


Review

A study on the application of convolutional neural networks for the maintenance of railway tracks

Mauro José Pappaterra¹  · María Lucía Pappaterra²  · Francesco Flammini^{3,4} 

Received: 11 December 2023 / Accepted: 28 April 2024

Published online: 02 May 2024

© The Author(s) 2024 

Abstract

This paper provides an overview of the applications of Convolutional Neural Networks (CNN) in the railway maintenance industry. Our research covers specifically the subdomain of railway track maintenance. In this study, we have analyzed the state-of-the-art of CNNs applied to railway track maintenance by conducting an extensive literature review, summarizing different tasks and problems related to the topic and presenting solutions based on CNNs with a special emphasis on the data used to create these models. The results of our research show different applications of CNNs within the scope, including the detection of defects in the surface of railway rails and railway track components, such as fasteners, joints, sleepers, switches and crossings, as well as the recognition of track components, and the continuous monitoring of railway tracks. The architecture of CNNs is fitting to learning spatial hierarchies of features directly from the input data, making them of great use for Computer Vision and other applications related to the topic at hand. The implementation of IoT devices and smart sensors aid the collection of real-time data which can be used to feed powerful CNN models to recognize patterns and identify complex events related to the maintenance of railway tracks. This and more insights are discussed in detail within the contents of this paper.

Keywords Railways · Maintenance · Artificial Intelligence · Convolutional Neural Networks · Literature review · Datasets

1 Introduction

Since the beginning of the decade, there has been a noticeably surge of interest in the research on the application of Artificial Intelligence (AI) technology for the maintenance of railways. The railway industry is a wide domain, and there are many tasks related to its maintenance that can greatly benefit from different AI applications [1]. In our previous research, we have investigated the application of these models in in different railway maintenance-related application, including railway tracks maintenance, rolling stock maintenance, infrastructure maintenance, monitoring, and maintenance planning [1]. We have also noticed that different types of data are implemented to feed AI models and an observable shift from—more traditional—symbolic AI models to Machine Learning (ML) and numerical models, in which different types of Artificial Neural Networks (ANN) have been implemented. This literature review aims for an overview of the applications of Convolutional Neural Networks (CNN) for railway track maintenance.

✉ Francesco Flammini, francesco.flammini@supsi.ch; francesco.flammini@mdu.se; Mauro José Pappaterra, mauro.pappaterra@ieee.com; María Lucía Pappaterra, lucia.pappaterra@unc.edu.ar | ¹Department of Computer Science, Uppsala University, Uppsala, Sweden. ²Facultad de Matemática, Astronomía, Física y Computación (FaMAF), Universidad Nacional de Córdoba (UNC), Córdoba, Argentina. ³IDSIA USI-SUPSI, University of Applied Sciences and Arts of Southern Switzerland, Lugano, Switzerland. ⁴School of Innovation, Design and Engineering (IDT), Mälardalen University, Eskilstuna, Sweden.



Fig. 1 A cross-section of a typical railway track

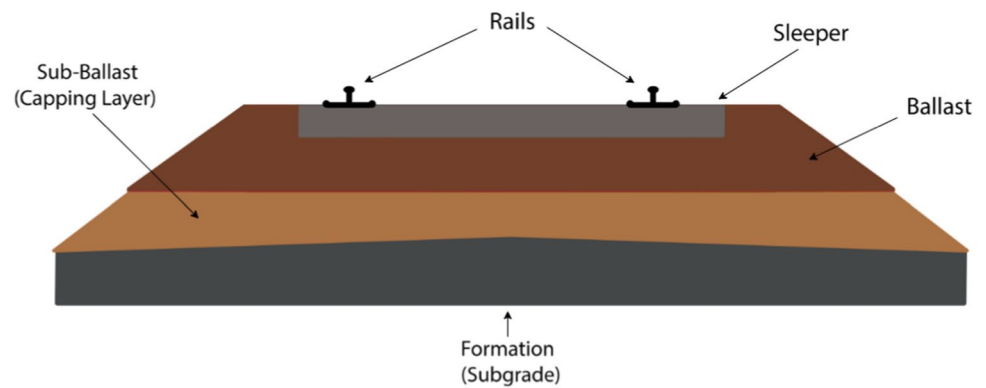


Fig. 2 A traditional ballasted railroad



1.1 Similar studies

We have consulted several literature reviews on specific AI applications in the railway industry. Some reviews contemplate different railway subdomains, including maintenance, safety and security, traffic planning, scheduling, logistics, optimization and more [2–4]. Other studies focus on very specific maintenance applications and AI methodologies. For instance, different authors have surveyed the literature on ML applications for track maintenance [5] and detection of wheel defects [6]; the study [7] presents an overview on fault detection in railway switch and crossing systems; other authors [8] investigated image processing approaches for track inspection; another study [9] presents a survey on the utilization of Deep Learning (DL) and audio and video sensors for railway maintenance. Our scope is specific to the implementation of CNNs techniques for different maintenance tasks in railway tracks, including not only those solutions related to Computer Vision (CV) and image and video data, but any applications of CNN with any type of data related to the preservation, inspection, maintenance and monitoring of railway tracks and their components.

1.2 Railway tracks

The traditional railway tracks or railroads consist of a set of parallel steel bars laid on a formation (earthwork) that provide a stable and continuous surface for the train trajectory. These are the most commonly implemented railroads worldwide (Fig. 1).

The railway consists of different components, including the rails (horizontal beams that provide a smooth surface for the wheels to roll upon); the sleepers (a fixed rectangular support to hold the rails in place perpendicularly while keeping a gauge distance, and bearing the train load to the track ballast); ballast (a layer of crushed stones or gravel implemented to bear the load from the railroad); fasteners (used to link rails with the sleepers and other railway parts); rail joints (implemented to hold adjoining rail beams) and Switches and Crossings (S&C) (movable sections that guide trains from one track to another allowing them to cross paths) [2] (Fig. 2).

In modern railways there exists different types and configurations of railroads. For instance, ballastless tracks or slab tracks replace the sleepers and ballast with a surface of concrete or asphalt. These types of railway tracks are mostly implemented in High-Speed Railways (HSR), metro lines and light rail; and concrete is the most implemented material in their construction [10] (Figs. 3, 4).

1.3 Convolutional neural networks

A Convolutional Neural Network (CNN) is a type of Artificial Neural Network that is specifically designed for Deep Learning and Computer Vision. CNNs imitate the human brain visual cortex by differentiating between parts and layers of an image through the application of relevant filters (kernels) and then adding each of the elements through the process of convolution to pass input to the next layer in the network. CNNs are widely used for image and video recognition as they are capable of recognizing different aspects and objects present in an image frame, they are also implemented in other fields such as Natural Language Processing (NLP) and time-series sequence analysis [2, 11] (Fig. 5).

The key components of a CNN are the convolutional layers (which apply convolution operations to the input data), pooling layers (used to down sample the spatial dimensions of the input data), activation function (applied after convolutional and pooling layers to add non-linearity), fully connected layers (connecting every neuron among layers), flattening (used to flatten feature maps into a one-dimensional vector) [11].

There are different types of CNN that present a wide variety of combinations, components, and architectures. Some of the most common CNN architectures today include: Region-based CNNs (R-CNN), Mask R-CNNs, Very Deep CNNs (VGGNet), Residual CNNs (ResNet), You Only Look Once (YOLO), GoogleNet, Encoder CNNs and more.

With the significant advance of CNNs in the last couple of years [1, 2] and an architecture that allows to learn spatial hierarchies of features automatically and adaptively from raw input data; CNNs become an ideal candidate for complex tasks, such as the one related to our topic of research.

Fig. 3 A cross-section of a modern ballastless rail track

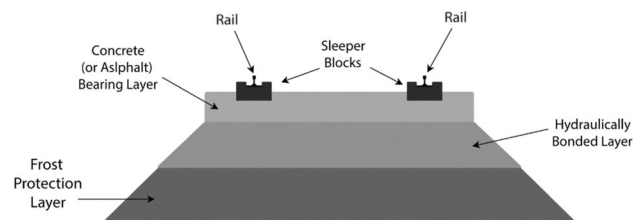
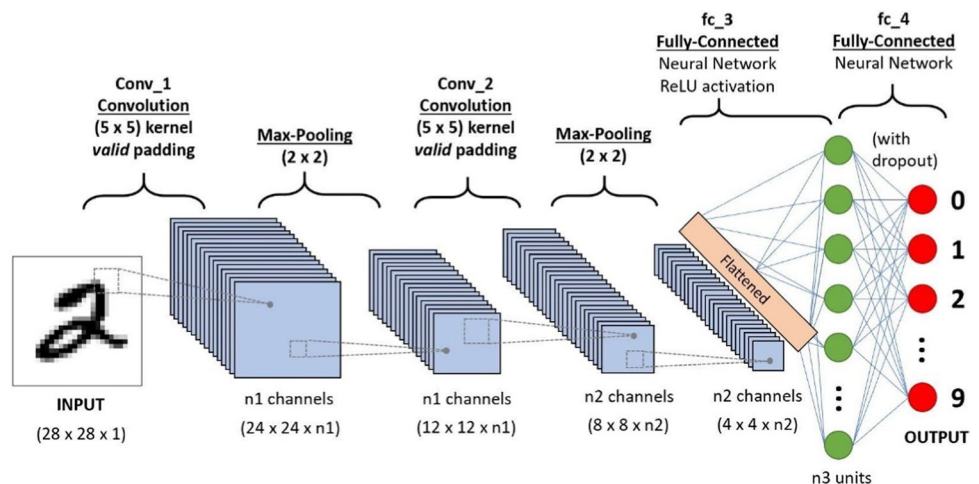


Fig. 4 A high-speed railway slab track



Fig. 5 Diagram of an example of a Convolutional Neural Network. (source: towardsdatascience.com [11])



2 Methodology

For this study, a wide variety of papers on the subject of CNNs applied to the maintenance of railway tracks were consulted. The portals used for the search include: IEEE Xplore Digital Library, SpringerLink, ACM Digital Library, ScienceDirect and Elsevier Scopus. These sources are briefly described in Table 1.

The search terms used to find the papers are a combination of the following key words: *AI, artificial intelligence, railway, railway tracks, railroads, convolutional neural networks, CNN, maintenance, predictive maintenance, condition-based maintenance, corrective maintenance, fault detection, fault diagnosis, and fault prognosis*. These keywords were combined alternately in different order. The searches were performed from June to August 2023.

For the selection of papers, the study's main content needed to be centered on the application of DL methods to solve any task related to the preservation, inspection, maintenance and monitoring of railway tracks. For the scope of this review, the range of the years of publication was set from 1990 to 2022. Papers that presented better quality (relevant content, clear methodology, critical analysis, reproducibility, sound interpretation and documentation, innovative ideas, etc.) were prioritized over papers of inferior characteristics. So is the case for papers that were more referenced from reputable sources or were published in creditable journals.

