

HSBI

Hochschule
Bielefeld
University of
Applied Sciences



iDaS

Institute for Data Science
Solutions of HSBI

Kongress KI205@HSBI Zukunft im Fokus -Posterbeiträge-

Schriftenreihe des Institute for Data Science Solutions

Nr. 2/ 2025

DOI: 10.60802/sidas.2025.2

ISSN: 2943-3509



Dieses Dokument ist lizenziert gemäß CC BY <https://creativecommons.org/licenses/by/4.0/> Ausgenommen von der Lizenz sind alle Wort-/Bildmarken und Logos

Die Autor*innen tragen die Verantwortung für die Einhaltung der urheberrechtlichen Bestimmungen. Zum Zeitpunkt der Drucklegung führten die Verweise auf Internetseiten zu den gewünschten Inhalten. Sollten zu einem späteren Zeitpunkt die Internetseiten verändert worden sein, distanzieren sich die Autor*innen von den inhaltlichen Aussagen der Internetseiten.

Poster auf dem Kongress KI@HSBI2025 an der Hochschule Bielefeld

Am 03. Juni 2025 fand an der Hochschule Bielefeld der Kongress KI@HSBI 2025 statt. Unter dem Motto „Zukunft im Fokus“ tauschten sich rund 340 Teilnehmende, darunter Vertreterinnen und Vertreter von Unternehmen als auch Studierende, Lehrende und Promovierende der HSBI, über aktuelle Entwicklungen sowie über Einsatzmöglichkeiten und Anwendungsbereiche von Künstlicher Intelligenz aus. Neben vielen verschiedenen Formaten wie Vorträgen, Workshops und einer KI-Expo, zeigten Studierende und Promovierende auf insgesamt 26 Postern ihre Projekte, Vorhaben oder Ergebnisse. Organisiert wurde der Kongress vom Institute for Data Science Solutions (IDaS) der HSBI, das sich in seiner Arbeit inhaltlich auf die Themen Data Science, KI, sowie Optimierung und Simulation fokussiert. Dabei stehen die Daten und deren Anwendung für den Einsatz von Künstlicher Intelligenz im Mittelpunkt. Die Anwendungsdomänen ergeben sich aus aktuellen gesellschaftlichen Herausforderungen. So entwickelt das IDaS gemeinsam mit Partnern aus der Wirtschaft und anderen Forschungseinrichtungen Lösungen mittels KI zum Beispiel zu den Themen Nachhaltigkeit, Klimaschutz, Sicherheit, Produktivität, u.v.m. Die große Anzahl an Teilnehmenden an dem Kongress verdeutlichte die Aktualität des Themas und den starken Bedarf an KI-gesteuerten Prozessen. Allerdings wurde auch aufgezeigt, wie viele Hürden und Herausforderungen noch bestehen. Das IDaS möchte dazu einen Beitrag leisten und mit den hier publizierten Erkenntnissen aus der Postersession einem breiteren Publikum die Möglichkeit geben, sich über aktuelle KI-Projekte zu informieren.

Inhaltsverzeichnis

AQ-Shuttle – Eine Studie über autonome Quartiers-Shuttle mit Mobilstation in Bielefeld	2
Digitaler Bahnhof Minden – Künstliche Intelligenz für den Bahnhof der Zukunft	4
SHARLY - A Smart Assistant for Intelligent Environments.....	6
Für eine inklusive digitale Zukunft: Ableismus-sensible Chatbots gestalten.....	8
Bildung gestalten in der KI-ndheit – Digitale Teilhabe in der Kita.....	10
Statistical Learning of Daily Routines in the Smart Home for Privacy Preserving Ambient Assisted Living	12
AutoDSE: Towards HW/AI co-design of ultra-low latency hardware accelerators for industrial applications	14
Edge4OD: An Automated Framework for Efficient Execution of Object Detection Algorithms on Edge Devices	16
Green by Design Nachhaltiger Tiefbau durch KI.....	18
Erforschung von Gefahren durch Adversarial Attacks beim Einsatz des LLM-getriebenen Roboters Furhat in der Lehre.....	20
AI-Powered Interactive Learning and Educational Analytics With Speech and Visual Integration	22
A Bio-Inspired Hybrid Neural Network Approach to Object Detection Using CNN-ViT and Spiking Networks.....	24
Informed Active Learning with Decision Trees to Balance Exploration and Exploitation	26
Bridging the Dataset Gap: Domain Adaptation for Fisheye Passenger Detection	28
Clustering-basierte Entwicklung von Haushaltslastprofilen nach sozioökonomischen Kategorien.....	30
Automatisierte Erkennung von Störstoffen in Bioabfall mit maschinellem Lernen – Ansätze und Ergebnisse aus dem Projekt TRACES.....	32
Facial Emotion Recognition.....	34

AQ-Shuttle – Eine Studie über autonome Quartiers-Shuttle mit Mobilstation in Bielefeld

Kaan Altun, Benjamin Eris, Marc Halbrügge, Thorsten Jungeblut, Rolf Naumann, Mats Jendrik Ottemeier, Riza Öztürk, Armin Pascanovic

Das Forschungsprojekt AQ-Shuttle ist ein interdisziplinäres Forschungsprojekt der Hochschule Bielefeld, der Stadt Bielefeld und der moBiel GmbH. Das Ziel ist die Erprobung eines autonomen Lieferroboters für die letzte Meile zwischen Mobilstationen und Wohnquartieren. Die Idee ist, den innerstädtischen Verkehr nachhaltiger zu gestalten, den motorisierten Individualverkehr zu reduzieren und damit die Lebensqualität im urbanen Raum zu steigern.

1. Projektkontext und Motivation

Bielefeld steht vor großen Herausforderungen im Bereich urbaner Mobilität. Laut einer Erhebung von 2017 liegt der Anteil des motorisierten Individualverkehrs bei 51 % (Stadt Bielefeld, 2022). Die Folgen sind bekannt: Staus, Luftverschmutzung und eine ineffiziente Nutzung städtischer Flächen. Die Mobilitätsstrategie 2030 der Stadt Bielefeld strebt eine Reduktion dieses Anteils auf 25 % an (Stadt Bielefeld, 2018). Hier setzt das AQ-Shuttle an: Mit autonomen Transportlösungen soll eine ökologisch und sozial nachhaltige Alternative geschaffen werden.

2. Technische Merkmale des AQ-Shuttles

Das Fahrzeug ist ein elektrisch betriebenes, autonom fahrendes Shuttle mit einer Zuladung von 14 kg. Die Akkukapazität beträgt 13 Stunden, die Selbstladung erfolgt über eine Dockingstation. Die Geschwindigkeit beträgt maximal 6 km/h.

3. Interdisziplinärer Forschungsansatz

Das Projekt vereint mehrere Disziplinen: Wirtschaftspsychologie (Erforschung der Wahrnehmung und Akzeptanz), Ingenieurwissenschaften (Analyse der infrastrukturellen und technischen Rahmenbedingungen), Recht (Untersuchung der rechtlichen Voraussetzungen), Wirtschaftswissenschaften (Kosten-Nutzen-Kalkulationen, Entwicklung von Geschäftsmodellen) und Logistik (Tourenplanung, Integration von Mikrodepots und Sharing-Stationen).

4. Fazit

Das AQ-Shuttle-Projekt verdeutlicht das Potenzial autonomer Mobilitätslösungen. Durch die Kombination technischer Innovation, interdisziplinärer Forschung und partizipativer Ansätze kann ein wertvoller Beitrag zur Verkehrswende geleistet werden. Die Integration autonomer Shuttles in bestehende Mobilitätsstrukturen eröffnet Perspektiven für eine lebenswertere, nachhaltigere Stadt. Nicht als Ersatz, sondern als sinnvolle Ergänzung multimodaler Verkehrssysteme.

Danksagung

Dieses Projekt wurde durch das Ministerium für Umwelt, Naturschutz und Verkehr des Landes NRW gefördert und zudem teilweise durch das Bundesministerium für Bildung und Forschung im Rahmen des Projekts Career@BI im Förderprogramm „FH Personal“ (03FHP106) unterstützt. Wir danken herzlich für die gewährte Unterstützung.

Literatur

Stadt Bielefeld. Mobilitätsstrategie 2030, 2018. URL <https://www.bielefeld.de/mobilitaetsstrategie>. Abruf 15.06.2025.

Stadt Bielefeld. Die Haushaltsbefragung zur Mobilität, 2022. URL <https://www.bielefeld.de/modal-split-befragung>. Abruf 15.06.2025.

Wir bieliefen.

AQ-Shuttle- Eine Studie über autonome Quartiers-Shuttle mit Mobilstation in Bielefeld

Projektüberblick

Das Forschungsprojekt AQ-Shuttle ist eine gemeinsame Initiative der Hochschule Bielefeld, der Stadt Bielefeld und moBiel. Es zielt darauf ab, autonome Quartiersshuttles für die Zustellung von Gütern auf der sogenannten 'letzten Meile' zu erproben. Dabei sollen zentrale Mobilitätsknotenpunkte (Mobilstationen) mit Wohnquartieren verbunden werden. Im Fokus steht die Verbesserung der urbanen Lebensqualität durch Reduktion des motorisierten Individualverkehrs.

Interdisziplinärer Forschungsansatz

Das Projekt ist breit interdisziplinär aufgestellt und berücksichtigt mehrere Perspektiven:

- Die wirtschaftspsychologische Begleitforschung untersucht die Einstellungen, Erwartungen und Bedenken der Bevölkerung gegenüber neuen Mobilitätskonzepten und Technologien. Ziel ist es, die Akzeptanz und mögliche Nutzungshürden besser zu verstehen
- Die ingenieurwissenschaftliche Begleitforschung analysiert infrastrukturelle Rahmenbedingungen und technische Anforderungen an Fahrzeuge und Strecken. Dazu zählen Untergründe, Lichtverhältnisse, Hindernisse und Sicherheitsaspekte (z. B. Antriebstechnik, Mechanik, Bordsteinbewältigung)
- Die juristische Forschung begleitet das Genehmigungsverfahren und analysiert die rechtlichen Rahmenbedingungen für den Einsatz autonomer Lieferfahrzeuge im öffentlichen Raum
- Die wirtschaftliche Begleitforschung erstellt betriebswirtschaftliche Kalkulationen, entwickelt Geschäftsmodelle und untersucht die wirtschaftliche Tragfähigkeit des Systems
- Die logistische Forschung befasst sich mit Tourenplanung, der Integration von Mikrodepots und Sharing-Stationen, Standortanalysen und statistischer Erhebung logistischer Kennzahlen

Städtischer Hintergrund

In Bielefeld lag der Anteil des motorisierten Individualverkehrs laut Modal-Split-Erhebung 2017 bei 51 %. Dies führt besonders in dicht besiedelten Quartieren zu erheblichen Verkehrsproblemen, Luftverschmutzung und Flächenverbrauch. Die Mobilitätsstrategie 2030 der Stadt Bielefeld sieht eine Reduktion dieses Anteils auf 25 % vor. Projekte wie das AQ-Shuttle sollen zur Erreichung dieses Ziels beitragen, indem sie attraktive, umweltfreundliche Alternativen für den Warentransport im Quartier bieten.

Ziele des Projekts

Das AQ-Shuttle-Projekt verfolgt mehrere zentrale Ziele:

- Etablierung eines autonomen Transportsystems für Güter auf der letzten Meile
- Entlastung städtischer Infrastrukturen und Umwandlung von Parkplatztflächen in Grün- und Aufenthaltsflächen
- Förderung multimodaler Mobilität durch Anbindung an den ÖPNV und Sharing-Angebote
- Förderung der gesellschaftlichen Akzeptanz autonomer Technologien
- Entwicklung nachhaltiger Geschäftsmodelle für den dauerhaften Betrieb

Qualitative Ergebnisse aus Fokusgruppen

In Fokusgruppenerhebungen mit insgesamt 18 Teilnehmer:innen wurden zentrale Einflussfaktoren auf das Mobilitätsverhalten identifiziert: Flexibilität, Witterungsbedingungen, Kosten, Umweltbewusstsein und Infrastrukturqualität. Die Stadtbahn wird wegen Umweltfreundlichkeit und niedriger Kosten geschätzt, aber häufig wegen Überfüllung, Unpünktlichkeit und schlechter Abendeverbindungen kritisiert. Das Fahrrad ist beliebt, aber durch unzureichende Wege und Wetter eingeschränkt. Der Pkw bietet Komfort, verursacht jedoch Parkplatzprobleme und widerspricht ökologischen Zielen.

Die vorgestellten autonomen Shuttles wurden als sympathisch, innovativ und alltagstauglich beschrieben. Es zeigte sich jedoch auch ein breites Spektrum an Bedenken: Datenschutz, Sicherheit des Laderaums, mögliche Störungen durch Tiere, Vandalismus, soziale Auswirkungen wie Arbeitsplatzverdrängung sowie Sorgen über die Zuverlässigkeit der Technik im öffentlichen Raum.

Einsatzszenarien

Während der Fokusgruppen entwickelten die Teilnehmer:innen konkrete Ideen für Einsatzmöglichkeiten:

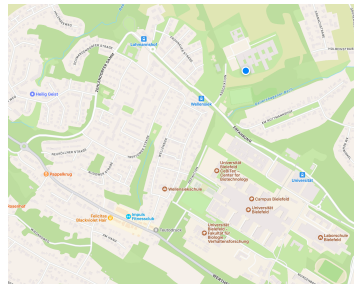
- Alltagshilfe: Lieferung von Briefen, Medikamenten oder kleinen Paketen
- Unterstützung für mobilitätseingeschränkte Menschen: Nutzung für Senior:innen oder Menschen mit Behinderung
- Einkaufstransport: Shuttle übernimmt Weg vom Supermarkt bis zur Haustür

Ausblick

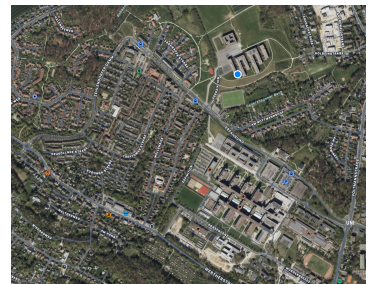
Im nächsten Schritt wird eine Online-Umfrage mit einer größeren Stichprobe durchgeführt. Die Ergebnisse aus den Fokusgruppen fließen in die Fragebogenentwicklung ein. Die Umfrage ist für Mai 2025 geplant, die Auswertung soll bis Juni erfolgen. Anschließend beginnt eine Pilotstudie, in der AQ-Shuttles unter realen Bedingungen im öffentlichen Raum eingesetzt und getestet werden. Dabei werden technische Funktionalitäten, Nutzerakzeptanz, Nutzungsrouinen und Verbesserungspotenziale systematisch erfasst.

Fazit

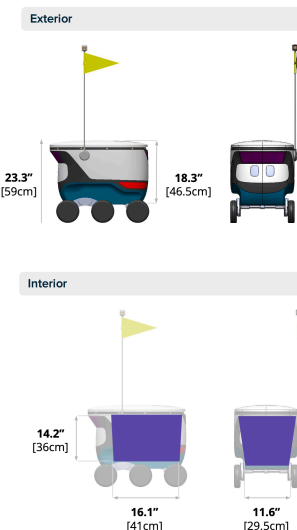
Das AQ-Shuttle-Projekt verbindet technologische Innovation mit einer nachhaltigen Vision für urbane Mobilität. Es zeigt auf, wie autonome Logistikkösungen zur Verkehrswende beitragen können – nutzerorientiert, stadtvträglich und zukunftsweisend.



Abbildungen: Bildschirmfotos aus Apple Karten (© Apple Inc.), Stand: Mai 2025.



Cartken Model C Specs



TUTOBOTS

- Robotergewicht (leer): 34 kg
- Maximale Zuladung: 14 kg
- Akkukapazität: über 1000 Wh
- Laufzeit: mehr als 13 Stunden mit einer einzigen Ladung
- Ladung: nutzt eine Selbstlade-Dockingstation
- Geschwindigkeit: 4,8 – 9,7 km/h
- Einsatzbereich: für Innen- und Außenbereiche geeignet
- Hindernisüberwindung: bis zu 20 cm
- Behälter: lebensmittelsicherer, isolierter Behälter

Digitaler Bahnhof Minden – Künstliche Intelligenz für den Bahnhof der Zukunft

Hasina Attaullah, Marc Halbrügge, Rolf Naumann & Thorsten Jungeblut

Hochschule Bielefeld, Fachbereich Ingenieurwissenschaften und Mathematik

Das Projekt Digitaler Bahnhof Minden (DiBaMi) verfolgt mehrere wichtige Ziele. Zum einen soll der Bahnhof Minden zu einem innovativen Showroom entwickelt werden, der als Ort der Kollaboration und außerschulischer Lernort dient. Ein Fokus liegt dabei auf dem Thema Mobilität der Zukunft, um in diesem Rahmen verschiedene Bildungs- und Dialogformate zu ermöglichen.

Autonomer Gepäckwagen

Im Projekt DiBaMi wird künstliche Intelligenz (KI) eingesetzt, um die Mobilität der Zukunft zu gestalten und gesellschaftliche Herausforderungen zu lösen. Ein Beispiel hierfür ist der autonome Gepäckroboter, der mit KI-gestützten Algorithmen ausgestattet ist. Diese Algorithmen ermöglichen dem Roboter, seine Umgebung zu erkennen, Hindernisse zu vermeiden und sicher seinen Weg zu einem bestimmten Zielpunkt zu finden. Durch den Einsatz von KI kann der Roboter selbstständig die optimale Route planen und während seiner Fahrt flexibel auf veränderte Situationen reagieren während er bspw. der Person folgt, die im Augenblick mit ihm interagiert (Beck & Bader, 2019).

Zusammenfassung und Ausblick

Die Entwicklung eines autonomen Gepäckwagens im Rahmen des Projekts DiBaMi stellt einen wichtigen Schritt zur Gestaltung einer inklusiveren Mobilität der Zukunft dar. Die Integration von künstlicher Intelligenz ermöglicht dem Roboter, seine Umgebung sicher zu navigieren, Gepäck sicher und effizient zu befördern, mit Nutzer:innen zu interagieren und ihren Bedürfnissen gerecht zu werden.

Durch die Fokussierung auf barrierefreie Designprinzipien kann der autonome Gepäckwagen sowohl Menschen mit eingeschränkter Mobilität als auch Senior:innen neue Möglichkeiten bieten und das Reiseerlebnis deutlich verbessern. Die Einbindung von Nutzer:innen aus verschiedenen Bevölkerungsgruppen in den Entwicklungsprozess sichert eine praxisnahe Umsetzung und trägt zur Schaffung einer zugänglicheren und inklusiveren Gesellschaft bei.

Danksagungen

Dieses Projekt ist gefördert durch das EFRE/JTF-Programm NRW, REGIONALE Ostwestfalen-Lippe – Vernetzte Mobilität und digitale Anwendungen. Des Weiteren ist dieses Projekt unterstützt durch das Bundesministerium für Bildung und Forschung (Förderkennzeichen 03FHP106) im Rahmen des Projekts Career@BI innerhalb der Förderrichtlinie FH Personal.

Referenzen

Beck, F. & Bader, M. (2019). Map based human motion prediction for people tracking. In 2019 IEEE/RSJ international conference on intelligent robots and systems (IROS) (S. 1–7). doi: 10.1109/IROS40897.2019.8967742

Digitaler Bahnhof Minden – Künstliche Intelligenz für den Bahnhof der Zukunft

Hasina Attaullah, Marc Halbrügge, Rolf Naumann & Thorsten Jungeblut



Bahnhofsgebäude Minden
© de:Benutzer:Aeggy, CC BY-SA 2.5 via Wikimedia Commons

Das Projekt **Digitaler Bahnhof Minden** (DiBaMi) verfolgt mehrere wichtige Ziele. Zum einen soll der Bahnhof Minden zu einem innovativen Showroom entwickelt werden, der als Ort der Kollaboration und außerschulischer Lernort dient. Ein Fokus liegt dabei auf dem Thema **Mobilität der Zukunft**, um in diesem Rahmen verschiedene Bildungs- und Dialogformate zu ermöglichen.

Ein weiterer wichtiger Aspekt ist die Erprobung neuer Technologien und Services wie beispielsweise autonomer Roboter. Im Schaufensterbahnhof werden Fahrgäste als Reallabors in die Nutzung neuer Technologien eingebunden und können somit direkt an der Weiterentwicklung des Bahnhofes mitwirken.

