefahren werden können. Ein Lastenrad in der Paketzustellung braucht z.B. pro Kilometer nur 10% der Energie eines E-Vans in der Zustellung.

Im Mobilitätsmarkt findet eine ökologische Transformation statt mit vielen KMU und Start-Ups, die neue, effiziente Mikromobile als nachhaltige Alternativen zu aktuellen Fahrzeugen anbieten. Die häufig jungen und eher kleinen bis mittleren Marktakteure haben sich dabei bisher auf die Entwicklung robuster, kostengünstiger Fahrzeugkonzepte mit manueller Bedienung konzentriert, um die Marktetablierung zu realisieren. Damit Mikromobile im Trend der Automatisierung von Kraftfahrzeugen in Bezug auf Komfort, Ergonomie und Kosten wettbewerbsfähig bleiben und das damit verbundene Potential zum Klimaschutz voll zu heben, braucht es auch hier langfristig automatisierte und autonome Fahrfunktionen.

Eine wesentliche Barriere für die Entwicklung von automatisierten mikromobilen Fahrzeugkonzepten ist, dass es anders als im Pkw-Bereich keine offenen Datensätze für unstrukturierte Verkehrsräume gibt. Bestehende Datensätze sind meist aus der Kfz-Perspektive oder aus Drohnenaufnahmen generiert und daher nur sehr eingeschränkt verwendbar. Zusätzlich beschränken sich viele Datensätze nur auf eine Objektklasse wie z.B. Fußgänger. Besonders die veränderte Verkehrsdomäne, weg von der Fahrbahn hin zu kombinierten Rad- und Fußwegen oder offenen Plätzen, mit dem veränderten Verhalten der umgebenden Verkehrsteilnehmer im unstrukturierten Verkehrsraum, macht einen spezialisierten Datensatz notwendig.

Das durch den mFund geförderte Vorhaben OPTmicro hatte zum Ziel einen hochwertigen, offenen multisensoriellen Datensatz als Trainings- und Validierungsbasis für KI-Modelle in unstrukturierten Verkehrsräumen zur Verfügung zu stellen.

### 2.2 Voraussetzungen des Vorhabens

Aus einem Vorprojekt stand dem Konsortium ein automatisiert fahrendes Lastenrad mit vollständiger Sensorik und Datenverarbeitungskette für autonome Fahrfunktionen zur

Verfügung. Im Rahmen des Projektes wurden grundsätzlichen Erkenntnisse über die Anforderung an Datensätze und deren Erhebung im Kontext automatisierter Mikromobile gewonnen, an die das Projekt OPTmicro direkt anknüpfen konnte.

### 2.3 Planung und Ablauf des Vorhabens

Das Vorhaben gliederte sich in 6 Arbeitspakete (siehe Abbildung 1)

- AP 1: Spezifikation und Konzeption
- AP 2: Aufbau Sensorik
- AP 3: Labeling Tool
- AP 4: Datenerhebung
- AP 5: Datenaufbereitung
- AP 6: Baseline, Toolchain und Datenveröffentlichung

Die wesentlichen Projektmeilensteine waren der Abschluss der Spezifikation, der Aufbau des Sensorträgers, der Abschluss der Datenerhebung sowie die finale Datenveröffentlichung.

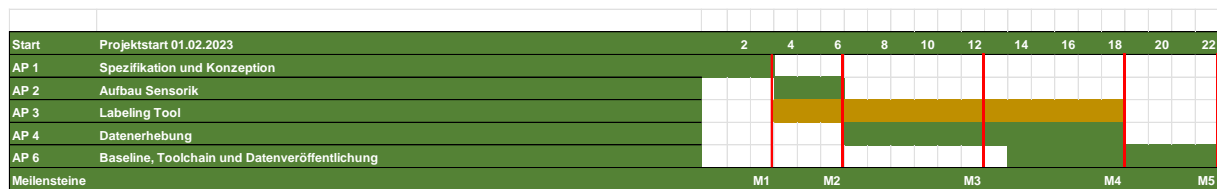


Abbildung 1: Projektablaufplan OPTmicro

Die verschiedenen Arbeitspakete wurden gemeinsam von den Projektpartnern Hochschule Merseburg und FusionSystems GmbH (kurz FS) durchgeführt.

### 2.4 Wissenschaftlicher und technischer Ausgangspunkt

Ein breites Spektrum an Unternehmen entwickelt weltweit Software für das autonome Fahren im Anwendungsbereich von Mobilität und Logistik. Im Mobilitätssektor liegt der Fokus der Entwicklung jedoch auf Einsatzszenarien mit Bussen, LKW und/oder PKW. Mikromobile mit deutlich geringerem Energie- und Platzbedarf haben ein sehr hohes Einsatzpotential, verringern Emissionen und Flächenverbrauch deutlich und tragen zu menschenzentrierten Straßengestaltungsformen bei. Die aktuellen fokussierten Kfz-Szenarien betreffen jedoch hauptsächlich definierte Fahrbahnen mit linearer Spurführung. Die dort verfolgten Fahrplanungsansätze sind für unstrukturierte Verkehrsräume von Mikromobilen nicht geeignet. Durch den Fokus auf die Kfz-Szenarien des automatisierten und autonomen Fahrens existieren in dieser Domäne auch offene Datensätze, die frei genutzt werden können, um Algorithmen zu entwickeln und zu trainieren. Bei vielen Datensätzen stehen dabei jedoch die hochgenaue Vermessung von Spurführungen, die korrekte Erkennung von Verkehrsschildern, die Erfassung des Verkehrsaufkommens oder die Erfassung von Lichtsignalanlagen im Fokus der Forschungsarbeiten. Die Sensordatenerhebung dafür wird vorrangig von Kraftfahrzeugflotten betrieben. Abseits der Kfz-Domäne bestehen Ansätze zur Datensatzgenerierung mit Fokus auf die korrekte Detektion und Erkennung von Verkehrsteilnehmenden, die Erfassung, Klassifizierung und Bewertung von (Rad-)Wegen, die jedoch häufig ohne Kamera- oder Lidardaten realisiert werden.

## 2.5 Zusammenarbeit mit anderen Stellen

Die im Projekt OPTmicro aufgeworfenen datenschutzrechtlichen Fragen in Bezug auf die Generierung und Bereitstellung eines Referenzdatensatzes für die Entwicklung von autonomen Mikromobilen wurden durch ein Gutachten der Kanzlei GSK Stockmann beantwortet.

Die manuelle Erstellung der Annotationen im Referenzdatensatz wurde als Unterauftrag von der Firma Intenta SEH GmbH übernommen.

## 3 Projektdurchführung und -ergebnisse (ausführlicher Bericht)

### 3.1 Verwendung der Zuwendung und erzielte Ergebnisse

Die folgenden Abschnitte dokumentieren die in den einzelnen Arbeitspaketen durchgeführten Arbeiten sowie die erzielten Projektergebnisse im Detail.

#### 3.1.1 AP 1 Spezifikation und Konzeption

##### **AP 1.1 – Sensorik und Rechenhardware**

Im Rahmen des AP 1.1 wurde der vorhandene Sensorträger „Eigen“ analysiert und die Eignung des Sensorkonzeptes für die veränderten Anforderungen überprüft. Es konnte dabei auf Erfahrung aus vorausgehenden Projekten (z.B. mit dem FS-Sensorfahrzeug) zurückgegriffen werden. Um das Projektziel möglichst unterschiedlicher Aufnahmen zu ermöglichen, wurde die Wetterfestigkeit der Sensorsuite als Anforderung in die Entwicklung eingebracht. Dies führte dazu, dass verschiedene neue Sensoren integriert wurden, dies umfasst drei ZED Zi Stereokameras im Heckbereich des Fahrzeuges, einen Trimble BX992 DPGS Receiver inkl. 2 Doppelantennen sowie einen Hokuyo UST-10LX Nahbereichs Laserscanner. Die Installation der Sensoren erforderte mechanische Anpassungen am Fahrzeug wie etwa die Konstruktion und Fertigung von Sensorhalterungen, aber auch eine Überarbeitung des elektrischen Systems zur Versorgung der Komponenten.

Die ursprünglich im Sensorträger verbauten Recheneinheiten wurden mit neuartigen, für die KI-Anwendung optimierten Einheiten ergänzt. Es handelt sich dabei um Stereolabs ZED Box Orin NX 16GB und NVIDIA Jetson AGX Orin 64GB. Die zusätzliche Hardware erlaubt die Aufnahme von mehr Umgebungsdaten und erweitert generell die Datenverarbeitungsfähigkeiten des Systems.

##### **AP 1.2 – Erstellung eines Annotationsguides inklusive Klassen und Klassifizierungsmerkmale**

Von den Partnern wurde in einem mehrschleifigen Abstimmungsprozess ein 20-seitiger Annotationsguide erstellt. Dem voraus ging eine detaillierte Analyse der möglichen Anwendungsfälle des Datensatzes. Im Detail enthält der Annotationsguide folgende Informationen:

- Beschreibung des allgemeinen Kontextes des Guides
- Sensordaten, auf welche sich der Guide bezieht
- Beschreibung der zu benutzenden Annotationsgeometrien
- Definition der betrachteten IDs
- Beschreibung der relevanten Attribute
- Erläuterung der Sensorkonfiguration, unter welcher die zu annotierenden Daten aufgenommen wurden
- Sensorübergreifende Annotationsregeln
- Sensorspezifische Annotationsregeln

Klassenspezifische Annotationsregeln (inklusive zu nutzender Annotationsgeometrie + Auflistung aller Attribute mit Attributsausprägungen und Beispielbildern)

Im Annotationsguide werden die Klassen path, vehicle lane, open space, lane boundary, restricted areas sowie person, vehicle, two-wheeler, animal, sign, traffic light, vertical barrier und rumble strip betrachtet. Abbildung 2 zeigt exemplarisch einen Ausschnitt der Annotationsregeln für die Klasse vehicle.

Table 13: Geometry requirements of the object class "vehicle"

sensor	geometry
camera	2D Bounding Box
lidar	3D Bounding Box

Table 14: Attributes and attribute values of the object class "vehicle"

attribute	attribute value
type	car
	truck
	bus
	tram
	ambulance
	police
	fire truck
	other
unknown	
occlusion	0-33 %
	34-67 %
	68-99 %
	100 %



Figure 7: Annotation example of the object class "vehicle" with "type" = "car" and occlusion = "0-33 %"

Abbildung 2: Ausschnitt aus den Annotationsregeln

Der Annotationsguide wurde projektbegleitend weiterentwickelt, insbesondere wenn im Rahmen des Annotationsprozesses festgestellt wurde, dass die Annotationsregeln bestimmte Grenzfälle nicht abdeckten. Dies war zum Beispiel der Fall bei verkehrsberuhigten Bereichen, die sowohl von Fußgängern, Radfahrern und Autos genutzt werden konnten. Weiterhin wurden Definitionen nachgeschärft, um zum Beispiel klar herauszuarbeiten, dass die Deckfläche eines Bordsteins noch zum Rad- bzw. Fußweg gehört, die Seitenfläche aber nicht.

