

# KI-basierte Analyse von Wirbelstrom- und Ultraschalldaten bei der Schienenprüfung

Thomas HECKEL<sup>1</sup>, Ralf CASPERSON<sup>1</sup>, Axel SIMROTH<sup>2</sup>

<sup>1</sup> Bundesanstalt für Materialforschung und -prüfung (BAM), Berlin

<sup>2</sup> Deutsches Zentrum für Schienenverkehrsforschung beim Eisenbahn-Bundesamt, Dresden

Kontakt E-Mail: thomas.heckel@bam.de

**Kurzfassung.** Die zerstörungsfreie Prüfung von Eisenbahnschienen auf betriebsbedingte Schädigungen wird mit Schienenprüfzügen mittels Ultraschall- und Wirbelstromprüfverfahren durchgeführt. Die Auswertung der Prüfdaten erfolgt überwiegend manuell, die eingesetzte Software unterstützt die Auswertenden lediglich durch eine Vorauswahl relevanter Anzeigen. Die Überprüfung der Ergebnisse erfolgt anschließend vor Ort mittels handgeführter Prüfgeräte. Instandhaltungsmaßnahmen werden auf Basis der Befundung vor Ort abgeleitet.

Ziel des durch das Bundesministerium für Digitales und Verkehr (BMDV) im Rahmen von mFUND unter dem Förderkennzeichen 19FS2014 geförderten Vorhabens AIFRI (Artificial Intelligence For Rail Inspection) ist es, durch den Einsatz von KI-Methoden den Automatisierungsgrad des Prüfprozesses von der Auswertung der Daten bis hin zur Planung von Instandhaltungsmaßnahmen zu erhöhen. Die Genauigkeit der Fehlerdetektion soll gesteigert werden, um eine automatisierte Einstufung der aufgefundenen Anzeigen in Risikoklassen zu ermöglichen. Hierfür werden Daten sowohl von Wirbelstromprüfungen als auch Ultraschallprüfungen verwendet.

Im Rahmen des IT-orientierten Projektes werden relevante Schienenschädigungen und in der Schiene vorhandene Artefakte analysiert und in einen parametrierbaren digitalen Zwilling übertragen. Mit diesem digitalen Zwilling werden virtuelle Schädigungsbilder generiert, mit denen KI-Algorithmen auf die Defekterkennung und -klassifizierung trainiert werden. Insbesondere werden hierbei Synergien genutzt, die durch die Verknüpfung von Daten der Wirbelstrom- und Ultraschallprüfung bei einer gemeinsamen Bewertung entstehen. Mit Hilfe von Zuverlässigkeitsbetrachtungen werden die entwickelten und trainierten Algorithmen hinsichtlich der Detektion und Charakterisierung von Schienenschädigungen bewertet.

Im Verlauf des Projektes soll mit dem entwickelten IT-Werkzeug ein Demonstrator aufgebaut und im Feld mit realen Datensätzen getestet werden.

## Einführung

Eine Analyse des Ist-Zustandes des Vorgehens bei der Instandhaltung der Schieneninfrastruktur zeigt, dass dieser Prozess heute zumeist nach präventiven Konzepten erfolgt. In regelmäßigen Zeitintervallen werden definierte Maßnahmen durchgeführt, um einen geforderten Zustand der Gleise zu erhalten. Grundlage für die Entscheidung, welche Maßnahmen geplant werden, bilden dabei Mess- und Prüfdaten zum Schienenzustand aus Inspektionsfahrten mit Prüfzügen. Diese werden bisher nur teilweise automatisiert ausgewertet und lassen nur ungenaue Rückschlüsse über die tatsächlichen Schädigungsbilder und deren Ausmaße zu, so dass jede Fehlervermutung durch eine Begutachtung der Stelle vor Ort verifiziert werden muss. Um den Gesamtprozess der Schienenprüfung und Instandhaltungsplanung zu optimieren, soll im Projekt AIFRI ein datenbasiertes IT-Werkzeug entwickelt werden, welches den gesamten Bearbeitungsprozess von der Analyse bis zur Entscheidungsfindung in digitaler Form unterstützt. Das zum 01.12.2021 gestartete Projekt setzt sich hierfür die folgenden konkreten Ziele:

- Anhebung des Automatisierungsgrades im Prüfprozess:

Der gesamte Prüfprozess, vor allem der aufwändige Schritt der Sichtung von Prüfdaten und Bewertung von Anzeigemustern bzw. Fehlerbildern, soll durch eine stark verbesserte, weitgehend automatisierte Datenanalyse mittels KI-Verfahren unterstützt werden.

