



Partnerspezifischer Schlussbericht

Berichtszeitraum: 01.04.2020 - 31.12.2023

dSPACE GmbH

Vorhabenbezeichnung	KI Data Tooling - Methoden und Werkzeuge für das Generieren und Veredeln von Trainings-, Validierungs- und Absicherungsdaten für KI-Funktionen autonomer Fahrzeuge
Laufzeit des Vorhabens	01.04.2020-31.12.2023
Förderkennzeichen	19A20001H
Fälligkeitsdatum	30.06.2024
Erstellungsdatum	30.06.2024
Version	1.0

Gefördert durch:



aufgrund eines Beschlusses
des Deutschen Bundestages

Dokumenteninformation

Autoren

Nikolas Hemion - dSPACE GmbH

Tim Hengsbach - dSPACE GmbH

Andre Skusa - dSPACE GmbH

Matthias Thureau - dSPACE GmbH

Reviewer

Nikolas Hemion - dSPACE GmbH

Kontakt

Nikolas Hemion

dSPACE GmbH

Rathenaustraße 26

33102 Paderborn

Deutschland

Phone: +49 5251 1638-2078

Email: nhemion@dspace.de

Revisionslog

Version	Datum	Kommentar	Autor	Partner
1.0	30.06.2024	Version zur Einreichung	Nikolas Hemion	dSPACE GmbH

Inhaltsverzeichnis

1 Zusammenfassung / Summary	6
2 Übersicht	7
2.1 Aufgabenstellung	7
2.2 Voraussetzungen für das Vorhaben	7
2.3 Planung und Ablauf	8
2.4 Stand der Wissenschaft und Technik	9
2.5 Zusammenarbeit mit anderen Stellen	10
3 Eingehende Darstellung	11
3.1 Erzielte Ergebnisse	11
3.1.1 TP1	11
3.1.2 TP3	27
3.2 Zahlenmäßiger Nachweis	28
3.3 Notwendigkeit der Förderung	28
3.4 Nutzen und Verwertbarkeit	28
3.5 Veröffentlichungen	29

Abbildungsverzeichnis

Abbildung 3-1: Rekonstruktion einer Kreuzung in Aschaffenburg	14
Abbildung 3-2: Konvertierung von extrahierten Trajektorien aus Messdaten zu logischen Simulationsszenarien	16
Abbildung 3-3: Schematische Darstellung eines Cut-in Manövers	16
Abbildung 3-4: Ein Fußgänger überquert die Straße in einem logischen Szenario	17
Abbildung 3-5: Ein Fußgänger läuft parallel zum Ego-Fahrzeug in einem logischen Szenario	17
Abbildung 3-6: Vergleich zwischen realem und simuliertem Kamerabild der Szene "20210517_184350_0tb000h_00033".	18
Abbildung 3-7: Vergleich zwischen realem und simuliertem Kamerabild der Szene "20210517_191055_0tb000h_00045".	18
Abbildung 3-8: Vergleich zwischen realem und simuliertem Kamerabild der Szene "20210517_205604_0tbv000_00094".	18
Abbildung 3-9: Beispielbildpaar aus einem Realbild (Kitti, oben) und dem dazu korrespondierenden synthetischen Bild des zu Kitti erstellten 3D-Modells VKitti (unten).	19
Abbildung 3-10: Zu Kameradaten einer Messfahrt (links) wurde mit den Tools der dSPACE-Toolkette ein 3D-Modell erzeugt und simuliert (rechts).	20
Abbildung 3-11: Ein Bildpaar der ersten Sequenz aus den Kitti (oben) und VKitti Daten. Dabei sind in beide Bilder die Ground Truth Bounding Boxen beider Datensätze eingezeichnet (grün: Kitti, gelb: VKitti)	23
Abbildung 3-12: Das gleiche Bildpaar aus der ersten Sequenz wie in Abbildung mit bereinigten Ground Truth Bounding Boxen. Auch hier sind wieder in beide Bilder die Ground Truth Bounding Boxen beider Datensätze eingezeichnet (grün: Kitti, gelb: VKitti)	24
Abbildung 3-13: Vergleich der mAP Werte (Y-Achse) für alle 6 in VKitti modellierten Kitti-Szenen (X-Achse) vor (durchgezogene Linien) und nach (gepunktete Linien) der Bereinigung. Vor der Bereinigung liegt der mAP für 4 der 6 Szenen bei den synthetischen Bildern höher als bei den realen.	24
Abbildung 3-14: Beispielsequenz (a) - (f) mit Bounding Boxen eines auf A2D2 vortrainierten Objektdetektors. Zu sehen sind bildtechnische Herausforderungen (Lichtkränze, Spiegelungen, durch Bewegung verschwommen dargestellte Objekte etc) und Fehlklassifikationen (Motorbike statt Bicycle und eine Nicht-Erkennung des Fahrradfahrers).	26
Abbildung 3-15: <i>Beispielbild in schwieriger Lichtsituation. Fahrsequenz im Halbdunkel bei schlechtem Wetter mit vielen zu erkennenden Objekten auf der Straße</i>	26
Abbildung 3-16: <i>Beispielbild in sehr ausgeprägter Dunkelheit.</i>	27

1 Zusammenfassung / Summary

Zusammenfassung

Das Projekt "KI Data Tooling" hatte die Entwicklung von Methoden und Werkzeugen zur synthetischen Generierung von hochqualitativen Daten für das Training, die Validierung und die Absicherung von KI-Funktionen für das automatisierte Fahren zum Ziel. Hiermit würde die Entwicklung und Validierung von KI-basierten automatisierten Fahrfunktionen beschleunigen und verbessern werden, unter anderem durch die Simulation und das Testen einer breiten Palette von Fahrsituationen. dSPACE konzentrierte sich insbesondere auf die Erstellung künstlicher Simulationsszenarien und die Evaluierung von synthetisch generierten Sensordaten.

Dieser Bericht fasst die Beiträge von dSPACE zum Projekt zusammen. Diese umfassten neben der Übernahme der Leitung mehrerer Arbeitspakete insbesondere die Entwicklung von Metriken, die Weiterentwicklung von Software zur automatisierten Rekonstruktion von Szenen und Szenarien für die Simulation, sowie die Generierung von logischen Szenarien basierend auf realen Messdaten.

Summary

The "KI Data Tooling" project aimed at developing methods and tools for the synthetic generation of high-quality data for the training, validation, and safeguarding of AI functions for automated driving. This would accelerate and improve the development and validation of AI-based automated driving functions, among other things through the simulation and testing of a wide range of driving situations. dSPACE focused in particular on creating artificial simulation scenarios and evaluating synthetically generated sensor data.

This report summarizes the contributions of dSPACE to the project. These included, in addition to taking over the management of several work packages, the development of metrics, the further development of software for the automated reconstruction of scenes and scenarios for simulation, and the generation of logical scenarios based on real measurement data.

2 Übersicht

2.1 Aufgabenstellung

Im Projekt "KI Data Tooling" wurde die Entwicklung von Methoden und Werkzeugen zur effizienten Erweiterung der Datenbasis für das Training, die Validierung und die Absicherung von KI-Funktionen für das automatisierte Fahren vorangetrieben. Die dSPACE GmbH konzentrierte sich dabei insbesondere auf die Erstellung künstlicher Simulationsszenarien, einschließlich automatisierter 3D-Rekonstruktion, sowie die Evaluierung von synthetisch generierten Sensordaten und deren Verwendung in KI-Anwendungsfällen.

Die drei Sensormodalitäten, Kamera, Lidar und Radar, sollten in einem integrierten Ansatz generiert werden, um erstmals die gezielte Verfügbarkeit von multimodalen Daten realer, synthetischer und synthetisch augmentierter Verkehrssituationen in Form eines hochqualitativen Datensatzes zu erreichen.

Hierfür mussten im Rahmen des Projekts Realdatenaufnahmen an zwei Forschungskreuzungen in Aschaffenburg und Braunschweig durchgeführt werden. Diese Aufnahmen bildeten die notwendige Grundlage für die Erstellung virtueller Zwillinge zur Generierung synthetischer Daten. Diese synthetischen Daten sollten in Kombination mit den Realdaten einen neuen, umfassenden und hochqualitativen Datensatz bilden, welcher zur Untersuchung von Anwendungsfällen für das Training, die Validierung und die Absicherung von KI-Funktionen für das automatisierte Fahren dienen sollte.

Ein wichtiger Aspekt des Projekts war die systematische Erzeugung der relevanten Daten für Training, Validierung, Test und Absicherung. Hierfür stand die Entwicklung von Methoden und Tools im Vordergrund, welche erstmals einen generischen und übertragbaren Ansatz für die Bereitstellung solcher Daten ermöglichen sollten. Die so generierten Daten mussten validiert werden, wofür zunächst geeignete Metriken und Methoden entwickelt wurden.

Insgesamt hatte das Projekt "KI Data Tooling" zum Ziel aufzuzeigen, dass eine Toolkette zur Erzeugung synthetischer Sensordaten von großem Wert für die Entwicklung von autonomen Fahrfunktionen ist. Sie ermöglicht es, eine breite Palette von Fahrsituationen zu simulieren und zu testen, die für das Training von KI-Methoden relevant sind, und bietet gleichzeitig die Möglichkeit, diese Daten effizient zu verarbeiten und zu verwalten. Dies trägt dazu bei, die Entwicklung und Validierung von KI-basierten automatisierten Fahrfunktionen zu beschleunigen und zu verbessern.

2.2 Voraussetzungen für das Vorhaben

Das geplante Vorhaben zielte darauf ab, die Generierung hochqualitativer Daten zur Anreicherung von Realdaten durch den Einsatz von Simulationen für Anwendungsfälle im Kontext der Entwicklung des autonomen Fahrens zu untersuchen. Hierfür kamen verschiedene Ansätze in Betracht, die sich durch unterschiedliche Kosten, Aufwand und Qualitätseigenschaften auszeichneten. Diese Ansätze praktisch in Form von Toolketten umzusetzen und in der Praxis zu evaluieren, versprach einen Erkenntnisgewinn darüber, welche Methode am besten für welche Anwendungsfälle geeignet wäre. Das KI Data Tooling Projekt bot eine hervorragende Plattform für dieses Vorhaben. Die Partnerzusammensetzung aus OEM, Tier-1s, Technologieanbietern und

Forschungseinrichtungen konnte die notwendige Vielfalt an Expertisen und Perspektiven zusammenbringen.

Die Durchführung des Vorhabens versprach daher, die durch das Förderprogramm „Neue Fahrzeug- und Systemtechnologien“ angestrebte Stärkung des Forschungs- und Entwicklungsstandorts Deutschland zu unterstützen. Darüber hinaus sollte es dazu beitragen, Netzwerke zu stärken und Bestrebungen zur Standardisierung zu fördern.

