A review of computer vision for railways

Bryan Olivier, Feng Guo, Yu Qian, David P. Connolly

***Abstract*—Modern railways continue to strive for remote and automated methods to improve the visual inspection procedures for their assets. In some cases, these inspections provide new information that could not previously be collected, while in other cases they help them improve upon the quality control, safety, time and costs associated with manual inspection. As such, computer vision continues to find applications for visually inspecting the track, earthworks, tunnels, overhead line equipment and rolling stock. Considering the recent pace of computer vision related developments, this paper seeks to review the state of the art of the field for railways. First, the hardware and data requirements are discussed, focusing on the unique challenges associated with operating optical equipment in a railway environment, such as contamination, power sources and lighting. This also discusses the most common mounting arrangements for camera hardware, including rolling-stock, satellites and way-side cameras. Next, image processing algorithms are discussed, comparing classical approaches and more modern artificial intelligence approaches, for example You Only Look Once (YOLO) and Region-Based Convolutional Neural Network (R-CNN). Then the most common applications for computer vision in the rail industry are analysed. First the track is studied considering computer vision analysis for the detection of different types of rail surface defects on plain line and turnouts, fastener defects, concrete track slab cracking and ballast particle characterisation. Next, the overhead line equipment is considered with applications related to detecting contact loss between pantograph and contact wire, stagger behaviour and defective catenary components. This is followed by discussion of other applications such as rail tunnelling subsidence, tunnel inspection, level crossings, trespass and on-track safety hazards. Finally, opportunities for future research are discussed such as hyperspectral imaging and generative AI, along with possible frontier technologies such as quantum computing.**

***Index Terms*—Railway Computer Vision, Unmanned Aerial Vehicles UAV Drone, optical-InSAR Satellite, YOLO R-CNN, Hyperspectral-Multispectral Imaging, Overhead Catenary Computer Vision, Vehicle-borne Camera Inspection,**

I. Introduction

A railway is a complex system comprising multiple interacting physical assets such as the trains, track, overhead lines, earthworks and stations. Traditionally these assets and their sub-components were, and in many cases still are, visually inspected by trained engineers to assess their physical condition. This approach started in 1840 with the introduction of 'Inspecting Officers' in response to safety incidents on the first commercial passenger train service. In the early days of inspections these involved rudimentary checks of the trains and tracks; for example, 'wheel-tappers' [1] would strike each train wheel with a hammer and depending upon the resulting sound, consider reporting it for removal to check for signs of stress or cracking. As railways became more complex and linespeeds increased the standardisation of visual inspection became commonplace. This often involved dedicated inspection tools and processes aimed at maximising safety and efficiency. For example, visually checking crossings for evidence of cracks or measuring track geometry using optical surveying equipment. More recently the rail industry has been moving towards replacing in-person visual inspection with remote condition monitoring [2]. This is partly due to the desire to: improve rail worker safety by removing them from live railway environments, achieve improved repeatability and to minimise the costs associated with regular manual inspections. This has been helped by advances in camera technology, with some of the assets and camera technologies currently used to inspect them shown in Figure 1.

In the 1900's the switch from steam to diesel and then electrical locomotive power allowed easier access to electricity for on-board camera systems. These have since been deployed, for example on the front of in-service trains to record the driver's view, below the train to detect track component issues, and above the train to monitor overhead line equipment. The advantage of vehicle-borne camera's is that the entire route can be monitored using a small number of camera's compared to placing multiple in the wayside. However, the camera positions are relatively fixed and cannot record all track angles/positions of interest.

A promising area of development to overcome some of the vehicle-borne camera limitations, is the use of aerial photography. One of the first railway applications was in the mid-1900's, when human photographers on-board airplanes and helicopters were shuttled back and forth along future potential railway routes to take black-and-white aerial photos of potential alignments. These individual photographs were then pieced together to map the area and make route planning decisions. Although helicopters are still used, more recently unmanned aerial vehicles (UAVs) have been introduced for the visual inspection of existing rail assets. They are advantageous because they can be remotely operated and have flexible flightpaths thus allowing them to record images of the track

Manuscript received , 2024; revised 2025.

(Corresponding author: Feng Guo)

Bryan Olivier is associated with the Department of Theoretical Mechanics, Dynamics and Vibrations, Faculty of Engineering, University of Mons, Place du Parc 20, B-7000 Mons (Belgium) and the School of Civil Engineering, University of Leeds, Woodhouse, LS2 9JT Leeds (UK). (e-mail: bryan.olivier@umons.ac.be)

Feng Guo is associated with the School of Qilu Transportation, Shandong University, No. 17923 Jingshi rd, 250061 Jinan, Shandong (China). (e-mail: fengg@sdu.edu.cn)

Yu Qian is associated with the Department of Civil and Environmental Engineering, College of Engineering and Computing, University of South Carolina, 300 Main Street, 29208 Columbia, South Carolina (USA). (e-mail: yuqian@sc.edu)

David P. Connolly is associated with the School of Civil Engineering, University of Leeds, Woodhouse, LS2 9JT Leeds (UK). (e-mail: d.connolly@leeds.ac.uk)

Digital Object Identifier 10.1109/TITS.2025.3552011

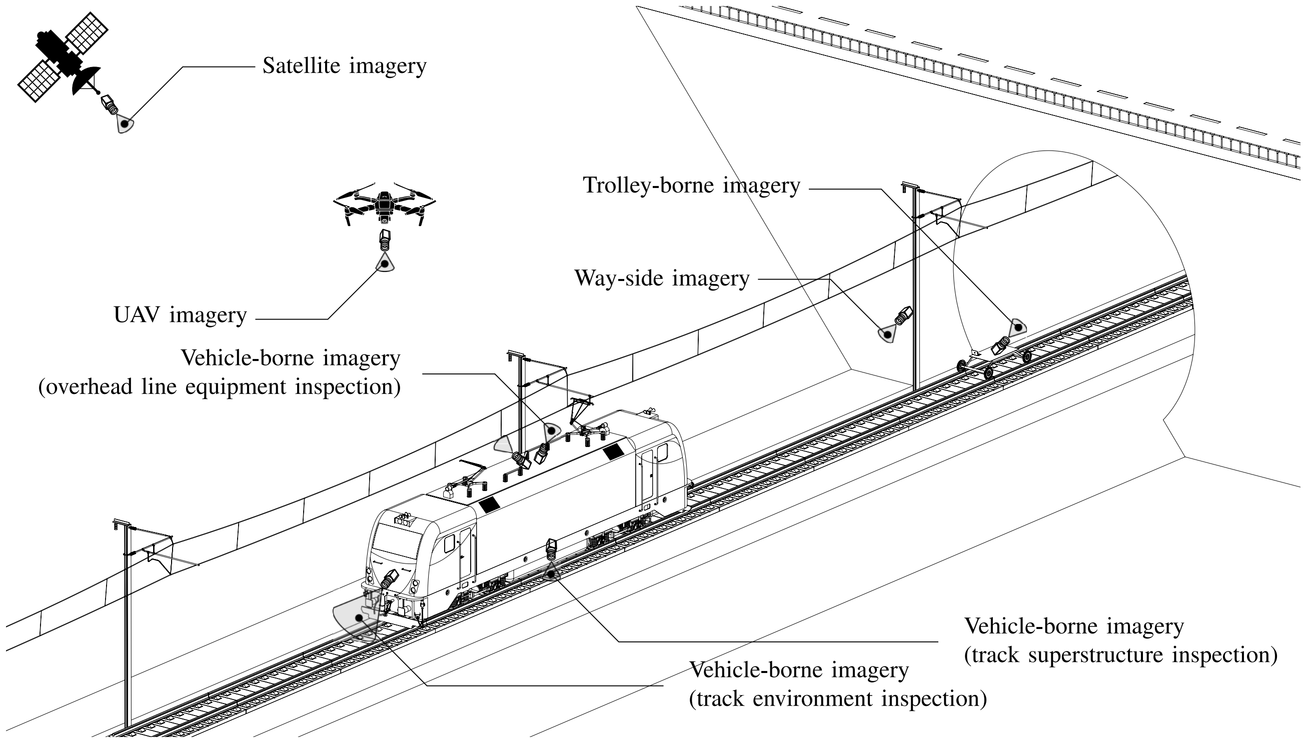


Fig. 1. Overview of camera mountings for railway applications.

infrastructure that would otherwise be difficult to access (e.g. overhead wires). Complimentary to UAVs, satellite monitoring is another promising field of aerial photography. This is attractive because human operators are not required, and it does not require any modification to the train or track systems.

Considering the rapid evolution of imaging technologies and processing techniques, their deployment in the rail industry is also ongoing at a fast pace. Thus, regarding previous review papers on railway computer vision, Oh et al. [3] reviewed computer vision methods, focusing on close railway infrastructure, the lower part of the vehicle, the operation in general and the stations. Further, in 2024, Aela et al. [4], [5] produced two relevant literature reviews. The first one [4] focused on the use of UAVs for railway monitoring while the other [5] focused on vision-based monitoring techniques for the railway superstructure. In addition, Chen et al. [6] briefly reviewed the use of camera sensors for monitoring of railway catenary, while Pappaterra et al. [7] published a review concerning the use of convolutional neural networks for track maintenance. Finally, Cao et al. [8] presented a review of vision-based techniques for the detection of railway intrusions. Compared to these previous reviews, the current paper presents several key advancements. Firstly it provides a more comprehensive review in terms of depth and breadth of railway applications, not just focusing on a single asset type. Secondly, it provides novel review of the practical considerations of camera hardware deployment, a topic absent from the above mentioned review papers.

II. Methodology

The methodology/structure of this paper is to first review different camera types and setups, such as wayside, train-borne, UAV and satellite. Next, image/video data requirements are discussed along with practical considerations for using camera hardware in a railway environment that is a naturally unclean environment, often exposed to harsh weather conditions. Image processing algorithms are then discussed, focusing on artificial intelligence and digital image correlation approaches. Next, applications for differing camera monitoring solutions are presented, including: pantograph/catenaries, track components, tunnels, earthworks and stations. Finally, future research trends are discussed.

