

TAHAI – Prototypische Evaluation (Arbeitspaket 6)

Forschungsprojekt TAHAI - TrustAdHocAI

Sandro Hartenstein | 20.5.2025 | HWR-Berlin

Agenda



- Analyse von Mediationstranskriptionen
- videobasierte Vandalismuserkennung an Bahnhöfen
- Fazit



Analyse von Mediationstranskriptionen

Analyse von Mediationstranskriptionen

Anwendungsfall



- Analyse von Mediationstranskriptionen mit KI-WebAPIs
 - Ziel: Analyse der Arbeitsweise von Mediatoren
 - Besondere Anforderungen
 - Schutz der personenbezogenen Daten der Medianten und Mediatoren
 - Beibehaltung der Kerninhalte und Sachverhalte



Analyse von Mediationstranskriptionen

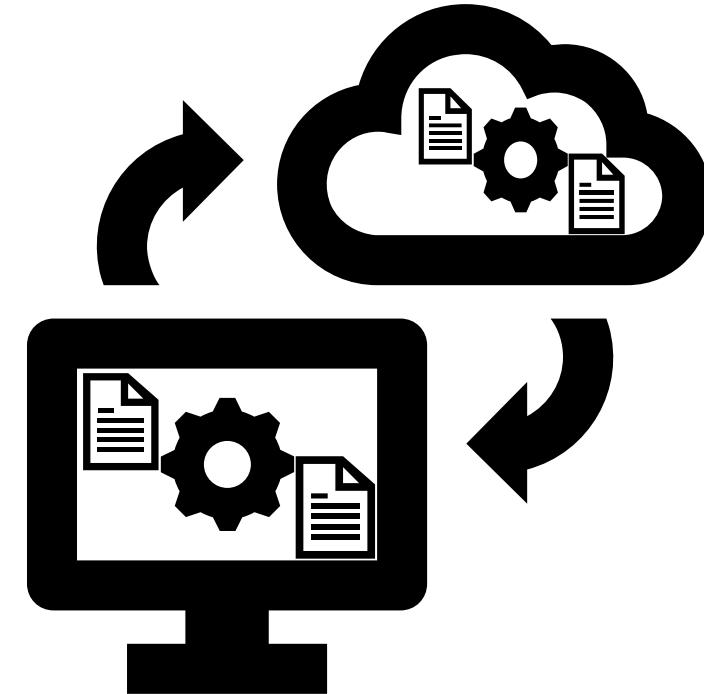
Technisches Management von Quelldaten



Analyse von Mediationstranskriptionen

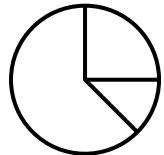
Anonymisieren von Transkriptionen

- Datenvorbereitung
 - automatischen Anonymisierung von Text
 - Beibehaltung der Kerninhalte und Sachverhalte
- Methodik
 - Praktischer Vergleich unterschiedlicher Ansätze
- Ergebnis
 - MS Presidio Framework am besten geeignet



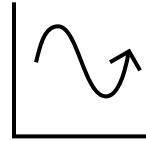
Analyse von Mediationstranskriptionen

Methodik



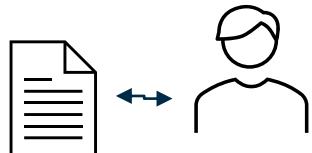
Statistische Analyse via RegEx

- Redeanteile



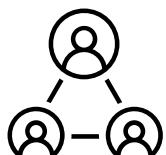
Semantische Analyse via NLP

- Emotionen im Zeitverlauf -> Vergleich unterschiedlicher KI Modelle



Direkte Befragung der Transkripte via RAG (retrieval-augmented generation)

- Konkrete Fragen nach Mediationsmethoden
- Hinweise für besondere Phasen



Indirekte Befragung der Transkripte via Agentic RAG

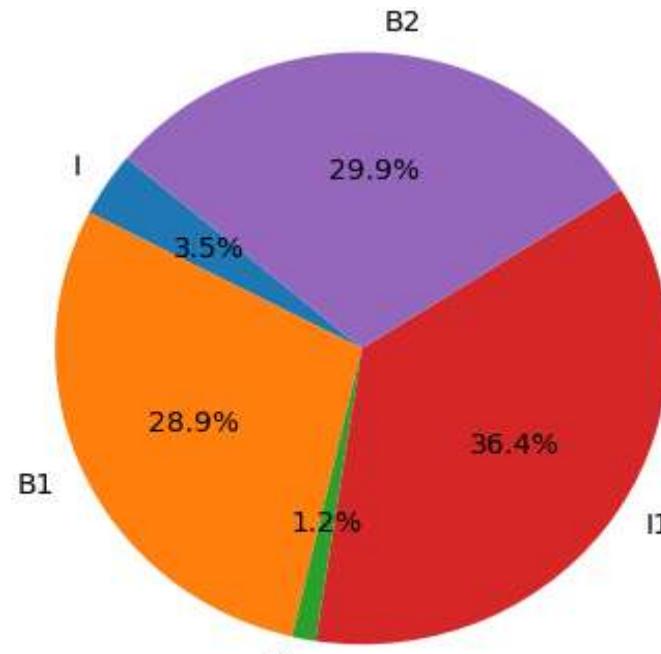
- Automatisierte Beurteilung der Mediationsleistung mithilfe von Rollenbeschreibungen und Zielebestimmungen, sowie bereitstellen von Tools