2.3 Planung und Ablauf

Die Arbeiten im Projekt erfolgten entsprechend der Vorhabensbeschreibung in der vorgegebenen Arbeitspaketstruktur. Dabei fokussierten sich die Beiträge der dSPACE GmbH auf die Teilprojekte 1 und 3. Im Teilprojekt 1 nahm dSPACE eine zentrale Rolle ein und übernahm die Leitung der Arbeitspakete 1.1 und 1.3. Die regelmäßigen Treffen im Rahmen von Arbeitspaket- und Teilprojekt-Meetings erwiesen sich als äußerst wertvoll. Sie ermöglichten einen kontinuierlichen Informationsaustausch und eine enge Koordination zwischen den beteiligten Parteien. Die verschiedenen Perspektiven von OEM, Tier-1, Technologiepartnern und Forschungseinrichtungen führten zu einer ganzheitlichen Betrachtung und einer effektiven Zusammenarbeit.

Die Arbeitspakete wurden in mehreren Phasen bearbeitet. In der frühen Projektphase im Jahr 2020 lag der Fokus bei dSPACE auf den Arbeitspaketen 1.1, 1.2 und 3.5. In dieser Phase übernahm dSPACE die Leitung des Arbeitspakets 1.1 und koordinierte das Ergebnis 3.5.2. Die Entwicklung von Metriken für den Bildvergleich und die Untersuchung der Automatisierung der Szenenerstellung standen im Mittelpunkt.

Im Jahr 2021 erweiterte sich der Fokus auf die Arbeitspakete 1.1, 1.2, 1.3 und 3.5. In dieser Zeit leitete dSPACE die Arbeitspakete 1.1 und 1.3, und entwickelte Metriken für den Bildvergleich. Gleichzeitig wurden Softwarekomponenten für die automatische Generierung von statischen und dynamischen Szenenanteilen untersucht. Im Arbeitspaket 3.5 wurden Methoden zur automatischen Manövererkennung zur Kontextidentifizierung erarbeitet. Parallel dazu beteiligte sich dSPACE an der Ausarbeitung einer gemeinsamen Vorgehensweise zur Ablage und Dokumentation von KI-Modellen im Rahmen der TPX Working Group Model Zoo.

In den weiteren Projektphasen führte dSPACE die Arbeitspaketleitung im AP1.3 fort. Inhaltlich konzentrierte sich dSPACE im Jahr 2022 auf Arbeiten, die auch ohne die Verfügbarkeit der projekteigenen Daten möglich waren, die zu diesem Zeitpunkt wider der ursprünglichen Projektplanung noch nicht zur Verfügung standen. Die methodischen Weiterentwicklungen im Bereich der Szenariengenerierung (AP1.2) wurden vorangetrieben. Im Jahr 2023 lag der Fokus dann auf der Datenerstellung im "Efficient Path". Die Software zur Szenario-Generierung und die automatische Asset-Erzeugung wurden weiterentwickelt, um eine effiziente Rekonstruktion der 3D-Geometrie der Forschungskreuzung Aschaffenburg sowie einiger Szenarien zu ermöglichen. Gleichzeitig schloss dSPACE seine Arbeiten zum Vergleich zwischen realen und synthetischen Sensordaten ab.

Die kostenneutrale Projektverlängerung war ein entscheidender Schritt, um die angestrebten Ergebnisse zu erzielen. Sie ermöglichte es, die Arbeiten im Bereich der Szenariengenerierung und der Qualitätsuntersuchungen zu synthetischen Daten weiterzuführen und abzuschließen. Durch diese Verlängerung konnte dSPACE die geplanten Ziele erreichen und den Projekterfolg sicherstellen.

2.4 Stand der Wissenschaft und Technik

Für die geplante Bewertung der im Projekt erzeugten synthetischen Daten wurden sowohl allgemein bekannte und etablierte Metriken aus der Bildverarbeitung herangezogen als auch kurz vor Projektbeginn veröffentlichte neuere Verfahren, die auf vortrainierten neuronalen Netzen beruhen. Damit wurde ein Bereich von einfachen pixelbasierten Vergleichsverfahren über solche, die Erkenntnisse über menschliche Wahrnehmung zu berücksichtigen versuchen bis hin zu denjenigen Methoden, die auf Ergebnissen von Bilderkennung aufbauen. Die Grundannahme bestand darin, dass wir es auch in diesem Projektkontext immer mit Bildpaaren zu tun haben, aufgrund deren Unterschiede oder Gemeinsamkeiten auf die Qualität der synthetischen Nachahmung geschlossen werden kann. Somit standen insgesamt Methodiken und Verfahren zu Bild-Bild-Vergleichen im Vordergrund. Es wurde dabei ausschließlich auf veröffentlichte Methoden zugegriffen, und diese entweder neu implementiert oder auf eine freie (Open Source) Implementierung ohne weitere Rechteinschränkungen zugegriffen.

Weitere Vorarbeiten bei dSPACE betreffen die eigene Implementierung und Verwendung tiefer neuronaler Netze, insbesondere Convolutional Nets, verschiedener Bauart zur Objekterkennung. Mit diesen Vorarbeiten war eine Verwendung der eben erwähnten Verfahren zum Bildvergleich basierend auf Objekterkennungsergebnissen einfach in einen Gesamtablauf zu integrieren.

Hinsichtlich der Generierung von dynamischen Simulationsszenarien aus Messdaten hatte dSPACE bereits vor Beginn der Projektlaufzeit erste Konzepte entwickelt und implementiert. Zu diesem Zeitpunkt war es möglich Replay Szenarien aus annotierten Messdaten zu generieren, die anschließend in der Simulation verwendet werden konnten. Da es an einigen Stellen allerdings Abweichungen der Fahrzeugpositionen zwischen Annotationen und Simulation gab, ist die Verbesserung der Genauigkeit von Replay Szenarien ein Ziel von dSPACE im Projekt.

Im Bereich Generierung von logischen Simulationsszenarien hatte dSPACE vor dem Projekt bereits erste Ansätze implementiert, um konkrete Trajektorien zu abstrahieren. Mithilfe von Algorithmen konnten die Trajektorien in longitudinale und laterale Segmente gegliedert werden. Ein Segment umfasst beispielsweise einen Teil der Trajektorie mit annähernd konstanter Beschleunigung. Die identifizierten Segmente wurden genutzt, um parametrisierbare Szenarien zu erstellen. Insbesondere die longitudinalen und lateralen Geschwindigkeiten konnten mithilfe dieser Szenarien parametrisiert werden. Zu diesem Zeitpunkt war es allerdings nicht möglich in den Messdaten konkrete Manöver wie z.B. einen „Cut-In“ zu detektieren und nur diese speziellen Manöver zu logischen Simulationsszenarien zu konvertieren. Daher ist ein weiteres Projektziel von dSPACE logische Simulationsszenarien für spezielle Manöver aus Messdaten zu generieren. Die generierten Szenarien sollen zugeschnittene Parameter besitzen, mit denen die jeweiligen Manöver geeignet variiert werden können.

Stand der Technik zur Generierung synthetischer statischer Umgebungen bei dSPACE war hauptsächlich manuelle 3D Modellierung. Hierfür wurden viele externe Ressourcen verwendet (Assets), die anschließend für eine Sensor Simulation aufbereitet wurden. Darunter fiel das Anpassen der Polygonanzahl (Echtzeitfähigkeit), sowie der Erweiterung sensorspezifischer Materialien. Hier wurde mit sogenannten MaterialID-Texturen gearbeitet, die manuell in einen zusätzlichen Texturkanal auf handelsübliche 3D Assets gelegt werden. Assets wurden anschließend an eine automatisch generierte Straße (OpenDRIVE/RD) platziert. Nachteil dieser Vorgehensweise war das Wiederverwenden einiger weniger Assets in verschiedenen Szenen, dass sich nachteilig auf KI-Ansätze auswirken könnte.

2.5 Zusammenarbeit mit anderen Stellen

Im Rahmen von Tagungen oder Konferenzen wie zum Beispiel der Konferenz „AUTO-AI Europe“ hat ein Austausch mit anderen für KI Data Tooling relevanten Personenkreisen aus Wirtschaft und Forschung stattgefunden. Weiter sind die dSPACE Kundenkontakte genutzt worden, um die Sicht von weiteren OEMs und Zulieferern auf zukünftige Anforderungen an die Entwicklung von autonomen Fahrzeugen im Projekt berücksichtigen zu können. In Bezug auf Standardisierungsaktivitäten war dSPACE, und ist weiterhin, beteiligt an Gremien der ASAM. Außerdem fand ein Austausch mit der Universität der Bundeswehr München statt, wo unabhängig vom KI Data Tooling Projekt an verwandten Themen geforscht wurde.

3 Eingehende Darstellung

3.1 Erzielte Ergebnisse

3.1.1 TP1

3.1.1.1 AP1.1

In diesem AP erarbeitete dSPACE Metriken zur Bewertung der Qualität von synthetischen Sensordaten. Hierbei lag der Fokus auf Metriken für die LiDAR- und Kamera-Modalitäten und mit einem besonderem Augenmerk auf solchen Metriken, die für die Anwendung von neuronalen Netzen als Perzeptionsalgorithmen geeignet sind. Des Weiteren untersuchte dSPACE Aspekte von Verkehrsteilnehmer-Trajektorien und erarbeitete Metriken, welche für die Bewertung des Realismus von synthetisch erzeugten Trajektorien herangezogen werden können.

Zur Bewertung der Qualität synthetischer Kameradaten können verschiedenste Kriterien herangezogen werden. Von entscheidender Bedeutung für die Entwicklung und den Test von Perzeptionsalgorithmen im Bereich des autonomen Fahrens ist der Grad an Fotorealismus, den die synthetischen Bilder aufweisen. Da „Fotorealismus“ ein zwar intuitiver, aber quantitativ und physikalisch schwer fassbarer Begriff ist, stellt sich die Frage, wie diese Eigenschaft in synthetischen Bildern sichergestellt werden kann. Ein mögliches Vorgehen dazu ist, das synthetische Bild mit seinem Kamerapendant zu vergleichen und eine Distanz zu berechnen, die eine quantitative Beschreibung des Unterschiedes beider Bilder ermöglicht.

Zur Berechnung solcher Bilddistanzen gilt es, geeignete Metriken zu finden. Hier bietet die gängige Forschungsliteratur bereits viele Verfahren. Diese reichen vom einfachen Durchschnitt der quadratischen Differenz aller Pixelwertwerte (Mean Squared Error, MSE) bis hin zu Verfahren, die die Verteilung der Bildpunkte berücksichtigen und dabei die menschliche Wahrnehmung (Human Vision System, HVS) nachahmen (Structural Similarity, SSIM).