III. Camera hardware and data requirements

A. General

When deploying a camera in a railway environment, hardware suitability must be considered. In recent years digital cameras have almost exclusively superseded film cameras in engineering applications. These have electronic sensors that react to light rather than photosensitive material meaning they are easily digitised and readily processed. When considering the choice of camera lens for a particular application, a variety of variables should be considered, including:

- Focal length: dictates the angle of view so should be tailored to capture the entire object under observation. Smaller focal lengths give a wider-angle view while longer lengths give a narrower view.

- Aperture size: dictates the size of the opening and thus how much light can enter the lens. Smaller apertures make more of an image in focus.

Typical consumer grade cameras use standard lenses with focal lengths in the 35mm range. However, specialist lenses are also available such as fish-eye lenses which are ultra-wide, making them sometimes useful when a wide field of view is needed, for example for counting passengers on a platform. Alternatively, infrared lenses filter out all light waves except infrared, making them useful for heat detection (e.g. lineside cable faults).

Technology type: Several imaging related technologies are commonly used in railways. The most common is a standard image sensor collects and measures the reflected or refracted light coming from the environment. In addition is lidar (Light Detection and Ranging) which emits light and then measures the differences between the emitted light and reflected light coming back to the sensor to deduce distance information. Also, classical image sensors may be used in combination with lasers. For example, by operating a registration of the camera with respect to the laser, it is possible to obtain pixel depth information and also perform 3D image reconstruction.

Image sensor type: There are two main types of sensors used in digital cameras, each one having its own characteristics that may be useful for specific railway applications.

- Charge-Coupled Device (CCD) sensors are composed of an array of capacitors connected in rows. Their charge depends on the exposure time and light intensity. When triggered, each capacitor transfers its charge to its neighbour in the row with the final one transferring its charge into an amplifier to be converted into a voltage. CCD can be designed in-line or as a full array of sensors. They are well-suited to line scanning, making them useful for the acquisition of long railway images such as the continuous monitoring of rails and catenary wire.
- Complementary Metal-Oxide-Semiconductor (CMOS) sensors comprise a matrix of photosensors, organized in rows and columns, with each one having its own amplifier. The difference compared to CCD is that the photosensors do not communicate their charge with their neighbours, but instead communicate directly with the register. A region of interest (ROI) can then be specified directly by selecting the relevant photosensors, leading to an increased acquisition speed. This improved efficiency and also reduced cost means CMOS have become widespread in recent years. Regarding railway infrastructure monitoring, a high sensor framerate may be required. Indeed, a certain number of discrete elements in track superstructure or overhead line equipment such as catenary components, fasteners or sleepers are checked with vehicle-mounted cameras that perform image acquisition adjusted to the vehicle speed.

Sensor spectral response: The spectral response of a camera sensor describes its sensitivity to different wavelengths. Classical sensors are usually more sensitive in the spectrum of visible light while other sensors (e.g. lidar or thermal cameras) are more sensitive in other ranges such as infrared. This is advantageous for monitoring phenomenon not observable by the human eye (heat changes for instance). Therefore, thermal cameras have been used to provide information about the condition of the vehicle and track. For example, the detection of lineside cable faults which cause overheating. Further, Chapman et al. [9] mounted a standalone thermal camera on a personnel carrier to monitor the rail temperature and determine the sections of the track which prone to buckling. Kim et al. [10], [11] used thermal cameras to monitor the brake discs and steel axles in railway vehicles.

Sampling frequency and pixel count: Data collected from imaging devices can be sampled at different frequencies and using a different number of pixels depending upon the application. For example, monitoring level crossings may use lower resolution cameras while the monitoring of tunnel lining cracks may use much higher resolution. Although this is related to the need to study cracks in fine detail, it is also a function of the application processing requirements. For example, level crossings require data to be processed immediately to give real-time monitoring, meaning the volume of data must be minimised. In contrast, tunnel lining analysis is less time sensitive meaning the data can be stored on local hard drives which are manually transported to an off-site location for detailed investigation, potentially using a combination of automated and trained expert analysis.

Calibration: Once the images are taken by the camera, two different calibrations, intrinsic and extrinsic, are necessary to allow taken images to represent the actual reality. While using the usual pin-hole camera sensors with lenses, the intrinsic calibration considers the intrinsic characteristics of the camera-lens assembly such as the focal length, lens properties or sensor resolution to correct the obtained images. Moreover, the extrinsic calibration will correct the different types of distortion of the images depending upon the position and orientation of the camera with respect to the observed objects. Usually, these calibrations are performed with a well-known pattern such as a chessboard. Another calibration may be performed when the camera provides colour images. Indeed, a white balance is sometimes necessary when the colour of the monitored object is relevant. When quantitative information is retrieved from the images, these calibration steps are often necessary. However, when qualitative information is observed these calibration steps may be relaxed since the actual geometry of the objects is not necessarily considered. For instance, for some AI-based post-processing techniques, the network can directly be fed with uncalibrated images.

Recording format: Firstly, the resolution of data is dependent upon the desired use case. For example, ballast endoscope technology will often be performed on a one-off type basis, perhaps with the generation of a dedicated site-investigation report. Therefore, the image format and decisions based upon these maybe more suited to trained human analysis rather than automated techniques. In contrast, track images recorded using vehicle-mounted cameras are often recorded in a well-defined and standardised format (e.g. in terms file format and size), that allows machine learning algorithms to readily operate upon it. For certain applications, a requirement of data collection is that it can be integrated within the infrastructure operator's existing database systems and application programming interfaces (API's). This is desirable because it allows for alignment of datasets and potentially the mining of multiple complimentary datasets to generate more powerful conclusions.

Data storage: Railway monitoring can generate large datasets that cause transmission and storage challenges. As a solution, edge computing can perform the data analytics as close to the monitoring data source as possible. For example, instead of sending all data to a processing centre, edge devices such as a smart camera can locally process the data on-board for example using machine learning algorithms. Simple alert signals can then be provided to the responsible asset engineers and the related datasets sent to a data centre for further investigation. In this way, the storage and processing pressures on the data centre are reduced.

Image recording location: It's important for imaging systems to be able to automatically record the location of each image. In the rail industry this is commonly performed using GPS coordinates, and without this information it can be difficult to align datasets along the route. However, railways can often be located in remote areas and/or with topography challenges that can make it difficult to obtain reliable GPS data. Therefore, depending upon the route, alternatives may need to be considered.

Region of Interest (ROI): The region of interest is a subset of an image in which the monitored phenomenon (e.g. cracking) is expected to be located. The choice of this region of interest is important to avoid confusion with non-relevant phenomena happening elsewhere in the image. The ROI can be a fixed region, meaning it's location doesn't change during subsequent image frames, or alternatively assigned dynamically, meaning it can shift between frames, as proposed by Zhang et al. [12] to manage railway obstacle intrusions.

Data failsafes: For applications where image analysis is important for safety critical operation (e.g. rail buckling monitoring), hardware failures must be detectable and impacts minimised. Regarding detection, if there is a hardware failure (e.g. power failure) this can be detected for example using a communication gateway. If instead there is a recording error, this can be detected during data processing, and potentially assisted by deploying pairs of complimentary cameras rather than individual ones, thus allowing for consensus checking. The severity of impact of hardware faults can be mitigated by designing the imaging hardware to be easily and quickly replaced once a fault has been detected.

Data security: For safety critical applications it may be necessary to consider cyber security. This may need to cover both the transfer of data and storage. To do so it is common to adhere to, 'IEC 62443 - Industrial communication networks - Network and system security' and 'ISO 27001 – Information Security Management'. Depending upon national regulations, data may also need to be maintained in a 'forensic ready' state for a given period (e.g. 2 years). This allows for investigations to study the records held in the event of asset and/or system failures.

B. Hardware performance requirements

Railway environments present different conditions to laboratory environments and any imaging hardware must be able to operate under extreme conditions. For example:

Contaminates: Hardware placed within the aerodynamic envelope of rolling stock is exposed to significant levels of dust, dirt and contaminates. If these collect on the lens the imaging ability can be degraded or possibly completely obscured. Therefore, regular cleaning and maintenance must be planned in advance. For certain applications a mechanical stutter can be automatically opened and shut when required.

Lighting: Railways require monitoring during both day and night meaning the image hardware may need to be self-illuminating. Similarly, light intensity can quickly change over short time periods, meaning the imaging solution must be able to handle extreme lighting conditions. An example is under-train camera technology, which can be subject to rapidly changing lighting conditions (e.g. shadows) when the sun is lying low in the sky, or in conditions when the sun is reflecting off a highly reflective railhead.

Weather conditions: Hardware solutions must be capable of performing their functions in the presence of extreme high and low ambient temperatures. For example, to meet industry standards, way-side systems may need to show acceptable operation in a range depending upon typical atmospheric conditions (e.g. -27 °C to +60 °C). They must also operate under conditions of rapid temperature changes, such as 0.5 °C / min. Additional weather conditions that may need to be considered depending upon the location(s) of operation include: wind which may cause camera movement, ice which may cover the lens, solar radiation if the hardware is directly exposed to the sun, and humidity which can cause condensation on the lens.

Mechanical/electrical interference: Camera hardware must be capable of performing its function under typical environmental vibrations experienced in a rail setting and shock loading. Further, the rail environment typically experiences a range of electromagnetic conditions, for which the system must be compatible.

Power Supply: Imaging hardware typically has higher power requirements compared to mechanical sensing methods such as accelerometers. This is because its multidimensional nature means it records and stores larger quantities of data. Vehicle-borne imaging systems have the potential to use on-board power, however way-side systems often lack the availability of a reliable power supply. Therefore, batteries are commonly relied upon, possibly combined with a remote charging technology such as solar panels (or possibly even mechanical energy harvesters in the future [13], [14]. However, current such solutions, unless large, struggle to provide sufficient power to record high resolution images at high frequency.