- Erhöhung der Genauigkeit der Fehlerdetektion:

Die Zuverlässigkeit des Prüfprozesses wird wesentlich von der Genauigkeit der Fehlerdetektion bestimmt. Hier soll eine Erhöhung vor allem durch Einsatz neuer Verfahren aus dem Bereich KI/Machine Learning sowie durch Kombination der verwendeten Datenquellen der Ultraschall- und Wirbelstromprüfung erreicht werden. Dadurch soll auch der Umfang der erkennbaren Schädigungsbilder erweitert werden.

- Klassifizierung und Einstufung der Fehler in „Risikoklassen“:

Neben der reinen Fehlerkennung soll eine neue Qualität des Erkenntnisgewinns aus den Datenanalysen heraus erreicht werden durch eine Klassifizierung und Einstufung der Schädigungen in „Risikoklassen“, die die Dringlichkeit der Behebung der Schädigung ausdrücken.

- Unterstützung bei der Maßnahmenplanung:

Basierend auf der Risikobewertung soll eine Entscheidungsunterstützung für die Ableitung, Bündelung und Planung der notwendigen Instandhaltungsmaßnahmen konzipiert werden. Damit kann das bisherige präventive Vorgehen in ein risikobasiertes Instandhaltungskonzept überführt werden.

- Datenmodellierung für Schienenprüfdaten:

Zur Weitergabe und potenziellen Veröffentlichung von Schienenprüfdaten sowie der daraus ermittelten Zustände soll ein generisches Datenformat entwickelt werden, welches sich an vorhandene Standards wie IDMVU oder railML anlehnt bzw. diese erweitert.

Die Erreichung dieser Ziele soll zum Abschluss des dreijährigen Projekts durch eine Demonstration des entwickelten IT-Werkzeugs auf einem ausgewählten Netzabschnitt nachgewiesen werden.

## **1. Ansatz**

Ultraschall- und Wirbelstromdaten der mechanisierten Schienenprüfung mit Prüfzügen werden heute meist noch manuell ausgewertet, um mögliche Schäden an den Schienen zu klassifizieren. Um Schäden automatisiert mit KI-Methoden zu identifizieren, bietet sich vor allem die Anwendung der Schadensklassifikation mittels neuronaler Netzwerke [1],[2] als gut geeignete Methode an.

Eine Herausforderung für die Anwendung des maschinellen Lernens in der Praxis stellt dabei die Bereitstellung geeigneter Trainingsdatensätze zum Anlernen (Training) der KI dar. Zur Anwendung des maschinellen Lernens werden gelabelte Prüfdatensätze benötigt, in denen Anzeigen von Schädigungen und Artefakten in Schienen bereits deren Ursache zugeordnet sind.

Es sind bislang keine Prüfdatensätze einer Referenzstrecke bekannt und in entsprechendem Umfang vorhanden, die eine ausreichend sichere statistische Basis an Schädigungen und Artefakten abbilden. Dezidierte Teststrecken für Zulassungsfahrten sind für die geplante Untersuchung im Projekt nicht vollumfänglich nutzbringend, da diese Teststrecken typischerweise mit speziellen geometrischen Artefakten nach DIN EN 16729-1 und -2 ausgestattet sind, die nur bedingt reale Schienenschädigungen nachbilden.

