Schlussbericht

Verbundprojekt: 6G Integrated Communication & Sensing for Mobility – 6G-ICAS4Mobility

Teilvorhaben: KI-basierte Verfahren zur Szenenrekonstruktion, Klassifikation und Lokalisierung auf Basis von Radarsignalen

Zuwendungsgeber:	Bundesministerium für Bildung und Forschung
Zuwendungsnehmer:	Merantix Momentum GmbH (im Folgenden kurz: Merantix)
Förderkennzeichen:	16KISK238
Laufzeit:	01.10.2022 – 31.01.2024

Gefördert durch:

aufgrund eines Beschlusses
des Deutschen Bundestages

Das diesem Bericht zugrundeliegende Vorhaben wurde mit Mitteln des Bundesministeriums für Bildung und Forschung unter dem Förderkennzeichen 16KISK238 gefördert. Die Verantwortung für den Inhalt dieser Veröffentlichung liegt bei der Autorin/beim Autor.

Inhaltsverzeichnis

Inhaltsverzeichnis **1**
Schlussbericht **2**
I. Kurze Darstellung **2**
1. Aufgabenstellung 2
2. Voraussetzungen, unter denen das Vorhaben durchgeführt wurde 2
3. Planung und Ablauf des Vorhabens 2
4. Wissenschaftlicher und technischer Stand, an den angeknüpft wurde 3
5. Zusammenarbeit mit anderen Stellen 3
II. Eingehende Darstellung **4**
1. Verwendung der Zuwendung, erzielte Ergebnisse 4
AP1: Management & Verbreitung der Projektergebnisse 4
AP2: Anwendungsfälle und Anforderungen 4
AP3: Gesamtarchitektur, Schnittstellen und technische Enabler 6
AP4: Kommunikationszentrierter Ansatz und AP5: Sensierungszentrierter Ansatz 6
Literaturübersicht 6
Radardatensätze 8
Experimenteller Versuchsaufbau 9
Diskussion der Ergebnisse 12
AP6: Praktische Untersuchungen und Proof-of-Concept Demonstration 14
2. Wichtigste Positionen des zahlenmäßigen Nachweises 14
3. Notwendigkeit und Angemessenheit der geleisteten Arbeit 14
4. Voraussichtlicher Nutzen 15
5. Während der Durchführung des Vorhabens dem ZE bekannt gewordenen Fortschritts auf dem Gebiet des Vorhabens bei anderen Stellen 16
6. Erfolgte oder geplante Veröffentlichungen des Ergebnisses 16

Schlussbericht

I. Kurze Darstellung

1. Aufgabenstellung

Das Ziel von 6G-ICAS4Mobility besteht darin, im Zuge der Entwicklungen hin zu 6G bislang getrennte Kommunikations- und Sensierungsfunktionalitäten (insb. Radar) enger miteinander zu koppeln und zu integrieren. Der Fokus liegt dabei auf mobilen Anwendungen im Kontext Automotive und Flugdrohnen, sowie auf der Endgeräteseite mit einer direkten Kommunikation („Sidelink“) zwischen den beteiligten Geräten.

Das Teilvorhaben von Merantix Momentum untersuchte den Einsatz von Radardaten als Basis für die KI- und ML-Methoden. Eine Herausforderung und gleichzeitig das Potential bestand in der Integration von Kommunikation in diesen Ansatz. Der Austausch von gewonnenen Daten zwischen Fahrzeugen macht zusätzliche Methoden zur Datenfusion notwendig, welche allerdings auch das Potential haben, deutlich zu einer verbesserten Erkennungsleistung beizutragen.

2. Voraussetzungen, unter denen das Vorhaben durchgeführt wurde

Das Vorhaben 6G-ICAS4Mobility bündelt die Kompetenzen von 15 Projektpartnern aus Wissenschaft, Forschung, Technologie und Industrie. Es ist eines von einer Reihe von 6G-Industrieprojekten, welche ganzheitliche Systeme und Teiltechnologien für den Mobilfunk der 6. Generation erforschen. Mit der Maßnahme soll erreicht werden, dass wissenschaftliche und wirtschaftliche Akteure aus Deutschland bei der Ausgestaltung der technologischen Grundlagen für den Mobilfunk der 6. Generation im weltweiten Vergleich eine starke Rolle einnehmen.

Parallel zu den 6G-Industrieprojekten wurde eine Reihe von 6G-Forschungshubs gestartet, um die Spitzenforschung für die Kommunikationssysteme von morgen zu unterstützen. Gemeinsam bilden die 6G-Industrieprojekte und die 6G-Forschungshubs das Rückgrat der 6G-Initiative des BMBF, eine der zentralen Initiativen im Rahmen des Forschungsprogramms “Digital. Souverän. Vernetzt.”

3. Planung und Ablauf des Vorhabens

Das Vorhaben 6G-ICAS4Mobility wurde als Verbundforschungsprojekt mit 36 Monaten Laufzeit ausgelegt. Die nachfolgende Darstellung zeigt die Arbeitsplanstruktur des Projekts, welche die Arbeit in 6 wesentliche Arbeitspakete untergliedert.

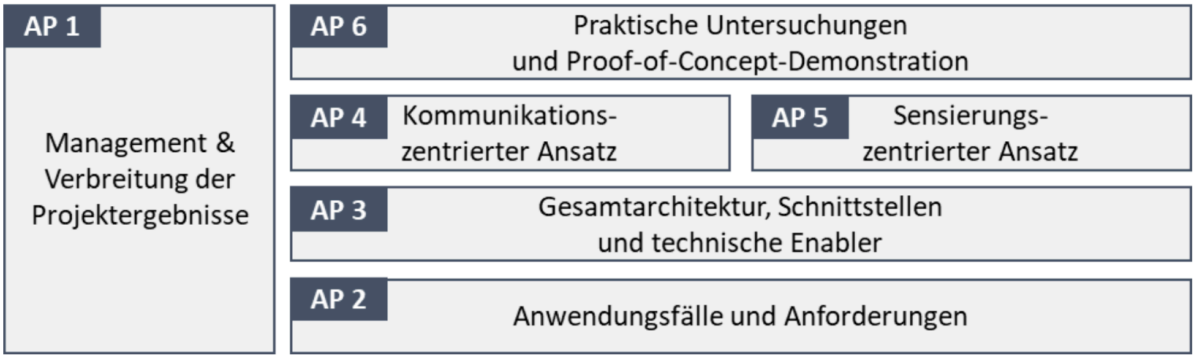


Abbildung: Projektstrukturplan.

