

Automatisierte Qualitätsüberprüfung von Sensordatenannotationen

Automatisierte Qualitätssicherung und Fehlererkennung von Sensordatenannotationen zur Umfeld- und Hinderniserkennung

NIKLAS FREUND | ZEKIYE-ILKNUR ÖZ |
TOBIAS KLOCKAU | PATRICK DENZLER |
PHILIPP NEUMAIER | MARTIN KÖPPEL

Die Digitalisierung bietet enorme Chancen, die Leistungsfähigkeit des Bahnsystems zu steigern. Dies gilt nicht nur unter technologischen Aspekten – auch auf die Herausforderungen des demografischen Wandels kann die Digitalisierung geeignete Antworten geben. Die Sektorinitiative Digitale Schiene Deutschland (DSD) arbeitet daher intensiv an der Entwicklung des vollautomatisierten, fahrerlosen Fahrens. Züge werden zukünftig über Umfeldwahrnehmungssysteme verfügen, die auf Basis des maschinellen Lernens arbeiten. Dafür werden große Mengen an Trainingsdaten benötigt, die aufgrund ihrer Sicherheitsrelevanz einen hohen Qualitätsstandard aufweisen müssen. Die DSD entwickelt aktuell automatische Verfahren zur Qualitätssicherung der Trainingsdaten, womit der Arbeitsaufwand deutlich reduziert und die Entwicklung dieser Systeme beschleunigt werden kann.

Einleitung

Ein vollautomatisiert fahrender Zug hat ein ganzes Paket moderner Technik an Bord. Eine zentrale Komponente dabei ist die Umfeld- und Hinderniserkennung, die gemäß dem Zielbild der DSD weitgehend zuseitig gelöst wird. An der Front und Seite des Zuges übernehmen Sensoren wie Radar, LiDAR, Kamera, Ultraschall und Infrarot die Funktion des menschlichen Auges. Diese Umfeldwahrnehmungssysteme überwachen die Umgebung und müssen Hindernisse im Fahrweg frühzeitig erkennen. Die Technologien dahinter basieren häufig auch auf Künstlicher Intelligenz (KI). Um die KI für die Aufgabe der Umfeldüberwachung zu trainieren, werden große Mengen an Daten benötigt. Diese werden in einer sogenannten „Data Factory“ gespeichert und datenschutzkonform zu Datensätzen für das maschinelle Lernen verarbeitet. Die Data Factory – ein weiteres Entwicklungsprojekt der DSD – stellt dabei eine Plattform zur Speicherung, Verarbeitung, Simulation und Annotation von Daten zur Verfügung [1]. Diese Daten bestehen zum einen aus Sensorrohdaten von Kameras, mehreren LiDAR- und Radarsystemen und zum anderen

aus den dazugehörigen manuellen Annotationen. Annotationen in Sensordatensets, speziell für die Umfelderkennung, sind Markierungen in Rohdaten zum Trainieren, Testen und Validieren von KI-Verfahren. Bei 2D-Sensoren wie Kameras umfassen diese Annotationen oft sogenannte Bounding-Boxes (Abb. 1a), Polylinien (Abb. 1b) und Polygone (Abb. 1c), um Objekte in Bildern zu kennzeichnen. Für 3D-Sensoren wie LiDAR werden Punktwolken annotiert, indem 3D-Bounding-Boxes (Abb. 1d) um Objekte gezogen oder bestimmte Punkte markiert werden. Die Eigenschaften einzelner Annotationen werden durch sogenannte Attribute repräsentiert, die aus einer Bezeichnung und einem zugehörigen Wert bestehen.

Da die Umgebungswahrnehmung im Schienenverkehr von sicherheitskritischer Relevanz ist, müssen die Annotationen einen sehr hohen Qualitätsstandard erreichen. Hierzu ist eine ausführliche Qualitätsprüfung erforderlich, die heute überwiegend manuell erfolgt. Aufgrund der stetig steigenden Datenmengen wird es jedoch notwendig, hierbei vermehrt auf automatisierte Verfahren der Fehlererkennung zurückzugreifen, um eine beschleunigte Datenaufbereitung zu ermöglichen. Von hier entwickelten Software werden zukünftig Fehler, wie z.B. Größen- oder Attributfehler, schneller erkannt.