Analyse von Mediationstranskriptionen

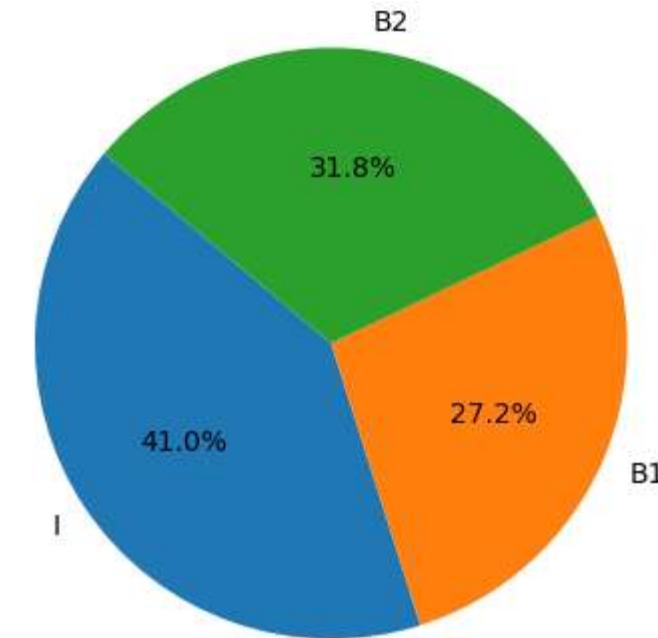
Ergebnisse | Statistische Analyse via RegEx



pattern = r'^(\w+):.*?#(\d{2}):(\d{2}):(\d{2})-(\d{1,2})#\$'



Redeanteile Fall 5

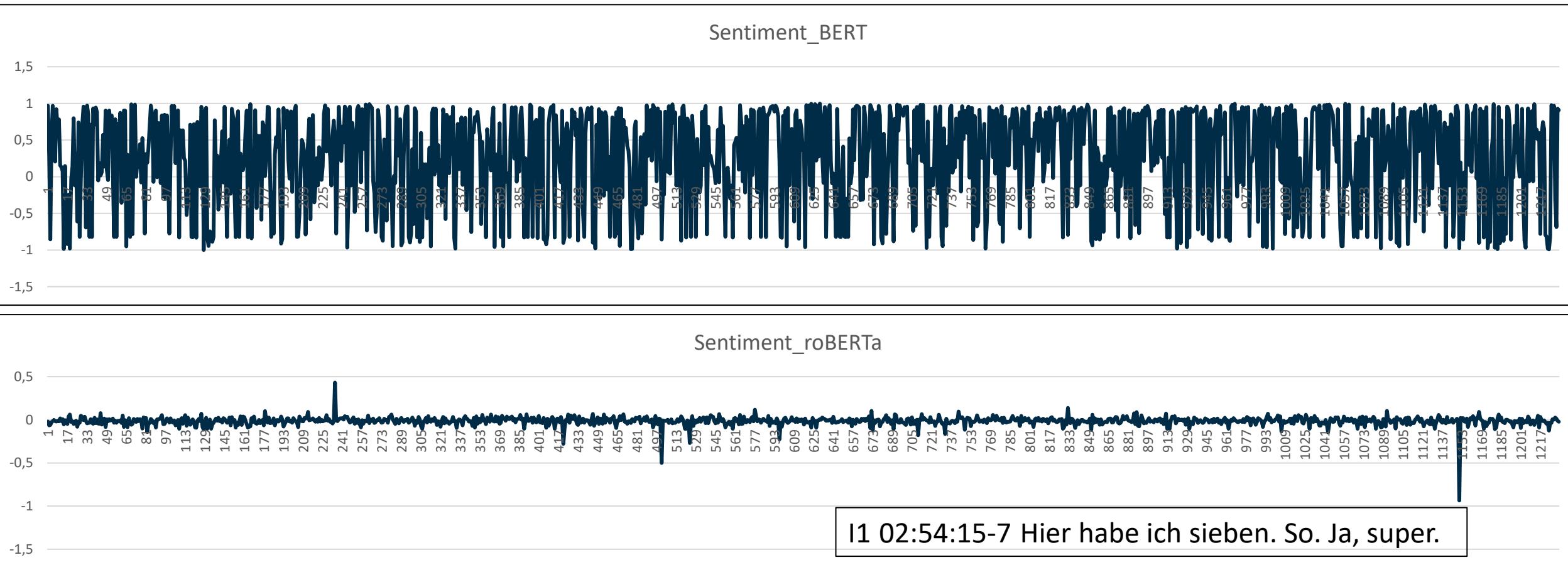


Redeanteile Fall 8

Analyse von Mediationstranskriptionen

Ergebnisse | Semantische Analyse via NLP

Fall 5 Session 3



Analyse von Mediationstranskriptionen

Ergebnisse | Direkte Befragung via RAG (anythingLLM)