Innerhalb des Arbeitsplans waren die wesentlichen Beiträge von Merantix Momentum als Teil der Arbeitspakete 4 und 5 geplant. Das Teilvorhaben adressierte die Fragestellung des Verbundprojekts aus den zwei wesentlichen Blickwinkeln durch:

1. Entwicklung von Methoden mit dem Schwerpunkt der Nutzung von kooperativen Systemen (Kommunikationszentrierter Ansatz)
2. Entwicklung von Methoden mit dem Schwerpunkt der Nutzung von neuartiger 6G-Radar-Sensorik (Sensierungszentrierter Ansatz)

Zunächst wurden Beiträge zur Anforderungsdefinition und zur Gesamtarchitektur geleistet. Darauf aufbauend wurden für die Beiträge in AP4 und AP5 eine Literatur-, Datensatz- und Methoden-Analyse durchgeführt, gefolgt von Experimenten zur Methodenentwicklung.

Zum 31.01.2024 ist Merantix Momentum (somit nach 16 Monaten Projektlaufzeit) aus dem Vorhaben ausgestiegen.

4. Wissenschaftlicher und technischer Stand, an den angeknüpft wurde

Das Projekt baut auf umfangreichen technischen Entwicklungen in den Bereichen der Funkkommunikation und Funksensorik auf. Bisher wurden diese Bereiche jedoch weitestgehend getrennt betrachtet und optimiert. Das Potential im Projekt lag damit vor allem in der neuartigen gemeinsamen Betrachtung von Funk und Sensorik aus den oben genannten zwei Blickwinkeln der kooperativen Systeme und der neuartigen 6G-Radar-Sensorik.

5. Zusammenarbeit mit anderen Stellen

Während der Projektbeteiligung fanden regelmäßige Austausche mit den weiteren geförderten 6G-Industrieprojekte und 6G-Forschungshubs statt, etwa im Rahmen der begleitenden 6G-Konferenzen.

II. Eingehende Darstellung

1. Verwendung der Zuwendung, erzielte Ergebnisse

Der wesentliche Teil der Zuwendung wurde für die anteilige Deckung der Personalkosten verwendet. Diese Aufwendungen spiegeln sich in den Ergebnissen der Arbeitspakete, an denen Merantix Momentum beteiligt war, wider. Zu einem geringen Teil fielen Reisekosten - resultierend aus Projekttreffen auf Gesamtprojektebene - an, sowie Kosten für den Betrieb von Recheninfrastruktur.

Das Teilvorhaben von Merantix Momentum (MXM) ist im Gesamtprojekt im Wesentlichen in AP4 (Kommunikationszentrierter Ansatz) und AP5 (Sensierungszentrierter Ansatz) beteiligt. Darüber hinaus leistet MXM Beiträge zu AP1 (Management & Verbreitung der Projektergebnisse), AP2 (Anwendungsfälle und Anforderungen), AP3 (Gesamtarchitektur, Schnittstellen und technische Enabler) und AP6 (Praktische Untersuchungen und Proof-of-Concept-Demonstration).

Im Nachfolgenden werden die Arbeiten nach den einzelnen APs sowie dem in der Vorhabenbeschreibung dargelegten Arbeitsplan gegliedert genauer beschrieben.

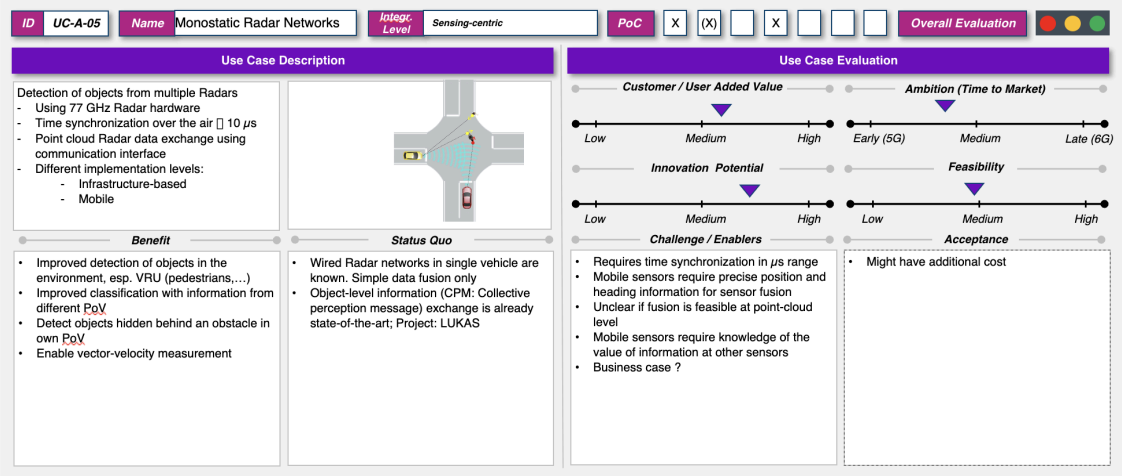
AP1: Management & Verbreitung der Projektergebnisse

Im Rahmen des Projektmanagements hat Merantix Momentum am Kickoff-Workshop bei Bosch in Stuttgart sowie zwei weiteren physischen Projekttreffen teilgenommen. Darüber hinaus fanden regelmäßige virtuelle Abstimmungsrunden der beteiligten Partner unter Teilnahme von Merantix Momentum statt.