### **AP 1.3 Recherche verfügbarer Annotationtools und Methoden für manuelles und automatisiertes Labeln**

In AP 1.3 hat FS eine Übersicht der wichtigsten öffentlich zugänglichen Labelingtools angefertigt. Dabei unterscheiden sich die Angebote hauptsächlich in der Verfügbarkeit (Open Source oder kommerziell) sowie hinsichtlich des Umfangs (manuelles und/oder automatisches Labeln, Objektgeometrien, Attribute, Multi-Sensor-Settings, ...). Beispielsweise wurden die Tools LabelMe, VGG Image Annotator und CVAT (Computer Vision Annotation Tool) untersucht. Auf Basis der Recherche wurden Anforderungen an das FS Labeling Tool abgeleitet

und mit den Projekterfordernissen abgeglichen. Die resultierenden Anforderungen umfassen unter anderem folgende Features:

- Laden und Speichern von Multi-Sensor-Sequenzen (Sensordaten + Annotationen)
- Annotationsgeometrien im 2D (Bilder aus Kameras): Zweidimensionale Bounding Box, Zweidimensionale Polygone
- Implementierte Geometrien im 3D (Lidar-Punktwolken): Dreidimensionale Bounding Box Dreidimensionale Polygone
- Automatische Vergabe von Annotations- und Objekt-IDs
- Informationsfenster zu Attributen und Geometrieparametern
- Einfache Punktwolkenavigation, verschiedene Färbungsmodi der Punkte
- Relevante Attributarten: Single-Select, Multi-Select
- Vergabe von „default“-Werten für Attribute zur Fehlerminimierung
- Projektion der Annotationen von Punktwolken in 2D-Bilder bei gegebenen Kalibrierparametern
- Kopieren von Annotationen von einem Multi-Sensor-Frame (M-Frame) in den nächsten möglich
- Interpolation zwischen sogenannten Key-M-Frames
- Automatische Überprüfung der Annotationskonsistenz während und nach dem Annotationsvorgang
- Vielfältige Review-Funktionalitäten für effiziente Überprüfung der erstellten Annotationen durch interne und externe Reviewer
- Tracking des Bearbeitungszustandes („Annotation gestartet“, „Annotation abgeschlossen“, „1. Review gestartet“, „1. Review abgeschlossen“, ...)
- Anzeige einer Labelstatistik je Sequenz
- Export der Annotationsdaten ins ASAM OpenLABEL Format
- Verschiede Färbungsmodi für Annotationsgeometrien: nach Objektklassen, nach Objekt-ID oder nach Zusammengehörigkeit laut „connectedTo“-Attributen
- Zoom für hochgenaues Setzen der Annotationsgeometrien
- Übersichtsvideo zum besseren „Zurechtfinden“ in einer Sequenz
- Filteroptionen für die unterschiedlichen Objektklassen
- Automatische Konvertierung von Bilddaten in speicherarmes Format zur Minimierung von Ladezeiten

#### **AP 1.4 Architektur und Toolchain**

Im AP 1.4 wurde das Konzept zu Architektur und Toolchain des Labeling-Tools entworfen. Auf das Tool kann online über eine entsprechende Webseite zugegriffen werden. Im Tool werden die je Sequenz hinterlegten Sensordaten angezeigt. Es bietet die Möglichkeit, verschiedene Labels einzuzichnen (z.B. Bounding Boxes etc.) und mit weiteren Informationen zu hinterlegen (z.B. Klassifikation und Attribute). Anschließend werden die Daten zum Speichern an den Server geschickt. Die Kommunikation der Webseite ist immer HTTPS-verschlüsselt. Das Tool basiert auf Code-Bausteinen aus mehreren Programmiersprachen:

- HTML/CSS/JavaScript für das Frontend
- PHP und C# für die Kommunikationsebene mit dem Server (Validation, Speichern, ...)

- C++/Python für diverse Backendprozesse (Interpolation, Projektion, ...)

Die Sequenzen bestehen dabei aus sogenannten Multisensor-Frames (M-Frames), wobei ein Multi-Sensor-Frame aus allen einem Zeitpunkt zugeordneten Sensor-Frames besteht. Dieser schematische Aufbau ist in Abbildung 3 dargestellt.

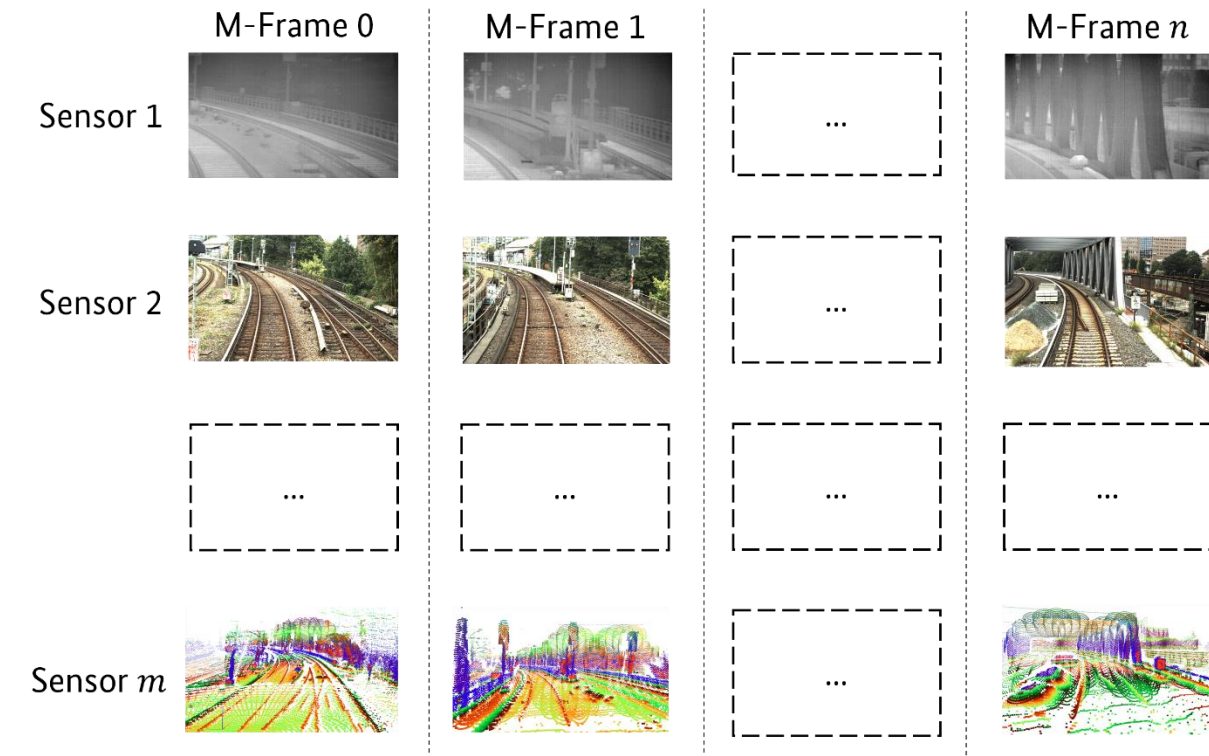


Abbildung 3: Aufbau einer Sequenz

Der Output des Labeling-Tools ist eine Annotationsdatei, die Eckpunkte der Annotationsgeometrien, die zugehörigen Attribute und auch die Objekt- und Annotations-IDs enthält.

### **AP 1.5 Referenzmodelle für die Objekterkennung und Verhaltensprädiktion**

Im AP 1.5. hat die HoMe den Part der Verhaltensprädiktion, der Partner FS die Arbeiten zur Objekterkennung übernommen. Es wurden durch FS viele verschiedene Modelle zur Objekterkennung evaluiert, darunter:

1. Region-based Convolutional Neural Networks (R-CNN): Diese Modelle verwenden Region Proposal Networks (RPNs), um potenzielle Regionen zu identifizieren, in denen sich Objekte befinden könnten, und verwenden dann CNNs, um die Objekte in diesen Regionen zu erkennen.
2. Fast R-CNN: Eine Weiterentwicklung von R-CNN, die den Prozess der Objekterkennung beschleunigt, indem sie die Region Proposal und die Objekterkennung in einem Schritt kombiniert.

Für die Verhaltensprädiktion wurden ebenfalls unterschiedliche potenzielle Referenzmodelle identifiziert und gegeneinander verglichen. Dabei wurden verschiedene Modellarchitekturen untersucht, darunter klassische Ansätze wie lineare Modelle sowie neuronale Netze,

insbesondere rekurrente Netzwerke (LSTMs) und Transformer-basierte Modelle. Im Ergebnis wurde eine methodische Pipeline identifiziert, die sich an gängigen Verfahren des maschinellen Lernens für zeitabhängige Vorhersagen orientiert.

### **AP 1.6 Datensatzbereitstellung – Veröffentlichung**

Die Veröffentlichung der annotierten Sensordaten erfordert eine entsprechende Infrastruktur. Um große Datenmengen speichern und verwalten zu können, wurde als Ergebnis einer Recherche der Amazon Simple Storage Service (Amazon S3) ins Auge gefasst. Dabei handelt es sich um einen Objektspeicher-Service mit branchenführender Skalierbarkeit, Datenverfügbarkeit, Sicherheit und Leistung. Amazon S3 speichert Daten als Objekte innerhalb von Buckets. Ein Objekt besteht aus einer Datei und optional aus Metadaten, die diese Datei beschreiben. Der Zugriff erfolgt über HTTP bzw. HTTPS. S3FS ist ein FUSE-Dateisystem, mit dem man einen Amazon S3-Bucket lokal einbinden kann. Dies wird von FusionSystems zusammen mit einem AWS-Client für die Verbindung zum objektbasierten Speicher genutzt. Die Verbindung wird über Access Keys, also ein Schlüsselsystem geregelt.

Dies bietet den Vorteil der Flexibilität aufgrund der Tatsache, dass FusionSystems keinen Wartungsaufwand für den Speicher hat, da er in der Cloud liegt. Weiterhin gibt es keine Begrenzung für die Datenmenge, wobei sich die Kosten aus der Menge des genutzten Speichers und dem ausgehenden Datentransfer zusammensetzen. Zudem verfügt die Cloud über eine Hochverfügbarkeit sowie eine bessere Performance-Anbindung für Nutzer außerhalb des FS-Netzes als physische Server bei FS und sie erfüllt die Anforderungen der DSGVO.

### **AP 1.7 Automatische Plausibilitätsüberprüfungen**

Zur Unterstützung des Reviews können verschiedene Verfahren zur automatischen Konsistenzprüfung angewendet werden. Im Projekt wurden folgende Verfahren implementiert:

1. Überprüfung der Plausibilität der Abmessungen der 3D-Bounding Boxes: Bei dieser Konsistenzüberprüfung wird die Plausibilität der Abmessungen der 3D-Bounding Boxes 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 Tiefe der Objekte definiert und geprüft werden.
2. 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 **Fehler! Verweisquelle konnte nicht gefunden werden.** dargestellt. In a) befinden sich die 2D-BB-Annotationen (2D Bounding Boxes) mit der entsprechenden Tracking-ID (in Rot) für einen Frame, in Abbildung b) die 2D-BB-Annotationen mit Tracking-ID für den Nachfolger-Frame (in Blau). Eine Assoziation über die ungarische Methode erlaubt es, 2D-BBs einander zuzuordnen, die sich in räumlicher Nähe zueinander befinden, wie in Abbildung c) dargestellt. In der Abbildung können 2D-BB  $b_1$  und  $b_4$ ,  $b_2$  und  $b_5$  sowie  $b_3$  und  $b_6$  einander zugeordnet werden. Anschließend kann überprüft werden, ob zugeordnete

2D-BBs 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.

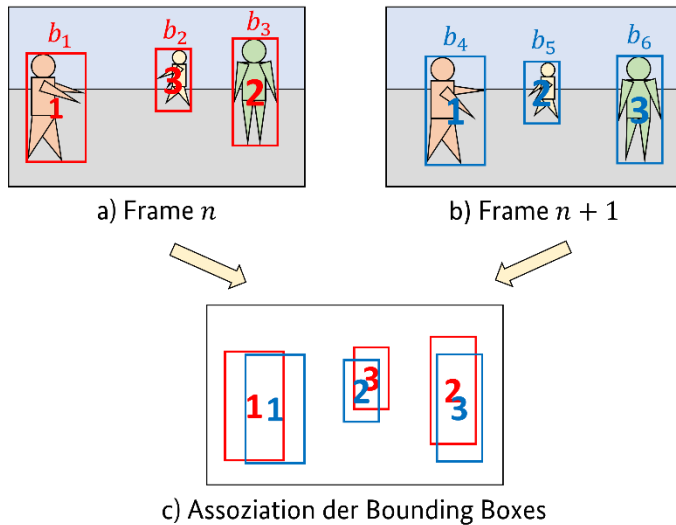


Abbildung 4: Assoziation von Bounding Boxes

- 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 der „type“ eines „vehicles“ nicht von einem Frame zum anderen oder von Sensor zu Sensor ändern. Ein „connectedObject“ einer „person“ kann sich zwar mit der Zeit ändern (wenn die Person beispielsweise das Objekt weglagt), allerdings ist diese zu einem bestimmten Zeitpunkt für alle Sensoren gleich. Die „occlusion“ 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. Entsprechend der Zuordnung kann die Konsistenz von Attributsausprägungen überprüft werden.