Stand der Technik

In der manuellen Prüfung werden alle Annotationen auf Fehler überprüft, was mit einem hohen Arbeitsaufwand verbunden ist. Um den Arbeitsaufwand zu reduzieren, kann beispielsweise nur ein möglichst repräsentativer Teil der Daten geprüft werden, was jedoch bei ungünstiger Wahl der geprüften Daten zu einer signifikanten Fehleinschätzung der Qualität führen kann.

Einige Fehlertypen lassen sich auch automatisiert erkennen [2]. Ziel ist, dass die automatisierte Qualitätsprüfung zukünftig nahezu alle Fehler erkennt. Durch die Aufstellung von Regeln zur Erkennung bestimmter Fehler kann der gesamte Datensatz effektiv auf diese geprüft werden [3]. Der nötige menschliche Aufwand zur Prüfung kann somit signifikant reduziert werden, sofern die Genauigkeit der Erkennungsverfahren nachgewiesen wurde. Die automatisch erkannten Fehler können dann als Basis für die meist manuelle Fehlerkorrektur verwendet werden [4].



Abb. 1: Übersicht über die Annotationstypen Bounding Box (a), Polyline (b), Polygon (c) und 3D-Bounding-Box (d)

Quelle: eigene Grafik unter Verwendung von [8]

