



Ressourceneffiziente KI für eingebettete Systeme in Landmaschinen

Schlussbericht 2021 - 2024

Zuwendungsempfänger: Deutsches Forschungszentrum für Künstliche Intelligenz (DFKI) GmbH	Förderkennzeichen: 28DK102D20
Ausführende Stelle: Standort Niedersachsen, Forschungsbereich Planbasierte Robotersteuerung (Prof. Dr. Joachim Hertzberg)	
Vorhabenbezeichnung: resKIL	
Laufzeit des Vorhabens: 01.01.2021 - 30.06.2024	
Berichtszeitraum: 01.01.2021 - 30.06.2024	

Inhaltsverzeichnis

Abkürzungsverzeichnis.....	3
I. Kurze Darstellung zu:	4
1. Aufgabenstellung.....	4
2. Voraussetzungen, unter denen das Vorhaben durchgeführt wurde	6
3. Planung und Ablauf des Vorhabens	6
4. Wissenschaftlicher und technischer Stand, an den angeknüpft wurde.....	8
4.1 Angabe bekannter Konstruktionen, Verfahren und Schutzrechte, die für die Durchführung des Vorhabens benutzt wurden	9
4.2 Angabe der verwendeten Fachliteratur sowie der benutzten Informations- und Dokumentationsdienste	10
5. Zusammenarbeit mit anderen Stellen	10
II. Eingehende Darstellung:	10
1. des erzielten Ergebnisses	10
2. des voraussichtlichen Nutzens, insbesondere der Verwertbarkeit des Ergebnisses und der Erfahrungen,	20
3. des während der Durchführung des Vorhabens dem Zuwendungsempfänger bekannt gewordenen Fortschritts auf diesem Gebiet bei anderen Stellen,	20
4. der erfolgten oder geplanten Veröffentlichung des Ergebnisses	21
5. Anhang.....	22
5.1 Literaturverzeichnis	Fehler! Textmarke nicht definiert.

Abkürzungsverzeichnis

DFKI	Deutsches Forschungszentrum für Künstliche Intelligenz (Projektpartner)
API	Application Programming Interface
IoU	Intersection over Union
CAN-Bus	Control Area Network Bus
FPS	Frame per second
KI	Künstliche Intelligenz
LL	Zauberzeug Learning Loop
PBR	Planbasierte Robotersteuerung (Forschungsbereich des DFKI – Labor Niedersachsen)

I. Kurze Darstellung zu:

1. Aufgabenstellung

Das Projekt resKIL (Ressourceneffiziente KI für eingebettete Systeme in Landmaschinen) hat das Ziel, durch detaillierte Analyse und Generalisierung der Problemstellungen, Entwicklung einer variablen Softwarearchitektur und Entwicklung einer KI-Toolchain, ganzheitlich Ressourcen (sowohl Personen als auch Hardware) zu optimieren. Dadurch wird eine skalierbare und variable Lösung entwickelt, damit maschinelles Lernen in naher Zukunft eine deutlich größere Ausbreitung in der Landwirtschaft erfährt. Es erfolgt dabei eine Fokussierung, auf die im Folgenden beschriebenen Teilbereiche. Die Entwicklungen werden am Beispiel der Qualitätsbestimmung von Erntegut und der Merkmalerkennung im Maschinenumfeld in Funktionsmuster integriert und umgesetzt. Hierbei wird eine agile Entwicklungsstrategie verfolgt, um eine stetige Steigerung der Qualität zu gewährleisten. Im Zentrum der Projektplanung stehen drei Ernteperioden. Diese legen den Ablauf und die Iterationen der einzelnen Arbeitspakete fest. Gerade der Transfer der Entwicklungen über mehrere Ernten hinweg ist ein weiterer Bestandteil, um die Generalisierbarkeit und darüber hinaus eine Steigerung der Robustheit und Qualität der Ansätze zeigen zu können.

Das Projekt orientiert sich an den strategischen Zielen des BMEL, die Potenziale KI-basierter Methoden und Technologien in der Landwirtschaft und Agrartechnik zu erschließen. Sie orientieren sich eng an den 17 Zielen für nachhaltige Entwicklung der Vereinten Nationen (SDG: Sustainable Development Goals). Das Projekt resKIL liefert unmittelbar und mittelbar Beiträge zu den folgenden Zielen: *Ziel 2*: kein Hunger, *Ziel 9*: Industrie, Innovation und Infrastruktur, *Ziel 12*: Nachhaltige Produktion, *Ziel 13*: Maßnahmen zum Klimaschutz, Verständnis komplexer Systeme der Natur, der Ökonomie und von gesellschaftlichem Fortschritt. Dafür sind aus der Grundlagenforschung stammende Verfahren und Techniken in der Praxis nutzbar zu machen, datenbasierte Anwendungen zu entwickeln und die Landwirtschaft nachhaltig zu gestalten. Das Projekt resKIL adressiert die definierten Förderschwerpunkte in unterschiedlicher Intensität:

Maschinelles Lernen: Es werden Verfahren des maschinellen Lernens zur Verwendung in landwirtschaftlichen Arbeitsmaschinen/Verfahren („on the edge“) entwickelt bzw. Adaptiert.

Deep Learning: Methoden des Deep Learnings werden intensiv bearbeitet, weil sie im Hinblick auf die Anwendungsfälle besonders erfolgversprechend erscheinen.

Intelligente Maschinen: CLAAS-Landmaschinen sind bereits heute durch diverse zeitlich und räumlich vorausschauende Assistenzsysteme gekennzeichnet. „System of Systems“ und „Digital Twin“ sind in Teilen realisiert. Die Förderung einer weitergehenden Autonomie und Verbesserung der Arbeitsqualität durch KI-basierte Methoden „on the edge“ führt tendenziell zu intelligenteren Maschinen, falls der Begriff in diesem Zusammenhang zulässig ist.

Maschinelles Planen und Handeln: Vorausschauendes Planen und Handeln ist in der Landwirtschaft eine besondere Herausforderung, die aufgrund des geringen Determinismus der Umgebung erschwert ist. Maschinelle Unterstützung ist in diesem Zusammenhang von großem Nutzen und wird im Rahmen des Projektes angestrebt.

Mustererkennung: Die beiden Anwendungsfälle zielen auf eine Mustererkennung. Im Fall der Gutflussanalyse sollen Bestandteile und deren Merkmale erkannt werden und darauf aufbauend verbesserte Einstellungen realisiert werden. Im Fall der Umgebungsanalyse sind Objekte zu klassifizieren.

Wissensbasierte Systeme: In die zu entwickelten KI-Methoden fließt Expertenwissen ein, so z.B. bei der Annotation von Trainingsdaten. Weitere Forderungen werden ebenfalls erfüllt:

Praxisorientierung: Die entwickelten KI-Algorithmen mit den zugehörigen eingebetteten Systemen werden unter Praxisbedingungen evaluiert. Dabei wird darauf geachtet, dass durch möglichst generische Ansätze und Vorgehensweisen auch weitere Anwendungsfälle zeitnah nach dem Projekt umsetzbar sein.

Veröffentlichungsfähige Ergebnisse: Im Projekt resKIL sind drei akademische Partner engagiert, die u.a. an der Qualität ihrer Veröffentlichungen gemessen werden. Aber auch die Industriepartner haben ein starkes Interesse an der Veröffentlichung von Ergebnissen in ihrer Community. Hier sind jedoch schutzwürdige Interessen zu beachten. Das haben wir aber auch in vergangenen Förderprojekten immer gut umgesetzt.

2. Wissenschaftliche und technische Arbeitsziele des Vorhabens

Um den aktuellen Fortschritt des Projektes erlebbar machen zu können, steht die kontinuierliche Entwicklung und Optimierung von Funktionsmustern im Mittelpunkt. Diese basieren direkt auf den gewonnenen Erkenntnissen und Daten der Ernteperioden. Somit wird garantiert, dass die angewandte Forschung auch in die Praxis und den praktischen Einsatz übertragen werden kann. Zur Sicherstellung einer Skalierbarkeit der entwickelten Systeme über unterschiedliche Problemstellungen hinweg, werden zwei deutlich unterscheidbare Funktionsmuster kontinuierlich entwickelt und evaluiert:

1. Bestimmung der Qualität von Erntegut beim Einsatz einer Erntemaschine.
2. Erkennung von Merkmalen im Umfeld einer landwirtschaftlichen Arbeitsmaschine

Im Projekt wird eine ganzheitliche Betrachtung und Optimierung der benötigten Ressourcen verfolgt. Dabei wird ein Schwerpunkt auf eine Steigerung der Effizienz der in den folgenden beschreibenden Bereichen gelegt.

Prototypische Integration von KI-Technologien in landtechnische Systeme: Die in diesem Projekt zu entwickelnden bzw. anzupassenden KI-Algorithmen werden in die von CLAAS zur Verfügung gestellten landtechnischen Systeme (CLAAS Landmaschinen) integriert und zur Erfüllung der definierten Funktionsmuster implementiert.

Effizienter Softwareentwurf von KI-Systemen: Eine variable KI-Architektur wird entwickelt, die auf eingebetteten, autarken Systemen, auf Edge-Systemen mit temporärer Cloud-Verbindung und auf vollständigen Cloud-Systemen angewendet werden kann.

Effiziente KI-Hardware: Gerade eingebettete Systeme bieten ein großes Potential zur Effizienzsteigerung. Durch eine genaue Analyse der zu lösenden Problemgruppen, der anzuwendenden Algorithmen und der geforderten Qualität wird ein Konzept entwickelt und umgesetzt, um eingebettete Systeme skalierbar und mit wenig Aufwand an die geforderten Gegebenheiten anpassen zu können.

Effiziente Datengewinnung: Auch die Datengewinnung bietet Potential der Ressourcenoptimierung. Es wird ein Konzept entwickelt und umgesetzt, das besonders bei der Planung von Datenaufnahmen angewendet werden kann. Durch eine statistische Analyse der Problemstellung, der verwendeten Algorithmen und des betrachteten Zielraumes kann die Datenaufnahme gezielt gelenkt und somit verbessert werden.