After applying this inclusion/exclusion criteria we collected a total of 71 papers from approximately ~ 120 preselected papers. Many papers were excluded because they did not present an AI solution based in CNN, they were not centered in railway track maintenance or simply did not pass our aforementioned criteria in any other way.

3 Literature review

This section includes the complete literature review. The full list of papers that have been examined can be found in the *References* section.

3.1 Papers overview

The maintenance and monitoring of railroads are a crucial aspect of railway infrastructure maintenance. We have identified a total of 71 papers on the application of CNNs within this domain. The following tasks or track maintenance subdomains were elicited based on the content of the papers: *surface defect detection, fasteners, joints, sleepers, switches and crossings, detection of objects and components and monitoring*. Notice that the scope of some papers can cover more than one of the subdomains enumerated resulting in some overlap. Table 2 displays the list of maintenance tasks and the papers identified; Table 3 displays the types of data found in the literature; and Figs. 6, 7 and 8 the distribution of papers according to year, maintenance task and data implemented.

3.2 Papers review

A review of the selected papers is presented below. The papers are divided into subsections according to the classification presented in Sect. 3.1.

Table 1 Search portals consulted for the collection of literature

Search Portal	Short Description	URL
ACM Digital Library	An academic research platform from the Association for Computing Machinery (ACM)	https://dl.acm.org/
Elsevier Scopus	Elsevier's abstract and citation database curated by experts in different fields	https://www.scopus.com/
IEEE Xplore Digital Library	IEEE's research database for academic publications	https://ieeexplore.ieee.org/
ScienceDirect	Elsevier's database of scientific publications	https://www.sciencedirect.com/
SpringerLink	Search platform for Springer academic publications	https://link.springer.com/

Table 2 Papers divided by railway track maintenance tasks

Maintenance Task/Area	Citations
Surface defect detection	[12–16, 18–39, 41–56]
Fasteners	[57–65]
Joints	[48, 66]
Sleepers	[59, 67–69]
Switches and crossings	[48, 70, 71, 73–78]
Detection of objects and components	[16, 28, 62, 64, 65, 79–86]
Monitoring	[22, 26, 30, 42, 45, 66, 73, 80]

Table 3 Papers from the literature review divided by type of data implemented

Type of data	Citations
Acoustic and sound data	[54, 55]
Image and video	[12–16, 18–22, 24, 25, 27–39, 41–45, 52, 57–60, 63–67, 69, 70, 73, 74, 79, 80, 82–86]
Speed and acceleration data	[46–50, 68, 77, 78]
Track geometry	[23]
Temperature	[50]
Vibration measurements	[43, 49, 51–53, 78, 81]
Electrical signals	[56, 75, 76]

Fig. 6 Number of papers in the literature review divided by year of publication

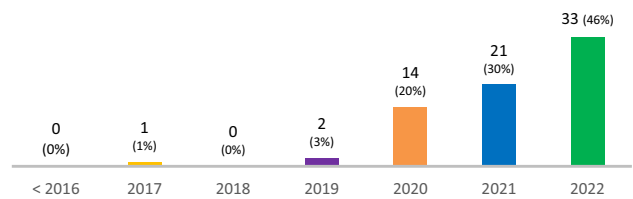
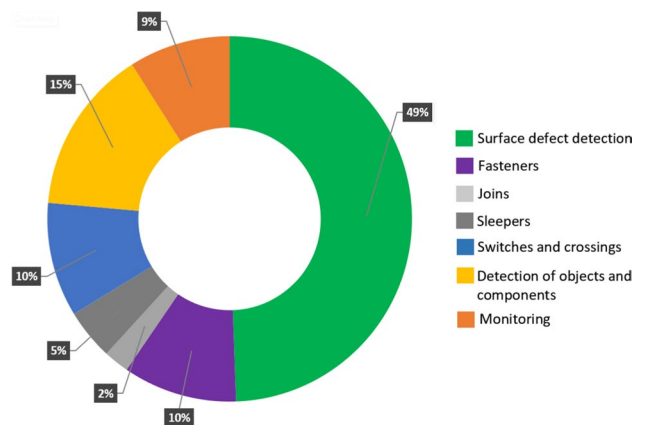


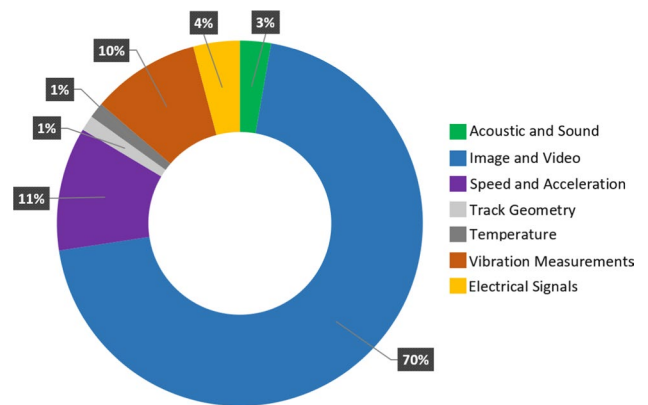
Fig. 7 Percentage of papers according to maintenance task



3.2.1 Surface defect detection

The maintenance of rails is crucial for enabling efficient transportation in railroads. One of the most prominent tasks in the literature is the inspection, early detection of faults and degradation analysis of railway tracks. CNNs have been

Fig. 8 Percentage of papers according to data implemented



successfully implemented in different CV and image classification studies related to Fault Detection (FD) contributing to the ongoing efforts to automate and improve railway track inspection processes.

3.2.1.1 Image-based fault detection using CNNs The authors in [12] address the challenge of detecting cracks on railway rail surfaces. They present a solution that involves the amalgamation of CNNs with image augmentation and transfer learning with VGG16 networks. This combined approach is applied to classify rail surfaces as *defective* (with cracks) or *non-defective*. In reference [13], the authors propose a CV-based system for efficient detection of cracks in rails using images captured by a rolling camera beneath a self-moving railway vehicle. The approach involves pre-processing, Gabor transform application, and extraction of first-order statistical features. These features serve as input for a DL ANN to differentiate between cracked and non-cracked track images. Another research presented in [14] employs a Faster Region-based Convolutional Neural Network (Faster R-CNN) for the detection of rail surface defects. The process includes establishing a dataset of rail surface defect images, creating a training set with random segmentation, and validating the model. In reference [15], the authors present a DL-powered rail surface multi-flaw identification framework. The framework consists of two main components: a rail extractor for isolating rails from the background and a cascading rail surface flaw. The latter includes a rail detector for recognizing the health status of rail surfaces and a flaw classifier for various flaw types. The framework incorporates feature joint learning from various feature extractors to enhance its accuracy. Current automated rail surface defect detection methods mainly use CNNs, but they often overlook early issues like rail surface cracks. To fill this gap, the authors in reference [16] propose a multitask learning model. It has a segmentation decoder for identifying crack defects and a decoder for detecting rail objects. During training, the object detection task improves the accuracy of the segmentation model. In testing, the object detection task doesn't slow down the process, ensuring swift inference. The model is evaluated on the Rail-5 k dataset [17], showing a good balance between accuracy and efficiency.

The study [18] introduces an image classification-based approach for flaw detection in railway tracks. They propose two models, a fully convolutional encoder-decoder type segmentation network, called U-Net, and dilated convolutions, known for their effectiveness in addressing pixel-based problems rather than class-based problems. These models aim to separate pixels containing flaws from background images to enhance the detection of railway track flaws. In reference [19], the authors utilize GoogleNet's CNN architecture to automatically detect rail defects, aiming to speed up the inspection process and reduce the risk of railway accidents. They trained the model using 2000 images categorized into two classes: *broken* and *intact*. The authors in [20] present the use of EfficientNet CNN architecture on railway track fault detection and classification as a novel technique in the field. The paper compares B0 to B7 network families and studies the effect of input image resolution on a small dataset. In reference [21], the authors introduce a high-performance fine-tuning CNN model for the detection of time or impact-dependent defects on railway surfaces used for train transportation. Their approach involves two stages: first, they focus on cropped images of the train tracks rather than large-area rail images. In the second stage, various CNNs models are applied to the cropped images for classification. Another study [22] introduces a system aimed to identify and detect faulty railway tracks. The system is built upon DL algorithms and conducts a comprehensive analysis of images depicting railway tracks to determine their condition. This innovative approach ensures high accuracy in fault detection while minimizing the utilization of features, addressing the limitations of previous surveillance systems. Squat defects are produced by wheel-rail dynamic impact, leaving a depression mark on the rail surface. In reference [23], the DL VGG16 network is used to classify thousands of kilometers of railway track into two binary classes: *squat(p)* and *no-squat(n)*. In order to mitigate large class imbalance, they considered several

natural sampling and data augmentation methods; which yielded a considerable improvement to traditional re-sampling approaches. In reference [24], the authors introduce a pixel-level defects segmentation method to enhance the detection and categorization of surface defects on railway tracks. Their network tessellates features at the channel level for denser information propagation across high-resolution layers. Dropout is used to address weak correlations during convolution, reducing computational redundancy and complexity. The study categorizes track datasets, preprocesses samples to grayscale, and trains the proposed network. Another study [25] introduces an onboard image detection system for high-speed magnetic levitation (Maglev) tracks, focusing on long stator tracks. The system captures accurate images under challenging conditions, addressing issues like limited space, low illumination, and fast vehicle movement. To overcome the scarcity of maglev track images, especially those with defects, they present a data enhancement method involving sample generation and image fusion. The research employs DL-based target detection algorithms for automatic detection, classification, and location of defects, specifically on the stator surface and cables.

3.2.1.2 You only look once (YOLO) The YOLO Deep Neural Network (DNN) provides real-time object detection with a notorious velocity and accuracy in comparison to many other CNN-based object detection algorithms. Authors in [26] propose a modified version of the YOLO algorithm, referred to as YOLOv3-M. This model is employed for the detection of railway track defects, and it is utilized for the monitoring of rail health. In reference [27], the authors introduce a YOLOv4-based DL algorithm, using a modified CSPDarknet-53 backbone for improved and quicker defect detection on complex rail surfaces. The study assesses the performance of the algorithm across six diverse datasets, comparing it with an alternative network to evaluate efficiency against existing techniques. In another study [28], the authors introduce YOLOv5s-VF, a rail surface defect detection network. It features a Sharpening Functional Attention Mechanism (V-CBAM) for efficient attention and a Microscale Adaptive Spatial Feature Fusion (M-ASFF) for capturing fine details. Notably, they manually collected real rail images to create an open-source object detection network, addressing the lack of publicly available labeled datasets for rail surface defects. While in [29], the authors introduce a lightweight method for detecting defects on bimodal rail surfaces, which they name Parallel-YOLOv4-Tiny. This method incorporates a parallel feature processing strategy designed specifically for extracting features from bimodal images. Experiments show that Parallel-YOLOv4-Tiny with CIoU loss function can achieve a much better performance than original YOLOv4-Tiny algorithm methods. Lastly, the study presented in [30] introduces intelligent methods for multi-target defect identification in railway tracks using image processing and DL techniques. Their contributions include a track and fastener positioning method, a Bag of Visual Words (BoVW) model for defect detection, an improved YOLOv3 network named Track Line Multi-target Defect Detection Network (TLMDDNet), and a lightweight design strategy named Dense Connection-Based TLMDDNet (DC-TLMDDNet), which employs DenseNet to optimize feature extraction layers. These methods demonstrate potential in enhancing railway track defect detection.