Prüfdaten aus dem Feld bilden die notwendigen Anzeigenmuster von Schienenschädigungen glücklicherweise nicht in ausreichendem Umfang ab, zudem ist bei realen Daten die Ursache der Anzeigen oft nicht hinreichend bekannt. Weiterhin wird die für ein Training notwendige Balance in den Prüfdaten aus dem Feld zwischen „gut“ und „schlecht“ nicht erreicht. Reale Prüfdaten sind daher zum Training einer KI nur unzureichend geeignet. Ohne die Zuhilfenahme von Simulationsrechnungen ist die Interpretation von Prüfdaten ausschließlich auf Basis der dargestellten Anzeigen zu unzuverlässig. Aus diesem Grund sollen mit Hilfe eines digitalen Zwillings synthetische Trainingsdaten generiert werden. Hierzu werden mit Hilfe von umfangreichen Simulationen und a-priori-Wissen die Signalantworten der Prüfsysteme von Musterfehler (z. B. Squats, Anrisse) und Artefakten (z. B. Stegbohrungen, Schweißungen) erzeugt, so wie diese von den Prüfsystemen generiert werden.

## **2. Simulation**

Mit Hilfe von Simulationen können entsprechend geeignete und gelabelte Datensätze für die Ultraschallprüfung [3] und die Wirbelstromprüfung [4] erzeugt werden.

In einem ersten Schritt werden die relevanten Artefakte in und Schädigungen an Schienen sowie deren Anzeigemuster erhoben. Hierzu müssen im Projekt zuerst die Schienenfehler und ihre resultierenden Anzeigemuster beschrieben und geeignete Eigenschaften aus den Prüfdaten identifiziert oder berechnet werden. Es werden dazu Features ermittelt, mit welchen Fehlerstellen ausreichend stochastisch beschrieben werden können.

In einem zweiten Schritt werden die Artefakte und Schädigungen in einen parametrierbaren digitalen Zwilling überführt, mit dessen Hilfe beliebig parametrisierte, virtuelle Prüffahrten erstellt werden können, so dass die volle Bandbreite der möglichen Schädigungen und Artefakte abgebildet werden kann. Hierdurch können die Signalantworten

von bekannten Fehlertypen und Artefakten nachvollzogen und deren Parameter (z. B. Fehlerlänge, Schädigungstiefe, Orientierung) beliebig variiert werden.

Die simulierten Daten werden in einem dritten Schritt in Laborversuchen an Vergleichskörpern validiert.

