

Artificial intelligence in predictive maintenance and fault detection on railway tracks

Amila Keško, Hamdo Puška

¹Faculty of Traffic and Communications, University of Sarajevo, Zmaja od Bosne br. 8, 71000 Sarajevo, Bosnia and Herzegovina

Abstract

Due to the high demand for railway transportation, which is one of the primary modes of passenger and freight transport, we encounter challenges in forming maintenance plans for vehicles and tracks and coordinating train movements on the railway network. Numerous maintenance methods and techniques have been developed, one of which is predictive maintenance. In fault detection through predictive maintenance, the application of artificial intelligence plays a significant role. Its capabilities are particularly exploited in diagnosing faults on railway tracks. Previous research has shown that artificial intelligence is used in predicting train delays, scheduling timetables, enhancing traffic safety, as well as in maintenance and diagnosis of potential faults that may compromise safety. Techniques such as machine learning, deep learning, neural networks, and the Internet of Things are employed for data collection and processing through smart sensors. By collecting and analyzing data, AI has developed innovative models capable of predicting potential faults and detecting current damages on the track. This paper will investigate advanced techniques based on deep neural networks (DNNs) and convolutional neural networks (CNNs) to support real-time fault identification via smart sensors and cameras. The research results contribute to better railway traffic organization, simplified maintenance planning, and minimizing unplanned train traffic disruptions.

Keywords: *artificial intelligence, railway track, predictive maintenance, machine learning*

1 Introduction

Rail transport, due to its characteristics, plays a significant role in passenger and freight transportation. Therefore, its systems influence the functioning, efficiency, service quality of transportation, and the level of safety perception. Safety is an extremely important factor in railway traffic and is also one of the most important criteria for the organization and functioning of the railway as a complex, technical-technological, and dynamic system.

Both fixed and mobile assets must be in proper technical condition to facilitate the movement of trains. Timely detection of faults is necessary to prevent accidents, breakdowns, or disruptions in train operations.

The rapid advancement of artificial intelligence presents unprecedented opportunities for enhancing the performance of various industries and enterprises, including the transportation sector. Artificial intelligence introduces highly advanced computational methods and techniques that mimic the functioning of the human brain. Railway transportation, as a complex dynamic system, strives to stay current by integrating the capabilities offered by artificial intelligence. It is predominantly applied in the areas of achieving autonomous driving, maintenance, and train traffic management.

The literature contains numerous review papers dedicated to the application of artificial

Corresponding author: Amila Keško (amila.kesko@fsk.unsa.ba)

© 2024 The Author(s). This work is licensed under a Creative Commons Attribution 4.0 International License.

intelligence in railway transportation. These papers focus on specific aspects of artificial intelligence relevant to various subsystems of the railway network.

In this paper, we provide an overview of existing literature on the application of deep learning in railway track condition monitoring. Several models are presented, and the quality of their application is still being examined.

2 Related Works

Modern technologies and research methods play an important role in detecting and

maintaining railway faults. Traditional approaches, such as preventive maintenance, are increasingly being supplemented or replaced by advanced techniques that include the use of sensors, machine learning, and big data analysis. These technologies enable earlier detection of potential problems, more precise diagnostics, and more efficient maintenance planning. The review of research methods and technologies for detecting and maintaining faults aims to provide a comprehensive insight into the current state of the art and highlight the latest innovations and future development directions

Table 1. Overview of Research Methods and Technologies for Detecting and Maintaining Railway Faults