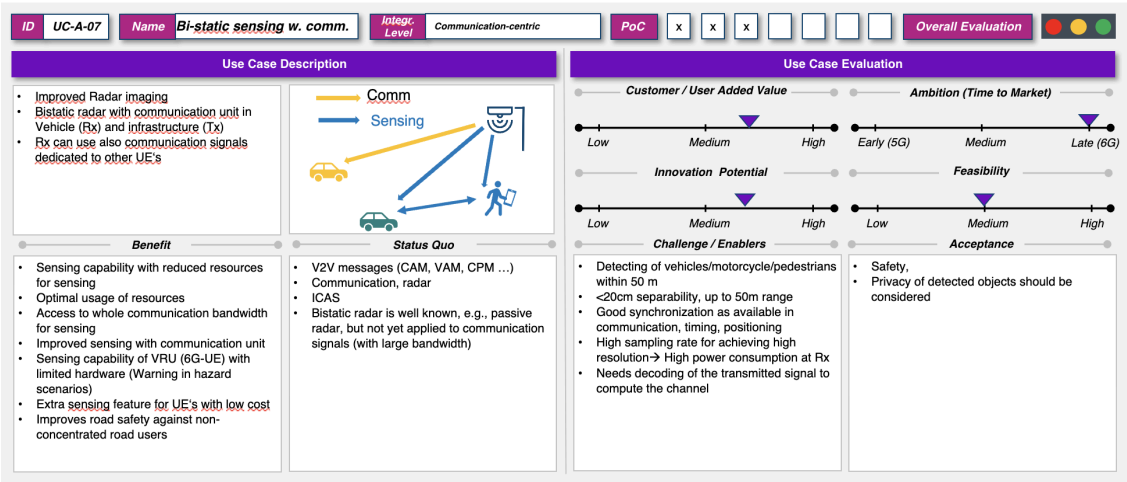
AP2: Anwendungsfälle und Anforderungen

Merantix Momentum hat an mehreren Abstimmungen teilgenommen, bei denen die im Projekt geplanten Use Cases weiter spezifiziert und durch die Partner bewertet wurden. Hierbei wurden die 10 genauer spezifizierten Use Cases betreffend PoC-2 (welcher später im Rahmen von AP6.2 umgesetzt werden wird) hinsichtlich ihrer Eignung für den Einsatz von Methoden des maschinellen Lernens bewertet. Auf Basis der initialen Bewertung wurden für die weiteren Arbeiten drei Use Cases priorisiert, die im Folgenden abgebildet werden.

Use Case UC-A-05 “Monostatic Radar Networks”



Use Case UC-A-07 “Bi-static sensing w. comm.”



Use Case UC-A-09 “Extended field of view”

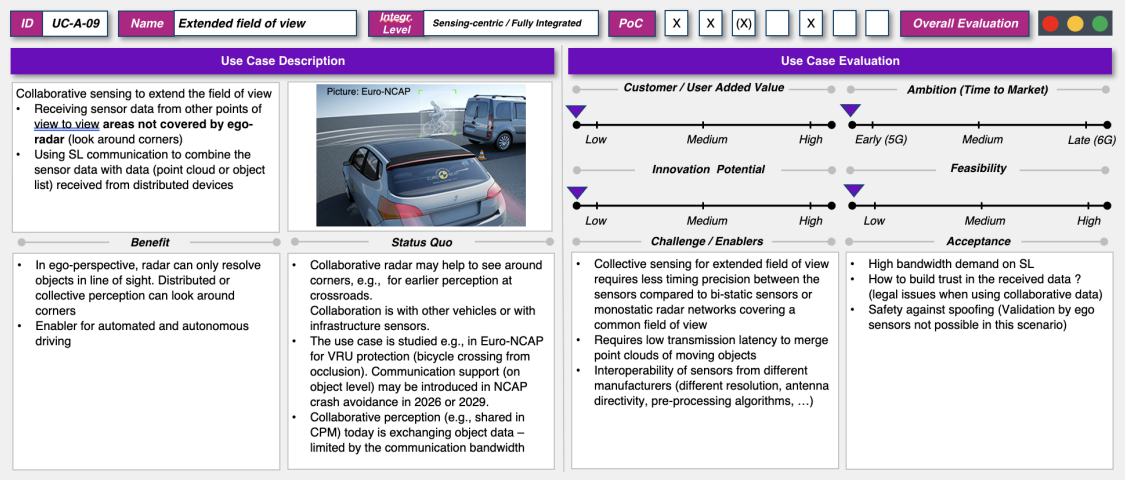


Abbildung 1: Priorisierte Use Cases.

AP3: Gesamtarchitektur, Schnittstellen und technische Enabler

Merantix Momentum hat an regelmäßigen Abstimmungen zur Systemarchitektur teilgenommen und notwendige Schnittstellen für die Integration der geplanten ML-Komponenten in die Gesamtarchitektur geleistet.

AP4: Kommunikationszentrierter Ansatz und AP5: Sensierungszentrierter Ansatz

Die wesentlichen geplanten Arbeiten des Teilvorhabens fallen in die Arbeitspakete 4 und 5. Während der Projektbeteiligung hat Merantix vor allem eine umfangreiche Literaturrecherche durchgeführt, Methoden auf ihre Eignung für das Projekt hin untersucht und erste Experimente durchgeführt. Die durchgeführten Untersuchungen zielen sowohl auf die in Task 4.2 geplante Szenenrekonstruktion auf Basis des 5G-Sidelinks, als auch auf die in AP5 geplanten Methoden zur Klassifikation und Lokalisierung von Objekten ab und werden daher im Folgenden gemeinschaftlich beschrieben.