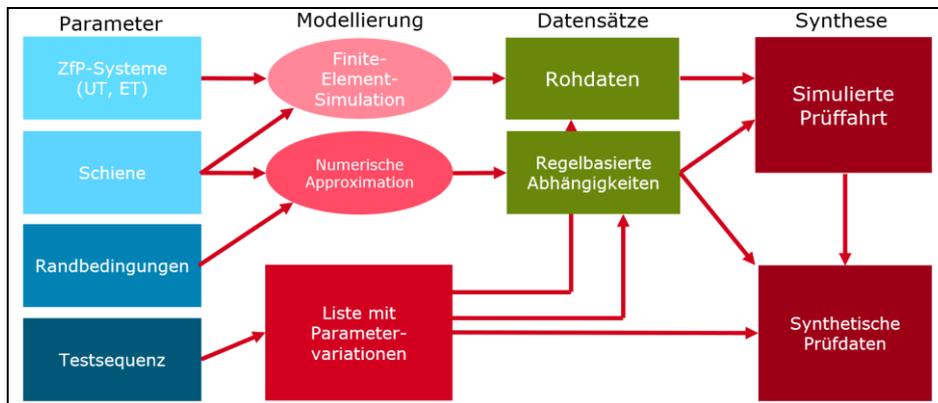


Abb. 1. Simulation einer Schienenprüfung

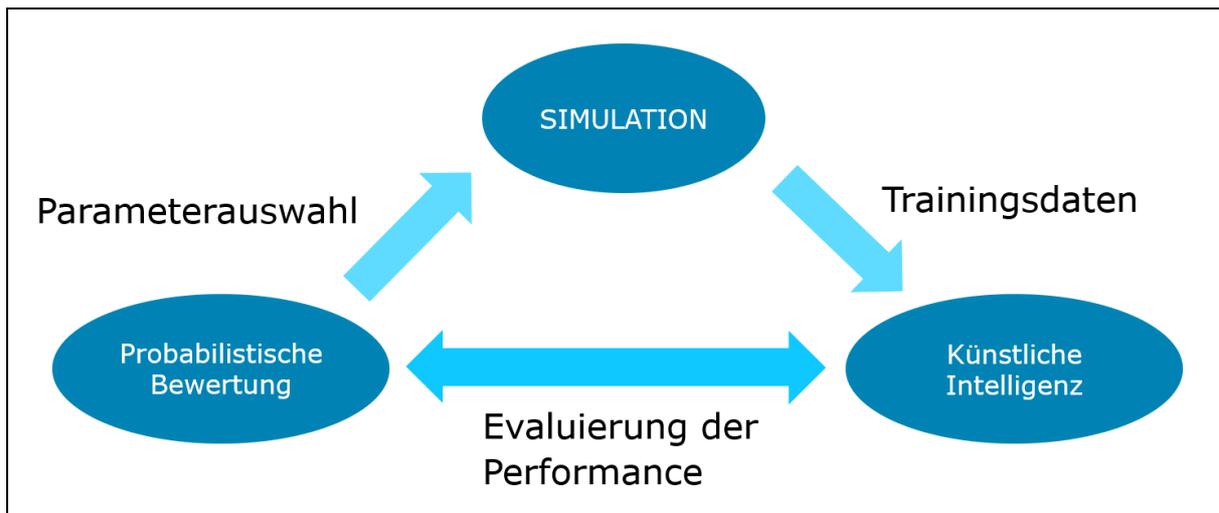
Die akustischen und elektromagnetischen Wechselwirkungen werden in Abhängigkeit der Parameter der ZfP-Prüfsysteme und der Schiene z. B. mit Finite-Element-Methoden (FEM) und analytischen Berechnungen simuliert und die Abhängigkeiten von den Randbedingungen mittels numerischer Approximationen beschrieben. Die erzeugten Daten werden als Rohdaten und Regeln in einer Datenbank abgelegt. Durch die Definition von Testsequenzen werden Listen mit Parametervariationen generiert. Mit Hilfe dieser Listen werden aus der Datenbank gelabelte Testdatensätze simulierter Prüffahrten zusammengestellt (Abbildung 1).

Die so gewonnenen Datensätze werden nicht nur projektintern verwendet, sondern auch über die offene Rechercheplattform für mobilitätsbezogene Daten mCLOUD des Bundesministeriums für Digitales und Verkehr für Forschungszwecke zur Verfügung gestellt.

### 3. Zusammenwirken von Simulation, Training der KI und probabilistischer Bewertung

Der Übergang von ZfP-Informationen zu sinnvoll nutzbaren Daten benötigt einen Kontext. Besonders im Fall der Einführung von KI kann dieser Kontext durch den KI-Ansatz, den Operator und durch die Kenntnis der Fähigkeit des Prüfverfahrens geschaffen werden. Daher ist ein wesentlicher Punkt des methodischen Ansatzes zur Bestimmung der Leistungsfähigkeit von Prüfverfahren (Detektion) und der Güte von Algorithmen (Klassifikation) sowie deren Validierung die Bewertung auf Basis probabilistischer Methoden.

Hier wird ein neuartiger Ansatz zur Evaluierung der Performance der Künstlichen Intelligenz verfolgt (siehe Abbildung 2). Die durch Simulationen bereitgestellten balancierten und gelabelten Trainingsdatensätze werden zum Training der KI verwendet und die damit erzielten Ergebnisse der Klassifikation mit Berücksichtigung der probabilistischen Bewertung ausgewertet. Die Ergebnisse werden dann für die Bereitstellung weiterer Trainingsdaten mit veränderten Parametern zur Optimierung und Schärfung der Klassifikation rückgespeist.



**Abb. 2.** Zusammenwirken von Simulation, KI-Training und probabilistischer Bewertung

Ohne die probabilistische Bestimmung der Zuverlässigkeit der ZfP ist eine prädiktive Instandhaltung nicht denkbar, da erst durch die probabilistische Bruchmechanik eine flexible Planung der Instandhaltungsintervalle unter genauer Kenntnis des Prüfverfahrens möglich ist.