3.2.1.3 Integration of CNNs with other types of AI Other authors implement different types of CNNs combined with other ANN and DL architectures. The authors in [31] propose a novel approach that integrates CNNs with Transformers, known for its self-attention mechanisms and global information processing capabilities. This combined approach is employed to detect *Type-I* and *Type-II* rail surface defects. In another study [32], the authors focus on classifying railway shelling defects using DL techniques, specifically the Residual Convolutional Neural Network (ResNet). Their ResNet model, with 41 convolutional layers and a residual learning block, outperforms other classifiers such as Very Deep Convolutional Networks (VGGNet) and Support Vector Machines (SVM) using various feature extraction methods. The ResNet achieves better accuracy on the testing dataset, surpassing VGGNet. Another study [33] presents a feature extraction module and a track surface defect identification framework based on an enhanced ResNet. The model comprises three modules: feature extraction, ROI extraction, and defect identification. It presents an improved ResNet-based feature extraction method. In reference [34], the authors present an efficient damage detection method for HSR rails, referred to as SCueU-Net. This method combines the U-Net graph segmentation network with the saliency cues approach for damage localization. Experimental results demonstrate that SCueU-Net achieves a high detection accuracy rate, outperforming recent methods in damage identification accuracy. The authors in [35] present the Collaborative Learning Attention Network (CLANet), for inspecting non-service rail surface defects. CLANet focuses on accurate defect identification and segmentation through three stages: feature extraction, cross-modal information fusion, and defect location and segmentation. It introduces a multimodal attention block and a dual-stream decoder to enhance feature representation and prevent information dilution during decoding. To address data scarcity, they created the NEU RSDDS-AUG dataset [35] and conducted a comparative analysis against nine existing methods, demonstrating the effectiveness of CLANet. CLANet also performs competitively across four public benchmark datasets. In reference [36], the authors introduce the Depth

Repeated-Enhancement RGB (DRER-Net) for rail surface defect inspection. This network optimally utilizes depth and RGB information in an encoder-decoder architecture to enhance defect inspection. The encoder incorporates a novel cross-modality enhancement fusion module, merging details from RGB maps and location information from depth maps. In the decoder, a multimodality complementation module repeatedly refines the DRER-Net prediction using details and location information. Extensive experiments compare DRER-Net with 10 state-of-the-art methods on the industrial NEU RSDDS-AUG RGB-depth dataset. In reference [37], the authors detected rail surface defects by fusing the features of two DL models, SqueezeNet and MobileNetV2. These models were chosen because they are smaller and faster compared to other DL models, albeit slightly less accurate. To address this accuracy issue, the authors propose a fusion model that combines the high-weighted features from both models. The process begins with a contrast adjustment applied to the original rail image, followed by determining the rail track's location. The most weighted features from each network are selected, and these reduced features are then used to identify defects with SVMs. Experimental results indicate that this method outperforms using a single DL model, particularly for detecting multiple rail surface defects under low-contrast conditions. The study [38] proposes a Sequential Updatable framework for Anomaly Detection (SUAD) on tracks. The model continues learning through a sequential knowledge update module without revisiting old data. This update is done using new information from false alarms, employing the Robbin-Monro algorithm and a swift Mahalanobis distance variant. The fast Mahalanobis distance calculation is based on principal component analysis, resulting in a quicker inference and a more compact model. SUAD's performance is evaluated with the Metro Anomaly Detection (MAD) dataset (build for the study) and three public datasets. In [39], a novel inspection scheme for Rail Surface Defects (RSDs) is introduced, utilizing limited samples with a line-level label. The approach treats defect images as sequence data, simplifying the labeling task by classifying pixel lines. The scheme comprises two methods: OC-IAN for express rail defects and OC-TD for common/heavy rail defects. Both methods leverage One-dimensional Convolutional Neural Networks (1-D CNN) for feature extraction and Long Short-Term Memory (LSTM) networks for context information based on Interactive Attention Network (IAN) and Target Dependent (TD) LSTMs respectively. OC-IAN adopts a single-branch structure with an attention module, while OC-TD employs a double-branch structure without the attention module. Experimental results on the Rail Surface Defects Dataset (RSDDS) [40] demonstrate the effectiveness of these methods, surpassing state-of-the-art techniques on defect-level metrics.

3.2.1.4 Integration of CNNs with the internet of things Some studies combine the power of CNN's CV capabilities with innovative Internet of Things (IoT) systems such as robots and autonomous drones. In reference [41], the authors introduce a system that combines robotics and visual inspection for defect detection in railway tracks. This system conducts on-the-spot image processing, stores images of defective tracks in the cloud, and localizes the robot within a 3 to 6-inch range. It utilizes a ML system to classify track images as normal or suspicious, enabling efficient, targeted inspection by a dedicated operator; reducing the need to inspect the entire track. In another study [42], authors introduce a rail-track monitoring system using robots and optical checks for surface fault detection. The system employs Two-dimensional Convolutional Neural Network (2-D CNN) for real-time local detection during inspection. Images are sent to the ANN for assessment during robot investigation. In reference [43], the authors introduce a multi-robot fault detection system for railway tracks, replacing manual inspection. The hardware prototype features a master-slave robot mechanism detecting rail surface defects using ultrasonic sensors, image processing techniques and DL. The proposed CNN outperforms other methods like ANN, Random Forest (RF) and SVM, based on metrics like accuracy, R-squared value, F1 score, and mean-squared error. To eliminate manual inspection, the system communicates fault location and status to a central location via Global System for Mobile Communications (GSM), Global Positioning System (GPS), and cloud storage. In a different study [44], the authors propose a method for the control of the rail track with an autonomous Unmanned Aerial Vehicles (UAV). The approach employs the deep Hough transform method for rail navigation without requiring preprocessing or parameter adjustments. After removing the rails from the obtained rail images, the rail defects are detected by semantic segmentation. Another study [45] investigates the use of U-Net for the task of segmenting rail track regions from UAV-based images in the railway sector. This task is crucial for ensuring safety and security by monitoring potential hazards on railway tracks. The rapid development of DL and CV techniques has enabled automated railway hazard detection systems based on UAV-based imagery. The paper demonstrates the effectiveness of U-Net in terms of mean Intersection over Union (IoU) through experimental evaluations using a real-world dataset. This research has practical applications in automated UAV navigation along rail tracks.

3.2.1.5 Applications of CNNs with acceleration data Other data sources can also be utilized to feed CNN. Other types of data found in the literature include acceleration and speed data. For instance, the authors in [46] propose a model that integrates Building Information Modeling (BIM) and ANNs for the localization of defects in railway infrastructure. The study presents a case study for wheel burns in railway tracks, and implements DNN, CNN and RNN to make predictions based on axle acceleration data generated by simulations which is also synced to the 3D BIM model. Another study [47], introduces a CNN framework for real-time prediction of railway track dynamic stiffness using accelerometer data from sensors on train axle boxes. To ensure computational efficiency, the framework incorporates dilated convolutional layers, suitable for implementation on compact devices. They utilize a calibrated nonlinear finite element model to create an unbiased dataset of axle box acceleration under various conditions. The fine-tuned CNN model achieves optimal R-squared values, offering a continuous, cost-effective, and fast method for track stiffness measurement. In reference [48], the authors introduce a classification-based method to detect various rail defects, including localized surface collapse, rail end batter, and different rail components such as joints, turning points, and crossings. They leverage acceleration data for this purpose. To enhance the practicality and performance of these classification-based models, the authors put forward a DL approach that employs CNN. These models aim to identify joints or defects on either the left or right rail. The authors investigate and evaluate two convolutional networks, ResNet and Fully Convolutional Networks (FCN), using acceleration data. Another study [49] implements DL for processing data harvested from smart sensors and IoT devices for FD in railway tracks. The solution combines traditional signal processing methods with Deep Convolutional Autoencoders (CAEs) and clustering algorithms to find anomalies and their patterns on railway tracks. The methods are applied to real world Axle Box Acceleration (ABA) data gathered with a multi-sensor measurement system on a shunter locomotive. In reference [50], the authors introduce a novel DL approach using daily monitoring data from in-service trains to predict rail breaks. The proposed model tackles data imbalance and preserves temporal dynamics. It employs Time-Series Generative Adversarial Network (TimeGAN) for handling imbalance and generating synthetic rail break data while maintaining temporal characteristics. The Feature-Level Attention-Based Bidirectional Recurrent Neural Network (AM-BRNN) enhances feature extraction and captures bidirectional dependencies for accurate prediction. The approach is validated with a three-year dataset from Australian railroads, covering up to 350 km that includes data on the characteristics of tracks: including rail age, annual tonnage, and train speed; as well as temperature data. Results show successful prediction of nine out of eleven rail breaks with 3 months in advance.

3.2.1.6 Applications of CNNs with other types of data A number of studies implement ultrasound and data from vibration waves. For instance, authors of reference [51] address the automation of ultrasound diagnostics for railway tracks using a CNN. They present a system architecture for real-time decoding of railway track defectograms, encompassing ultrasound data processing, CNN classifiers, and a decision block. Data preprocessing involves transforming measurements into a format suitable for ANNs and combining information based on the defect type. While in [52], the authors aim to enhance efficiency, accuracy, and cost-effectiveness in ultrasonic flaw detection, focusing on B-scan image classification. Inspired by successful transformer models in NLP, they propose using the Vision Transformer (ViT) model for training on rail defect B-scan images. This approach explores the practicality and effectiveness of employing ViT, a breakthrough in computer image processing, for the classification of rail defects. In another study [53], the authors propose a DL model that integrates CNNs and LSTM models for track quality evaluation in HSR networks. The proposed solution relies on the prediction of vehicle-body vibrations and takes advantage of the powerful feature extraction capacities of CNN and LSTM models.