Das Projekt strebt eine hohe **Nutzerakzeptanz** an, indem neue Lösungen entwickelt werden, die sowohl den Bedürfnissen der Fahrgäste als auch den Anforderungen der modernen Mobilität gerecht werden. Durch die Einbindung der Bahnreisenden in den Entwicklungsprozess soll eine nachhaltige und praxisnahe Umsetzung garantiert werden (Green, Billingham, Chen & Chase, 2008).

Autonomer Gepäckwagen

Im Projekt DiBaMi wird **künstliche Intelligenz** (KI) eingesetzt, um die Mobilität der Zukunft zu gestalten und gesellschaftliche Herausforderungen zu lösen. Ein Beispiel hierfür ist der autonome Gepäckroboter, der mit KI-gestützten Algorithmen ausgestattet ist. Diese Algorithmen ermöglichen dem Roboter, seine Umgebung zu erkennen, Hindernisse zu vermeiden und sicher seinen Weg zu einem bestimmten Zielpunkt zu finden. Durch den Einsatz von KI kann der Roboter selbstständig die optimale Route planen und während seiner Fahrt flexibel auf veränderte Situationen reagieren während er bspw. **der Person folgt**, die im Augenblick mit ihm interagiert (Beck & Bader, 2019).

Durch eine breite Beteiligung von Nutzer:innen unterschiedlicher Altersgruppen, Kulturen und Mobilitätsebenen kann sichergestellt werden, dass der Roboter für alle zugänglich und benutzerfreundlich ist. **Co-Creation** mit potenziellen Nutzer:innen ermöglicht es, Bedürfnisse frühzeitig zu identifizieren und in die Designentscheidungen einzubeziehen (Beck & Bader, 2019).

Darüber hinaus sind politische Rahmenbedingungen von Bedeutung. Die Anpassung bestehender Vorschriften an neue Technologien und die Förderung einer barrierefreien Infrastruktur sind essentiell, um den Einsatz autonomer Gepäckroboter zu ermöglichen und gleichzeitig die Rechte aller Bürger:innen zu schützen. Eine transparente Kommunikation und **Partizipation der Öffentlichkeit** sind dabei unerlässlich (Mintrom et al., 2022).



Robotikplattform AgileX Ranger, die im Projekt DiBaMi verwendet wird.
© AgileX <https://www.agilex.ai/>

Diskussion

Der Einsatz eines autonomen Gepäckroboters bietet die Chance, Mobilität für alle Menschen zugänglicher zu gestalten. **Inklusives Design** spielt dabei eine zentrale Rolle, um sicherzustellen, dass der Roboter von möglichst vielen Menschen genutzt werden kann (Anvari & Wurde-mann, 2020; Weiss, Bader, Vincze, Hasenhütl & Moritsch, 2014).

Die Bedürfnisse verschiedener Nutzergruppen müssen berücksichtigt werden, z. B. Personen mit unterschiedlichen körperlichen Fähigkeiten, Seh- oder Hörbeeinträchtigungen sowie Menschen mit altersbedingten Einschränkungen. Intuitive Bedienoberflächen, klare Signalisierung und anpassbare Funktionen sind essenziell, um den Roboter für alle nutzbar zu machen (Reddy, Malviya & Kala, 2021).

Durch die Einbeziehung von Nutzer:innen aus verschiedenen Bevölkerungsgruppen in die Entwicklungsphase kann sichergestellt werden, dass der autonome Gepäckroboter wirklich barrierefrei und inklusiv ist. Daneben müssen auch die rechtlichen und politischen Rahmenbedingungen berücksichtigt werden. Die Synthese dieser verschiedenen Bedürfnisse ist für den Erfolg des Projekts essentiell.

Zusammenfassung und Ausblick

Die Entwicklung eines autonomen Gepäckwagens im Rahmen des Projekts DiBaMi stellt einen wichtigen Schritt zur Gestaltung einer inklusiveren Mobilität der Zukunft dar. Die Integration von künstlicher Intelligenz ermöglicht dem Roboter, seine Umgebung sicher zu navigieren, Gepäck sicher und effizient zu befördern, mit Nutzer:innen zu interagieren und ihren Bedürfnissen gerecht zu werden (Silva, Verdezoto, Pailacho, Millan-Norman & Hernández, 2023).

Durch die Fokussierung auf barrierefreie Designprinzipien kann der autonome Gepäckwagen sowohl Menschen mit eingeschränkter Mobilität als auch Senior:innen neue Möglichkeiten bieten und das Reiseerlebnis deutlich verbessern. Die Einbindung von Nutzer:innen aus verschiedenen Bevölkerungsgruppen in den Entwicklungsprozess sichert eine praxisnahe Umsetzung und trägt zur Schaffung einer zugänglicheren und inklusiveren Gesellschaft bei.

Die erfolgreiche Implementierung des autonomen Gepäckwagens am Bahnhof Minden hat das Potenzial, ein **Vorbild für andere Bahnhöfe** und öffentliche Verkehrsmittel zu werden. Die Weiterentwicklung dieser Technologie trägt dazu bei, die Mobilität der Zukunft nachhaltig und inklusiv zu gestalten.

Danksagungen

Dieses Projekt ist gefördert durch das EFRE/JTF-Programm NRW, REGIONALE Ostwestfalen-Lippe – Vernetzte Mobilität und digitale Anwendungen. Des Weiteren ist dieses Projekt unterstützt durch das Bundesministerium für Bildung und Forschung (Förderkennzeichen 03FHP106) im Rahmen des Projekts **Career@BI** innerhalb der Förderrichtlinie **FH Personal**.

Literatur

- Anvari, B. & Wurde-mann, H. A. (2020). Modelling social interaction between humans and service robots in large public spaces. In IEEE/RSJ international conference on intelligent robots and systems (IROS) (S. 11189–11196). IEEE. doi: 10.1109/IROS45743.2020.9341133
- Beck, F. & Bader, M. (2019). Map based human motion prediction for people tracking. In 2019 IEEE/RSJ international conference on intelligent robots and systems (IROS) (S. 1–7). doi: 10.1109/IROS40897.2019.8967742
- Fernaues, Y. (2009). Human action and experience as basis for the design and study of robotic artefacts. In RO-MAN 2009-the 18th IEEE international symposium on robot and human interactive communication (S. 522–527). doi: 10.1109/ROMAN.2009.5326218
- Green, S. A., Billingham, M., Chen, X. & Chase, J. G. (2008). Human-robot collaboration: A literature review and augmented reality approach in design. International journal of advanced robotic systems, 5 (1), 1. doi: 10.5898/IJHRI.1.2.Glas
- Mintrom, M., Sumartojo, S., Kulić, D., Tian, L., Carreno-Medrano, P. & Allen, A. (2022). Robots in public spaces: Implications for policy design. Policy design and practice, 5 (2), 123–139. doi: 10.1080/25741292.2021.1905342
- Reddy, A. K., Malviya, V. & Kala, R. (2021). Social cues in the autonomous navigation of indoor mobile robots. International Journal of Social Robotics, 13 (6), 1335–1358. doi: 10.1007/s12369-020-00721-1
- Silva, S., Verdezoto, N., Pailacho, D., Millan-Norman, S. & Hernández, J. D. (2023). Online social robot navigation in indoor, large and crowded environments. In 2023 IEEE international conference on robotics and automation (ICRA) (S. 9749–9756). doi: 10.1109/ICRA48891.2023.10160603
- Weiss, A., Bader, M., Vincze, M., Hasenhütl, G. & Moritsch, S. (2014). Designing a service robot for public space: An action-and-experiences-approach. In Proceedings of the 2014 ACM/IEEE international conference on human-robot interaction (S. 318–319). doi: 10.1145/2559636.2559791

SHARLY - A Smart Assistant for Intelligent Environments

Justin Baudisch, Thorsten Jungeblut

SHARLY is a software application which addresses the increasing demand for innovative health-care solutions in times of demographic change. As life expectancy rises and the shortage of caregivers grows, smart environments offer a promising pathway to support independent living and relieve care structures. To approach this, SHARLY collects, processes and evaluates smart home sensor data in order to provide reliable insights into daily living of its residents. SHARLY is built upon different components. The presented work shows two of them, the „Anomaly Detection“ and „Activity Index“.

The „Activity Index“ transforms raw sensor data into a single, easy-to-understand score that reflects a person's quantitative activity level and can be used for health monitoring in smart homes. To do so, each event is assigned a weight depending on its origin (e.g., a motion sensor in a hallway counts more than a pressure sensor on a sofa). To avoid artificial inflation of activity, repeated events from the same source are dampened through a capacitor mechanism that reduces their contribution over time, which reduces artificial inflation of activity scores, for example through repetitive triggering of the same sensor.

The „Anomaly Detection“ identifies deviations from learned event sequences [1]. By establishing stable temporal patterns of daily routines, SHARLY can recognize when unusual events occur, which may signal health risks or behavioral changes. This mechanism reflects the system's capacity to combine robustness with adaptability, with consideration for privacy and data protection requirements.

Together, these components show how SHARLY can be applied in practice within intelligent environments. It links prior algorithmic research with an applied framework that translates sensor-based analyses into interpretable indicators for health monitoring. By bridging algorithmic foundations with intuitive feedback, SHARLY illustrates how sensor-based monitoring can contribute to maintaining autonomy and safety in daily living while offering opportunities to reduce workload in health-care systems.

References:

- [1] Baudisch, J., Richter, B. & Jungeblut, T. A Framework for Learning Event Sequences and Explaining Detected Anomalies in a Smart Home Environment. *Künstl Intell* 36, 259–266 (2022).
<https://doi.org/10.1007/s13218-022-00775-5>

SHARLY - [S]mart [H]ome [A]gent, [R]ea[L]I[Y]?! A Smart Assistant for Intelligent Environments

Justin Baudisch, Thorsten Jungeblut



Demographic change puts **enormous pressure** on societies and **healthcare systems**

Life expectancy increases, a growing number of frail older people will require long-term care^{1,2}

Growing **shortage of caregivers** to meet the needs of all older people³

→ **Smart Home sensors offer a solution** to the demographic challenge⁴ by enabling older people to live longer independently and safely in their own homes⁵



Source: Generated by ChatGPT 4o, prompt: "Create image: An 80-year-old man has fallen in his retirement home. A red light is flashing – emergency."

SHARLY is a software solution designed for smart environments to provide support and supervision in daily life, analyzing sensor information and acts as a reliable co-pilot and caregiving assistant.

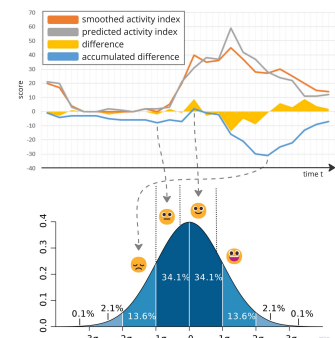
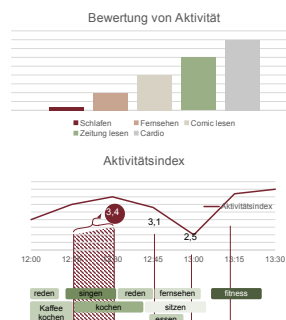
Current Components

Activity Index

General level of activity (physical and cognitive)

Calculated by accumulating all activities within a time window (e.g. every 30 minutes)

Displayed as smiley for simplicity



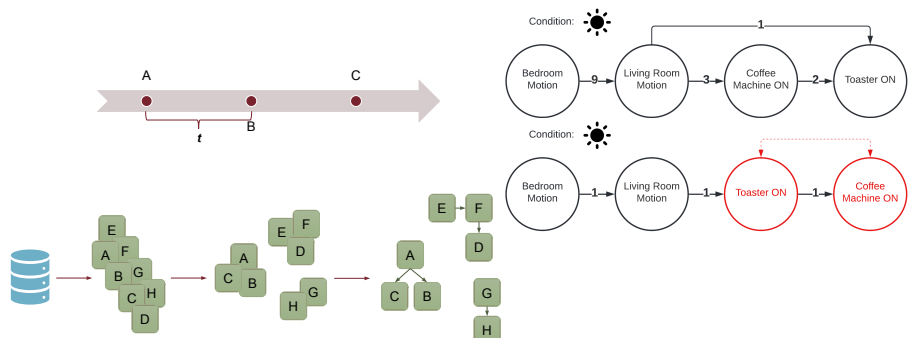
Anomaly Detection

Learn activities based on time

- Try to find a stable event-sequence by maximizing the time-delta t and minimizing the amount of events n

Record new sequences based on the learned time parameter t

Find deviations between recorded and learned sequences



Für eine inklusive digitale Zukunft: Ableismus-sensible Chatbots gestalten

Forschungsansatz

Das Kooperationsprojekt der Hochschule Bielefeld, Aktion Mensch e.V. und wonk.ai GmbH untersucht systematisch ableistische Muster und Diskriminierung aufgrund des Merkmals Behinderung in KI-Systemen. Dabei wurde ein partizipativer Ansatz gewählt, der sicherstellen soll, dass Menschen, die Behinderung erfahren, direkt eingebunden werden. Im Kontext einer Forschung zu Diskriminierungserfahrungen und einer angestrebten Verbesserung der Situation ist dieses Vorgehen unerlässlich (Behrisch 2022).

Bisherige Forschung konnte bereits andere Diskriminierungsformen wie etwa rassistische Diskriminierung oder Diskriminierung auf Grund von Geschlecht und sexueller Identität in KI-Systemen aufzeigen (Gadiraju et al. 2023). Insbesondere zu ableistischer Diskriminierung besteht jedoch ein Forschungsdesiderat (ebd.).

Innovatives Evaluationstool

Entwickelt wurde ein automatisiertes Tool, das KI-Chatbots systematisch auf diskriminierende Antworten prüft. Es führt automatisiert hunderte Gespräche und wertet diese strukturiert aus – effizient, skalierbar und objektiv. Acht Evaluationskategorien werden benotet und mit Begründungen versehen.

Mit Daten und Regelwerken gegen Ableismus in LLM

Ein umfassendes Regelwerk adressiert verschiedene sprachliche Dimensionen, wobei insbesondere die Verständlichkeit, die Verwendung respektvoller Begriffe und eine klare Textgestaltung von zentraler Bedeutung sind. Dabei wird ebenfalls darauf geachtet, stereotype Darstellungen zu vermeiden und eine differenzierte Repräsentation von Menschen mit Behinderungen zu erreichen. Der Fokus liegt zudem auf einer Stärkenorientierung, die darauf abzielt, statt auf zugeschriebene „Defizite“ den Blick sachlich auf die Fähigkeiten und Ressourcen der Individuen zu lenken.

Um sicherzustellen, dass KI-Systeme keine diskriminierenden Inhalte reproduzieren, werden strukturierte Testdatensätze entwickelt. Diese Datensätze enthalten Textbeispiele, Bewertungen, Kontextinformationen sowie technische Metadaten und stehen als Open-Source-Ressource zur Verfügung. Dies ermöglicht es Entwickler*innen, ihre Systeme frühzeitig zu testen und zu überprüfen, ob diese alle Menschen gleichwertig adressieren.

Ausblick

Das Projekt etabliert neue Qualitätsstandards für diskriminierungssensible KI. Die Kombination aus wissenschaftlicher Expertise, technischer Innovation und direkter Beteiligung von Menschen mit Erfahrungswissen schafft praktische Werkzeuge für die ethische Technologieentwicklung. Die Ergebnisse helfen, eine inklusivere digitale Zukunft zu gestalten.

Behrisch, Birgit (2022): Partizipatorische und emanzipatorische Forschung in den Disability Studies. In: Anne Waldschmidt (Hg.): Handbuch Disability Studies. Wiesbaden: Springer VS, S. 109–124. <https://doi.org/10.1007/978-3-531-18925-3>

Gadiraju, Vinitha; Kane, Shaun; Dev, Sunipa; Taylor, Alex; Wang, Ding; Denton, Emily; Brewer, Robin (2023):

"I wouldn't say offensive but...": Disability-Centered Perspectives on Large Language Models. In: 2023 ACM Conference on Fairness, Accountability and Transparency. FaccT '23: the 2023 ACM Conference on Fairness, Accountability, and Transparency. Chicago IL USA, 12 06 2023 15 06 2023. New York, NY, USA: ACM, S. 205–216. <https://doi.org/10.1145/3593013.3593989>

Für eine inklusive digitale Zukunft: Ableismus-sensible Chatbots gestalten

Generative Sprachmodelle beeinflussen unser Leben täglich und in Zukunft – doch nicht alle profitieren gleichermaßen. Viele Menschen mit Behinderungserfahrungen stoßen auf umwelt- und einstellungsbedingte Barrieren, auch KI-Systeme können ableistische Einstellungen reproduzieren.

Unser Kooperationsprojekt von HSBI, Aktion Mensch e. V. und wonk.ai analysiert moderne Sprachmodelle und entwickelt innovative Lösungen, um Barrieren abzubauen. Unser Ziel: Ableistische Sprache in generierten Texten sichtbar, messbar und vermeidbar zu machen.

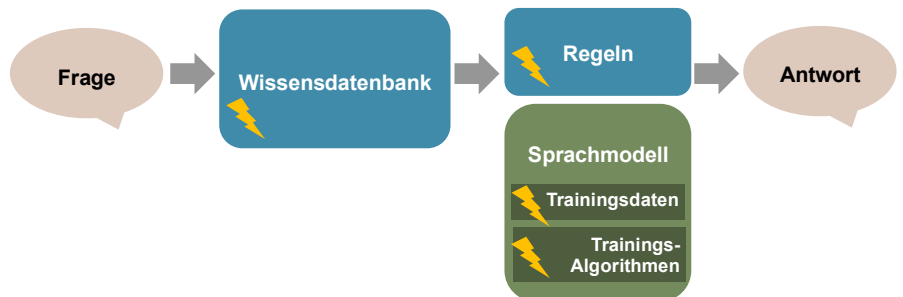


Workshop mit Menschen mit Behinderungserfahrung im Projektdurchlauf 2024

Ableismus und Sprachmodelle

Ableismus bezeichnet die gesellschaftliche Abwertung, Ausgrenzung und Benachteiligung von Menschen mit Behinderungen – oft basierend auf der Annahme, dass ein Leben ohne Behinderung wertvoller oder erstrebenswerter sei.

Ist KI ableistisch? Ja, aus mehreren Gründen: KI-Systeme lernen aus Daten, die oft versteckte Vorurteile enthalten. Besonders in RAG-Systemen (Retrieval-Augmented Generation), die bei der Beantwortung von Fragen auf umfangreiche Datenquellen zugreifen, können Verzerrungen sich wie in einer Kette vervielfachen.



Diskriminierung in KI-Systemen unter Einbindung von Retrieval-Augmented Generation (RAG) kann an zahlreichen Stellen entstehen.

Wo entsteht ableistische Sprache in Sprachmodellen?

Unsere Analyse zeigt vier kritische Punkte, an denen Ableismus entstehen kann:

- In der **Wissensdatenbank**: Wenn das Grundwissen der KI ableistische Informationen enthält.
- In den **Antwortregeln**: Wenn die Formulierungsregeln bestimmte Gruppen benachteiligen.
- In den **Trainingsdaten**: Wenn die Daten, mit denen KI lernt, nicht divers genug sind.
- In den **Algorithmen**: Wenn die Lernmethoden selbst bestimmte Muster bevorzugen.

Konkret untersuchen wir folgende Fragen:

- Was ist Ableismus in und durch KI?
- Wie manifestiert** sich Ableismus in der Sprache von KI-Systemen konkret für Menschen mit Erfahrungen von Behinderungen?
- Wie können wir ableistische Sprache in KI-Systemen **messbar** machen?

Ableismus ist komplex

Ein Voruntersuchungsergebnis zeigt: Ableistische Sprache wird unterschiedlich wahrgenommen.

Ein Test mit und ohne Menschen mit Behinderungserfahrungen verdeutlicht, wie herausfordernd es ist, Ableismus zu identifizieren.

Ein Beispiel: 33,3 % der Teilnehmenden bezeichneten die Beschreibung einer Kandidatin, die *trotz* Amputation Lebensmut zeigt, als ableistisch. Wiederum 33,3 % haben dies als eher diskriminierend eingestuft, während 14,8 % meinten, sie sei eher nicht diskriminierend und 18,5 % sie sogar als nicht diskriminierend einstufen.

Klar wird, wer Ableismus bearbeiten will, muss Menschen einbinden, die Behinderungen erfahren – und das tun wir.

Ziel I

Partizipative Entwicklung von Werkzeugen

Wir entwickeln Bewertungskriterien und bewerten Testdatensätze – in einem partizipativen Prozess, der die Perspektiven von Menschen mit Behinderungserfahrungen relevant einbezieht. In Workshops werden Kriterien diskutiert und Datensätze auf potenziell ableistische Sprache geprüft.