The screenshot shows a conversational AI interface. On the left, a sidebar lists various cases: medi, Buch, Trust, Fall8, Fall_5 (selected), default, Sandro Validation, Optimierter Promt, New Thread, Fall-5_Sitzung-1, and Fall-5_Sitzg-2. The main area displays a message from a blue circular icon: "Was sind die Streitpunkte". Below it, a yellow circular icon provides an analysis: "Basierend auf den gegebenen Kontexten lassen sich folgende Streitpunkte identifizieren: 1. Betreuung vs. Freiheit in der Entscheidung, wer die Betreuung übernimmt (Context 0): Hier geht es um die Präferenz, wer die Betreuung einer bestimmten Person übernehmen soll, und die Abwägung dieser Entscheidung gegenüber dem Wunsch oder der Freiheit, diese Betreuung selbst zu übernehmen. 2. Klärung finanzieller Aspekte (Context 1 & 2): In diesen Kontexten wird ein Streitpunkt über die Klärung oder Abklärung finanzieller Details angesprochen, speziell um den Anruf bei einer bestimmten Person namens Roman, um den genauen Betrag zu erfragen, der für eine bestimmte Sache oder Leistung relevant ist. 3. Bedeutung von Erklärungen für das Verständnis von Reaktionen (Context 3): Der Streitpunkt hier dreht sich um die Notwendigkeit, Erklärungen für bestimmte Reaktionen oder Verhaltensweisen zu liefern. Eine Partei fühlt sich missverstanden ohne die Bereitstellung von Kontext oder Erklärung für ihr Verhalten, während die andere Partei der Meinung ist, keine Erklärungen zu benötigen. Diese Streitpunkte reflektieren unterschiedliche Bedürfnisse und Perspektiven der Beteiligten, die während der Mediation adressiert und möglichst aufgelöst werden sollen." At the bottom, there is a "Send a message" input field with a blue arrow icon.

Analyse von Mediationstranskriptionen

Ergebnisse | Indirekte Befragung via Agentic RAG



#####

In dem vorliegenden Mediationstranskript KA88 werden verschiedene **Mediationsmethoden** angewendet und interessante Wendepunkte erreicht, die beispielhaft für die Arbeitsweise des Mediators stehen.

Zu Beginn der Mediation, markiert durch die Begrüßung "Hallo" (B1, #00:00:20-0#) und "Guten Tag" (I, #00:00:21-3#), etabliert der Mediator eine höfliche und formelle Atmosphäre. Dies dient dazu, ein professionelles Umfeld zu schaffen, das für vertrauensvolle Gespräche essentiell ist. Die einführenden Worte und die kurze Pause vorher (#00:00:08-2# bis #00:00:18-3#) helfen, die Teilnehmer auf die Sitzung einzustimmen.

Ein **Wendepunkt** im Transkript ist in der Aussage von I2 zu erkennen: "Ich hätte sie auch reinlassen können" (#00:00:19-3#). Diese Bemerkung könnte auf eine vorherige Unsicherheit oder ein Missverständnis hinweisen, welches der Mediator aufgreifen muss, um Klarheit und Verständnis unter den Parteien zu fördern.

Weiterhin zeigt sich in der Äußerung "Mediation und das ist ein/ nicht, hm, hm (nachdenkend) meine Befindlichkeiten irgendwie zu ähm diskutieren" (#00:09:22-9#), wie der Mediator den Fokus der Mediation steuert. Hier leitet er von persönlichen Befindlichkeiten zu sachlichen Diskussionen über, was wesentlich für die Lösungsfindung ist.

Ein weiterer **entscheidender Moment** ist die Erörterung der Möglichkeit, "in der Mediation rechtsverbindliche Verträge zu machen, wie Sie das gedenken aufzuteilen" (#02:21:38-0#). Hier zeigt sich, dass der Mediator darauf abzielt, praktische und rechtlich haltbare Lösungen zu erarbeiten, die von den gerichtlichen Entscheidungen abweichen können. Dies ist ein Beispiel dafür, wie in der Mediation kreative und individuell angepasste Lösungen gefunden werden können, die den Bedürfnissen aller Parteien entsprechen.

Zusammengefasst nutzt der Mediator in diesem Transkript gezielt Pausen, formelle Begrüßungen und die Lenkung von persönlichen zu sachlichen Themen, um eine effektive Mediation zu führen. Diese Methoden und Wendepunkte sind entscheidend für den Fortschritt und Erfolg der Mediationssitzung.

Analyse von Mediationstranskriptionen

Erkenntnisse



Zentrale Erkenntnisse der KI-gestützten Mediationsanalyse

- **KI ermöglicht** systematische Auswertung anonymisierter Mediations-Transkripte.
- **Redeanteile, Stimmungen und Methoden** wurden automatisiert erkannt.
- **RAG und Agentic RAG** lieferten kontextsensitive Analysen und Interpretationen.
- **Multi-Agenten-Systeme** unterstützen komplexe Aufgaben wie Sequenzanalyse.
- **Fazit:** KI unterstützt die Forschung und Professionalisierung von Mediation



Vandalismuserkennung

videobasierte Vandalismuserkennung an Bahnhöfen

Vandalismuserkennung

- Ziel:
 - Erkennen von Vandalismus (insbesondere Graffitysprayer)
- Herausforderungen:
 - Nutzung vorhandener Videoanlagen (100+ Anzahl Kameras)
 - Echtzeiterkennung
 - Geringe False Positives notwendig



Bild von [Gerd Altmann](#) auf [Pixabay](#)