Effiziente Datenaufbereitung und -annotation: Die Aufbereitung und Auswahl der zu verwenden Daten soll weitestgehend automatisiert erfolgen. Des Weiteren soll zum einen der Bedarf an Expertenwissen bei der Annotation der Daten auf ein Minimum reduziert, zum anderen das Annotieren selbst optimiert werden. Hierbei werden Ansätze verfolgt, die den Nutzer

bestmöglich während der Annotation unterstützen. So wird beispielsweise eine Mensch-Maschine-Schnittstelle umgesetzt, die eine möglichst hohe Qualität der Annotationen gewährleistet. Außerdem werden im Laufe des Modelltrainings, nur tatsächlich benötigte Annotationen (hohe Modellunsicherheit) vom System beim Nutzer angefragt.

Effiziente Berechnung / optimierte KI-Algorithmen: Ziel hierbei ist das Zusammenbringen der Genauigkeitsanforderungen aus Sicht der Automatisierung und des technisch Machbaren. Ein möglicher gangbarer Weg ist dabei beispielsweise die Reduktion des neuronalen Netzes um einige Layer unter Akzeptanz von Genauigkeitsverlust bei gleichzeitiger Optimierung der Inferenzzeit¹. Zu beachten sind hierbei die grundsätzlichen Anforderungen an die Genauigkeit des Systems.

3. Voraussetzungen, unter denen das Vorhaben durchgeführt wurde

Die Integration von KI-Technologien in landwirtschaftliche Arbeitsmaschinen und Arbeitsverfahren ist mit verschiedenen Restriktionen verbunden, z.B. eingebettete Systeme und Netzwerke mit vergleichsweise geringer Leistung, widrige Umgebungsbedingungen (Temperatur, Vibration, Staub, Feuchtigkeit, Bauraum, Stromversorgung, elektromagnetische Strahlung, Funknetzverfügbarkeit, domänenspezifische Standards). Die Funktionsfähigkeit soll unter Berücksichtigung der genannten Restriktionen erfolgen. Gerade in Hinblick auf die Konnektivität bewegen sich Landmaschinen in einer stark variablen Umgebung. Diese reicht von keiner Verbindung bis hin zu Umgebungen mit 4G-Verfügbarkeit. Dies erschwert nicht nur die Bereitstellung eines aktualisierten Modells auf der Edge nach dem Training, sondern auch die Notwendigkeit einer Cloud-Edge-Konnektivität für die Inferenz.

Bei den Ressourcenbeschränkungen der eingebetteten Systeme (Edge) spielt die begrenzte Speicherkapazität eine entscheidende Rolle. Sie beeinträchtigt sowohl die Speicherung des Modellgewichts als auch die Ausführung von Deep-Learning-Modellen. Die Genauigkeit dieser Modelle ist jedoch direkt von der Anzahl der Schichten bzw. der Tiefe des Modells abhängig, welche sich in der Größe des Speicherplatzes der Gewichte widerspiegelt. Des Weiteren limitiert die begrenzte Rechenleistung die Möglichkeit zur Echtzeit-Ausführung von Deep-Learning-Modellen, was eine optimierte Mensch-Maschine-Interaktion erschwert. Schließlich weisen die eingebetteten Hardware-Systemkomponenten im Vergleich zur Cloud schwache KI-Recheneinheiten (AI-Accelerator) auf. Was die Datenerhebung für die Verbesserung der KI-Modelle betrifft, stellt die ständig wechselnde Umgebung (Beleuchtung, Pflanzenwachstumsstadien, Staub, Verschmutzung usw.) eine Herausforderung dar. Darüber hinaus ist die Datenerhebung in der Landwirtschaft eingeschränkt und kann nicht kontinuierlich stattfinden, da einigen Pflanzarten nur einmal pro Jahr wachsen.

4. Planung und Ablauf des Vorhabens

Das Vorhaben besteht aus insgesamt 6 Arbeitspaketen, in denen die Arbeit der verschiedenen Partner organisiert ist (Tabelle 1 und Tabelle 2). Das DFKI ist an den Arbeiten in den Arbeitspaketen 2, 3, 5 und 6 beteiligt. Ein weiteres Arbeitspaket (AP 7) beschreibt die Koordination.

Im Laufe des resKIL-Projekts fanden 3 Meilensteine nach jedem vollen Jahr statt. Dies entspricht gut den im Projekt 3 geplanten Ernten, die für die Datenaufnahme, Anwendung des entwickelten Konzepts sowie die Auswertung der Ergebnisse dienen. Darüber hinaus finden auch während des Projekts monatlich Core-Meetings statt. Dabei trafen alle beteiligten Partner in einem kleinen Kreis zusammen und diskutierten über den Stand des Projekts. **Fehler! Verweisquelle konnte nicht gefunden werden.** zeigt den Arbeitsplan des Vorhabens mit einer

Aufstellung der verschiedenen Arbeitspakete und deren Laufzeit. Ein grobes Blockdiagramm der verschiedenen Arbeitspakete und deren Beziehungen ist in Abbildung 1 dargestellt.

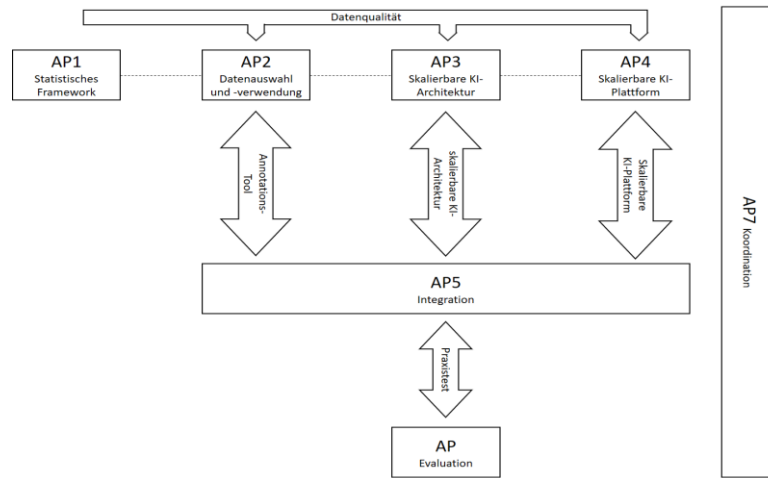


Abbildung 1: Zusammenfassung der Arbeitspakete und der wichtigsten Beziehungen zwischen ihnen

Projektjahr = Kalenderjahr		Jahr 1												Jahr 2												Jahr 3											
		1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	4	5	6	7	8	9	10	11	12
AP 1	Statistisches Framework zur Datenakquise und -aufbereitung																																				
AP 1.1	Stichproben- und Versuchsplanung																																				
AP 1.2	Datenqualitätsbewertung und Datenaufbereitung																																				
AP 2	Framework zur Datenauswahl/ -verwendung																																				
AP 2.1	Prozessplanung																																				
AP 2.2	Implementierung effizienter Annotations-Methoden																																				
AP 2.3	Prozessintegration																																				
AP 3	Skalierbare KI-Architektur																																				
AP 3.1	Anforderungsanalyse																																				
AP 3.2	Entwicklung der Abstraktionsschicht																																				
AP 3.3	Entwicklung der Architektur																																				
AP 4	Skalierbare KI-Plattform																																				
AP 4.1	Entwurfsraumanalyse																																				
AP 4.2	Effizienzbestimmung und Optimierung																																				
AP 5	Integration und Demonstratorentwicklung																																				
AP 5.1	Umsetzung der KI-Plattform																																				
AP 5.2	Umsetzung der KI-Architektur auf der Edge																																				
AP 5.3	Umsetzung der KI-Architektur in der Cloud																																				
AP 5.4	Kontinuierlicher Prozess der KI-Modellverbesserung																																				
AP 6	Evaluation																																				
AP 6.1	Erntekampagnen																																				
AP 6.2	Evaluation der Datenauswahl und -verwendung																																				
AP 6.3	Evaluation Human-in-Loop Labeling / Active Learning																																				
AP 6.4	Evaluation der Architektur-/Plattformintegration																																				
AP 6.5	Evaluation der trainierten Modelle																																				
AP 7	Projektkoordination																																				

Tabelle 1: Arbeitsplan des Vorhabens

Meilenstein	Monat	Beschreibung
M1	12	Auf Basis der statistischen Versuchsplanung wurden in der ersten Erntekampagne umfassende Trainingsdaten gewonnen. Außerdem liegen Umsetzungskonzepte für Framework, Architektur und Plattform vor.
M2	24	Die aufgenommenen Daten wurden mit dem entwickelten Tool annotiert und der quantitative Nachweis für eine Reduktion des Annotationsaufwands erbracht. Framework, Architektur und Plattform liegen als Funktionsmuster/Demonstrator vor und wurden in der zweiten Ernte erprobt.
M3	36	Der Funktionsnachweis für die Anwendung von KI-Verfahren zur Ausführung auf eingebetteten Systemen ist erbracht. Die Skalierbarkeit und Generalisierbarkeit der Methoden wurden durch verschiedene Anwendungen sowohl auf der Landmaschine als auch in der Cloud gezeigt.

Tabelle 2: Meilensteinplanung des Vorhabens

5. Wissenschaftlicher und technischer Stand, an den angeknüpft wurde

Die folgenden Absätze beschreiben den wissenschaftlichen und technischen Stand in Bezug auf AP2: das Framework zur Datenauswahl und -verwendung, AP3: skalierbare KI-Architektur sowie AP5: Integration und Funktionsmusterentwicklung. Diese Arbeitspakete stellen für das DFKI den Ort dar, an dem die meisten technischen Konzeptionierungen und Entwicklungen stattgefunden haben.

Die Qualität der Daten spielt eine wichtige Rolle für die Leistung des trainierten maschinellen Lernmodells. In der Landwirtschaft können zwar Daten für einen kurzen Zeitraum erfasst werden, aber die Menge der gesammelten Daten kann umfangreich sein. Daher sind leistungsfähige Systeme wie Cloud-Instanzen für die Speicherung, Verwaltung und Verarbeitung dieser Daten erforderlich. Eine intelligente Auswahl informativer und repräsentativer Daten für das Training eines KI-Modells ist daher unerlässlich. Active Learning wurde von (Settles, 2009) als alternatives Lernverfahren zum klassischen Supervised Learning eingeführt. Beim Active Learning wählt das Modell aus, welche Daten für das Training wichtig sind, und nur diese Daten werden von Menschen gelabelt und für das Training verwendet. Es gibt verschiedene Ansätze zur intelligenten Auswahl der Daten. *Informative Datenauswahl*: Hierbei werden Daten ausgewählt, bei denen das Modell unsicher ist. Beispiele sind die Cereals-Methode (Mackowiak, 2018), „Region-based Active Learning“ (Kasarla, 2019, January), ViewAI (Siddiqui, 2020) und „Distributed Active Learning“ (Puck, 2021). Bei der *Repräsentative Datenauswahl* werden Datenpunkte anhand der Struktur der unlabeled Daten ausgewählt, um eine repräsentative Stichprobe zu erhalten. Ein Beispiel ist die Core-Set-Methode (Sener, 2017).