Category	Description	Research work
Diagnostics of faults using deep neural networks	Methods for diagnosing faults using deep neural networks.	[1]
Advanced ultrasonic techniques	Utilization of advanced ultrasonic techniques and eddy currents for inspection.	[2]
Computer vision with transfer learning for defect detection on railway tracks	Deep learning transfer models, which combine training on large datasets (ImageNet, COCO) with fine-tuning on specific railway track data, enabling defect detection.	[3]
Maintenance modules utilizing Artificial Neural Networks (ANN) with dynamic time series modelling.	They involve data acquisition, pre-processing, network training, feature extraction, and prediction modules using AI for condition monitoring and failure prediction of ball bearings.	[4]
Monitoring the condition of railway tracks using fault detection and diagnosis (FDD) method based on neuro-fuzzy logic.	The monitoring system is based on neuro-fuzzy logic method for fault detection and diagnosis, aiming to improve the reliability and availability of railway tracks while reducing operational costs.	[5]
System for detecting surface damage on rails.	The paper proposes a new network for detecting surface damage on rails, YOLOv5s-VF, which utilizes the Visual Channel-wise Attention Mechanism (V-CBAM) and Adaptive Spatial Feature Fusion (M-ASFF) to achieve both accuracy and speed in detection.	[6]
Deep learning for signal analysis and fault diagnosis.	In the paper, methods such as CNN with a foundation on Hidden Markov Models, dislocated time series, and 2D CNN are utilized for classification and learning of various fault characteristics. These methods have provided accurate assessment of abnormalities.	[7]
Artificial Neural Networks (ANN) and Support Vector Machine (SVM) methods for predicting failures in railway switches.	In the study, data from railway switch points were collected and analysed to provide insights into the causes of failures and to propose necessary corrective and preventive measures to maintenance personnel.	[8]
Predictive maintenance using Internet of Things (IoT) and Radio-Frequency Identification (RFID).	The paper discusses new trends and techniques in the field of predictive maintenance.	[9]
Implementation of Distributed Acoustic Sensing (DAS) system and semi-supervised learning (SSL) mechanism.	An analysis of new trends and techniques in rail defect detection was conducted, highlighting the advantages of deep learning in detecting multiple types of defects.	[10]

3 Maintenance of Railway Systems

Rail transportation plays a crucial role in the national economy and development and directly impacts the lifestyle of residents. There is a low

tolerance level among the public towards any accidents or negative events that occur in rail transportation. A working group of the European Union Commission has determined what would be an appropriate definition of the term "Railway

System Safety": "Railway safety represents a socially necessary level of absence of risk of danger in the railway transport system, risks related to accidents and incidents resulting in fatalities or serious injuries and significant material damage" [11].

Maintenance, defined as the combination of all technical, administrative, and managerial activities undertaken throughout the lifecycle of an element to support or restore it to a state in which it can perform a specific function [12], is one of the most crucial factors in ensuring the safety of vehicle movement on the tracks.

Safe railway operation requires effective maintenance, which heavily relies on inspection and monitoring of the railway track condition. Monitoring the condition is crucial for the safety, reliability, and cost-effectiveness of railway transportation [13].

To perform maintenance and track condition monitoring, it is necessary to use various means, devices, and methods. When procuring or using new equipment, it must undergo appropriate testing. According to Directive (EU²) 2016/798, each member state establishes a national safety authority responsible for approving and authorizing the use of railway traffic management and control subsystems, signaling and safety subsystems, electro-energetic and construction subsystems, as well as supervising the compliance of constituent parts of interoperability with the basic requirements that constitute the Union's railway system by Directive (EU) 2016/797 [14].

3.1 Diagnostic Methods

The classification of diagnostic methods can be varied and depends on the aspect of observation. The general classification according to the desired objective includes two types of diagnostic methods:

- General diagnostic methods (basic, functional, express)
- Local diagnostic methods (detailed, individual).

A technical system in operation can be in one of two possible states: "operational" or "failed." If it is "operational," it means the system is

functional and capable of performing the assigned task in the prescribed manner and time. If it is in a "failed" state, the task and function are not performed as required [15].

Today, a wide range of materials is used, and they must meet the required quality standards. While materials are not perfect, it is important to detect any defects, and for this purpose, various testing methods are conducted. Materials are tested using destructive (DT) and non-destructive testing (NDT) methods. NDT methods are applied in practice for detecting various types of surface and subsurface defects without damaging the tested object. Non-destructive methods are also known as defectoscopy. These tests include penetrant testing (PT), ultrasonic testing (UT), magnetic testing (MT), visual inspection (VT), and radiographic testing (RT). DT methods are used for testing materials under different loads, where the material is damaged during testing, necessitating additional material samples [16].

3.2 Defects on the Railway Track

Railway systems around the world operate in different environments, with their most critical infrastructure being the steel track, which includes rails, sleepers, ballast, fasteners, and subgrades. Defects on the track and rail failures can lead to unintended consequences such as derailments, fatalities, injuries, economic losses, and loss of public trust. Safe railway traffic requires efficient maintenance, which heavily relies on the inspection and monitoring of track conditions. Monitoring track conditions is fundamentally important for the safety, reliability, and cost-effectiveness of railway operations [13].

Rails are an indispensable element of track construction, accounting for a large portion of the overall costs of building and maintaining railway infrastructure. Detecting defects on the track surface is crucial in railway transportation. However, edge defects on both sides of the rail and multi-scale variation among different defects pose challenges in defect detection [6].