There are two studies that implement audio and sound data for the detection of railway defects. The authors in [54] present a smart railway cart for detecting cracks on rail through acoustic analysis. For this work, acoustic signals were collected from real railway operations and several well-known ML algorithms were applied such as SVMs, Logistic Regression (LR), RF and Decision Tree (DT) classifier, in addition to DL models like multilayer perceptron and CNNs. Results suggest that acoustic data can successfully determine the presence of faults in tracks. And the authors in [55] introduce WSCNN-GRU, a novel method for HSR fatigue crack signal classification using Acoustic Emission (AE) technology. It combines DL principles, leveraging Self-Normalizing CNNs with Wide First-Layer Kernels (WSCNN) to extract local features, Gated Recurrent Unit (GRU) for capturing temporal relevance, SeLU activation for stability, and Adaptive Batch Normalization (AdaBN) for versatile domain adaptation.

Lastly, there is one study that implements electrical signals from monitoring events. The authors in [56] present an innovative track circuit fault detection framework using 1-D CNN with multiscale feature fusion. The framework addresses locality dependencies among monitoring variables and introduces three multiscale feature fusion methods (parallel, serial, and dense) to enhance feature learning and diagnosis. Experimental results, compared against

traditional classifiers, affirm the effectiveness of the proposed approach. The dense feature fusion method stands out as the most robust, significantly improving fault diagnosis performance.

3.2.2 Fasteners

Railway fasteners are integral components that serve multifaceted purposes essential for the correct functioning of railways. The maintenance of these fasteners is pivotal, as they must withstand the heavy loads and vibrations associated with train traffic, ensuring the safety, stability, and efficiency of railroads.

3.2.2.1 Integration of CNNs with other types of AI There are many studies that implement CNNs combined with other DL models for this purpose. For instance, the study in [57] addresses imbalanced data in diagnosing faults in rail fastener by employing a combination of Generative Adversarial Networks (GAN) and ResNets. The GAN is utilized to generate additional fault data, balancing the dataset, which is then used to train the ResNet. The study evaluates the fault diagnosis method by calculating the average accuracy from multiple experiments, demonstrating enhanced accuracy in fault detection, notably given the significant shortage of fault data. While in reference [58], the authors aim to improve the performance of DL-based defective fastener inspection methods. They introduce a novel image generation method called Four-Discriminator Cycle-consistent Adversarial Network (FD-Cycle-GAN) to generate defect fastener images using a substantial number of defect-free ones. Through extensive experiments on both real and generated images, results demonstrate that the defect-fastener images produced by the proposed method exhibit higher quality and greater diversity compared to other state-of-the-art methods. Moreover, training the fastener inspection model on the expanded dataset (which includes defect fastener images generated by FD-Cycle-GAN) significantly improves performance compared to the CNN-only baseline. Another approach [59] proposes a DL solution for railroads inspection that implements collaborative CNN-based detection models. This solution combines multiple detectors within a multi-task learning framework to find defects on railway sleepers and fasteners based on image data. The authors in reference [60] investigate the combined use of image processing and DL algorithms for detecting missing clamps within a rail fastening system. They use image processing techniques to improve the fastener's location and remove redundant information from the images. Then a CNN and ResNet-50 networks are used for classification purposes. The images used for this study were acquired during field inspections and enhanced with data augmentation techniques.

3.2.2.2 CNNs for the detection of fasteners Other studies implement different types of CNN such as YOLO or different varieties of Region-based Convolutional Neural Networks (R-CNN) for the detection and classification of railway fasteners. For instance, authors in [61] propose the implementation of DL along with image processing for FD in railway track fasteners. The implementation relies on image detection, feature extraction and a classifier algorithm, and it implements a VGG-16 CNN and a Faster R-CNN for the positioning and recognition of fasteners. The authors in [62] present a method to detect targets in fastening systems in different railway sections. They improved the Faster R-CNN model, using multi-scale feature map fusion for small targets; modified predefined anchors to generate region proposals; and an additional attention module for focusing on meaningful features. They used railway inspection images as the dataset and carried out labeled work. In reference [63], a Mask R-CNN architecture and a DL framework for image segmentation, were employed to differentiate between intact and missing rail fasteners. Railway images captured by an autonomous drone were annotated to identify the healthy and absent fasteners. A model was then trained using this labeled dataset, and its performance was assessed using a separate test dataset. The experimental results demonstrated that this approach could detect both healthy and missing fasteners, achieving high levels of accuracy. A different study [64] presents a track fasteners classification system based on a YOLOv4-Tiny DNN model, that can perform real-time identification with two cameras. The datasets used to train the model were provided by the Taiwan Railway Administration. Finally, the authors in [65] propose to use a one-stage RetinaNet DCN to detect rail fasteners in aerial imagery of rail tracks. The proposed model was tested on a one-kilometer railway stretch: It was determined that the learned deep features promise a robust method that can help to identify rail anomalies, though more extensive training data is needed.

3.2.3 Joints

Railway joints allow for the connection of individual rail segments, creating a continuous track network. These components are fundamental, as they provide flexibility, accommodate thermal expansion, ensure proper alignment, and facilitate the maintenance of railway tracks. Only two papers in the literature present CNN-based solutions for their maintenance. The first study [48] presents a model based on a ResNet and a FCN that can classify and detect various rail defects—including joints—based on acceleration data. Another study [66], implements a DL-based approach for the conditional monitoring of thermite-welded joints in rails CNNs. This solution proposes the implementation of an onboard camera or sensor to collect the images necessary for the visual inspection of joints in real-time. With YOLOv3 as the detector, the proposed model can accurately classify and determine the location boundary box for each thermite-welded joint image.

3.2.4 Sleepers

Railway sleepers, also known as railroad ties or crossties, are essential components in the physicality of railway tracks. Sleepers provide structural support, distribute loads, maintain track alignment, absorb vibrations, facilitate maintenance, anchor ballast, and withstand environmental conditions. Their maintenance is important for the functioning of railroads. Nonetheless, few papers in the literature refer to the applications of CNNs for the maintenance of railway sleepers. The authors in [67] introduce a two-stage algorithm to detect railway sleeper cracks. The initial stage applies a 3×3 edge detection in a neighborhood range, identifying potential crack areas. The subsequent stage employs a CNNs for accurate classification of detected edges. The model implements 500 images of cracked sleepers and 500 images of healthy sleepers and data augmentation techniques. In reference [68], the authors develop multiple CNN-based models for predicting and diagnosing defect severities in unsupported sleepers, aligned with track inspection guidelines. Using data from a validated finite element model and real field measurements, the study covers various scenarios, considering sleeper locations, quantities, and operational parameters. Key indicators rely on axle box accelerations, and multiple DL techniques (including CNNs, RNNs, ResNet, and fully CNNs) are explored. Notably, the CNN model exhibits the highest accuracy in predicting unsupported sleeper conditions and identifying defect severities. Another solution [59], implements collaborative CNN-based detection models within a multi-task learning framework to find defects on railway sleepers (and fasteners) based on image data. Finally, the authors in [69] propose an improved YOLOv3 algorithm for detecting sleeper defects, aiming to enhance track maintenance efficiency and reduce manual risks. They optimize loss function weights based on sleeper image characteristics, utilize the K-means algorithm for sleeper data clustering, and employ multi-scale training for robustness. Experimental results demonstrate significant improvements in recall, precision, and Mean Average Precision (MAP), highlighting the effectiveness of the enhanced YOLOv3 algorithm.

3.2.5 Switches and crossings

The railway Switches and Crossings (S&C), commonly known as turnouts or points, are vital components in railroad infrastructure as they contribute to the overall efficiency and reliability of railway systems by providing the necessary infrastructure for routing, directional control, traffic management, yard operations, maintenance access, and safety.

3.2.5.1 CNNs and computer vision techniques There are different CNN-based approaches that implement different types of data for the maintenance of S&C in railroads. Most approaches are based on image and video data. For instance, in reference [70], the authors introduce a CV method for detecting wear on railway automatic switch Stationary Contacts (SC) under few-shot occasions. They propose a two-stage approach for achieving state-of-the-art performance in Railway Automatic Switch Stationary Contacts (RASSC) wear detection. The first stage involves a Few Shot SC Detection (FSDet) module that employs a Particle Swarm Optimization (PSO) based Weighted Non-Maximum Suppression (PWNMS) algorithm for decision-making in multi-template detectors based on deep features matching. The second stage, Contour-based Size Measurement (CSMea), utilizes unique area features for wear detection and measurement. The authors in [71] investigate a DNN based environment perception using vehicle-borne camera images from the rail domain. They utilized two datasets, RailSem19 [72] and a smaller self-labeled dataset. The DNN is applied to railway switch detection and classification, using Transfer Learning (TL), anchor box optimization and appropriate architecture to address the lack of suitable training data and class imbalance. Another study [73] presents a new method based on YOLOv4 for simultaneous condition monitoring and fault detection in the railway switch and level crossing sections. A YOLOv4 DNN (with a

Darknet53 backbone) was trained using four class label datasets consisting of real railway visual data collected with an autonomous drone. In reference [74] the authors present an automatic isolating switch segmentation and state recognition framework called ISSSR-Net using multitask learning to address the issue of isolating switch accurate localization and state recognition simultaneously. ISSSR-Net comprises two stages: Firstly, ISS-Net, an isolating switch segmentation network. It features a novel structure including strip pooling, channel attention, and three pyramid pooling modules, enhancing performance under challenging conditions like rain, snow, and fog; Secondly, ISS-Net's segmentation map and the shared backbone's feature map are input into ISR-Net, the isolating switch recognition network. An additional global context block further enhances state recognition accuracy. The dataset implemented for training the model is composed of self-collected isolating switch images under five different conditions and states.

3.2.5.2 Implementing electrical signal data with CNNs Other models in the literature implement data from electrical signals. For instance, the authors of reference [75] propose a novel hybrid DL framework combining Deep Convolutional Autoencoder (DCAE) and LR for turnout fault diagnosis. The current signal is converted into a 2-D grayscale image, the DCAE is used for automatic feature extraction, and LR is used for fault diagnosis. They use a historical field dataset composed of turnout curves (current vs. time), that were converted into grayscale images to feed the model. In reference [76], the authors present an improved technique for extracting effective features from faulty signals using Energy-Based Thresholding Wavelets (EBTW) and ANNs. EBTW applies a wavelet transform coefficient threshold, leveraging localized and redistributed signal energy analysis. This method is versatile across sensors with varying fault sensitivities. The threshold effectively eliminates measurement noise, calculated based on the conserved signal energy ratio with physical meanings. Comparative analysis with different feature extraction methods and ANNs classifiers, using real-world railway switch data, reveals EBTW outperforming conventional methods like Discrete Wavelet Transform (C-DWT) and Soft-Thresholding DWT (ST-DWT) in dimension reduction ratio and diagnosis accuracy.