Die Werkzeuge stellen wir der Öffentlichkeit als Open-Source zur Verfügung.

Ziel II

Von der Theorie zur Praxis

Unsere Software ermöglicht es, KI-Service-Chats systematisch auf ableistische Antworten zu prüfen. Die Analyseergebnisse zeigen konkret, wo Verbesserungen nötig sind.

Heute ist eine systematische Prüfung noch nicht möglich – morgen könnte sie zum Standard werden. **Gemeinsam etablieren wir neue Qualitätsstandards für inklusive KI.**

Bildung gestalten in der KI-ndheit – Digitale Teilhabe in der Kita

Die Bildungswerkstatt am Fachbereich Sozialwesen der HSBI ist eine Lernumgebung, in der Studierende mit analogen und digitalen Materialien arbeiten, um Konzepte für frühe Bildungsprozesse im digitalen Wandel zu entwickeln. Ziel ist, Kinder zum reflektierten Umgang mit Digitalität und so zur Teilhabe an einer digitalen Gesellschaft zu befähigen.

Kinder wachsen in einer Welt auf, die geprägt von **digitalen Technologien und KI** ist. Tablets, Smartphones und digitale Spielzeuge sind Teil ihres Alltags. Damit verändern sich u.a. Lern-, Denk- und Interaktionsformen (Stalder 2016). Frühe Bildung sollte an der Lebenswelt von Kindern anknüpfen und auf diese Bedingungen reagieren, ohne ihre Eigen- und Interaktionsaktivität einzuschränken.

Kindertageseinrichtungen als frühe Bildungsorte spielen hierbei eine zentrale Rolle. Erste Ergebnisse einer Dissertation (HSBI) zeigen, dass Fachkräfte Bildung als die Förderung von **Selbstständigkeit, Interaktionsfähigkeit und Selbstbildung** verstehen. Diese empirisch gewonnenen Kategorien bilden eine Grundlage, um Potenziale und Herausforderungen digitaler Materialien in Kitas zu reflektieren.

In der Bildungswerkstatt wird erprobt, wie diese Bildungsziele mit digitalen und analogen Materialien verbunden werden können. Von Studierenden entwickelte Praxis-Konzepte zeigen, dass digitale Medien im Sinne einer reflektierten Teilhabe genutzt werden können – nicht als Ersatz, sondern als Erweiterung klassischer Bildungsgelegenheiten.

Eine theoretische Rahmung bietet das Konzept der **digitalen Kompetenz nach Knauf (2024)**. Sieben Kompetenzbereiche sind demnach für Kinder in einer digitalisierten Welt besonders relevant (Abb. 1, Knauf 2024). Sie sind keine technischen Fertigkeiten, sondern dienen als **pädagogisch fundierte Lernziele**, die Selbstständigkeit, Partizipation und kritisches Denken in der Kindheit ermöglichen.

Digitale Teilhabe in der Kindheit eröffnet Möglichkeiten, Bildungsprozesse und Ausdrucksformen zu individualisieren, erfordert aber eine reflektierte pädagogische Begleitung, damit digitale Angebote nicht zu Selbstbedienung, sondern zu aktiver Mitgestaltung führen.

Literatur

Knauf, H. (2024). *Förderung digitaler Kompetenzen von Kindern in Kindertageseinrichtungen. Empirische Befunde und konzeptionelle Grundlagen*. Weiterbildungsinitiative Frühpädagogische Fachkräfte, WiFF Expertisen, Band 57. München.

Stalder, F. (2016). *Kultur der Digitalität*. Suhrkamp.

Bildung gestalten in der KI-ndheit – aber wie?

Digitale Teilhabe in der Kita: von Selbstbildung bis Selbstbedienung.



Aktivität „Der menschliche Körper“, © Foto: Alina Ratsak

Digitalität in der Kindheit – muss das sein?

Kindheit in digitalen Zeiten

- Kinder wachsen in einer Welt auf, die geprägt ist von KI, digitalen Technologien und medienkulturellem Wandel. Tablets, Smartphones und interaktive Medien sind Teil ihres Alltags. In der Familie sowie im öffentlichen Raum – sie sind allgegenwärtig (Stalder 2026).
- Digitalität verändert so nicht nur das Medienverhalten, sondern auch Denkweisen, Lernformen und soziale Beziehungen – dieser Wandel berührt auch die Kindheitspädagogik (Burkhardt 2018).
- Digitale Teilhabe in der Kindheit bedeutet nicht nur Zugang zu Technik, sondern auch die Befähigung zu einem mündigen und reflektierten Umgang damit – also zur aktiven Teilnahme an einer digitalen Gesellschaft. Dazu gehört der Aufbau digitaler Kompetenzen (Knauf 2024).

→ Digitale Teilhabe in der Kindheit – wie passt das zu frühpädagogischen Bildungszielen in Kitas?

Was ist Bildung?

Frühe Bildungsprozesse – ein empirischer Blick

- Frühkindliche Bildung wird bildungspolitisch und theoretisch vielfach diskutiert, doch die Konzepte sind oft heterogen (König 2025). In Bildungsgrundsätzen werden normative Leitlinien formuliert. Daraus gehen allerdings keine eindeutigen Handlungskonzepte hervor.
- Damit bleibt die Umsetzung des Bildungsauftrags in Kitas bisher offen und ist durch die subjektiven Deutungsmuster der Fachkräfte geprägt (Betz 2015).
- Ein Dissertationsprojekt (laufend, HSB!) untersucht diese Deutungen mit der Frage, welche Bildungsverständnisse Fachkräfte in ihrer pädagogischen Praxis leiten.

→ Qualitative Interviews mit Fachkräften in Kitas zeigen: Pädagogische Fachkräfte verstehen Bildung als die Förderung von ...

... Selbstständigkeit

→ selbstständig denken & handeln

... Interaktionsfähigkeit

→ Ausdrücken eigener Denkweisen & Meinungen

... Selbstbildung

→ Ganzheitliche Nutzung von Material- & Raumbedingungen

Wie kann digitale Teilhabe frühe Bildung stärken?

Digitale Werkzeuge für die Kita

- Die Bildungswerkstatt am Fachbereich Sozialwesen der HSB! bietet Raum für die forschende Auseinandersetzung mit frühkindlichen Bildungsprozessen im Kontext von Digitalisierung. Studierende entwickeln hier praxisnahe Konzepte zur Förderung früher Bildungsprozesse in der Kita.
- Hierfür steht ein Repertoire an digitalem Material bereit. Was ist digitales Material? Dazu zählen digitale Spielzeuge wie z. B. programmierbare Roboter, Tablets sowie analoge Materialien, die die Entwicklung digitaler Kompetenzen anregen.



Aktivität „Interaktives Roboter-Labyrinth“, © Foto: Alina Ratsak

→ Beispiele aus der Bildungswerkstatt (Seminar im WiSe 2024/2025) zeigen, wie analoge und digitale Materialien zur Förderung früher Bildung unter Einbezug digitaler Kompetenzen in Kitas eingesetzt werden können:

Aktivität	Material	Inhalte	Digitale Kompetenz	Bildungsziel
Der menschliche Körper	Tablet inklusive Simulationsapp, Anatomiepuppe, Stethoskop	Die körperlichen Funktionen verstehen Entwicklung von Ableitungen für den Alltag (z. B. Trink- & Bewegungsverhalten)	Bedienen & Anwenden Informieren & recherchieren	Selbstständigkeit
Interaktives Roboter-Labyrinth	Mindestens zwei programmierbare Roboter, aufgemaltes oder konstruiertes Labyrinth	Ausdruck eigener Gedanken Perspektivübernahme Interaktive Entwicklung von Kompromissen & Problemlösungen	Kommunizieren & kooperieren Analysieren & reflektieren Problemlösen & Handeln	Interaktionsfähigkeit
Erforschen von Materialien und ihren Eigenschaften	Digitale Lupe, ggf. Tablet zum Übertragen der Bilder, verschiedene Alltagsmaterialien	Präzise Beobachtung von Phänomenen Logisches Denken Ästhetische Erfahrungen	Analysieren & reflektieren Bedienen & Anwenden	Selbstbildung

Diskussion & Ausblick

- Chancen vs. Risiken:** Frühe Bildungsprozesse und -ziele können mit digitalem Material angeregt werden. Ein wenig reflektierter Einsatz kann aber das Risiko bergen, Eigenaktivität der Kinder einzuschränken, Interaktion mit anderen und die digitale Nutzung auf reine (Selbst-)Bedienung zu reduzieren.
- Autonomie vs. Kontrolle:** Der Einbezug digitaler Teilhabe in die Kita-Praxis bietet Potenzial, Bildungsprozesse zu differenzieren und zu individualisieren. Dies erfordert eine Balance zwischen Unterstützung und dem Freiraum für Selbstständigkeit in der Auseinandersetzung mit digitalen Materialien.

Bildungswerkstatt an der HSB! – noch nie gehört?



Literatur

Betz, T. (2015). Pädagogische Fachkräfte im Spannungsfeld zwischen Selbstverständnis und externen Erwartungen. In: Springer, U., Eckmann, D. & König, A. (Hrsg.), *Erziehungswissenschaftliche Perspektiven in frühpädagogischer Theoriebildung und Forschung* (S. 221–243). Weinheim: Beltz Juventa.

Burkhardt, Y. (2023). *Mediale Möglichkeitenräume. Digitale Materialität in der frühen Kindheit*. In: Leineweber, C., Waldmann, M. & Wunder, M. (Hrsg.), *Materialität – Digitalisierung – Bildung* (S. 85–99). Bad Heilbrunn: Klinkhardt.

König, A. (2025). *Pädagogik der Kindheit. Erziehung und Bildung im soziokulturellen Wandel*. Stuttgart: Kohlhammer.

Knauf, H. (2024). *Förderung digitaler Kompetenzen von Kindern in Kindertageseinrichtungen. Empirische Befunde und konzeptionelle Grundlagen*. Weiterbildungsinitiative Frühpädagogische Fachkräfte, WiFF Expertisen, Band 57. München.

Stalder, F. (2016). *Kultur der Digitalität*. Suhrkamp.

Fachbereich Sozialwesen
Milena Michelle Förster (M.A.)
milena_michelle.foerster@hsb!.de

Betreuung der Promotion:
Prof. Dr. Helen Knauf (Hochschule Bielefeld) &
Prof. Dr. Markus Walber (Universität Bielefeld)

Statistical Learning of Daily Routines in the Smart Home for Privacy Preserving Ambient Assisted Living

Marc Halbrügge & Thorsten Jungeblut

Bielefeld University of Applied Sciences and Arts

Ambient Assisted Living focuses on enabling Ageing in Place, a goal widely supported by seniors who prefer to remain in their homes (Brittain, Corner, Robinson, & Bond, 2010). Transitioning to elderly care often leads to cognitive and emotional overload, which may cause abrupt declines in health and functional abilities.

IoT Devices for Assisted Living

Off-the-shelf smart home devices can provide helpful insights in the context of AAL. However, using cloud services with smart home devices raises privacy concerns. The inhabitants' privacy can be preserved by analyzing the data locally on edge devices instead.

While there is much research about predicting activities of daily living (ADL: eating, sleeping, reading, etc.) from sensor data (Ahad, Antar, & Ahmed, 2021), we concentrate on daily rhythms instead. While ADL are important and interesting, they may not fit the assisted living scenario best.

Our goal is to automatically notify caregivers if an older person's daily routine changes substantially. This is achieved by a) learning daily routines from observation, and b) using a model of the routines to detect anomalies.

Discussion and Conclusions

Our results show that applying machine learning to sensor data gathered from off-the-shelf smart home devices can provide valuable insights to caregivers of older adults. This can help improve their daily life and enable them to stay in their personal homes. In the future, we plan to improve the algorithms and validate them in cooperation with care professionals.

Acknowledgements

This work was supported by the German Federal Ministry of Education and Research (BMBF) as part of the *Career@BI* project within the funding program *FH Personal* (03FHP106).

References

- Ahad, M. A. R., Antar, A. D., & Ahmed, M. (2021). IoT sensor-based activity recognition. Cham: Springer. DOI: 10.1007/978-3-030-51379-5
- Brittain, K., Corner, L., Robinson, L., & Bond, J. (2010). Ageing in place and technologies of place: The lived experience of people with dementia in changing social, physical and technological environments. *Sociology of health & illness*, 32(2), 272–287. DOI: 10.1111/j.1467-9566.2009.01203.x

Statistical Learning of Daily Routines in the Smart Home for Privacy-Preserving Assisted Living

Marc Halbrügge & Thorsten Jungeblut

Ambient Assisted Living (AAL)

Ambient Assisted Living focuses on enabling **Ageing in Place**, a goal widely supported by seniors who prefer to remain in their homes (Brittain, Corner, Robinson, & Bond, 2010). Transitioning to elderly care often leads to cognitive and emotional overload, which may cause abrupt declines in health and functional abilities.

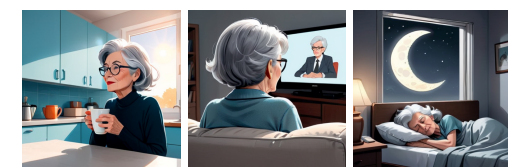
Dementia represents a significant public health challenge, with its prevalence growing exponentially with age (Kruse et al., 2002). In Germany, dementia is the fourth leading cause of death among individuals over 65 (Statistisches Bundesamt, 2024).

Mild cognitive decline is increasingly viewed as an early indicator of dementia, prompting interest in detecting subtle changes in daily routines through technological means. Smart home devices offer promising avenues for monitoring such degradation, as demonstrated by systems that track lifestyle patterns and health metrics (Wang et al., 2021; Zwierenberg et al., 2018). These innovations aim to support early intervention and improve the quality of life for aging populations.

IoT Devices for Assisted Living

Off-the-shelf smart home devices can provide helpful insights in the context of AAL. However, using cloud services with smart home devices raises **privacy concerns**. The inhabitants' privacy can be preserved by analyzing the data locally on edge devices instead. Choosing edge computing may pose additional constraints for the machine learning algorithms regarding their computational complexity, though.

- Pros of using IoT for assisted living
 - Inexpensive ubiquitous hardware
 - Can help older adults live independently and safely in their own homes (Morris et al., 2013)
 - Other benefit for the users (e.g., smart lighting)
- Cons of using IoT for assisted living
 - Potential intrusion into personal space
 - Video / audio recordings → privacy concerns



Activities throughout the day (AI generated)

Research Goal

While there is much research about predicting activities of daily living (ADL: eating, sleeping, reading, etc.) from sensor data (Ahad, Antar, & Ahmed, 2021), we concentrate on daily rhythms instead. While ADL are important and interesting, they may not fit the assisted living scenario best. From a caregiver's point of view, detailed daily activities may contain rather irrelevant and too much information. From the older adult who is taken care of, revealing the details of their daily life may intrude too much into their private space. Therefore, our research goal can be stated as such:

Caregivers should be notified if an older person's daily routines change substantially.

We want to reach this goal by

- Learning daily routines from observations
- Using a model of the routines to detect anomalies

Smart Home Data

Different algorithms were used with data from the CASAS project (Cook, Crandall, & Thomas, 2019). The data was recorded in a flat with a single elderly occupant. Sensor readings are provided for 12 months; two are additionally annotated with activities of daily living (ADL).

Typical Day (based on cluster analysis of activities)

Sleep (Bed) · Hygiene · Meds (Kitchen) · Read (Chair) · Sleep (Chair) · Read (Chair) · Prepare Food (Kitchen) · Eat (Chair) · Watch TV (Chair) · Wash Dishes (Kitchen) · Hygiene · Get Dressed (Bedroom) · Watch TV (Chair) · Sleep (Chair) · Step Out · Sleep (Chair) · Step Out · Watch TV (Chair) · Sleep (Chair) · Watch TV (Chair) · Meds (Kitchen) · Watch TV (Chair)

Markov Process

The sensor readings are modeled as a time discrete Markov process (Georgii, 2015).

- finite set of states s
- s emit observed sensor vectors o
- o may need to be enriched to ensure Markov properties (e.g., add derivation)
- Use discrete time steps. Here: 60 seconds per step
- Learning problem: State transition probabilities

Learning Algorithms

Four different algorithms have been evaluated.

- Use previous value (don't expect any change)
$$P(O_t = o_t) = P(O_{t-1} = o_t) \quad (1)$$
- Nearest neighbor based on previously observed transitions $o_j \rightarrow o_i$
$$P(O_t = o_i) = P(o_i \mid \arg \min_o \text{distance}(o, o_{t-1})) \quad (2)$$
- Use value from the previous days at that time
$$P(O_t = o_i) = P(O_{t-24h} = o_i) \quad (3)$$
- Statistical regression based on current state
$$P(O_t = o_i) = \text{expit}(o_{t-1} \cdot w) \quad (4)$$

where the weights w have been estimated from observed data with a limited time horizon.

Anomaly Score per Day

Each of the N IoT sensors is weighted based on how well the models' predictions fit to the new observations. The weighted squared differences per sensor between predictions p and observations o are summed to obtain the daily anomaly score AS .

$$AS = \sum_{t=1}^{24h} \sum_{k=1}^N w_k (p_{k,t} - o_{k,t})^2$$

The score was computed based on each of the four learning algorithms using only the annotated data to establish some ground truth. The following steps were taken:

- Use CSH101 data set, 17 binary sensors
- Discretize time to 60s steps
- Use exponentially smoothed sensor value
- Apply algorithms and compute fit (in logit domain)
- Determine day with highest Anomaly Score AS
- Determine sensor with best prediction (highest r^2)

Goodness of Fit Results

Method	r^2	$RMSE$	AUC	$\max_{AS}$	$\max_{r^2}$
Random (baseline)	.000	11.46	.500	—	—
1 Last value	.040	1.13	.577	09/07	Fridge
2 Neighbors (5d)	.009	1.98	.552	09/09	Fridge
3 Day curve (5d)	.001	1.68	.589	09/09	Entry
4 Regression (5d)	.006	2.95	.570	09/14	Entry

What happened on the days with high anomaly score?

Hypotheses based on the ADL annotations:

- 09/07 Entertaining Guests, absences from the flat
- 09/09 Long absence from the flat
- 09/14 ?? (nothing special in the annotations)

Discussion

On a minute-by-minute basis, the statistical fit is not very good. Still, the daily Anomaly Scores are viable because they are aggregated across all data points of a day. Sensors that show the strongest statistical relationship (e.g., the front door sensor and hallway motion sensor) provide the least benefit regarding the AAL use case. The degradation of daily routines would better be spotted when comparing a sensor reading to its typical value at that time of the day (Eq. 3).

All algorithms used here can run efficiently on edge devices. This means better **privacy protection** as no data needs to be transferred over the internet.

Establishing ground truth is a hard problem in this domain. Within the 10 months of data without human annotations, we have been looking at the Anomaly Scores for public holidays. These yielded promising trends that are worth further investigation.

Conclusion and Future Work

We have shown that applying machine learning to sensor data gathered from off-the-shelf smart home devices can provide valuable insights to caregivers of older adults. This can help improve their daily life and enable them to stay in their personal homes. In the future, we plan to improve the algorithms and validate them in co-operation with care professionals.

Acknowledgements

This work was supported by the German Federal Ministry of Education and Research (BMBF) as part of the *Career@BI* project within the funding program *FH Personal* (03FHP106).

References

- Ahad, M. A. R., Antar, A. D., & Ahmed, M. (2021). *IoT sensor-based activity recognition*. Cham: Springer.
- Brittain, K., Corner, L., Robinson, L., & Bond, J. (2010). Ageing in place and technologies of place: The lived experience of people with dementia in changing social, physical and technological environments. *Sociology of health & illness*, 32(2), 272–287.
- Cook, D., Crandall, A., & Thomas, B. (2019). *Human activity recognition from continuous ambient sensor data*. UCI Machine Learning Repository. doi: 10.24432/C5D60P
- Georgii, H.-O. (2015). *Stochastik* (5th ed.). Berlin: De Gruyter.
- Kruse, A., Gaber, E., Heuft, G., Oster, P., Re, S., & Schulz-Nieswandt, F. (2002). *Gesundheit im Alter*. Robert Koch-Institut.
- Morris, M. E., Adair, B., Miller, K., Ozanne, E., Hansen, R., Pearce, A. J., ... Said, C. M. (2013). Smart-home technologies to assist older people to live well at home. *Journal of Aging Science*, 1(1), 1–9. Retrieved from <http://dx.doi.org/10.4172/jasc.1000101> doi: 10.4172/jasc.1000101
- Statistisches Bundesamt. (2024). *Todesursachen in Deutschland 2023*. (Statistischer Bericht 2120400237005)
- Wang, J., Spicher, N., Warnecke, J. M., Haghi, M., Schwartz, J., & Deserno, T. M. (2021). Unobtrusive health monitoring in private spaces: The smart home. *Sensors (Switzerland)*, 21(3), 1–23. doi: 10.3390/s21030864
- Zwierenberg, E., Nap, H. H., Lückien, D., Cornelisse, L., Finnema, E., Hagedoorn, M., & Sander-man, R. (2018). A lifestyle monitoring system to support (in) formal caregivers of people with dementia: Analysis of users need, benefits, and concerns. *Gerontechnology*, 17(4), 194–205.