The rail consists of the rail head, rail web, and rail bottom. The distribution of defects in the rail head area includes nuclear defects, rail ends, and welded joints, which are also distributed in the same location. Defects in the rail web area include

² EU – European Union

cracks around the screws, which are located around screw holes, including upper, lower, and horizontal cracks. The distribution of defects in the rail bottom area includes transverse and longitudinal cracks [17].

To utilize a specific model or system, it must comply with regulations and directives, meaning it must be a compatible and interoperable system. Commission Regulation (EU) No 1299/2014 on the technical specification for interoperability relating to the subsystem 'infrastructure' of the rail system within the European Union applies to all new, modernized, or renewed rail infrastructure within the European Union. The elements covered by the TSI³ infrastructure, as described in Directive 2008/57/EC, include the following: track alignment; track parameters, switches, and crossings, track resistance to loads, structure resistance to traffic loads, intervention maintenance limits in case of track geometry damage, platforms; health, safety, and environmental protection, equipment for stable work, servicing facilities for trains [14].

4 Deep Learning and Railway Track Condition Monitoring

Artificial Intelligence (AI) is the study of how to build or program computers to enable them to perform tasks that humans can do [18]. The rapid pace of AI development provides unprecedented opportunities for enhancing the performance of various industries and businesses, including the transportation sector. The application of AI in the field of transportation aims to overcome challenges such as increased travel demand, CO₂ emissions, safety issues, and environmental degradation. Examples of AI methods finding their place in the transportation field include Artificial Neural Networks (ANN), Genetic Algorithms (GA), Simulated Annealing (SA), Artificial Immune System (AIS), Ant Colony Optimization (ACO), Bee Colony Optimization (BCO), as well as Fuzzy Logic Model (FLM) [19].

Computer program-based methods have been explored in the past for detecting defects on railway tracks, but fully automated monitoring has always been a challenge because traditional image processing methods and deep learning models fail to perform well enough due to limited data availability. Some models effectively detect defects on tracks such as cracks, looseness, some signaling damages, and so on. These models have been tested on different continents to assess their performance under various weather conditions [3].

Deep learning methods have achieved success in various applications with recent advances in techniques. Deep learning utilizes neural networks as functional units that mimic how the human brain solves complex problems based on data. Methods such as Long Short-Term Memory (LSTM) as a type of Recurrent Neural Network (RNN) and Convolutional Neural Networks (CNN) have driven the development of deep learning and artificial intelligence fields and have been reported to show convincing performances in monitoring conditions for tools, machinery, and turbines. The effectiveness of prediction and learning of these methods improves with the increase in available data volume. Deep learning methods have been adopted for monitoring the condition of railway tracks and detecting and classifying anomalies [13].

4.1 InSAR and GPR for Infrastructure Monitoring

SAR⁴, a satellite remote sensing tool, continuously monitors ground deformations around infrastructure, including railway tracks, using the amplitude and phase of recorded electromagnetic wave signals. Temporal variations in the phase recorded by sensors over time are used to distinguish changes in the topography of the Earth's surface, through a tool known as SAR interferometry (InSAR), resulting in the creation of interferograms and deformation maps. The flexibility of this technique enables monitoring of various conditions, with regular data updates and global coverage.