Die meisten dieser Verfahren zielen jedoch darauf ab, Unterschiede in den Veränderungen eines Bildes zu finden, beispielsweise um Störungen oder Bildrauschen zu identifizieren. Ein Beispiel für ein solches Standardverfahren ist das auf den MSE aufbauende Signal-Rausch-Verhältnis (Peak Signal-to-noise ratio, PSNR). Solche Vorgehensweisen setzen jedoch voraus, dass die beiden zu vergleichenden Bilder vom Aufbau und Inhalt sehr ähnlich sind.

Für die hier angestrebten Vergleiche kann das jedoch gerade nicht vorausgesetzt werden. Im Gegenteil: Die bezüglich ihrer Bildqualität zu vergleichende synthetische (3D) Szene wird auch bei großem Modellierungsaufwand nie exakt mit dem realen Vorbild übereinstimmen, beispielsweise was natürliche Objekte wie Bäume oder Sträucher betrifft. Ebenso wird die Synchronisation von Real- und Simulationsbild in einer Fahrsequenz nie so exakt sein, dass die bewegten Objekte der Szene sich in beiden Bildern an genau der gleichen Stelle befinden.

Gesucht sind also auch Maße, die dazu in der Lage sind, von Detailunterschieden im Bild- und Objektaufbau zu abstrahieren und Bildaspekte vergleichen, die zum visuellen Eindruck des Fotorealismus beitragen. Ebenso denkbar ist eine Auswahl und Kombination einer Menge von geeigneten Maßen, die jeweils einen Teilaspekt des Vergleichsspektrums abbilden.

Interessante Kandidaten hierfür sind Metriken, die auf der Auswertung interner Eigenschaften tiefer neuronaler Netze beruhen (Deep Neural Network, DNN). Die Grundannahme bei solchen

Verfahren ist, dass aufgrund der vortrainierten Objekterkennung die „Reaktion“ des Netzes bei ähnlichen Bildern auch ähnlich verhalten sollte. Basierend auf dieser Annahme sind unterschiedliche Vorgehensweisen denkbar, die sich im Wesentlichen durch die Repräsentation eines Bildes in den jeweiligen Netzwerkarchitekturen und deren Auswertung unterscheidet.

So wird zum Beispiel bei der *Perceptual Similarity* Metrik ein zur Bilderkennung vortrainiertes DNN nacheinander mit beiden zu vergleichenden Bildern als Eingabe ausgeführt. Die während einer Ausführung entstehenden Feature-Vektoren, die die Aktivierungen in den einzelnen Schichten des Netzes beinhalten, werden im Anschluss für beide Bilder normiert, sodass deren Differenz als Distanzwert gebildet werden kann.

Andere DNN-basierte Verfahren nutzen den latenten Raum eines *Variational Autoencoders* (VAE). Wird das Netz auf Bildern ausgeführt, stellt dieser Raum eine kompakte Repräsentation des Eingangsbildes dar, da VAEs darauf trainiert werden, ein Eingangssignal so zu komprimieren, dass es anschließend wieder korrekt „entpackt“ werden kann. Wird darüber hinaus die diskrete, vektorisierte Variante VQ-VAE angewandt, kann die entstehende Repräsentation ohne weitere Dekodierung zum Vergleich genutzt werden. Kontrastive Verfahren, die dafür sorgen, dass ein solches Netz beim Training auch negative Beispiele in ausreichender Zahl präsentiert bekommt, erhöhen die Repräsentationsfähigkeit noch einmal. Eine Bilddistanz beruht hierbei auf dem Vergleich der von einem trainierten Netz identifizierten wesentlichen Bildmerkmale.

Eine Übersicht dieser vergleichenden Betrachtung findet sich in den Projektergebnissen unter E1.1.1 (Metriken zur Bewertung der Qualität von synthetischen Kamera-, Lidar- und Radar-Daten). Diese Metriken wurden im Anschluss daran auf bereits vorhandenen öffentlichen Datensätzen evaluiert (siehe 3.1.1.3)

3.1.1.2 AP1.2

Ziel des AP1.2 ist die Entwicklung einer Toolchain, die für die Generierung von synthetischen Daten genutzt werden kann. Anschließend soll die Qualität der synthetischen Daten analysiert werden, indem geprüft wird, ob diese sich für das Training von KI-Algorithmen eignen. Das Training von KI-Algorithmen mit synthetischen Daten hat den Vorteil, dass der Annotierungsaufwand vermieden wird, der bei der Vorbereitung von Realdaten für das Training anfällt. Bei der Verwendung von synthetischen Daten entfällt dieser Annotierungsaufwand, da die Simulation die benötigten Ground Truth Daten direkt mitliefert.

Um eine hohe Qualität der synthetischen Daten zu erzielen, arbeitet dSPACE an einer sensorrealistischen Simulation. Durch die Rekonstruktion realer Verkehrssituationen mithilfe von aufgezeichneten Messdaten entsteht ein digitaler Zwilling, der für die Generierung synthetischer Daten genutzt werden kann.

Für die Validierung von ADAS/AD Software wird eine Vielzahl von realistischen Simulationsszenarien benötigt. Da deren manuelle Erstellung sehr zeitaufwändig ist, automatisiert dSPACE die Konvertierung von realen Messungen zu Simulationsszenarien. Für die Konvertierung werden sowohl Algorithmen für die Nachbildung der statischen als auch der dynamischen Szenarianteile benötigt.

Nachbildung der statischen Anteile

Zur Generierung statischer Umgebungen wurde ein Tool als Addon in der 3D Modellierungssoftware „Blender“ (blender.org) entwickelt. Ziel war es fertige 3D Umgebungen zu erhalten, die möglichst realistisch sind. Realismus bezieht sich dabei auf zwei Metriken:

- **Sensor-Realistisch:** Die 3D Umgebungen beinhalten neben den Texturen für das optische Rendern (Kamera-Simulation), auch Materialeigenschaften für andere Sensoren (Lidar, Radar), so dass die simulierten Sensordaten den Daten einer realen Messung ähneln.
- **Realitätsnah:** Im Idealfall ist der digitale Zwilling eine Rekonstruktion der echten Welt, bzw. eines Teilabschnitts, so dass eine Überführung vom simulierten zum realen Testen problemlos verläuft.

Sensor-Realismus der statischen Assets wurde verbessert, indem eine große Materialien-Datenbank angelegt wurde. Aufwändig erstellte Material-Datenbanken für Kamera-Simulationen (PBR: Physikalisch basiertes Rendering) existieren bereits, da dies ein gängiger Use-Case in der Spielebranche ist. Wir bedienen uns so einer Datenbank und erweiterten die dort vorhandenen Materialien mit Eigenschaften, die für eine realistische Sensor-Simulation benötigt werden.

Die Hauptarbeit steckte allerdings darin, das Ziel der realitätsnahen Rekonstruktion zu Erreichen. Es wurden einige KI-basierte Methoden zur Extraktion von Geometrien aus Realmessungen ausprobiert. Erste Untersuchungen ergaben allerdings, dass automatisierte Methoden aufgrund ungenauer, unvollständiger oder schlichtweg unterschiedlicher Eingangsdaten nicht zielführend erschienen. Unsere Lösungsstrategie strebte daher eine Lösung an, die sowohl mit wenigen als auch mit vielen Eingabedaten funktioniert. Wenige Eingabedaten heißt in dem Fall beispielsweise, dass nur Häusergrundrisse bekannt sind und der Rest prozedural entschieden werden muss. Viele Eingabedaten heißt, dass neben den Häusergrundrissen auch die Etagenanzahl, Fensterlayouts, Fenstertypen, Fensterrahmen-Materialien, Fassaden-Materialien, Dachformen, etc. bekannt sind.

Unsere Weiterentwicklungen zur Generierung statischer Inhalte soll also einerseits in der Lage sein aus wenigen Mess- oder Eingabedaten klassentypische Rekonstruktionen zu erstellen, indem viele Entscheidungen prozedural entschieden werden. Auf der anderen Seite sollen viele Eingabedaten verwendet werden können, so dass auch realitätsnähere Rekonstruktionen möglich sind. Diese Lösungsstrategie war von Erfolg gekrönt, da wir nun in der Lage sind den Datenaufbereitungsaufwand zu definieren (auch Annotierungsaufwand), abhängig vom Budget und Erwartungen an eine Szene, und dennoch in jedem Fall eine realistische Rekonstruktion erhalten.



Abbildung 3-1: Rekonstruktion einer Kreuzung in Aschaffenburg

Unsere 3D-Generatoren kamen zum Einsatz in der Rekonstruktion der Aschaffenburg-Kreuzung. In unserem Tool können Messdaten im Hintergrund visualisiert werden, so dass ein Modellierer entscheidende Features aus den Messdaten erkennt und für die Generatoren annotiert. So war es möglich diese Kreuzung mit wenig Aufwand realitätsnah zu rekonstruieren (5 Tage). Die von den Generatoren generierten Assets sind automatisch vorbereitet für eine Sensor-Simulation, da sie die erweiterte Material-Datenbank verwenden.

Der Generatoren Ansatz hilft insbesondere auch darin mehr Variation in der Erstellung synthetischer Daten zu erhalten. Vor diesem Projekt wurden üblicherweise dieselben Gebäudemodelle immer wieder in verschiedenen Szenen verwendet. Nun ist es auch möglich Grundrisse aus *OpenStreetMap* zu integrieren um große prozedurale Umgebungen zu erhalten, die grob auf der Realität basieren und ohne weiteren Annotierungs-Aufwand realistische und unterschiedliche Variation erstellen kann. Umso mehr Aufwand ein Modellierer anschließend in eine erweiterte Annotierung spendet, desto zielgerichteter ist auch das Ergebnis.

Neben den Gebäudegeneratoren sind eine Vielzahl weiterer Generatoren entwickelt worden, wie beispielsweise Generatoren für Zäune, Mauern, Leitplanken, Bordsteine und Fußgängerwege, Vegetation, Profile und Oberflächen.

Nachbildung der dynamischen Anteile

Für die Erstellung der dynamischen Szenarienanteile, werden zwei Datenquellen verwendet, die vom Projekt zur Verfügung gestellt wurden. Zum einen wird das GPS-Signal des Ego-Fahrzeugs verwendet und zum anderen werden die gelabelten 3D Bounding Boxen verwendet. dSPACE hat einen Konverter geschrieben, der die Daten einliest und in ein Format umwandelt, das von dSPACE's Szenario-Generierungs-Software verarbeitet werden kann, die während der Projektlaufzeit weiterentwickelt und verbessert wurde.