### 3.1.2 AP 2 Aufbau Sensorik

#### **AP 2.1 – Feinplanung Sensorik und Hardware**

Zur detaillierten Planung der Sensorik wurde das separate Dokument „Sensorkonzept“ erstellt. Dieses Dokument enthält Informationen zu allen verbauten Sensoren, wie etwa Sensorprinzip und Verwendung, Hardware-Konfiguration, Software-Konfiguration,

Anbauposition und Erfassungsbereich. Der Erfassungsbereich des Gesamtsystems ist in der Abbildung 5 dargestellt.

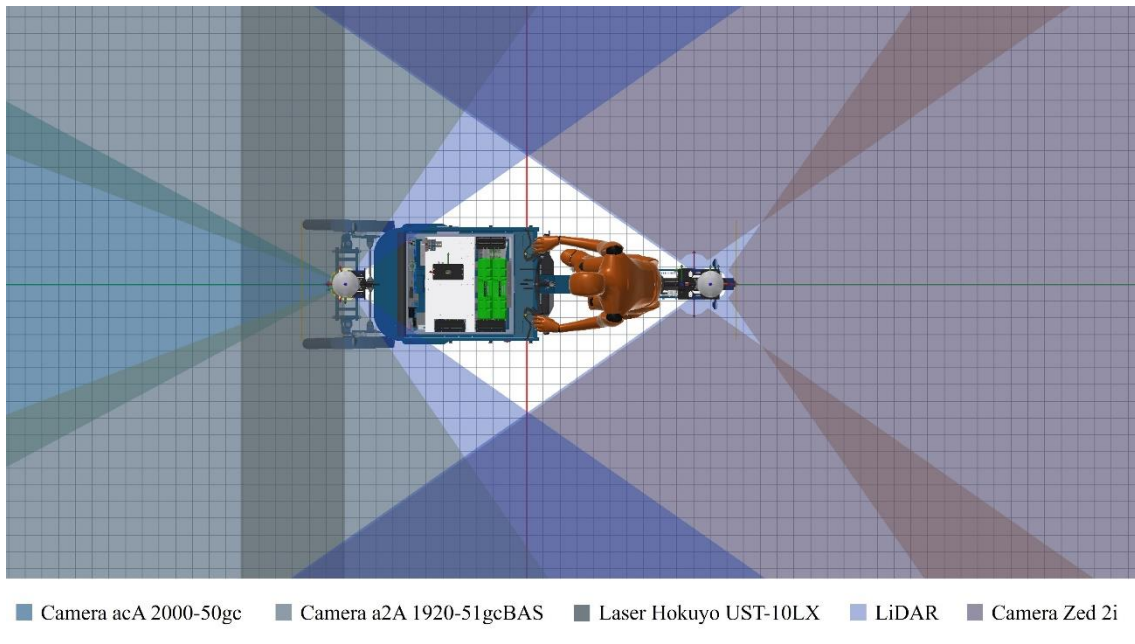


Abbildung 5: Erfassungsbereich der Sensoren auf dem Sensorträger Eigen

## **UAP 2.2 – Integration und Test-Sensorik und Rechenhardware in Versuchsträger**

Alle Sensoren wurden erfolgreich in den Sensorträger integriert und deren Funktionsfähigkeit mit dem ROS-Tool Rviz2 getestet. RGB-Bilder wurden erfolgreich durch drei Basler-Kameras aufgenommen. Die zwei verbauten 360-Grad-Lidare nehmen Punktwolken auf, die erfolgreich zu einer gemeinsamen Point-Cloud fusioniert werden konnten, die den gesamten Bereich um den Versuchsträger abdeckt. Die Funktionalität der Zed-2i-Kameras konnte erfolgreich evaluiert werden, inklusive ihrer Fähigkeit Tiefeninformationen oder hoch verarbeitete Daten etwa zum „Skeleton Tracking“ bereitzustellen.

Des Weiteren erfolgte gemeinsam mit dem Partner FS eine Definition der relevanten aufzunehmenden Informationen sowie eine einheitliche Benennung der ROS-Topics. In diesem Zuge wurde das Gesamtsystem von ROS auf ROS2 migriert, was einen erheblichen Mehraufwand im Vergleich zum ursprünglichen Antrag zur Folge hatte, das Gesamtsystem aber robuster und zukunftsfähiger macht.

Die folgenden Abbildungen zeigen einige Beispiele der erfolgreichen Sensortests.

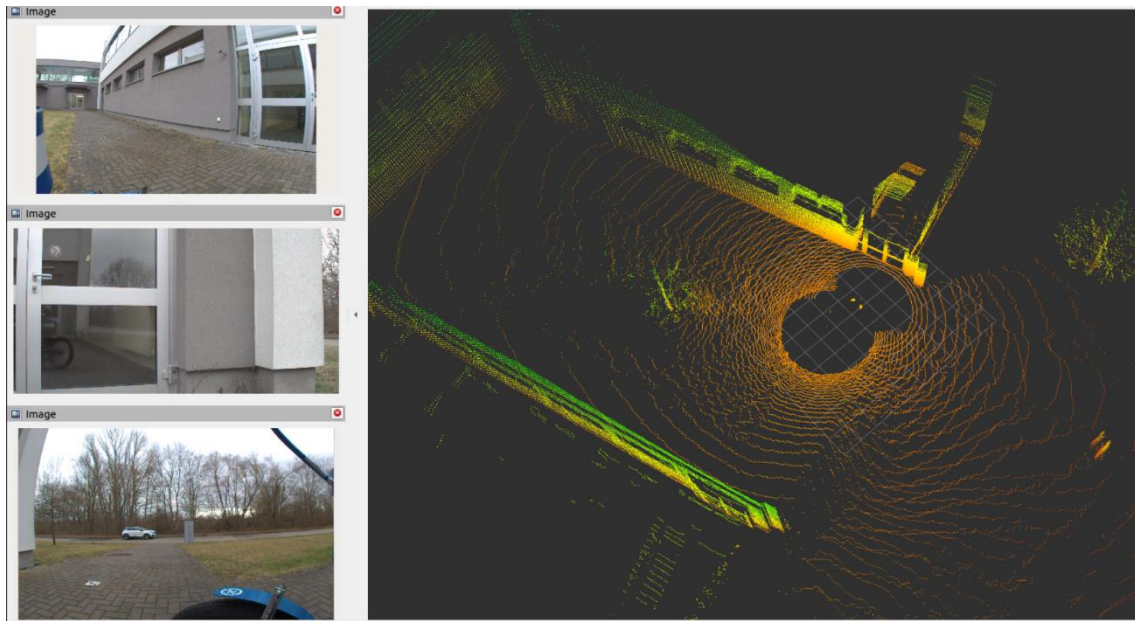


Abbildung 6: Erprobung der Basler Kameras (links) und der Fusionierten Punktwolke (rechts)

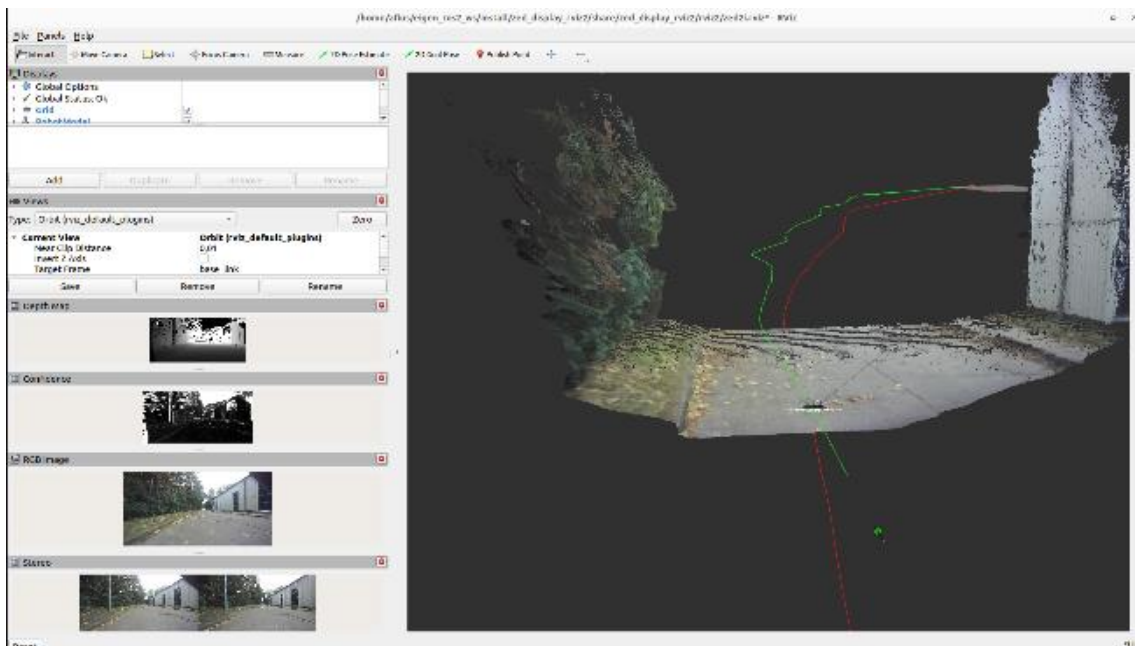


Abbildung 7: Erprobung Zed-2i-Kamera

### **UAP 2.3 – Kalibrierung der Sensorik**

Die drei Basler-Kameras wurden umfassend kalibriert, um das bestmögliche Bild unter den jeweiligen Arbeitsbedingungen zu erzeugen. Für die Zwecke dieses Projekts wurde die Bildqualität durch die folgenden Parameter definiert: Helligkeit, Schärfe, Bildgröße, Farbgenauigkeit, Bewegungsunschärfe und Framerate. Der Kalibrierungsprozess bestand aus drei Schritten: Kalibrierung der Blende jedes Objektivs, Kalibrierung des Fokusbereichs jedes Objektivs (siehe Abbildung 8) und Kalibrierung der Aufnahmeparameter jeder Kamera. Um eine optimale Bildqualität bei Tag- und Nachtaufnahmen zu erreichen, wurde für jede Kamera ein eigenes Kalibrierungsprofil erstellt. Eine intrinsische Kalibrierung der Kameraparameter

wurde durch die Analyse von Bildern eines Schachbretts mit einem auf der OpenCV-Bibliothek basierenden Programm durchgeführt (siehe Abbildung 9).