AutoDSE: Towards HW/AI co-design of ultra-low latency hardware accelerators for industrial applications

Farjana Jalil, Julien Leuering, Qazi Arbab Ahmed, Wolfram Schenck, Thorsten Jungeblut

The increasing use of artificial intelligence (AI) in industrial environments introduces new challenges when deploying machine learning (ML) models, due to strict constraints on latency, power, and computational resources. Traditional neural networks often exceed these limitations. The AutoDSE project addresses this challenge through a holistic and automated HW/AI co-design approach that integrates approximation techniques to enable efficient AI deployment on platforms like ASICs and FPGAs.

AutoDSE begins with training baseline full-precision neural networks, then combines AI-level optimizations with hardware-level approximations and utilizes design space exploration (DSE) to efficiently guide the selection of Pareto-optimal configurations. AI-level strategies include static quantization (e.g., UINT16, UINT8, UINT4), while hardware-level optimizations leverage approximate arithmetic components from the EvoApprox^[1] library. Design candidates are evaluated based on accuracy, area, power, and latency, then synthesized and tested on real-world systems.

Initial results show that using approximate arithmetic in a 16-bit quantized model leads to measurable reductions in area and power, while maintaining similar inference latency (62 ms vs. 65 ms) and minimal accuracy loss.

This approach lays the groundwork for scalable, energy-efficient AI systems tailored to embedded and industrial use cases. Future work will include techniques such as active learning^[2] to tackle challenges like sparse data, and investigate lower-precision quantization (e.g., 8-bit) to further enhance hardware efficiency.

References

- [1] Mrazek, Vojtech, Radek Hrbacek, Zdenek Vasicek, and Lukas Sekanina. 2017. "Evoapprox8b: Library of approximate adders and multipliers for circuit design and benchmarking of approximation methods." Design, Automation & Test in Europe Conference & Exhibition (DATE). IEEE. <https://doi.org/10.23919/DATe.2017.7926993>.
- [2] Tharwat, Alaa, and Wolfram Schenck. 2020. "Balancing Exploration and Exploitation: A Novel Active Learner for Imbalanced Data." Knowledge-Based Systems 210: 106500. <https://doi.org/10.1016/j.knosys.2020.106500>.

AutoDSE: Towards HW/AI co-design of ultra-low latency hardware accelerators for industrial applications

Farjana Jalil, Julien Leuring, Qazi Arbab Ahmed, Wolfram Schenck, Thorsten Jungeblut

Motivation

Advances in intelligent technical systems (ITS) and the integration of artificial intelligence (AI) have introduced new challenges for deploying machine learning (ML) models in industrial environments, where low latency, energy efficiency, and real-time processing are critical. These constraints often conflict with the high computational demands of traditional neural networks. This project addresses these challenges through a holistic, automated design approach—from model architecture to real-world inference—centered on a hardware (HW)/AI co-design strategy. By integrating approximation techniques such as algorithmic-level quantization and HW-level optimizations, the proposed methodology aims to efficiently map AI models onto resource-constrained platforms like ASICs and FPGAs. The automated design space exploration ensures scalable, low-power AI systems that meet strict latency budgets while adapting to diverse industrial requirements.

Methodology

Figure 1 illustrates AutoDSE, the holistic and automated approach with methods spanning both the AI and HW levels to be able to fulfill the specific requirements of the user and hardware.

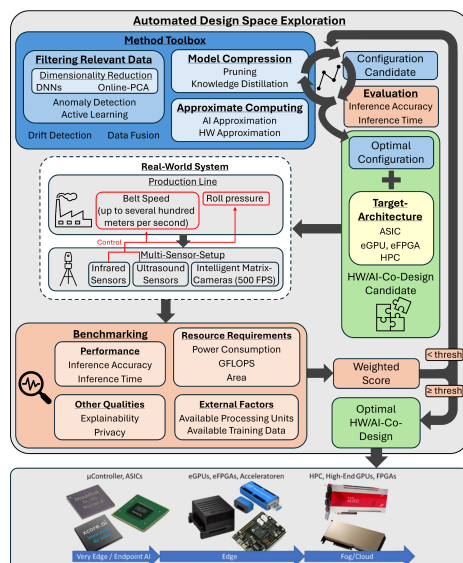


Figure 1: Operational flow of the system. blue=software; yellow=hardware; green=co-design; orange=evaluation; white=not directly part of the system

A **Method Toolbox** aims to provide solutions like Model Compression and Approximate Computing. In this study, we look at the possibilities of **HW/AI co-approximation** (Figure 2).

At the **AI-level**, the process begins with training baseline models (e.g., feed-forward neural networks (FFNNs)) in full-precision (FP32), followed by approximation techniques such as static quantization (e.g., UINT16, UINT8, UINT4) to reduce memory and computation cost.

At the **HW-level**, exact arithmetic components (e.g., adders and multipliers) are replaced with approximate alternatives from a pre-characterized library^[1], targeting area and energy savings.

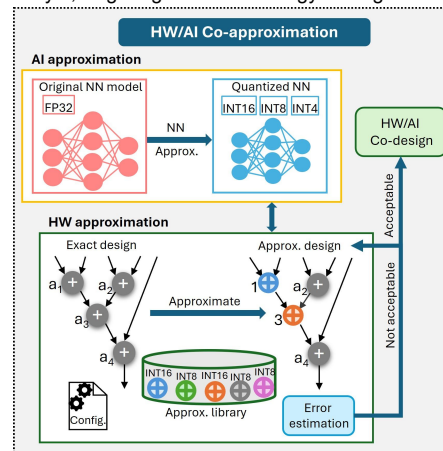


Figure 2: Overview of HW/AI co-approximation for efficient neural network (NN) deployment on target platforms.

Candidate approximated configurations are evaluated through simulations. Furthermore, error estimations guide trade-offs between accuracy, area and power. If the derived result is found feasible, a **HW/AI co-design candidate** is synthesized for the **target architectures**.

Afterwards, the candidate is benchmarked on a **real-world system** for the predefined design metrics. Based on the resulting **weighted score**, the system either deploys the **HW/AI co-design** on the target platform, or searches for another suitable configuration.

This project intends to optimize the entire processing chain for efficient embedded AI execution. This includes all levels from intelligent sensors to the cloud. The automated flow enables efficient design space exploration, ensuring the deployment of Pareto-optimal solutions for embedded AI across diverse industrial scenarios.

[1] Mrazek, Vojtech, Radek Hrbacek, Zdenek Vasicek, and Lukas Sekanina. 2017. "EvoApprox8b: Library of Approximate Adders and Multipliers for Circuit Design and Benchmarking of Approximation Methods." Design, Automation & Test in Europe Conference & Exhibition (DATE). IEEE. <https://doi.org/10.23919/DATE.2017.7926993>.

Results

As an initial proof of concept, we applied HW/AI co-approximation to a 16-bit quantized FFNN, using approximate adders and multipliers from the EvoApprox library^[1]. Figure 3 shows that with increasing error thresholds, the co-approximated design consistently reduces area and power compared to the original implementation.

For the time being, inference time was measured at the AI level. Importantly, the approximated model (16-bit) achieved 62 ms latency, which is moderately faster than 65 ms for the original. These early results reveal the efficacy of approximation for optimal co-design development with minimal accuracy loss, yet lower latency.

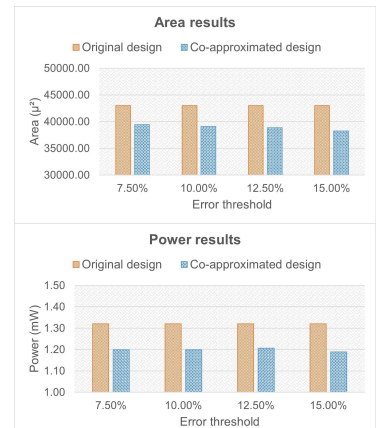


Figure 3: Preliminary results of co-approximated FFNN model to reduce area and power consumption.

Conclusion and Outlook

The project aims to develop a holistic, automated system for ultra-low latency AI deployment on resource-constrained hardware through HW/AI co-design, leading to development time reduction. The initial results demonstrate the prospective of approximation techniques at both HW and AI levels to reduce resource usage and latency. This co-approximation approach lays a strong foundation for scalable, energy-efficient and ultra-low latency AI accelerators.

Future work will explore techniques such as active learning^[2] to tackle challenges like sparse data and further optimize the latency, as well as investigate lower-precision quantization (e.g., 8-bit) to enhance hardware efficiency.

[2] Tharwat, Alaa, and Wolfram Schenck. 2020. "Balancing exploration and exploitation: A novel active learner for imbalanced data." Knowledge-Based Systems 210: 106500. <https://doi.org/10.1016/j.knsys.2020.106500>.

Supported by:

Career@BI (Grant no. 03FHP106),
Edge4SparseML (Grant no. 13FH572KX2),
funded by BMFTTR

With funding from the:
Federal Ministry
of Research, Technology
and Space

BST
an eLEXIS company

hsbi.de

Edge4OD: An Automated Framework for Efficient Execution of Object Detection Algorithms on Edge Devices

Mohammad Ibrahim Köse, Qazi Arbab Ahmed, Thorsten Jungeblut

1. Introduction

Artificial intelligence on edge devices is gaining importance in areas such as autonomous driving, smart homes, robotics, and surveillance. Particularly in computer vision applications, real-time image processing is crucial but difficult to achieve due to the limited computational resources of embedded hardware. While deep neural networks offer high accuracy, their high resource requirements make them unsuitable for direct deployment. Cloud-based inference, on the other hand, introduces latency and poses privacy risks, making it inappropriate for many real-world applications. Edge AI provides a solution by enabling local data processing, thereby reducing latency and protecting sensitive information. However, efficiently deploying AI models on edge hardware demands careful trade-offs between speed, accuracy, model size, and power consumption.

2. Research Objectives

This project investigates how various combinations of object detection models, quantization techniques, deployment toolchains, and hardware platforms influence the performance of edge AI systems. The primary research objective is to identify optimal trade-offs between accuracy, inference speed, energy efficiency, and model compactness.

3. Framework Concept

To address these challenges, we propose a fully automated framework for edge AI deployment. The framework explores a wide design space by varying critical parameters such as **Target Hardware**: (FPGA, STM32 NPU), **Model Architecture** (YOLOv8n, REQ-YOLO), **Quantization** (PTQ, QAT, 8-bit or 4-bit precision), **Deployment Toolchains** (FINN, STM32Cube.AI, TVM). Each parameter combination defines a “design point” that is deployed on real hardware to collect key performance metrics. These include accuracy (mAP), inference time (FPS), power consumption, model size, and hardware resource utilization. Unlike purely simulation-based evaluations, our framework incorporates physical measurements on embedded devices to reflect real-world constraints. The optimization goal is not to find a single best configuration but rather a Pareto front - a set of configurations that offer optimal trade-offs, such as speed versus accuracy. Initial experiments on the STM32N6 platform confirm the platform's strength in terms of low latency, power efficiency, and compact model deployment. These advantages make it a promising candidate for real-time edge AI applications where privacy and energy constraints are crucial.

5. Outlook

Future work will extend the framework to support additional models, quantization techniques, and hardware platforms. Emphasis will be placed on improving automation, evaluation reproducibility, and integrating the results into a user-friendly design tool for practitioners and researchers in embedded AI.

Edge4OD: An Automated Framework for Efficient Execution of Object Detection Algorithms on Edge Devices

Mohammad Ibrahim Köse, Qazi Arbab Ahmed, Thorsten Jungeblut

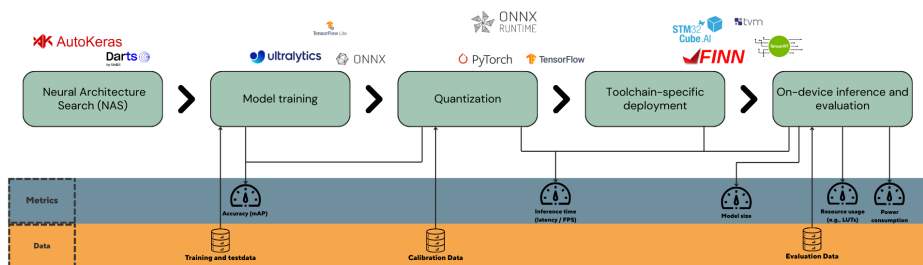


Figure 1: Overview of the deployment workflow for neural networks on edge devices. At different stages of the pipeline, specific performance metrics such as accuracy, inference time or model size are systematically collected using appropriate datasets to guide design decisions and optimize deployment.

1. Motivation

Artificial intelligence on edge devices is becoming increasingly relevant in domains like autonomous driving, smart homes, robotics, and surveillance. Especially in the field of computer vision, real-time image processing is crucial—but remains challenging due to hardware limitations. Conventional deep neural networks offer high accuracy, but their computational and memory demands make them unsuitable for direct deployment on resource-constrained hardware. Offloading to the cloud introduces latency and privacy concerns, which is unacceptable in many real-world use cases. Edge AI allows data to be processed locally, reducing latency and preserving sensitive information. However, efficient deployment of models on embedded systems requires careful trade-offs between model complexity, speed, accuracy, and power consumption.

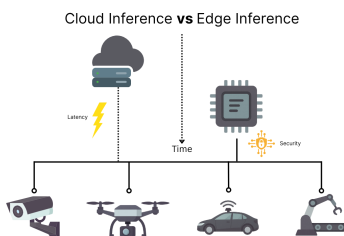


Figure 2: An exemplary scenario where a cloud-based processing introduces latency and privacy concerns, whereas edge inference enables low-latency and local data handling with improved security.

To address these challenges, the concept of Design Space Exploration (DSE) combined with AI/Hardware Co-Design has emerged as a promising approach.

2. Research Objectives

This project explores how different combinations of object detection models, quantization techniques, deployment toolchains, and hardware platforms impact the efficiency of edge AI deployments. The central research question focuses on identifying trade-offs between accuracy, inference speed, model size, and energy consumption. Additional aspects include model-to-hardware mapping, quantization effectiveness, toolchain compatibility, and the feasibility of a reproducible evaluation pipeline.

3. Framework Concept

Efficient deployment of deep learning models on edge devices involves many interdependent factors—model architecture, quantization, toolchain, and hardware. These create a large design space with many valid configurations. The goal is to develop an automated framework that explores this space systematically and deploys real configurations on physical devices to measure performance.

The framework systematically varies the following parameters:

- Target Hardware: e.g. FPGA, STM32 NPU, GPU, MCU
- Model: e.g. YOLOv8n, REQ-YOLO
- Quantization: QAT and PTQ (e.g., 8-bit, 4-bit)
- Deployment Toolchain: e.g. FINN, STM32Cube.AI, TVM

Each configuration forms a design point that will be evaluated

While simulation-based evaluation is useful for quick exploration, it often fails to reflect real-world constraints on embedded hardware.

That's why this framework includes prototype measurements on real devices for selected configurations. Prototype measurements of:

- Accuracy (mAP)
- Inference time (latency / FPS)
- Power consumption
- Model size
- Resource usage (e.g., LUTs)

This is a multi-objective optimization problem—no single best solution. Instead, a Pareto front is sought: a set of non-dominated designs that represent optimal trade-offs (e.g., speed vs. accuracy).

4. Preliminary Results

The STM32N6 platform demonstrates notable performance for edge AI applications, especially when model complexity is reduced through quantization. It supports both TFLite and ONNX models, leveraging the STM32Cube.AI toolchain to compile and deploy models as efficient C-code using the STM internal NPU compiler and STM-AI generation flow. A DNN will be deployed on the new STM32N6570-DK board and the possibilities and results will be evaluated.



Figure 3: Example deployment of a CNN using ST's toolchain: The model is first converted using ST Edge AI to generate optimized C code, which is then integrated in STM32CubeIDE for deployment to the target hardware.

Preliminary results have successfully confirmed the STM32N6's capabilities in terms of inference speed, low power consumption, and overall resource efficiency, making it a strong candidate for real-time, privacy-sensitive edge deployments. As such, this microcontroller will be actively integrated into the project's upcoming phases. Table 1 presents initial benchmarking results across all models, each quantized to INT8 PTQ and evaluated at a fixed input resolution of 416×416×3.

Model	Test-Dataset	mAP@30 (on-target)	Performance (FPS)	Model size (MB)	Activations (MB)
yolov8n_416_voc	voc_person	85.0%	15	2.926	2.749
st_yolo_x_nano_416_coco	voc_person	80.2%	18	1.004	2.579
tiny_yolo_v2_416_coco	voc_person	76.6%	50	10.575	1.836

Table 1: Comparison of three YOLO models in terms of accuracy (mAP), inference speed (FPS), model size, and memory usage (activations) on the STM32N6570-DK using the voc_person dataset.

Green by Design Nachhaltiger Tiefbau durch KI

Lena Golin, André Matutat, Prof. Dr.-Ing. Hans Brandt-Pook Institute for Data Science Solutions (IDaS) der Hochschule Bielefeld

Ziel

Der Tiefbau ist ein besonders ressourcenintensiver Bereich der Bauwirtschaft (vgl. Hamilton 2022). Eingriffe in Boden, Wasser und Infrastruktur verursachen hohe CO₂-Emissionen und erfordern zugleich eine präzise Planung und Dokumentation. Vor diesem Hintergrund wird im Rahmen eines laufenden Projekts gemeinsam mit der Naue GmbH ein KI-basiertes System entwickelt, das ökologische Potenziale im Planungsprozess besser sichtbar machen und den Einsatz umweltfreundlicher Geobaustoffe gezielt unterstützen soll. Ziel ist es, unternehmensinternes Fachwissen nutzbar zu machen, Planungs- und Beratungsprozesse effizienter zu gestalten und die CO₂-Bilanz von Bauvorhaben frühzeitig zu berücksichtigen.

Methoden

Als technologische Grundlage dient ein unternehmensspezifisches Retrieval-Augmented-Generation-System (RAG), das große Sprachmodelle mit einem gezielten Zugriff auf interne Dokumente kombiniert. Dieser Ansatz folgt dem Prinzip, bei dem relevante Informationen zunächst über einen Retriever identifiziert und anschließend von einem generativen Sprachmodell kontextsensitiv verarbeitet werden (vgl. Barnett et al. 2020). Das entstehende *Naue-Wiki* fungiert als zentrale, durchsuchbare Wissensbasis mit rollenbasierter Zugriffsteuerung. Ein besonderer Fokus des laufenden Entwicklungsprozesses liegt auf der strukturierten Aufbereitung technischer Dokumente sowie der Integration grafischer Inhalte. Dabei werden unter anderem Methoden erprobt, um semantische Relationen zwischen Text, Bildern und Bildgruppen automatisiert zu erkennen und korrekt abzubilden.

Ergebnisse(vorläufig)

Die bisher entwickelten Prototypen zeigen, dass RAG-Systeme eine vielversprechende Grundlage für den gezielten Wissenstransfer im Bauwesen bieten. Erste Anwendungsszenarien deuten darauf hin, dass durch die schnelle Bereitstellung relevanter Informationen Planungsentscheidungen nicht nur beschleunigt, sondern auch nachhaltiger getroffen werden können. Ergänzend wurde ein CO₂-Kalkulator in das System integriert, der Eingaben zu Materialien, Mengen und Transportwegen verarbeitet und auf Basis verfügbarer Emissionsdaten (z. B. aus EPDs) vorläufige CO₂-Bilanzen erstellt. Diese Funktion wird aktuell weiterentwickelt und validiert.

Ausblick

In den kommenden Projektphasen soll die semantische Verarbeitung visueller Inhalte weiter verbessert werden, um Grafiken und Bilder nicht nur als Anhänge, sondern als eigenständige Wissensquellen nutzbar zu machen. Dabei sollen aktuelle Fortschritte im Bereich multimodaler Bild-Text-Verarbeitung (z. B. Li et al. 2022) aufgegriffen und für den Einsatz in RAG erweitert werden. Auch die Genauigkeit und Anwendbarkeit des CO₂-Kalkulators wird weiter ausgebaut, um belastbare ökologische Bewertungen bereits in frühen Planungsphasen zu ermöglichen. Langfristig zielt das Projekt darauf ab, KI-Systeme als integralen Bestandteil eines nachhaltigen Planungsprozesses im Tiefbau zu etablieren.