³ TSI – Technical Specifications for Interoperability

⁴ SAR – Synthetic Aperture Radar

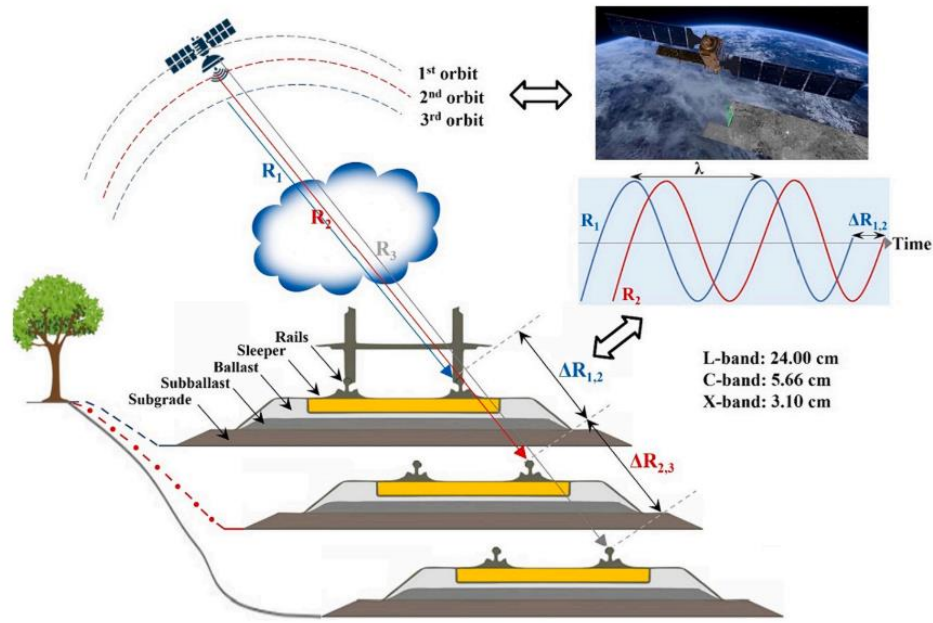


Fig. 1 Schematic representation of ground surface displacement extraction near a railway track using the InSAR technique [20]

GPR, or Ground Penetrating Radar, is a device that utilizes electromagnetic waves to probe subsurface conditions of various infrastructures, including layers of ballast, subgrades, and embankments of railway tracks. Analysis of GPR data allows for the recognition of changes in reflection intensity and time of maximum reflection intensity, facilitating the identification of thickness and variations in dielectric constant, demonstrating the feasibility of GPR application in assessing underground layers of railway tracks.

Koohmishi, Kaewunruen, Chang, and Guo explored the application of InSAR and GPR fusion in railway track monitoring, along with the integration of machine learning, data processing, and prediction, focusing on condition-based maintenance of railway tracks. The fusion of InSAR and GPR proves useful for locating the technical limitations of standalone health monitoring methods. Continuous measurements by InSAR can produce large amounts of data that can be processed by machine learning to formulate predictive models. Further development of integrating these techniques is necessary for monitoring and detecting defects and anomalies along railway tracks [20].

4.2 DAS – Distributed Acoustic Sensing

DAS is considered an ideal model for railway track detection using deep learning methods. Deep learning methods are highly popular due to their high sampling frequency and low implementation costs. The use of a spare optical cable for video enables detection without additional sensor installation along the track. DAS records the amplitude of vibrations along the optical cable, allowing for the analysis of defects such as cracks, voids, protrusions, and fractures, using specific vibrations as "defects."

This detection method using semi-supervised deep learning achieves an accuracy of 97.91%. Although traditional methods show better spatial accuracy, this new method opens up possibilities for expanding datasets in engineering situations. For DAS, higher spatial resolution may result in higher costs, but integrating defect detection systems can improve the actual sampling rate and accuracy. However, the drawback of this method is the inability to assess the severity of defects and their evolution, suggesting the need to combine deep learning with traditional machine learning methods to explain the detection process. Integrating deep learning into engineering tasks requires adapting existing models and finding more suitable ways to represent data, which will be crucial for the future of this technology [10].

4.3 The application of CNN and RNN in railway maintenance

The most commonly mentioned neural network names are CNN and RNN. CNN, also known as ConvNet, is inspired by the organization of the visual cortex and resembles the way neurons are connected in the human brain. Individual neurons respond to stimuli within a limited area of the visual field, known as the receptive field, and the collection of such fields overlaps to cover the entire visual area. CNN processes input images, assigning importance to different aspects/objects in the image, enabling their distinction.

RNN, derived from feedforward neural networks, has nodes connected in a directed graph along a time sequence to demonstrate temporal dynamic behavior. The internal states (memory) of RNN are used to process sequences of inputs of variable length. A typical RNN architecture is LSTM⁵, which has feedback connections and can process individual data points (such as images) and entire data sequences (such as speech or video).

Applications of CNN and RNN for railway maintenance operations are common, but CNN is more widely accepted. Detection and classification of defects are common tasks, and images are the most common type of input data. Traditional methods are combined with deep learning for track condition analysis. CNN is popular for feature extraction, while RNN/LSTM is useful for sequential data. The purpose of this research is focused on defect detection, classification, and localization on the track surface, including various components such as rails, insulators, valves, fasteners, and switches. Different deep learning methods show promising results, indicating the increasing popularity and usefulness of deep learning in track condition monitoring [13].

4.4 TrackNet

Deep convolutional neural networks are popular models for image classification and error detection. State-of-the-art convolutional neural networks like ResNet and DenseNet have been adopted as fundamental techniques for performance comparison with the proposed TrackNet method.