Abbildung 8: Kalibrierung des Fokus der Basler Kameras

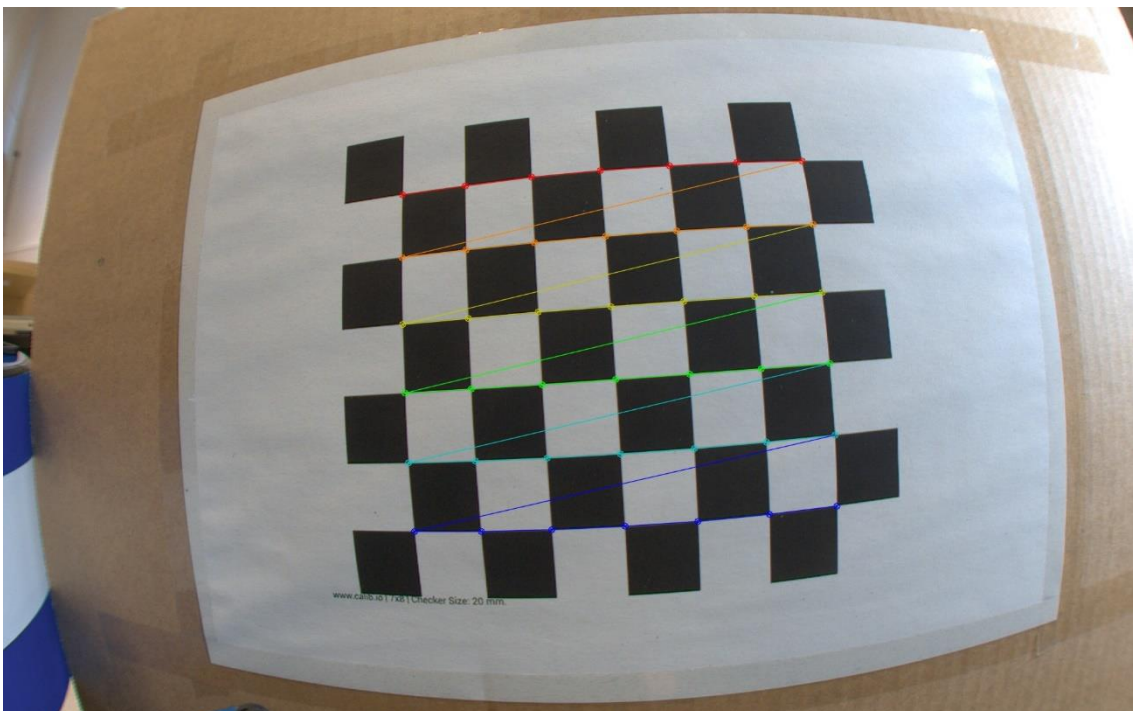


Abbildung 9: Kalibrierung der intrinsischen Parameter des Kamerasystems

### 3.1.3 AP 3 Labeling- Tool

#### **AP 3.1 Anpassung GUI und Softwarearchitektur**

Die GUI des FS Labeling Tools ist projektunabhängig konzipiert und wurde auf die Projektspezifika angepasst. Dazu wurden Details wie der Aufbau der Sequenzen und die verschiedenen genutzten Sensoren berücksichtigt. Weiterhin können Einschränkungen der zu annotierenden Multisensor-Frames getroffen und die möglichen Objektklassen sowie die dazugehörigen Attribute inklusive möglicher Attributsausprägungen festgelegt werden. Das Format und der Inhalt der zu exportierenden Dateien wurde ebenso gemäß der Projektabsprachen festgelegt (siehe AP 3.5). All diese Parameter lassen sich im FS Labeling Tool mittels einer Konfigurationsdatei für jede Sequenz einstellen.

Abbildung 10 gibt einen Eindruck über die Nutzeroberfläche des Tools.

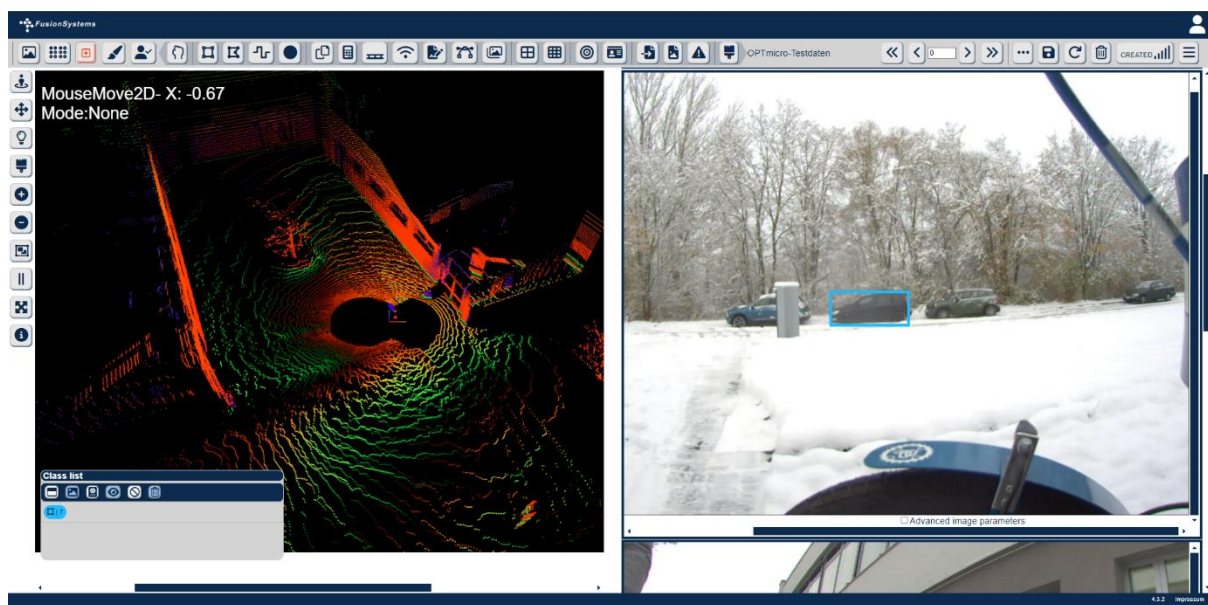


Abbildung 10: Nutzeroberfläche des Labeling-Tools

#### **AP 3.2 Sensor-zu-Sensor-Projektionen**

Die Projektion von Annotationen aus dem Lidar in die Kameras wurde implementiert und erfolgreich getestet. Dabei werden folgende Schritte nacheinander ausgeführt:

1. Transformation vom Lidar- in das optische Kamerakoordinatensystem:

$$\begin{pmatrix} x'_n \\ y'_n \\ z'_n \end{pmatrix} = R^{-1} \cdot \left( \begin{pmatrix} x_B \\ y_B \\ z_B \end{pmatrix} - \vec{t} \right) \text{ mit } \vec{t} = \begin{pmatrix} d_x \\ d_y \\ d_z \end{pmatrix} \text{ und}$$
$$R = \begin{pmatrix} 1 - 2q_y^2 - 2q_z^2 & 2q_xq_y - 2\omega q_z & 2q_xq_z + 2\omega q_y \\ 2q_xq_y + 2\omega q_z & 1 - 2q_x^2 - 2q_z^2 & 2q_yq_z - 2\omega q_x \\ 2q_xq_z - 2\omega q_y & 2q_yq_z + 2\omega q_x & 1 - 2q_x^2 - 2q_y^2 \end{pmatrix}$$

2. Anwendung der distortion coefficients:

$$\begin{pmatrix} x_n''' \\ y_n''' \end{pmatrix} = \begin{pmatrix} x_n''(1 + k_1 r_n^2 + k_2 r_n^4 + k_3 r_n^6) + 2p_1 x_n'' y_n'' + p_2(r_n^2 + 2(x_n'')^2) \\ y_n''(1 + k_1 r_n^2 + k_2 r_n^4 + k_3 r_n^6) + 2p_2 x_n'' y_n'' + p_1(r_n^2 + 2(y_n'')^2) \end{pmatrix} \text{ mit}$$

$$\begin{pmatrix} x_n'' \\ y_n'' \end{pmatrix} = \begin{pmatrix} x_n' \\ y_n' \\ z_n' \end{pmatrix} \text{ and } r_n^2 = (x_n'')^2 + (y_n'')^2$$

3. Transformation in das Bildkoordinatensystem:

$$\begin{pmatrix} x_n \\ y_n \\ 1 \end{pmatrix} = C \cdot \begin{pmatrix} x_n''' \\ y_n''' \\ 1 \end{pmatrix} \text{ mit } C = \begin{pmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{pmatrix}$$

Hierbei ist  $\begin{pmatrix} x_B \\ y_B \\ z_B \end{pmatrix}$  der Ortsvektor eines Punktes im Lidarkoordinatensystem,  $\begin{pmatrix} x_n' \\ y_n' \\ z_n' \end{pmatrix}$  der

Ortsvektor eines Punktes im optischen Kamerakoordinatensystem,  $\vec{t} = \begin{pmatrix} d_x \\ d_y \\ d_z \end{pmatrix}$  der

Translationsvektor zwischen den beiden Koordinatensystemen,  $R$  die Rotationsmatrix

zwischen den beiden Koordinatensystemen (mit dem Rotationsquaternion  $\vec{q} = \begin{pmatrix} q_x \\ q_y \\ q_z \\ \omega \end{pmatrix}$ ),

$k_1, k_2, k_3, p_1, p_2$  die distortion coefficients,  $\begin{pmatrix} x_n \\ y_n \end{pmatrix}$  der Ortsvektor eines Punktes im Bildkoordinatensystem und  $C$  die Kameramatrix mit den Brennweiten  $f_x, f_y$  und dem Bildhauptpunkt  $(c_x, c_y)$ .  $\vec{t}, R$  und  $C$  wurden durch eine Kalibrierung ermittelt.

### **AP 3.3 Teilautomatische Erzeugung von Annotationen**

Die Annotation von Sensordaten ist ein aufwendiger Prozess, deswegen ist die teilautomatisierte Erstellung von Annotationen wichtig, um einen möglichst effizienten Prozess zu ermöglichen. Im Projekt wurden folgende Prozesse implementiert und erfolgreich getestet:

1. 3D-zu-2D-Projektionen: Die Annotationen im Lidar werden mittels extrinsischer und intrinsischer Kalibrierparameter aus dem Lidar-Koordinatensystem in die Kamera-Koordinatensysteme projiziert.
2. Zeitliche Interpolation: Die starke zeitliche Korrelation aufeinanderfolgender M-Frames (Multisensor-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 (kubische Splines) zu erzeugen und gegebenenfalls manuell zu korrigieren.
3. Vorhersage mittels KI-Verfahren: Wenn ein gewisser Prozentsatz der Sensordaten annotiert ist, dann kann dieser als Lerndatensatz für einen Detektionsalgorithmus (welcher auf Neuronalen Netzen basiert) genutzt werden, welcher auf die noch unannotierten Sensordaten angewendet wird. Der Detektor macht so Vorschläge, die vom Labeler noch angepasst oder gelöscht werden müssen.

### AP 3.4 Automatische Überprüfung der Annotationen

Die in AP 1.7 entwickelten Verfahren wurden implementiert und erfolgreich getestet. Dazu wurden folgende Schritte durchgeführt:

1. Testplanung: In diesem Schritt wurden die Testziele und -strategien definiert. Testziel war die Ermittlung der Falsch-Positiv- bzw. Falsch-Negativ-Rate der drei entwickelten Konsistenzüberprüfungen. Zudem wurde ein Zeitplan für die Testaktivitäten erstellt.
2. Testfallentwurf: In diesem Schritt wurden detaillierte Testfälle erstellt, die verschiedene Szenarien abdeckten, einschließlich bewusst gesetzter Annotationsfehler. Die erwarteten Ergebnisse für jeden Testfall wurden ebenfalls dokumentiert.
3. Testdurchführung: Die Tests wurden gemäß den definierten Testfällen durchgeführt, indem im Labeling Tool die automatische Überprüfung gestartet wurde. Die Testergebnisse (korrekt gefundene Annotationsfehler, nicht gefundene Annotationsfehler, falsch als Annotationsfehler klassifizierte Annotationen) sowie alle aufgetretenen Software-Fehler (Absturz des Softwaremoduls) wurden dokumentiert.
4. Fehlerberichterstattung: Alle gefundenen Software-Fehler wurden in einem Fehlerverfolgungssystem erfasst. Detaillierte Informationen zu jedem Fehler, einschließlich der Schritte zur Reproduktion, wurden bereitgestellt.
5. Regressionstests: Regressionstests wurden durchgeführt, um sicherzustellen, dass neue Änderungen oder Fehlerbehebungen keine bestehenden Funktionen beeinträchtigen.
6. Testauswertung: Die Testergebnisse wurden analysiert und die Qualität der Software bewertet. Ein Testbericht wurde erstellt, der die Ergebnisse sowie die gefundenen Software-Fehler zusammenfasst.
7. Inkrementelle Anpassung des Codes: Basierend auf der Testauswertung wurde der Code so angepasst, dass die Software-Fehler beseitigt wurden und die Falsch-Positiv sowie die Falsch-Negative-Rate der erkannten Annotationsfehler gesenkt werden konnte. Dann wurden die Schritte 3 bis 7 wiederholt, bis keine Software-Fehler mehr auftraten und die Raten konvergierten.