3.2.5.3 Implementing acceleration data with CNNs Acceleration data can be an important indicator of S&C health status. For instance, in the study with reference [77], the authors explore train type identification in railway S&C using accelerometer data and contemporary ML methods including ANNs. They use data from accelerometers placed around four S&C structures at different locations, and test various ANNs architectures. Their findings show that models trained in one location can typically be applied to another, even when there are variations in geometry, substructure, and the direction of passing trains. Another study [78] implements DL methods based on the LSTM and ResNet to predict the amount of wear in the entire S&C infrastructure, using medium-range accelerometer sensors. Vibration data were collected, processed, and used for developing accurate data-driven models. The first task in this study was to confirm the assumption that it was possible to measure vibrations that would reflect the amount of wear in the entire S&C. The second task was to investigate how the measurement accuracy would be affected by the distance between the sensor and the wear location. Lastly, the study [48] introduces a classification-based method to detect various rail defects based on ResNet and FCNs, including—among other tasks—the detection of crossings defects based on acceleration data.

3.2.5.4 Object and component detection The detection and classification of specific components and external objects is essential for automating intelligence maintenance in railways. This contributes not only to the maintenance (and maintenance planning) of railroads but to the overall safety and efficiency of train operations. The models discussed in this section are integrated in some of the solutions we have presented in the previous subtasks. For instance: the study presented in [65] can detect rail fasteners in aerial imagery of rail tracks by implementing a RetinaNet DCN; The solution presented in [64] implements YOLOv4-Tiny of real-time identification and classification of track fasteners; In [62] a Faster R-CNN and multi-scale feature map fusion technique is implemented to detect targets in the fastening systems. The detection of anomalies and foreign objects can also be leverage to CNNs. As we have seen in studies previously presented: In [16], for the detection of rail surface cracks, implements a multitask learning model and an object detection decoder for detecting rail objects; And the authors of [28] introduce a YOLOv5s-VF for rail surface defect and object detection.

3.2.5.5 CNNs for the detection of railroad parts We have also reviewed new studies that are specifically targeted at the detection of different railroad components. For instance, in reference [79], the authors introduce an attention-powered deep CNN, AttnConv-net, for detecting multiple rail components, like rails, clips, and bolts. This method integrates Cascading Attention Blocks (CAB), two Feed-Forward Networks (FFN), and positional embeddings into a deep CNN core. The CAB module focuses on learning local context, while the FFN generates final categories and bounding boxes. The model

is trained with various data augmentation techniques to improve the detection of small components. The AttnConv-net simplifies the detection pipeline, eliminating the need for pre and post-processing, offering a faster and more accurate detection system. The solution proposed in [80] uses drone images and YOLOv3 for the detection of various track assets with the purpose of monitoring the health of railroads. In a different study [81], the authors proposed a CNN-based detection method for damaged steel-spring vibration isolators (SVIs) in a floating-slab. A 1-D Deep ResNet has been designed for feature extraction and data classification. Using vibration responses generated via vehicle floating-slab track (FST) coupled dynamic simulations, the network extracts damage-sensitive features from raw data to identify the damaged SVIs. For network training and testing, multiple data sets are constructed under various scenarios. The key contribution of this work is to investigate the sensor deployment for CNN with a good performance and adaptability to different scenarios.

3.2.5.6 CNNs for the detection of anomalies Other studies are centered in the detection of anomalies in railway tracks for maintenance purposes. For instance, the authors in [83] develop a multi-stage framework to automatically inspect the railway during the night, in order to detect, localize and classify objects or anomalies that could affect the safety of the train transport. The framework is able to predict the presence, the image coordinates and the class of obstacles. The framework is based on thermal images and consists of three different modules that address the problem of detecting anomalies, predicting their image coordinates and classifying them. The authors introduce a new multi-modal dataset, acquired from a rail drone, used to evaluate the proposed framework. In reference [84], the authors propose a novel DL application for detecting anomalies on railway tracks using camera data. Their two-stage approach involves binary semantic segmentation to extract rails and a self-supervised learning-based Autoencoder to identify anomalies. The Autoencoder is trained on patches from the segmented rails, aiming to reconstruct non-anomalous data. During inference, larger reconstruction errors indicate anomalies, detected by applying a predefined threshold. The first stage achieves a high mean Intersection over Union (IoU), and the Autoencoder network performs well on real scenario test images, successfully detecting anomalies without false positives or false negatives. Another study [85] proposes a semi supervised algorithm for detecting foreign objects in ballastless beds based on the improved deep SVDD (Support Vector Data Description) algorithm. First, they use the improved Mask R-CNN algorithm to extract the rail and fastener areas in images, assuming that no foreign object exists in the rail and fastener areas. Second, they deepen the backbone network of the deep SVDD to enhance its ability to extract deep semantics from complex images. They perform pure color coverage processing with different colors and mean blur processing with different blur kernels on the rails and fastener regions extracted by the improved Mask R-CNN. The study in [82] presents an automated real-time system developed for the maintenance of vegetation on and near railway tracks. This system enables precise spraying of herbicides in specific locations where needed, in contrast to older systems that apply herbicides uniformly along the tracks. The system comprises a locomotive-mounted camera that records the area in front of the train. The video stream from the camera is transmitted to a standard computer, where it undergoes processing using CNNs. This software facilitates the detection of weeds and bushes along the railway tracks and communicates with the Programmable Logic Controller (PLC) to activate herbicide sprayers at the appropriate moments. Finally, the authors in [86] propose a new semi-supervised anomaly detection method based on GANs to detect foreign objects in a railway environment. This approach tackles the problem that datasets of railways with foreign objects are scarce, using normal railway images for training.

3.2.6 Monitoring

The continuous monitoring in railroads is important for the prevention, early detection and prognosis of faults; to ensure infrastructure integrity; for maintenance planning purposes; and support of the overall efficient railway maintenance tasks and operations. There are many CNN-based solutions presented in the literature for the monitoring of railroads. Some of the models discussed in this section have been presented in previous subtasks.

3.2.6.1 Computer vision monitoring Most monitoring approaches leverage on the CV capabilities of CNNs: The YOLO single-shot CNN-based detection method has been implemented in many studies for continuous monitoring in railway tracks. For instance, in [80], the authors focus on detecting various track assets in drone images for the monitoring of the health status of railroads. For this purpose, they have utilized a pre-trained model. The image classes for categorization include *construction*, *power junction/brick*, *cement slabs*, *transformer wires*, *garbage*, and *person*. To enhance detection accuracy, they compute color spaces for images based on selected separability index values. The YOLOv3 framework is implemented for track asset detection. In the study presented in [30], the authors highlight the importance of monitor-

ing railway track conditions by leveraging on an improved YOLOv3 model named TLMDDNet and a lightweight design strategy named DC-TLMDDNet which optimize feature extraction layers using a DenseNet. While the study [26] introduces YOLOv3-M for rail health monitoring and detection of railway defects. The results demonstrate that the YOLOv3-M-based method can effectively monitor the state of railway tracks, enabling early detection of defects before they lead to failures. In reference [66], the authors develop an application prototype for the detection and monitoring of thermite-welded rail joints using the YOLOv3 DL algorithm. With YOLOv3 as the detector, it can accurately classify and pinpoint the location boundary box for each thermite-welded joint image. The experimental training and validation of this algorithm have yielded promising results, affirming the capability of this application prototype. When integrated with the proposed camera monitoring system, it becomes a valuable tool for detecting and monitoring critical components of the rail track system. Finally, the authors in [73] present a new method based on YOLOv4 (and data collected with an autonomous drone) for the condition monitoring and FD in the railway S&C system.

3.2.6.2 Monitoring through other CNNs Other types of CNNs are also used in similar monitoring applications. For instance, the authors of [42] proposed a monitoring system based on 2-D CNNs and robots for real-time FD during inspection of railway tracks. Detected flaws are promptly reported to the cloud, including the corresponding region, for additional analysis. Similarly, the study [22] presents a continuous and efficient monitoring of railway tracks with the aim of reducing the occurrence of accidents caused by poor condition of railway tracks. In a different study [45], the authors implemented (UAV-obtained) aerial images and a U-Net to automate the detection of hazards in railway tracks.

4 Discussion

Our study has reviewed several implementations of Convolutional Neural Networks (CNN) models for different tasks related to the preservation, inspection, maintenance and monitoring of railway tracks and their components. Results show a sustained increase in research literature produced from 2019 onwards. Most papers revised were published in 2022, constituting 46% of the total, this is followed by 2021 with 30% of papers, and 2020 with 20% of papers. In earlier years, there were fewer publications, with two papers (3%) in 2019, no papers (0%) in 2018, and only one paper (1%) in 2017. There were no publications from 1990 to 2016 (0%). When it comes to the classification of papers according to the maintenance task, there is a noticeable unbalance in their distribution: *surface defect detection* constitutes the majority with 49% of the papers, followed by *detection of objects and components* (15% of papers), *fasteners* (10%), *monitoring* (9%), *switches and crossings* (10%), *sleepers* (5%), and *joints* (2%). The distribution of papers according to the data implemented is as follows: The vast majority of papers focus on image and video analysis, comprising 70% of the total, followed by speed and acceleration analysis (which represents 11% of papers), vibration measurements (10% of papers), electrical signals analysis (4%), acoustic and sound (3%), track geometry (1%), and temperature (1%). In the rest of this section, we briefly delve into our discoveries, highlighting certain trends and insights we have identified.

Most of the papers we have revised present solutions for the detection of defects in the surface of railway rails. Within this category many studies implement solutions based on Computer Vision (CV) to discern between images of healthy tracks and defected tracks that present cracks and other anomalies. Apart from the regular CNN implementations [13, 15, 16, 22, 25] different architectures are proposed, including VGG16 [12, 23], Faster Region-based Convolutional Neural Network (R-CNN) [14], Residual Network (ResNet) [32, 33], U-Net [18, 34, 45], GoogleNet's CNN [19], 1-D [39] and 2-D [42] CNNs, along with other novel applications based on CNNs [20, 21, 34–36]. A recurrent implementation in image and video-based solutions is the You Only Look Once (YOLO) network. This Deep Neural Network (DNN) is implemented any many real-time approaches, in most cases presenting their own modified version of the network [26–30], which is improved for the detection of railway defects. Some novel approaches to image and video data acquisition include the use of onboard cameras [13, 25] and—most notably—robots [42–44] and aerial railway track images captured with autonomous Unmanned Aerial Vehicles (UAVs) [44, 45]. For the purpose of identifying defects in the railways, CNN have also notable results with raw data collected from railway operations including acceleration [46–48], vibration [50, 52–54], and audio signals [54, 55].