Literaturübersicht

Die erste Phase des Teilprojekts konzentrierte sich auf eine umfassende Literaturrecherche über die Anwendung von Deep Learning (DL) auf Modalitäten außerhalb des visuellen Spektrums. Da es keine Referenzen für die Anwendung von DL auf 6G-Daten gibt, wurden vergleichbare Ansätze untersucht, die im Rahmen des ICAS4Mobility-Projekts von Bedeutung sein könnten – einschließlich Radar- und WiFi-Sensorik. Ein Schlüsselaspekt, der diesen Bereich vom klassischen Deep Learning im Bereich Computer Vision unterscheidet, ist die Einzigartigkeit der (Funk-)Datenmodalitäten. Im Gegensatz zu herkömmlichen bildbasierten Datensätzen bestehen die Daten in diesen Bereichen oft aus komplexen Signalmustern, dünn besetzten und verrauschten Punktwolken und mehrdimensionalen Sensordaten, was eine

besondere Herausforderung für die Analyse und Interpretation darstellt. Dies erfordert neuartige Deep-Learning-Ansätze, die auf diese spezifischen Datenmerkmale zugeschnitten sind und sich von den konventionellen Techniken der Computer Vision unterscheiden.

Die Literaturrecherche ergab sieben thematische Cluster:

1) Umwandlung von Rohsignalen in abstrakte Repräsentationen,
2) Deep-Learning-Methoden,
3) Sensorfusion,
4) menschliche Gesten- und Bewegungserkennung,
5) Karierung und Lokalisation (*Simultaneous Localization and Mapping*, SLAM),
6) Lokalisierung in Innenräumen und
7) weitere Ansätze.

Für die jeweiligen Bereiche wurde jeweils eine ausführliche Literaturanalyse durchgeführt, siehe auch die in diesem Abschnitt angefügte Literaturliste. Aufbauend auf der Literaturübersicht konzentrierte sich Merantix in der nächsten Phase auf zwei Schlüsselbereiche. Erstens die Sensorfusion, insbesondere die Integration von Daten aus verschiedenen Sensoren wie Radar, Kameras und LiDAR. Zweitens öffentlich verfügbare Radardatensätze, um geeignete Methoden des maschinellen Lernens für bestimmte Radardatentypen zu identifizieren.

Die gewonnenen Erkenntnisse werden im Blogbeitrag "Multimodal Fusion for Object Detection"[1] näher beschrieben und im Kontext des Projekts eingeordnet. Hier werden verschiedene DL-Methoden zur Verbesserung der integrierten Kommunikation und Sensorik (ICAS) für die Sensorfusion beleuchtet. Auch wird erläutert, wie verschiedene Sensoren mit ihren jeweiligen Stärken zu einem umfassenden Verständnis der Szene beitragen, wenn sie kombiniert werden. Beispielsweise liefern Kameras hochauflösende Bilder, LiDAR bietet präzise Entfernungs-messungen, und Radar eignet sich hervorragend für ungünstige Wetterbedingungen.

Die Kombination der Verfahren wird anhand von Fallstudien zur Verwendung etablierter multimodaler Lernstrategien in ML für die Sensorfusion vorgestellt. Dies wirft die Frage auf, wann und wie die Informationen fusioniert werden sollen. Dabei werden drei primäre Fusionsverfahren vorgestellt: frühe, späte und hybride Fusion (*early*, *late*, und *hybrid fusion*).

[1] URL: https://medium.com/merantix-momentum-insights/sensing-outside-the-visible-spectrum-3b78e96a6472

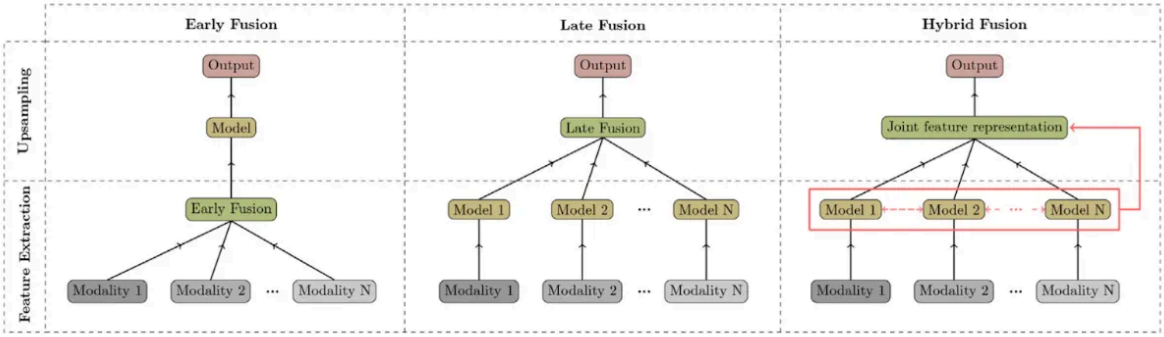


Abbildung 2: Eine Taxonomie der Fusionsverfahren [13]

Diese Taxonomie legt den Grundstein für die Nutzung der 6G-Sensorfusion zur robusten Objekterkennung. Sie unterstreicht den Bedarf an verbesserten Fusionsmethoden und die Bedeutung des Trainings von Modellen auf verschiedenen Datensätzen, um deren Anwendbarkeit in der Praxis zu verbessern. Auch wird das Potenzial für selbstüberwachtes Lernen aufgezeigt, um gemeinsame Embeddings für Radardarstellungen zu erstellen, die Modellvergleiche über verschiedene Datensätze hinweg ermöglichen und so die Robustheit verbessern.

Radardatensätze

Aufbauend auf die Literaturstudie und die Vertiefung von Sensorfusionsansätzen für multimodale Daten wurden experimentelle Arbeiten für die Objekterkennung in Radardaten durchgeführt, welche zunächst auf öffentlich verfügbaren Datensätzen ausgeführt wurden.