Als Ergebnis des Testes konnten folgende Raten ermittelt werden:

<b>Automatische Prüfung</b>	<b>Falsch-Positiv-Rate</b>	<b>Falsch-Negativ-Rate</b>	<b>Begründung</b>
Abmessungen 3D Bounding Boxes	1 %	15 %	FP: Es kann zu Extremfällen bei den Abmessungen kommen (z.B. ist ein Lastenrad sehr lang für ein Fahrrad), sodass diese außerhalb der angedachten Grenzen liegen. FN: Wenn die Abmessungen eines Objektes innerhalb des physikalisch plausiblen Bereichs liegen, kann nicht direkt geschlussfolgert werden, ob diese tatsächlich passend (im Sinne von eng anliegend) sind

Tracking-ID	0,1 %	0,1 %	Wenn sehr viele Objekte sehr nah beieinander liegen, dann kann es vorkommen, dass die Assoziation der Bounding Boxes nicht korrekt erfolgt.
Konsistenz der Attribute	0 %	0 %	Verletzungen der Attributkonsistenz können stets eindeutig festgestellt werden.

### **AP 3.5 Export Annotationsdaten**

Im Projekt wurde sich auf ASAM OpenLABEL

(<https://www.asam.net/standards/detail/openlabel/>) als Exportformat für die

Annotationsdaten geeignet. Dieses zeichnet sich unter anderem durch folgende Merkmale aus:

1. Standardisierung: OpenLABEL ist ein offenes und standardisiertes Format, das von der ASAM (Association for Standardization of Automation and Measuring Systems) entwickelt wurde. Dies fördert die Interoperabilität zwischen verschiedenen Systemen und Anwendungen.
2. Flexibilität: Das Format ist flexibel und anpassbar, sodass es für verschiedene Anwendungsfälle im Mobilitätsbereich, insbesondere im Bereich der Datenannotation für maschinelles Lernen und autonome Systeme, verwendet werden kann.
3. Unterstützung für verschiedene Datentypen: OpenLABEL ermöglicht die Annotation von verschiedenen Sensordatentypen, einschließlich Bild-, Video- und LiDAR-Daten.
4. Einfache Integration: Das Format ist so konzipiert, dass es leicht in bestehende Workflows und Tools integriert werden kann, was die Effizienz bei der Datenverarbeitung und -analyse erhöht.
5. Rich Metadata: OpenLABEL unterstützt die Verwendung von umfangreichen Metadaten, die zusätzliche Informationen über die annotierten Daten bereitstellen. Dies kann die Qualität der Datenanalyse und das Training von Algorithmen verbessern.
6. Community-getrieben: Da es sich um ein offenes Format handelt, wird es von einer aktiven Community unterstützt, die kontinuierlich an der Weiterentwicklung und Verbesserung des Formats arbeitet.
7. Zukunftssicherheit: Durch die Standardisierung und die Unterstützung durch die ASAM ist das OpenLABEL Format zukunftssicher und wird voraussichtlich auch in den kommenden Jahren relevant bleiben.

Die Exportfunktionalität des Labeling Tools wurde so angepasst, dass es alle Anforderungen des ASAM OpenLABEL Formats erfüllt. Über einen zusätzlich implementierten Validator wurde die Korrektheit der exportierten Daten verifiziert.

#### 3.1.4 AP 4 Datenerhebung

### **AP 4.1 Planung – Auswahl Operationsgebiet und -zeitraum**

Drei der grundlegenden Anforderungen an KI-Datensätze sind Größe, Vielfalt und Ausgewogenheit. Die Größe des Datasets beeinflusst die Leistung des trainierten Modells. Eine unzureichende Größe kann zu einem über- oder unterangepassten Modell führen. Der

Datensatz muss zudem vielfältig sein, d.h. er muss Beispiele für verschiedene Objekte, Umgebungsbedingungen und Szenarien enthalten, damit das Modell in der Lage ist, gut auf unbekannte Daten zu verallgemeinern. Die Ausgewogenheit der Datensätze, d.h. die Datenmenge für jede Klasse, ist entscheidend, um das Training eines verzerrten Modells zu verhindern.

Um diesen Anforderungen gerecht zu werden, wurde ein Aufnahmekonzept weiterentwickelt dieses umfasst nun unter anderem sieben Aufnahmestrecken in der Stadt Merseburg. Die Aufzeichnungsrouten wurden speziell für die Zwecke autonomer Kleinstfahrzeuge ausgewählt, um gemeinsame Betriebsbereiche wie Fahrrad- und Fußwege einzubeziehen. Eine kombinierte Karte dieser Routen findet sich in Abbildung 11.

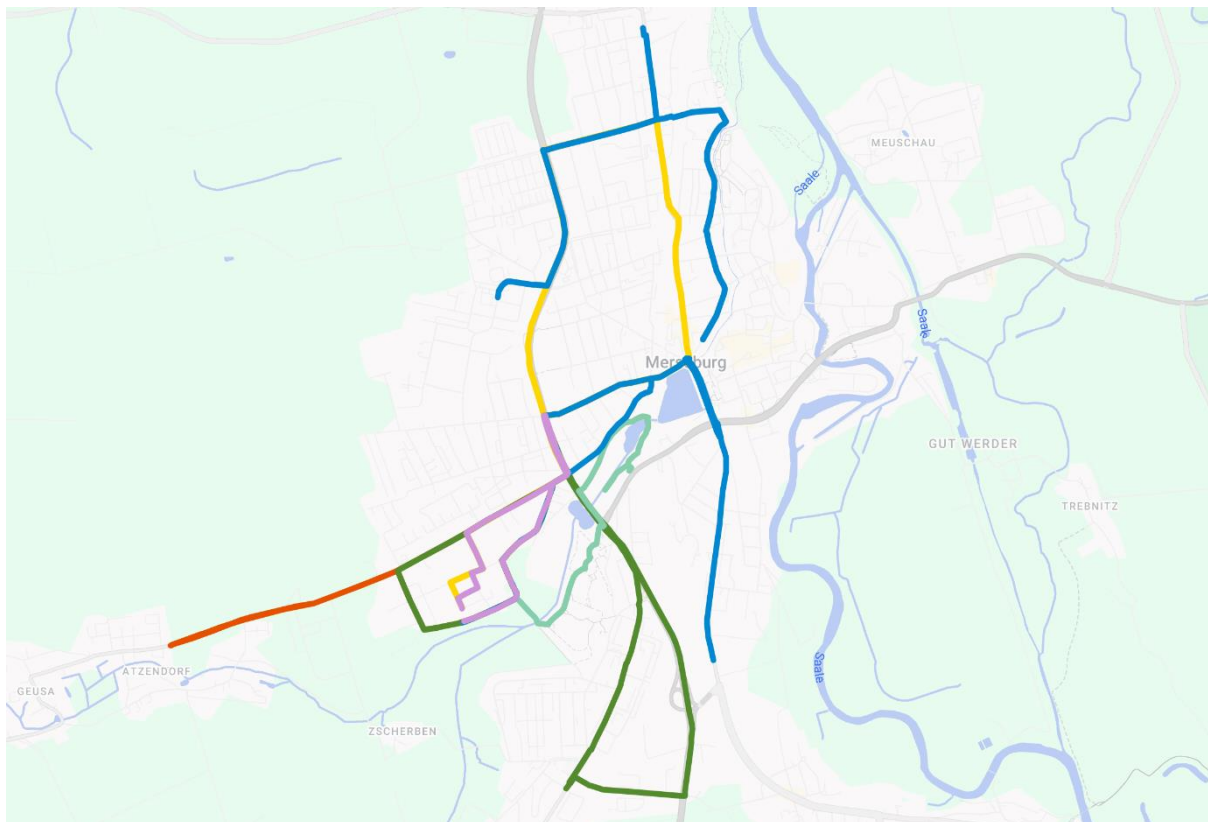


Abbildung 11: Aufnahmerouten in Merseburg

Der geplante Erfassungszeitraum betrug über ein Jahr, um die unterschiedlichen Wetterbedingungen und Umweltveränderungen insbesondere die vier Jahreszeiten abzubilden. Um den OPTmicro-Datensatz weiter zu diversifizieren, wurde die gleiche Aufnahmestrategie auf die Städte Köthen, Schkeuditz, Freiberg und Altenau angewandt. Im Speziellen wurde Altenau im Harz ausgewählt, da so die Möglichkeit zur Aufzeichnung von Winterdaten möglich war.

Zur Kategorisierung der Daten wurde eine Strategie entwickelt, bei der jeder aufgenommenen Sequenz eine Wertzahl zugewiesen wurde. Sequenzen, die besondere Ereignisse oder Objekte wie Polizeifahrzeuge, Krankenwagen oder Feuerwehrautos enthalten, erhielten eine höhere Wertzahl als Sequenzen mit wenig bis gar keiner Objektvielfalt. Die Bewertung ermöglichte es, Sequenzen zu priorisieren, die für den OPTmicro-Datensatz besonders relevant oder

wertvoll waren. Um die Vielfalt und Ausgewogenheit des Datensatzes zu garantieren wurden für jede Aufnahme Saison, Datum, Ort, Wetterbedingungen und Uhrzeit der Aufzeichnung in einer separaten Aufzeichnungsauswertungsdatei dokumentiert. Die Informationen wurden dann in einer kombinierten Excel-Pivot-Tabelle zusammengefasst, als Ganzes analysiert und eine Annotationsreihenfolge festgelegt.

#### **AP 4.2 Durchführung der Datenerhebung**

Als Ergebnis der in AP 4.1 beschriebenen Aufzeichnungsstrategie wurden 1.191 Minuten Rohdaten mit einer Gesamtgröße von 10.234 GB aufgezeichnet.

Um die Aufzeichnung all dieser Daten zu erleichtern, musste der Hauptrechner des Sensorträgers mit einem neuen Intel X550-T2-Netzwerkadapter aufgerüstet werden, der Übertragungsgeschwindigkeiten von bis zu 10 Gbit/s ermöglichte, und einer neuen Samsung 990 Pro NVME SSD, die 4 TB Speicher mit einer Lesegeschwindigkeit von bis zu 7450 MB/s und einer Schreibgeschwindigkeit von bis zu 6900 MB/s zur Verfügung stellte. Diese Upgrades verbesserten die Aufzeichnungsfunktionen des Systems. Des Weiteren wurde ein Lenovo Tab M8 Tablet am Lenker des Sensorträgers verbaut. Das Tablet wurde vom Bediener des Sensorträgers verwendet, um Aufzeichnungen zu starten und zu stoppen, den Status der Sensoren zu überwachen, den Status der aktiven Aufzeichnung zu überwachen sowie zu Navigationszwecken. Die im Aufzeichnungskonzept definierten Routen wurden mit Google Maps erstellt und auf das Tablet geladen, so dass der Bediener sie wie geplant verfolgen konnte. Eine Anleitung zur Bedienung des Fahrrades und zur Durchführung einer Recording-Session wurden dokumentiert und für das Training der Hilfwissenschaftler verwendet.

Wie in AP 4.1 geplant, wurden verschiedene Orte und Objekte unter unterschiedlichen Wetterbedingungen aufgenommen, siehe Abbildung 12.