We have analyzed a good number of papers related to the maintenance of railway components such as fasteners, sleepers, and joints. Among these components, railway fasteners are addressed the most; there are solutions that implement different CNNs: for instance ResNets for the detection of missing clamps [60]; the use of Region-based CNN (R-CNN) such as Faster R-CNN [61, 62] for detection of faults and targets, Mask R-CNN [63] for discerning between intact and missing fasteners; YOLO [64] and RetinaNet with aerial images [65] for fasteners detection and classification; and

the implementation of collaborative multiple-CNN models for the inspection of railway fasteners [59, 61]. There are approaches that implement CNNs in combination with other Deep Learning (DL) models. One example is the use of Generative Adversarial Networks (GANs) for the generation of additional data [57, 58]. Only a few papers address the maintenance of railway sleepers. These studies implement for example, CNNs for the detection of cracks in sleepers [67], multiple collaborative CNN models [59, 68], and YOLO [69] for detection of sleeper defects. Finally, only two papers implement CNN-based solutions for the early fault detection in railway joints: one study implements ResNet and Fully Convolutional Networks (FCN) to detect defects in joints based on acceleration data [48]; and another CNNs for the monitoring of welded railway joints using images collected with onboard cameras [66].

Another subtask identified is the maintenance of railway switches and crossings (S&C). Among this classification we found studies that implement Deep Neural Networks (DNN) [73], YOLO [73] and other CNN-based solutions [70, 74] for the detection and classification of S&C [74], as well as detection of wear and other possible faults based on S&C images. Other solutions implement electrical signals [75, 76] or acceleration data [48, 77, 78] for the same purpose.

The detection and classification of specific railways components is an essential part of intelligent maintenance and crucial for early Fault Detection (FD). Some of the components include fasteners, joints and sleepers among others, overlapping with the identified subtasks present in this study. There are approaches found in the literature that implement R-CNNs [62], YOLO [64, 80], DCN [65, 79], ResNet [81] and other traditional models for this purpose. Other studies are centered in the identification of anomalies, such as external objects in the rails [16, 28, 83–86] or plants growing in the tracks [82].

The last task is the real-time monitoring of railway tracks for maintenance purposes. The YOLO single-shot CNN-based detection is a recurring method for continuous monitoring of railroads [26, 30, 66, 73, 80]. Other applications include 2-D CNNs [42] and U-Net [45]. The use of robots and Unmanned Aerial Vehicles (UAVs) is also a frequently discussed alternative for the monitoring of railway tracks [42, 73, 80].

There are some recurring methodologies that are common to all areas of railway track maintenance. For instance, the YOLO network has been implemented in many approaches, that leverage on the CV capabilities of CNNs for the detection and tracking of objects in real-time [26–30, 64, 66, 69, 73, 80]. Other popular models that are seen among different subtasks are R-CNNs (and they different variations) [14, 61–63], U-Nets [18, 34, 45] and ResNets [32, 33, 48, 60, 81]. The most common type of data used for CNN-based approaches is image and video data as observed in surface defect detection [12–16, 18–22, 24, 25, 27–39, 41–45, 52]; maintenance of fasteners [57–60, 63–65], joints [66], sleepers [59, 67, 69], switches and crossings [70, 73, 74]; detection of objects and components [16, 28, 64, 65, 79, 80, 82–86]; and monitoring [30, 42, 45, 66, 73, 80]. This is not surprising given the inherited qualities of CNN for CV. The implementation of IoT devices and smart sensors aid the collection of real-time data which can be used to feed CNN models. For this purpose, various IoT devices, robots [41–43] and autonomous drones [44, 45, 63, 65, 73, 80, 83] are implemented to obtain different types of data. To overcome lack of training data CNNs are used in combination with GANs to generate new data [50, 57, 58, 86] that can be used later for training purposes. Other types of data include speed and acceleration [46–50, 68, 77, 78], vibration [43, 49, 51–53, 78, 81], audio [54, 55] and electrical signals [56, 75, 76].

5 Conclusion

In previous studies, we have discussed the implementation of powerful Artificial Intelligence (AI) models for different maintenance tasks within the railway industry. The advance of Artificial Neural Networks (ANN) and other numerical models have the potential to greatly contribute to smart-maintenance activities in the railway sector. In this study, we have addressed the potential of Convolutional Neural Networks (CNN) for different tasks related to the preservation, inspection, maintenance and monitoring of railway tracks and their components. Our research showed a steady increase in the number of research papers on the topic in the last years. We have analyzed the state-of-the-art by summarizing different tasks and problems belonging to the maintenance of railway tracks and common CNN-based models implemented for their solution. Within the scope of our research, we have identified the following tasks: surface defect detection; maintenance of fasteners, joints, sleepers, switches and crossings; detection of objects and components; and railway monitoring.

Most of the papers we have revised present solutions for the detection of defects in the surface of railway rails including cracks, shelling defects, and the detection of other damages and anomalies. Many studies leverage the Computer Vision (CV) capabilities of CNN for this task. So is the case of the maintenance of railway track components, such as fasteners, joints, sleepers, switches and crossings and the recognition of track components, in which CV techniques are widely implemented for the detection of defects and anomalies. These models are also able to recognize patterns and identification of complex

events aiding in the continuous monitoring of railway tracks. Monitoring ensures the integrity of the railway infrastructure by helping identify potential issues such as track deformations, misalignments, or defects; as well as wear and tear, fatigue, or damage in track components, among other problems that could compromise the stability and durability of railroads. Not only limited to CV approaches, CNNs have been used with different types of data including image and video, vibration, audio, and electrical signals. Moreover CNN-based models have the advantage of working directly with raw data, reducing the need for extensive feature engineering preprocessing. The implementation of IoT devices and smart sensors aid the collection of real-time data which can be used to feed powerful CNN-based models to recognize patterns and identification of complex events related to the maintenance of railway tracks. These approaches have the potential to significantly transform railway track engineering by promptly identifying problematic tracks within extensive railway networks and (promptly) accommodating pertinent maintenance tasks.

There are great advantages to using CNN for the maintenance of railway tracks. The architecture of CNNs allows these models to learn spatial hierarchies of features automatically and adaptively from the input data making them an ideal candidate for different maintenance tasks related to railway tracks. On the other hand, CNN models lack explainability and transparency in contrast to traditional AI approaches which may be more intrinsically comprehensible. The explainability, comprehension and interpretability of CNN models become more obscure as complexity increases. This issue can be overcome with exhausting testing and various techniques to make these models more transparent and easier to understand. It is important that models are tested thoroughly before they can be deployed, especially in applications such as the maintenance of railway tracks, where errors can be translated into fatalities.

Acknowledgements The work presented in this study was partially supported by the *Swiss State Secretariat for Education, Research and Innovation* (SERI) under contract number 23.00321 (*Academics4Rail* project). The project has been selected within the European Union's *Horizon Europe research and innovation* program under grant agreement: HORIZON-ER-JU-2022-ExpIR-04. The views and opinions expressed are, however, those of the authors only and do not necessarily reflect those of the funding agencies, which cannot be held responsible for them.

Special thanks to Daniel Rytenberg from the *School of Innovation, Design and Engineering* (IDT) at Mälardalen University for his collaboration with the technical drawings presented in this work.

Author contributions Conceptualization Francesco Flammini, Mauro José Pappaterra; Investigation Mauro José Pappaterra, María Lucía Pappaterra; Data Curation Mauro José Pappaterra; Writing (original draft preparation) Mauro José Pappaterra, María Lucía Pappaterra; Writing (review and editing) Francesco Flammini, Mauro José Pappaterra, María Lucía Pappaterra; Statistical Analysis Mauro José Pappaterra, María Lucía Pappaterra; Supervision Francesco Flammini; Funding Acquisition Francesco Flammini. All authors have read and agree to the published version of the manuscript.

Funding HORIZON EUROPE Framework Programme, HORIZON-ER-JU-2022-ExpIR-04, HORIZON-ER-JU-2022-ExpIR-04, HORIZON-ER-JU-2022-ExpIR-04, HORIZON-ER-JU-2022-ExpIR-04

Data availability No datasets were generated or analysed during the current study.

Declarations

Competing interests The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Appendix

A tabulated list of all papers consulted for the literature review can be found online: <https://github.com/mauropappaterra/list-literature-CNNs-Springer-Discover-AI>.