Eine wichtige Erkenntnis aus der Literaturrecherche war, dass Radardaten oft in verschiedenen Formaten vorliegen, die jeweils spezielle Verarbeitungsmethoden erfordern. Zu diesen Formaten, die in verschiedenen Stadien der Radardatenverarbeitung anzutreffen sind, gehören:

1. Analog-Digital-Wandler (ADC): Rohdaten von Radarsensoren, in der Regel unverarbeitet und in einem Format, das dem vom Radar erfassten Originalsignal am nächsten kommt.

2. Range-Doppler (RD): 2D-Tensordarstellung mit Informationen über die Entfernung und Geschwindigkeit (Doppler-Verschiebung) von Objekten.

3. Entfernungs-Winkel (RA): 2D-Tensordarstellung, welche die Entfernung und den Azimutwinkel der erkannten Objekte relativ zum Radarsensor darstellt.

4. Bereich-Winkel-Doppler (RAD): Ein umfassenderes Datenformat, das Entfernungs-, Winkel- und Dopplerinformationen kombiniert und eine detaillierte 3D-Perspektive der Umgebung liefert. Neuere Radarsysteme sind sogar in der Lage, einen 4D-Tensor mit zusätzlichen Höhenwinkelinformationen auszugeben.

5. Punktwolke: Eine Darstellung der Umgebung, bei der jeder Datenpunkt in der Wolke der Position eines Objekts im 3D-Raum entspricht.

Die Wahl der DL-Techniken muss auf den Modalitätstyp abgestimmt werden, da unterschiedliche Radardatendarstellungen (z. B. 2D-Bilder vs. Punktwolken) spezielle Architekturen und Vorverarbeitungen erfordern.

Angesichts dieser unterschiedlichen Radardatenformate wurde eine umfassende Übersicht zusammengestellt, in der öffentlich verfügbare Datensätze aus der Literatur verglichen werden. Diese Referenztabelle dient als wichtige Ressource für das Verständnis der Vielfalt von Radardaten und die Auswahl geeigneter Datensätze für spezifische Forschungsfragen und das Training oder die Evaluierung von Machine Learning Modellen. Die Datensätze wurden insbesondere auf diejenigen eingegrenzt, die zur Objekterkennung dienen und die somit für die im Rahmen des 6G-ICAS4Mobility-Projekts durchgeführten Aufgaben relevant sind.

Dataset	Year	Scale	ADC	RA or RD	RAD	PC	Annotation Type
astyx	2019	small	✗	✓	✗	✓	3d boxes
nuScenes	2019	large	✗	✓	✗	✓	3d boxes
Boreas	2020	large	✗	✗	✓	✗	pose, 3d boxes
CARRADA	2020	small	✗	✓	✓	✓	2d boxes, segmentation masks
DENSE	2020	large	✗	✗	✗	✓	3d boxes
GhentVRU	2020	medium	✗	✗	✓	✗	segmentation mask
Mulran	2020	medium	✗	✓	✗	✗	pose
Pointillism	2020	medium	✗	✗	✗	✓	3d boxes
RADIATE	2020	medium	✗	✓	✗	✗	2d boxes
Zendar	2020	small	✗	✓	✗	✓	2d boxes
CRUW	2021	medium	✗	✓	✗	✗	point location
PixSet	2021	medium	✗	✗	✗	✓	3d boxes
RaDICal	2021	large	✓	✗	✗	✗	2d boxes
RadarScenes	2021	large	✗	✗	✗	✓	point-wise
RaDDet	2021	small	✗	✓	✓	✗	2d boxes
K-Radar	2022	large	✗	✗	✓	✗	3d boxes
RaDIal	2022	medium	✓	✓	✓	✓	2d boxes, segmentation masks
TJ4DRadSet	2022	medium	✗	✗	✗	✓	3d boxes
View-of-Delft	2022	small	✗	✗	✗	✓	3d boxes

Tabelle 1: Vergleichstabelle der Radardatensätze für Erkennungsaufgaben.

Basierend auf den ermittelten Datensätzen wurde die Forschungsfrage formuliert: "Kann ein einzelnes Modell trainiert werden, um Informationen aus verschiedenen Darstellungen von Radardaten mit unterschiedlichen Datensätzen zu extrahieren?" Diese Frage zielt darauf ab, die Herausforderung der Entwicklung vielseitiger ML-Modelle anzugehen, die in der Lage sind, mehrere Radardatenformate zu verarbeiten, ohne dass datensatzspezifische Architekturen erforderlich sind. Zur weiteren Untersuchung der Frage wurde der Datensatz RaDIal [14] als am geeignetsten bewertet.

Experimenteller Versuchsaufbau

In jüngster Zeit wurden Transformer, die für ihren Erfolg bei der Verarbeitung natürlicher Sprache bekannt sind, für Computer-Vision-Aufgaben angepasst und bieten ein neues Paradigma für die Objekterkennung. Ein Beispiel für solche Transformer-Architekturen ist YOLOS (You Only Look at One Sequence), ein einfaches, aber effektives Modell zur Objekterkennung. YOLOS [16] ist ein Objekterkennungsmodell, das auf der Vision Transformer

(ViT)-Architektur mit minimalen Änderungen basiert. Die primäre Transformation in YOLOS besteht darin, das [CLS]-Token, das in ViT für die Bildklassifikation verwendet wird, durch hundert [DET]-Token für die Objekterkennung zu ersetzen und den Verlust bei der Bildklassifikation durch einen bipartiten Matching-Verlust für die Objekterkennung in Form einer Mengenvorhersage zu ersetzen. YOLOS übernimmt die Eigenschaften der Transformer-Modelle und wendet sie für Objekterkennung an. Im Rahmen der Projektarbeit wurde schließlich ein Finetuning-Ansatz basierend auf YOLOS implementiert.

Für die Experimente wurde eine Huggingface-Bibliothek mit Implementierungen des YOLOS-Modells verwendet. Das Modell benutzt als Basis einen Vision Transformer, der auf dem ImageNet-1k-Datensatz trainiert wurde. Tabelle 2 zeigt die verfügbaren Modell-Checkpoints, die mit der huggingface-Bibliothek geladen werden können.