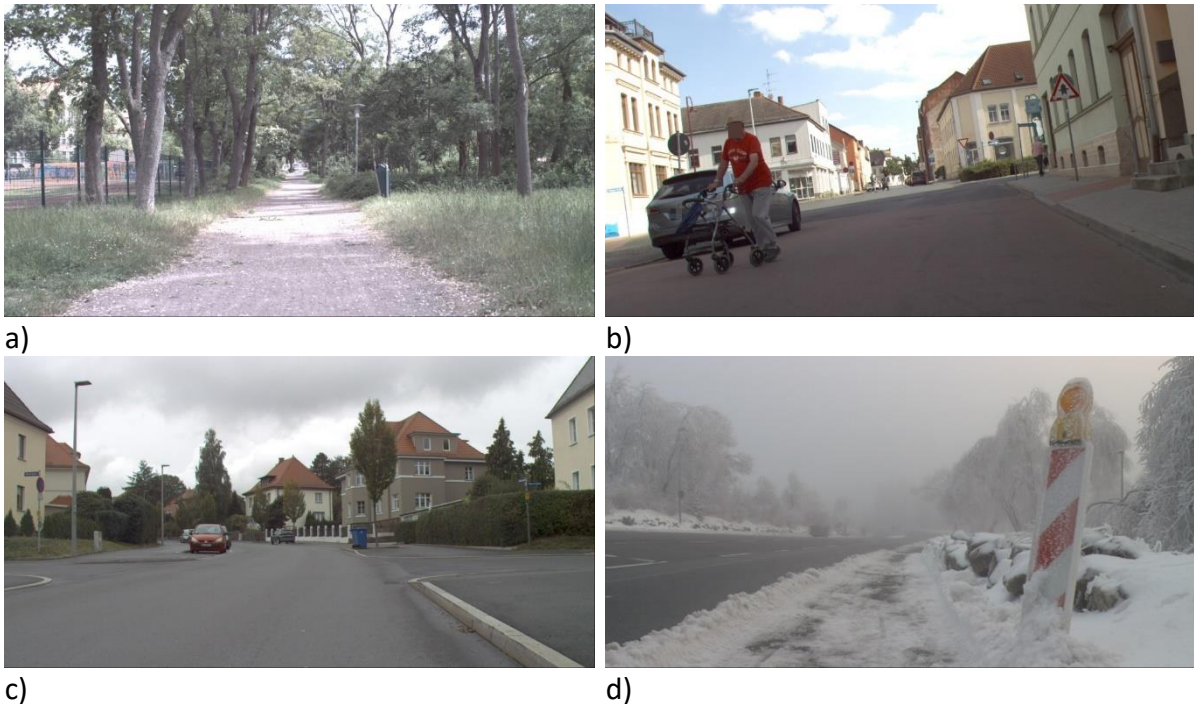


Abbildung 12: Bilder aus dem OPTmicro-Datensatz, aufgenommen in verschiedenen Jahreszeiten und Städten: (a) Frühling in Merseburg, (b) Sommer in Köthen, (c) Herbst in Freiberg und (d) Winter in Altenau

Möglich wurden die Aufnahmen, die auch in anderen Städten als Merseburg durchgeführt wurden, durch einen Anhänger, mit dem der Sensorträger sicher und effizient transportiert konnte.

Um den Aufzeichnungsprozess zu erleichtern, wurde eine Webanwendung entwickelt und auf dem Hauptrechner des Sensorträgers bereitgestellt, die dem Bediener eine grafische Benutzeroberfläche bot, mit der dieser eine Aufzeichnung starten und stoppen, die zu erfassenden ROS-Topics auswählen, den Status der Topics und den Status der aktiven Aufzeichnung überwachen und die Länge der Sequenzen für jede Aufzeichnung auswählen konnte. Der Vorteil der Erstellung der GUI für die Aufzeichnungssoftware als Webanwendung besteht darin, dass sie aus der Ferne über einen üblichen Webbrowser geöffnet werden kann, wobei die einzige Voraussetzung darin besteht, dass sich das Endgerät im privaten WLAN-Netzwerk des Sensorträgers angemeldet.

#### **AP 4.3 Datentransfer und Überprüfung der Datenerhebung**

Aufgrund der großen Größe der aufgezeichneten Daten (über 10 TB) mussten die Daten aus dem lokalen Speicher des Sensorträgers auf einen NAS-Server übertragen werden, der im Local Area Network der Hochschule Merseburg eingerichtet wurde. Der sichere Zugriff auf den NAS-Server erfolgte über eine Ethernet-Verbindung nach Übermittlung der erforderlichen Anmeldedaten.

Die Überprüfung der gesammelten Daten erfolgte manuell. Statistische Daten wie die Größe der Rohdaten in GB, die aufgezeichnete Zeit, die Anzahl der Sequenzen und das Datum der Aufzeichnung wurden in einem Tabellendokument gesammelt. Jede Sequenz wurde auch manuell ausgewertet und entsprechend dem in AP 4.1 beschriebenen Verfahren priorisiert und bewertet.

#### 3.1.5 AP 5 Datenaufbereitung

##### **AP 5.1 Aufbereitung der Sensordaten**

In AP 5.1 wurde durch die Projektpartner verschiedene Verfahren zur Datenaufbereitung und -verbesserung umgesetzt. Die Sensordaten wurden in 150-Sekunden-Intervallen aufgezeichnet, um die Dateigrößen zu limitieren und die Verarbeitung zu erleichtern, sowie als ROS2-Bagdateien (.db3-Format) gespeichert. Nach der zeitlichen Synchronisation besteht ein weiterer wichtiger Schritt im Extraktionsprozess im Data Parsing, bei dem Rohnachrichten, die aus ROS2-Bag-Dateien (.db3-Format) abgerufen werden, in strukturierte und manipulierbare Datenformate deserialisiert. Die können dann für die weitere Verarbeitung, Analyse und Annotation verwendet werden können. Der Deserialisierungsschritt zielt speziell auf mehrere ROS2-Topics ab, die unterschiedlichen Sensoren entsprechen. Jedes deserialisierte Topic wird in einem bestimmten Dateiformat gespeichert, abhängig von den Daten und ihrer beabsichtigten Verwendung in nachfolgenden Verarbeitungs- und Anmerkungsphasen. Nachfolgend findet sich eine detaillierte Zuordnung der ROS2-Themen zu ihrem jeweiligen gespeicherten Format:

*Tabelle 1 ROS2-Topics und ihr jeweiliges gespeichertes Format*

---

Sensor-Typ	Art der Meldung	Gespeichertes Format
------------	-----------------	----------------------

---

Kamera	sensor_msgs/Bild	JPEG-Bild (.jpg)
Lidar	sensor_msgs/Punktwolke2	PCD-Datei (.pcd)

Die Daten werden dann in einer hierarchischen Verzeichnisstruktur organisiert, um eine einfache Navigation und ein effizientes Datenmanagement zu ermöglichen, siehe Abbildung 13.

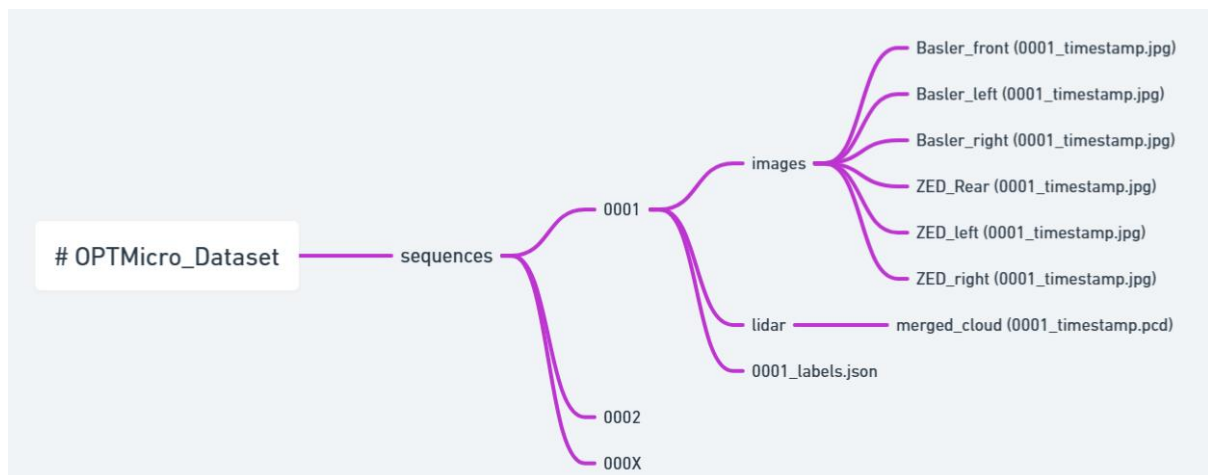


Abbildung 13: Struktur des OPTmicro-Datensatzes

Jede Aufnahme besteht aus mehreren 150-Sekunden-Sequenzen, die als 0001, 0002 usw. indiziert sind. Die Anzahl der Sequenzen hängt von der Länge der Aufnahme ab. Jede Sequenz enthält die synchronisierten Bilder und LiDAR-Daten.

Um die Ladegeschwindigkeit und Verarbeitungsgeschwindigkeit in der Erfassungs- und Annotationsphase zu erhöhen und die Produktivität des Annotators zu steigern, wurden die Bilder der Frontkamera auf  $1024 \times 540$  Pixel verkleinert, während die linke und rechte Kamera auf  $960 \times 600$  Pixel verkleinert wurden. Zusätzlich wurden die LiDAR-Punktwolken auf eine horizontale Auflösung von 512 Punkten pro Scan eingestellt. Diese Anpassungen waren entscheidend für die Verwaltung der Speicheranforderungen, ohne wesentliche Details zu verlieren.

Um die Datenschutzbestimmungen einzuhalten und die Privatsphäre des Einzelnen zu gewährleisten, wurden die Bilder im Datensatz einem zweistufigen Anonymisierungsprozess unterzogen, bei dem Gesichter und Kfz-Kennzeichen unkenntlich gemacht wurden. Die Notwendigkeit einer Anonymisierung ergab sich aus einem Rechtsgutachten zur datenschutzkonformen Sensordatenaufzeichnung und -speicherung.

In der ersten Phase wurden KI-basierte Algorithmen eingesetzt, um die Bilder zu durchschauen und Gesichter und Kfz-Kennzeichen zu finden. Das InsightFace-System wurde verwendet, um Gesichter zu finden. Es verwendet tiefe Convolutional Neural Networks (CNNs), um Teile des Gesichts zu finden. Das Modell verarbeitete jedes Bild und gab einen Satz von Begrenzungsrahmen aus, wobei jeder Begrenzungsrahmen die Position und Größe eines erkannten Gesichts darstellte.

In ähnlicher Weise wurde das YOLOv8-Objekterkennungsmodell zur Identifizierung von Nummernschildern verwendet, das einen Satz von Bounding-Boxes erzeugte, die den erkannten Nummernschildern in jedem Bild entsprachen. Nachdem alle Bounding-Boxes für Gesicht und Nummernschild erhalten waren, wurden die entsprechenden Bildbereiche extrahiert und dann verpixelt.

Trotz der Wirksamkeit dieser KI-basierten Methoden kann es zu falsch positiven und falsch negativen Ergebnissen kommen, z. B. kann ein Nicht-Gesichtsbereich fälschlicherweise als Gesicht identifiziert werden oder ein tatsächliches Gesicht könnte übersehen werden. Um dieses Problem zu beheben, wurde eine zweite Phase der manuellen Überprüfung durchgeführt. Hierzu wurde ein interaktives Tool entwickelt, das es erlaubt jedes Bild visuell zu inspizieren und manuell alle verbleibenden Gesichter oder Nummernschilder auszuwählen, die eine Nach-Anonymisierung erforderten. Dadurch wird sichergestellt, dass alle identifizierbaren personenbezogenen Daten in Übereinstimmung mit den Datenschutzbestimmungen unkenntlich gemacht werden.