Deployment/Installation: The deployment of imaging systems should have minimal negative impact on the scheduling of trains. For vehicle-borne systems installation can be performed in depot's and it is also not a concern for satellite systems. However, for any semi-permanent wayside systems, the installation and maintenance times should be low (e.g. $<$15 minutes) to minimise disruption to the lines, and achievable under poor lighting conditions (i.e. at night when track access

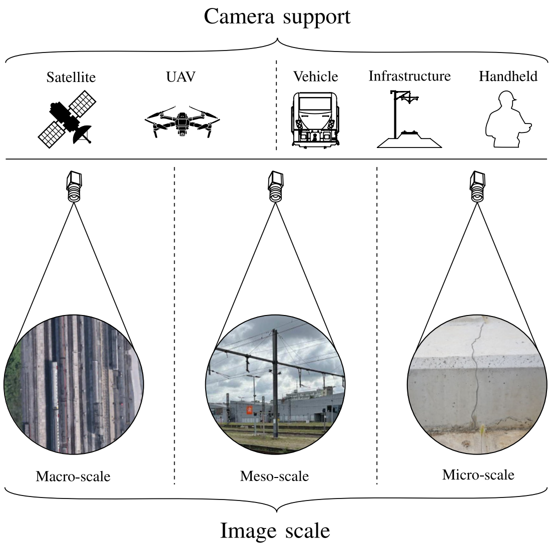


Fig. 2. Different scales of the images taken for railway applications.

is commonly available). Also, unless planned far in advance, access to railway sites is often by foot with limited vehicular accessibility. Therefore, the hardware should be modular where possible, with no individual parts weighing more than Health and Safety regulations.

IV. Camera setup types

A variety of camera setups are used for railway imager analysis. Depending on the application, cameras can be either fixed, handheld or mounted on a vehicle. To distinguish these camera setups, one way to classify them is the size of the ROI. As illustrated in Figure 2, different scales of monitoring can be identified: the macro-scale includes pictures in which large zones (km^2) are monitored, the micro-scale denotes the monitoring of small phenomenon's (mm^2) while the meso-scale contains all monitoring types that are comparable to the human scale (m^2). This section covers the major camera setup types corresponding to this classification.

A. Vehicle-borne cameras

Vehicle-borne imaging is attractive because track access is not required to setup way-side cameras, the entire route can be monitored in a short timeframe using a small number of cameras and the images are relatively repeatable. In Figure 2, vehicles are classified either in the meso and micro scales since they can either be used to detect, as detailed further in the paper, meter-centimeter scale defects such as low-ballast heights or milimeter scale defects such as a crack in the rail. Imaging systems can be installed on either existing passenger trains rather or on a dedicated train. They must be configured to be able to handle the harsh environment in the vicinity of the running gear which contains contaminates that could obscure the lens, and to handle greatly varying light conditions. An example of this is show in Figure 3 (right), where the sleepers on one side of the rail are well illuminated, the sleepers on the other side are much darker. This indicates the presence of a low hanging light source during measurement, compared to Figure 3 (left) which was recorded when the sun was higher in the sky.

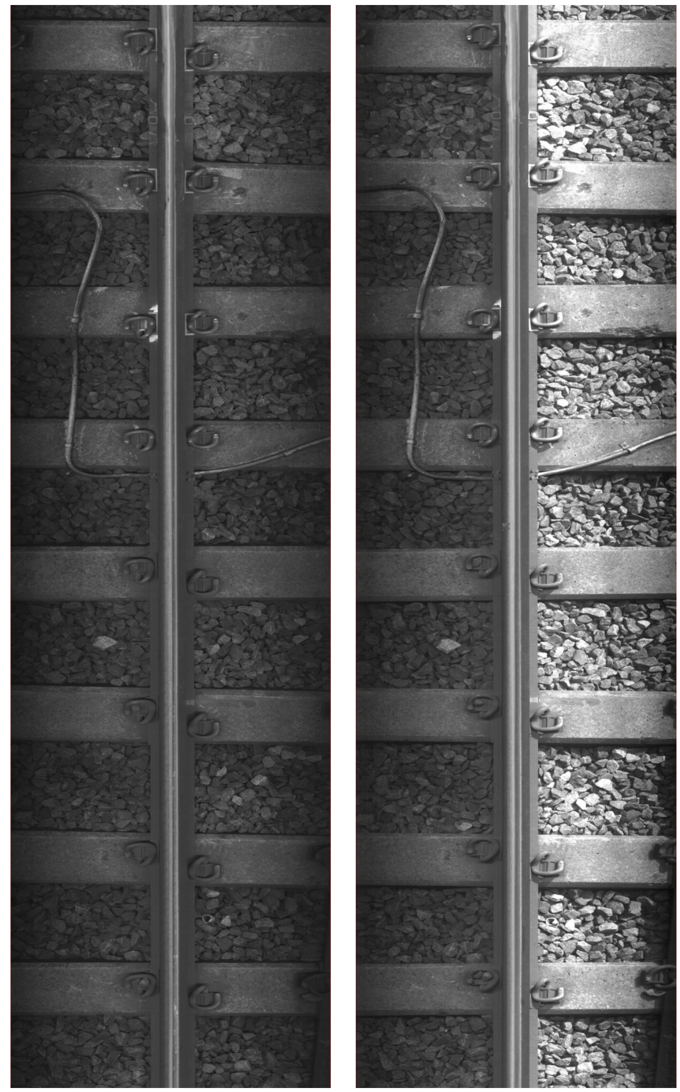
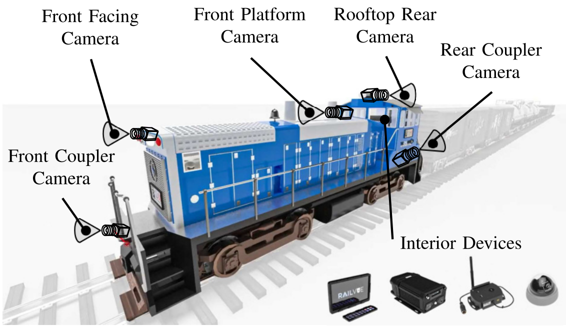


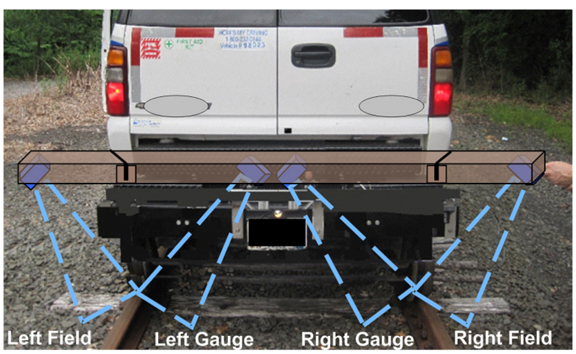
Fig. 3. The effect of lighting on track images recorded using vehicle-borne measurement. Left: recorded during the middle of the day, Right: recorded early in the morning.

Type of vehicle: Depending on the application, it is possible to use different types of vehicles adapted to railways for track monitoring. As illustrated in Figure 4, four example categories are identified. The first category is in-service trains [18] which are passenger vehicles, retrofitted with cameras. However, since they are in-service, it provides a frequent and low-cost way to collect data without having to schedule train movements around the existing timetable. The second category is dedicated measurement trains [19], [20] which can be equipped with a wide range of sensors to monitor the railway environment. However, a challenge is that their scheduling needs fitting around existing train timetables. The third category is self-propelled road-rail vehicles [21]–[23] which use rubber tyres to arrive at the test site and position themselves on the rails, using a secondary set of steel wheels to traverse along the track. These vehicles are often used to



(a)

(b)



(c)

(d)

Fig. 4. Categories of vehicles for track component monitoring. (a) Camera equipped in-service train [15]. (b) Dedicated measurement train (Shinkansen's "Doctor Yellow"). (c) Self-propelled road-rail vehicles [16]. (d) Instrumented trolley [17].

test and validate algorithms or methods developed to monitor track components before being integrated into a completely functional solution designed to be used with the first category of vehicles. Trolleys are another category of vehicle [24]–[27]. Unlike the self-propelled vehicles, trolleys are not necessarily self-propelled and may be manually operated. Other types of vehicles and robots are were discussed by Jing et al. [28].

Type of camera: Depending on the asset to be monitored and the type of vehicle on which it is mounted, the type of camera can vary. Most of the vehicle-borne inspection is performed in a continuous manner, for example to inspect the rails or catenary along the route. For the rails, a laser-scanning camera-based system (different from regular laser-scanning devices without cameras) is commonly used to project a laser line on the surface of the rail. The reason of this is the possibility to retrieve the 3D profile of the rail thanks to the monitoring of the laser beam with the camera. Using a camera alone would only provide a two-dimensional image of the rail surface. Doing so, both top and side of the rail are visible and reconstructed to detect possible surface defects. For such applications the sampling rate must be carefully considered with respect to the train speed. Also, it is noted that other systems such as thermal cameras can be mounted on the vehicle to monitor the heating condition of some internal elements of the vehicle [10].

Location and number of cameras: Depending on the application, the position of cameras on the vehicle is commonly on the roof (e.g. overhead line equipment monitoring), the under-carriage (e.g. track superstructure monitoring) or the front of the vehicle (e.g. route monitoring). Further, for each application the quantity and the way they are grouped and attached may vary. The relative position of the cameras (and their laser beams) is also important to avoid any blind spots. For example, Figure 5 shows the arrangement of multiple cameras and lasers attached on the under-carriage. In this example both sides of the measurement device (for the imaging of both rails) are coupled in an all-in-one device. Alternative devices also exist for which both sides of the track are inspected using independent cameras whose fixation is independent rather than via a common beam. The Figure depicts camera-based laser scanning devices meaning that the camera captures the image of the rail on which a laser-beam is projected. Again, depending on the application, accelerometers, gyroscopes or other sensors may be used with the cameras to retrieve additional data such as the position of the cameras with respect to the track. For instance, grinding check operations can be independently operated by different devices on each side of the track. Independant devices with several cameras are used to control the fishplates and joint bars on both rails and even rail corrugation detection. All-in-one device with one camera on each side may be used for detection of gauge-side defects on the head of the rail. Finally, the rail profile can be measured using an all-in-one device with two cameras with their lasers on each side. Cameras in the centre of all-in-one devices are also used to detect defects on fastenings, sleepers and even trackbed (e.g. [29]).