In der Bildverarbeitung ist die Pool-basierte Active-Learning-Methode weit verbreitet. Dabei wird aus einem großen Pool von ungelabelten Daten ein entsprechender Batch für das Training ausgewählt. Bekannte Ansätze sind DEAL (Xie, 2020), Task-Aware (Kim, 2021), View AL (Siddiqui, 2020), und DIAL (Lenczner, 2022). Es ist wichtig zu beachten, dass klassische veröffentlichte Active-Learning-Ansätze in der Bildverarbeitung oft den Nachteil haben, dass der Aufwand für das Labeling aller Bilder gleichgestellt ist. Das bedeutet, dass die Kosten für das Labeling eines Bildes konstant gesehen werden, obwohl dies nicht immer der Fall ist.

Insbesondere in der Landwirtschaft ist der Aufwand für das Pixel-Labeling eines Baums deutlich höher als der Aufwand für das Pixel-Labeling einer geraden Fahrspur. Dies wurde in der Arbeit von (Mackowiak, 2018) und (Kasarla, 2019, January) ausführlich untersucht. Daher werden Methoden vorgeschlagen, die die Kosten eines Bildes schätzen. Außerdem wird in allen verwandten Arbeiten dieses Projekts manuelles Labeling für das Pixel-Labeling der Bilder durchgeführt. Dieser Labeling-Prozess erfordert vom Labeler, dass er Punkte oder Klicks rund um den Rand eines Objekts im Bild eingibt. Dies ist besonders aufwändig für Objekte mit unregelmäßigen Formen wie Bäumen, Himmel, Menschen usw.

In einem parallelen und unabhängigen Forschungsfeld wird das Konzept der "interaktiven semantischen Segmentierung" von (Xu, 2016) eingeführt und weiterhin von (Sofiiuk K. P., 2016) und (Sofiiuk K. P., 2022) untersucht. In diesem Bereich wird ein KI-Modell für das Labeling verwendet. Der Labeler wählt nur anhand von Vordergrund- und Hintergrundklicks ein Objekt aus, und dies wird automatisch von einem interaktiven semantischen KI-Modell segmentiert oder anders gesagt mit einer Maske annotiert. Interaktive semantische KI-Modelle können nicht direkt für die Klassifikation von Feldmerkmalen verwendet werden, da sie als Eingabe ein Bild und einen Vorder- und Hintergrundpunkt für das zu annotierende Objekt erwarten.

Für die Erkennung von Merkmalen im Feld und die einfache Bereitstellung von KI-Modellen sowohl in der Cloud als auch auf Edge-Geräten wird auf Arbeiten im Bereich der semantischen Bildsegmentierung in Echtzeit zurückgegriffen. Dabei handelt es sich um spezielle KI-Architekturen, die im Gegensatz zu bekannten KI-Architekturen wie DeepLab (Chen, 2017) und HRNET (Wang, 2020) relativ klein und schnell sind. Diese Modelle eignen sich gut für eingebettete Systeme und können auch auf Cloud-Infrastrukturen leicht skaliert werden. Um höhere frames per seconds (FPS) bei gleichzeitig höherem Durchsatz zu erreichen, setzen die Autoren von DDRNET (Hong, 2021) und Bisenet (Yu, 2018) auf eine Architektur mit zwei parallelen Verzweigungen und leichten Encodern. In verwandten Arbeiten wie Fasdnet (Rosas-Arias, 2021) und ERFNET (Romera, 2017) werden hingegen einzelne Branches verwendet, die bestimmte Feature-Fusion-Module und Factorization Convolution nutzen.

Aktuelle Machine-Learning-Frameworks wie PyTorch (Pytorch, 2024) und TensorFlow (TensorFlow, 2024) ermöglichen es nun, KI-Modelle schnell und effizient für verschiedene Plattformen zu entwickeln und bereitzustellen. Alle Modelle, die in diesen Frameworks implementiert sind, können leicht im von der Community akzeptierten Exportformat ONNX (Bai, 2019) exportiert werden. ONNX lässt sich problemlos in plattformunabhängigen Code transformieren. Zum Beispiel ist es bei allen von NVIDIA unterstützten Hardwarekomponenten möglich, ONNX-exportierte Modelle zu verwenden.

5.1 Angabe bekannter Konstruktionen, Verfahren und Schutzrechte, die für die Durchführung des Vorhabens benutzt wurden

Die im Vorhaben entwickelte Software für die Datenauswahl und die KI-Architektur wurde in der Programmiersprache Python entwickelt, einer universellen höheren Programmiersprache, die von der gemeinnützigen Python Software Foundation unterstützt wird. Für die Umsetzung bestimmter Basisfunktionen wurden verschiedene etablierte KI Python-Bibliotheken verwendet, darunter PyTorch, TensorFlow, ONNX, NumPy (Hoyer, 2020) und Pandas. Die Programmierung erfolgte in der integrierten Entwicklungsumgebung Visual Studio Code (Microsoft, 2024) von Microsoft. Für die Versionsverwaltung des geschriebenen Codes wurde die Open-Source-Software GitLab (GitLab, 2024) eingesetzt.

Für die Implementierung der Module zur KI-gestützten Datenauswahl (Active Learning) und des effizienten Datenlabelings im Zauberzeug Zauberzeug Learning Loop Frontend-Weboberfläche (Zauberzeug, 2024) wurde neben den oben genannten KI-Frameworks auch das von Zauberzeug entwickelte Python-Framework Learning-Loop-Node für die Backend-Funktionalitäten verwendet. Alle Backend-Funktionalitäten sind als REST-API definiert, die mit dem Python Web-Framework FastAPI implementiert wurde. FastAPI (FastAPI, 2024) ist ein modernes Framework für schnelle und hochleistungsnahe Kommunikation zwischen Server und Clients. Alle verwendeten Computer verwenden Ubuntu als Betriebssystem. Unter Punkt 5 im Anhang finden sich die Verweise auf alle oben genannten Softwaretools.

5.2 Angabe der verwendeten Fachliteratur sowie der benutzten Informations- und Dokumentationsdienste

Unter Punkt **Fehler! Verweisquelle konnte nicht gefunden werden.** im Anhang findet sich ein ausführliches Literaturverzeichnis der verwendeten Fachliteratur, ebenso wie Hinweise und links auf technische Dokumentation und andere Informationen.

6. Zusammenarbeit mit anderen Stellen

Es fand eine enge Zusammenarbeit zwischen den Projektpartnern statt.

II. Eingehende Darstellung:

1. des erzielten Ergebnisses

Im Folgenden werden die Arbeiten und wichtigsten Ergebnisse zu den AP2, AP3 und AP5 des Projektpartners DFKI vorgestellt.

AP2: Framework zur Datenauswahl und -verwendung

Im Arbeitspaket 2 ist es geplant, ein ressourceneffizientes Annotationsverfahren zu konzipieren und umzusetzen. Dieses Verfahren zielt darauf ab, die Qualität der generierten Ground-Truth-Bilder zu verbessern und gleichzeitig den Aufwand für den Benutzer bei der Annotation zu minimieren. Zusätzlich sollen durch die Implementierung von Ressourcen-Einsparmethoden auf Basis des entwickelten Annotationsprozesses bestehende Prozesse innerhalb der "Zauberzeug Learning Loop" (LL) optimiert werden.

Das Deutsche Forschungszentrum für Künstliche Intelligenz (DFKI) hat zusammen mit weiteren Partnern ein ressourceneffizientes Annotationsverfahren vorgeschlagen und entwickelt. Dieses Verfahren baut auf zwei zentralen Ansätzen der Künstlichen Intelligenz (KI) auf, nämlich dem Active Learning und der Interactive Semantic Segmentation. Beide Ansätze dienen dazu, die Kosten für die Ressourcennutzung bei der Annotation zu senken. Das Konzept des Active Learning beinhaltet die Auswahl von Bildern aus einem großen Datensatz für die Annotation und das Training eines Modells. Durch diese Methode wird die Anzahl der zu annotierenden Bilder signifikant reduziert. Mit der Anwendung des Konzepts der Interactive Semantic Segmentation wird schließlich der Benutzeraufwand für die Annotation gesenkt. Hierfür unterstützt ein maschinelles Lernmodell den Benutzer im Hintergrund.

Abbildung 2 illustriert das resKIL Active Learning Loop Framework. Der Prozess beginnt direkt am Rand der Landmaschine, wo Daten aus dem Feld für das Training des KI-Modells aufgenommen werden. Sobald die Maschine Zugang zum Internet hat, werden alle relevanten Bilder vom Umfeld an die Cloud gesendet. Diese Bilder werden in der Cloud gespeichert, bereit für

spätere Verarbeitung. Die neuen eintreffenden Bilder werden als „unlabeled“ gekennzeichnet und anschließend im Active Learning Loop verwendet, um die zu annotierende Bilder auszuwählen. Verschiedene Active Learning-Ansätze, wie Entropy, Margin und MC-Dropout, können je nach Bedarf im Learning-Loop vordefiniert werden. Unsere Tests haben gezeigt, dass Entropie besonders effektiv ist im Vergleich zu anderen Methoden. Nach der Auswahl der zu annotierenden Bilder für das Labeling werden diese in der Zauberzeug Learning Loop Frontend-Weboberfläche an die Labeler angezeigt. Die Bilder werden in einer sortierten Reihenfolge präsentiert, wobei die informativsten Bilder zuerst erscheinen.