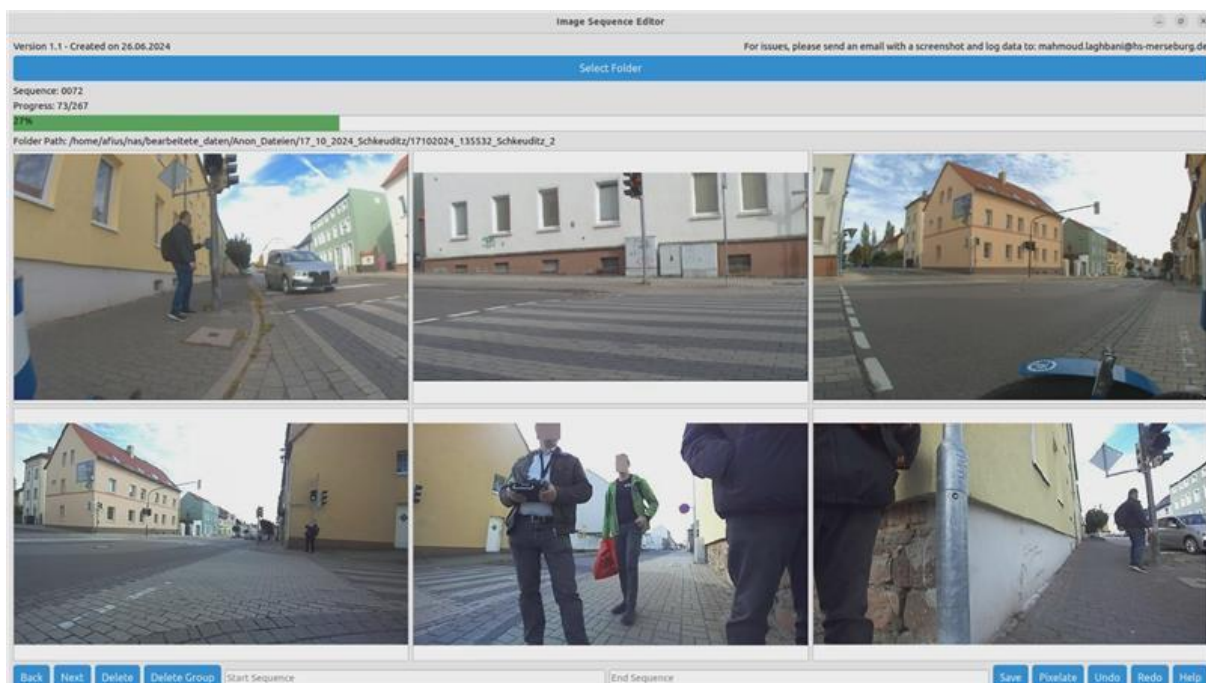


Abbildung 14 Tool zur Anonymisierung personenbezogener Daten

## **AP 5.2 Erstellen der manuellen Annotationen (Labeln)**

Der Großteil der manuellen Annotationen wurden im Unterauftrag durch die Intenta SEH GmbH durchgeführt. Hierfür war es notwendig, dass FusionSystems den Annotationsprozess anhand diverser Testsequenzen durchtestete, um Fehler und Bugs im Annotationstool zu identifizieren sowie die Widerspruchsfreiheit und Präzision der Annotationsregeln zu testen. Des Weiteren wurde ein Metabase-Fortschrittsbericht aufgesetzt, welcher den Annotationsfortschritt visualisiert, siehe Abbildung 15. Dieser Fortschrittsbericht beinhaltet unter anderem

1. Die Aufteilung der Sequenzen und Label nach aktuellem Zustand (Labeling, Review, Final)
2. Die Gesamtzahl an Labeln über der Zeit (hieraus lässt sich auch die mittlere Annotationsgeschwindigkeit ermitteln)
3. Die Aufteilung der Label auf Sequenzen und Sensor
4. Die Klassenverteilung der Annotationen
5. Eine Sequenzübersicht mit allen relevanten Informationen je Sequenz
6. Ein Sequenzsteckbrief mit Metainformationen je Sequenz (Jahreszeit, Wetterbedingungen, Ort, ...)

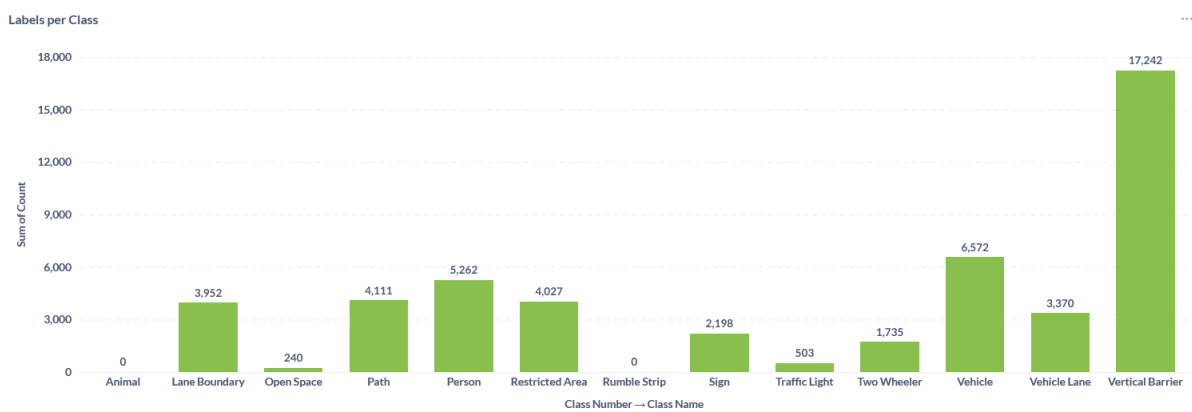


Abbildung 15: Ausschnitt aus dem Metabase-Fortschrittsbericht

FS unterwies die Labeler von Intenta in den Annotationsanforderungen und der Nutzung des Labeling Tools und beantwortete zahlreiche Fragen dazu. Abschließend wurde ein Review der Annotationen durch Hilfwissenschaftler der Hochschule Merseburg durchgeführt, um eine möglichst hohe Qualität des Referenzdatensatzes zu garantieren.

### **AP 5.3 Automatische Erzeugung der restlichen Annotationen**

Die in AP 3.3 entwickelten Verfahren zur automatischen Erzeugung von Annotationen wurden implementiert und erfolgreich getestet. Dazu wurden folgende Schritte durchgeführt:

1. Testplanung: In diesem Schritt wurden die Testziele und -strategien definiert. Testziel war die Ermittlung der Positionsgenauigkeit der automatisch erzeugten Annotationen. Zudem wurde ein Zeitplan für die Testaktivitäten erstellt. Es wurde entschieden, dass die Positionsgenauigkeit über die Metrik „Intersection over Union“ beschrieben wird. Dies setzt die Schnittmenge zweier Flächen ins Verhältnis zu deren Vereinigungsmenge. Sind zwei Flächen in Lage und Abmessungen identisch, dann beträgt diese 100%.
2. Testfallentwurf: In diesem Schritt wurden detaillierte Testfälle erstellt, die verschiedene Szenarien abdeckten. Die erwarteten Ergebnisse für jeden Testfall wurden in Form einer Ground-Truth-Annotation ebenfalls dokumentiert.
3. Testdurchführung: Die Tests wurden gemäß den definierten Testfällen durchgeführt, indem im Labeling Tool die jeweilige automatische Erzeugung gestartet wurde. Die

Testergebnisse (Abweichung von automatisch erstellter Annotation mit Ground truth) sowie alle aufgetretenen Software-Fehler (Absturz des Softwaremoduls) wurden dokumentiert.

4. Fehlerberichterstattung: Alle gefundenen Software-Fehler wurden in einem Fehlerverfolgungssystem erfasst. Detaillierte Informationen zu jedem Fehler, einschließlich der Schritte zur Reproduktion, wurden bereitgestellt.
5. Regressionstests: Regressionstests wurden durchgeführt, um sicherzustellen, dass neue Änderungen oder Fehlerbehebungen keine bestehenden Funktionen beeinträchtigen.
6. Testauswertung: Die Testergebnisse wurden analysiert und die Qualität der Software bewertet. Ein Testbericht wurde erstellt, der die Ergebnisse (Positionsgenauigkeit) sowie die gefundenen Software-Fehler zusammenfasst.
7. Inkrementelle Anpassung des Codes: Basierend auf der Testauswertung wurde der Code so angepasst, dass die Software-Fehler beseitigt wurden und die Positionsgenauigkeit maximiert werden konnte. Dann wurden die Schritte 3 bis 7 wiederholt, bis keine Software-Fehler mehr auftraten und die Genauigkeit konvergierte.

Als Ergebnis des Testes konnten folgende Positionsgenauigkeiten ermittelt werden:

Automatische Erzeugung	Positionsgenauigkeit	Begründung
3D-zu-2D-Projektionen	95 %	Die Genauigkeit der Projektion hängt im Wesentlichen von der Genauigkeit des Kalibriermodells und der Kalibrierparameter ab.
Zeitliche Interpolation	94 %	Die Genauigkeit der Interpolation hängt im Wesentlichen davon ab, inwieweit die relevanten Objekte eine Bewegung durchführen, die im Bild durch ein kubisches Polynom beschrieben werden kann.
Vorhersage mittels KI-Verfahren	85 %	Die Genauigkeit der mittels KI-erstellten Annotationen hängt von der Lerndatenmenge und -varianz sowie der Architektur des Neuronalen Netzes ab. Der Einfluss dieser Größen auf die Positionsgenauigkeit ist jedoch aufgrund des Black-Box-Charakters von Neuronalen Netzen nicht direkt interpretierbar.

#### **AP 5.4 Stichprobenartige und automatische Evaluierung der erstellten Annotationen (Review)**

Der Reviewprozess für die erstellten Annotationen erfolgte in mehreren Stufen:

1. Automatische Überprüfung der Annotationen anhand der in AP 3.4 implementierten Algorithmen
2. Bewertung/Einarbeitung der gefundenen Fehler der automatischen Überprüfung durch den Labeler
3. Interner Review durch Mitarbeiter des Labeldienstleisters

4. Bewertung/Einarbeitung der gefundenen Fehler des internen Reviews durch den Labeler
5. Externer Review durch Mitarbeiter des Auftraggebers
6. Bewertung/Einarbeitung der gefundenen Fehler des internen Reviews durch den Labeler

Diskussionswürdige gefundene Fehler wurden mit allen Beteiligten diskutiert und das Ergebnis in Form einer Überarbeitung der Annotationsvorschrift dokumentiert.

### 3.1.6 AP 6 Baseline, Toolchain und Datenveröffentlichung

#### **AP 6.1 Umsetzung Baseline Modelle – Training, Validierung, Test**

Der Annotationsprozess stellte sich als aufwendiger als ursprünglich geplant heraus. Deswegen standen erst zu Projektende die annotierten Sensordaten zur Verfügung. Das Training des Baseline-Modells für die Objekterkennung und anschließende Validierung bzw. Test wird nach Projektende durch FS durchgeführt. Die Trainingspipeline wurde hierfür bereits implementiert. Dafür wurde die YOLO-Version in der Version v9 genutzt. Diese wurde auch bereits an öffentlich verfügbaren Annotationen getestet, sodass ein nachträgliches Training mit den Projektannotationen bestmöglich vorbereitet ist.

Im Rahmen der Entwicklung von Baseline-Modellen für die Trajektorienvorhersage wurde ebenfalls eine methodische Pipeline aufgebaut, die sich an gängigen Verfahren des maschinellen Lernens für zeitabhängige Vorhersagen orientiert. Dabei wurden verschiedene Modellarchitekturen untersucht, darunter klassische Ansätze wie lineare Modelle sowie neuronale Netze, insbesondere rekurrente Netzwerke (LSTMs) und Transformer-basierte Modelle.