Datum and device registration: For most applications it is required to precisely geolocate images to provide relevant

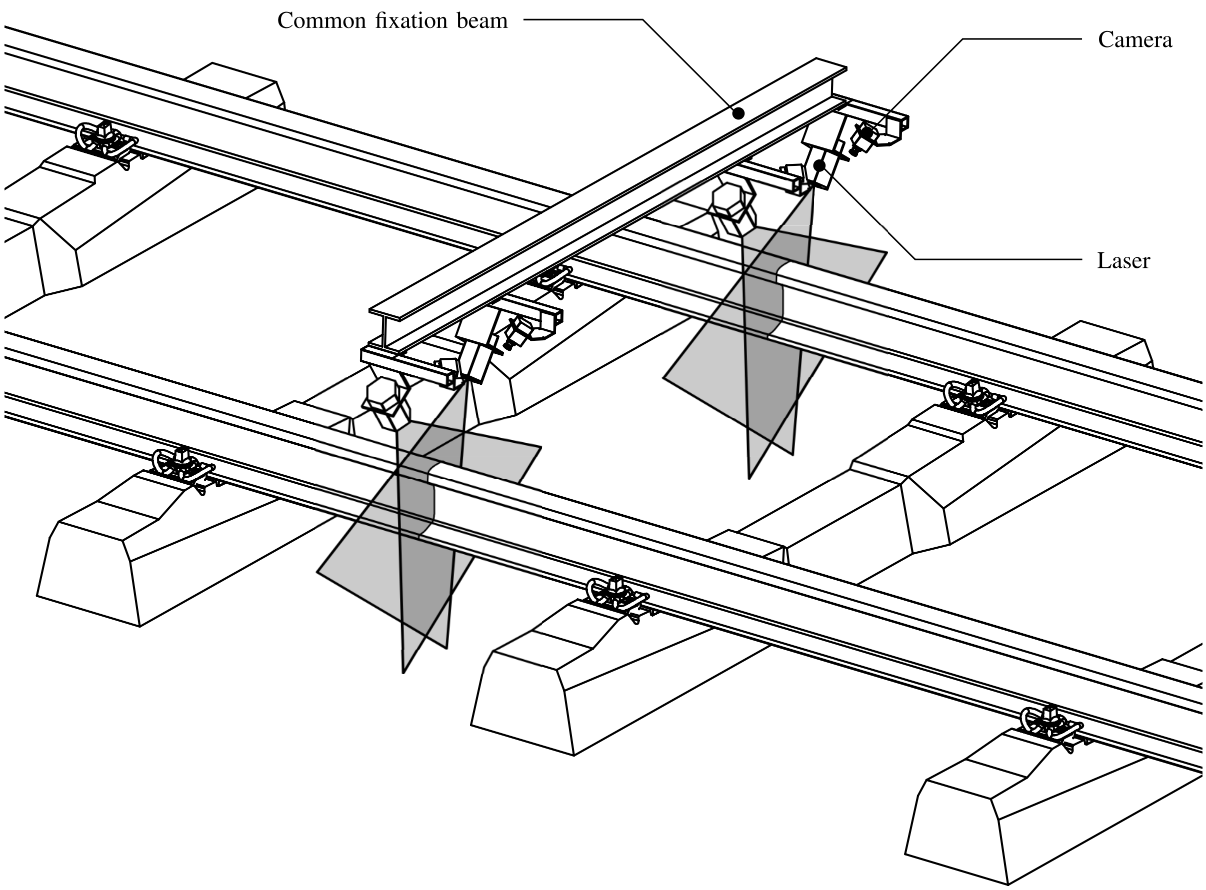


Fig. 5. Illustration of camera-based laser-scanning devices for track superstructure inspection.

information to maintenance teams. This may be via curvilinear abscissa along the track from a reference point or alternatively GPS coordinates. In some cases, more advanced techniques can be used (e.g. SnakeGrid [30]).

B. Unmanned aerial vehicle cameras

Unmanned aerial vehicles (UAVs) have the potential to provide flexible image recording of railways [4], both in terms of positioning and orientation (Figure 6). For example, compared to a vehicle-borne system, a UAV can capture images while hovering at an almost arbitrary position. Compared to satellites, UAVs can also be used on arbitrary dates/times unrelated to the earth's movement. In Figure 2, UAVs are classified both in the macro and meso scales because their monitoring range can vary with the height of the UAV as well as the camera resolution. For example, tens of meters of railway track length can be monitored for vegetation growth analysis or alternatively centimetre-scale parts of the track can be observed such as the overhead line equipment.

UAVs are commonly used to perform ah-hoc track inspection. For example, to detect trespass, vegetation growth and artefacts (e.g. leaves) on the line. They are also used for surveying sites after incidents (e.g. landslips as shown in Figure 7) to determine the most appropriate response. They have applications such as inspecting overhead wires and detecting drainage problems (e.g. standing water and/or blocked drains). However, there are challenges related to UAV inspection. For example, they are negatively affected by weather conditions such as strong winds and have relatively short battery lives. Also, aviation regulations can be complex, often confining operation to visual-line-of-sight (VLOS) operation which can be limited to circa 500m from the operator. Further, regulations usually place restrictions on flying near airfields.

From an imaging perspective, UAVs fly at low heights which can cause challenges in post-processing. For example, the images can be affected by camera lens aberration, while the GPS location accuracy induces lateral and vertical error positioning errors. To reduce certain types of errors, ground sampling points can be used, however these require significant additional time and effort in data collection. Further, ground sampling distance must also be considered to ensure the correct resolution is captured. If too high then adequate image detail will not be obtained, while if too low then the processing and storage time/cost of the data will increase. In an attempt to overcome some of these post-processing challenges, Tong et al. [32] developed an anchor-adaptive railway track detection network, consisting of a dual-branch structure for the full-angle railway track detection. A balanced transpose co-training

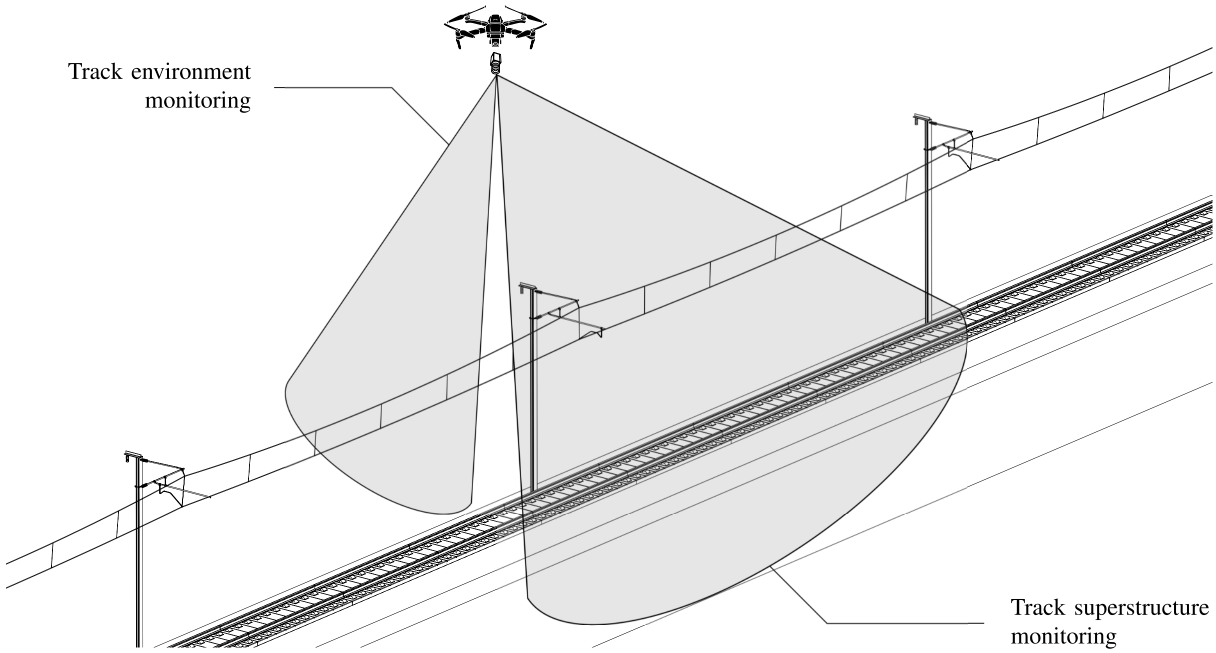


Fig. 6. Using a UAV to monitor the track environment.

Fig. 7. Example landslip captured using UAV surveillance [31].

strategy and customised transposed consistency loss were used to reduce the computational complexity.

Another challenge relates to weather conditions. For example, under conditions such as haze or fog, it can be difficult for UAVs to record images. Thus, Wu et al. [33] designed a dehazing network for restoring the acquired inspection images. Improved residual blocks were developed, aiming to perceive features at different scales. A loss function considering structural similarity was proposed to preserve more information in the image.

Most commonly UAV images are post-processed offline from the drone, because UAVs have limited on-board processing power, however there are potential advantages of performing processing on-board. For example, Tong et al. [34] developed a deep fully decoupled residual convolutional network with the aim of minimising the processing requirements of the recorded images. The processing used three sequential convolutions and a customised auxiliary line loss to reduce the computational requirements. Similarly, Huang et al. [35] developed an intrusion detection system for railways, using a fused LSTM-convolutional block.