Model	Pre-train Epochs	ViT (DeiT) Weight / Log	Fine-tune Epochs	Eval Size	YOLOS Checkpoint / Log	AP @ COCO val
`YOLOS-Ti`	300	FB	300	512	Baidu Drive, Google Drive / Log	28.7
`YOLOS-S`	200	Baidu Drive, Google Drive / Log	150	800	Baidu Drive, Google Drive / Log	36.1
`YOLOS-S`	300	FB	150	800	Baidu Drive, Google Drive / Log	36.1
`YOLOS-S (dWr)`	300	Baidu Drive, Google Drive / Log	150	800	Baidu Drive, Google Drive / Log	37.6
`YOLOS-B`	1000	FB	150	800	Baidu Drive, Google Drive / Log	42.0

Tabelle 2: Übersicht der verfügbaren vortrainierten YOLOS-Modelle.

Auf Basis dieses Setups wurden Experimente durchgeführt, um einen Finetuning-Ansatz zu bewerten. Hierzu wurden drei verschiedene Finetuning-Schemata evaluiert:

1) vollständiges Finetuning aller Modellparameter,
2) Finetuning nur der Normalization Layers des vortrainierten Modells und
3) Hinzufügen von Adapterschichten zum parametereffizienten Finetuning.

Bei der ersten Methode, dem **vollständigen Feintuning,** werden alle Gewichte des Modells während des Trainingsprozesses für den neuen Datensatz aktualisiert. Dieser Ansatz kann zu einem Modell führen, das hochgradig auf die Zielaufgabe spezialisiert ist, allerdings auf Kosten von Rechenressourcen und einer potenziellen Überanpassung an den neuen Datensatz.

Die zweite Methode, das Finetuning **nur der Normalization Layers,** ist ein parameterschonenderer Ansatz. Hier werden während des Trainings nur die Parameter der Normalisierung des vortrainierten Modells aktualisiert, während der Rest der Gewichte

unverändert bleibt. Diese Methode stützt sich auf die bereits gelernten Repräsentationen des vortrainierten Modells und nimmt geringfügige Anpassungen vor, um sich der neuen Datensatzstatistik anzupassen.

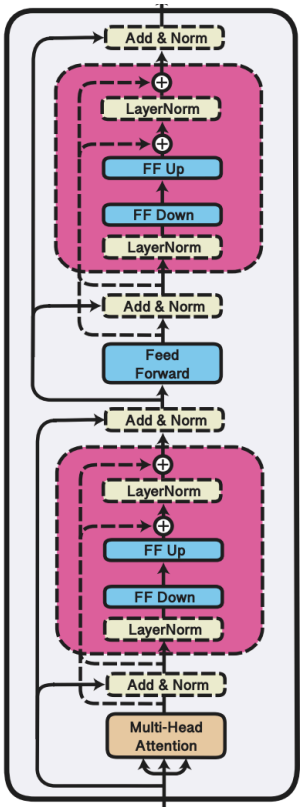


Abbildung 4: Vorgeschlagene Konfiguration der Adapterschichten (rosa Kästchen) innerhalb der traditionellen Transformer-Architektur [17]. Nur die gestrichelten Kästen stellen Schichten mit trainierbaren Parametern dar.

Bei der dritten Methode werden **Adapterschichten** hinzugefügt [18]. Bei den Adaptern handelt es sich um kleine, trainierbare Module, die in jede Schicht des Transformer-Modells eingefügt werden und an die jeweilige Aufgabe angepasst werden können, während die übrigen Parameter des Modells fest bleiben. Diese Adapterschichten bestehen aus einer Abwärtsprojektion, die die Dimensionalität reduziert, einer Nichtlinearität und einer Aufwärtsprojektion, die die ursprüngliche Dimensionalität wiederherstellt, wodurch eine Engpassstruktur entsteht, die aufgabenspezifische Repräsentationen lernt. Adapter ermöglichen ein effizientes Transferlernen, da sie nur wenige zusätzliche Parameter trainieren müssen und leicht für verschiedene Aufgaben und Modelle verwendet werden können.

Diskussion der Ergebnisse

Die folgenden Abbildungen zeigen die Ergebnisse der durchgeführten Experimente.

Abbildung 5: Trainingslossfunktion der untersuchten Finetuningansätze. Links: gesamte Lossfunktion. Rechts: Klassifizierungslossfunktion

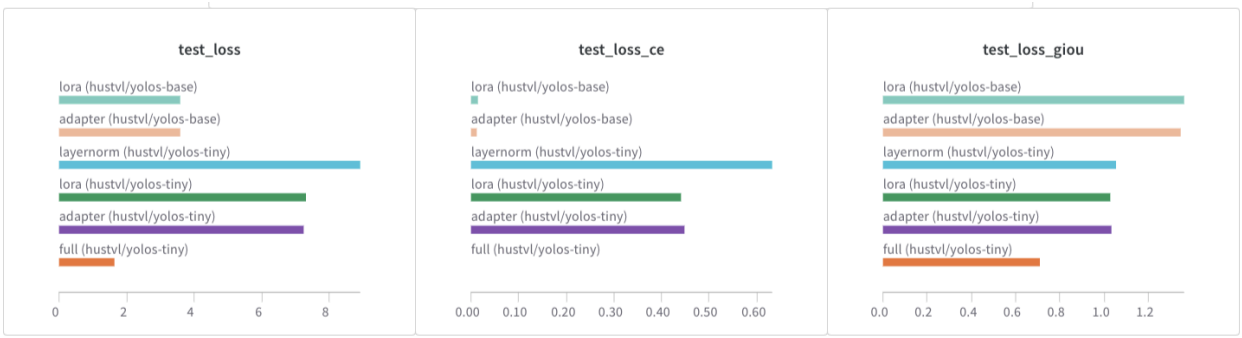


Abbildung 6: Testloss der untersuchten Finetuningansätze.