Basierend auf dem digitalen Zwilling können relevante Einflüsse für die Zuverlässigkeitsanalyse identifiziert und somit das passende Zuverlässigkeitsbewertungsmodell ausgewählt werden. Gibt es beispielsweise mehrere Fehlerparameter, die die Auffindwahrscheinlichkeit stark beeinflussen, kann von einer monometrischen Auffindwahrscheinlichkeit (Probability of Detection, POD) zu einer multiparametrischen POD gewechselt werden. Auch die Einteilung in Risikogruppen für die Klassifizierungswahrscheinlichkeit (Probability of Classification, POC) wird auf den Kenntnissen des digitalen Zwillings aufbauen.

#### 4. KI-Methoden

In einem nächsten Schritt werden KI-Algorithmen basierend auf verschiedenen Ansätzen entwickelt und trainiert. Die in den Experteninterviews gesammelten Daten dienen zur Definition von Klassifikatoren für das Training der KI (siehe Darstellung des Workflows in **Abb. 2**).

Eine konsequente Bereitstellung gelabelter, balancierter und anonymisierter Datensätze sowohl durch Simulation von Schädigungen und Artefakten (synthetische Datensätze) als auch durch Auswahl von Prüfdaten bekannter Netzabschnitte (reale Datensätze) ist eine wesentliche Neuerung gegenüber den bisherigen Ansätzen.

Mit den Prüfaussagen, dem sogenannten „Feature Engineering“, können dann eine große Menge an Prüfdaten „gelabelt“ werden. Mit den gelabelten Daten können dann Experimente mit geeigneten Methoden des maschinellen Lernens durchgeführt werden, um datenbasierte Modelle zur automatischen Klassifizierung von Fehlerstellen zu trainieren. Ein Teil dieser Experimente wird sich mit den Möglichkeiten der Anwendung von neuartigen Methoden wie z. B. Deep Feedforward Networks oder Convolutional Neural Networks beschäftigen.

## 5. Synergien

Die Ergebnisse der Ultraschall- (UT) und der Wirbelstromprüfung (ET) werden derzeit nicht gemeinsam ausgewertet. Dies ist aus historischer Sicht dem Umstand geschuldet, dass beide Prüfverfahren im Produktivbetrieb unabhängig voneinander eingesetzt werden können.

Da die eingesetzten Prüfverfahren auf Basis unterschiedlicher physikalischer Wechselwirkungen basieren (elektromagnetisch, akustisch), die zudem unterschiedliche Bereiche der Schiene abdecken (Oberfläche, Volumen), können durch eine kombinierte Auswertung der Prüfdaten im Sinne von Datafusion zusätzliche Informationen gewonnen und die Zuverlässigkeit der Prüfaussagen verbessert werden. Auch die Einbeziehung weiterer Daten wie z. B. die Gleisgeometrie, Untergrund, klimatische Bedingungen und Verkehrsbelastung können hierzu einen Beitrag leisten.

Im Gegensatz zu der bisherigen Bewertung der Prüfdaten mit jeweils verfahrensbezogenen Auswertewerkzeugen können durch einen Ansatz mit künstlicher Intelligenz die Daten der verschiedenen Prüfsysteme und zusätzliche streckenbezogene Informationen gemeinsam und automatisiert bewertet werden. Durch eine gemeinsame Befundung der Daten der Ultraschall- und Wirbelstromprüfung kann die Trennschärfe bei der Erkennung verschiedener Schädigungstypen und Artefakte signifikant erhöht werden, z. B. für die Erkennung von Schweißverbindungen in Schienen. Weiterhin können Daten von Schädigungen auf der Fahrfläche mit der Gleisgeometrie abgeglichen werden, um z. B. betriebsbedingte Aufhärtungen zu detektieren.

Bei UT können Oberflächenschäden die Ausbreitung der Ultraschallwellen so beeinflussen, dass eine Klassifizierung nicht mehr möglich ist. Die automatisierte Auswertung erkennt hier nur einen Ankoppel- oder Rückwandecheverlust. Erst durch die Synergie von UT und ET kann auch hier eine Klassifizierung der Anzeigen gelingen [5].