Literatur

Hamilton, Ian, et al. 2022 Global Status Report for Buildings and Construction: Towards a Zero-Emission, Efficient and Resilient Buildings and Construction Sector. Nairobi: United Nations Environment Programme. <https://globalabc.org/resources/publications/2022-global-status-report-buildings-and-construction>.

Barnett, Scott, Stefanus Kurniawan, Srikanth Thudumu, Zach Brannelly, und Mohamed Abdelrazek. 2024. „Seven Failure Points When Engineering a Retrieval Augmented Generation System.“ CAIN 2024. doi: 10.1145/3644815.3644945

L. H. Li et al., 2022 "Grounded Language-Image Pre-training," 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), New Orleans, LA, USA., pp. 10955-10965, doi: 10.1109/CVPR52688.2022.01069.



Abb. 1: Verlegte Geogitter aus gedünntem, monolithischem Polyester (PET) oder Polypropylen (PP) kombinieren hohe Stabilität und geringe Kriechneigung mit extremer Robustheit.

Ausgangssituation

Der Tiefbau ist ein besonders ressourcenintensiver Sektor der Bauwirtschaft - Eingriffe in Boden, Wasser und Infrastruktur sind nicht nur technisch anspruchsvoll, sondern mit erheblichen CO₂-Emissionen verbunden.

Vor dem Hintergrund wachsender Anforderungen an Nachhaltigkeit, Qualitätssicherung und Dokumentation sind neue Ansätze zur digitalen Unterstützung gefragt.

Die Naue GmbH ist ein mittelständisches Unternehmen mit Fokus auf Geobaustoffe (vgl. Abbildung 1) für Umweltgeotechnik, Wasserwirtschaft und Verkehrswegebau untersucht seit mehreren Jahren den Einsatz von KI-Systemen im Planungsprozess.

Ziel ist es, unternehmensspezifisches Fachwissen effizienter verfügbar zu machen, den Beratungsprozess zu optimieren und die ökologischen Einsparpotenziale durch den Einsatz von Naue-Produkten nachvollziehbar aufzuzeigen.

Fachbereich 5
Lena Golln
Andre Matulat
lena.golln@hsbi.de
andre.matulat@hsbi.de
hans.brandt-pook@hsbi.de

Hochschule Bielefeld
Interaktion 1, 33619 Bielefeld

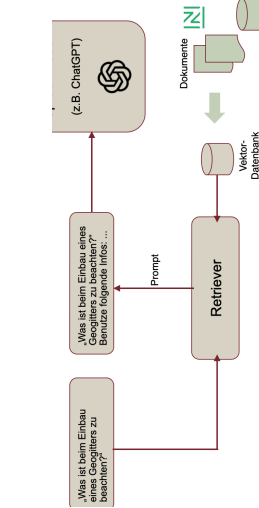


Abb. 2: Darstellung eines RAG-Systems mit verschiedenen Wissensquellen.

Grüner planen durch RAG

Retrieval-Augmented Generation (RAG) kombiniert große Sprachmodelle mit gezielter Informationsabfrage aus firmeninternem Wissen. So entstehen Antworten, die sprachlich kohärent und zugleich inhaltlich fundiert sowie kontextbezogen sind.

Im Rahmen der KI-Initiative bei Naue wurde ein unternehmensinternes RAG-System entwickelt, das in Form des „Naue-Wikis“ die zentrale Wissensbasis bildet. Eine rollenbasierte Zugriffssteuerung sorgt dafür, dass Nutzergruppen - etwa Technik, Vertrieb oder Planung - jeweils nur die für sie relevanten Informationen erhalten.

Abbildung 2 zeigt schematisch den Ablauf eines Nutzerprompts im RAG-System - von der Anfrage über die Wissensabfrage bis zur generierten Antwort.

Durch die schnelle, zielgerichtete Informationsbeschaffung unterstützt das System nicht nur den Planungsprozess, sondern auch den gezielten Einsatz umweltfreundlicher Geobaustoffe - und trägt so zur CO₂-Reduktion bei.

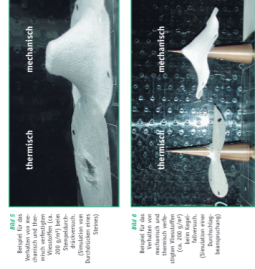


Abb. 3: Beispiel grafischer Inhalte aus der Wissensquelle mit Bildbeschreibungen neben den Bildern.

Grafiken im RAG

Die Einbindung von Grafiken in das RAG-System ist mit spezifischen Herausforderungen verbunden, da durch die technische Aufbereitung der Dokumente für die Wissensbasis visuell klare Zuordnungen von Text und Bildern verloren gehen. Diese müssen dann technisch neu ermittelt werden. Im Projekt wurden verschiedene solcher Szenarien untersucht und Lösungen erarbeitet.

Abbildung 3 zeigt ein Beispiel, bei dem die Bildunterschrift direkt neben dem Bild steht. Diese räumliche Nähe macht die Zuordnung für den Menschen einfach, doch für das RAG-System muss diese Beziehung durch die strukturierte Dokumentenrepräsentation korrekt erfasst werden.

Abbildung 4 zeigt eine weitere Herausforderung: Hier geht es um eine Bilderkette, bei der die Herausforderung darin besteht, die Abbildungen als zusammenhängende Gruppe zu erkennen und sicherzustellen, dass der Fließtext korrekt auf diese Bilder verweist.

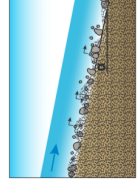
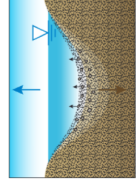


Abb. 4: Beispiel von Bilderkarten mit Bildunterschrift und Fließtext.

Tools für KI: CO2-Rechner

Im Projekt wurde ein CO₂-Kalkulator an das RAG-System angebunden, um den CO₂-Fußabdruck von Bauvorhaben zu berechnen. Der Chatbot erkennt relevante Eingaben und leitet diese an interne Kalkulationswerkzeuge weiter.

Mit einer regelbasierten Logik werden Materialien, Mengen und Transportwege erfasst, und Emissionsdaten aus EPDs sowie Emissionsfaktoren genutzt, um die CO₂-Bilanz zu berechnen. So unterstützt das System nicht nur bei der Planung, sondern auch bei der nachhaltigen Bewertung von Bauprojekten.

Ausblick

Zukünftige Entwicklungen zielen darauf ab, Bilder als eigenständige Informationsquelle in RAG-Systeme zu integrieren und deren semantische Durchsuchbarkeit zu ermöglichen. Der CO₂-Kalkulator soll verfeinert werden, um noch genauere und belastbarere Ergebnisse zu liefern.

Erforschung von Gefahren durch Adversarial Attacks beim Einsatz des LLM-getriebenen Roboters Furhat in der Lehre

Thema

Large Language Models (LLMs) haben mit der Veröffentlichung von ChatGPT erheblich an Popularität gewonnen. Die Einsatzgebiete dieser Produkte sind weitreichend. Auch im universitären Kontext haben diese Modelle bereits Einzug erhalten.

KI und IT-Sicherheit

Generative KI bedroht die IT-Sicherheit in zweierlei Maß: Angreifenden wird es deutlich vereinfacht mit Tools wie ChatGPT gezieltere Phishing- oder Scam-Attacken durchzuführen. Und auch die Modelle selbst weisen inhärente Schwachstellen auf, die Angreifende dafür nutzen können die Firmen, die die Modelle bereitstellen, Nutzende dieser Modelle oder Unbeteiligte anzugreifen.

Fragestellung

Untersucht werden soll, wie verschiedenste Adversarial Attacks gegen das verwendete LLM eingesetzt werden könnten. Es geht dabei um die Frage: Welche Risiken entstehen bei der Nutzung eines solchen LLM-getriebenen Roboters durch Adversarial Attacks von außen?

Vorgehen

Es kommt ein Roboter der Firma Furhat Robotics zum Einsatz. Der Roboter verfügt über eine Spracherkennung, deren Ausgaben an die API von OpenAI übergeben werden. Zum Einsatz kommt das Modell ChatGPT 4o. Die Ausgaben des Sprachmodells werden als Antworten des Roboters verwendet.

Die Schwachstellen werden in verschiedenen Szenarien behandelt. Dabei wird in Konversationen mit dem Roboter versucht, die jeweilige Schwachstelle sichtbar zu machen. Damit können Aussagen über die Umsetzbarkeit und Effektivität eines Angriffs, der die jeweilige Schwachstelle ausnutzt, getroffen werden.

Vorläufige Ergebnisse

Es lassen sich einige vorläufige Ergebnisse aus der bisherigen Arbeit wie folgt zusammenfassen: Misinformation scheint weiterhin ein großes Problem zu sein, denn bei geringer oder keiner Datengrundlage beantwortet das Modell Fragen häufiger falsch. Darüber hinaus können Ausgaben, die mit der Sprachausgabe des Roboters nicht harmonisieren, Unterbrechungen hervorrufen. Abschließend sollte vor allem die Sensibilisierung von Nutzenden über die Schwächen von LLMs als unabdingbar hervorgehoben werden.

Erforschung von Gefahren durch Adversarial Attacks beim Einsatz des LLM-getriebenen Roboters Furhat in der Lehre

Tjark Nitsche, Prof. Dr. Thomas Süße

Thema

Large Language Models (LLMs) haben mit der Veröffentlichung von ChatGPT erheblich an Popularität gewonnen. Die Einsatzgebiete dieser Produkte sind weitreichend. Auch im universitären Kontext haben diese Modelle bereits Einzug erhalten, bspw. als Schreibunterstützung bei wissenschaftlichen Arbeiten [1].

Roboter, die spezifische menschliche Aufgaben erfüllen und dabei eine gewisse soziale Interaktion simulieren sollen, sogenannte **Social Robots**, erleben insbesondere im Angesicht eines weitreichenden Fachkräftemangels ebenfalls einen Aufschwung in der Forschung.

KI und IT-Sicherheit

Generative KI bedroht die IT-Sicherheit in zweierlei Maß: Angreifen wird es deutlich vereinfacht mit Tools wie ChatGPT **gezielte Phishing- oder Scam-Attacken** durchzuführen. Und auch die Modelle selbst weisen **inhärente Schwachstellen** auf, die Angreifende dafür nutzen können die Firmen, die die Modelle bereitstellen, Nutzende dieser Modelle oder Unbeteiligte anzugreifen

Andererseits wird KI bereits seit längerem als Werkzeug in der IT-Sicherheit eingesetzt. KI kann vom **Risk Assessment** über die **Erkennung** bis hin zur aktiven **Abwehr** von Angriffen unterstützen. [2]

Angriffsmethoden

Orientiert wird sich bei den Angriffsmethoden an den **OWASP Top 10 for LLM Applications 2025**:

Schwachstelle	Beschreibung
Misinformation	Sogenannte „ Halluzinationen “ des Sprachmodells, durch die Informationen falsch wiedergegeben oder erfunden werden. Nutzende rufen diese meist unabsichtlich hervor
Direct Prompt Injection	Versuch eines Angreifenden, durch gestellte Prompts, eingebaute Sicherheitsvorkehrungen des Modells zu umgehen , um unerwünschte Inhalte zu (re-) produzieren.
Unbounded Consumption	Lahmlegen des Modells durch das stellen komplexer Anfragen oder Fluten des Servers mit Anfragen.
System Prompt Leakage	Durch Prompt Injection oder ähnliche Methoden unterstützter Versuch, an nichtöffentliche Informationen im System Prompt zu gelangen.
Indirect Prompt Injection	Das manipulieren des Outputs aus einem Sprachmodell bspw. durch versteckte Befehle in einem Dokument, dass Nutzende an das Sprachmodell geben.

In dem Projekt behandelte Schwachstellen nach den OWASP Top 10 [3].

Fragestellung

Eine Zusammenführung der Themen LLMs und Social Robots ist ein nächster logischer Schritt in der Entwicklung der beiden Felder. Diese Kombination kommt auch in diesem Projekt zum Einsatz. Konkret, wird der Social Robot **Furhat** als Unterstützung für die Dozierenden in der Lehre eingesetzt.

Untersucht werden soll, wie verschiedenste Adversarial Attacks gegen das verwendete LLM eingesetzt werden könnten. Es geht dabei um die Frage: **Welche Risiken entstehen bei der Nutzung eines solchen LLM-getriebenen Roboters durch Adversarial Attacks von außen?**



Einsatz des Furhat-Roboters in einer Vorlesung

Vorgehen

Es kommt ein Roboter der Firma Furhat Robotics zum Einsatz. Der Roboter verfügt über eine Spracherkennung, deren Ausgaben an die **API eines Sprachmodells** übergeben werden. Als Sprachmodell wird die API von OpenAI angebunden. Zum Einsatz kommt das Modell ChatGPT 4o. Die Ausgaben des Sprachmodells werden als Antworten des Roboters verwendet.

Die Schwachstellen werden in verschiedenen Szenarien behandelt. Dabei wird in Konversationen mit dem Roboter versucht, die jeweilige **Schwachstelle sichtbar zu machen**. Damit können Aussagen über die Umsetzbarkeit und Effektivität eines Angriffes, der die jeweilige Schwachstelle ausnutzt, getroffen werden.

Genutzt werden hierfür sowohl Methoden, die sich bereits bewährt haben, als auch grundsätzlich neue Herangehensweisen, die speziell auf den Anwendungsfall zugeschnitten sind.

Vorläufige Ergebnisse

Es lassen sich einige vorläufige Ergebnisse aus der bisherigen Arbeit wie folgt zusammenfassen:

- **Misinformation** scheint weiterhin ein großes Problem zu sein
 - Bei geringer oder keiner **Datengrundlage** beantwortet das Modell Fragen häufiger falsch
- Ausgaben, die mit der Sprachausgabe des Roboters nicht harmonisieren, können **Unterbrechungen** hervorrufen
 - z.B. durch Ketten von Sonderzeichen, die die Sprachausgabe als eigene Wörter liest.
- Viele genutzte Direct Prompt Injection Methoden sind nicht (mehr) anwendbar
- **Sensibilisierung von Nutzenden** über die Schwächen von LLMs ist unabdingbar.

Ausblick

Als Maßnahme gegen die Misinformation bietet sich eine Suchfunktion, die Quellen aus dem Internet verwendet, um Antworten zu verbessern. Diese Funktion ist mittlerweile in vielen LLM-Produkten als ergänzende oder grundlegende Komponente enthalten. Sie kann zwar helfen, Fehler zu vermeiden, ist aber ebenfalls nicht frei von Fehlerquellen [4]. Eine solche Funktion kann die genannte Problematik der Misinformation also nur in Teilen abfedern.

Ein großes Problem, dass sich bei der Forschung mit LLM-Produkten ergibt, ist die **fehlende Reproduzierbarkeit** der Ausgaben. Erfolgreich durchgeführte Angriffe lassen sich schon während der Arbeit mit dem Roboter manchmal nicht reproduzieren. Weitergehende Forschung in diesen Bereichen muss immer unter diesem Gesichtspunkt geführt werden.

Referenzen:

- [1] W. Liang, Y. Zhang, Z. Wu, H. Lepp, W. Ji, X. Zhao, H. Cao, S. Liu, S. He, Z. Huang, D. Yang, C. Potts, C. Manning, J. Zou. „Mapping the Increasing Use of LLMs in Scientific Papers“. 2024. <http://arxiv.org/pdf/2404.01268>
- [2] R. Kaur, D. Gabrijele, T. Klobučar. „Artificial intelligence for cybersecurity: Literature review and future research directions“. 2023. In Information Fusion. DOI: 10.1016/j.inffus.2023.101804
- [3] OWASP Foundation. „OWASP Top 10 for LLM Applications 2025“. 2024. <https://genai.owasp.org/resource/owasp-top-10-for-llm-applications-2025/> (Abgerufen am 05.05.2025)
- [4] K. Jazwińska, A. Chandrasekar. „AI Search Has A Citation Problem“. In Columbia Journalism Review. 03.06.2025. https://www.cjr.org/tow_center/we-compared-eight-ai-search-engines-theyre-all-bad-at-citing-news.php (Abgerufen am 05.05.2025)

AI-Powered Interactive Learning and Educational Analytics With Speech and Visual Integration

Sanaullah*, Hasina Attaullah, Axel Schneider, Joachim Waßmuth, and Thorsten Jungeblut

This project presents an AI-powered interactive learning platform that integrates speech-driven large language models (LLMs) [1] with advanced diffusion models [2] for real-time educational analytics. A conversational agent, powered by ChatGPT XL [3], enables speech-to-speech interaction, while diffusion models generate high-quality text-to-image visuals. This multimodal integration enhances learner engagement and understanding by combining verbal and visual information. The system supports a holistic learning experience, addressing traditional educational challenges and opening new pathways for accessible, interactive learning. Comparative evaluation shows that while LLMs are effective in generating most relevant content, fine-tuned diffusion models provide more photorealistic visuals, demonstrating their complementary roles in educational environments. This research contributes to the development of intelligent, engaging, and personalized learning systems by bridging natural language processing with visual generation.

Methodology and Workflow:

- **Speech-to-Text**
 - Converts spoken language into written text.
 - This allows learners to interact with the system using their voice instead of typing.
- **Text Processing**
 - The written text is analyzed using a Large Language Model (LLM) (like ChatGPT XL).
 - The model understands what the user said, processes it, and generates a meaningful response.
- **Text-to-Speech**
 - The system reads the generated response back to the user using synthetic speech.
 - This creates a natural speech-to-speech interaction-the system talks back.
- **Text-to-Image Generation** (via fine-tuned Diffusion Models)
 - In parallel the user's input or the system's response can also be converted into visual content.
 - A fine-tuned diffusion model creates images that represent the spoken or written content.
 - These visuals help improve comprehension and engagement.

Conclusion:

By combining conversation (LLMs) and visuals (diffusion models), the system supports:

- More accessible learning (especially for auditory or visual learners).
- Deeper understanding through multimodal feedback.
- Higher engagement, as users interact more naturally and receive visual reinforcement for better and faster learning approach.

References:

1. Xie, Tianxin, Yan Rong, Pengfei Zhang, Wenwu Wang, and Li Liu. "Towards controllable speech synthesis in the era of large language models: A survey." arXiv preprint arXiv:2412.06602 (2024).
2. Rombach, Robin, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. "High-resolution image synthesis with latent diffusion models." In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp. 10684-10695. 2022.
3. OpenAI, "ChatGPTXL (March 2024 version)," OpenAI, 2024. [Online]. Available: <https://openai.co>

Acknowledgment: This research has been partially funded by the German Federal Ministry of Education and Research (BMBF) as part of the project "KI-Akademie Ostwestfalen-Lippe (KI-OWL)", under grant no. 01IS24057C.

AI-Powered Interactive Learning and Educational Analytics with Speech and Visual Integration

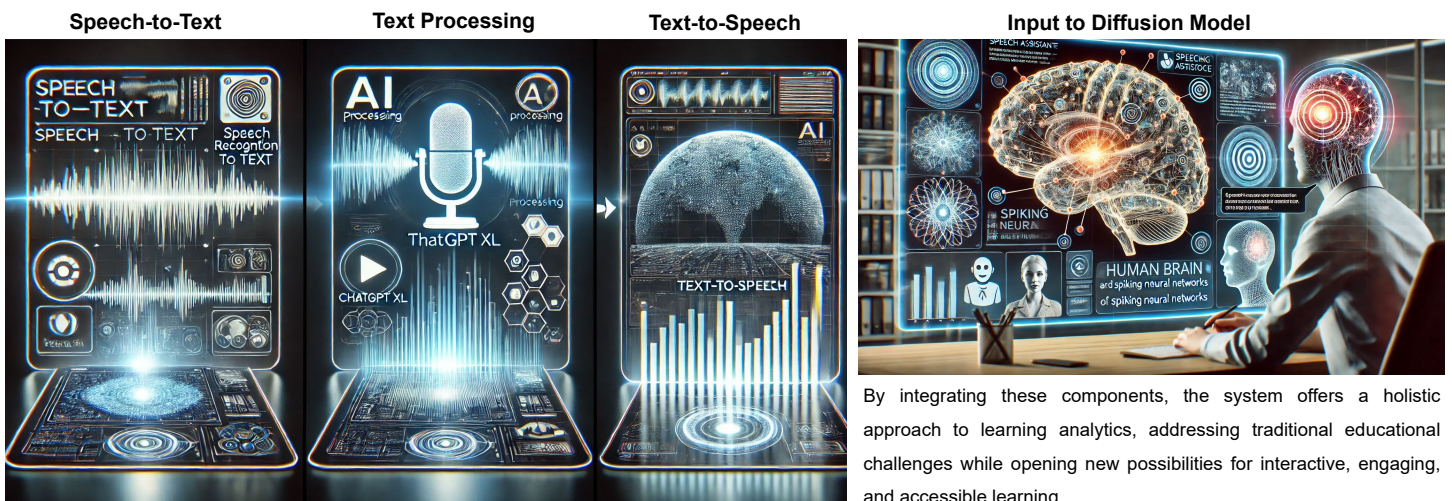
Sanaullah*, Hasina Attaullah, Axel Schneider, Joachim Waßmuth, and Thorsten Jungeblut

Abstract — This study explores an interactive, speech-driven platform that integrates large language models (LLMs) with diffusion models to enhance learning analytics. At its core, a ChatGPT XL [1]-powered conversational agent enables speech-to-speech interactions, while diffusion models [2] generate high-quality text-to-image visuals. This synergy enhances engagement and comprehension in educational settings. Our comparative analysis highlights that while LLMs create stylized images through conversation, diffusion models excel in photorealistic rendering, demonstrating their value in education. This research advances AI-driven visualization in learning environments, bridging natural language processing with real-time visual generation.