C. Helicopters

Compared to UAVs, helicopters typically have longer operating ranges and can stay airborne for longer periods of time. Also, they are not bound by line-of-sight regulations because they have a trained pilot on-board. Further, they are less effected by adverse weather conditions (e.g. strong winds) compared to UAVs and are useful for surveying hilly or mountainous terrain. Due to their size, they typically have high performance camera systems. For example, a forward-looking gyro-stabilised infra-red camera is shown in Figure 8, with multiple changeable lenses, which provides high-definition images.

Helicopters are commonly used for performing visual inspections over long lengths of track, for example, overhead lines and earthworks. According to Figure 2, helicopters would then be classified in meso and macro scales of monitoring. They are also well suited for inspecting assets which are ample in number across a network, but often at large distances between them (e.g. heated points systems). Compared to UAVs they cannot be used for internal inspections (e.g. tunnels) and are less frequently used for ad-hoc inspections and incident response.

D. Satellite imaging cameras

Two families of satellite imaging techniques are predominantly used in railway monitoring. There are optical imagery and radar imaging. Satellites like Sentinel-1 [36] are dedicated to Synthetic Aperture Radar (SAR) imagery while satellites like Sentinel-2 [37] or Pléiades [38] are dedicated to optical

Fig. 8. Left: Example of an helicopter equipped for railway monitoring, Right: the helicopter camera, a Star SAFIRE 380-HD camera [31].

imaging. The difference between them majorly comes from the type of wavelengths used and their active or passive characteristic as detailed below. Satellite imagery is usually used to monitor large scale phenomenon and is therefore classified in the macro-scale section of Figure 2.

Optical remote sensing techniques. These use optical sensors to measure the light naturally emitted by a portion of the earth [39]–[42] . This kind of sensing can be described as passive since it is measuring the reflection of ambient illumination or the reflection of the sun by the earth. The resolution, that often depends on the wavelength ranges used [37], of the provided images can vary from coarse to very high. Previously, the IKONOS satellite [43] provided high resolution images, however newer satellites can achieve so-called Very High-Resolution imagery (VHR) (e.g. the Pléiades Néo satellite [44]). Image data has been used for applications such as assessing the risk of vegetation (e.g. trees and leaves) falling on railway lines. For example, [42] used Sentinel-2 image data with a normalised vegetation index which included factors such as tree type, height and distance from the railway to determine the health of nearby vegetation. Some example railway satellite images, overlaid with InSAR data (see below) are shown in Figure 9, produced by Berhard et al. [45] that conducted PS-InSAR analysis to evaluate the surface deformations along railway tracks. A challenge with optical imagery is that images are frequently obscured by cloud cover. Therefore although a satellite revisit time might be quite regular, depending upon the climate of the region under study, not all images will be suitable for analysis.

Synthetic Aperture Radar (SAR) imagery techniques. Instead of using visible light wavelengths, these use specific wavelengths of energy that can pass through objects such as clouds [46]. The SAR sensor can be described as active compared to classical optical satellite imagery because it emits wave energy and then measures the reflection (backscatter) produced by the earth. Each pixel represents a small portion of the Earth's surface using a complex number with amplitude and phase. The working principle is the measurement of the displacement of an object at the surface of the earth using signal phase variation between two images taken by a single satellite at different times [47]. Interferograms are the difference between two SAR images taken at different times

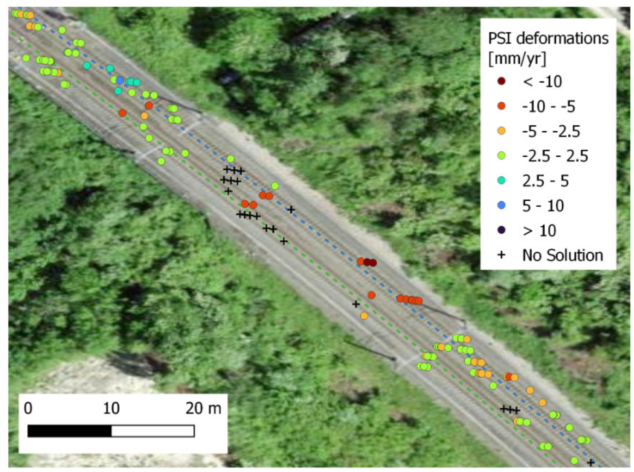


Fig. 9. Example of combining satellite imagery and PS-InSAR data. Reproduced from [45].

coming either from multiple measurements in different points of the same pass or from multiple measurements along the same trajectory but for different passes. The interferogram, containing the phase shift for each pixel of the image, is used to determine the variation of the distance between the satellite and the measured object. If enough temporal images are available, it can be possible to achieve measurement accuracy in the order of millimetres [46], [48]–[53]. Two sub-categories of the InSAR methods are commonly used to monitor railway infrastructures: persistent scatter (PS) [46], [48], [49], [51] and small baseline subsets (SBAS). Persistent Scatter uses a set of SAR interferograms of the same area captured from different passes of the same satellite. It isolates pixels with consistent high signal-to-noise ratio (called persistent scatters) in all interferograms to compute the change in displacement of these specific regions of the inspected area. It is commonly used in urban areas where persistent scatters are more densely available and in cases where corner reflectors can be installed. Alternatively, SBAS creates interferograms between all image pairs which meet certain criteria (e.g. days between recordings). It is commonly used for larger and non-urban areas where the number of persistant scatters is minimal. It is also possible to combine PS and SBAS, for example to identify suitable pixels for analysis. Some applications include detecting landslides along railway tracks [52], studying bridge collapse [54], mapping earthwork subsidence over an entire railway network [49] and studying rail embankments formed from peat [55].

Qualitative or quantitative imaging: Wickramasinghe et al. [40] studied the possibility of using high (2m resolution multispectral with several wavelengths bands) and high-resolution (0.5m panchromatic, visible light wavelengths range) images to monitor a railway construction project. In their work, they used Computer-Aided Designs CAD models as references to establish an approximation of the construction locations. Their aim was qualitative inspection, meaning sub-millimetre resolution was not necessarily required and classic VHR images was sufficient. Another qualitative example of the usage of passive optical satellite imagery was developed by Kucera to monitor the threats of falling vegetation on railway infrastructure [42]. The idea was not to measure precisely any characteristics of the infrastructure, but to qualify a threat index thanks to the satellite images. In certain cases, quantitative measurements are performed and require a high precision technique. For instance, Chang et al. [48] proposed a method based on InSAR imagery to determine, with millimetre precision, the creep of railway tracks in lateral and normal directions.

Weather: Clouds and rain can negatively affect the accuracy of InSAR because the refract the beam causing a phase delay. Because clouds and rain are closer to the ground than the satellite, the phase shifts appear as surface deformation. Weather models can be used to correct for these shifts.

Surface water, snow vegetation: SAR accuracy can be affected by vegetation on the ground surface, which is likely to be constantly growing and changing density. Also, if there is even a thin film (e.g. several millimetres) of water on the surface, for example after a rainfall event, this can also affect accuracy. Similar is true for snowfall.

E. Handheld cameras

Types of devices: A wide variety of handheld devices have been proposed for railway inspection, particularly for track applications. For example, Digital Single Lens Reflex (DSLR) cameras have been used to take high resolution 2D images of track components such as ballast [56]. Alternatively, smartphones typically have on-board Global Positioning Systems (GPS), inertial sensors and more recently lidar, all of which can be used for railway inspections. For example, Zhang et al. [57] monitored ballast condition. Moreover, smartphones have the possibility to directly communicate with a cloud server, through GPRS to 5G communication protocols, to perform cloud computing for image post-processing. They have also been fixed to the windshield of railway vehicles in order to take pictures and videos of the environment when the vehicle is moving (e.g. [58]). Alternatively, Bowness [59], [60] used a portable webcam connected to a small telescope to monitor rail movements. Instead of taking snapshots, a video was taken and then split into images to allow post-processing. While DSLR and smartphones can store images in their internal memory, webcams usually require an external device allowing local or cloud storage of the data. While Bowness studied track deflection along track, Le Pen [61] used the same technique, with a USB handheld camera, to monitor track deflection of level crossings specifically. Wilson [62] used an infrared camera to perform pulse eddy current thermography of rail to detect surface defects. According to Gartrell [63], US Army uses Electronic Rail Inspection Data System (ERIDS) to perform track inspection as well as defect detection. This kind of handheld devices combines GPS and geographical software with a built-in camera which provides photo documentation of track defects linked to their exact geographic locations. Handheld 3D scanners can also be used to study rail and wheel defects, using vertical-cavity surface-emitting lasers (VCSEL) and white light to generate 3D point clouds. Mishra et al. [64] used an autonomous dynamic penetrometer (PANDA) to monitor the conditions of track substructure layers. It consisted of hollow metal rods instrumented with strain gauges that computed, with the help of a standardised hammer, the depth of penetration and the energy input from each impact to drive the rods into the ground. Then, geo-endoscopy was performed by inserting an endoscope camera inside the hole the rods to monitor the condition of the ballast and its sublayers. Other studies [65]–[68] used the same device for track-bed maintenance and characterisation purposes. Another method dedicated to ballast sampling is Automatic Ballast Sampling (ABS), where heavy duty metallic tubes are sunk into the trackbed up to several meters' depth using a hydraulic jack. Once the cores are extracted, a transparent plastic sleeve containing the sample is extracted from the metallic tube. Images are then taken and studied to determine the depths of each trackbed layer and ballast fouling characteristics. Considering these applications and according to Figure 2, handheld cameras are mostly used for meso and micro scale monitoring.

Datum and device registration: Compared to devices mounted on vehicles or trolleys whose position references are given by both rails, difficulties to register the local position of the device with respect to the image taken which may lead distance estimation problems. However, this is not necessarily required for every application when a qualitative inspection is performed instead of a quantitative one. Bowness [59], [60] that used a portable webcam to monitor rail deflection chose to provide an absolute datum to the webcam by attaching it to a rod grouted into the bottom of a 20m deep hole. Moreover, this hole was also located sufficiently far away from the track to avoid any disturbance coming from the ground-borne vibration generated by a running vehicle. For measurements along long distances, the registration of the data according to its position on earth becomes necessary to localise precisely the defects that may arise along the track. As stated by Zhang [57], the GPS position of images taken along the track can be sent through network communication to a cloud server. The position along track remains similar as the problems encountered by vehicle-borne imagery.