Nach der Konvertierung der Eingangsdaten liegt eine Beschreibung des dynamischen Verhaltens aller Verkehrsteilnehmer vor. Aus dieser Beschreibung generiert die Software das Simulationsszenario sowohl im dSPACE Format als auch in den OpenSCENARIO und OpenDRIVE Standards, die in der Industrie weit verbreitet sind.

Hinsichtlich der dynamischen Anteile kann in zwei Arten von Simulationsszenarien unterschieden werden. Zum einen ermöglichen Replay Szenarien präzise Rekonstruktionen der realen Messfahrten und zum anderen lassen sich reale Messfahrten mithilfe von logischen Szenarien abstrahieren.

In beiden Fällen ist es sinnvoll die annotierten Trajektorien zu filtern, bevor sie für die Generierung der Simulationsszenarien verwendet werden, da die Qualität der Trajektorien abhängig von den Sensoren, Algorithmen und Umgebungsbedingungen ist. Durch das Filtern wird vorhandenes Rauschen reduziert und die Plausibilität der Trajektorien erhöht.

Über eine Automatisierungsschnittstelle können die gefilterten Trajektorien in die Simulationsumgebung importiert werden. Auf diese Weise werden Replay Simulationsszenarien erzeugt, die das reale Verhalten der Verkehrsteilnehmer präzise reproduzieren. Dieser Ansatz ermöglicht den Vergleich zwischen realen und synthetischen Daten. dSPACE hat den Import der Trajektorien optimiert und die Genauigkeit der Replay Szenarien verbessert.

Eine Metrik, die die Genauigkeit von Replay Szenarien beschreibt, ist die Abweichung zwischen den Fahrzeugpositionen der Messung und der Simulation. Da das Ego Fahrzeug in der dSPACE Simulation mit einem komplexen Fahrzeugmodell simuliert wird, lassen sich Abweichungen nicht vollständig vermeiden. dSPACE hat die Regelung des Ego Fahrzeugs erweitert, sodass die Positionsvorgaben genauer getroffen werden. Eine zusätzlich implementierte Regelschleife minimiert den Positionsfehler in longitudinaler Bewegungsrichtung. Die Abweichungen der Replay Simulationsszenarien zur realen Messfahrt werden auf diese Weise verringert.

Neben dem Verhalten des Ego Fahrzeugs wurde außerdem das Verhalten der übrigen Verkehrsteilnehmer (sog. Fellows) angepasst. Zusätzlich zur Vorgabe des Bewegungsverhaltens in Straßenkoordinaten ist nun eine Vorgabe in Weltkoordinaten möglich. Auf diese Weise können die Trajektorien der Fellows mit sehr geringen Abweichungen simuliert werden.

Im Gegensatz zu Replay Szenarien liegt der Fokus von logischen Szenarien nicht auf einer exakten Rekonstruktion des Bewegungsverhaltens, sondern auf dessen Abstraktion. Die Abstraktion vereinfacht die Beschreibung des Bewegungsverhaltens und ermöglicht es das Szenario mithilfe unterschiedlicher Parametrisierungen in einer Vielzahl von Variationen zu simulieren.

Für die Ermittlung der Parameter werden die Geschwindigkeitsprofile der Fahrzeuge stark vereinfacht (vgl. Abbildung 3-2), indem die Beschleunigungen abschnittsweise konstant gesetzt werden. Diese Vereinfachung führt dazu, dass das Verhalten in der Simulation nicht exakt dem der realen Situation entspricht, allerdings ermöglicht es das generierte Szenario durch Parametrisierung zu verändern und gezielt Situationen mit einem hohen Mehrwert für die Validierung des SUT abzuleiten.

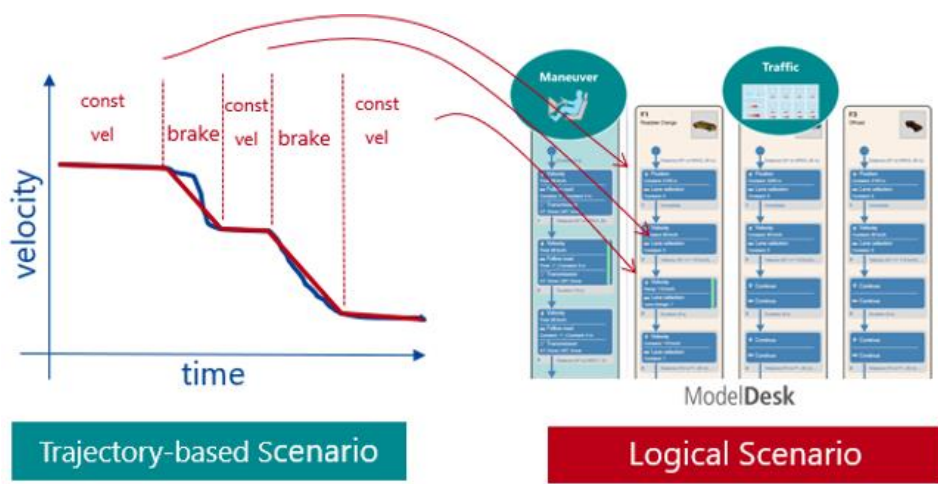


Abbildung 3-2: Konvertierung von extrahierten Trajektorien aus Messdaten zu logischen Simulationsszenarien

Häufig steht eine sehr große Menge an Messdaten zur Verfügung, wobei nur relativ wenige Sequenzen der Daten für die Validierung einer speziellen ADAS/AD-Funktion relevant sind. Beispielsweise kann es für den Test eines Autobahnpiloten interessant sein „Cut-in“ Manöver eines überholenden Fahrzeugs zu simulieren (vgl. Abbildung 3-3).

Um Tests mit einer hohen Relevanz für das SUT zu generieren, hat dSPACE Software entwickelt, die definierte Situationen in den realen Messdaten detektiert. Die detektierten Sequenzen können anschließend als Basis für die Generierung von logischen Szenarien verwendet werden.

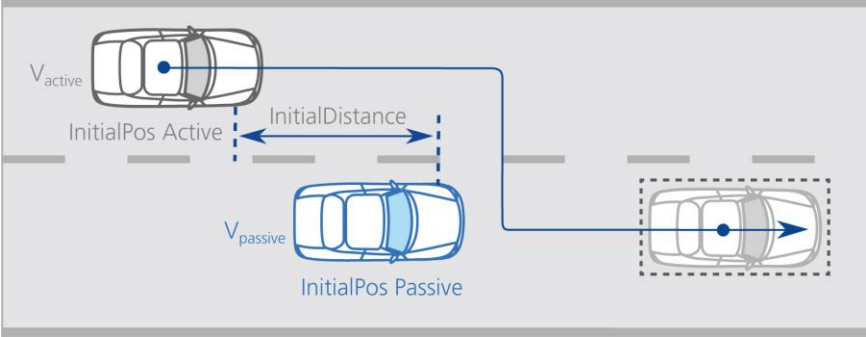


Abbildung 3-3: Schematische Darstellung eines Cut-in Manövers

Zusätzlich zu „Cut-In“, „Cut-Out“, „Drift-In“ und „Drift-Out“ Manövern, wurden die Detektion und Abstraktion von Manövern zwischen Fahrzeugen und Fußgängern implementiert. Die entwickelte Software sucht in den Messdaten nach Interaktionen zwischen dem Ego-Fahrzeug und einem Fußgänger. Dabei wird zwischen zwei Konstellationen unterschieden. In der ersten Konstellation überquert ein Fußgänger vor dem Ego-Fahrzeug die Straße (s. Abbildung 3-4). Nach der Detektion der Situation, wird das Bewegungsverhalten der beiden Verkehrsteilnehmer abstrahiert, sodass es mit einem logischen Szenario beschrieben werden kann. Sowohl das Bewegungsverhalten vom Ego-Fahrzeug als auch das Bewegungsverhalten vom Fußgänger lassen sich mithilfe von Parametern verändern.

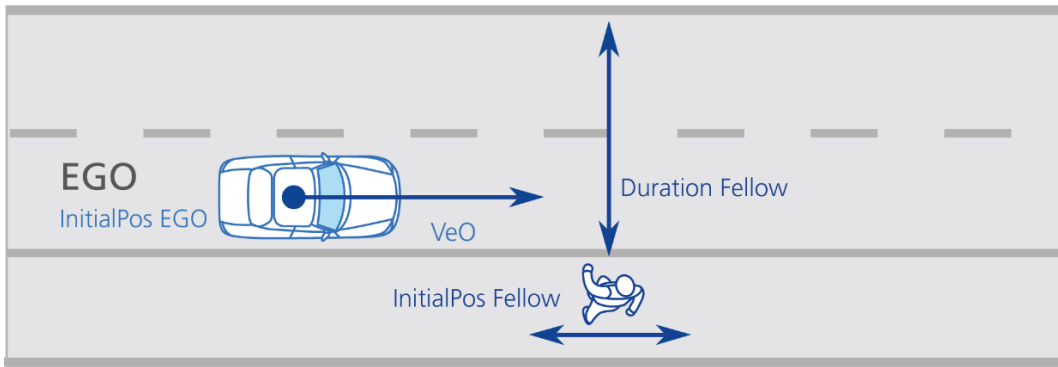


Abbildung 3-4: Ein Fußgänger überquert die Straße in einem logischen Szenario

Während das Verhalten des Ego-Fahrzeugs durch die Variation der initialen Position und Geschwindigkeit angepasst werden kann, lassen sich für den Fußgänger die Startposition entlang der Straße und die benötigte Zeit für das Überqueren der Straße variieren. Je nach Parametrisierung des Szenarios kann die Simulation mehr oder weniger herausfordernd für das Ego-Fahrzeug gestaltet werden.

Die Software detektiert außerdem eine zweite Konstellation zwischen dem Ego-Fahrzeug und dem Fußgänger in den Messdaten. In dieser Konstellation werden Situationen identifiziert, in denen der Fußgänger parallel zur Straße läuft (s. Abbildung 3-5). Nach der Generierung des entsprechenden logischen Szenarios, lässt sich die Simulation mit definierten Parametern variieren.