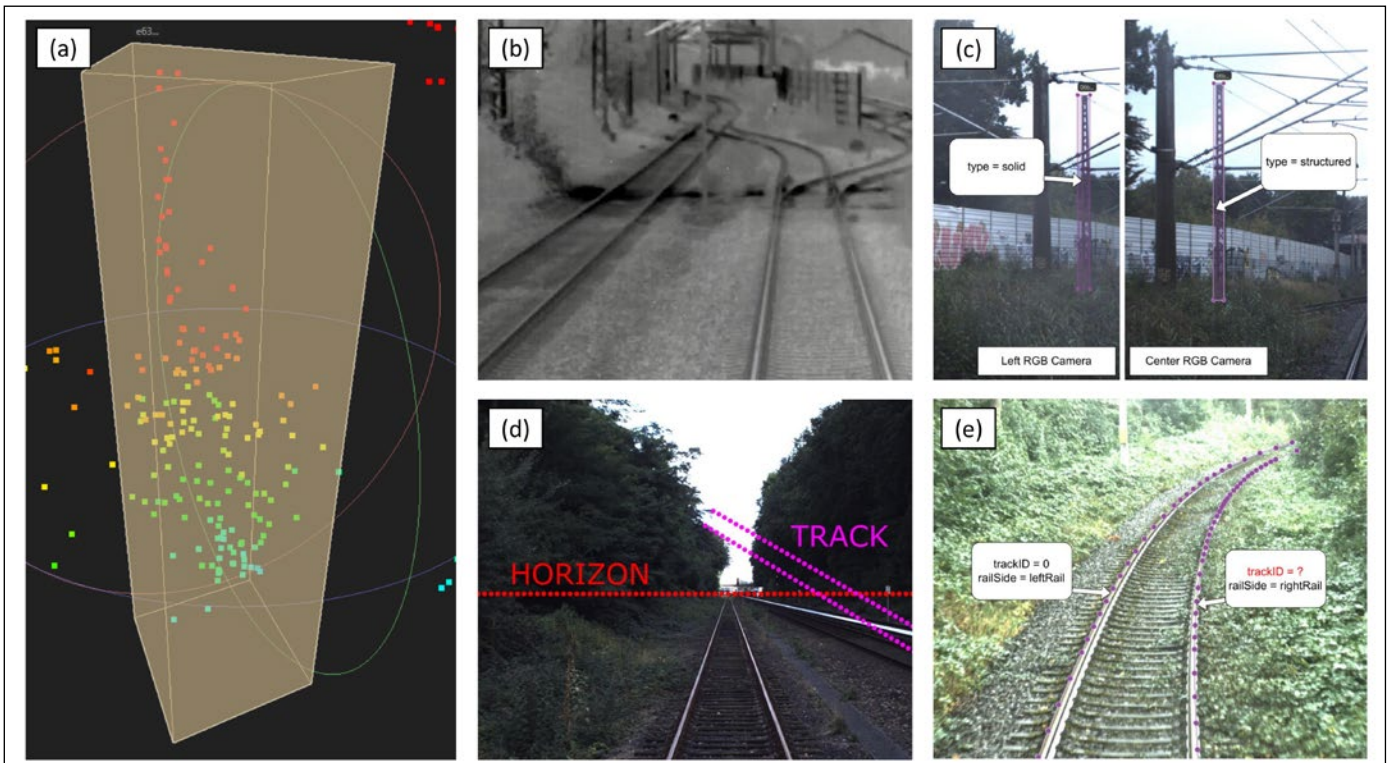


Abb. 2: Beispiele für die Fehlertypen DimensionInvalid (a), EmptySensor (b), InconsistentAttributeScope (c), AnnotationAboveHorizon (d) und MissingAttribute (e)
 Quelle: eigene Grafik unter Verwendung von [8]

Aufbauend auf der automatisierten Fehlererkennung ist es unter bestimmten Voraussetzungen möglich, Fehler auch automatisiert zu korrigieren. Ein einfacher Ansatz hierfür ist,

alle fehlerhaften Annotationen zu löschen, sodass auf Kosten der Vollständigkeit nur noch korrekte Annotationen übrigbleiben [5]. Alternativ können die Annotationen automatisiert

modifiziert werden. Hierbei ist es jedoch wichtig, die Korrektheit der automatischen Fehlerkorrektur nachzuweisen, damit keine zusätzlichen Fehler in den Datensatz einfließen.