Safety considerations: Monitoring track components using handheld devices requires an operator to take either images or videos and therefore implies the presence of a person on or near the track. As stated by Rui [69], the current monitoring of cracks in concrete sleepers may be performed using an optical comparator or a handheld microscope. Replacing these devices by trolley-mounted cameras with an automatic algorithm allowing crack detection could lead to a safer monitoring of the track.

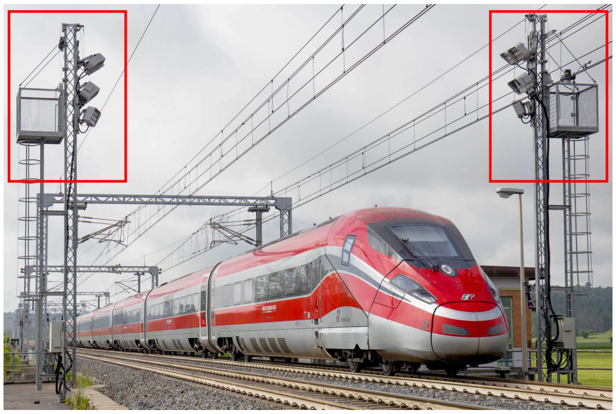


Fig. 10. Way-side pantograph imaging. Cameras and illumination equipment shown in red boxes. Reproduced from [70].

F. Way-side cameras

Way-side cameras are most commonly used for rolling-stock monitoring or track monitoring. Regarding rolling-stock, high resolution camera systems can be placed in the way-side to monitor aspects such as overheating ball bearings [71], pantographs (e.g. chipped contact strip or missing horns) and running gear (e.g. missing bolts or open ports). This can be cost effective compared to fitting individual trains with dedicated camera systems. The cameras and their illumination equipment are typically supported by permanent structures with dedicated steelwork. For example, using towers on either side of the track, as shown in Figure 10, or even a portal frame spanning over the railway with imaging equipment in the span centre to capture high quality recordings. Commonly, due to the expense of such structural modifications and the minimal number of such camera systems required on a network, a reliable power supply is routed to the system.

Regarding track monitoring, way-side cameras are sued for applications which require semi-continuous monitoring over longer timespans, for example, rail buckle detection, level-crossing monitoring ([72]–[74]) and trespass. Due to the need for long-term deployment power supply can be a challenge. In some locations photovoltaic cells can be used, however even when using larger cells, resolution and framerate will likely be constrained due to camera power consumption requirements. Alternatively, in some strategic locations it can be possible to connect to mains power, for example via a nearby signalling system. Another issue is the physical placement of the hardware because cameras typically require some degree of elevation to capture their surroundings. If they are self-supporting then they must resist the vibrations induced by train movement and wind, both of which can negatively affect image quality. These can be similar to the structures used for rolling-stock monitoring, however for many way-side applications such cost is prohibitive. Alternatively, cameras can attempt to take advantage of existing periodically spaced furniture along the railway line. This may include fencing, electrical boxes and uprights supporting overhead line equipment. Depending upon the likely vibration expected from these structures, for certain applications, it is possible to fix cameras to them.

V. Image processing algorithms

This section aims to provide general information about the algorithms used for railway image post-processing. This non-exhaustive list covers the some of the more common algorithms used to perform object or defect detection.

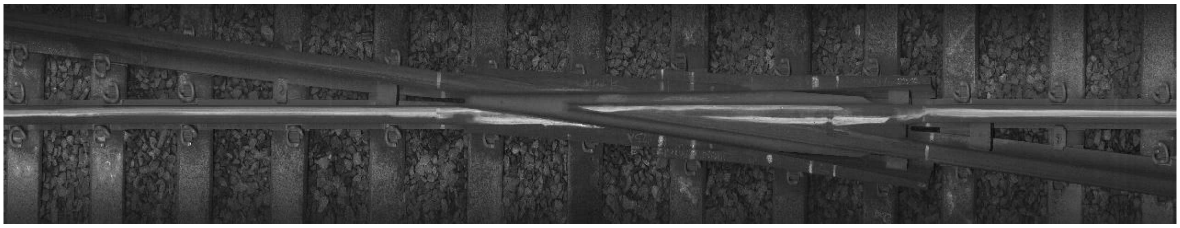
A. Classical methods

Classical post-processing techniques are used to modify images in such a way that the desired objects or defects are more clearly visible. Most of these post-processing techniques consist of a modification or filtering of the pixel information within an image. The basics of these image processing algorithms are listed below. Some examples of these algorithms are illustrated in Figure 11, as applied to a railway crossing.

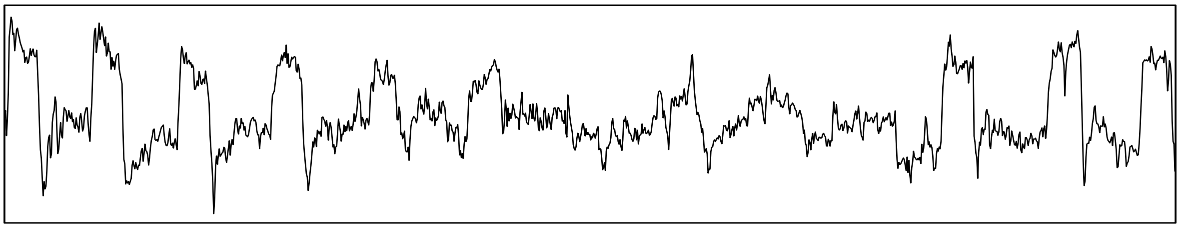
Grey histogram detection: Jie et al. [75] proposed a real-time image processing method based on the grey histogram variation on the columns of an image to monitor rail defects along the track with a vehicle-borne camera. In this case, each bar of the grey histogram represented the average value for a rail section whose width corresponded to a pixel. By studying this histogram, it is possible to study the position of defects along the rail.

Image thresholding: Once an image is transformed into its greyscale equivalent, thresholding can be performed. This means that depending on the intensity of grey, each pixel will be either set to either white or black, depending upon whether it is above or below the specific threshold. While a constant threshold across the entire image is the simplest thresholding methodology, more complex variations exist. For instance, the threshold may be set locally or there might be a threshold for each RGB channel of a colour image. As an example, an enhanced segmentation method was developed by Otsu [76] consisting of a locally optimized intensity threshold for a grey scale image. Further, Li et al. [77] developed a real-time visual inspection system for discrete rail surface defects and compared its efficiency with other thresholding methods such as Otsu's algorithm. Similarly, Aydin [78] used an algorithm based upon Otsu's thresholding method to detect pantograph arcing that may occur with the catenary while they are nearly in contact or when the contact is slightly disrupted.

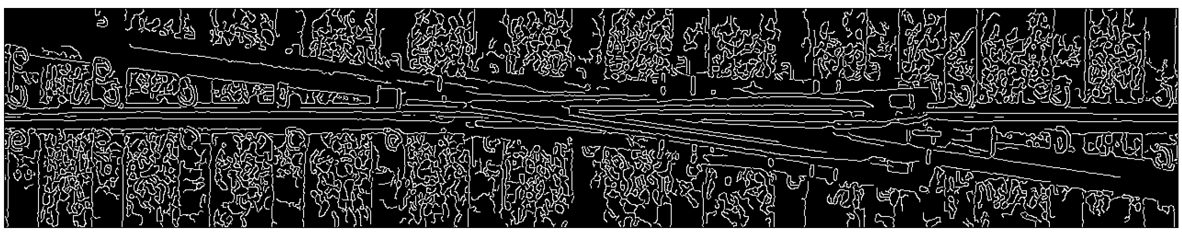
Edge Detection: This technique is used to detect image imperfections or discontinuities, often related to the presence of an object in an image. For example, on a greyscale image, edges are spotted by the computation of the gradient between adjacent pixels. In the simplest form of edge detection algorithms, the gradients can be computed either along the pixel rows or the columns of the image. Therefore, two different transformed images are obtained which provide information on discontinuities with respect to the pixel rows and the columns. More advanced techniques that use filters and convolution



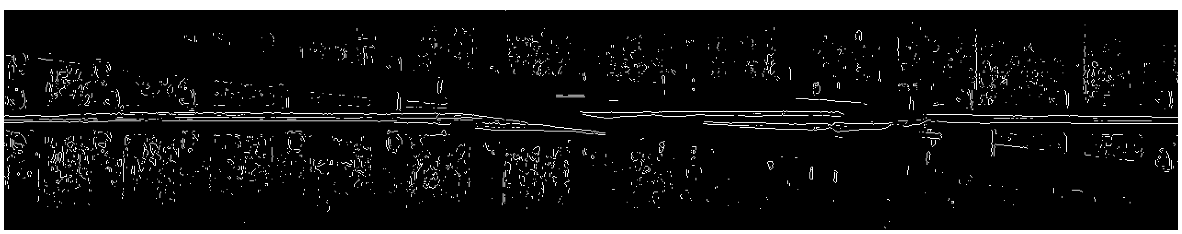
(a)



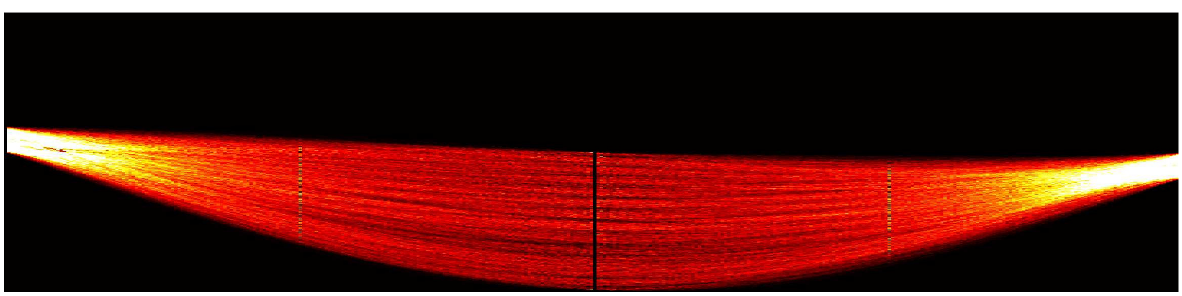
(b)



(c)



(d)



(e)

Fig. 11. Example outputs of several raw image processing techniques applied for railway crossing analysis. (a) Original railway crossing. (b) Mean grey histogram. (c) Canny edge detection. (d) Sobel edge detection. (e) Hough transform.

operations are also available such as Sobel or Canny edge detection algorithms. For instance, Min et al. [27] used Canny edge detection to isolate defects on the rail surface. Shi et al. [79] also used edge detection to monitor rail surface defects but with either Sobel or their own enhanced Sobel detection algorithm.