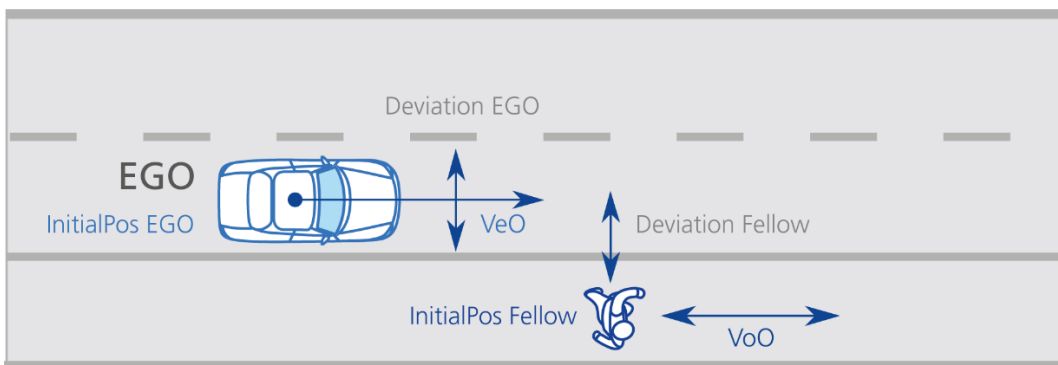


Abbildung 3-5: Ein Fußgänger läuft parallel zum Ego-Fahrzeug in einem logischen Szenario

In dieser Konstellation ist die laterale Distanz zwischen dem Ego-Fahrzeug und dem Fußgänger entscheidend. Daher lassen sich für beide Verkehrsteilnehmer laterale Abweichungen in der Position einstellen. Um die Vielfalt der Szenarien weiter zu erhöhen, können außerdem die Geschwindigkeiten von Ego-Fahrzeug und Fußgänger parametrisiert werden.

Je nach Parametrisierung eines logischen Szenarios kann sich das Verhalten der Verkehrsteilnehmer stark verändern. Eine inkonsistente Parametrisierung kann beispielsweise dazu führen, dass unbeabsichtigte Kollisionen zwischen Fahrzeugen entstehen. Um derartige Kollisionen zu vermeiden, ist es möglich die Verkehrsteilnehmer mit einem Fahrermodell auszustatten, das auf andere Verkehrsteilnehmer reagiert und Kollisionen vermeidet.

Neben der Generierung von logischen Simulationsszenarien können die extrahierten Parameterwerte außerdem zur Erstellung von Statistiken genutzt werden. Anhand der Statistiken können beispielsweise Verteilungen der relativen Geschwindigkeiten und relativen Distanzen zwischen den Fahrzeugen über eine große Anzahl von detektierten Cut-in Manövern identifiziert werden.

Simulation der Szenarien

Um die Ergebnisse zu visualisieren, hat dSPACE die realen Kamerabilder und die simulierten Kamerabilder in Form von Bildern und Videos gegenübergestellt. Abbildung 3-6, Abbildung 3-7 und Abbildung 3-8 zeigen beispielhaft den Vergleich zwischen Kamera und Replay Simulation.



Abbildung 3-6: Vergleich zwischen realem und simuliertem Kamerabild der Szene "20210517_184350_0tb000h_00033".



Abbildung 3-7: Vergleich zwischen realem und simuliertem Kamerabild der Szene "20210517_191055_0tb000h_00045".



Abbildung 3-8: Vergleich zwischen realem und simuliertem Kamerabild der Szene "20210517_205604_0tbv000_00094".

Nachdem die automatisierte Generierung der Trajektorien abgeschlossen ist, werden die 3D Assets der Fußgänger und Fahrzeuge manuell anhand der Bilddaten ausgewählt. Dieser Schritt

kann zu diesem Zeitpunkt nicht automatisiert ausgeführt werden, da die Annotationen keine Informationen über das Modell und die Farbe der gelabelten Fahrzeuge beinhalten.

Bei genauerer Betrachtung lassen sich einige Unterschiede zwischen den realen und synthetischen Kamerabildern erkennen. In der Simulation werden Standardeinstellungen für den Sonnenstand, das Wetter und die Kameraparameter verwendet. Dies führt dazu, dass sich die Ausrichtung der Schatten, die Helligkeit, die Wetterverhältnisse und der Öffnungswinkel der Kamera von der realen Situation unterscheiden. Diese Parameter müssten an die realen Bedingungen angepasst werden, um die Übereinstimmung zwischen Realität und Simulation weiter zu erhöhen. Außerdem werden die 3D Assets der Fußgänger und Fahrzeuge nicht exakt anhand der realen Objekte nachgebaut, sondern aus einer vorhandenen Bibliothek ausgewählt.

3.1.1.3 AP1.3

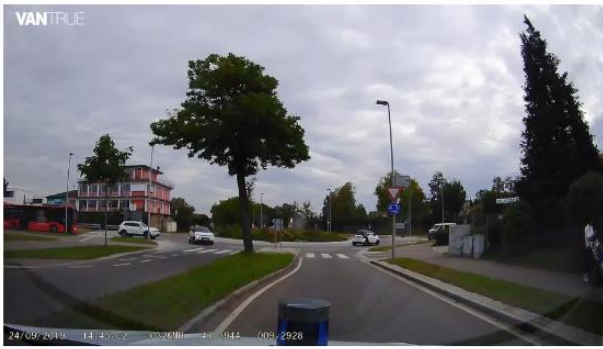
Die in 3.1.1.1 vorgestellten Metriken zum Bildvergleich wurden implementiert oder vorhandene Open Source - Implementierungen genutzt und auf öffentlich zugänglichen oder bei dSPACE bereits vorhandenen Testdaten evaluiert.

Dabei wurden auch die Eigenheiten der unterschiedlichen Metriken untersucht. Es ist naheliegend, dass es keine „one rules them all“ Metrik gibt, die alle Betrachtungsebenen, die sich im Bildvergleich ergeben, in eine einzige Zahl zusammenfasst. Vielmehr wird sich jede Anwendung der Metriken auf einen anderen Bildaspekt fokussieren, ohne wieder in ein einziges numerisches Ergebnis verrechnet zu werden.

Als Testdaten wurden die öffentlich verfügbaren Datensätze Kitti/VKitti (Abbildung 3-9) und ein von dSPACE erzeugter Datensatz einer Kreisverkehrsszene (Abbildung 3-10) verwendet. Bei Kitti handelt es sich um einen weit verbreiteten Datensatz des KIT (Karlsruhe Institute of Technology), zu dem in VKitti von den Neverlabs ein 3D-Modell erstellt und in Unity 3D simuliert wurde. Die interessanteste Eigenschaft dieser Datensätze für die vorliegende Anwendung ist die exakte Synchronizität auf Frame-Ebene, d.h. zu jedem Frame der realen Kameraszene gibt es genau einen synthetischen Frame mit der gleichen Nummerierung. So lassen sich große Menge von Bildpaaren auf einfache Art und Weise automatisiert und sequenziell vergleichen. Ein ähnliches Vorgehen wurde auch bei der Erstellung der dSPACE-Daten gewählt.



Abbildung 3-9: Beispielbildpaar aus einem Realbild (Kitti, oben) und dem dazu korrespondierenden synthetischen Bild des zu Kitti erstellten 3D-Modells VKitti (unten).



dSPACE: Measurement data

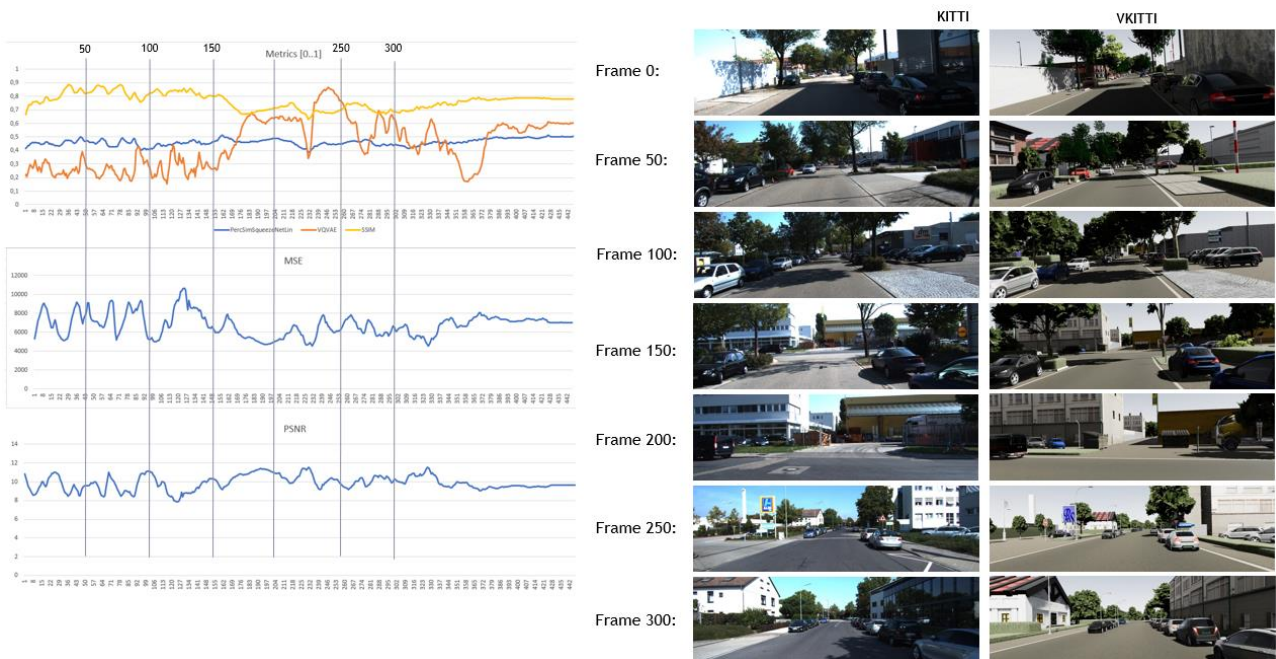


dSPACE: Generated simulation scenario

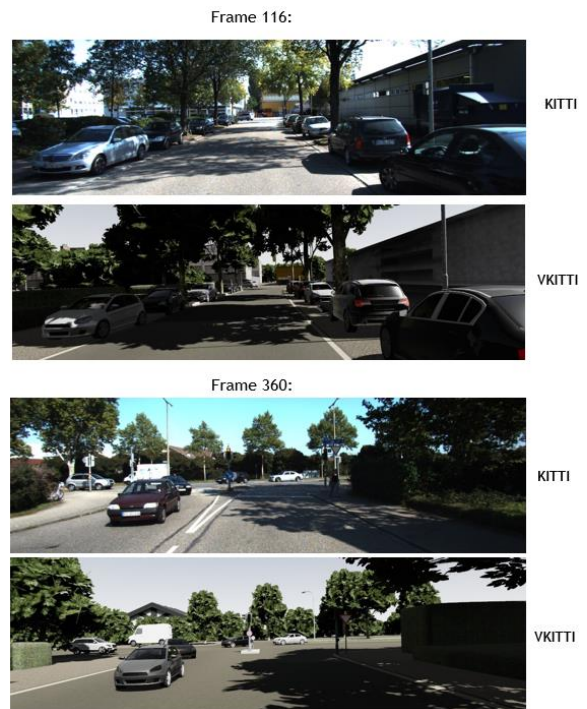
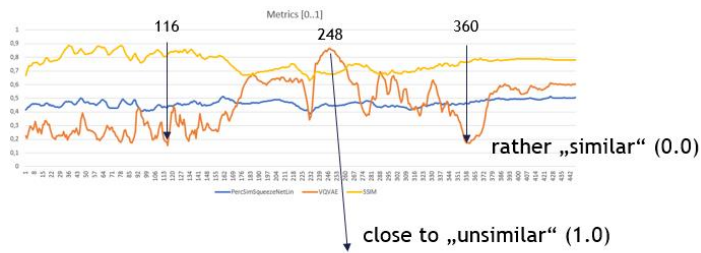
Abbildung 3-10: Zu Kameradaten einer Messfahrt (links) wurde mit den Tools der dSPACE-Toolkette ein 3D-Modell erzeugt und simuliert (rechts).