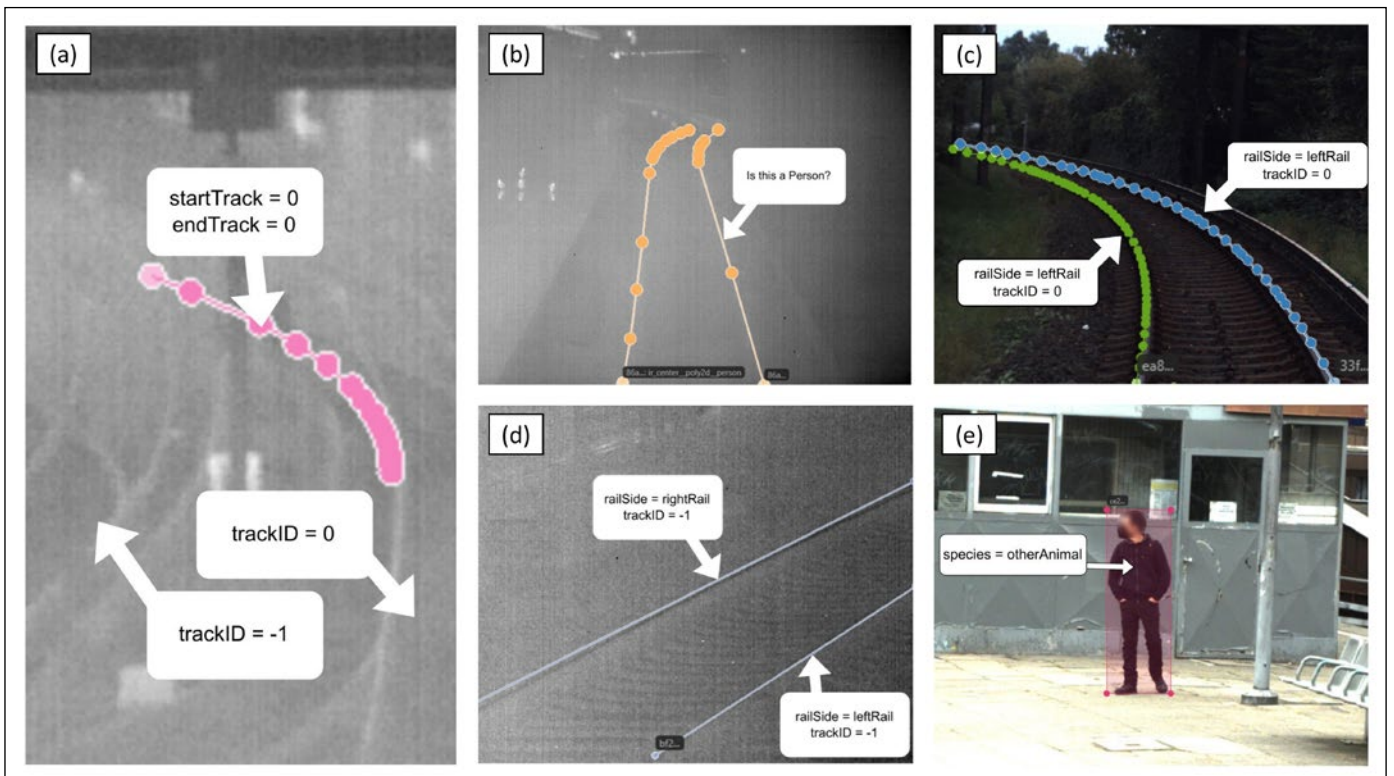


Abb. 3: Beispiele für die Fehlertypen TransitionIdenticalStartAndEnd (a), MissingEgoTrack (b), RailSideCount (c), RailSideOrder (d) und UnexpectedAttribute (e)
 Quelle: eigene Grafik unter Verwendung von [8]

Die DSD hat 2023 einen frei verfügbaren Multisensordatensatz für maschinelles Lernen zur Entwicklung des vollautomatisierten Fahrens erstellt und in einem standardisierten Annotationsformat veröffentlicht (OSDaR23) [6, 7, 8]. Er wurde in Kooperation zwischen dem Deutschen Zentrum für Schienenverkehrsforschung beim Eisenbahn-Bundesamt (DZSF), DSD sowie FusionSystems GmbH erstellt. Der Datensatz enthält 45 Sequenzen synchronisierter Multisensordaten. Verwendet wurden drei RGB-Kameras mit mittlerer und drei mit hoher Auflösung, drei Infrarot-Kameras sowie ein Radar und sechs zusammengerechnete LiDAR-Sensoren. Insgesamt enthält der Datensatz 204 091 Annotationen. Die Anforderungen an die Annotationen sind als Labeling-Guide öffentlich verfügbar [9]. Zur Erprobung automatischer Fehlererkennung entsprechend dem Stand der Technik nutzen wir diesen Datensatz.

Konzept

Anhand einer Pre-Release-Version des OSDaR23-Datensatzes, bei der noch eine Iteration einer manuellen Qualitätsprüfung fehlte, erprobte die DSD die Entwicklung und den Einsatz vollautomatischer Fehlererkennungsverfahren zur Qualitätsprüfung. Hierzu wurden zunächst häufig beobachtete Fehlertypen definiert. Für jeden Fehlertyp wurden Regeln zur Erkennung entwickelt, welche in Softwarealgorithmen implementiert wurden. Die Software wurde mit der Programmiersprache „Python“ entwickelt, wobei zum Laden der Annotationen auf die RailLabel-Bibliothek [10] zurückgegriffen wird.

Problematische Annotationen und Erkennung

In Abb. 2 und 3 sind genannte Fehlertypen detailliert dargestellt:

- **DimensionInvalid** (Abb. 2a): Falls die Dimension einer 3D-Bounding-Box in x-, y- oder z-Richtung zu groß oder zu klein für seine Objektklasse ist, handelt es sich um den Fehlertyp `DimensionInvalid`. Im Beispiel ist die 3D-Bounding-Box für einen Menschen 3,3 m hoch.
- **EmptySensor** (Abb. 2b): Beinhaltet ein Sensor über die komplette Sequenz hinweg keine Annotationen, handelt es sich um den Fehlertyp `EmptySensor`. Hier wurde z.B. vergessen, die Gleise zu annotieren.
- **InconsistentAttributeScope** (Abb. 2c): Ändert sich ein Attribut, welches über Bilder oder Sensoren hinweg unveränderlich sein soll, handelt es sich um den Fehlertyp `InconsistentAttributeScope`. Im Beispiel wurde der „Type“ eines „Catenary Pole“ auf einer Kamera als „structured“, aber auf einer anderen Kamera als „solid“ markiert.
- **AnnotationAboveHorizon** (Abb. 2d): Wurde ein Gleis zu weit gezogen und ragt über den Horizont hinaus, so handelt es sich um den Fehlertyp `AnnotationAboveHorizon`. Im Beispiel ragt ein Gleis in den Himmel.
- **MissingAttribute** (Abb. 2e): Beim Fehlertyp `MissingAttribute` fehlt einer Annotation ein erwartetes Attribut. Im Beispiel fehlt einer „Track“-Annotation die „TrackID“.
- **TransitionIdenticalStartAndEnd** (Abb. 3a): Startet und endet ein Übergangsgleis, welches eigentlich zwei verschiedene Gleise miteinander verbinden sollte, auf demselben Gleis, handelt es sich um den Fehlertyp `TransitionIdenticalStartAndEnd`. Im Beispiel star-