References

1. Pappaterra MJ. A literature review for the application of artificial intelligence in the maintenance of railway operations with an emphasis on data. In: Presented at: 18th European dependable computing conference—EDCC2022; 2022. <https://doi.org/10.13140/RG.2.2.25282.66249>.
2. Pappaterra MJ. A review of literature and public datasets for the application of artificial intelligence in the railway industry [dissertation]. Uppsala Universitet. Sweden; 2022. <https://doi.org/10.13140/RG.2.2.24091.13600/2>.
3. Tang R, De Donato L, Besinović N, Flammini F, Goverde RM, Lin Z, Liu R, Tang T, Vittorini V, Wang Z. A literature review of artificial intelligence applications in railway systems. *Transp Res Part C Emerg Technol*. 2022. <https://doi.org/10.1016/j.trc.2022.103679>.
4. Pappaterra MJ, Flammini F, Vittorini V, Bešinović N. A systematic review of artificial intelligence public datasets for railway applications. *Infrastructures*. 2021;6(10):136. <https://doi.org/10.3390/infrastructures6100136>.
5. Nakhaee MC, Hiemstra D, Stoelinga M, Noort MV. The Recent Applications of Machine Learning in Rail Track Maintenance: A Survey. In: *International Conference on Reliability, Safety, and Security of Railway Systems*; June 4–6, 2019; Lille, France. Springer, Germany; 2019:91–105.
6. Thilagavathy N, Harene J, Sherine M, Shanmugasundari T. Survey on railway wheel defect detection using machine learning. *AutAut Res J*. 2020;11(4):4.
7. Hamadache M, Dutta S, Olaby O, Ambur R, Stewart E, Dixon R. On the fault detection and diagnosis of railway switch and crossing systems: an overview. *Appl Sci Open Access J*. 2019;9(5129):1–32.
8. Liu S, Wang Q, Luo Y. A review of applications of visual inspection technology based on image processing in the railway industry. *Transp Saf Environ*. 2019;1:185–204.
9. De Donato L, Flammini F, Marrone S, Mazzariello C, Nardone R, Sansone C, Vittorini V. A survey on audio-video based defect detection through deep learning in railway maintenance. *IEEE Access*. 2022. <https://doi.org/10.1109/ACCESS.2022.3183102>.
10. Michas G. Slab track systems for high-speed railways [dissertation]. KTH, School of Architecture and the Built Environment (ABE), Transport Science, Highway and Railway Engineering; 2012.
11. Saha S. A comprehensive guide to convolutional neural networks—the ELI5 way. Towards data science. 2018. <https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-theeli5-way-3bd2b1164a53>. Accessed 17 Nov 2023.
12. Bhat S, Karegowda AG, L. R. A, V. S, D. G. Classification of Rail Track Crack using CNN with Pre-Trained VGG16 Model. In: *2022 IEEE International Conference on Distributed Computing and Electrical Circuits and Electronics (ICDCECE)*; Ballari, India; 2022. pp. 1–6. <https://doi.org/10.1109/ICDCECE53908.2022.9793318>.
13. Thendral R, Ranjeeth A. Computer vision system for railway track crack detection using deep learning neural network. In: *2021 3rd International conference on signal processing and communication (ICPSC)*; Coimbatore, India; 2021. pp. 193–196. <https://doi.org/10.1109/ICSPSC51351.2021.9451771>.
14. Chen X, Zhang H. Rail Surface Defects Detection Based on Faster R-CNN. In: *2020 International conference on artificial intelligence and electromechanical automation (AIEA)*; Tianjin, China; 2020. pp. 819–822. <https://doi.org/10.1109/AIEA51086.2020.00179>.
15. Zhuang L, Qi H, Zhang Z. The automatic rail surface multi-flaw identification based on a deep learning powered framework. *IEEE Trans Intell Transp Syst*. 2022;23(8):12133–43. <https://doi.org/10.1109/TITS.2021.3109949>.
16. Meng S, Kuang S, Ma Z, Wu Y. MtlrNet: an effective deep multitask learning architecture for rail crack detection. *IEEE Trans Instrum Meas*. 2022;71:1–10. <https://doi.org/10.1109/TIM.2022.3181940>.
17. Zhang Z. Rail-5k: a real-world dataset for railway surface defects detection (Version 1). 2021. Zenodo. <https://doi.org/10.5281/zenodo.4872618>.
18. Sevi M, Aydin İ. Detecting Flaws on Railways Using Semantic Segmentation. In: *2021 International Conference on Information Technology (ICIT)*; Amman, Jordan; 2021. pp. 179–183. <https://doi.org/10.1109/ICIT52682.2021.9491736>.
19. Aydin İ, Kirat SS, Akin E. Detection of rail surface defects with two deep learning methods: comparative analysis. In: *2022 30th signal processing and communications applications conference (SIU)*; Safranbolu, Turkey; 2022. pp. 1-4. <https://doi.org/10.1109/SIU55565.2022.9864863>.
20. Rengel J, Santos M, Pandit R. EfficientNet architecture family analysis on railway track defects. In: Yin H, Camacho D, Tino P, editors. *Intelligent data engineering and automated learning—IDEAL 2022*. IDEAL 2022. Lecture notes in computer science, vol. 13756. Cham: Springer; 2022. https://doi.org/10.1007/978-3-031-21753-1_46.
21. Aydin A, Salur MU, Aydin İ. Fine-Tuning Convolutional Neural Network Based Railway Damage Detection. In: *IEEE EUROCON 2021—19th International Conference on Smart Technologies*; Lviv, Ukraine; 2021. pp. 216–221. <https://doi.org/10.1109/EUROCON52738.2021.9535585>.
22. Rakshit S, BS S, E. F. S. S. Railway Track Fault Detection using Deep Neural Networks. In: *2022 IEEE 6th Conference on Information and Communication Technology (CICT)*; Gwalior, India; 2022. pp. 1–5. <https://doi.org/10.1109/CICT56698.2022.9997883>.
23. Balogun I, Leadingham M, Gulliot D, Attah-Okine N. Deep learning approach towards squat isolation in a multi-embedded track geometry defects. In: *2021 IEEE International Conference on Big Data (Big Data)*; Orlando, FL, USA; 2021. pp. 4491–4495. <https://doi.org/10.1109/BigData52589.2021.9671701>.
24. Yang H, Wang Y, Hu J, He J, Yao Z, Bi Q. Segmentation of track surface defects based on machine vision and neural networks. *IEEE Sens J*. 2022;22(2):1571–82. <https://doi.org/10.1109/JSEN.2021.3133280>.
25. He Y, Wu J, Zheng Y, Zhang Y, Hong X. Track defect detection for high-speed maglev trains via deep learning. *IEEE Trans Instrum Meas*. 2022;71:1–8. <https://doi.org/10.1109/TIM.2022.3151165>.
26. Wan ZJ, Chen SQ. Railway Tracks Defects Detection Based on Deep Convolution Neural Networks. In: Liang Q, Wang W, Mu J, Liu X, Na Z, Cai X, editors. *Artificial intelligence in China*. Lecture notes in electrical engineering, vol. 653. Singapore: Springer; 2021. https://doi.org/10.1007/978-981-15-8599-9_15.

27. Anwar N et al. YOLOv4 Based Deep Learning Algorithm for Defects Detection and Classification of Rail Surfaces. In: 2021 IEEE International Intelligent Transportation Systems Conference (ITSC); Indianapolis, IN, USA; 2021. pp. 1616–1620. <https://doi.org/10.1109/ITSC48978.2021.9565036>
28. Wang M, Li K, Zhu X, Zhao Y. Detection of surface defects on railway tracks based on deep learning. *IEEE Access*. 2022;10:126451–65. <https://doi.org/10.1109/ACCESS.2022.3224594>.
29. Zheng J, Han Q, Wang L, Wang S, Zhang Z, Wang L, Wang L. 2022. Lightweight Bimodal Rail Surface Defect Detection Method: Parallel-YOLOv4-Tiny. In: Proceedings of the 2021 5th International Conference on Video and Image Processing (ICVIP '21). Association for Computing Machinery, New York, NY, USA; 98–103. <https://doi.org/10.1145/3511176.3511192>.
30. Wei X, Wei D, Suo D, Jia L, Li Y. Multi-target defect identification for railway track line based on image processing and improved YOLOv3 model. *IEEE Access*. 2020;8:61973–88. <https://doi.org/10.1109/ACCESS.2020.2984264>.
31. Hu Q, Tang B, Jiang L, Zhu F, Zhao X. Rail Surface Defects Detection Based on YOLOv5 Integrated with Transformer. In: 2022 IEEE 5th International Conference on Electronics Technology (ICET); Chengdu, China; 2022. pp. 1131–1135. <https://doi.org/10.1109/ICET55676.2022.9824255>.
32. Song X, Chen K, Cao Z. ResNet-based Image Classification of Railway Shelling Defect. In: 2020 39th Chinese Control Conference (CCC); Shenyang, China; 2020. pp. 6589–6593. <https://doi.org/10.23919/CCC50068.2020.9189112>.
33. Yang H, Bi Q, Yao Z, Wang Y. Accurate and effective framework for identifying track defects. *Measurement*. 2022;190:110625. <https://doi.org/10.1016/j.measurement.2021.110625>.
34. Lu J, et al. SCueU-Net: efficient damage detection method for railway rail. *IEEE Access*. 2020;8:125109–20. <https://doi.org/10.1109/ACCESS.2020.3007603>.
35. Wang J, Song K, Zhang D, Niu M, Yan Y. Collaborative Learning Attention Network Based on RGB Image and Depth Image for Surface Defect Inspection of No-Service Rail. *IEEE/ASME Transactions on Mechatronics*. 2022;27(6):4874–4884. Dec. <https://doi.org/10.1109/TMECH.2022.3167412>.
36. Wu J, Zhou W, Qiu W, Yu L. Depth repeated-Enhancement RGB network for rail surface defect inspection. *IEEE Signal Process Lett*. 2022;29:2053–7. <https://doi.org/10.1109/LSP.2022.3211199>.
37. Aydin I, Akin E, Karakose M. Defect classification based on deep features for railway tracks in sustainable transportation. *Appl Soft Comput*. 2021;111:107706. <https://doi.org/10.1016/j.asoc.2021.107706>.
38. Zheng Z, et al. Anomaly detection of metro station tracks based on sequential updatable anomaly detection framework. *IEEE Trans Circuits Syst Video Technol*. 2022;32(11):7677–91. <https://doi.org/10.1109/TCSVT.2022.3181452>.
39. Zhang D, Song K, Wang Q, He Y, Wen X, Yan Y. Two deep learning networks for rail surface defect inspection of limited samples with line-level label. *IEEE Trans Ind Inform*. 2021;17(10):6731–41. <https://doi.org/10.1109/TII.2020.3045196>.
40. Niu M, Song K, Huang L, Wang Q, Yan Y, Meng Q. Unsupervised saliency detection of rail surface defects using stereoscopic images. *IEEE Trans Ind Inform*. 2021;17(3):2271–81. <https://doi.org/10.1109/TII.2020.3004397>.
41. AlNaimi N, Qidwai U. IoT Based on-the-fly Visual Defect Detection in Railway Tracks. In: 2020 IEEE International Conference on Informatics, IoT, and Enabling Technologies (ICIoT); Doha, Qatar; 2020. pp. 627–631. <https://doi.org/10.1109/ICIoT48696.2020.9089560>.
42. Maheswari BU, Nithya P, Vijay S, Tamilarasi K, G A, Muthukumaran N. An Automated Technology for IoT based Rail-Track Inspection to Locate Surface Flaws by Robotics and Neural Networks. In: 2022 International Conference on Inventive Computation Technologies (ICICT); Nepal; 2022. pp. 873–878. <https://doi.org/10.1109/ICICT54344.2022.9850453>.
43. Iyer S, Velmurugan T, Gandomi AH, et al. Structural health monitoring of railway tracks using IoT-based multi-robot system. *Neural Comput Applic*. 2021;33:5897–915. <https://doi.org/10.1007/s00521-020-05366-9>.
44. Aydin I, Sevi M, Sahbaz K, Karakose M. Detection of Rail Defects with Deep Learning Controlled Autonomous UAV. In: 2021 International Conference on Data Analytics for Business and Industry (ICDABI); Sakheer, Bahrain; 2021. pp. 500–504. <https://doi.org/10.1109/ICDABI53623.2021.9655796>.
45. Mammeri A, Siddiqui AJ, Zhao Y. UAV-assisted Railway Track Segmentation based on Convolutional Neural Networks. In: 2021 IEEE 93rd Vehicular Technology Conference (VTC2021-Spring); Helsinki, Finland; 2021. pp. 1–7. <https://doi.org/10.1109/VTC2021-Spring51267.2021.9448887>.
46. Sresakoolchai J, Kaewunruen S. Integration of building information modeling and machine learning for railway defect localization. *IEEE Access*. 2021;9:166039–47. <https://doi.org/10.1109/ACCESS.2021.3135451>.
47. Huang J, Yin X, Kaewunruen S. Quantification of dynamic track stiffness using machine learning. *IEEE Access*. 2022;10:78747–53. <https://doi.org/10.1109/ACCESS.2022.3191278>.
48. Yang C, Sun Y, Ladubec C, Liu Y. Developing machine learning-based models for railway inspection. *Appl Sci*. 2021;11(1):13. <https://doi.org/10.3390/app11010013>.
49. Niebling J, Baasch B, Kruspe A. Analysis of Railway Track Irregularities with Convolutional Autoencoders and Clustering Algorithms. In: Proc Dependable Computing—EDCC 2020 Workshops: AI4RAILS, DREAMS, DSOGR1, SERENE 2020, EDCC 2020. Communications in Computer and Information Science, vol 1279. Springer, Cham. https://doi.org/10.1007/978-3-030-58462-7_7.
50. Zeng C, Huang J, Wang H, Xie J, Huang S. Rail break prediction and cause analysis using imbalanced in-service train data. *IEEE Trans Instrument Meas*. 2022;71:1–14. <https://doi.org/10.1109/TIM.2022.3214494>.
51. Igonin A, Vitaliy U. Machine learning in problems of automation of ultrasound diagnostics of railway tracks. arXiv preprint [arXiv:2006.16844](https://arxiv.org/abs/2006.16844). 2020.
52. Lu S, Wang J, Jing G, Qiang W, Rad MM. Rail defect classification with deep learning method. *Acta Polytech Hung*. 2022;19(6):225.
53. Ma S, Gao L, Liu X, Lin J. Deep learning for track quality evaluation of high-speed railway based on vehicle-body vibration prediction. *IEEE Access*. 2019;7:185099–107. <https://doi.org/10.1109/ACCESS.2019.2960537>.
54. Shafique R, Siddiqui HUR, Rustam F, Ullah S, Siddique MA, Lee E, Ashraf I, Dudley S. A novel approach to railway track faults detection using acoustic analysis. *Sensors*. 2021;21:6221. <https://doi.org/10.3390/s21186221>.
55. Chang Y, Hao Q, Hu X, Zhang X, Shen Y. Rail Fatigue Crack Classification by WSCNN-GRU using Acoustic Emission. In: 2022 41st Chinese Control Conference (CCC); Hefei, China; 2022. pp. 3901–3906. <https://doi.org/10.23919/CCC55666.2022.9902249>.