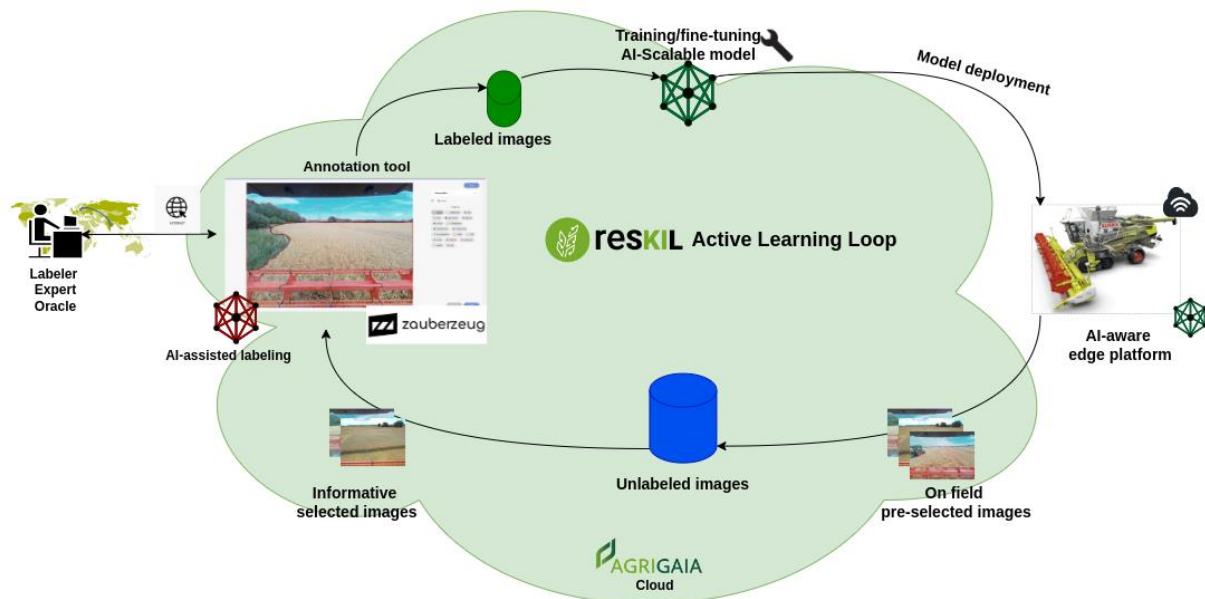


Abbildung 2: resKIL Active Learning Loop Framework

Diese informativen Bilder sind solche, die das Modell als unbekannt klassifiziert oder Bilder, bei denen das Modell Schwierigkeiten hat, Klassen für die Pixel in den Bildern zuzuordnen. Ein Administrator oder Stakeholder kann die Reihenfolge anpassen, indem er für bestimmte Bilder mit Tags hinzufügt. Abbildung 3 zeigt die Ansicht der Zauberzeug Learning Loop mit bereits annotierten Bildern. Nach dem Labeling werden die annotierten Bilder in einem „Labeled-Pool“ gespeichert, um das Modell nachzutrainieren. Das Training kann ausschließlich auf Cloud-Instanzen oder Computern mit GPU-Unterstützung stattfinden. Im resKIL wurde das Training entweder auf dem Zauberzeug-Rechner oder auf den Agri-Gaia Cloud-Instanzen durchgeführt. Nach Abschluss des Trainings wird das aktualisierte Modell auf die Landmaschine im Feld bereitgestellt.

Während des Projekts wurde das Zauberzeug Learning Loop Tool fortlaufend verbessert, um das Labeling von Bildern auf Pixel-Ebene zu ermöglichen. Dies wurde erreicht durch die Implementierung zweier neuer Funktionen: des "Manual Labeling Modes" und des "Click-Interaktiven Labeling Modes". Im "Manual Labeling Mode" kann der Labeler Objekte oder Merkmale in den Bildern ohne KI-Unterstützung annotieren. Dabei gibt der Labeler Klicks um das Rand eines Objekts ein, um dieses zu kennzeichnen. Abbildung 4 veranschaulicht ein Beispiel für das manuelle Labeling. Die schwarzen gefüllten Kreise repräsentieren die Klicks des Labelers, die bereits eingegeben wurden.

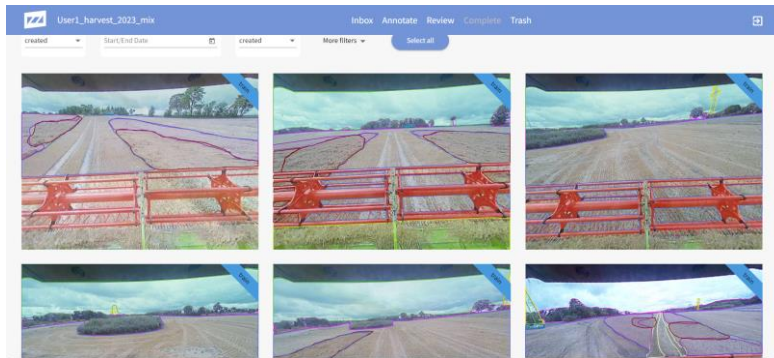


Abbildung 3: Bereits annotierten Bilder in "Zauberzeug Learning Loop" Webschnittstelle

Ein weiteres Feature, das im Zauberzeug Learning Loop eingeführt wurde, ist der "Click-Interaktive Labeling Mode". Für diesen Modus wurde während des Projekts der State-of-the-Art-Ansatz von ISEG implementiert und auf die landwirtschaftliche Branche zugeschnitten, da der ursprüngliche Ansatz ausschließlich auf Daten außerhalb der Landwirtschaft trainiert und getestet worden war. Der Click-Interaktive Labeling Mode nutzt das ISEG KI-Modell (Sofiiuk K. P., 2022), das die Klicks eines Labelers analysiert und versteht, um daraus eine Maske für die geklickte Region zu vorschlagen. Abbildung 5 zeigt ein Beispiel für das Labeling im Click-Interaktiven Modus. Das Bild zeigt das Originalbild mit einer Maske nach Eingabe von Klicks durch den Benutzer.



Abbildung 4: Annotiertes Lagergetreide im „manual labeling mode“ (ohne KI und mit ca. 20 Klicks)

Die beiden oben erwähnten Features wurden in einem speziellen Tool umgesetzt und in das Zauberzeug Learning Loop integriert. Dieses Tool, bekannt als die "reskil-Annotation-Node", besteht aus einem KI-Modell sowie weiteren kleineren Modulen, die zur Funktionsweise des Tools beitragen. Die reskil-Annotation-Node wurde in einen Docker-Container integriert, der von unserem Partner Zauberzeug verwendet wird. Diese Integration ermöglicht eine flexible Bereitstellung des Tools in verschiedenen Umgebungen, unabhängig von der zugrunde liegenden Hardware, und eine automatische Verknüpfung mit der Zauberzeug Learning Loop

Weboberfläche. Dadurch steht das Tool den Benutzern unabhängig von ihrem Standort während des Annotationsprozesses zur Verfügung.



Abbildung 5: Annotiertes Lagergetreide im „Click-Interactiven Labeling Mode“ (mit KI und ca. 8 Klicks)

Im Sommer 2022 wurde die erste Version des Tools während der Ernteperiode getestet. Das primäre Ziel dieses Tests war die Validierung des Tools in Bezug auf die im Projektplan festgelegten Anforderungen. Zwei Fern-Annotatoren von CLAAS Indien wurden eingesetzt, um die aufgenommenen Erntebilder zu annotieren. Sie haben das Tool genutzt, um eine Vielzahl von Bildern zu annotieren. Die Rückmeldungen und die Ergebnisse der Labeler waren im grundsätzlich positiv. Die reskil-Annotation-Node zeigte während ihrer ersten Anwendung die Fähigkeit, komplexe Objekte wie Bäume, Personen und Mähdrescher mit durchschnittlich drei Klicks zu annotieren. Trotz des insgesamt positiven Feedbacks gab es jedoch Situationen, in denen das Tool nicht optimal funktioniert, etwa bei der Annotation von Bildobjekten wie "Lagergetreide". In solchen Fällen benötigte das KI-Modell innerhalb der reskil-Annotation-Node mehrere Klicks, um die gesamte Fläche des Lagergetreides zu annotieren. Zudem trübte die Stabilität des Tools seine Anwendbarkeit, da es sporadisch abstürzte. Basierend auf diesen Erfahrungen wurde nach der Ernte eine umfangreiche Überarbeitung des Tools durchgeführt, um es zu verbessern. Dies umfasste die Behebung von Fehlern, die Sicherung der Stabilität des Tools und die Implementierung von Anmerkungen sowie Vorschlägen der Labelers. Im Sommer 2023 wurde die reskil-Annotation-Node weiterentwickelt, um die Bildannotation effizienter zu gestalten und die Integration in den Learning Loop von Zauberzeug zu optimieren. Eine wichtige Verbesserung war die Fähigkeit des Tools, während der Annotation Statistiken über die Labeler zu erfassen. Diese Funktion soll bei der Bewertung der in AP2.2 eingeführten effizienten Annotationsmethode in AP6 hilfreich sein. Um eine fundierte Bewertung zu ermöglichen, wurde die reskil-Annotation-Node erweitert, um die Anzahl der Klicks, die Zeit pro Bild und die Zeit pro Kategorie während der Annotation zu messen.

AP3: Skalierbare KI-Architektur

Arbeitspaket 3 zielt darauf ab, eine variable und skalierbare KI-Architektur zu entwickeln, die sich flexibel an moderne KI-Workflows in der Landwirtschaft anpassen kann und alle verfügbaren Ressourcen effizient nutzt. Dies beinhaltet die Implementierung einer Abstraktionsschicht und die Gestaltung der skalierbaren Architektur. Die Untersuchung wissenschaftlicher Literatur und industrieller Quellen hat gezeigt, dass etablierte Frameworks, die eine Abstraktionsschicht bieten, für die Ausführung von KI-Modellen unabhängig von deren Architektur auf verschiedenen Plattformen (Cross-Platform Execution) sehr vielversprechend sind. Drei der bekanntesten und häufig verwendeten Frameworks in der KI-Community sind TensorFlow XLA, OpenVINO und ONNX Runtime. Für das resKIL-Projekt wurde ONNX Runtime und ONNX als Ersatz für die Abstraktionsschicht ausgewählt, da diese in vielen KI-Frameworks wie TensorFlow und PyTorch gut integriert sind, einfach zu verwenden sind und alle notwendigen Funktionen bieten. Dies bedeutet, dass keine projektspezifische Entwicklung der Abstraktionsschicht erforderlich war. Dies trägt auch zur Stabilität und Übertragbarkeit der Software über das Projekt hinaus bei. Um diese Bibliotheken im Rahmen des resKIL-Projekts integrieren zu können, wurden einige Module um die PyTorch- und ONNX-Abstraktionsschicht herum entwickelt und implementiert. Dafür wurde die geplante Kapazität für AP3.3. genutzt. Zum Beispiel wurde ein Modul um ONNX und PyTorch herum implementiert, das es ermöglicht, das skalierbare Modell sowohl auf kleinen Edge-Geräten wie dem Jetson Xavier als auch auf leistungsfähigen Servern bereitzustellen. Abbildung 6 veranschaulicht die Schichtenarchitektur der resKIL skalierbaren KI-Architektur mit der ONNX-Schicht.