tet es auf trackID 0 (eigenes Gleis) und sollte auf dem Nachbargleis (trackID -1) enden. Fälschlicherweise endet es hier auf trackID 0.

- **MissingEgoTrack** (Abb. 3b): Wurde das Gleis, auf dem der Zug fährt, auf einer mittleren Kamera nicht als eigenes Gleis (Ego-Track) annotiert, handelt es sich um den Fehlertyp `MissingEgoTrack`. Im Beispiel wurde ein Gleis stattdessen mit der Objektklasse „Person“ annotiert.
- **RailSideCount** (Abb. 3c): Beim Fehlertyp `RailSideCount` hat ein Gleis mehr als eine linke oder rechte Schiene. Typischerweise sollte ein Gleis eine linke und eine rechte Schiene haben. Im Fall einer Verdeckung kann es auch nur eine linke oder eine rechte Schiene geben.
- **RailSideOrder** (Abb. 3d): Ist die Reihenfolge der Schienen eines Gleises vertauscht, handelt es sich um den Fehlertyp `RailSideOrder`. Im Beispiel befindet sich das linke Gleis auf der rechten und das rechte Gleis auf der linken Seite.
- **UnexpectedAttribute** (Abb. 3e): Hier hat eine Annotation ein für ihre Objektklasse unerwartetes Attribut. Beispielsweise wurde bei einem Menschen das Attribut „Species“ gesetzt, welches eigentlich zur Objektklasse „Animal“ gehört (Abb. 3e).

Auswertung

Die entwickelte Software wurde auf der o.g. Pre-Release-Version des OSDaR23-Datensatzes getestet, auf der im Vergleich zur veröffentlichten Version noch eine manuelle Qualitätsprüfung fehlte. Um eine Aussage über die Präzision der entwickelten Verfahren zur Fehlererkennung treffen zu können, wurden die von den automatischen Qualitätsprüfungen erkannten Fehler manuell gegengeprüft. Auf diese Weise konnte ermittelt werden, ob es sich um einen tatsächlichen Fehler handelt, oder ob die Software fälschlicherweise einen Fehler erkannte. Abb. 4 zeigt für jede Fehlerklasse, wieviel Prozent der automatisch gefundenen Fehler in der manuellen Nachprüfung auch tatsächlich einen Fehler markiert haben.

Für fünf Fehlertypen konnten die entwickelten Erkennungsverfahren eine Präzision von 100 % erzielen. D.h., dass alle automatisch erkannten Fehler in den Annotationen auch tatsächliche Fehler darstellen. Für drei Fehlerklassen war die Präzision über 89 % (Abb. 4). Lediglich der „DimensionInvalid“- und der „EmptySensor“-Fehlertyp wiesen eine unzureichende Präzision auf. So waren nur 38 % bzw. 24 % der von der Software gefundenen Fehler tatsächlich Fehler. Eine eingehende Untersuchung hat gezeigt, dass für die Fehlerklasse „DimensionInvalid“ die Schwellwerte bei der Aufstellung der Regeln (s. o.) zu streng gesetzt wurden. Oft wurden die Schwellwerte nur um wenige Zentimeter überschritten und haben somit fälschlicherweise einen Fehler signalisiert. Weiterhin haben Spezialobjekte