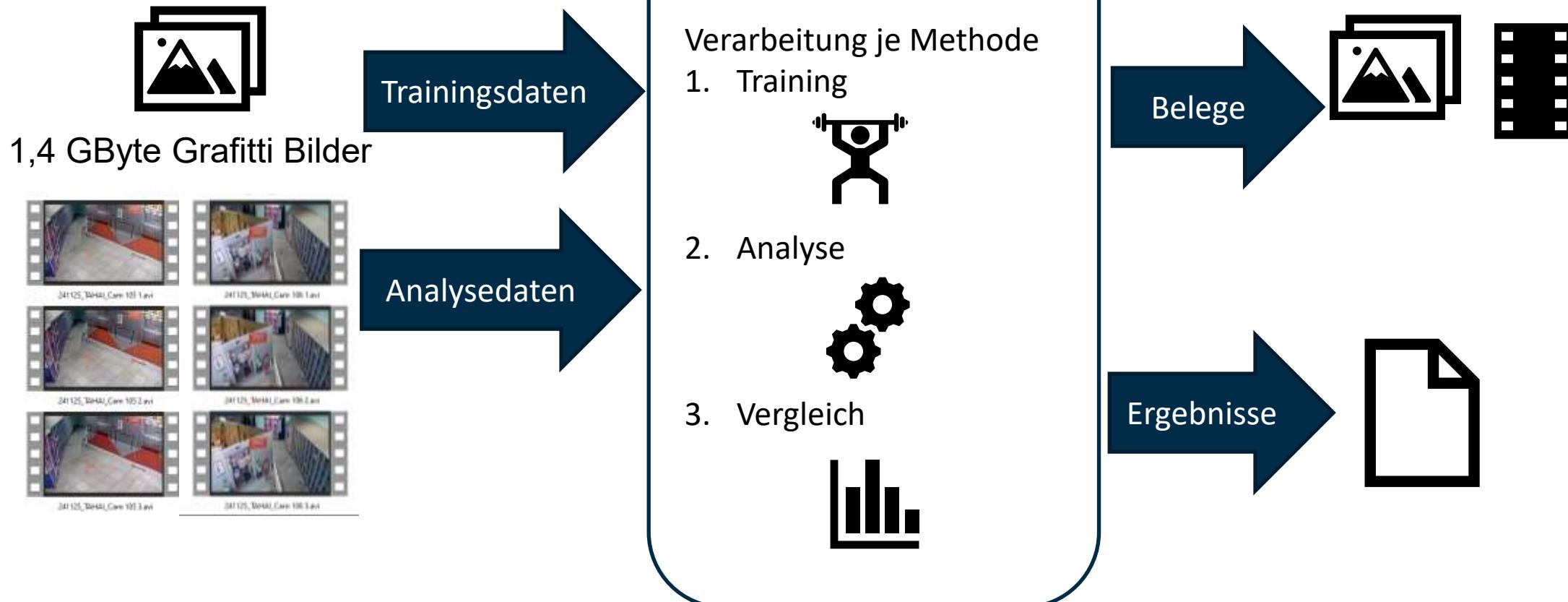
Vandalismuserkennung

Medodik



Methode	Prinzip	Vorteile	Nachteile
OpenCV	Klassische Algorithmen	Einfach, geringe Rechenanforderungen	Ungenau, anfällig für Fehlalarme
YOLO	Echtzeit-Objekterkennung	Schnell, genau	Hohe Hardware-Anforderungen
CRNN	CNNs + RNNs	Lernt räumliche und zeitliche Informationen	Rechenintensiv, große Datensätze erforderlich
Conv-LSTM-SVM	CNNs + LSTMs + SVMs	Erkennt komplexe Verhaltensmuster	Rechenintensiv, erfordert Parametereinstellung