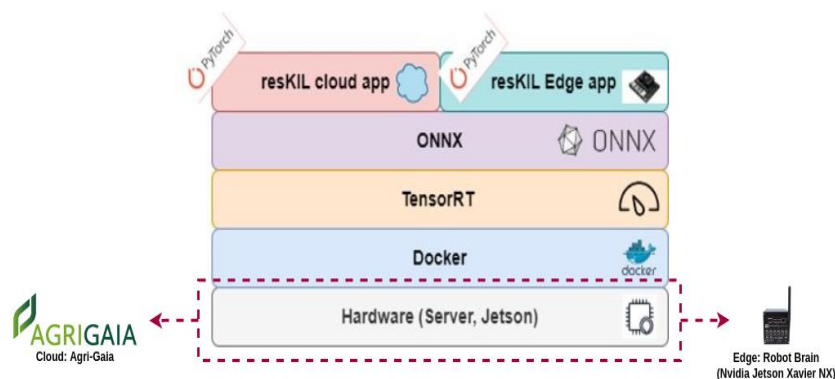


Abbildung 6: Schichtenarchitektur der resKIL skalierbaren KI-Architektur

Die Literaturuntersuchung zur Entwicklung einer skalierbaren Architektur hat gezeigt, dass Modelle für Real-Time Semantic Segmentation gut zu den Anwendungsfällen des resKIL-Projekts passen würden. Diese Modelle sollen nicht nur aufgrund ihrer kompakten Struktur höhere frames per second (FPS) erreichen, sondern auch eine hohe Leistung (Intersection over Union, IoU) liefern, unabhängig davon, ob sie auf ressourcenarme (Edge) oder ressourcenintensive (Cloud) Instanzen ausgeführt werden. Die Suche nach der geeigneten Architektur konzentrierte sich auf drei Modelle: Bisenet, DDRNET-slim und FASSDNET. Alle drei Modelle zeigten beeindruckende Ergebnisse beim Cityscapes Dataset, mit mIoUs von 73.4%, 77.8% und 78.2%. Zusätzliche Experimente wurden durchgeführt, um die FPS auf der Zielplattform zu berechnen, speziell auf der Nvidia Jetson Xavier Hardwareplattform, die für resKIL ausgewählt wurde. Die Ergebnisse dieser Experimente sind in Tabelle 3 dokumentiert, die die erreich-

ten FPS für die drei Modelle auf der ausgewählten KI-Hardwareplattform darstellt. Offensichtlich bietet das FASSDNET-Modell die besten Ergebnisse unter den getesteten Architekturen, mit einer FPS von 19. Aufgrund dieser herausragenden Leistung wurde FASSDNET als die skalierbare Architektur für das resKIL-Projekt ausgewählt. Trotz seiner kompakten Struktur bietet FASSDNET eine schnelle Ausführungsgeschwindigkeit auf Edge-Geräten und liefert gleichzeitig eine hohe Leistung.

Device	Architecture	Framework	Image size	FPS
Nvidia Jetson Xavier	DDRNET-Slim	Pytorch	1024x512	14
	Bisenet_V2	Pytorch	1024x512	10,6
	Fassdnet	Pytorch	1024x512	19,2

Tabelle 3: Ergebnisse des Experiments zur Messung der FPS auf dem Gerät NVIDIA Jetson NX

Für das Nachtrainieren des skalierbaren KI-Modells auf den annotierten Bildern aus Arbeitspaket 2 (AP2) wurde ein spezielles Trainingstool entwickelt, bekannt als die "reskil-Trainer-Node". Diese Node wurde so konzipiert, dass sie jedes KI-Modell unterstützt, das den festgelegten Spezifikationen oder Signaturen des resKIL-Projekts entspricht. Ähnlich wie bei früheren Entwicklungen wurde die reskil-Trainer-Node in einen Docker-Container integriert, was ihre Bereitstellbarkeit in jeder Umgebung erleichtert, in der eine Docker-Container-Runtime installiert ist und ausreichend Ressourcen (wie GPUs) für das Training zur Verfügung stehen. Durch diese Integration ist die reskil-Trainer-Node über die Learning-Loop-Web-Oberfläche unseres Partners Zauberzeug erreichbar. Dies ermöglicht es dem Nutzer des Learning Loops, das Training komfortabel zu starten, sobald er mit der Bildannotation fertig ist (siehe Abbildung 7).

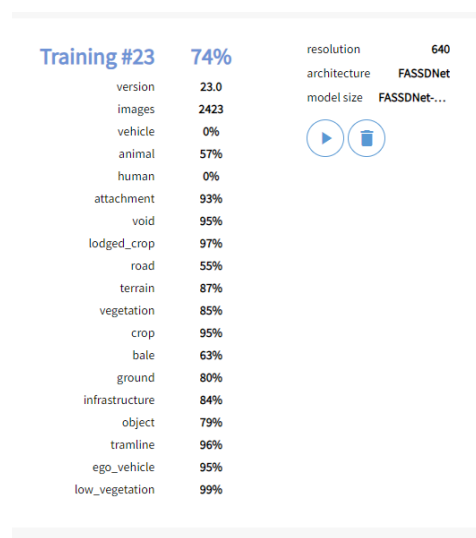


Abbildung 7: Ansicht der resKIL Trainingsoberfläche auf Zauberzeug Learning Loop

AP 5: Integration und Funktionsmusterentwicklung

In Arbeitspaket 5 soll die entwickelte KI-Architektur in der Praxis erprobt und auf der KI-Plattform sowie die Cloud bereitgestellt. Für die Cloud Plattform wird im resKIL-Projekt die zur Verfügung gestellte Agri-Gaia Cloud Plattform und die leistungsstarken Rechner von Zauberzeug verwendet. Im Gegensatz dazu wurde für die Edge-Lösung das Nvidia Jetson Xavier NX Development Kit ausgewählt. Hierbei wurde die Development Kit Board von Nvidia nicht direkt verwendet, da diese nicht in einem soliden Gehäuse für den Einsatz im Feld abgedeckt ist und viele Schnittstellen zum Anschluss an Landmaschinen, wie CAN-BUS oder eine externe Antenne, fehlen.

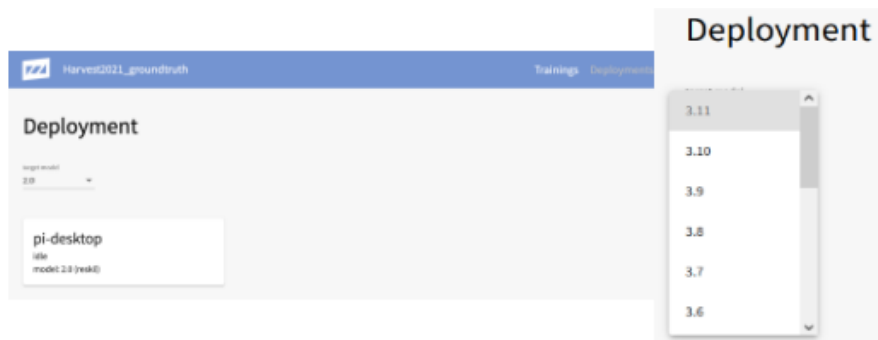


Abbildung 8: ResKIL Model Deployment Interface auf Zauberzeug Learning Loop

Deshalb wurde eine verbesserte Variante des Jetson Xavier NX von Zauberzeug, sogenanntes „Robot Brain“, verwendet. Dies hat den Original-Jetson Xavier NX-Board mit zusätzlichen Schnittstellen, die die Bereitstellung und Integration auf verschiedene Geräte, wie Landmaschinen, vereinfachen. Die skalierbare Architektur wurde auf dem Jetson-Board über das Zauberzeug Learning Loop bereitgestellt (siehe Abbildung 8), der auf der Landmaschine eingekoppelt und montiert wurde (siehe Abbildung 9).

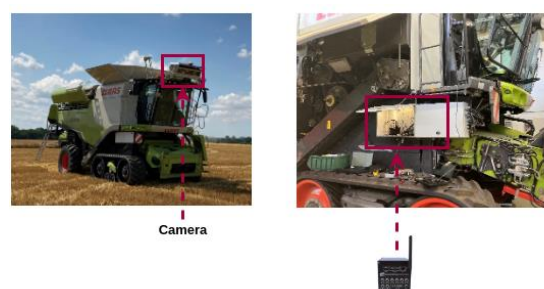


Abbildung 9: Montage von Kamera und Edge-Device an der Landmaschine

Die skalierte KI-Architektur kann problemlos auf verschiedene Cloud-Plattformen mit hoher Rechnerleistung bereitgestellt werden. Im Projekt resKIL wird beispielsweise die Agri-Gaia-Plattform für sein 32GB GPU für das Training genutzt (siehe Abbildung 10). Darüber hinaus wurden auch herkömmliche PCs oder Server als Rechenressourcen für das Training und die Bereitstellung der skalierbaren Architektur sowie der reskil-Annotation-Node verwendet. Dies

gilt insbesondere für die Bereitstellung der reskil-trainer-node, reskil-inference-node und reskil-annotation-node auf den Rechnern von Zauberzeug.