Ein geplanter Test auf dem projektspezifischen Datensatz konnte aufgrund unzureichender Datenmenge nicht durchgeführt werden. Um dennoch eine fundierte Einschätzung der möglichen Modellperformance zu erhalten, wurden theoretische Evaluierungsmöglichkeiten erarbeitet. Hierzu zählten:

- Analyse existierender Benchmarks: Vergleich mit etablierten Testverfahren aus der Automobilbranche.
- Übertragung auf verwandte Datensätze: Untersuchung der Möglichkeit, die Trainingspipeline auf automobiler Datensätze anzuwenden.

Für die Implementierung wurde eine modulare Toolchain auf Basis von **TensorFlow** entwickelt. Zur Evaluierung der Modelle kamen standardisierte Metriken wie **ADE (Average Displacement Error)** und **FDE (Final Displacement Error)** zum Einsatz. Zudem wurde ein Überblick über aktuelle Methoden und deren Anwendung im Bereich der Trajektorienvorhersage erstellt.

Zukünftig ist geplant, die erarbeitete Pipeline direkt auf den projektspezifischen Datensatz anzuwenden.

#### **AP 6.2 Review Usability Framework, Daten**

Um die Annotationsdaten schnell und übersichtlich visualisieren zu können, wurde eine Visualisierungssoftware entwickelt, siehe Abbildung 16, und auf die Projektspezifika angepasst.

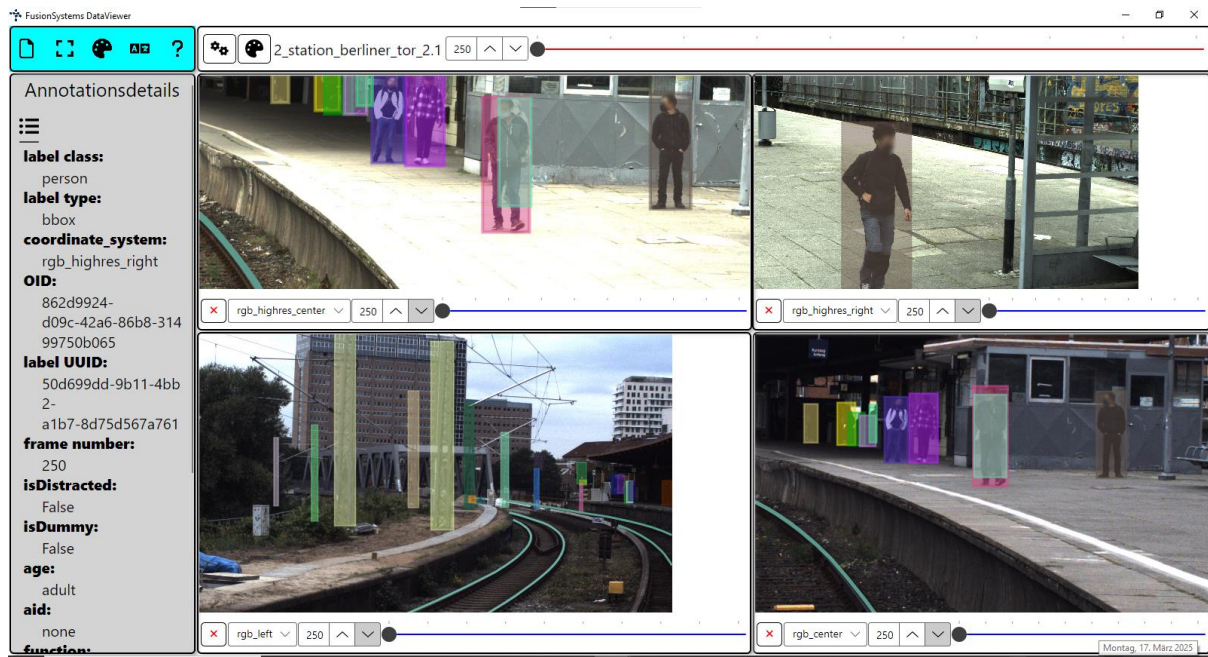


Abbildung 16: Visualisierungstool "Data Viewer"

Die Bewertung der Nutzerfreundlichkeit des Data Viewers ergab folgende Stärken:

- Intuitive Bedienbarkeit durch Nutzung üblicher Tasturlayouts (Nutzung Mausrad für Scrollen, WASD-Steuerung im Lidar, ...)
- Übersichtliche Darstellung der relevanten Annotationsinformationen (Objektklasse, Annotationsgeometrie, IDs, M-Frame-Nummer, Attributsausprägungen)
- Flexible Annotation der Sensor-Frames möglich und speicherbar
- Übersichtliches Hervorhebungskonzept angewählter Annotationen
- Einstellbare Färbungs- und Filteroptionen

Daneben konnte folgendes Optimierungspotential festgestellt werden:

- Abhängig von der Bildschirmauflösung und der Anzahl an dargestellten Sensor-Frames können die Sensordaten zu klein erscheinen, um Details zu erkennen.
- Die Navigation im Lidar erfordert ein gewisses Maß an Eingewöhnung.

### **AP 6.3 Veröffentlichung eines ersten Datensatzes**

Laut Projektantrag war es geplant, die Sensor- und Annotationsdaten frühzeitig auf einem Server von FusionSystems zu veröffentlichen. Allerdings lagen die Annotationsdaten erst zu Projektende vor, weswegen die projektbegleitende Veröffentlichung eines ersten Datensatzes in die Mobiltheke nicht möglich war.

## **AP 6.4 Datensatz veröffentlichen - Datenbank, Plattform für Softwareverwaltung**

Wie in AP 6.1 dargelegt, stand das Referenzmodell noch nicht zu Projektende zur Verfügung, weswegen es noch nicht veröffentlicht ist. Der Referenzdatensatz enthält ca. 100.000 Annotationen und ist in Form eines Beispieldatensatzes in der Mobiltheke unter <https://mobiltheke.info/offers/859106565693345792> abrufbar. Aus Gründen des Datenschutzes und des Umfangs des Rohdatensatzes wurde auf eine Gesamtveröffentlichung verzichtet. Weitere annotierte Sequenzen sowie die zugehörigen ROS-bags können auf Anfrage bei der HoMe bezogen werden.

### 3.2 Wichtigste Positionen des zahlenmäßigen Nachweises

Der Schwerpunkt des Projektes lag auf den Personalkosten sowie der Ausrüstung des Versuchsträgers. Des Weiteren wurden Geldmittel für die Beauftragung eines Annotationsdienstleisters sowie der Einholung eines Rechtsgutachtens zu datenschutzrechtlichen Fragestellungen verwendet.

### 3.3 Notwendigkeit und Angemessenheit der geleisteten Arbeit

Das Vorhaben OPTmicro stellt einen neuen, multisensoriellen Datensatz für das Training von Algorithmen im Kontext des autonomen Fahrens von Mikromobilen bereit. Damit ermöglicht es die Entwicklung von innovativen Ansätzen für Mikromobilen sowohl im Mobilitäts- wie auch im Logistiksegment. Mit der Erhebung und Bereitstellung des Datensatzes und den Begleitprodukten wird besonders der Datenzugang zu entsprechenden Datensätzen für das automatisierte/autonome Fahren, besonders für die Wissenschaftscommunity, deutlich verbessert. Diese müssen für entsprechende Entwicklungsarbeiten keine eigenen Datensätze aufwendig erheben und teure Sensorträger aufbauen, sondern können direkt in ihre Spezialdisziplin einsteigen. Damit wird ein Datenloch geschlossen, das aktuell noch viele F&E-Aktivitäten in dem Bereich des automatisierten/autonomen Fahrens von Mikromobilen deutlich hemmt.

Insgesamt waren die begleitenden Projektarbeiten notwendig, um Anforderungen und Eigenschaften des Referenzdatensatzes zu definieren, einen Versuchsträger entsprechend vorzubereiten und auszurüsten, die Datenerhebungskampagnen vorzubereiten und durchzuführen sowie abschließend den Datensatz zu annotieren und dessen Qualität und Anwendbarkeit sicherzustellen.

### 3.4 Verwertbarkeit der Ergebnisse

Die wissenschaftlichen Ergebnisse werden Einzug in verschiedene studentische Abschlussarbeiten und wissenschaftliche Veröffentlichungen finden, in denen auf Basis des erhobenen Datensatzes neue Ansätze zur Objekterkennung und Trajektorienprädiktion verfolgt werden. Das Konsortium ist mit verschiedenen Akteuren aus dem Mikromobilitätsbereich in Kontakt, die den finalen Datensatz für Ihre eigenen Arbeiten und Entwicklungen nutzen wollen.

Auf der wirtschaftlichen Seite wird FusionSystems die gewonnenen Erkenntnisse nutzen, um sein Labeling Tool in Bezug auf Nutzbarkeits- und Effizienzaspekte kontinuierlich weiterzuentwickeln. Des Weiteren erhofft sich FusionSystems, durch den Kompetenzausbau

in der Sensordatenverarbeitung neue Kundenprojekte (Auftragsentwicklung) aus dem Bereich Automatisierten Fahren durchführen zu können.

Auf der Mobilithek wurden folgende Daten und Dokumente veröffentlicht:

- Ein annotierter Testdatensatz
- Ausführliche Beschreibung des Sensorsetups und der Datenerfassung
- Annotationsleitfaden
- Skripte zur Visualisierung des Testdatensatzes

Mit dem Testdatensatz kann ein erster Eindruck von Aufbau und Qualität des Testdatensatzes gewonnen werden. Interessierte Anwender können somit den Nutzen für ihren jeweiligen Einsatzzweck abschätzen und bei Bedarf des Gesamtdatensatz anfragen.

Nicht enthalten sind:

- Der vollständig annotierte Datensatz, inkl. die zugehörigen Rohdaten. Gründe sind die Größe des Datensatzes (>2 TB) sowie Datenschutzgründe. Der vollständige Datensatz sowie die zugehörigen Rohdaten können jederzeit bei der HoMe nach Abschluss einer separaten Datenschutzvereinbarung angefragt werden.
- Rechtsgutachten, aufgrund einer Mandatsvereinbarung zwischen der Hochschule Merseburg und der beratenden Kanzlei

Link zur Mobilithek: <https://mobilithek.info/offers/859106565693345792>

### 3.5 Fortschritt auf dem Gebiet

Seit Antragstellung hat sich insbesondere das Gebiet der Objektdetektion und Trajektorienprädiktion mittels KI und maschinellem Lernen rasant weiterentwickelt. Bestehende Ansätze wurden verfeinert und neue entwickelt. Weiterhin bleibt die Verfügbarkeit der Trainingsdaten von entscheidender Bedeutung. Dem Projektkonsortium ist zum Berichtszeitraum kein weiterer multisensorieller Datensatz in unstrukturierten Verkehrsräumen oder mit Fokus auf die Mikromobilitätsentwicklung bekannt, so dass das Projektziel und die erreichten Ergebnisse weiterhin von höchster Relevanz sind.

### 3.6 Veröffentlichung der Ergebnisse

Im Rahmen des Projektes wurden folgende Veröffentlichungen erstellt:

1. S. Schmidt, K. Schwalbe: „OPTmicro: Lern- und Trainingsdaten für automatisierte Mikromobile“ in „MobilityAGENDA“ (Dezember 2024)
2. M. Laghbani, D. Iliev, K. Schwalbe, R. Konradt and S. Schmidt: „OPTmicro Dataset: Open Training Data for Autonomous Microvehicles“ in „Proceedings of the 5th Winter IFSA Conference on Automation, Robotics & Communications for Industry 4.0/5.0“ (February 2025)

Nach Projektende sind weitere Veröffentlichungen geplant, die die auf dem erstellten Sensordatensatz trainierten Referenzmodelle evaluieren.