Vandalismuserkennung Implementierung



Vandalismuserkennung

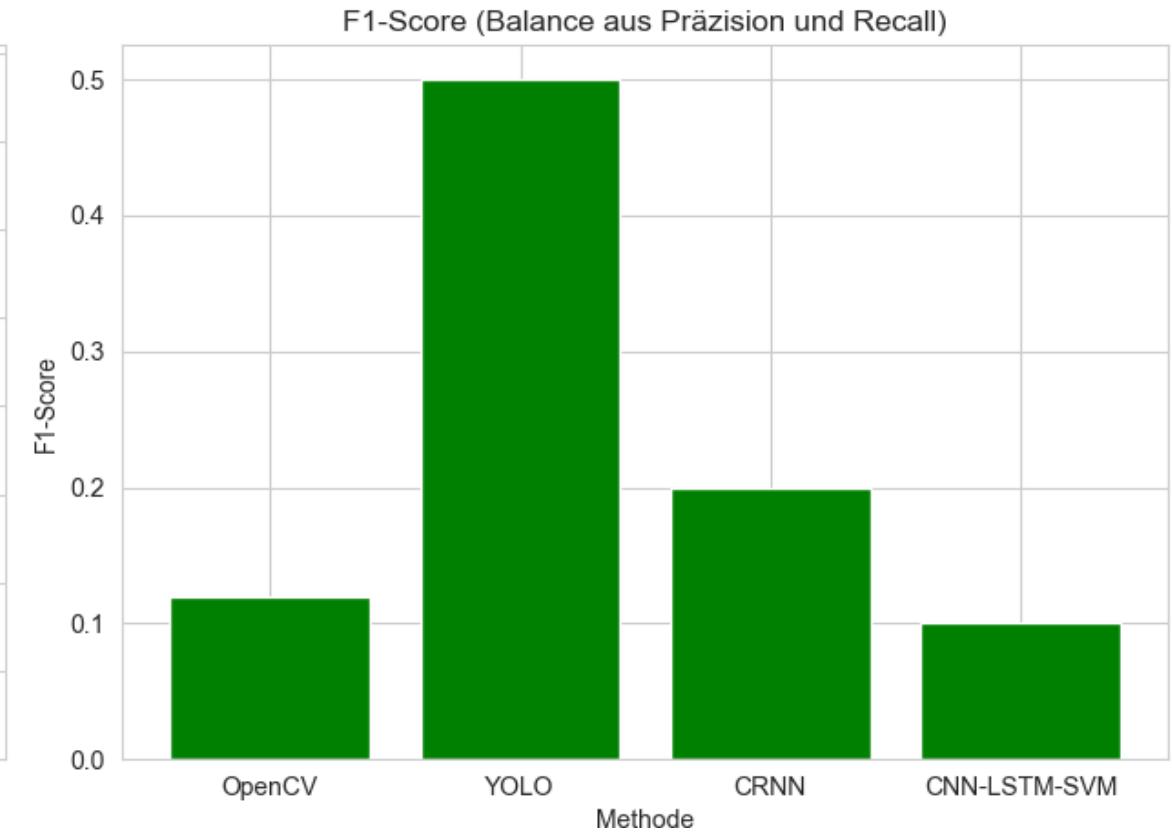
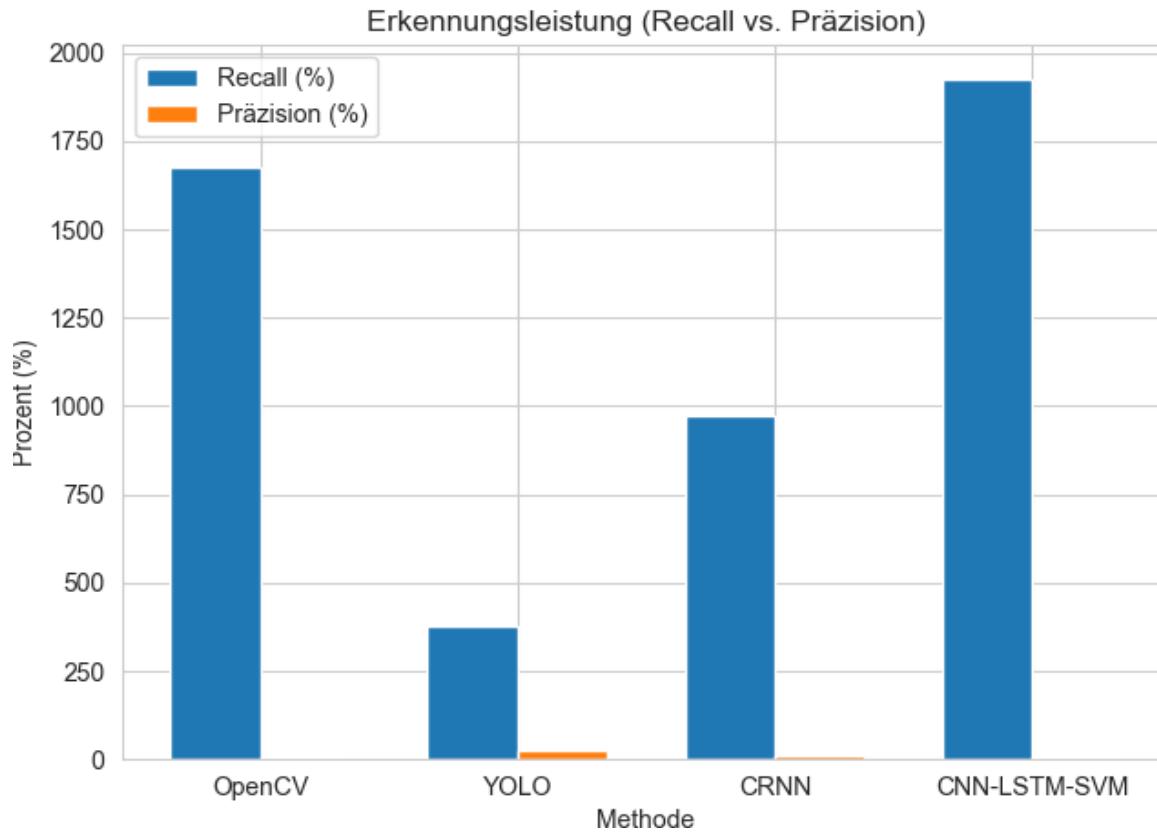
Ergebnisse