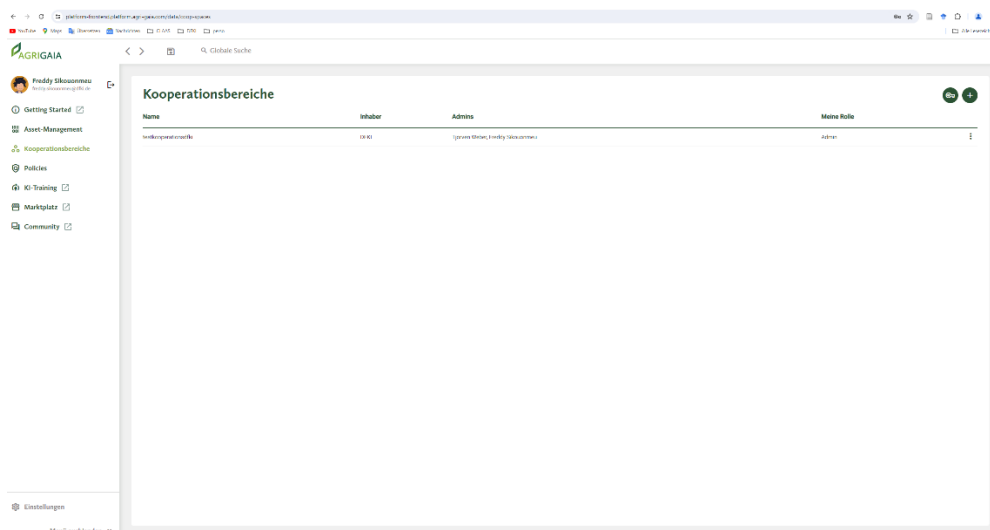
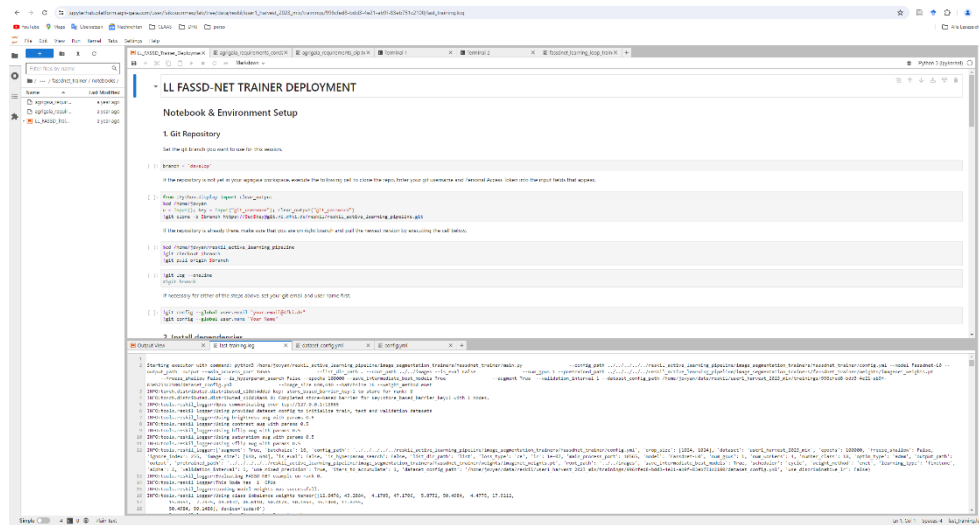


Abbildung 10: Agri-Gaia Plattform

AP6: Evaluation

Während der verschiedenen Erntekampagne wurde das Gesamtsystem resKIL evaluiert und verbessert. Dabei wurden alle Komponenten getestet und Metriken zur Bewertung des Systems erfasst. Für die Evaluierungsphase wurde in AP2 höchsten 2 Remote-Annotatoren zur Verfügung gestellt. Diese sollten eine Reihe von Bildern aus dem Feld annotieren. Um den im resKIL vorgeschlagenen aktiven Lernprozess besser evaluieren zu können, wurden zwei Labeling-Experimente geplant. In einem soll der Labeler bzw. Annotator die Bilder nur mit manuellen Methoden annotieren, d.h. es werden hier keine KI-Methoden eingesetzt.

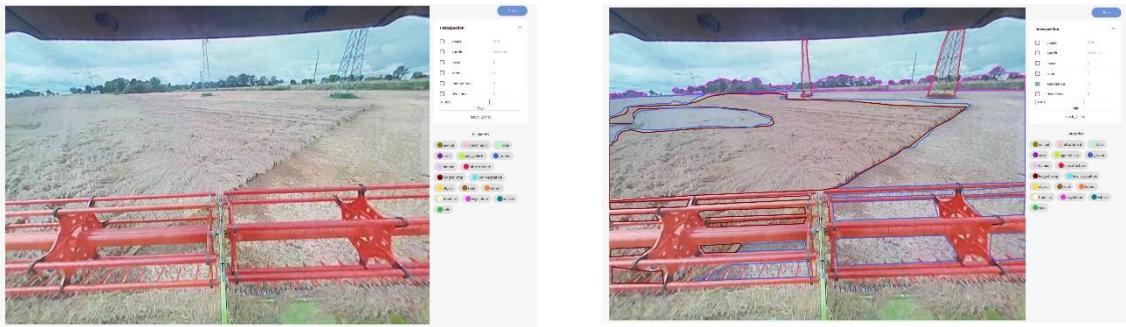


Abbildung 11: Anwendung der reskil-Annotation-Node in der Zauberzeug Learning Loop

Im zweiten Experiment soll der Labeler die gleichen Bilder in der gleichen Reihenfolge annotieren, aber diesmal mit dem KI-Modell im Hintergrund (mit der reskil-Annotation-Node), das den Labeler unterstützt. Die Bilder, die annotiert werden, sind Bilder, die von der reskil-Inferenz-Node auf dem Edge mit niedriger Konfidenz markiert wurden. Diese wurden in kleine Batches aufgeteilt und dem Labeler als Aufgabe zugewiesen. Während der Annotation wurden Metriken wie die Anzahl der Klicks und die Zeit pro Bild und Objektkategorie gemessen. Diese wurden nach der Erntekampagne detailliert ausgewertet.

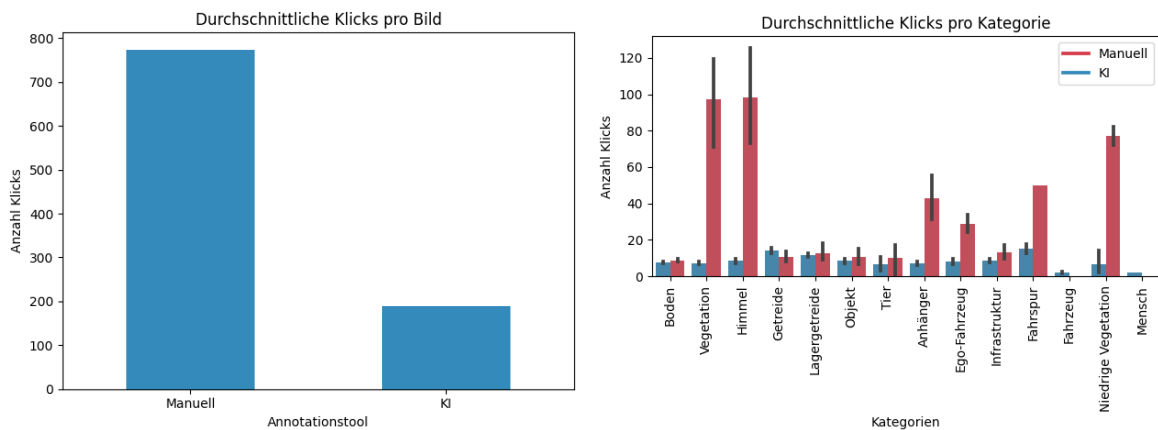


Abbildung 12: Durchschnittliche Klicks pro Bild und Kategorie

Die Bilder in Abbildung 12 zeigen die Anzahl der Klicks, die der Labeler pro Bild und pro Kategorie verwendet hat. Das Balkendiagramm zeigt deutlich, dass der Labeler mehr Klicks gebraucht hat, wenn er nicht unsere KI-Methode verwendet hat. Im Gegensatz dazu hat er im Durchschnitt 4-mal weniger Klicks gebraucht, wenn die reskil-Annotations-Node während der Annotation angewendet wurde. Ähnlich zeigt die rechte Abbildung einen ähnlichen Trend, wobei Kategorien wie Vegetation, Himmel schlechter abschneiden als mit der KI-gestützten Methode.

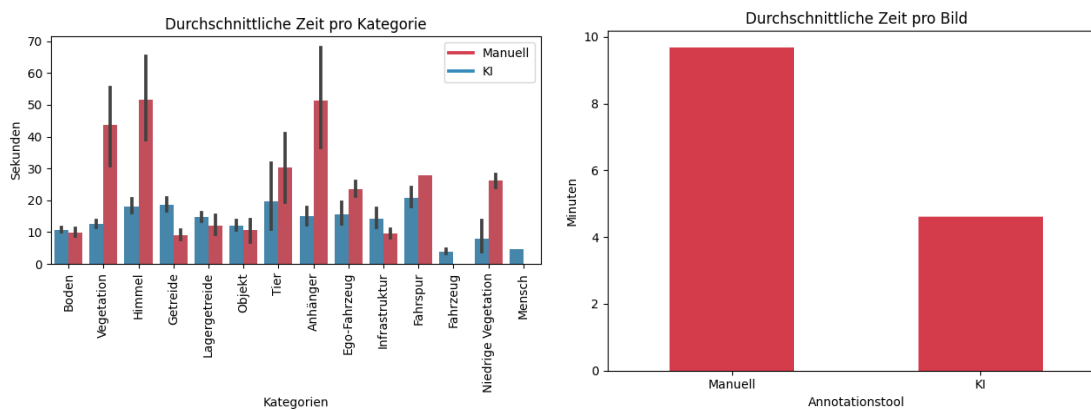


Abbildung 13: Durchschnittliche Zeit pro Bild und Kategorie

Es ist nicht verwunderlich, dass die Ergebnisse der Labeling-Experimente, die die Anzahl der Klicks auf die annotierten Bilder analysieren (siehe Abbildung 12), ähnlich den Daten in Abbildung 13 entsprechen. Im Allgemeinen deuten die Daten in Abbildung 13 darauf hin, dass der Labeler ohne KI-Unterstützung mehr Zeit benötigt, um Bilder zu annotieren, was sich auch in den Zeiten pro Kategorie widerspiegelt. Außerdem wurde während der Labeling-Experimente festgestellt, dass die Konsistenz der Annotation eine wichtige Rolle spielt. Genauer wurde herausgefunden, dass im Laufe des Labeling-Experiments vermutlich aufgrund von Müdigkeit, Ablenkung oder der Qualität der Originalbilder (wie Auflösung, Beleuchtung etc.) die Qualität der von den Labelern annotierten Bilder manchmal abgestiegen oder inkonsistent zwischen ähnlichen Bildregionen in unterschiedlichen Bildern war. Zum Beispiel wurden beim Labeling ähnliche Bildregionen (Lagergetreide und Getreide) in verschiedenen Bildern verwechselt annotiert. Die Lösung hier ist die Verwendung eines „Review“-Schrittes, bei dem alle annotierten Bilder von einem Experten geprüft werden und nach Feststellung von Fehlern an den Labeler zur Wiederholung zurückgeschickt werden.