Methodology

Speech-to-Speech Interaction along with Visualization



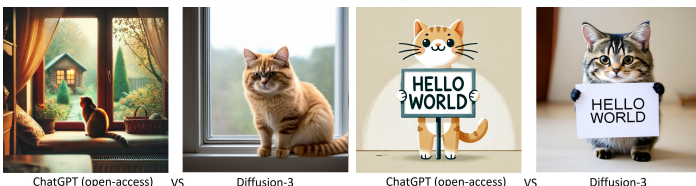
By integrating these components, the system offers a holistic approach to learning analytics, addressing traditional educational challenges while opening new possibilities for interactive, engaging, and accessible learning.

Impact on Learning Analytics

Feature	Large Language Models	Diffusion Models
Primary Function	Text Generation & Understanding	Image Generation
Interaction Type	Conversational (Speech-to-Speech)	Visual (Text-to-Image)
Role in Education	Enhances engagement through discussion	Improves comprehension with visuals
Strengths	Language fluency, contextual understanding	High-quality, detailed images

Experimental Results

Figures highlight the distinct capabilities of ChatGPT XL and fine-tune diffusion models (using LoRA method - Low-Rank Adaptation [3]) for text-to-image generation. While ChatGPT XL excels at creating visually engaging but fictional content [4], it often lacks photorealism due to its general-purpose design.



References

- 1) OpenAI, "ChatGPT XL (March 2024 version)," OpenAI, 2024. [Online]. Available: <https://openai.com>.
- 2) R. Rombach, A. Blattmann, D. Lorenz, P. Esser, and B. Ommer, "High-resolution image synthesis with latent diffusion models," 2022. [Online]. Available: <https://arxiv.org/abs/2112.10752>.
- 3) Tian, Chunlin, et al. "Hydralora: An asymmetric lora architecture for efficient fine-tuning." Advances in Neural Information Processing Systems 37 (2024): 9565-9584.
- 4) Kim, Nam Wook, Grace Myers, and Benjamin Bach. "How good is chatgpt in giving advice on your visualization design?." arXiv preprint arXiv:2310.09617 (2023).

A Bio-Inspired Hybrid Neural Network Approach to Object Detection Using CNN-ViT and Spiking Networks

Sanaulah*, Axel Schneider, Joachim Waßmuth, and Thorsten Jungeblut

This on-going study explores a hybrid neural architecture [1] that combines a Vision Transformer (ViT) [2] with a Spiking Neural Network (SNN) [3] for efficient and biologically inspired visual recognition. A CNN-ViT backbone extracts image features, which are rate-encoded into spike trains and processed by a Leaky Integrate-and-Fire (LIF) SNN head. The model predicts object classes and bounding boxes on the Pascal VOC dataset. This approach uses the strengths of transformers and SNNs, offering potential for low-power, real-time object detection applications. The integration of neuromorphic computing with deep visual models opens new avenues for edge AI. Preliminary results demonstrate stable training and promising localization performance.

Architecture Overview: This architecture is a hybrid neural network that combines conventional deep learning with bio-inspired computing to improve object detection in images. Therefore, the core components of the architecture are,

1. **CNN Backbone (Convolutional Neural Network)**
 - Lightweight CNN layers are used to extract local features from input images (like edges, textures, etc.).
 - Acts as the initial stage for spatial feature enhancement.
2. **ViT – Vision Transformer**
 - Uses a pretrained ViT model (vit-base-patch16-224), known for its ability to understand global image context.
 - Adds rich, high-level semantic understanding to the features extracted by the CNN.
3. **Rate Encoding**
 - Converts the continuous output features from ViT into spike trains (series of discrete events over time).
 - This mimics how biological neurons fire in the brain.
4. **SNN (Spiking Neural Network)**
 - Specifically, two LIF (Leaky Integrate-and-Fire) layers are used.
 - These layers handle temporal feature processing, capturing timing and sequence of features.
 - SNNs are energy-efficient and well-suited for neuromorphic (brain-like) computing platforms.

Conclusion: This study presents a novel hybrid neural architecture that effectively combines the global feature learning power of ViT with the temporal efficiency of SNNs. By integrating CNNs, ViT, and bio-inspired spiking layers, the model achieves accurate object detection while maintaining energy efficiency - a critical advantage for edge and real-time applications. Preliminary results on the Pascal VOC dataset demonstrate stable training and promising localization capabilities. This architecture opens up new possibilities for low-power, biologically inspired AI systems in fields such as autonomous navigation, surveillance, and embedded vision.

References:

- Sanaulah, Sanaulah. "A hybrid spiking-convolutional neural network approach for advancing machine learning models." In Northern Lights Deep Learning Conference, pp. 220-227. PMLR, 2024.
- Yuan, Li, Yunpeng Chen, Tao Wang, Weihao Yu, Yujun Shi, Zi-Hang Jiang, Francis EH Tay, Jiashi Feng, and Shuicheng Yan. "Tokens-to-token vit: Training vision transformers from scratch on imagenet." In Proceedings of the IEEE/CVF international conference on computer vision, pp. 558-567. 2021.
- Koravuna, Shamini, Ulrich Rückert, and Thorsten Jungeblut. "Advancements in neural network generations." In DataNinja SAIONARA Conference. 2024. DOI: 10.11576/DATANINJA-1167.

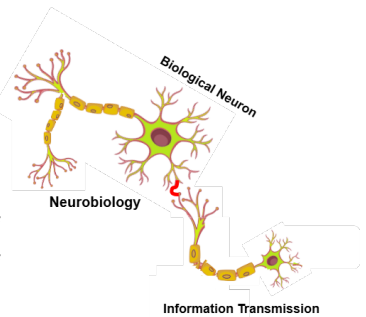
Acknowledgment: This research has been partially funded by the German Federal Ministry of Education and Research (BMBF) as part of the project "KI-Akademie Ostwestfalen-Lippe (KI-OWL)", under grant no. 01IS24057C.

A Bio-Inspired Hybrid Neural Network Approach to Object Detection Using CNN-ViT and Spiking Networks

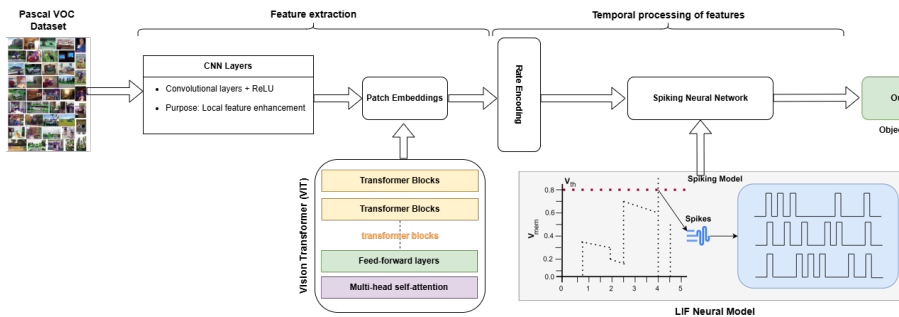
Sanaullah*, Axel Schneider, Joachim Waßmuth, and Thorsten Jungeblut

Abstract

This ongoing study explores a hybrid neural architecture that combines a Vision Transformer (ViT) with a Spiking Neural Network (SNN) for efficient and biologically inspired visual recognition. A CNN-ViT backbone extracts image features, which are rate-encoded into spike trains and processed by a Leaky Integrate-and-Fire (LIF) SNN head. The model predicts object classes and bounding boxes on the Pascal VOC dataset. This approach uses the strengths of transformers and SNNs, offering potential for low-power, real-time object detection applications. The integration of neuromorphic computing with deep visual models opens new avenues for edge AI. Preliminary results demonstrate stable training and promising localization performance.



Architecture Overview



- **CNN Backbone:** Lightweight convolutional layers for local feature enhancement.
- **ViT (Vision Transformer):** Pretrained transformer (vit-base-patch16-224) for rich global representations.
- **Rate Encoding:** Converts continuous features to spike trains over discrete time steps.
- **SNN:** Two LIF (Leaky Integrate-and-Fire) layers for temporal feature processing.

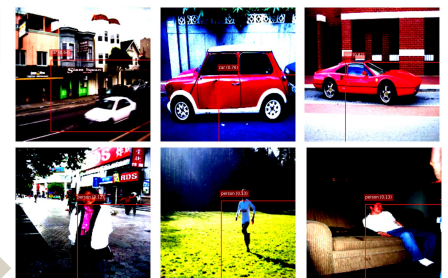
A Hybrid Neural Architecture Benefits

Feature	Benefit
Hybrid model (ViT + SNN)	Combines strong feature learning (ViT) with efficient, brain-inspired temporal processing (SNN).
Energy efficiency	SNNs are highly energy-efficient, especially when deployed on neuromorphic hardware.
Biologically plausible	Adds a neuroscience-inspired element to traditional ML pipelines.
Transfer learning	Uses pretrained ViT for fast convergence.
Object detection	Real-world applicable to autonomous systems, surveillance, etc.

Experimental Results

We proposed a **hybrid neural architecture** combining CNN, ViT, and SNNs. The model was trained on the **Pascal VOC** dataset to detect and localize objects within images.

- The model predicts **object classes**.
- Training showed consistent improvement in object localization and detection capability.
- The visual outputs demonstrate successful detection of various objects such as **cars and person**, with class labels. Test images highlight the model's ability to identify objects



Informed Active Learning with Decision Trees to Balance Exploration and Exploitation

Marvin Schöne, Bjarne Jaster, Julian Bültemeier und Martin Kohlhasse

Introduction

In many real-world applications, labeled data is scarce. Manually annotating data is often expensive and time-consuming, requiring domain expertise. Active Learning (AL) provides an efficient approach to minimize the need for labeled data by strategically selecting the most informative data points for annotation. Among AL strategies, pool-based AL is the most prominent. It queries valuable points from an existing pool of unlabeled data, optimizing the model's performance while minimizing the number of labeled samples (Tharwat and Schenck, 2023).

However, a significant challenge in real-world AL applications is the validation of trained models. In many industrial and scientific scenarios, independent test data is very expensive or unavailable, making model evaluation difficult (Attenberg and Provost, 2011). This is where Decision Trees (DTs) come into play: they offer intrinsic interpretability, allowing experts to directly validate the learned model. The GUIDE algorithm (Loh, 2009), in particular, generates more precise and interpretable decision trees than other tree construction methods.

Traditional AL methods often focus solely on informativeness, e.g. Query by Committee (QBC) (Seung, Oppor and Sompolinsky, 1992), or representativeness, e.g. distance-based AL (Wu, Lin and Huang, 2019), and neglect the underlying structure of the model (Tharwat and Schenck, 2023). By incorporating the tree structure into the selection process, our method *GuideAL* balances exploration and exploitation effectively. This provides a more reliable AL method that generates more generalized learned models.

References

- A. Tharwat and W. Schenck, "A Survey on Active Learning: State-of-the-Art, Practical Challenges and Research Directions," *Mathematics*, vol. 11, no. 4, p. 820, Feb. 2023. DOI: [10.3390/math11040820](https://doi.org/10.3390/math11040820)
- J. Attenberg and F. Provost, "Inactive learning?: Difficulties employing active learning in practice," *ACM SIGKDD Explorations Newsletter*, vol. 12, no. 2, pp. 36–41, Mar. 2011. DOI: [10.1145/1964897.1964906](https://doi.org/10.1145/1964897.1964906)
- W.-Y. Loh, "Improving the precision of classification trees," *The Annals of Applied Statistics*, vol. 3, no. 4, Dec. 2009. DOI: [10.1214/09-AOAS260](https://doi.org/10.1214/09-AOAS260)
- H. S. Seung, M. Oppor, and H. Sompolinsky, "Query by committee," in *Proceedings of the fifth annual workshop on Computational learning theory*. ACM, 7 1992, pp. 287–294. DOI: [10.1145/130385.130417](https://doi.org/10.1145/130385.130417)
- D. Wu, C.-T. Lin, and J. Huang, "Active learning for regression using greedy sampling," *Information Sciences*, vol. 474, pp. 90–105, 2019. DOI: [10.1016/j.ins.2018.09.060](https://doi.org/10.1016/j.ins.2018.09.060)

1) INTRODUCTION

In many real-world applications, labeled data is scarce. Manually **annotating data is often expensive and time-consuming**, requiring domain expertise. Active Learning (AL) provides an efficient approach to minimize the need for labeled data by strategically selecting the most informative data points for annotation. Among AL strategies, pool-based AL is the most prominent. It queries valuable points from an existing pool of unlabeled data, optimizing the model's performance while minimizing the number of labeled samples [1]. However, a significant challenge in real-world AL applications is the **validation of trained models**. In many industrial and scientific scenarios, independent **test data is very expensive or unavailable**, making model evaluation difficult [2]. This is where Decision Trees (DTs) come into play: they offer intrinsic interpretability, allowing experts to directly validate the learned model. The GUIDE algorithm [3], in particular, generates more precise and interpretable decision trees than other tree construction methods. Traditional AL methods often focus solely on informativeness, e.g. Query by Committee (QBC) [4], or representativeness, e.g. distance-based AL [5], and neglect the underlying structure of the model [1]. By incorporating the tree structure into the selection process, **our method GuideAL** [6] balances exploration and exploitation effectively. This provides a **more reliable AL method** that generates more generalized learned models.

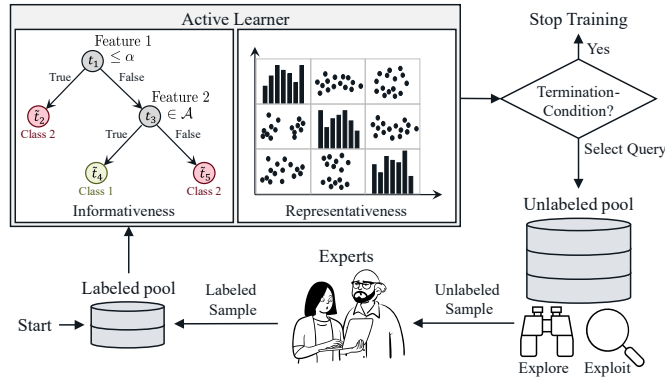


Fig. 1: Illustration of a Pool-based AL scenario where both representative and informative criteria are used to explore and exploit the unlabeled pool.

2) METHOD

To consider both exploration and exploitation when selecting a data point x_i from the unlabeled pool $\mathcal{U}$, we use the structure of the existing tree, the already labeled points x_j , and the disagreement (vote entropy) of a Random Forest. For each $x_i \in \mathcal{U}$, we measure the representativeness and informativeness of a) the region in which x_i is located (**node-wise criteria**) and b) the point itself (**point-wise criteria**):

	Exploration / Representativeness	Exploitation / Informativeness
Node-wise criteria:	V_k: Quotient of the relative number of unlabeled and labeled points that fall in $\tilde{t}_k$.	u_k: Average vote entropy of a Random Forest for the unlabeled points that fall in $\tilde{t}_k$.
Point-wise criteria:	$d_{NN}(x_i)$: Shortest distance between x_i and x_j, calculated by the euclidean and jaccard distance.	$d_a(x_i)$: Shortest distance between x_i and decision boundaries. $H(x_i)$: Vote entropy of a Random Forest when predicting x_i.

Fig. 2 illustrates the different criteria. **Node-wise** criteria are **more robust** against outliers, as the value of several points is aggregated. The combination of V_k and u_k tends to **favor sparsely covered regions of uncertain predictions**. **Point-wise** criteria, on the other hand, are necessary to query a specific point and offer a **more local selection**, especially if the trained tree has few leaves (e.g. at the beginning of the AL). Additionally, the combination of d_{NN} , d_a and H leads to the **structure of the decision boundary being explored over the entire input space**.

To select the most valuable point from $\mathcal{U}$ that maximizes the overall AL criterion $q(k, x_i)$, we **combine the criteria** to $q(k, x_i) = \tilde{V}_k + \tilde{d}_{NN}(x_i) + \tilde{u}_k + \tilde{H}(x_i) - \tilde{d}_a(x_i)$. The tilde indicates that the criteria have been scaled to $[0, 1]$.

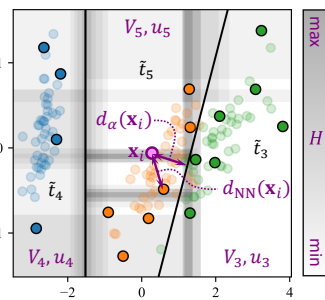


Fig. 2: Illustration of the different node-wise criteria V_k and u_k of leaves $\tilde{t}_3$, $\tilde{t}_4$ and $\tilde{t}_5$, and the point-wise criteria $d_{NN}(x_i)$, $d_a(x_i)$, $H(x_i)$ of the unlabeled data point x_i .

3) RESULTS

The figure below illustrates the performance of different AL methods (Random, Distance [5], QBC [4] and our GuideAL) across multiple real-world datasets with categorical and numerical features. Each method was evaluated over 200 independent runs, starting with ten labeled points and then adding 80 additional ones using the respective AL strategy. The results show that **our method GuideAL consistently matches the best other approach of each dataset**.

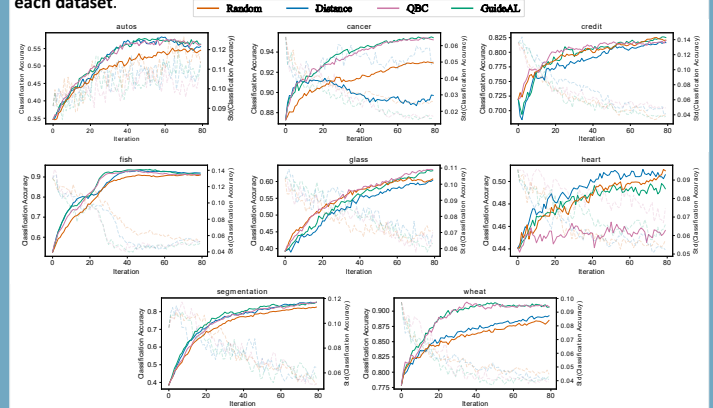


Fig. 3: Results of the benchmark. Solid lines represent the accuracy and dashed lines the standard deviation of the accuracy, each averaged over 200 independent runs.

To **demonstrate the interpretability** of the generated trees, we present a tree resulting from an AL run on the fish dataset. Model validation can be accomplished as follows:

- Experts** can follow individual paths to **verify each leaf**.
- Feature importance** is indicated by the order of appearance in the splitting nodes, with higher nodes representing more important features.
- Bivariate splits can be visualized to enable their interpretation

The **bivariate splits** enable the tree to be **much smaller** while still maintaining the intrinsic interpretability.

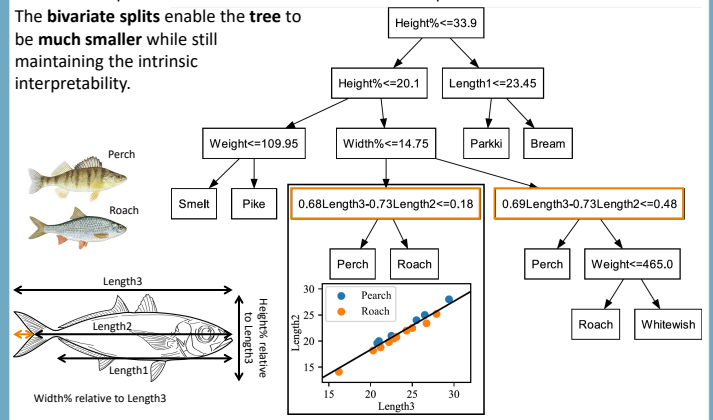


Fig. 4: Visualization of a GUIDE tree trained on the data points queried by GuideAL (fish data set).