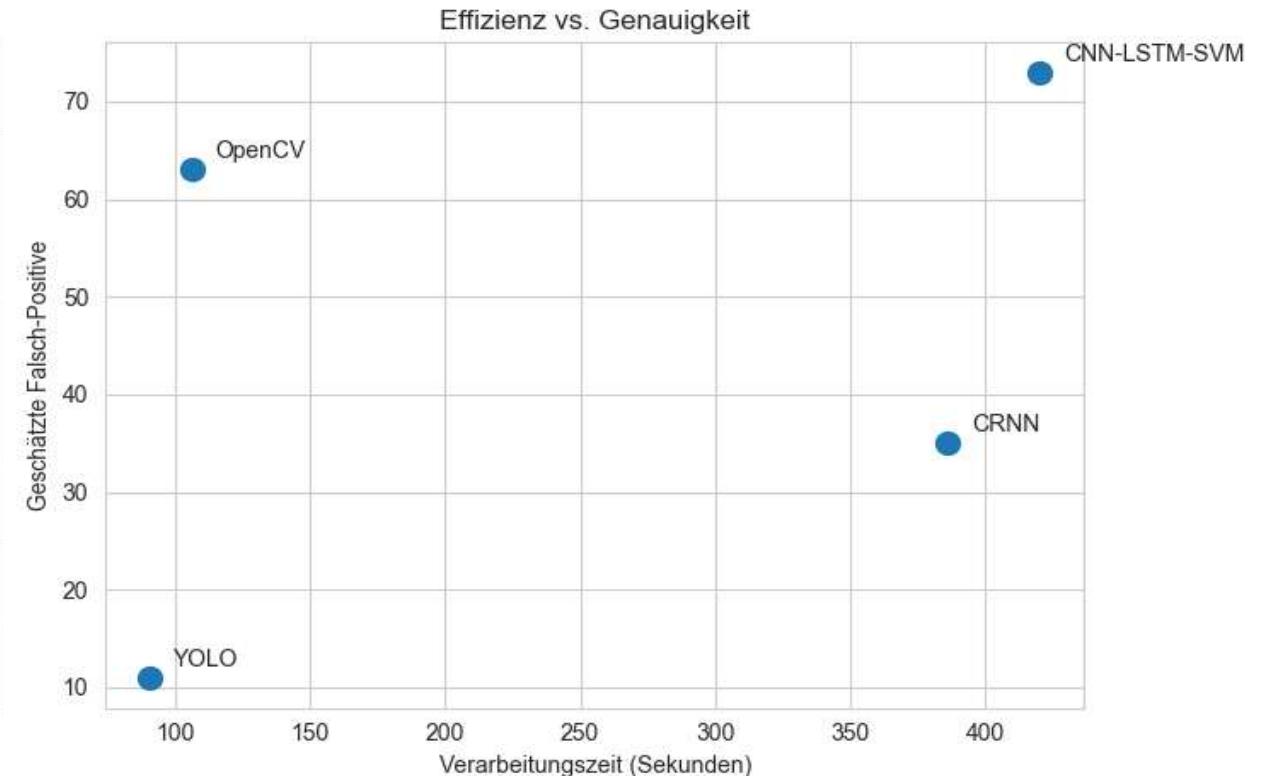
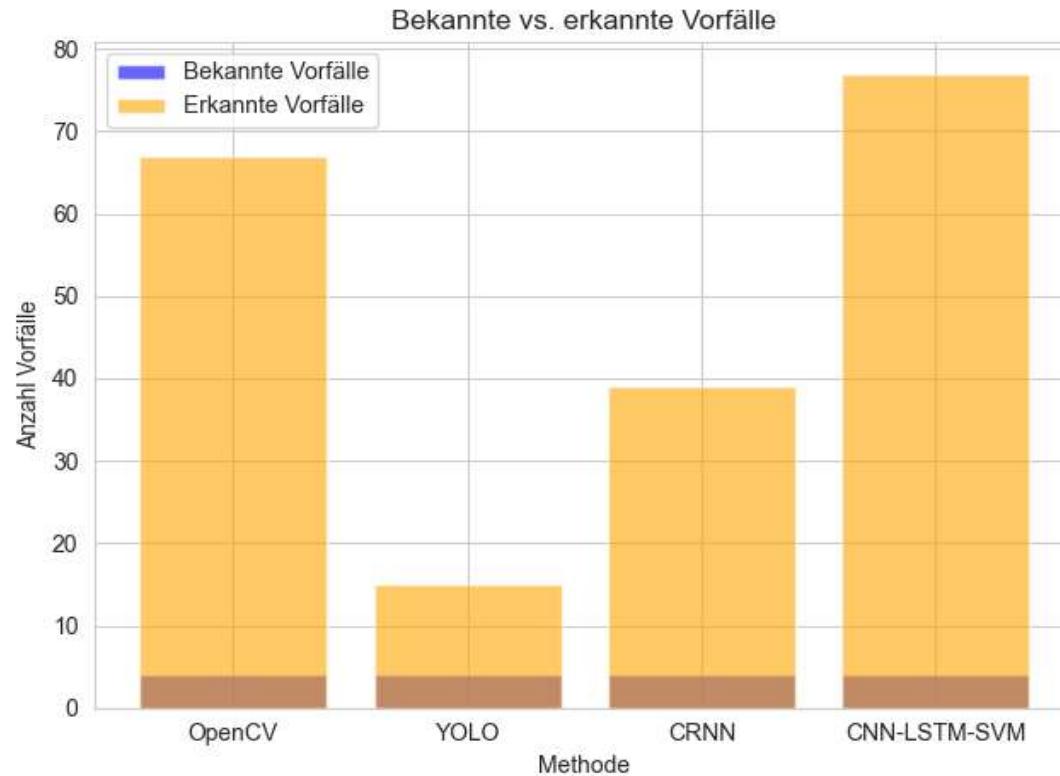
Methode	F1-Score	Erkennungsrate	Falsch-Positive	Verarbeitungszeit
YOLO v11	0,50	11,4%	11	90,7s
OpenCV	0,12	58,0%	63	106,1s
CRNN	0,20	49,7%	35	386,2s
CNN-LSTM-SVM	0,10	100,0%	73	420,2s