The technique called TrackNet is a multi-phase approach to deep learning that integrates rail segmentation and classification of true/false alarms. This is achieved through two neural networks, where one handles semantic rail segmentation, and the other classifies segmented images as true or false alarms.

For semantic segmentation, the U-Net architecture, a convolutional neural network architecture that enables precise localization through a symmetric expanding path, is utilized. Proper shortcuts are employed to preserve pixel-level information for different image resolutions.

After initial segmentation by U-Net and extraction of regions of interest (ROI), an image processing tool is used to crop out a portion of potentially faulty regions in images marked with rectangular frames. The size of the cropping window is approximately 64x64 pixels.

The cropped images are saved and forwarded to the second phase for classification as true or false alarms. This new technique enables the integration of rail segmentation and alarm classification through deep learning. In VTIS, images are first captured via the image acquisition subsystem and then forwarded to the image analysis subsystem for detection and classification of anomalies on the rail surface. Defects are then analyzed by a human controller, where it has been observed that there is a large number of false alarms. An example of a true and false alarm (caused by graffiti on the rail) labeled by the COTS VTIS is shown in Fig. 2.

⁵ LSTM – Long Short-Term Memory

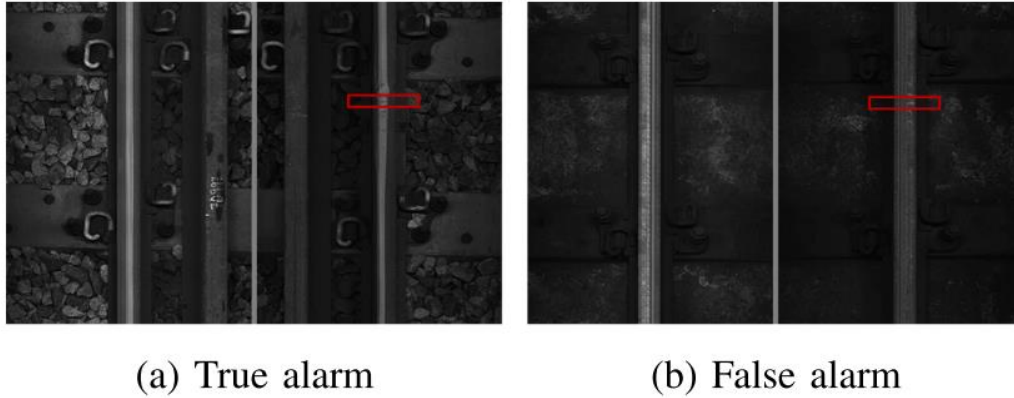


Fig. 2 Images of true and false alarms from a COTS VTIS [2]

In the final phase of TrackNet, cropped images with ROI are forwarded to a neural network that classifies the images as true or false alarms. For image classification, ResNet and DenseNet are used, which are currently leading architectures in convolutional neural networks for image classification. Among them, DenseNet has several compelling advantages, such as alleviating the vanishing gradient problem, enhancing feature propagation, promoting feature reuse, and significantly reducing the number of parameters. This allows for the optimization of very deep neural networks that are both efficient and robust [2].

5 Conclusion

In this research, we explored the application of advanced artificial intelligence techniques in the context of railway transportation, focusing on diagnostic methods, maintenance of railway systems, and monitoring the condition of railway infrastructure. Through the analysis of deep learning applications, such as Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN), as well as innovative approaches like TrackNet, we have determined that these techniques can provide significant advantages in improving the efficiency, reliability, and safety of railway transportation.

Our research has shown that deep learning enables efficient modeling and analysis of large volumes of data generated from railway systems, facilitating rapid detection and prediction of issues in infrastructure or vehicles. Specifically, CNN and RNN offer powerful tools for processing images and time series data, enabling

precise real-time monitoring of infrastructure and vehicles.

Furthermore, the application of innovative approaches like TrackNet demonstrates the potential for personalized diagnostics and monitoring tailored to the specific needs of railway systems. These results underscore the potential of artificial intelligence to revolutionize railway transportation, laying the groundwork for the development of smart, autonomous systems that can enhance the performance and safety of railway transport in the future.