Zur Bewertung der Skalierbarkeit der KI-Architektur wurde der reskil-Inference-Node auf verschiedenen Plattformen getestet (siehe Abbildung 14). Für die Evaluierung wurden ausschließlich Geräte von Robot-Brain Edge Device (Zauberzeug), Cloud (Agri-Gaia & Zauberzeug) und Desktop-Computern verwendet, die jeweils unterschiedliche Rechenleistungen aufweisen. Zum Beispiel verfügt das Edge-Gerät über begrenzte Rechenkapazität mit einer 8GB GPU für 64 Tensor Cores, im Vergleich zur Agri-Gaia Cloud mit einer 32GB GPU, die 640 Tensor Cores umfasst. Die Tensor Cores sind für die Berechnungen zuständig. Die Testergebnisse haben gezeigt, dass die KI-Architektur dank einer Abstraktionsschicht auf verschiedenen Plattformen bereitgestellt werden kann. Besonders bemerkenswert ist, dass die einzige Differenz zwischen den Plattformen in der Inferenzgeschwindigkeit (frames per second - FPS) besteht. Wie erwartet liefert die Cloud-Plattform die höchste Geschwindigkeit, gefolgt vom Desktop-Computer. Das Edge-Gerät zeigte die geringste Leistung.

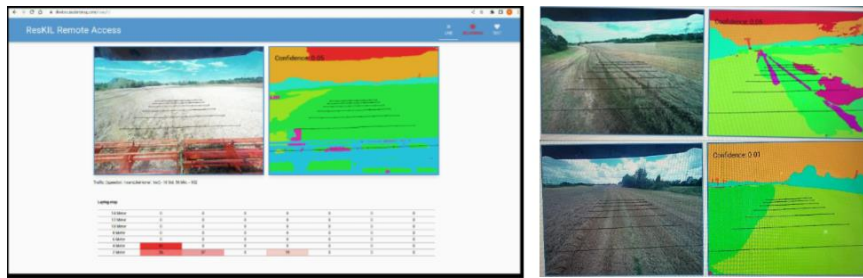


Abbildung 14: Visualisierung der reskil-Inference-Node zur Merkmalserkennung im Maschinenumfeld

2. des voraussichtlichen Nutzens, insbesondere der Verwertbarkeit des Ergebnisses und der Erfahrungen,

Das DFKI konnte in diesem Projekt umfangreiche Erfahrungen im Bereich der ressourceneffizienten KI und der menschenzentrierten KI sammeln. Dies ermöglichte eine praxisnahe Bewertung der technologischen Möglichkeiten und das Wissen darüber, wie diese in Folgeaktivitäten genutzt werden können. Die reskil-Annotation-Node mit seinem KI-Modell zur interaktiven Segmentierung kann leicht in viele andere Tools als backend für die KI-gestützte Annotation von landwirtschaftlichen Bildern integriert werden. Eine Anbindung an andere Labeling-Tools, die lediglich nur das Kommunikationsprotokoll entwickeln müssen, ist möglich. Die skalierbare Architektur von Fassednet ist kompakt und schnell genug, um auf vielen Edge- und Cloud-Plattformen eingesetzt zu werden. Die reskil-Inference-Node kann auch einfach mit anderen Tools verknüpft werden. Die Node wird als Docker Container bereitgestellt und erwartet Bilder für die Inferenz, die über eine REST API Schnittstelle gesendet werden. Das KI-Modell innerhalb der reskil-Inference-Node kann einfach lokal aktualisiert werden, ohne dass auf den Learning Loop zugegriffen werden muss. Abschließend fand ein Austausch mit einigen Mitarbeitern von CLAAS statt, die nicht am Projekt beteiligt waren. Dabei wurden die resKIL-Tools und die erzielten Ergebnisse vorgestellt. Dies soll als Ausgangspunkt für CLAAS-interne Projekte in ähnlicher Richtung wie resKIL dienen. Darüber hinaus ist noch in diesem Jahr ein Vortrag/Workshop zum Thema Data Labeling am Thünen-Institut, Zentrum für Informationsmanagement / KIDA" geplant, um die Ergebnisse des Projektes resKIL vorzustellen.

3. des während der Durchführung des Vorhabens dem Zuwendungsempfänger bekannt gewordenen Fortschritts auf diesem Gebiet bei anderen Stellen,

Während des Vorhabens wurden inhaltlich ähnliche Themen auch von anderen Stellen untersucht. Besonders hervorzuheben ist die Arbeit von Puck et al (Puck, 2021) mit dem Titel „Distributed active learning for semantic segmentation on walking robots“, in der ein ähnliches, aber etwas anderer Ansatz für Active Learning vorgeschlagen ist und die Arbeit von Lenczner et al. (Lenczner, 2022) „DIAL: Deep Interactive and Active Learning for Semantic Segmentation in Remote Sensing“, in der interaktives Learning mit Active Learning vereinigt ist. Die Erkenntnisse aus den veröffentlichten Untersuchungen ergänzen sich inhaltlich mit den Ergebnissen dieses Vorhabens. Zum Beispiel ist die Arbeit von Puck et al. kein Ansatz zur effizienten Annotation wie bei diesem Vorhaben anhand von der reskil-Annotation-Node, aber sie bieten jedoch wertvolle Einblicke in das Feld des Active Learning auf Edge device. Außerdem ist die Arbeit von Lenczner et al. nur für Task-spezifische Aufgaben geeignet und kann nicht einfach auf andere Domänen ohne ein volles Finetuning angewendet werden. Dennoch gibt der von den Autoren vorgeschlagene Ansatz einen guten Einblick in den Einsatz von interaktivem Learning kombiniert mit Active Learning für Task-spezifische Aufgaben.

4. der erfolgten oder geplanten Veröffentlichung des Ergebnisses

Die Ergebnisse des resKIL-Projekts wurden auf verschiedenen Veranstaltungen präsentiert. Im November 2023 stellte das DFKI das Projekt als Mitaussteller am Stand des Agrotech Valley e.V. auf der Agritechnica Messe vor. Während dieser Messe wurde das Vorhaben zahlreichen Besuchern präsentiert. Außerdem wurde während der Messe das resKIL-Projekt bei dem Besuch der parlamentarischen Staatssekretärin für Ernährung und Landwirtschaft, Frau Dr. Ophelia Nick, und ihrem Team vorgestellt. Im April 2024 wurde das resKIL-Projekt innerhalb der X-KIT (Gaia-X und KI-Projekte: Transfer & Vernetzung) als Teil eines GAIA-X-Vernetzungsprojekts und eines von BMEL geförderten Projekts vorgestellt. Diese Präsentation diente als Impuls für einen bevorstehenden Fachvortrag zum Thema „Labeling/Annotieren der Daten“, der im späten September stattfinden sollte. Bei diesem Fachvortrag werden die Ergebnisse dieses Vorhabens präsentiert und eine Einführung in effizientes Data Labeling mit KI gegeben. Als Abschluss wurde ein Papier zur vorgeschlagenen Annotationsmethode im Rahmen der GIL-2023-Jahrestagung eingereicht und angenommen. Dieses Papier untersucht, wie Agrardaten kosteneffizient annotiert werden können, indem das Active-Learning-Konzept mit interaktiven Bildsegmentierungsmethoden kombiniert wird. Die Experimente im Papier wurden auf öffentlichen Datensätzen durchgeführt, einschließlich Zuckerrüben und Blumenkohl. Der Titel des Papiers lautet: „Active-Learning-Driven Deep Interactive Segmentation for Cost-Effective Labeling of Crop-Weed Image Data“, verfasst von Sikouonmeu, Freddy und Atzmueller, Martin. Darüber hinaus wurde im Dezember eine Masterarbeit mit dem Titel „Density-Aware Active Domain Adaptation with Self-Training for Agricultural Image Segmentation“ von Frau Weber abgeschlossen. Sie wurde betreut von Prof. Dr. Joachim Hertzberg und Freddy Sikouonmeu vom DFKI. Das Ziel der Arbeit war es, ein KI-Modell während des Active Learning-Prozesses mit Bildern aus neuen Erntekampagnen mit wenigen Labels und Pseudo-Labels anzupassen.

5. Anhang

5.1 Literaturverzeichnis

- Bai, J. a. (2019). ONNX: Open Neural Network Exchange. *Github*. <https://github.com/onnx/onnx>: Github.
- Chen, L. C. (2017). Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE transactions on pattern analysis and machine intelligence*, (S. 40(4), 834-848). IEEE.
- FastAPI. (Juli 2024). *FastAPI*. Von FastAPI: <https://github.com/tiangolo/fastapi> abgerufen
- GitLab. (Juli 2024). *GitLab*. Von GitLab: <https://about.gitlab.com/> abgerufen
- Hong, Y. P. (2021). Deep dual-resolution networks for real-time and accurate semantic segmentation of road scenes. *arXiv preprint*. arXiv:2101.06085.
- Hoyer, C. R. (2020). Array programming with {NumPy}. *Natur*. 585. Natur: Springer Science and Business Media LLC.
- Kasarla, T. N. (2019, January). Region-based active learning for efficient labeling in semantic segmentation. *IEEE winter conference on applications of computer vision (WACV)*, (S. pp. 1109-1117). IEEE .
- Kim, K. P. (2021). Task-aware variational adversarial active learning. *Proceedings of the IEEE/CVF conference on computer vision and pattern*, (S. pp. 8166-8175). IEEE.
- Lenczner, G. C.-H.-T. (2022). DIAL: Deep interactive and active learning for semantic segmentation in remote sensing. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, (S. 15, 3376-3389.). IEEE.
- Mackowiak, R. L. (2018). Cereals-cost-effective region-based active learning for semantic segmentation. *arXiv preprint*. arXiv:1810.09726.
- Microsoft, C. (Juli 2024). *Visual Studio Code*. Von <https://code.visualstudio.com/>. abgerufen
- Puck, L. S. (Dezember 2021). Distributed active learning for semantic segmentation on walking robots. *In 2021 20th International Conference on Advanced Robotics*, S. pp. 461-467.
- Pytorch. (Juli 2024). *Pytorch*. Von <https://pytorch.org/> abgerufen
- Romera, E. A. (2017). Erfnet: Efficient residual factorized convnet for real-time semantic segmentation. *IEEE Transactions on Intelligent Transportation Systems*, (S. 19(1), 263-272.). IEEE.
- Rosas-Arias, L. B.-G.-P.-M.-P. (2021). FASSD-Net: Fast and accurate real-time semantic segmentation for embedded systems. . *IEEE Transactions on Intelligent Transportation Systems*, (S. 23(9), 14349-14360).
- Sener, O. &. (2017). Active learning for convolutional neural networks: A core-set approach. *arXiv preprint*. arXiv:1708.00489.
- Settles, B. (2009). Active learning literature survey.
- Siddiqui, Y. V. (2020). Viewal: Active learning with viewpoint entropy for semantic segmentation. *In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* , (S. pp. 9433-9443). IEEE.