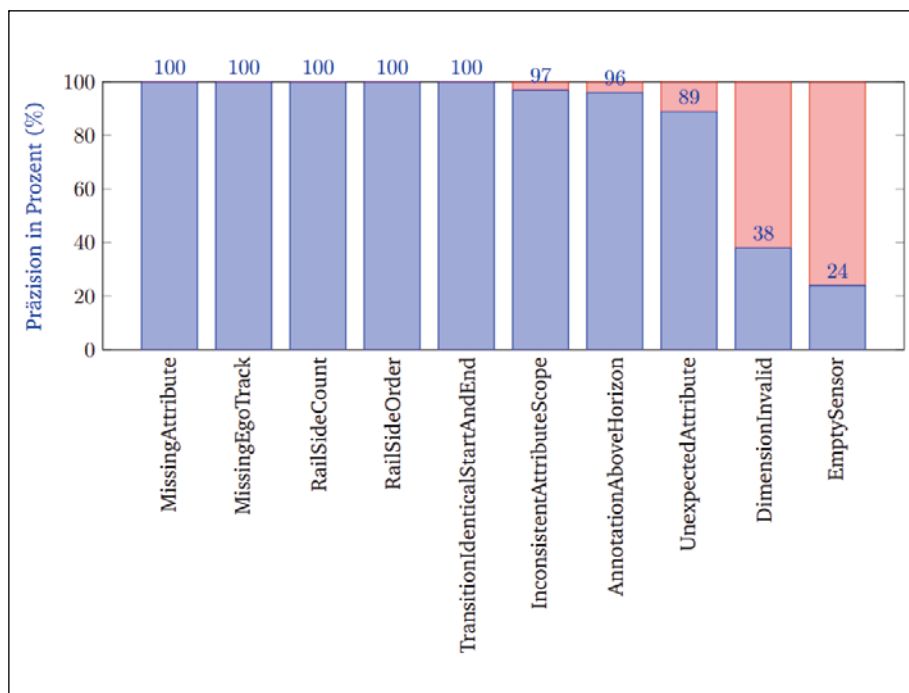


Abb. 4: Präzision der Erkennung verschiedener Fehlertypen

Quelle: eigene Grafik

größere Dimensionen aufgewiesen, als es bei der Definition der Regeln berücksichtigt wurde. Ein Beispiel hierfür ist ein Baufahrzeug mit einer Baggervorrichtung, die komplett nach oben ausgefahren war. Die Baggervorrichtung hatte in diesem Fall die vorgesehene Dimension in z-Richtung überschritten. Zur Verbesserung der Fehlererkennung werden zukünftig die erlaubten Größen der 3D-Bounding-Boxen überarbeitet.

Der Fehlertyp „EmptySensor“ prüft für jeden Sensor individuell, ob mindestens eine Annotation in der Sequenz dem Sensor zugeordnet ist. Nach der Überprüfung der Ergebnisse wurde festgestellt, dass die nach links oder rechts ausgerichteten Kameras häufig keine der gesuchten Objekte erfassen. Ein Grund dafür sind beispielsweise Lärmschutzwände im Sichtfeld der Kamera. Hierdurch wurden viele falsche positive Erkennungen generiert, sodass dieser Fehlertyp zukünftig grundlegend überarbeitet werden muss.

Zusammenfassung und Ausblick

Im Rahmen dieser Arbeit wurde eine Qualitätsprüfung für zehn häufig auftretende Fehlertypen in annotierten Sensordaten für

die Umfeldwahrnehmung im Bahnbereich entwickelt. Insgesamt zeigen die implementierten, automatischen Qualitätsprüfverfahren sehr gute Ergebnisse. Acht der entwickelten Prüfverfahren weisen eine sehr hohe bzw. vollständige Präzision bei der Fehlererkennung auf. Zwei der entwickelten Qualitätsprüfverfahren weisen eine etwas zu hohe Falscherkennungsrate auf. Daher muss hier aktuell noch ein zusätzlicher manueller Prüfschritt durchgeführt werden.