## 6. Risikoklassen

Zusätzlich soll im Projekt erforscht werden, wie gut sich verschiedene Fehlerklassen identifizieren und unter Zuhilfenahme von stochastischen Risikobewertungsmethoden quantifizieren lassen.

Unter Verwendung der Informationen über zu bewertende Fehler kann die Planung der Instandhaltungsmaßnahmen nach risikobasierten Konzepten durchgeführt werden. Dazu soll im Projekt untersucht werden, inwiefern ein auf mathematischer Optimierung basierendes Modell zur Ableitung, Bündelung und zeitlicher Terminierung von Maßnahmen eingesetzt werden kann. Eine Risikobewertung nach Eintrittswahrscheinlichkeit und Auswirkung von klassifizierten Fehlern erfolgt dabei anhand einer Risikomatrix wie in Abbildung 3 dargestellt.

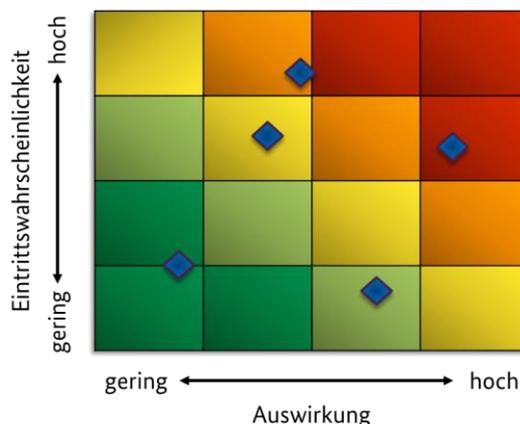


Abb. 3. Risikobewertung nach Eintrittswahrscheinlichkeit und Auswirkung

## 7. Weitere Ziele des Projektes

Neben den beschriebenen konkreten wissenschaftlich-technischen Arbeitszielen verfolgt das Projekt auch übergeordnete Zielstellungen im Einklang mit den Richtlinien des Förderprogramms mFUND: die Erschließung von Nutzungsmöglichkeiten und die Veröffentlichung von Daten zu Forschungszwecken.

Das Projekt beabsichtigt, verbesserte und erweiterte Nutzungsmöglichkeiten von Daten der Schienenverkehrsinfrastruktur zu entwickeln. Ausgehend von bereits heute regelmäßig erhobenen großen Datenmengen aus Schienenprüfzügen wird eine neuartige, datenbasierte Anwendung zur Optimierung von Management und Unterhalt des Schienennetzes bereitgestellt. Die Entwicklung wird wesentlich von der Kombination der Daten aus verschiedenen Quellen (Ultraschall- und Wirbelstromprüfsystemen) profitieren, da erst dadurch erweiterte Erkenntnisse bei der Datenanalyse möglich werden und genauere Aussagen zum Zustand der Schiene sowie notwendiger Instandhaltungsmaßnahmen getroffen werden können. Eine Identifikation zusätzlicher Datenbedarfe und Bewertung hinsichtlich deren potenziellen Nutzens für die Schienenfehlerdetektion und Zustandsbewertung – etwa Daten zur Gleislage, zur Streckenbelastung, zur Umgebung – wird ebenfalls Bestandteil des Projekts.

Eine gemeinsame Aufgabe aller in der mFUND-Förderinitiative bearbeiteten Projekte ist die Generierung von mobilitätsbezogenen Daten und deren Bereitstellung über die offene Rechercheplattform mCLOUD des BMDV. Die über die mCLOUD verfügbaren Daten aus den Bereichen Verkehr, Klima und Wetter, Luft- und Raumfahrt sowie Infrastruktur stehen der Allgemeinheit offen und sollen Forschern und Entwicklern entsprechendes Datenmaterial für innovative Ideen liefern [6]. Die im Projekt AIFRI angestrebte Datenveröffentlichung über die mCLOUD umfasst einen vollständigen Datensatz zu einem ausgewählten Netzabschnitt, beginnend bei Rohdaten aus dem Schienenprüfzug über Ergebnisse der Fehlerdetektion zu abgeleiteten Risikobewertungen. Aus Gründen bestehender Betriebsgeheimnisse seitens des Netzbetreibers DB Netz erfolgt eine Anonymisierung der Daten, um einen konkreten Orts- und Zeitbezug zu entfernen. Zusätzlich werden mittels Simulation gewonnene Trainingsdaten veröffentlicht. Der im Projekt zur Veröffentlichung ausgewählte Netzabschnitt befindet sich vollständig im Lausitzer Braunkohlerevier und ist dem Offenen Digitalen Testfeld des DZSF zuzuordnen [7].