Insgesamt kann eine leichte bis moderate Verbesserung der Verlustwerte während des Trainings beobachtet werden, was die prinzipielle Eignung des Ansatzes bestätigt.

Im Allgemeinen kann man feststellen, dass das vollständige Finetuning des `tiny` Basismodells bei allen gemessenen Metriken am besten abschneidet. Im Vergleich zu den parametereffizienten Finetuning-Strategien `layer-norm`, `adapter` und `lora` ist dies bei identischen Backbone-Modellen nicht überraschend (da das Finetuning aller Parameter als Obermenge in Bezug auf das zugrundeliegende Optimierungsproblem angesehen werden kann). Die Ergebnisse sind jedoch überraschend, wenn ein größeres Backbone verwendet wird, d.h. das Basismodell, das nicht vollständig finegetuned werden kann und nur mit den oben genannten Adapteransätzen übertragen werden kann. Während (erwartungsgemäß) starke Verbesserungen gegenüber identischen Adapteransätzen des Tiny-Modells beobachtet werden können, schneidet das mit dem Adapter abgestimmte Basismodell immer noch schlechter ab als das vollständig fein abgestimmte Tiny-Modell.

Wie erwartet, schneidet unter den parametereffizienten Finetuning-Ansätzen der Layernorm-Ansatz am schlechtesten ab, während die Adapter- und LoRA-basierten Methoden ähnlich gut abschneiden – unabhängig vom verwendeten Backbone-Modell.

Für alle Modelle werden jedoch nicht-ideale Bounding-Box-Vorhersagefähigkeiten sichtbar. Während der L1-basierte Verlust bbox gut konvergiert, insbesondere im Fall des vollständig fein abgestimmten Tiny-Modells, zeigt die entsprechende *Generalized Intersection over Union (gIoU)* nur geringe Verbesserungen im Laufe des Trainings. Dies könnte darauf hindeuten, dass zwar die allgemeine Größe der Bounding-Boxes korrekt geschätzt wird, die exakten Positionen jedoch nur schlecht geschätzt werden. Überraschenderweise ist bei den auf den Adapter abgestimmten Basismodellen sogar eine Zunahme von gIoU während des Trainings zu beobachten. Man beachte, dass gIoU zwischen -1 und 1 liegt (wobei 1 einer perfekten Anpassung entspricht) und gegen -1 tendiert, wenn sich die beiden Boundingboxen nicht überlappen, sondern der geometrische Abstand ihrer Zentren zunimmt.

Schließlich zeigt der gemessene Kardinalitätsfehler keine Verbesserungen durch Training. Das heißt, die Anzahl der falsch vorgeschlagenen Übereinstimmungen, d. h. der Objekte, konvergiert im Durchschnitt nicht wie gewünscht gegen Null. Daher ist im Durchschnitt ein gewisses "Erkennungsrauschen" in allen Vorhersagen vorhanden. Diese Beobachtung ist überraschenderweise unabhängig vom gewählten Backbone-Modell und vom Finetuning-Ansatz.

Literatur

[1] P. Zhao, C. X. Lu, B. Wang, N. Trigoni, and A. Markham, ‘CubeLearn: End-to-end Learning for Human Motion Recognition from Raw mmWave Radar Signals’, *arXiv*, 2021, doi: 10.48550/arxiv.2111.03976.

[2] M. Ulrich *et al.*, ‘Improved Orientation Estimation and Detection with Hybrid Object Detection Networks for Automotive Radar’, *2022 IEEE 25th Int. Conf. Intell. Transp. Syst. ITSC*, vol. 00, pp. 111–117, 2022, doi: 10.1109/itsc55140.2022.9922457.

[3] D. Maturana and S. Scherer, ‘3D Convolutional Neural Networks for Landing Zone Detection from LiDAR’, *2015 IEEE Int. Conf. Robot. Autom. ICRA*, pp. 3471–3478, 2015, doi: 10.1109/icra.2015.7139679.

[4] B. Yang, W. Luo, and R. Urtasun, ‘PIXOR: Real-Time 3D Object Detection from Point Clouds’, *2018 IEEECVF Conf. Comput. Vis. Pattern Recognit.*, pp. 7652–7660, 2018, doi: 10.1109/cvpr.2018.00798.

[5] A. Prabhakara *et al.*, ‘High-Resolution Point Clouds from mmWave Radar’. [Online]. Available: https://arxiv.org/pdf/2206.09273.pdf

[6] R. Wiersma, A. Nasikun, E. Eisemann, and K. Hildebrandt, ‘DeltaConv: anisotropic operators for geometric deep learning on point clouds’, *ACM Trans. Graph.*, vol. 41, no. 4, pp. 1–10, 2022, doi: 10.1145/3528223.3530166.

[7] D.-H. Paek, S.-H. Kong, and K. T. Wijaya, ‘K-Radar: 4D Radar Object Detection Dataset and Benchmark for Autonomous Driving in Various Weather Conditions’. [Online]. Available: https://arxiv.org/pdf/2206.08171v2.pdf

[8] Y. Sun, Z. Huang, H. Zhang, Z. Cao, and D. Xu, ‘3DRIMR: 3D Reconstruction and Imaging via mmWave Radar based on Deep Learning’, *arXiv*, 2021, doi: 10.48550/arxiv.2108.02858.

[9] T. Li, Y. Fu, X. Han, H. Liang, J. J. Zhang, and J. Chang, ‘DiffusionPointLabel: Annotated Point Cloud Generation with Diffusion Model’, 2022, [Online]. Available: https://app.readcube.com/library/10a9c178-0b2d-4970-9231-33fbc56c588f/item/946f7911-4c96-4aea-b0e1-0d36cae69d7b

[10] A. Dubey, A. Santra, J. Fuchs, M. Lübke, R. Weigel, and F. Lurz, ‘HARadNet: Anchor-free target detection for radar point clouds using hierarchical attention and multi-task learning’, *Mach. Learn. Appl.*, vol. 8, p. 100275, 2022, doi: 10.1016/j.mlwa.2022.100275.