Betrachtungsebenen, auf denen Bilder ausgewertet und verglichen werden können, sind beispielsweise Farbverteilungen, Kanten und Strukturen oder auch Transformationen in den Frequenzbereich.

Ein Beispiel hierfür ist in der folgenden Abbildung zu sehen: Sie zeigt den Verlauf verschiedener Bildvergleichsmetriken auf allen Bildpaaren der ersten in Kitti/VKitti verfügbaren Szene (linke Seite). Um eine Vorstellung davon zu bekommen, wie die Bildpaare dazu aussehen, sind exemplarisch einige Frames zu sehen (rechte Seite) und deren Position im Zeitverlauf (senkrechte Linien in den Plots auf der linken Seite).



Im nächsten Bild werden etwas detaillierter die Frames, die die größeren Ausschläge in einer der Metriken (VQVAE, orange Linie im ersten Plot) verursachen, betrachtet:



Dabei fällt auf, dass die Bereiche, dass die VQVAE-Metrik (ein DNN-basiertes, trainiertes Maß) Bildpaare als besonders ähnlich bewertet, wo die Bilder stark verschattet sind. Gut ausgeleuchtete, aber für den menschlichen Betrachter auf den ersten Blick sehr ähnliche Bilder werden hingegeben als unähnlich betrachtet.

Neben der Tatsache, dass die Schattenbereiche große Bereiche aus dem Bild für den Vergleich herausnehmen (und das Bildpaar dadurch insgesamt ähnlicher werden lassen), sieht man bei genauerer Betrachtung, dass in dem zunächst sehr ähnlich wirkenden Bildpaar die Ähnlichkeit eher strukturell als farblich bedingt ist. Die Geometrien der Straße, Autos und Gebäude sind im 3D-Modell gut nachempfunden, wohingegen das Farbschema im synthetischen Bild eher grau und „flau“ wirkt, also tatsächlich Unterschiede bestehen.

An diesem Beispiel werden die angesprochenen Bildebenen, die zum Vergleich von den Algorithmen herangezogen werden, deutlich. Ebenso wie die Perspektive, verschiedene Maße auf den unterschiedlichen Ebenen gleichzeitig einzusetzen und den daraus resultierenden Metriktor zum Vergleich zu nutzen.

Von Bilddistanzmetriken zu Metriken der Objektdetektion

Es ergibt sich heraus aber die Frage, ob die genannten Metriken für eine Bewertung der Qualität der synthetischen Daten ausreichend sind. Auch ohne eine detaillierte weitere Analyse, wird bei einem ersten Vergleich klar, dass es Bildpaare mit einer für den menschlichen Betrachter recht hohen Ähnlichkeit zwischen realem und synthetischem Bild gibt, und andere, wo der Unterschied größer ist. Diese offensichtlichen Ähnlichkeiten und Unterschiede spiegeln sich aber auf der anderen Seite nicht bzw. nicht in gleichem Maße in allen verwendeten Metriken wieder. Eine wichtige Frage, die daraus folgt, ist aber nicht nur, ob und wie man Metriken findet, die diese visuellen Unterschiede adäquat abbilden, sondern inwieweit diese Unterschiede für eine Bewertung der Qualität der synthetischen Bilder überhaupt relevant sind.

Eines der Ziele zur Erstellung qualitativ hochwertiger synthetischer Daten besteht insbesondere darin, dass die verwendeten technischen Bilderkennungsverfahren (d.h. in der Regel DNNs), auf realen oder synthetisch erzeugten Bildern gleichermaßen funktionieren.

Das wiederum bedeutet, dass ein geeigneter Vergleichsmaßstab für Bilddistanzmetriken als wesentliche Komponente *die Auswertung von Bilderkennungsergebnissen auf den zu vergleichenden Bildern beinhalten sollte*. So kann besser eingeordnet werden, ob die von den Distanzmetriken diagnostizierten Unterschiede Auswirkungen auf die Qualität der Erkennung haben. Oder ob es sich eventuell sogar umgekehrt verhält, und gleich scheinende Bilder in entscheidenden Details so unterscheiden, dass dies eine Bilderkennung sogar verschlechtert.

Somit kann aus einem quantitativen Vergleich der Bilderkennungsergebnisse pro Bildpaar eine weitere, bilderkennungs-basierte Distanzmetrik definiert werden, die - parallel zu den oben gezeigten Verläufen, als Baseline herangezogen werden kann.

Performance-Indikatoren der Bilderkennung

Grundsätzlich basiert die Abschätzung, wie gut eine Bilderkennung auf einem Bild funktioniert, auf der *Intersection of Unit* (IoU), also der Größe der Schnittmenge der prognostizierten *Bounding Box* um ein Objekt mit der *Bounding Box* aus dem *Ground Truth* zu diesem Bild. Je größer die Schnittmenge ist, desto sicherer handelt es sich dabei um eine korrekte Prädiktion.

Überschreitet der IoU einen vorgegebenen Schwellwert (meist werden 50% oder 60% zugrunde gelegt), zählt die Vorhersage als *True Positive*. Alle anderen statistischen Klassifizierungen (*False Positive* etc) ergeben sich daraus.

Mit Metriken wie beispielsweise der *Mean Average Precision* (mAP) werden für eine gegebene Menge von Bildern diese statistischen Kennzahlen zusammengefasst und beschreiben die generelle Performance eines Bilderkennungsnetzwerkes auf einer Menge von Bildern. Basierend auf dem IoU können dann weitere Metriken, die auch auf der Ebene einzelner Bilder und Bildpaare funktionieren und keine Durchschnittswerte ermitteln, aufgebaut werden.

Ergebnisse und dadurch identifizierte Besonderheiten des VKitti-Datensatzes

Menge und Qualität von Ground Truth Annotationen sind in synthetischen Daten aufgrund der technischen Verfügbarkeit deutlich in der Regel deutlich größer. Ein Beispiel ist in den Bildern in Abbildung 3-11 zu sehen: In beiden Bildern sind die 2D Bounding Boxen aus beiden Datensätzen eingetragen. Gut zu erkennen ist, dass es deutlich mehr VKitti Bounding Boxen (gelb) zu teilweise nicht mehr zu erkennenden Autos gibt als das bei den Kitti Bounding Boxen (grün) der Fall ist.



Abbildung 3-11: Ein Bildpaar der ersten Sequenz aus den Kitti (oben) und VKitti Daten. Dabei sind in beide Bilder die Ground Truth Bounding Boxes beider Datensätze eingezeichnet (grün: Kitti, gelb: VKitti)

Das würde statistische Auswertungen wie den mAP (mean average precision) zugunsten der synthetischen Daten beeinflussen und einen Vergleich unfair machen.

Daher wurde über die Anwendung der IoU-Metrik eine Methode entwickelt, in einem Preprocessing-Schritt den Ground Truth in beiden Datensätzen anzugleichen und die Performance-Messung nur auf der Schnittmenge beider Ground Truth Werte auszuführen. Damit wird sichergestellt, dass der Vergleich tatsächlich fair ist und belastbare Ergebnisse liefert. Nach Anwendung des Preprocessings ist gut sichtbar, dass nur noch Bounding Boxes übrig geblieben sind, die in beiden Datensätzen vorkommen (Abbildung 3-12):



Abbildung 3-12: Das gleiche Bildpaar aus der ersten Sequenz wie in Abbildung mit bereinigten Ground Truth Bounding Boxes. Auch hier sind wieder in beide Bilder die Ground Truth Bounding Boxes beider Datensätze eingezeichnet (grün: Kitti, gelb: VKitti)

Es konnte darüberhinaus auch anhand von Auswertungen gezeigt werden, dass die Performance-Messung ohne eine solche Bereinigung tatsächlich in einigen Fällen Ergebnisse für die synthetischen Daten erbracht hat, die näher an den Realdaten liegen (Abbildung 3-13).

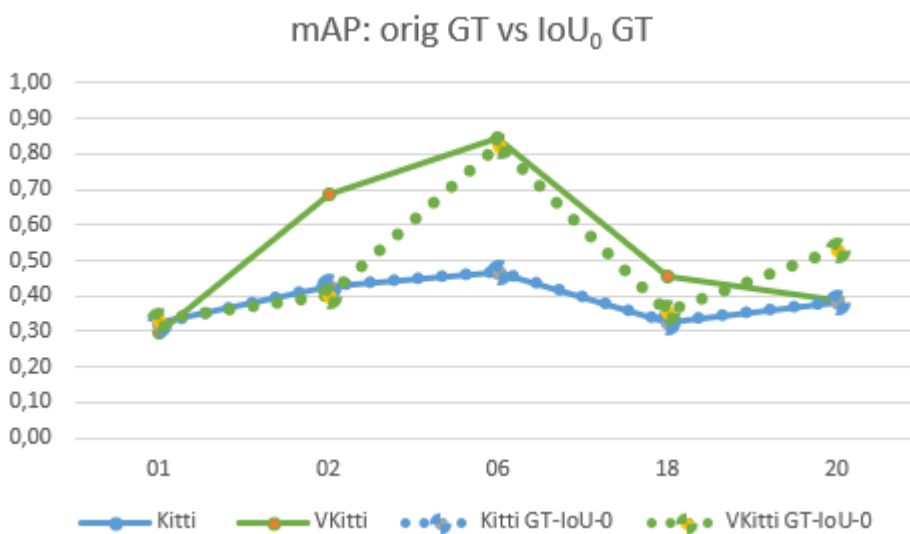


Abbildung 3-13: Vergleich der mAP Werte (Y-Achse) für alle 6 in VKitti modellierten Kitti-Szenen (X-Achse) vor (durchgezogene Linien) und nach (gepunktete Linien) der Bereinigung. Vor der Bereinigung liegt der mAP für 4 der 6 Szenen bei den synthetischen Bildern höher als bei den realen.

Mit der Bereinigung näherten sich die mAP-Werte an diesen Stellen an, und eine Schlussfolgerung ist an dieser Stelle, dass die Qualität der synthetischen Daten bei VKitti in einigen Szenen ausreicht, um zumindest aus Sicht des neuronalen Netzes dieselben Objekte zu erkennen bzw. nicht zu erkennen.