56. Xing Y, Wang J, Peng C, Zhu L. Railway track circuit fault diagnosis using 1DCNN considering locality dependencies and multiscale feature fusion. In: 2022 41st Chinese Control Conference (CCC); Hefei, China; 2022. pp. 4081–4086. <https://doi.org/10.23919/CCC55666.2022.9902536>.
57. Dechen Y, Qiang S, Jianwei Y, Hengchang L, Jiao Z. Railway fastener fault diagnosis based on generative adversarial network and residual network model. *Shock Vib.* 2020;2020:8823050. <https://doi.org/10.1155/2020/8823050>.
58. Liu J, Ma Z, Qiu Y, Ni X, Shi B, Liu H. Four discriminator cycle-consistent adversarial network for improving railway defective fastener inspection. *IEEE Trans Intell Transp Syst.* 2022;23(8):10636–45. <https://doi.org/10.1109/TITS.2021.3095167>.
59. Gibert X, Patel VM, Chellappa R. Deep multitask learning for railway track inspection. *IEEE Trans Intell Transp Syst.* 2017;18(1):153–64. <https://doi.org/10.1109/TITS.2016.2568758>.
60. Chandran P, Asber J, Thiery F, Odelius J, Rantatalo M. An investigation of railway fastener detection using image processing and augmented deep learning. *Sustainability.* 2021;13:12051. <https://doi.org/10.3390/su132112051>.
61. Wei X, Yang Z, Liu Y, Wei D, Jia L, Li Y. Railway track fastener defect detection based on image processing and deep learning techniques: a comparative study. *Eng Appl Artif Intell.* 2019;80:66–81.
62. Peng X, Tong S, Lu X, Wei Y. An improved faster R-CNN for railway fastening system detection. In: Proceedings of the 5th international conference on computer science and application Engineering (CSAE '21). New York, NY, USA: Association for Computing Machinery; 2021:Article 109. <https://doi.org/10.1145/3487075.3487184>.
63. Yilmazer M, Karakose M. Mask R-CNN architecture based railway fastener fault detection approach. In: Proceedings of the 2022 international conference on decision aid sciences and applications (DASA); Chiangrai, Thailand; 2022. p. 1363–1366. <https://doi.org/10.1109/DASA54658.2022.9765024>.
64. Hsu TY, Hsieh CC. A real-time track fasteners classification system based on Yolov4-Tiny model. In: 2022 IEEE international conference on consumer electronics—Taiwan, Taipei, Taiwan, 2022. pp. 41–42. <https://doi.org/10.1109/ICCE-Taiwan55306.2022.9869046>.
65. Ranyal E, Jain K. Detection of rail fasteners from aerial images using deep convolution neural networks. In: IGARSS 2020—2020 IEEE international geoscience and remote sensing symposium, Waikoloa, HI, USA, 2020. pp. 2217–2220. <https://doi.org/10.1109/IGARSS39084.2020.9323596>.
66. Liu Y, Sun X, Hock Lye Pang J. A YOLOv3-based deep learning application research for condition monitoring of rail thermite welded joints. In: Proceedings of the 2020 2nd international conference on image, video and signal processing (IVSP '20), New York, NY, United States. Association for Computing Machinery; 2020. pp. 33–38. <https://doi.org/10.1145/3388818.3388827>.
67. Wang G, Liu Y, Xiang J. A two-stage algorithm of railway sleeper crack detection based on edge detection and CNN. In: Proceedings of the 2020 Asia-Pacific International Symposium on Advanced Reliability and Maintenance Modeling (APARM), Vancouver, BC, Canada. 2020. pp. 1–5. <https://doi.org/10.1109/APARM49247.2020.9209351>.
68. Sresakoolchai J, Kaewunruen S. Prognostics of unsupported railway sleepers and their severity diagnostics using machine learning. *Sci Rep.* 2022. <https://doi.org/10.1038/s41598-022-10062-w>.
69. Zheng Y, Wu S, Liu D, Wei R, Li S, Tu Z. Sleeper defect detection based on improved YOLO V3 algorithm. In: 2020 15th IEEE conference on industrial electronics and applications (ICIEA); Kristiansand, Norway; 2020. pp. 955–960. <https://doi.org/10.1109/ICIEA48937.2020.9248299>.
70. Hu X, Cao Y, Sun Y, Tang T. Railway automatic switch stationary contacts wear detection under few-shot occasions. *IEEE Trans Intell Transp Syst.* 2022;23(9):14893–907. <https://doi.org/10.1109/TITS.2021.3135006>.
71. Jahan K, Niemeijer J, Kornfeld N, Roth M. Deep neural networks for railway switch detection and classification using onboard camera images. In: 2021 IEEE symposium series on computational intelligence (SSCI), Orlando, FL, USA; 2021. pp. 01–07. <https://doi.org/10.1109/SSCI50451.2021.9659983>.
72. Zendel O, Murschitz M, Zeilinger M, Steininger D, Abbasi S, Beleznai C. RailSem19: A Dataset for Semantic Rail Scene Understanding. In: *Computer Vision and Pattern Recognition (CVPR)*; 2019. p. 9.
73. Yilmazer M, Karakose M. Railway condition monitoring and fault detection based on YOLOv4. In: 2022 International conference on innovation and intelligence for informatics, computing, and technologies (3ICT); Sakheer, Bahrain; 2022. pp. 439–443. <https://doi.org/10.1109/3ICT56508.2022.9990640>.
74. Lu X, et al. A segmentation-based multitask learning approach for isolating switch state recognition in high-speed railway traction substation. *IEEE Trans Intell Transp Syst.* 2022;23(9):15922–39. <https://doi.org/10.1109/TITS.2022.3146338>.
75. Chen X, Zhang G, Dong W, Sun X, Cheng X. A novel fault diagnosis method for high-speed railway turnout based On DCAE-logistic regression. In: Proceedings of the 3rd international conference on artificial intelligence and big data (ICAIBD), Chengdu, China; 2020. p. 318–323. <https://doi.org/10.1109/ICAIBD49809.2020.9137500>.
76. Chen Q, Nicholson G, Roberts C, Ye J, Zhao Y. Improved fault diagnosis of railway switch system using energy-based thresholding wavelets (EBTW) and neural networks. *IEEE Trans Instrum Meas.* 2021;70:1–12. <https://doi.org/10.1109/TIM.2020.3029365>.
77. Krč R, Podrouzek J, Kratochvílová M, Vukušič I, Plášek O. Neural network-based train identification in railway switches and crossings using accelerometer data. *J Adv Transp.* 2020. <https://doi.org/10.1155/2020/8841810>.
78. Najeh T, Lundberg J, Kerrouche A. Deep-learning and vibration-based system for wear size estimation of railway switches and crossings. *Sensors.* 2021;21:5217. <https://doi.org/10.3390/s21155217>.
79. Wang T, Zhang Z, Yang F, Tsui K. Automatic rail component detection based on AttnConv-Net. *IEEE Sens J.* 2022;22(3):2379–88. <https://doi.org/10.1109/JSEN.2021.3132460>.
80. Saini A, KG K, Sriram K, Singh D, Singh KP. Machine learning approach for detection of track assets for railroad health monitoring with drone images. In: IGARSS 2022—2022 IEEE International Geoscience and Remote Sensing Symposium; Kuala Lumpur, Malaysia; 2022. pp. 4891–4894. <https://doi.org/10.1109/IGARSS46834.2022.9883490>.
81. Zhu S, Zhang Q, Zhai W, Yuan Z, Cai C. Sensor deploying for damage identification of vibration isolator in floating-slab track using deep residual network. *Measurement.* 2021;183:109801. <https://doi.org/10.1016/j.measurement.2021.109801>.
82. Ilić V, Romić M, Ilić S. Vegetation suppression system on and near the railway tracks based on PLC and deep learning. In: 2022 IEEE zooming innovation in consumer technologies conference (ZINC), Novi Sad, Serbia, 2022. pp. 251–254. <https://doi.org/10.1109/ZINC55034.2022.9840689>.

83. Gasparini R et al. Anomaly Detection, Localization and Classification for Railway Inspection. In: 2020 25th International Conference on Pattern Recognition (ICPR), Milan, Italy, 2021. pp. 3419–3426. <https://doi.org/10.1109/ICPR48806.2021.9412972>.
84. Jahan K, Umesh JP, Roth M. Anomaly detection on the rail lines using semantic segmentation and self-supervised learning. In: 2021 IEEE symposium series on computational intelligence (SSCI), Orlando, FL, USA, 2021. p. 1-7. <https://doi.org/10.1109/SSCI50451.2021.9659920>
85. Chen Z, Wang Q, Yu T, Zhang M, Liu Q, Yao J, Wu Y, Wang P, He Q. Foreign object detection for railway ballastless trackbeds: a semi-supervised learning method. *Measurement*. 2022;190:110757. <https://doi.org/10.1016/j.measurement.2022.110757>.
86. Chen Y, Tong S, Lu X, Wei Y. A semi-supervised railway foreign object detection method based on GAN. In: Proceedings of the 5th international conference on computer science and application engineering (CSAE '21). New York, NY, USA: Association for Computing Machinery; 2021:58. <https://doi.org/10.1145/3487075.3487133>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.