Hough Transform: The Hough Transform was initially developed for line detection in images and is typically applied after an image has already been transformed, for example using Canny edge detection. The output resulting from the Hough transformation of an image is no more an image, but a graphical representation of the Hough matrix associated with the image. This matrix can be displayed as a picture where the x-axis corresponds to an angle and y-axis to a distance (as depicted in Figure 11). These angles and distances are representative of lines with gradient and distance from the origin of the image, which are given by the distance/angle couples of the Hough matrix. Landi et al. [80] used the Hough transform of images taken by vehicle-borne cameras to monitor the pantograph/catenary system under a thermal point of view. The Hough transformation was used to detect straight lines on the pantograph to determine the pantograph/catenary contact point. As in Figure 11, Li et al. [23] used the Hough transform to detect near-vertical lines in track images, corresponding to vertical edges of sleepers, thus facilitating their detection.

Gabor Filter: Initially introduced as a one-dimensional transform, the Gabor transform aims to study the local behaviour of harmonic functions by windowing them using a Gaussian function. While generalised in two-dimensions to be applied to an image, the Gabor filter is associated with an axis going through the image to determine whether there is any repeatable content along this specified direction. Mandriota et al. [81] used the Gabor transform to detect repeated wear patterns on the rail surface such as rail corrugation. Similarly, Huber et al. [82] used the Gabor filter to analyse the surface of the rail, combining it with 2D image texture and a 2.5D analysis of surface disruption to detect rail surface defects. Resendiz et al. [24] used the Gabor filter on the lateral and top views of the track to detect periodically occurring objects such as fastening systems and sleepers. Alternatively, Guclu et al. [83] used the Gabor transform to detect the railway track as an input to the trajectory control of a UAV.

Colour detection: Depending on the type of camera sensor, the acquired images can either be grey-monochrome or colour images. The pixels of monochrome images contain an intensity which corresponds to light exposure of a physical sensor. Using a filter composed of red, green and blue cells (or Bayer pattern) on a monochrome image sensor, it is possible to retrieve an image with lower resolution, but with pixels containing an exposure intensity corresponding to all red, green and blue colours. While most classical post-processing techniques operate on converted grey-scale images, it is also possible to work with colour. For example, while searching for pixels with a specific colour in an image, it is possible to perform a thresholding on one of the three RGB image channels. If the colour of the object is constant and known, it is also possible to transform the RGB intensities into the HSV representation which provides information on the hue, saturation and value. Again, using thresholding on one of these three parameters, it is possible to isolate certain details of an image. For example, Marmo et al. [84] performed railway sign detection and classification using classical image processing techniques. They used RGB-HSV colour transformation after the sign detection to determine the colour of the light included in the sign. Tastimur [85] used a combination of YCbCr, another colour representation of images, and other classical image processing techniques such as edge detection and the Hough transform to detect, from vehicle-borne cameras, level-crossings in railway paths. In the same work, Tastimur also used an RGB-HSV transformation as well as other image processing techniques to perform on-track foreign object recognition.

Change detection: If an analysis cannot be performed on a single image, it is sometimes possible to work with a comparison between two or more images. This comparison results in a new image on which post-processing techniques (e.g. edge detection) can be applied. For example, Sacchi et al. [86] proposed a simple subtraction between the pixels of two monochromatic images, one being the reference image and the other one being the tested image. The aim was to develop a method for detecting abandoned objects in certain railway environments. Alternatively, Zheng et al. [87] developed a methodology to monitor the rail track gauge using image comparison to eliminate non-relevant information and objects present on images.

Stereovision: When using multiple independent cameras to take pictures of the same object, it is not really an image transformation but more an image combination that leads to a computation of the depth of the object in both images. A 3D point cloud can then be reconstructed for each pixel in a specific zone covered by both cameras. For example, Niu et al. [88] used this stereoscopic technique to reconstruct the surface of the rail to deduce three-dimensional information about rail surface defects.

Digital Image Correlation: This is the comparison between images taken using the same camera position. It can be performed either in 2 or 3 dimensions (for example using stereovision with two fixed cameras). The change detection compares only, pixel by pixel, the pixels of two or more images in order to detect what changed exactly between a series of images. The DIC however requires texture on the monitored object to properly identify the regions that experience displacement between two different images. Moreover, DIC focuses on the determination of a displacement rather than the pixel changes in the images. For example, Farahani et al. [89] used 3D-DIC to monitor the deformation of a printed model of a railway tunnel. They added a black dot target near a region of interest to allow for DIC analysis to be performed.

Independently, edge detection, Hough transforms, Gabor Filters and colour detection are useful for the detection and identification of an object in an image. Further, thresholding and digital image correlation are useful for object change detection. Other techniques such as grey histogram detection or stereovision can be suitable for both. In any case, the combination of multiple of these techniques oftens leads to better results. As an example, Li et al. [23] first used the Hough transform to detect a tie in an image and thus limited the region of interest in the next step searching for the anchors. Moreover, these classical processing techniques can be used to post-process the images to prepare them for the use of AI-based techniques for potential improvement of the detection in terms of accuracy and/or efficiency. For instance, Emoto et al. [90] used a combination of grayscale thresholding, canny edge detection as well as Hough transform to prepare images to be used within a YOLO algorithm to detect wheel surface defects.

B. AI-based algorithms

In recent years, deep learning has shown to offer advantages compared to the traditional image processing techniques used

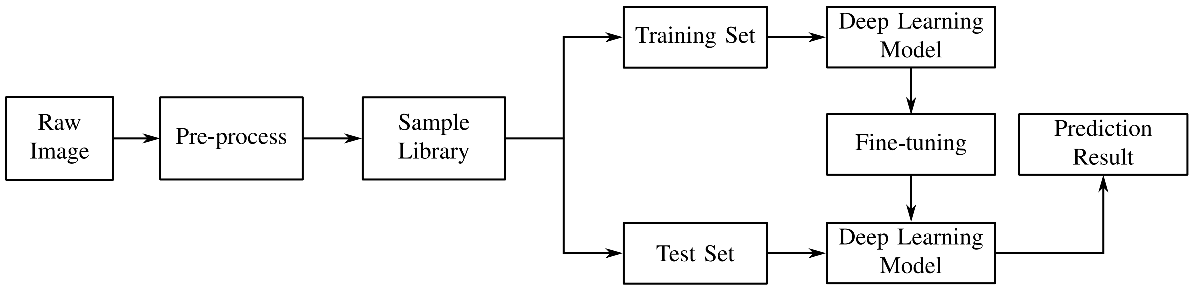


Fig. 12. The general workflow of object detection and semantic segmentation.

for railway applications. There are two commonly adopted deep learning-based approaches including the object detection method and the semantic segmentation method which excel at automatically learning features from raw image data without the manual feature engineering and the learning process can be executed in an end-to-end manner. Furthermore, these sub-approaches have a superior scalability that can assist the developed model to be deployed on edge computing-based devices with limited computing resources. The general workflow of these two models can be found in Figure 12. Besides, Vision Transformers apply the Transformer architecture which is originally developed from natural language processing tasks to image or video data. They offer competitive performance to CNN while providing superior capacity in modelling long-range dependencies within images through self-attention mechanisms.

Object detection-based methods: These detectors aim for object features to be automatically learned and extracted from raw data without the need for labelling or complex algorithm configurations. They are commonly divided into either one-stage or two-stage detectors. The major difference between them is that the one-stage object detector executes the prediction directly without generating potential regions or areas within an image that are likely to contain objects of interest. Alternatively, the two-stage object detector assigns categorical labels to inputs as well as the refinement in the proposal stage. In railway applications, the most used approaches for identifying and locating multiple targets in a single image, with classified categories and regressed location coordinates in a bounding box, are Region-based Convolutional Neural Network (R-CNN) and You Only Look Once (YOLO) detectors. For instance, He et al. [91] developed a R-CNN model to detect train obstacles through the adding of context extraction module, content-aware reassembly of features module, and the parallel design into the feature extraction network. In addition, based on the basic YOLO network, Xu et al. [92] proposed the AED-YOLO network to locate small components in the catenary system with the improved bidirectional feature pyramid network and the asymmetrically effective decoupled head. Wang et al. [93] combined Swin Transformer with Mask R-CNN for rail components diagnosis in railway system. Along with the Swin Transformer, the target can be located and detected. While with the Mask RCNN structure, the object can be segmented. Gao et al. [94] developed Cas-VSwin Transformer for rail surface defect detection. A new window shift scheme was developed to improve the feature transfer between neighbour windows. Through the comprehensive tests, the developed model achieves satisfying results on both of the bounding box and the mask accuracies.