In zukünftigen Arbeitsschritten wird die Präzision der Prüfverfahren geschärft und auf weitere Fehlertypen ausgeweitet. Durch die Integration der automatischen Qualitätsprüfung in ein Annotationsprojekt kann die Qualität von Datensätzen für den Bahnsektor dadurch enorm verbessert werden. Zusätzlich wird der manuelle Arbeitsaufwand der Qualitätsprüfung durch die automatischen Verfahren reduziert. Die auf diese Weise qualitätsgeprüften Daten werden über die Data Factory [7] zukünftig Unternehmen und wissenschaftlichen Instituten für die Entwicklung von KI-Verfahren zur Umfelderkennung im Bahnbereich zur Verfügung gestellt. ■

QUELLEN

- [1] <https://digitale-schiene-deutschland.de/de/projekte/DataFactory>, 28.08.2024 um 17:00
- [2] Klie, J.-C.; de Castilho, R. E.; Gurevych, I.: „Analyzing Dataset Annotation Quality Management in the Wild“, Computational Linguistics, S. 1–48, März 2024, doi: 10.1162/coli_a_00516
- [3] Klie, J.-C.; Webber, B.; Gurevych, I.: „Annotation Error Detection: Analyzing the Past and Present for a More Coherent Future“, Computational Linguistics, Bd. 49, Nr. 1, S. 157–198, März 2023, doi: 10.1162/coli_a_00464
- [4] Weber, L.; Plank, B.: „ActiveAED: A Human in the Loop Improves Annotation Error Detection“, 31. Mai 2023, arXiv: arXiv:2305.20045. doi: 10.48550/arXiv.2305.20045
- [5] Huang, J.; Qu, L.; Jia, R.; Zhao, B.: „02U-Net: A Simple Noisy Label Detection Approach for Deep Neural Networks“, in 2019 IEEE/CVF International Conference on Computer Vision (ICCV), Okt. 2019, S. 3325–3333. doi: 10.1109/ICCV.2019.00342
- [6] Tilly, R.; Neumaier, P. et al.: „Offener Multisensordatensatz für die Entwicklung der Umfeldwahrnehmung beim vollautomatisierten Fahren“, EI 04/2023, S. 24–29, online unter <https://digitale-schiene-deutschland.de/Downloads/ETR-OSDaR23.pdf>, 07.08.2024 um 15:02 Uhr
- [7] Marsch, P.; Neumaier, P. et al.: „Pan-European Railway Data Factory – Infrastruktur und Ökosystem für einen vollautomatisierten Bahnbetrieb“, SIGNAL+DRAHT 04/2024, S. 6–14, online unter <https://digitale-schiene-deutschland.de/de/aktuelles/2024/Pan-European-Railway-Data-Factory>, 07.08.2024 um 15:51 Uhr
- [8] <https://data.fid-move.de/dataset/osdar23>, 07.08.2024 um 17:13
- [9] Kaden, K.-U.; Karsten, S.; Scheithauer, M.: „Labeling Guide for Multi-sensor Data for Object Recognition in the Railway Domain“, German Centre of Rail Traffic Research at the Federal Railway Authority, Okt. 2023. doi: 10.48755/DZSF.230012.05
- [10] <https://github.com/DSD-DBS/raillabel>, 07.08.2024 um 15:11



Niklas Freund

Software Developer
Digitale Schiene Deutschland
DB InfraGO AG, Berlin
niklas.freund@deutschebahn.com



Zekiye-Ilknur Öz

Annotation Quality Managerin
Digitale Schiene Deutschland
DB InfraGO AG, Berlin
zekiye-ilknur.oez@deutschebahn.com



Tobias Klockau

Computer Vision & Data Engineer
Digitale Schiene Deutschland
DB InfraGO AG, Berlin
tobias.klockau@deutschebahn.com



Patrick Denzler

Product Owner Data Management
Digitale Schiene Deutschland
DB InfraGO AG, Berlin
patrick.denzler@deutschebahn.com



Dr. Philipp Neumaier

Head of Data Factory
Digitale Schiene Deutschland
DB InfraGO AG, Berlin
philipp.neumaier@deutschebahn.com



Dr. Martin Köppel

Product Owner
GoA4 Datasets & Annotation Tool
Digitale Schiene Deutschland
DB InfraGO AG, Berlin
martin.koepfel@deutschebahn.com



Archiv

über 44.000 Beiträge

laufende Aktualisierung

individuelle Suchoptionen

Volltextsuche

Sofort-Download

Abonnenten erhalten bis zu

50% Rabatt



www.eurailpress.de/fachartikel