[11] J. Yang, X. Chen, H. Zou, D. Wang, and L. Xie, ‘AutoFi: Towards Automatic WiFi Human Sensing via Geometric Self-Supervised Learning’, *arXiv*, 2022, doi: 10.48550/arxiv.2205.01629.

[12] M. Gadd, D. D. Martini, and P. Newman, ‘Contrastive Learning for Unsupervised Radar Place Recognition’, *arXiv*, 2021, doi: 10.48550/arxiv.2110.02744.

[13] Y. Zhang, D. Sidibé, O. Morel, and F. Mériaudeau, ‘Deep multimodal fusion for semantic image segmentation: A survey’, *Image Vis. Comput.*, vol. 105, p. 104042, 2021, doi: 10.1016/j.imavis.2020.104042.

[14] J. Rebut, A. Ouaknine, W. Malik, and P. Pérez, ‘Raw High-Definition Radar for Multi-Task Learning’, *arXiv*, 2021, doi: 10.48550/arxiv.2112.10646.

[15] N. Carion, F. Massa, G. Synnaeve, N. Usunier, A. Kirillov, and S. Zagoruyko, ‘End-to-End Object Detection with Transformers’, *arXiv*, 2020, doi: 10.48550/arxiv.2005.12872.

[16] Y. Fang *et al.*, ‘You Only Look at One Sequence: Rethinking Transformer in Vision through Object Detection’, *arXiv*, 2021, doi: 10.48550/arxiv.2106.00666.

[17] J. Pfeiffer *et al.*, ‘AdapterHub: A Framework for Adapting Transformers’. arXiv, Oct. 06, 2020. Accessed: Dec. 21, 2023. [Online]. Available: http://arxiv.org/abs/2007.07779

[18] A. GIurgiu *et al.*, ‘Parameter-Efficient Transfer Learning for NLP’. [Online]. Available: http://proceedings.mlr.press/v97/houlsby19a/houlsby19a.pdf

AP6: Praktische Untersuchungen und Proof-of-Concept Demonstration

Die Arbeiten an AP6 wurden während der Dauer der Projektbeteiligung noch nicht begonnen.

2. Wichtigste Positionen des zahlenmäßigen Nachweises

Das Gros des Budgets entfiel auf Personalkosten, die etwa 97% der verausgabten Gesamtkosten ausmachen. Die restlichen Aufwände waren bedingt durch Reisekosten (ca. 1%) sowie Aufwendungen für Cloud-Rechenleistung (ca. 2%).

3. Notwendigkeit und Angemessenheit der geleisteten Arbeit

Die durchgeführten Forschungsarbeiten im Verbundprojekt 6G-ICAS4Mobility sowie die dafür aufgewandten Ressourcen waren notwendig und angemessen, da sie der im Projektantrag bis zum Ausstieg formulierten Planung entsprechen und alle wesentlichen im Arbeitsplan für die ersten 16 Projektmonate formulierten Aufgaben erfolgreich bearbeitet wurden.

Die durch Merantix Momentum erbrachten Arbeiten und erzielten Ergebnisse hätten ohne die bereitgestellte Förderung nicht durchgeführt werden können, da nur dadurch der umfassende Austausch mit den Projektpartnern im Konsortium ermöglicht wurde.

4. Voraussichtlicher Nutzen

Merantix Momentum ist darauf spezialisiert, aktuelle Forschungsergebnisse im Bereich des maschinellen Lernens (ML) in industrielle Anwendungen zu überführen. Durch Merantix Momentum konnten im Vorhaben 6G-ICAS4Mobility die folgenden Projektergebnisse erreicht werden:

- Aufbau von Domänenwissen im Bereich von Sensorik im Automotive-Umfeld
- Beiträge zur Use-Case-Definition, Anforderungsanalyse und Systemarchitektur
- Vergleich von ML-Methoden zur Erkennung von Objekten in Radar- und vergleichbaren Datenmodalitäten
- Verwendung von öffentlichen Datensätzen zur Verbesserung des ML-Trainings
- Finetuning von Transformer-basierten ML-Netzarchitekturen zur Objekterkennung

Der Nutzen des Vorhabens 6G-ICAS4Mobility für die Merantix Momentum GmbH schlägt sich im Kern in der Weiterverwertung der gewonnenen Erkenntnisse und Ergebnisse im Kundengeschäft, also der Umsetzung von Machine-Learning-Lösungen für (Industrie-)Kunden, nieder.

Konkret stellen sich die folgenden Verwertungsoptionen dar:

- Beratung von Kunden bei der Ideation, Konzeption und Umsetzung von KI- bzw. ML-Projekten
- Technische Realisierung von Proof-of-Concepts mithilfe der im Projekt gewonnenen methodischen Erkenntnisse
- Weiterentwicklung der methodischen Erkenntnisse zu wiederverwendbaren Komponenten mit der Option, diese als Open Source zur Verfügung zu stellen

5. Während der Durchführung des Vorhabens dem ZE bekannt gewordenen Fortschritts auf dem Gebiet des Vorhabens bei anderen Stellen

Während des Projektzeitraums hat Merantix Momentum intensiv den wissenschaftlichen Fortschritt im Bereich des Maschinellen Lernens gemonitort und die identifizierten relevanten Arbeiten in die Konzeption der eigenen Forschungsansätze einfließen lassen. Wesentliche Quellen hierfür waren öffentliche Publikationsplattformen. Die gewonnenen Erkenntnisse bildeten beispielsweise die Grundlage für die Literaturliste, welche in Kap. II.1 dargestellt und erläutert wird.

6. Erfolgte oder geplante Veröffentlichungen des Ergebnisses

Aus dem Teilvorhaben heraus wurde seitens Merantix der in Kap. II.1 genannte Blogbeitrag veröffentlicht. Darüber hinaus sind keine weiteren Veröffentlichungen geplant.