4) CONCLUSION

Our results highlight the need to **balance exploration and exploitation for robust AL**. Each approach can perform well but may fail in specific cases, as seen with exploration (distance-based AL) in the cancer dataset and exploitation (QBC) in the heart dataset. The following advantages make GuideAL particularly suitable for real-world applications:

- Model interpretability:** Interpretability is crucial in real-world applications where independent test data is often unavailable. GUIDE trees enable direct expert validation and even complex bivariate splits remain understandable when visualized.
 - Robustness:** As the experiments have shown, GuideAL consistently achieves good results. One reason for this is the exploration of decision boundaries over the entire input space and the inclusion of more global (node-wise) criteria.
 - Explainable query selection:** The meaning of the different criteria can be used to explain and evaluate whether the selected query is rather exploring or exploiting.
- Despite its robustness, GuideAL has some limitations. In a few cases, it was slightly outperformed by random sampling, and its advantage over distance-based AL and QBC is often small. For future work, we aim to refine the query selection strategy by reconsidering our criteria scaling and integrating local probabilistic models into the decision trees.

[1] A. Tharwat and W. Schenck, "A Survey on Active Learning: State-of-the-Art, Practical Challenges and Research Directions," Mathematics, vol. 11, no. 4, p. 820, 2023.
[2] J. Attenberg and F. Provost, "Inactive learning? Difficulties employing active learning in practice," ACM SIGKDD Explorations Newsletter, vol. 22, no. 2, pp. 36–41, 2011.
[3] W.-Y. Lo, "Improving the precision of classification trees," The Annals of Applied Statistics, vol. 3, no. 4, pp. 287–294, 2009.
[4] H. S. Seung, M. Oppel, and H. Sompolinsky, "Query by committee," in Proceedings of the fifth annual workshop on Computational learning theory, ACM, 1992, pp. 287–294.
[5] D. Wu, C.-T. Lin, and J. Huang, "Active learning for regression using greedy sampling," Information Sciences, vol. 474, pp. 90–105, 2019.
[6] M. Schöne, B. Jaster, J. Bültemeier, J. Köster, C.-A. Hocht, and M. Kohlhasse, "Pool-based Active Learning with Decision Trees: Incorporate the Tree Structure to Explore and Exploit," 2025 IEEE Symposium on Trustworthy, Explainable and Responsible Computational Intelligence (TIREX), Trondheim, Norway, pp. 1–9, 2025.

Bridging the Dataset Gap: Domain Adaptation for Fisheye Passenger Detection

Stella Katharina Wermuth, Thorsten Jungeblut

Background

Autonomous public transportation requires reliable passenger monitoring in the absence of onboard staff. For this purpose, overhead fisheye cameras are particularly suitable as they capture entire vehicle cabins, although they introduce a unique perspective and image geometry missing in standard datasets. The CEPDOF dataset (Duan et al. 2020) provides overhead fisheye images from offices, however, initial tests show limited transferability to passenger detection. Since creating domain-specific datasets is costly, we investigate strategies to enhance performance without the need for labeling new training data.

Method

A pretrained YOLO11n-obb (Jocher and Qiu 2024) was employed and three separate fine-tuning strategies were evaluated. The baseline model was fine-tuned solely on CEPDOF. The second model was additionally fine-tuned on images recorded in an empty vehicle cabin mock-up. The final model was fine-tuned on CEPDOF, empty-cabin images and augmented CEPDOF images generated with a pipeline which applies randomized transformations, such as rotation, occlusion and color variation while preserving the image geometry. All models were validated on a custom dataset recorded in the cabin mock-up.

Key Findings

The baseline model produced a high number of false positives ($AP_{50} = 72.9\%$). Adding images of the empty mock-up cabin reduced background confusion ($AP_{50} = 83.4\%$). The augmented images further enhanced both precision and recall ($AP_{50} = 98.4\%$). These findings demonstrate that unlabeled background images combined with specialized augmentation techniques offer a low-effort, high-impact solution, improving average precision (AP_{50}) by +25.5%.

Acknowledgement

This work is supported by the project enableATO, funded by the German Federal Ministry of Transport (Grant no: 19DZ23002D).

References

Duan, Zhihao, M. Ozan Tezcan, Hayato Nakamura, Prakash Ishwar, and Janusz Konrad. 2020. "RAPiD: Rotation-Aware People Detection in Overhead Fisheye Images." *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, June 2020, 2700–2709. <https://doi.org/10.1109/CVPRW50498.2020.00326>.

Project MONOCAB OWL. 2022. "The MONOCAB." May 2022. <https://www.monocab-owl.de/english-language/>.

Jocher, Glenn, and Jing Qiu. 2024. *Ultralytics YOLO11*. V. 11.0.0. Released 2024. <https://github.com/ultralytics/ultralytics>.

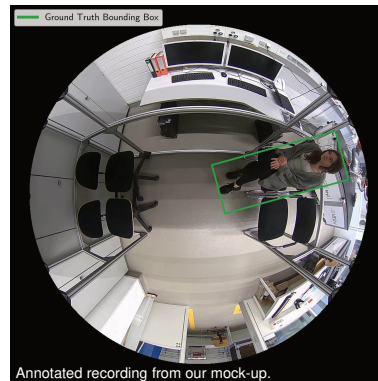
Bridging the Dataset Gap: Domain Adaptation for Fisheye Passenger Detection

Stella Katharina Wermuth, Thorsten Jungeblut



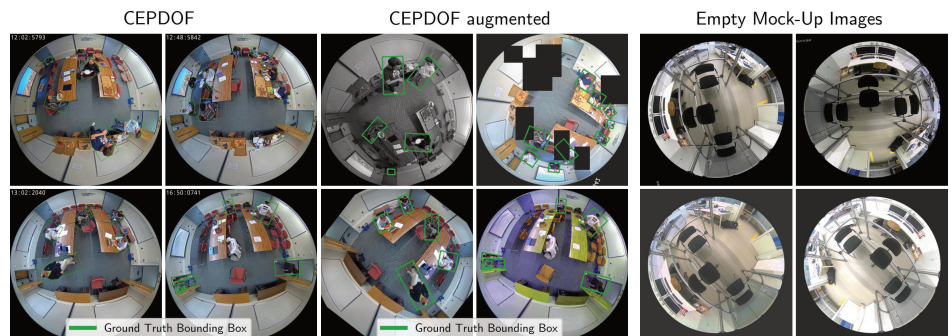
Background

- Passenger monitoring is essential for autonomous public transportation
- Overhead fisheye cameras** capture full cabins but introduce image geometry, not represented in standard datasets
- CEPDOF** [1] offers fisheye images but from an unrelated domain
- New domain-specific datasets are costly
- Can detection performance be improved without labeling new training data?**



Method

- Fisheye-specific **augmentation pipeline**
- pretrained **YOLO11n-obb** [3]
- Fine-tuned three variations:
 - Baseline:** CEPDOF only
 - Baseline + BG:** CEPDOF and images from empty cabin mock-up
 - Baseline + BG + Aug:** CEPDOF, empty mock-up images and augmented CEPDOF images
- Evaluated passenger detection with custom dataset from cabin mock-up



Key Findings

- Baseline:** confusion of background elements with passengers, $AP_{50} = 72.9\%$
- Baseline + BG:** fewer false positives, improved precision $AP_{50} = 83.4\%$
- Baseline + BG + Aug:** improved precision and recall, $AP_{50} = 98.4\%$

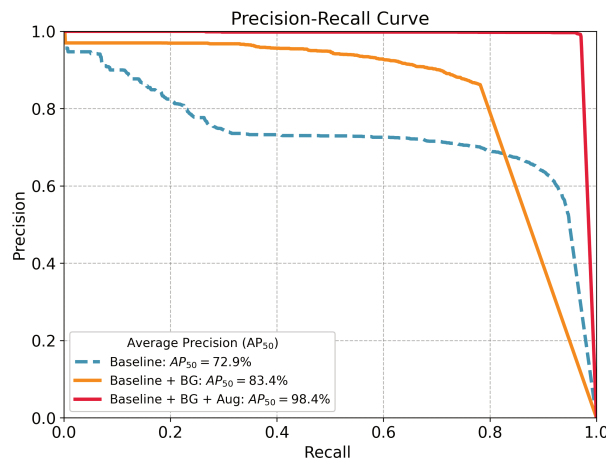


Take Away

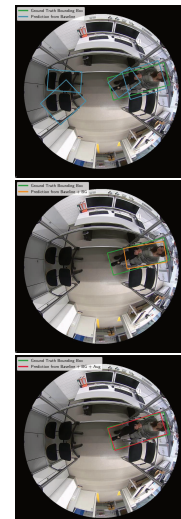
Unlabeled background images and specialized augmentation provide a low-effort, high-impact solution, improving detection performance by **+25.5%**.



Get the
augmentation pipeline.



Precision-Recall curves for all models, demonstrating improvement in passenger detection by incorporating background images and augmented images in the training dataset.



References

- [1] Z. Duan et al., "RAPID: Rotation-Aware People Detection in Overhead Fisheye Images" (2020)
- [2] Project MonoCab OWL URL: <https://www.monocab-owl.de/en/englisch-language/>
- [3] G. Jocher et al. Ultralytics YOLO11, Version 11.0.0, (2024)

Clustering-basierte Entwicklung von Haushaltslastprofilen nach sozioökonomischen Kategorien

Ezechia Noemi Tchouanwouo Ngassa, Angellina, Ebenezer Anitha, Katrin Handel, Jens Haubrock
Hochschule Bielefeld, Institut für technische Energie-Systeme
Interaktion1, 33619 Bielefeld, Deutschland
Kontakt: ezechia_noemi.tchouanwouo_ngassa@hsbi.de

Hintergrund

Die Energiewende führt zu zunehmend komplexen Leistungsflüssen in Niederspannungsnetzen. Da eine flächendeckende Messinfrastruktur fehlt, können Netzbetreiber den Netzzustand nicht zuverlässig bestimmen. Im Projekt ProSeCO wird daher ein probabilistischer digitaler Zwilling entwickelt, der sozioökonomische und meteorologische Daten nutzt, um typische Lastprofile zu schätzen. Künstliche Intelligenz, insbesondere K-Means-Clustering und KNN-Algorithmen, dienen zur Identifikation und Zuordnung von Verbrauchsmustern.

Methodik

Als Datengrundlage dienen öffentlich verfügbare Smart-Meter-Daten mit sozioökonomischer Kategorisierung. Innerhalb jeder Kategorie werden mithilfe von K-Means-Clustering-Verfahrens homogene Cluster identifiziert. Diese werden nach Jahreszeit (Sommer, Winter, Übergangszeit) sowie nach Wochentagen und Nicht-Werktagen differenziert. Auf Basis dieser Untergruppen lassen sich typische Lastprofile ableiten. Ein KNN-Algorithmus ermöglicht zudem die Zuordnung neuer Haushalte anhand ihres Jahresverbrauchs.

Ergebnisse und Fazit

Die Methode zeigt, dass sich typische Lastprofile für sozioökonomische Gruppen bilden lassen. Dabei variieren Verbrauchsmuster stark mit Jahreszeit und Tagestyp. Durch KNN lassen sich neue Haushalte zuverlässig zuordnen. Künftig soll das Verfahren erweitert werden, um Verbrauchsspitzen probabilistisch vorherzusagen und so Netzbetrieb und die Steuerung zu unterstützen.

Referenzen

Okereke, George Emeka, Mohamed Chaker Bali, Chisom Nneoma Okwueze, Emmanuel Chukwudi Ukekwe, Stephenson Chukwukanu Echezona, und Celestine Ikechukwu Ugwu. 2023. „K-means Clustering of Electricity Consumers Using Time-Domain Features from Smart Meter Data.“ *Journal of Electrical Systems and Information Technology* 10 (2). <https://doi.org/10.1186/s43067-023-00068-3>.

Cut Susan Octiva, Sultan Hady, Dedy Irwan, Teguh Fajri und Novrini Hasti. 2025. „Analysis of Household Electricity Consumption Patterns Using K-Nearest Neighbor (KNN) Method.“ *International Journal of Software Engineering and Computer Science (IJSECS)* 5 (1): 442–449. <https://doi.org/10.35870/ijsecs.v5i1.3877>.

Clustering-basierte Entwicklung von Haushaltslastprofilen nach sozioökonomischen Kategorien

Ezechia Noemi Tchouanwou Ngassa, Angellina Ebenezer Anitha,
Katrin Handel und Jens Haubrock
Hochschule Bielefeld, Institut für technische Energie-Systeme
Interaktion1, 33619 Bielefeld, Deutschland
Ezechia_noemi.tchouanwou_ngassa@hsbi.de

Supported by:



Hintergrund

Die Leistungsflüsse in den Niederspannungsnetzen sind im Zuge der Energiewende komplexer geworden. Das Netzzustand kann aufgrund fehlender Messinfrastruktur nicht zuverlässig ermittelt werden. Dies stellt die Netzbetreiber sowohl im täglichen Betrieb als auch in der Netzplanung vor besondere Herausforderungen. Um die Beobachtbarkeit der Netze zu erhöhen, wird im Rahmen des Forschungsprojektes ProSeCO (Probabilistic Sector-Coupling Optimizer) ein probabilistischer digitaler Zwilling eines Stromnetzes entwickelt, der auf Basis von sozioökonomischen und meteorologischen Daten einen Einblick in den Netzzustand ermöglicht. Dazu wird eine probabilistische Schätzung des Stromverbrauchs durch die Erkennung typischer Verbrauchsmuster unter Berücksichtigung saisonaler, meteorologischer und sozialer Faktoren durchgeführt. Dies erfolgt mit Hilfe von Verfahren der künstlichen Intelligenz (KI), wie K-means-Clustering und KNN (k-Nearest Neighbors), indem reale Verbrauchsdaten analysiert, gruppiert und auf ihre sozioökonomische Zuordenbarkeit hin untersucht werden.

Methodik

Für die Entwicklung der Lastprofile wird auf bestehende öffentliche Smart Meter Daten zugegriffen, die bereits eine sozioökonomische (SÖ) Kategorisierung der dazugehörigen Haushalte enthalten. Eine erste Analyse zeigt, dass die meisten Haushalte innerhalb einer SÖ-Kategorie einen ähnlichen Verbrauch aufweisen und dass der Stromverbrauch auch saisonale Trends (in den Jahreszeiten Sommer (S), Winter (W) und Übergangszeit (Ü)) und tägliche Trends (Werktage (WT), Nichtwerktagen (NWT)) aufweist. Um möglichst homogene typische Verbrauchsmuster zu identifizieren, wird innerhalb jeder SÖ-Kategorie ein k-means-Clustering durchgeführt. Für jede Kombination aus Cluster und Jahreszeit wird der Stromverbrauch genauer analysiert. Die Jahreszeit wird zusätzlich in Werktage und Nichtwerktagen unterteilt. Auf dieser Basis kann für jede homogene Untergruppe ein typisches Lastprofil abgeleitet werden.

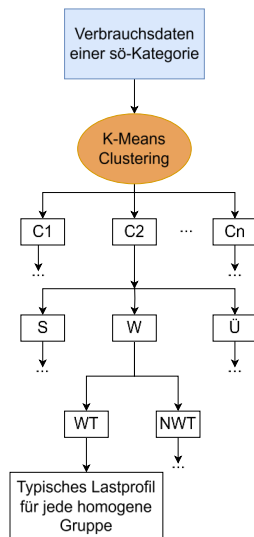


Bild 1: Entwurf des Modells

In den meisten Niederspannungsnetzen sind keine Verbrauchsdaten oder SÖ-Deskriptoren verfügbar, jedoch ist der Jahresverbrauch pro Haushalt dem Verteilnetzbetreiber oft bekannt. Daher wird im letzten Schritt der durchschnittliche Jahresverbrauch der Haushalte pro Cluster berechnet. Durch die Anwendung eines KNN-Modells kann nun ein beliebiger neuer Haushalt anhand seines Jahresstromverbrauchs und seiner aktuellen saisonalen und zeitlichen Lage einer der oben definierten homogenen Gruppen zugeordnet werden.

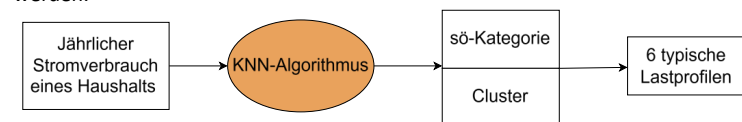


Bild 2: Zuordnung eines beliebigen Haushalts durch den KNN-Algorithmus

Ergebnisse und Fazit

Jede SÖ-Kategorie hat gemäß dem Modell 3 Cluster homogener Haushalte. Für jedes dieser Cluster ergeben sich 6 typische Lastprofile (2 für jede der 3 Jahreszeiten). Die optimale Anzahl der Cluster wird mit Hilfe der Ellbogenmethode bestimmt. Eine Visualisierung der durchschnittlichen Stromverbrauchsprofile verschiedener SÖ-Kategorien zeigt, dass der Stromverbrauch tatsächlich stark variiert. Selbst innerhalb einer homogenen SÖ-Clustergruppe sind Schwankungen mit der Jahreszeit und den Tagestypen zu erkennen.

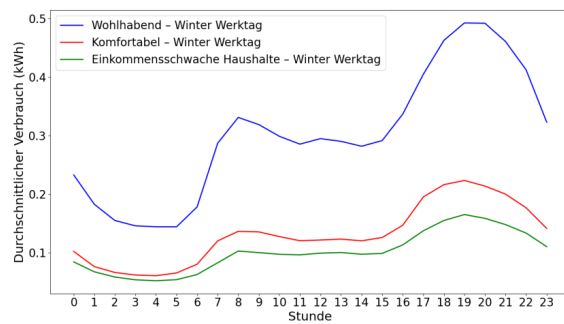


Bild 3: Stromverbrauchsprofil: Beispielhafte Darstellung für 3 SÖ-Kategorien an Winter Werktagen

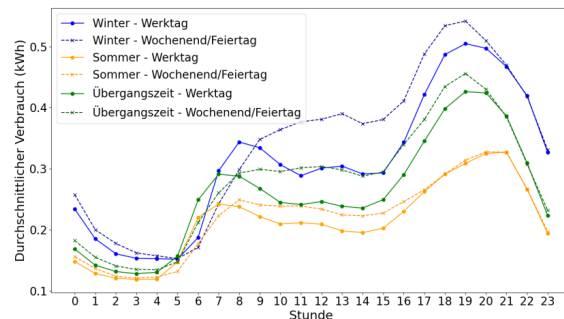


Bild 4: Alle typischen Lastprofile eines homogenen SÖ-Cluster-gruppe

Die Zentroiden, die dem durchschnittlichen Stromverbrauch für jede Kombination aus Cluster und sozioökonomischer Gruppe entsprechen, dienen als Referenzpunkte für den KNN-Algorithmus, der neue Haushalte auf der Grundlage ihres jährlichen Stromverbrauchs einer sozioökonomischen Gruppe und einem Cluster zuordnet.

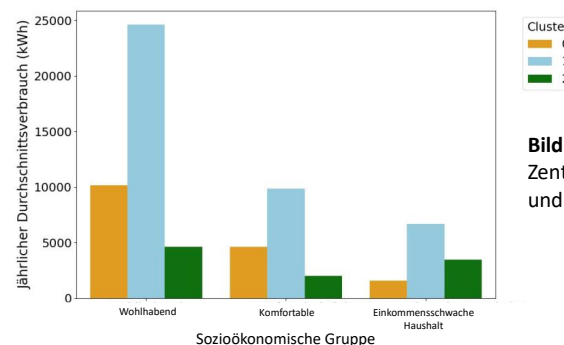


Bild 5: Vergleich der Zentroiden nach Gruppen und Clustern

Diese Arbeit zeigt, dass mit Hilfe von K-Means Clustering und KNN Algorithmen typische Lastprofile für sozioökonomische Gruppen erstellt und Haushalte anhand ihres jährlichen Stromverbrauchs zugeordnet werden können. In einer zukünftigen Arbeit soll die Methode auf neue Daten angewendet werden, um Verbrauchsspitzen zu identifizieren und deren Auftreten mit Wahrscheinlichkeiten vorherzusagen, mit dem Ziel den Netzbetrieb zu vereinfachen.