Dort, wo keine Annäherung stattgefunden hat, ist das in der Regel aus der Struktur der Szene heraus zu erklären. Dazu passen auch die Ergebnisse der Analysen des Projektpartners Bosch, die im Vergleich von Bilddistanzmetriken und Objekterkennungsperformance auf den sechs verschiedenen Szenen eine starke Tendenz zur Clusterung festgestellt haben: Innerhalb einer Szene sind liegen die Ergebnisse der Metriken und Performanbewertung oft relativ dicht beieinander, unterscheiden sich aber zwischen den Szenen stärker.

Bewertung von Realdaten zur Auswahl der Erstellung der synthetischen Szenen

Ein weiteres Arbeitspaket lag in der Bewertung der bereits vorliegenden Realdaten von der Forschungskreuzung in Aschaffenburg und einer darauf basierenden Auswahl von Szenen, die aus Sicht der Sensormodalität „Kamera“ besonders interessant für eine Replikation als digitaler Zwilling wären. Um diese Vorauswahl zu unterstützen, haben wir für die vorliegenden Bilder von einem bereits existierenden und von uns trainierten Objektdetektor 2D Bounding Boxes erzeugen lassen.

Ein Beispiel dazu ist in Abbildung 3-13 zu sehen: Zu erkennen sind Herausforderungen an die Bildererkennung durch die Beleuchtungssituation (Lichtkränze um Lichtquellen, Spiegelungen in der Straße, unscharfe dynamische Objekte). Bei den Autos scheint die Erkennung recht gut und stabil zu funktionieren, der Fahrradfahrer hingegen wird sowohl falsch erkannt (c-d) als auch gar nicht wahrgenommen (b, e), bevor er im letzten Bild (f) wieder wie im ersten (a) korrekt erkannt wird. Ob dieses Verhalten durch die gegebene Lichtsituation verursacht wurde oder auch in anderen Ausleuchtungen der Fall gewesen wäre, ist schwer festzustellen. Wichtig ist jedoch, dass ein digitaler Zwilling und dessen dynamisches Verhalten diese Situation so wiedergeben müssten, dass derselbe Objektdetektor auf dieser Sequenz auch auf den synthetischen Bildern das gleiche Verhalten zeigen würde.

Wäre das der Fall, könnte man durch Änderungen in der synthetischen Welt Ursachenforschung betreiben und versuchen festzustellen, woher die Fehlklassifikationen kommen und unter welchen Bedingungen sie verschwinden. Oder andersherum, welche Art von Objektdetektor in dieser Situation ein besseres Ergebnis liefern würde. Erst durch die Übereinstimmung von realen und synthetischen Daten in Bezug auf die Objekterkennungsergebnisse (im Detail und in der Gesamtperformance) wird eine sinnvolle Nutzung der synthetischen Szenen möglich.

Auch die Entscheidung, ob für nicht sicher erkannte Objekte ein Nachtraining auf synthetischen Bildern mit diesen Objekten erfolgen sollte, kann auf solchen Auswertungen beruhen. Solche und andere Szenen wurden aus diesen Gründen von uns als Kandidaten für die digitalen Zwillinge vorgeschlagen.

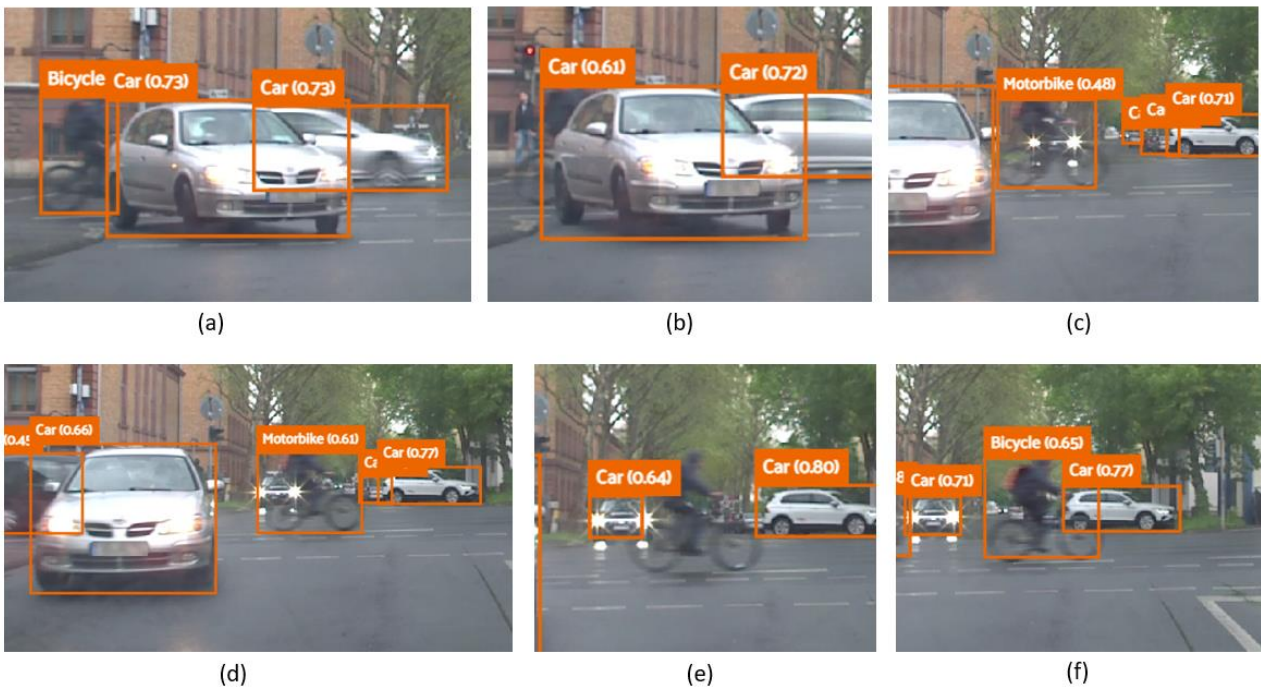


Abbildung 3-14: Beispielsequenz (a) - (f) mit Bounding Boxen eines auf A2D2 vortrainierten Objektdetektors. Zu sehen sind bildtechnische Herausforderungen (Lichtkränze, Spiegelungen, durch Bewegung verschwommen dargestellte Objekte etc) und Fehlklassifikationen (Motorbike statt Bicycle und eine Nicht-Erkennung des Fahrradfahrers).

Eine weitere Auffällig in weiteren Szenen der Realdaten ist, wie gut die Objekterkennung trotz schlechter Lichtverhältnisse noch funktionieren kann (Abbildung 3-15). Auch kleine und weit entfernte Objekte werden (teilweise) noch erkannt und korrekt zugeordnet, ebenso gelingt die Unterscheidung dicht zusammenstehende, parkender Autos am Straßenrand.



Abbildung 3-15: Beispielbild in schwieriger Lichtsituation. Fahrsequenz im Halbdunkel bei schlechtem Wetter mit vielen zu erkennenden Objekten auf der Straße

Für eine synthetische Zwillingsszene besteht die Erwartung nicht, diese Szene in ihrer Lichtsituation ganz exakt zu spiegeln, aber durch die gewählte Beleuchtung eine Bilderkennung zu ermöglichen, die im Durchschnitt dem hier gezeigten Verhalten entspricht: nicht alle Objekte

müssen sicher erkannt werden, aber ein großer Anteil, auch von entfernten oder zum Teil verdeckten Fahrzeugen.

Ein Extrembeispiel zeigt die folgende Abbildung 3-16:

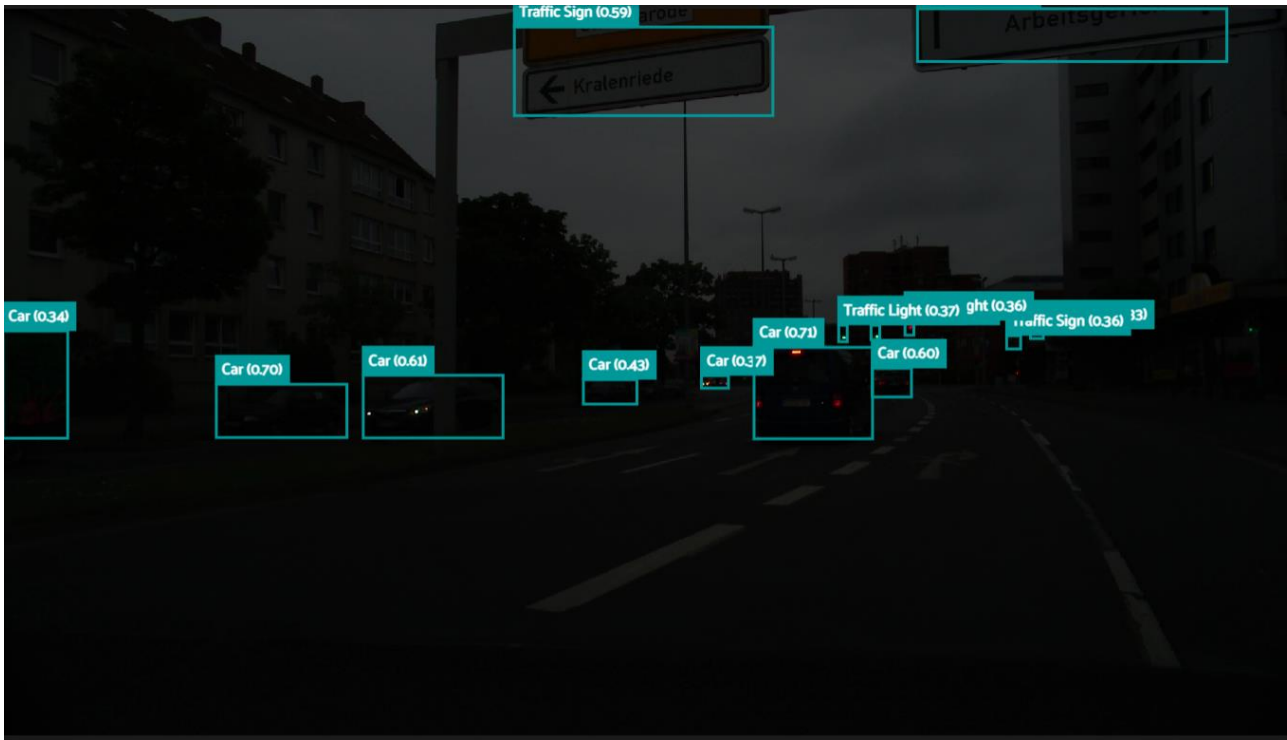


Abbildung 3-16: *Beispielbild in sehr ausgeprägter Dunkelheit.*

Hier ist besonders auffällig, wie gut die Objekterkennung auch noch in sehr abgedunkelten Situationen funktioniert.