- **YOLO-basierter Ansatz:** Beste Balance zwischen Präzision und Geschwindigkeit
- **Hohe Falsch-Positiv-Raten:** Alle Methoden weisen zu viele Fehlalarme auf
- **Schwellenwerte zu niedrig:** Durchgehend übermäßige Sensitivität festgestellt
- **Effizienz vs. Komplexität:** Komplexere Modelle zeigen keine bessere Leistung

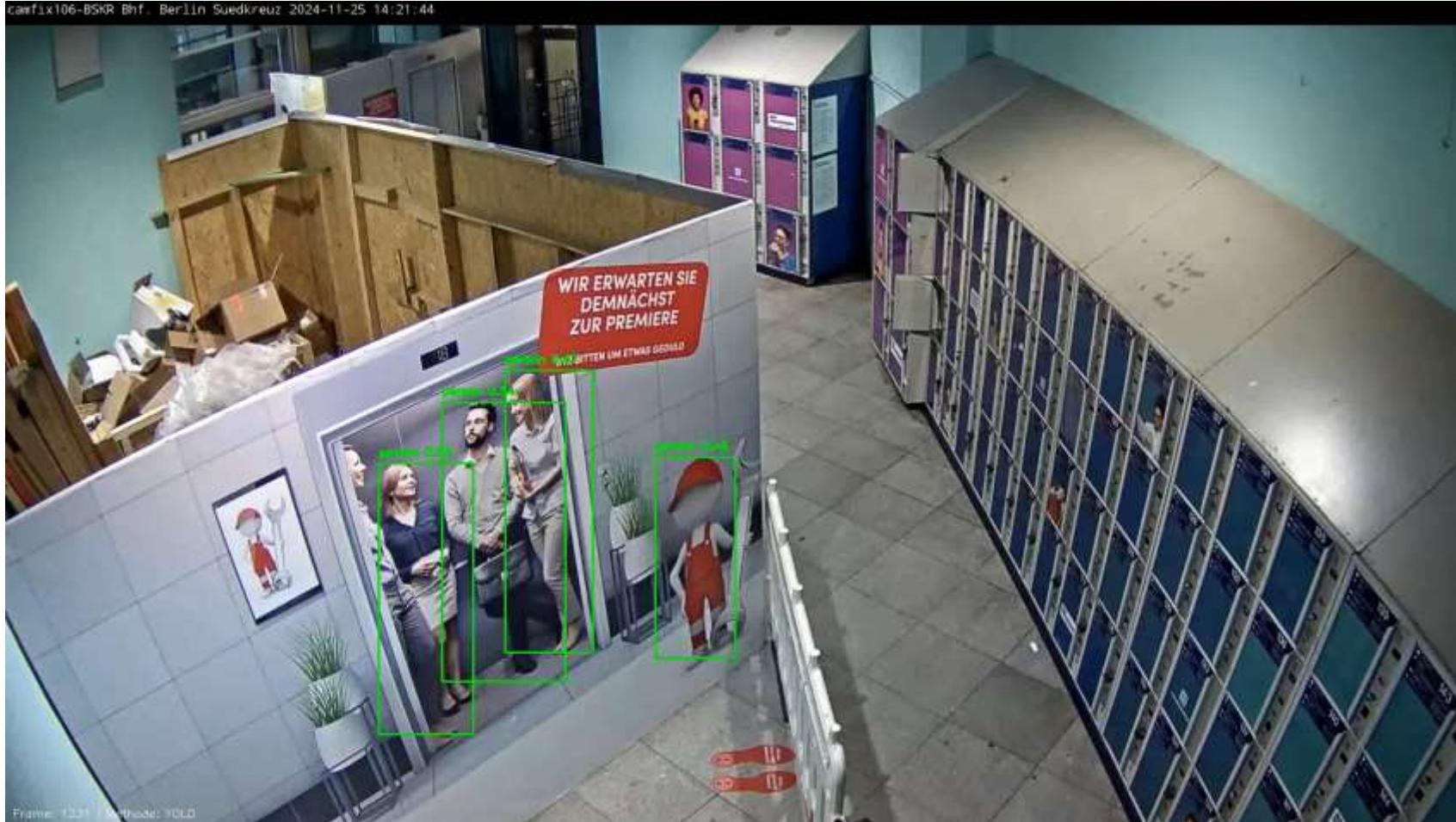
Vandalismuserkennung Ergebnisse



Vandalismuserkennung Ergebnisse



Vandalismuserkennung Ergebnisse | Yolo11



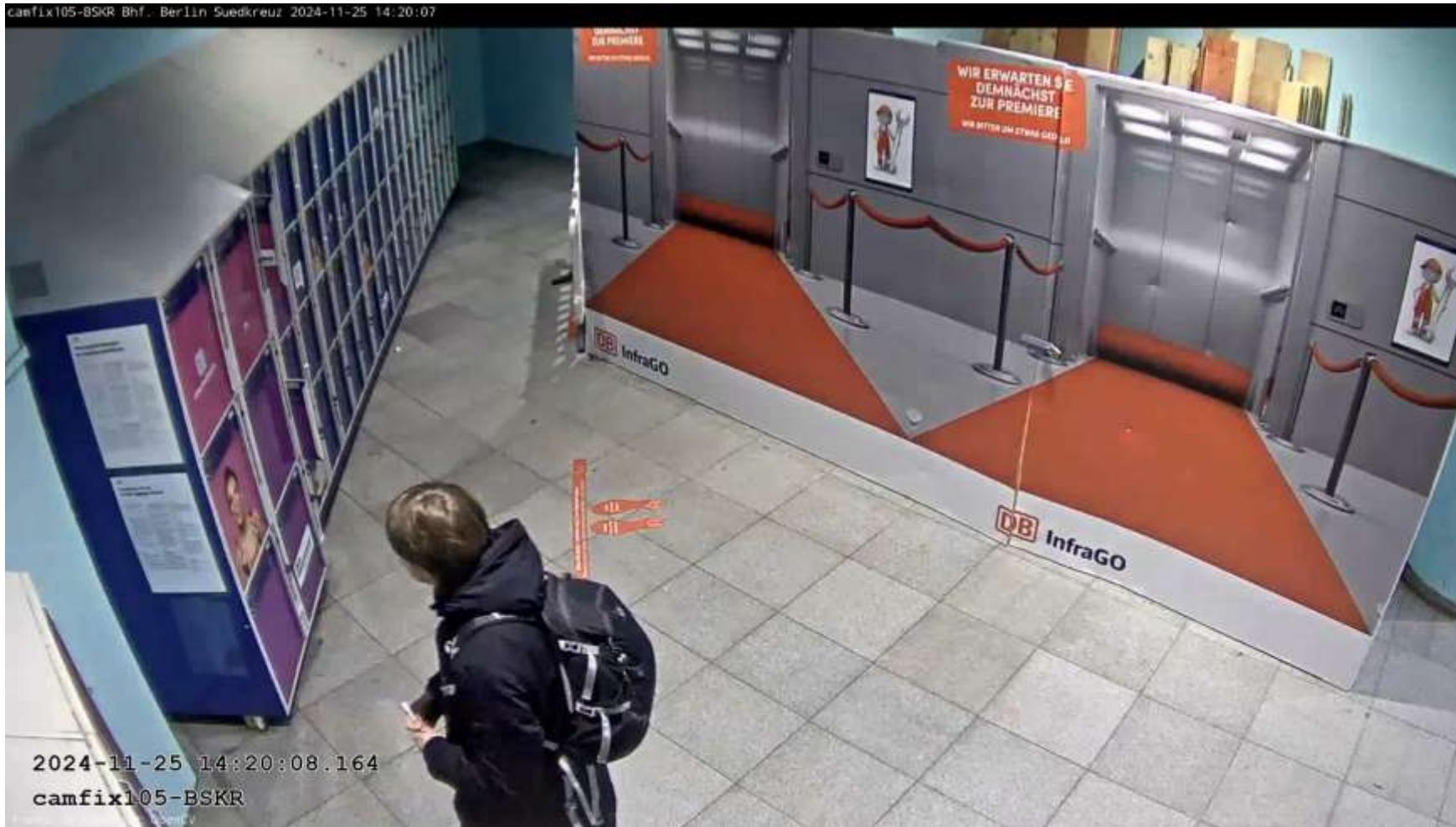


Vandalismuserkennung Ergebnisse | Yolo11





Vandalismuserkennung Ergebnisse | OpenCV





Vandalismuserkennung

Erkenntnisse

Überragende Bedeutung der Präzision

- Zu viele Falschmeldungen führen zu "Alarmermüdung"
- YOLO mit nur 11 Falsch-Positiven klar überlegen

Konservative Ansätze überlegen

- Trotz niedrigster Erkennungsrate bietet YOLO die beste Gesamtperformance

Komplexität ≠ bessere Leistung

- Das komplexeste Modell zeigt die schlechteste Gesamtleistung
- Einfachere, gut optimierte Methoden übertreffen komplexere Ansätze

Schwellenwertoptimierung entscheidend