- Sofiiuk, K. P. (2016). f-brs: Rethinking backpropagating refinement for interactive segmentation. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, (S. pp. 8623-8632). IEEE.
- Sofiiuk, K. P. (2022). Reviving iterative training with mask guidance for interactive segmentation. . *IEEE International Conference on Image Processing (ICIP)*, (S. pp. 3141-3145). IEEE.
- TensorFlow. (Juli 2024). *TensorFlow*. Von TensorFlow. abgerufen
- Wang, J. S. (2020). Deep high-resolution representation learning for visual recognition. *IEEE transactions on pattern analysis and machine intelligence*, (S. 43(10), 3349-3364.). IEEE.
- Xie, S. F. (2020). Deal: Difficulty-aware active learning for semantic segmentation. *Proceedings of the Asian conference on computer vision*.
- Xu, N. P. (2016). Deep interactive object selection. *Proceedings of the IEEE conference on computer vision and pattern recognition*, (S. pp. 373-381). IEEE.
- Yu, C. W. (2018). Bisenet: Bilateral segmentation network for real-time semantic segmentation . *Proceedings of the European conference on computer vision (ECCV)*, (S. pp. 325-341).
- Zauberzeug. (Juli 2024). *An intelligent web application to iteratively improve AI models*. Von Learning Loop: <https://zauberzeug.com/products/learning-loop> abgerufen

resKIL - Kurzer inhaltlicher Bericht DFKI

I – 1: Kurze Darstellung zur Aufgabenstellung

Das Projekt resKIL zielt darauf ab, die Ressourceneffizienz in der Landwirtschaft durch die Entwicklung einer flexiblen Softwarearchitektur und einer KI-Toolchain zu verbessern. Es konzentriert sich auf die Integration von maschinellem Lernen in Landmaschinen, um die Qualitätsbestimmung von Erntegut und die Erkennung von Merkmalen im Maschinenumfeld zu verbessern. Durch eine agile Entwicklungsstrategie und die Fokussierung auf drei Ernteperioden soll die Qualität und Robustheit der Lösungen kontinuierlich verbessert werden. Darüber hinaus werden Verfahren des maschinellen Lernens und des Deep Learning entwickelt bzw. angepasst, um landwirtschaftliche Arbeitsmaschinen zu verbessern und die Landwirtschaft nachhaltiger zu gestalten. Die Ergebnisse werden sowohl wissenschaftlich als auch praktisch evaluiert, um ihre Anwendbarkeit und Skalierbarkeit sicherzustellen.

I – 2: Kurze Darstellung zu den Voraussetzungen, unter denen das Vorhaben durchgeführt wurde

Die Integration von KI in landwirtschaftliche Maschinen steht vor erheblichen Herausforderungen, einschließlich begrenzter Ressourcen wie Speicher und Rechenleistung, die die Ausführung komplexer KI-Modelle erschweren. Widrige Umgebungsbedingungen wie Temperatur, Vibration und Staub sowie eine variable Konnektivität von keiner Verbindung bis hin zu 4G erschweren die Aktualisierung von Modellen und die Nutzung von Cloud-Diensten für Inferenzprozesse. Darüber hinaus ist die Datenerhebung aufgrund der saisonalen Natur der Landwirtschaft und der ständig wechselnden Umgebung eingeschränkt, was die Verbesserung von KI-Modellen weiter kompliziert.

I – 3: Kurze Darstellung zur Planung und Ablauf des Vorhabens

Das Vorhaben bestand aus insgesamt 6 Arbeitspaketen, in denen die Arbeit der verschiedenen Partner organisiert war. Das DFKI war an den Arbeiten in den Arbeitspaketen 2, 3, 5 und 6 beteiligt. Ein weiteres Arbeitspaket (AP 7) beschreibt die Koordination. Im Laufe des resKIL-Projekts fanden 3 Meilensteine nach jedem vollen Jahr statt. Dies entspricht gut den im Projekt 3 geplanten Ernten, die für die Datenaufnahme, Anwendung des entwickelten Konzepts sowie die Auswertung der Ergebnisse dienen. Darüber hinaus finden auch während des Projekts monatlich Core-Meetings statt. Dabei trafen alle beteiligten Partner in einem kleinen Kreis zusammen und diskutierten über den Stand des Projekts.

I – 4: Kurze Darstellung zum wissenschaftlichen und technischen Stand, an den angeknüpft wurde

Der wissenschaftliche und technische Stand, an den in den Arbeitspaketen AP2, AP3 und AP5 angeknüpft wurde, konzentriert sich auf fortschrittliche Methoden zur Datenauswahl und -verwendung, die Entwicklung skalierbarer KI-Architekturen und die Integration sowie Entwicklung von Funktionsmustern. Im Mittelpunkt steht die Optimierung der Datenqualität für maschinelle Lernmodelle, wobei Active Learning und intelligente Datenauswahlmethoden wie informative und repräsentative Selektion eine Schlüsselrolle spielen. Diese Ansätze berücksichtigen die Variabilität der Kosten für das Labeling verschiedener Datenpunkte, was besonders in der Landwirtschaft relevant ist. Zudem wird die "interaktive semantische Segmentierung" als innovative Methode zur Vereinfachung des Labeling-Prozesses hervorgehoben. Für die Implementierung einer skalierbaren KI-Architektur in der Praxis werden Echtzeit-Semantik-Bildsegmentierungsmodelle bevorzugt, die sowohl für Edge-Geräte als auch Cloud-Infrastrukturen geeignet sind. Aktuelle Machine-Learning-Frameworks wie PyTorch und TensorFlow sowie ONNX ermöglichen eine flexible und effiziente Entwicklung und Anpassung von KI-Modellen für verschiedene Plattformen.

I – 5: Kurze Darstellung zur Zusammenarbeit mit anderen Stellen

Es fand eine enge Zusammenarbeit zwischen den Projektpartnern statt.

II – 1: Eingehende Darstellung des erzielten Ergebnisses

Das reskil-Projekt hat ein ressourceneffizientes Annotationsverfahren entwickelt, das auf fortschrittlichen KI-Technologien für Active Learning und interaktive Segmentierung setzt, um die Bildannotation in der Landwirtschaft zu optimieren. Mit der Einführung der reskil-Annotation-Node, die in einem Docker-Container verfügbar ist, wurde ein Werkzeug geschaffen, das den Benutzeraufwand erheblich reduziert und nahtlos in den Zauberei-Learning-Loop integriert ist. Durch die Auswahl einer skalierbaren KI-Architektur, die auf der Integration in bestehende KI-Frameworks basiert (reskil-Training-Node und reskil-Inference-Node), wurde ein System entwickelt, das sowohl auf Edge-Geräten als auch in Cloud-Umgebungen effizient funktioniert.

II – 2: Eingehende Darstellung des voraussichtlichen Nutzens, insbesondere der Verwertbarkeit des Ergebnisses und der Erfahrungen

Das DFKI konnte in dem Vorhaben umfangreiche Erfahrungen in der ressourceneffizienten KI und Human-Centered AI sammeln, was eine praktische Einschätzung der Technologie und deren mögliche Anwendung in zukünftigen Aktivitäten ermöglichte. Das entwickelte resKIL-Annotation-Tool mit seinem interaktiven Segmentierung-KI-Modell bietet eine flexible Plattform für das Labeling von Agrarbildern und kann als Backend für andere Tools eingesetzt werden. Die skalierbare Architektur ist für die Bereitstellung auf verschiedenen Plattformen geeignet, und das resKIL-Inference-Node kann leicht mit anderen Tools integriert werden, was durch seine Bereitstellung als Docker-Container erleichtert wird.

II – 3: Eingehende Darstellung des während der Durchführung des Vorhabens dem Zuwendungsempfänger bekannt gewordenen Fortschritts auf diesem Gebiet bei anderen Stellen

Während des Vorhabens wurden ähnliche Themen von anderen Forschern untersucht, wie die Studien von Puck et al. und Lenczner et al. zeigen. Puck et al. fokussierten sich auf Distributed Active Learning für die Semantische Segmentierung auf Gehen-Robotern, während Lenczner et al. DIAL präsentierten, eine Kombination aus interaktivem und Active Learning für die Semantische Segmentierung in Fernerkundungsaufgaben. Obwohl diese Studien unterschiedliche Ansätze verfolgen, ergänzen sie die Ergebnisse dieses Vorhabens, indem sie wertvolle Einblicke in das Feld des Active Learning und interaktiven Lernens bieten.

II – 4: Eingehende Darstellung der erfolgten oder geplanten Veröffentlichung des Ergebnisses

Das Projekt resKIL wurde auf mehreren Veranstaltungen präsentiert, unter anderem auf der Agrotechnica im November 2023 und im Rahmen des X-KIT im April 2024, wo es als Teil eines GAIA-X-Verbundprojektes und eines vom BMEL geförderten Projektes vorgestellt wurde. Diese Präsentationen dienten als Impuls für einen Fachvortrag im September 2024 auf der KIDA-Fachtagung, der sich auf das Thema Data Labeling und Einführung in die effizientes Data Labeling mit KI konzentrierte. Ein Papier zur vorgeschlagenen Annotationsmethode wurde im Rahmen der GIL-2023-Jahrestagung eingereicht und diskutierte die Kombination von Active-Learning-Konzepten mit interaktiven Bildsegmentierungsmethoden für kosteneffizientes Labeling von Agrardaten. Darüber hinaus wurde eine Masterarbeit mit dem Titel "Density-Aware Active Domain Adaptation with Self-Training for Agricultural Image Segmentation" abgeschlossen, die ein KI-Modell während des Active-Learning-Prozesses mit wenigen annotierten und vielen unannotierten Bildern aus früheren Erntekampagnen anpasste.