Andererseits besteht die Vermutung, dass die Herausforderung bei den synthetischen Daten auch darin liegt, nicht zu viele Objekte zu erkennen. Synthetische Bilddaten zeichnen sich in der Regel dadurch aus, wenig verzerrt und verschwommen zu sein. Durch diese größere Klarheit, auch in den Kanten und Abgrenzungen der Objekte, könnte es sogar sein, dass auf solchen halbdunklen bis dunklen Bildern der Objektdetektor im synthetischen Fall sogar deutlich mehr erkennt. Das müsste sich aber ebenso negativ auf die Bewertung der synthetischen Bilder auswirken, da größtmöglicher (Foto-) Realismus gerade auch die (für die Bilderkennung) problematischen Eigenschaften abbilden soll.

3.1.2 TP3

3.1.2.1 AP3.5

Im Arbeitspaket 3.5 untersuchte dSPACE Kontextinformationen, die aus Fahrdaten, Metadaten oder annotierten Labels stammen können. Durch die Anreicherung der Datenbasis mit Kontextinformationen würde eine effiziente Durchsuchung großer Datenmengen ermöglicht. Dies ist nützlich für Entwickler, die spezifische Daten zur KI-gestützten Problemlösung suchen.

Wichtige Kontextinformationen, wie die Straßengeometrie (Krümmung, Neigung, Breite usw.), wurden identifiziert. Es wurde empfohlen, diese Informationen, einschließlich Fahrspurgeometrie und -breite, aus realen Daten oder HD-Karten zu extrahieren. Für Kreuzungen, welche als besonders relevant für kritische Szenarien des autonomen Fahrens anzusehen sind, sollte auch der Ampelstatus als Kontextinformation aufgenommen werden, gekennzeichnet als Ampelkategorie und Lichtfarbe.

Bei dSPACE stand des Weiteren die Entwicklung von Methoden zur Erkennung spezifischer Manöver aus den Daten im Fokus. Dafür wurde der öffentlich zugängliche Argoverse-Datensatz zur Entwicklung und zum Testen dieser Methoden genutzt. Da dieser keine Manöverlabels enthält, wurden solche mit Hilfe der Karten- und Trajektorieninformationen automatisch abgeleitet und zum Training von ML-Algorithmen verwendet. Zur Ergänzung der Realdaten von Projektpartnern, für die zunächst keine Karteninformationen verfügbar waren, wurden verschiedene Ansätze erprobt, darunter die Verwendung von OpenStreetMaps sowie alternative Ansätze basierend auf Fahrspurerkennung basierend auf Kamerabildern. Für das Generieren von Ampelkontextinformationen, bestehend aus dem Zustand und der Kategorie der Ampel, entwickelte dSPACE außerdem einen CNN-basierten Klassifikator.

3.2 Zahlenmäßiger Nachweis

Wurde bereits separat eingereicht.

3.3 Notwendigkeit der Förderung

Die Förderung des Projekts KI Data Tooling ermöglichte es dem Konsortium, wegweisende Arbeiten zu Methoden und Werkzeugen zur effizienten Erweiterung der Datenbasis für das Training, die Validierung und die Absicherung von KI-Funktionen für das automatisierte Fahren durchzuführen, die ohne diese Unterstützung in dieser Form nicht möglich gewesen wären.

Durch das Projekt wurde ein intensiver Austausch unter den Partnern, darunter OEM, Tier1s, Technologieanbieter und Forschungsinstitutionen, ermöglicht. Dieser Austausch, der ohne die Förderung wahrscheinlich nur in bilateralen Zusammenarbeiten, wie zum Beispiel in Engineering-Projekten, stattgefunden hätte, ermöglichte es uns, die volle Breite der Thematik zu beleuchten.

Darüber hinaus hat die Förderung die Realisierung der aufwändigen "Performance-Path" Toolkette ermöglicht. Ohne die Förderung und die Zusammenarbeit im Rahmen des Projekts wäre dies nicht umsetzbar gewesen, was bedeutet, dass ein Vergleich zu Toolketten einzelner Anbieter, wie zum Beispiel der Toolkette der dSPACE GmbH, nicht möglich gewesen wäre.

Die Förderung hat auch dazu beigetragen, Standardisierungsvorhaben voranzutreiben. Schließlich hat die Förderung die Bearbeitung einer Vielzahl von Forschungsfragen durch die Partner ermöglicht, die ohne die Förderung vermutlich nicht hätten finanziert werden können.

3.4 Nutzen und Verwertbarkeit

Das Projekt KI Data Tooling hat dSPACE eine Reihe von Vorteilen gebracht. Es hat wirtschaftliche Chancen eröffnet, indem es den Automatisierungsgrad bei der Erstellung von 3D-Szenen erhöht und die Sensorsimulation erweitert hat. Dies ermöglichte es dSPACE, passgenauere Simulationsszenarien für Kunden zu erstellen und synthetische Sensordaten zu erzeugen.

Darüber hinaus ermöglichen die im Projekt gewonnenen Erkenntnisse über die Einsetzbarkeit von synthetischen Daten für das Training und die Validierung von KI-Modellen es der dSPACE GmbH,

ihre Produkte weiter zu verbessern und zu erweitern. Dies wird durch die entwickelten Methoden zur Validierung von simulierten Sensordaten ergänzt, wodurch für die Kunden von dSPACE im Bereich von KI-Anwendungsfällen der Nutzen der dSPACE-Produkte klarer ersichtlich wird.

Bereits jetzt zeichnet sich der Einfluss des Projekts auf einige dSPACE Produkte und Dienstleistungen ab. Beispiele hierfür sind der TrafficVirtualizer zur Erstellung von Replay-Simulationsszenarien und die verbesserte Automatisierung zur Erstellung von 3D-Szenen.

In Bezug auf die Verwertung der Ergebnisse plant dSPACE kurz- und zu evaluieren, wie die im Projekt gemachten Erfahrungen und entwickelten Methoden in die verschiedenen dSPACE Produkte einfließen können. Für die Vermarktung zukünftiger Weiterentwicklungen wird dSPACE Pressemitteilungen, Newsletter und die dSPACE-Homepage nutzen. Es wird erwartet, dass die im KI Data Tooling erarbeiteten Konzepte langfristig auf weitere Anwendungsfelder übertragen werden können.

3.5 Veröffentlichungen

Die Projektergebnisse wurden im Rahmen des Abschlussevents vorgestellt.

Berichtsblatt

1. ISBN oder ISSN	2. Berichtsart (Schlussbericht oder Veröffentlichung) Schlussbericht
3. Titel KI Data Tooling - Methoden und Werkzeuge für das Generieren und Veredeln von Trainings-, Validierungs- und Absicherungsdaten für KI-Funktionen autonomer Fahrzeuge	
4. Autor(en) [Name(n), Vorname(n)] Nikolas Hemion Tim Hengsbach Andre Skusa Matthias Thureau	5. Abschlussdatum des Vorhabens 31.12.2023
	6. Veröffentlichungsdatum
	7. Form der Publikation Sonstiges
8. Durchführende Institution(en) (Name, Adresse) dSPACE GmbH Rathenaustr. 26 33102 Paderborn	9. Ber.-Nr. Durchführende Institution
	10. Förderkennzeichen 19A20001H
	11. Seitenzahl 29
12. Fördernde Institution (Name, Adresse) BMWK	13. Literaturangaben
	14. Tabellen
	15. Abbildungen 16
16. DOI (Digital Object Identifier)	
17. Vorgelegt bei (Titel, Ort, Datum)	
18. Kurzfassung Das Projekt "KI Data Tooling" hatte die Entwicklung von Methoden und Werkzeugen zur synthetischen Generierung von hochqualitativen Daten für das Training, die Validierung und die Absicherung von KI-Funktionen für das automatisierte Fahren zum Ziel. Hiermit würde die Entwicklung und Validierung von KI-basierten automatisierten Fahrfunktionen beschleunigen und verbessern werden, unter anderem durch die Simulation und das Testen einer breiten Palette von Fahrsituationen. dSPACE konzentrierte sich insbesondere auf die Erstellung künstlicher Simulationsszenarien und die Evaluierung von synthetisch generierten Sensordaten. Dieser Bericht fasst die Beiträge von dSPACE zum Projekt zusammen. Diese umfassten neben der Übernahme der Leitung mehrerer Arbeitspakete insbesondere die Entwicklung von Metriken, die Weiterentwicklung von Software zur automatisierten Rekonstruktion von Szenen und Szenarien für die Simulation, sowie die Generierung von logischen Szenarien basierend auf realen Messdaten.	
19. Schlagwörter Autonomes Fahren, synthetische Daten, künstliche Intelligenz, neuronale Netze, Sensorsimulation, Szenariogenerierung	
20. Verlag	21. Preis

Nicht änderbare Endfassung mit der Kennung 2369826-6

Document control sheet

1. ISBN or ISSN	2. type of document (e.g. report, publication) Veröffentlichung (Publikation)	
3. title KI Data Tooling - Methoden und Werkzeuge für das Generieren und Veredeln von Trainings-, Validierungs- und Absicherungsdaten für KI-Funktionen autonomer Fahrzeuge		
4. author(s) (family name, first name(s)) Nikolas Hemion Tim Hengsbach Andre Skusa Matthias Thurau	5. end of project 31.12.2023	6. publication date
	7. form of publication Sonstiges	
	8. performing organization(s) name, address dSPACE GmbH Rathenastr. 26 33102 Paderborn	
12. sponsoring agency (name, address) BMWK	9. originators report no.	
	10. reference no. 19A20001H	
	11. no. of pages 29	
	13. no. of references	
	14. no. of tables	
	15. no. of figures 16	
16. DOI (Digital Object Identifier)		
17. presented at (title, place, date)		
18. abstract The "KI Data Tooling" project aimed at developing methods and tools for the synthetic generation of high-quality data for the training, validation, and safeguarding of AI functions for automated driving. This would accelerate and improve the development and validation of AI-based automated driving functions, among other things through the simulation and testing of a wide range of driving situations. dSPACE focused in particular on creating artificial simulation scenarios and evaluating synthetically generated sensor data. This report summarizes the contributions of dSPACE to the project. These included, in addition to taking over the management of several work packages, the development of metrics, the further development of software for the automated reconstruction of scenes and scenarios for simulation, and the generation of logical scenarios based on real measurement data.		
19. keywords Autonomous driving, synthetic data, artificial intelligence, neural networks, sensor simulation scenario generation		
20. publisher	21. price	

Nicht änderbare Endfassung mit der Kennung 2369823-3