- Alle Methoden benötigen signifikant höhere Erkennungsschwellen

Kamerasynchronisationsherausforderung

- Mehrfachzählung von Vorfällen durch verschiedene Kameras
- Derselbe Vorfall wird von 2 Kameras als 2 separate Vorfälle gewertet
- Verfälscht die tatsächliche Anzahl an Vorfällen



Fazit

- **Praxisnahe Erprobung:** Direkte Evaluation von KI-Ansätzen in realitätsnahen Szenarien
- **Methodenvergleich:** Quantitativer und Qualitativer Vergleich geeigneter Technologien
- **Problemidentifikation:** Aufdeckung praxisrelevanter Herausforderungen (z. B. Fehlalarme, Synchronisation)
- **Konkrete Ergebnisse:** Ableitung belastbarer Empfehlungen für Architektur und Weiterentwicklung
- **Fundierte Entscheidungen:** Datengrundlage für den gezielten KI-Einsatz
- **Zusammengefasst:** Prototyping war essenziell, um im TAHAI-Projekt die Herausforderungen und Benefits von KI-Lösungen domänenspezifisch aufzuzeigen



https://github.com/hardstoneed/ai_mediation_analyse

https://github.com/hardstoneed/ai_vandalism_detection_public

TAHAI – Prototypische Evaluation (Arbeitspaket 6)

Forschungsprojekt TAHAI - TrustAdHocAI

Sandro Hartenstein | 20.5.2025 | HWR-Berlin