Für die zu verwendenden Formate der Datenveröffentlichung werden im Projekt Vorschläge erarbeitet, die an gängige Standards für Infrastruktur- bzw. Forschungsdaten angelehnt sind, und, wo nötig, diese erweitern (z. B. IDMVU, railML, IDS). Die veröffentlichten anonymisierten und gelabelten Datensätze bieten anderen forschenden Organisationen die Möglichkeit z. B. im Bereich der KI eigene algorithmische Entwicklungen zu validieren. Zwecks einer Erweiterung der Methodik kann so auch der Nutzen weiterer Datenquellen zur Verbesserung der Detektionsergebnisse untersucht werden.

## 8. Zusammenfassung und Ausblick

Das Projekt hat zum Ziel, verbesserte Nutzungsmöglichkeiten von Daten der Schienenverkehrsinfrastruktur zu entwickeln. Ausgehend von bereits heute regelmäßig erhobenen großen Datenmengen aus Schienenprüfzügen mit Ultraschall- und Wirbelstromprüfsystemen wird eine neuartige, datenbasierte Anwendung zur Optimierung von Management und Unterhalt des Schienennetzes bereitgestellt. Die Bewertung der Prüfdaten erfolgt auf Basis künstlicher Intelligenz. Hierbei wird ein neuartiger Ansatz zum

Trainieren der eingesetzten künstlichen Intelligenz angewendet. Das Training erfolgt einerseits mit gelabelten Datensätzen, die mit Hilfe simulierter virtueller Prüffahrten erzeugt werden, andererseits werden zur Bewertung der Ergebnisse der KI und zur Rückkopplung probabilistische Methoden angewendet. Der Automatisierungsgrad der Schienenprüfung wird im Verlauf des Projektes signifikant erhöht. Durch die Einführung von Risikoklassen bei der Bewertung der Prüfdaten kann mit den Ergebnissen des Projektes die Instandhaltung von einem präventiven Ansatz hin zu einem risikobasierten Konzept weiterentwickelt werden.

## Referenzen

- [1] E. Rohwerder, Clusterbasierte Analyse von Daten aus Ultraschall-Schienenprüfung auf Basis der Schienenhöhe, Bachelorarbeit, TU-Berlin, 2020
- [2] M. Neumann Analyse von Ultraschalldaten aus der Schienenprüfung mit Methoden des Supervised Machine Learnings, Bachelorarbeit, TU-Berlin, 2018
- [3] T. Heckel, Y. Wack, G. Mook, Simulation von mechanisierten Eisenbahnschienenprüfungen mit Ultraschall, Tagungsbeitrag, DGZfP Jahrestagung 2019, Friedrichshafen
- [4] T. Vahlsing et al., FE-simulation of eddy current signals produced from basic model cracks for running surface rail defects, Railway Engineering 2019, Edinburgh, UK
- [5] H.-M. Thomas, T. Heckel, G. Hanspach, Advantage of a combined ultrasonic and eddy current examination for railway inspection trains, Insight Vol. 49, 6, pp. 341-344, Juni 2007.
- [6] Bundesministerium für Digitales und Verkehr, mCLOUD – Öffentliche Daten in Bewegung, <https://www.mcloud.de> (abgerufen am 08.04.2022)
- [7] Deutsches Zentrum für Schienenverkehrsforschung, , Das Offene Digitale Testfeld des DZSF, [https://www.dzsf.bund.de/DZSF/DE/DasDZSF/Forschungsinfrastruktur/ODT/odt\\_node.html](https://www.dzsf.bund.de/DZSF/DE/DasDZSF/Forschungsinfrastruktur/ODT/odt_node.html) (abgerufen am 08.04.2022)