References

- [1] J. Yin and W. Zhao, "Fault diagnosis network design for vehicle on-board equipments of high-speed railway: A deep learning approach," *Eng. Appl. Artif. Intell.*, vol. 56, no. October, pp. 250–259, 2016, doi: 10.1016/j.engappai.2016.10.002.
- [2] A. James *et al.*, "TrackNet - A Deep Learning Based Fault Detection for Railway Track Inspection," *2018 Int. Conf. Intell. Rail Transp. ICIRT 2018*, 2018.
- [3] S. Mittal and D. Rao, "Vision Based Railway Track Monitoring using Deep Learning," 2017, [Online]. Available: <http://arxiv.org/abs/1711.06423>
- [4] I. Daniyan, K. Mpofu, M. Oyesola, B. Ramatsetse, and A. Adeodu, "Artificial intelligence for predictive maintenance in the railcar learning factories," *Procedia Manuf.*, vol. 45, no. 2019, pp. 13–18, 2020, [Online]. Available: <https://doi.org/10.1016/j.promfg.2020.04.032>
- [5] J. Chen, C. Roberts, and P. Weston, "Fault detection and diagnosis for railway track

- circuits using neuro-fuzzy systems,” *Control Eng. Pract.*, vol. 16, no. 5, pp. 585–596, 2008.
- [6] M. Wang, K. Li, X. Zhu, and Y. Zhao, “Detection of Surface Defects on Railway Tracks Based on Deep Learning,” *IEEE Access*, vol. 10, no. November, 2022.
- [7] C. Chellaswamy, S. Ponraj, and K. Venkatachalam, “Deep learning based intelligent rail track health monitoring system,” *Int. J. Innov. Technol. Explor. Eng.*, vol. 8, no. 12, pp. 693–702, 2019.
- [8] B. Arslan and H. Tiryaki, “Prediction of railway switch point failures by artificial intelligence methods,” *Turkish J. Electr. Eng. Comput. Sci.*, vol. 28, no. 2, pp. 1044–1058, 2020.
- [9] S. Selcuk, “Predictive maintenance, its implementation and latest trends,” *Inst. Mech. Eng.*, 2016.
- [10] S. Wang, F. Liu, and B. Liu, “Semi-Supervised Deep Learning in High-Speed Railway Track Detection Based on Distributed Fiber Acoustic Sensing,” *Sensors*, vol. 22, no. 2, p. 27, 2022.
- [11] G. Rosenthal, “Economic and Social Council - Safety and security in railway transport,” *Oxford Handb. United Nations*, 2009.
- [12] B. Gojković, R. Đuričić, V. Malčić, and N. Čabrić, *Održavanje željezničkih vozila*. Saobraćajni fakultet Doboj, 2021.
- [13] A. Ji, W. L. Woo, E. W. L. Wong, and Y. T. Quek, “Rail track condition monitoring: a review on deep learning approaches,” *Intell. Robot.*, vol. 1, no. 2, pp. 151–175, 2021.
- [14] The European Commission, “Commission Regulation EU No 1299/2014,” *Off. J. Eur. Union*, vol. 356, no. 881, pp. 1–109, 2012.
- [15] H. Avdić and D. Tufekčić, *Terotehnologija i. Tuzla: “PrintCom” d.o.o. grafički inženjering Tuzla*, 2007.
- [16] M. Balog and L. Knapčíková, “Smart techniques for technical conditions monitoring of railway wagons,” Prešov, Slovak Republic, 2018.
- [17] F. Wu, Q. Li, S. Li, and T. Wu, “Train rail defect classification detection and its parameters learning method,” *Meas. J. Int. Meas. Confed.*, vol. 151, p. 12, 2020, [Online]. Available: <https://doi.org/10.1016/j.measurement.2019.107246>
- [18] M. A. Boden, *Artificial Intelligence*. Brighton, England: School of Cognitive and Computing Sciences, University of Sussex. [Online]. Available: [https://books.google.ba/books?id=_ixmRIL9jcIC&lpg=PP1&ots=JRKJWSvEUT&dq=artificial intelligence&lr&hl=hr&pg=PR15#v=onepage&q=artificial intelligence&f=false](https://books.google.ba/books?id=_ixmRIL9jcIC&lpg=PP1&ots=JRKJWSvEUT&dq=artificial%20intelligence&lr&hl=hr&pg=PR15#v=onepage&q=artificial%20intelligence&f=false)
- [19] R. Abduljabbar, H. Dia, S. Liyanage, and S. A. Bagloee, “Applications of Artificial Intelligence in Transport: An Overview,” p. 24, 2019.
- [20] M. Koohmishi, S. Kaewunruen, L. Chang, and Y. Guo, “Advancing railway track health monitoring: Integrating GPR, InSAR and machine learning for enhanced asset management,” 2024.