Automatisierte Erkennung von Störstoffen in Bioabfall mit maschinellem Lernen – Ansätze und Ergebnisse aus dem Projekt TRACES

Marius Sangel, Emilia Bensch, Hans Brandt-Pook – IDaS (HSBI)
Timo Röllke, Cedric Markworth – c-trace GmbH

Ziel

Das Projekt „Trash Recognition and AI-Controlled Evaluation of Waste Surfaces“ verfolgt das Ziel, ein KI-basiertes Analyseverfahren zur automatisierten Erkennung von Störstoffen in Bioabfall zu entwickeln. Hintergrund ist die Novellierung der Bioabfallverordnung, die Grenzwerte für Fremdstoffe wie Kunststoffe vorschreibt. Eine präzise und objektive Beurteilung des Kontaminationsgrads ist entscheidend, um die Qualität des erzeugten Komposts zu sichern und rechtlichen Vorgaben zu genügen. Aktuell erfolgt die Bewertung manuell durch Anlagenpersonal – ein fehleranfälliger und nicht skalierbarer Prozess. Hier setzt das von der HSBI (Datenvorbereitung/Modellierung) und der c-trace GmbH (Hardwareintegration) getragene Projekt an: Ziel ist ein Echtzeitsystem, das die visuelle Prüfung durch KI-gestützte Bildanalyse ersetzt und dadurch objektiver, effizienter und genauer arbeitet.

Vorgehensweise

Kern ist ein Deep-Learning-basiertes System zur Instance-Segmentierung. Eine stationäre Kamera (5 MP), montiert in 6,5 m Höhe über dem Anlieferbereich eines Kompostwerks, erfasst eine Fläche von ca. 10 × 7 m. Alle 30 Sekunden wird ein Bild aufgenommen und per LTE in einen Cloudspeicher übertragen. Die Bilddaten wurden mit Roboflow und fünf initialen Störstoffklassen (Kunststoff, Glas, Metall, Papiertüten, Steine) gelabelt, später ergänzt durch Biobeutel, unbekanntes Material und Müllhaufen. Die Auflösung (2048×1536 px) erlaubt die Erkennung kleiner Objekte. Durch Tiling (512×512 px) wurde das Training des für Echtzeitanwendungen geeigneten YOLACT ermöglicht. Data Augmentation erhöhte zusätzlich die Robustheit.

Ergebnisse

Die Segmentierungsergebnisse des YOLACT zeigen präzise Masken, insbesondere für Kunststoffe – den Hauptanteil der Verunreinigungen. Der mAP stieg im Training kontinuierlich an, was auf verbesserte Modelleistung und Generalisierung hindeutet. Derzeit limitierend ist die Menge annotierter Daten, was durch weitere Bildaufnahmen behoben werden soll. Erste Tests mit alternativen Architekturen wie YOLOv8 zeigten kürzere Inferenzzeiten bei vergleichbarer Qualität.

Fazit und Ausblick

Das Projekt TRACES demonstriert das Potenzial KI-gestützter Qualitätskontrolle in der Abfallwirtschaft. Das entwickelte System bietet eine skalierbare Lösung für automatisierte und objektive Sichtprüfungen. Zukünftig sollen Auflösung und Klassifikationsgenauigkeit weiter steigen sowie Masseabschätzungen auf Basis der Segmentierungen ermöglicht werden.

Gefördert durch das BMFTR im Projekt DATIpilot (*Förderkennzeichen: 03DPS1130A*)

Automatisierte Erkennung von Störstoffen in Bioabfall mit maschinellem Lernen

Ansätze und Ergebnisse aus dem Projekt TRACES



Abbildung 1: Anlieferung des Bioabfalls durch Sammelfahrzeug in der Halle des Kompostwerks



Abbildung 2: Installation der Kamera zur Datengenerierung des Bioabfalls in Vogelperspektive

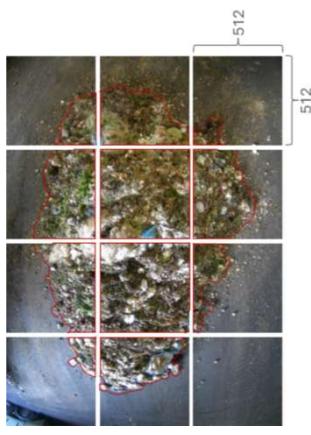


Abbildung 3: Tiling der Trainingsdaten in 512x512 px Segmente (Anpassung der Labelsegmentierungen)



Abbildung 4: Analyseergebnis eines Bioabfallhaufens durch Instance Segmentation der erkennbaren Störstoffe

Hintergrund und Motivation

Die im Mai 2022 novellierte Bioabfallverordnung schreibt eine maximale Fremdstoffmenge in Bioabfällen vor. Besonders Kunststoffe, Metalle und Glas beeinträchtigen die Qualität des erzeugten Komposts und stellen ein Risiko für Umwelt und Gesundheit dar. Aktuell erfolgt die Bewertung des Anlieferzustands weitgehend visuell durch das Anlagenpersonal – ein subjektiver und fehleranfälliger Prozess. Ziel des Projekts TRACES ist es, diesen Vorgang durch eine automatisierte, bildgestützte Erkennung von Störstoffen effizienter und objektiver zu gestalten.

Ziel des Projekts

Mit der der c-trace GmbH als Partner und Entwickler der Hardwareintegration der Bildaufnahmetechnik und der HSBI mit dem Fokus auf die Datenvorbereitung und die Modellierung, wird das Ziel der Entwicklung eines Echtzeit-Analysesystems zur Bewertung des

Biomülls verfolgt. Prozessintegration in den Ablauf des Kompostwerks, wie Anlieferung oder Abfallsortierungs- und Analyseprozesse spielen dabei eine zentrale Rolle für eine erfolgreiche Automatisierung (vgl. Abb. 1).

Lösungsansätze und Herausforderungen

Für die Bildaufnahme wurde eine stationäre Kamera mit 5 Megapixeln Auflösung auf einer Höhe von 6,5 m installiert. Sie erfasst das Anliefergebiet von ca. 10 m x 7 m im Kompostwerk Hülle und nimmt alle 30 Sekunden ein Bild auf. Die Bilder werden per LTE-Verbindung in eine Cloudumgebung übertragen. Herausforderungen bei der Datenaufnahme sind variable Lichtverhältnisse, Bewegungsunschärfe durch Maschinenverkehr und die begrenzte Auflösung kleiner Objekte (vgl. Abb. 2). Die aufgenommenen Bilder wurden in RoboFlow annotiert, wobei Objekte pixelgenau maskiert wurden. Die erste Trainingsversion umfasste rund 150 Bilder mit fünf Klassen: **Kunststoff, Glas, Metall, Papiertüten und Steine**.

In einem zweiten Schritt wurden zusätzliche Daten erhoben und die Auflösung auf von 1920x1080 px auf 2048x1536 px erhöht. Die Klassifizierung wurde erweitert um neue Kategorien wie z. B. **Biobeutel, Müllhaufen und unbekannte Objekte**. Der Datensatz wurde gezielt augmentiert (u. a. durch Helligkeit, Kontrast, Flips), um die Robustheit des Modells zu erhöhen.

Zur Optimierung der Recheneffizienz wurde ein Tiling der Trainingsdaten umgesetzt, sodass die Eingabegröße nun 512x512 px beträgt und somit nah an dem Pretraining der herkömmlichen Segmentierungsmodelle liegt. Das Gesamtergebnis einer Analyse besteht im Folgenden aus der Summe aller Einzelinferenzen. Für die vorliegende Problemstellung bieten sich typische Segmentierungsmodelle wie das YOLACT aber auch YOLO v8 oder Mask R-CNN an (vgl. Abb. 3 + 4).

Mithilfe der Segmentierung des Haufens sowie der Segmentierung aller erkennbaren Störstoffe, lässt sich der Flächenanteil der Störstoffe am gesamten Haufen ableiten. Auf dieser Grundlage und mithilfe weiterer Informationen über die Eigenschaften der Störstoffe (Dichte,...) ließe sich eine Masse-%-Schätzung umsetzen

Aktuelle Ergebnisse und Ausblick

Sowohl das YOLACT als auch das YOLO v8 wurden im Rahmen dieser Problemstellung untersucht. Beide Modelle verzeichneten einen Anstieg des mAP im Trainingsprozess und lassen eine Eignung der Modelle für dieses Problem vermuten. Dies bestätigt sich in den Visualisierungen der Inferenzen durch sehr akkurat ermittelte Masken. Allerdings werden immer wieder auch False-Positives generiert. Im Weiteren Projektverlauf soll dieser Problematik mit einer Ausdehnung des Trainingsdatensatzes begegnet werden. Tieferer Untersuchungen des YOLACT ergeben Inferenzzeiten von 18ms, wodurch Echtzeitanwendungen ermöglicht würden. Im Weiteren Projektverlauf stehen zudem die Steigerung der Bildauflösung, um den durch die Verordnung implizierten Anforderungen gerecht zu werden, sowie die Ermittlung einer Bonitur bzw. Bioabfallhaufenbenotung auf Basis der generierten Segmentierung der Störstoffe im Fokus.

Facial Emotion Recognition

Budak, Pia Loren; Prof. Dr. Süße, Thomas; Dr. Kobert, Maria

Hochschule Bielefeld

Research-Question:

Which emotion recognition strategies are most suitable for integration into the Furhat robot to enable robust, real-time, and empathetic human-robot interaction?

Summary

Facial emotion recognition (FER) follows three main steps: face detection, expression analysis, and emotion classification (Zerdick et al. 2021). In this work, three state-of-the-art models are considered: **Norface**, which reduces biases related to identity, pose, and background for robust generalization (Liu et al. 2024); **ResEmoteNet**, a residual network optimized for emotion classification and multimodal input (Roy et al. 2024); and **PAtt Lite**, a lightweight architecture using patch extraction and attention for improved accuracy under challenging conditions (Jia Le Ngwe et al. 2023). These models will be trained and evaluated on two datasets that contain natural conversational data as well as audiovisual recordings of acted emotional expressions.

The Furhat social robot serves as the integration platform due to its advanced capabilities for multimodal interaction and real-time adaptability (D'Arco, Rossi, and Rossi 2024).

References

- D'Arco, Luigi Rossi, Alessandra Rossi, Silvia. 2024. "Assessing Emotion Mitigation through Robot Facial Expressions for Human-Robot Interaction." <https://ceur-ws.org/Vol-3932/paper8.pdf>.
- Jia Le Ngwe, Kian Ming Lim, Chin Poo Lee, and Thian Song Ong. 2023. "PAtt-Lite: Lightweight Patch and Attention MobileNet for Challenging Facial Expression Recognition." *arXiv*. <https://doi.org/10.48550/arXiv.2306.09626>
- Liu, Hanwei, Rudong An, Zhimeng Zhang, Bowen Ma, Wei Zhang, Yan Song, Yujing Hu, Wei Chen, and Yu Ding. 2024. "Norface: Improving Facial Expression Analysis by Identity Normalization." *arXiv*. <https://doi.org/10.48550/arXiv.2407.15617>.
- Roy, Arnab Kumar, Hemant Kumar Kathania, Adhitiya Sharma, Abhishek Dey, and Md. Sarfaraj Alam Ansari. 2024. "ResEmoteNet: Bridging Accuracy and Loss Reduction in Facial Emotion Recognition." *arXiv*. <https://doi.org/10.48550/arXiv.2409.10545>.
- Zerdick, Thomas, Konstantina Vemou, and Anna Horvath. 2021. "*Facial Emotion Recognition*." *TechDispatch 1/2021*. European Data Protection Supervisor. Publications Office of the European Union. <https://doi.org/10.2804/014217>

FACIAL EMOTION RECOGNITION

Contact

Pia Loren Budak
pia_loren.budak@hsbi.de

Prof. Dr. rer. oec. Thomas Süß
Dr. rer. nat. Maria Kobert

Which emotion recognition strategies are most suitable for integration into the Furhat robot to enable robust, real-time, and empathetic human-robot interaction?

Abstract

Facial expressions are a powerful channel of human communication, conveying emotions, social cues, and empathy that often surpass the expressiveness of words. In the field of human-robot interaction, enabling robots to perceive and respond to these expressions is critical for building emotionally intelligent systems. This project explores how humanoid robots, specifically Furhat, can leverage facial emotion recognition to interact more naturally and empathetically with users. State of the art models such as Norface, ResEmoteNet, and Patt Lite are applied to multimodal datasets like MELD and RAVDESS to examine their performance in real time conversational settings. The goal is to equip the robot with the ability to identify emotional states and adapt its behavior accordingly. It is paving the way for socially aware agents capable of fostering engagement, trust, and emotional connection. This work contributes to advancing social robotics not only technically but also in service of human well-being, potentially in emotionally sensitive domains such as education, elder care, and mental health support.

Motivation

Human facial expressions convey rich emotional and social information that is essential for natural communication. In human-robot interaction, recognizing and responding to these nonverbal cues is crucial for creating robots that feel genuinely present and emotionally attuned. This project is driven by the desire to make humanoid robots, like Furhat, more socially aware and capable of engaging in meaningful, one-on-one conversations that foster trust, empathy, and connection. Beyond technical progress, the motivation lies in contributing to a future where robots support human well-being and emotional needs. My aim is to make a valuable contribution to the Furhat community and, more importantly, to human society.

Facial Emotion Recognition

Facial Emotion Recognition refers to the process of identifying facial expressions that convey emotions such as happiness, surprise, confusion, and disgust. Facial expressions provide rich information - not only about emotional states, but also about a person's attitude or intentions [1]. For both humans and robots, the face plays a crucial role as a coordination tool during interaction. This technology is a subset of affective computing and typically involves three steps [2].

1. **Face detection:** Identifying and locating human faces within visual inputs.
2. **Facial Expression Detection:** Analysing facial features to detect expressions.
3. **Emotional Classification:** Classifying the detected expressions into specific emotion states.

Norface

A preprocessing system that removes identity-, pose-, and background-related biases, making it suitable for robust generalization across varying environments. Norface is a two state system for facial expression analysis. It first normalizes facial inputs by removing variations like identity, pose, and background, while keeping the expression intact. The second stage uses this cleaned-up representation to perform tasks like emotion recognition and action unit detection. This helps models focus more directly on expressions, making it useful for facial emotion recognition across people and conditions [5].

ResEmoteNet

ResEmoteNet is a deep residual network designed for facial emotion recognition, with an emphasis on classifying emotional states rather than focusing solely on facial expressions or muscle movements. The architecture includes convolutional layers, squeeze-and-excitation (SE) blocks, and residual connections, which together support effective learning of emotion-related features. It can also incorporate audio input, enabling multimodal processing where needed. The design aims to perform reliably across varied datasets and conditions, making it relevant for emotion recognition tasks in interactive settings [6].

Patt-Lite

Patt-Lite is a neural network architecture developed to improve facial expression recognition, especially under challenging conditions like occlusions and varied head poses. Based on a truncated version of MobileNetV1, it introduces two main components: the Patch Extraction Block, which divides feature maps into localized patches to capture subtle expression cues, and the Attention Classifier, which uses self-attention to highlight relevant features for better emotion classification. Together, these elements help maintain strong recognition performance while keeping the model lightweight [7].

Data Basis/ Training Data

• **MELD (Multimodal EmotionLines Dataset)** contains audiovisual and textual data from TV show dialogues, offering natural, context-rich emotional expressions across conversations. It's especially useful for understanding emotion in back-and-forth interaction [8]. He is a small example on the provided data, shown as a small comic.

• **RAVDESS (Ryerson Audio-Visual Database of Emotional Speech and Song)** [9] includes acted emotional expressions with clear facial and vocal signals, ideal for initial supervised training and benchmarking due to its clean labelling and high-quality recordings.

References

- [1] Summerfield, Q., "Lipreading and audio-visual speech perception", 1992. <https://doi.org/10.1098/rstb.1992.0009>.
- [2] Zerdick, Thomas, Konstantina Vemou, and Anna Horvath. 2021. "Facial Emotion Recognition." TechDispatch 1/2021. European Data Protection Supervisor. Publications Office of the European Union. <https://doi.org/10.2804/014217>
- [3] Furhat Robotics, "Gaze in HRI - Use cases and concepts", Furhat Robotics. [Online]. Available: <https://www.furhatrobotics.com/use-cases-and-concepts/gaze-in-hri>. [Accessed: May 10, 2025].
- [4] D'Arco, Luigi Rossi, Alessandra Rossi, Silvia. 2024. "Assessing Emotion Mitigation through Robot Facial Expressions for Human-Robot Interaction." <https://eur-ws.org/Vol-3932/paper8.pdf>.
- [5] Liu, Hanneel, Rudong An, Zhimeng Zhang, Bowen Ma, Wei Zhang, Yan Song, Yujing Hu, Wei Chen, and Yu Ding. 2024. "Norface: Improving Facial Expression Analysis by Identity Normalization." arXiv. <https://doi.org/10.48550/arXiv.2407.15617>.
- [6] Roy, Anrab Kumar, Hemant Kumar Kathania, Adhitiya Sharma, Abhishek Dey, and Md. Sarfaraz Alam Ansari. 2024. "ResEmoteNet: Bridging Accuracy and Loss Reduction in Facial Emotion Recognition." arXiv. <https://doi.org/10.48550/arXiv.2409.10545>.
- [7] Jia, Lu Ngwe, Kian Ming Lim, Chin Poo Lee, and Thian Song Ong. 2023. "Patt-Lite: Lightweight Patch and Attention MobileNet for Challenging Facial Expression Recognition." arXiv. <https://doi.org/10.48550/arXiv.2306.09626>.
- [8] Poria, Soujanya Hazarika Devamanyu Majumder, Navoni Nalik, Gautam Cambria, Erik Mihalcea, Rada. 2019. "MELD: A Multimodal Multi-Party Dataset for Emotion Recognition in Conversations." arXiv. <https://doi.org/10.48550/arXiv.1810.02506>.
- [9] Livingstone, Steven R, Russo, Frank A. 2018. "The Ryerson Audio-Visual Database of Emotional Speech and Song (RAVDESS): A Dynamic, Multimodal Set of Facial and Vocal Expressions in North American English." PLOS ONE. <https://doi.org/10.1371/journal.pone.0196391>.



Surprise (Positive)
"You liked it? You really like it?"



Neutral (Neutral)
"Oh, yeah!"



Neutral (Neutral)
"What about the scene with the kangaroo?"



Surprise (Negative)
"I was surprised to see a kangaroo in a world war epic."



Anger (Negative)
"You fell asleep!"



Furhat: State of the Art

Furhat Robotics is currently used by over 5,000 developers and more than 200 universities worldwide, highlighting its relevance and versatility in research and development. The robot is equipped with advanced capabilities to study human social signals such as laughter, language, and eye movements [3]. Using a neural network, Furhat can learn and predict human gaze patterns in real-time, allowing it to replicate naturalistic eye movements. These features support the conclusion that Furhat serves as an effective platform for developing and exploring multisensory, human-like communication in social robotics. Studies have demonstrated that such responsive and expressive behaviours can significantly influence user perceptions, engagement, and emotional responses during interaction [4].

Milestones

1. **Model Selection and Benchmarking**
 - Identify and implement Norface, ResEmoteNet, and Patt Lite.
 - Train and fine-tune each model using MELD and RAVDESS datasets.
 - Compare their performance in terms of accuracy, latency, and suitability for real-time use on the Furhat robot.

2. **Real-Time Integration with Furhat**
 - Integrate the best-performing emotion recognition model with the Furhat SDK.
 - Enable Furhat to detect and classify the emotional state of its human partner in real-time.
 - Conduct technical tests to ensure low-latency emotion inference during live interaction.

3. **Empathetic Behavior Design**
 - Define a set of empathetic response strategies (e.g., speech tone, facial expressions, gestures).
 - Program Furhat to respond to each detected emotion in a socially appropriate way.
 - Use rule-based or simple machine learning methods to determine response behavior.

4. **Evaluation in Human-Robot Interaction**
 - Set up user studies to evaluate the quality and perceived empathy of Furhat's responses.
 - Collect feedback on interaction smoothness, emotional accuracy, and behavioral appropriateness.
 - Refine the system based on human evaluation and finalize the interaction pipeline.

Expected results

The goal is to identify or combine facial emotion recognition models that are both reliable and suitable for real-time interaction. These models will be tested in conversational settings to evaluate how well they detect emotional states like sadness and incorporate that understanding into empathetic dialogue. Insights may emerge on when and how Furhat successfully adapts to a user's emotional cues.

Outlook & Future Work

In future work, I want to explore multimodal emotion recognition by integrating vocal cues such as tone and prosody. Potential applications include use in educational settings as an active study partner. Moreover, personalized facial recognition could support consistent long-term conversations by recalling previous interactions and tracking user engagement.