Semantic segmentation-based methods: Semantic segmentation involves assigning a specific class label to every pixel within a given image, effectively representing each pixel according to its corresponding category. This process typically uses three steps to perform segmentation. Firstly, classification is used to recognise various categories present within the image. Secondly, localisation is used to identify target objects, often accompanied by the creation of bounding boxes. Lastly, segmentation is used to group pixels with matching categories. In applications within transportation infrastructure and surface defect inspection, popular semantic segmentation models like SegNet, UNet and DeepLab are commonly employed. For example, Chen et al. [110] developed the ERTNet for railway track region segmentation aiming to detect the intrusion behaviour in a real-time manner. With depthwise convolution, channel shuffle, and feature-matching-based cross-fusion decoder, ERTNet attempts to achieve a balance with respect to segmentation accuracy and computational efficiency. In addition, Feng et al. [111] developed a lightweight, yet efficient railway region segmentation module based on the encoder-decoder structure. Based on the self-correcting feature fusion module and the cascade structure, context and spatial information was extracted. Wan et al. [112] developed a awo-stream Swin Transformer Network (TSSTNet) for salient detection of no-service rail surface defects. From the design of a two-stream encoder and a three-stream decoder, the defect features can be effectively extracted and fused. With the addition of the Swin Transformer, the global information of defect can be captured.

Concerning railway image analysis methods, a timeline of the methods and a classification of them is proposed in Figure 13. Chronologically speaking, they are mainly subdivided into three categories, classical image processing, machine learning and deep-learning. Besides, for machine and deep, learning methods, a general three-step procedure is applied. As illustrated in the Figure, after acquiring the data, the model needs to be trained with relevant known situations to allow a proper detection.

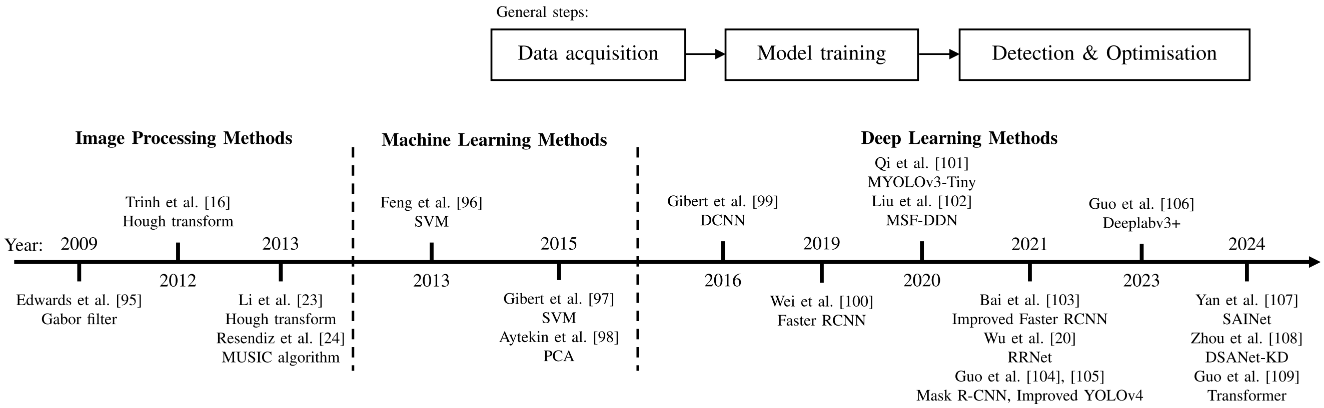


Fig. 13. The evolution of image processing techniques.

C. Decision making

Computer vision methods can detect abnormalities and can be incorporated within warning systems. However, their role in the decision-making process related to actioning appropriate remedial steps needs careful consideration. For example, railway safety is typically governed using Safety Integrity Levels (SIL) which indicate the relative risk reduction provided by a safety function of a device. The SIL must be determined for each application, with a score of 1 being the lowest level of safety protection and a score of 4 being the highest. For example, a camera system for detecting missing track components might require a higher SIL compared to detecting slab track cracking.

Secondly, the choice of real-time or delayed decision-making needs consideration. Delayed processing is typically performed off-site with the assistance of engineers trained to look for signs of specific defects. This can be effective for applications where the decision-making process is not time sensitive and remedial action is not urgent. It can also be applicable for images which are complex, the application is safety critical or there are high costs associated with an adverse decision.

In contrast, real-time decision making is usually performed on-board the camera system hardware and is thus capable of signalling decisions autonomously and instantly. This eliminates the disjointed stages of data handling, from collection, transmission, to centralised processing, data storage management and potential data loss, associated with transferring the data off-site. It also eliminates the expertise required for technicians to interpret the images, and the associated experience/training needed to do so. Further, it eliminates the probability of human error. For example, Li et al. [23] developed a real-time automatic vision-based rail inspection system and using the Hough transform, which inspected ties and tie plates at a frame rate of 20 fps while moving at 16 km/h. Further, Tastimur et al. [113] designed a real-time interface for the vision inspection of rail components. Using preprocessing, morphological feature extraction, fault and deficiency detection, missing rail components and rail surface faults were determined. Extending this, Guo et al. [114] showed the ability to perform real-time instance segmentation with high accuracy using a single GPU, exceeding 30 frames per second (FPS). This was further extended using AI-based track inspection methodologies to reach an inference speed of nearly 100 FPS, using edge-computing devices characterised by their limited power and computational resources [115], [116].

VI. Railway applications

There are, as summarized in Table I, a wide variety of components that comprise the railway system (Figure 14). Different types of cameras and processing algorithms are suitable for each [5], [117]. Therefore, this section discusses the main railway applications for imaging analysis, considering both AI and classical image analysis approaches. First, applications related to the track are discussed, followed by the overhead line equipment and finally several miscellaneous applications are explored.

A. Track applications

Track components detection: One of the most important tasks in a computer vision algorithm is the detection of the objects within an image. For example, Resendiz et al. [24] developed a methodology able to detect, and post-process these track components. The algorithm uses the periodicity of the occurrence of these track components, such as the tie, tie plate, rail, anchor or fastening components, to perform the detection on trolley-borne images. The images are processed firstly with a Gabor filter and are then combined within a one-dimensional signal to be analysed by a multiple signal classification algorithm. Most of the applications subsequently discussed in this section reply on some form of component detection.

Rail surface defects: Railway wheels roll, slip and slide along the rail inducing a range of rail defects, such as corrugation, squats and cracking. Figure 15 illustrates a non-exhaustive list of rail surface defects and their typical appearance on the rail surface.

To detect these, Mandriota et al. [81] developed a technique able to detect the rail corrugation using a line-scanning camera mounted on a maintenance trolley. If surface defects are

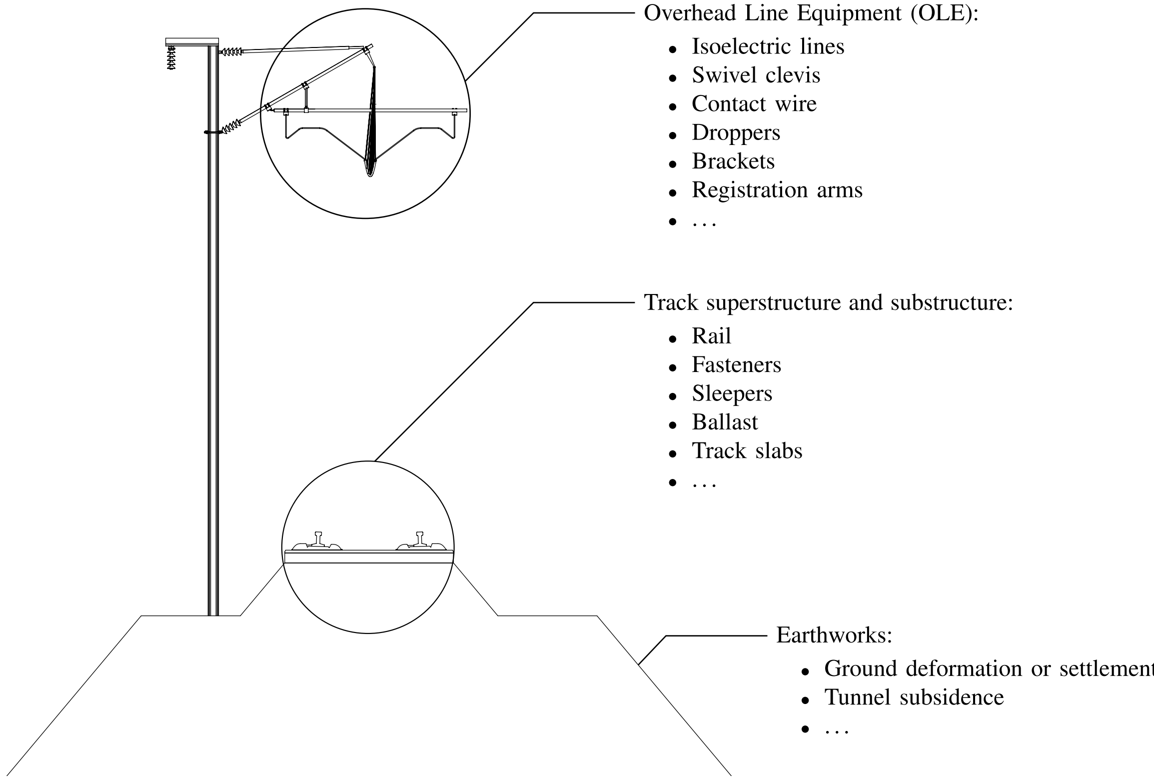


Fig. 14. Railway components commonly used for image inspection

TABLE I
COMMON USES FOR COMPUTER VISION IN RAILWAYS

Track applications	**Overhead-line applications**	**Other applications**
Track component detection	Pantograph arcing	Earthworks shrink-swelling
Rail surface defect detection	Pantograph/catenary contact point	Tunnel subsidence
Crossing nose monitoring	Pantograph stagger	Tunnel inspection
Wheel profile monitoring	Catenary components detection	On-track safety hazard monitoring
Wheel-rail contact position detection	Catenary failure	Level-crossing monitoring
Clip/fastener absence/condition detection		Vegetation management
Track slab defect detection		Station monitoring
Ballast particle characterisation		
Track deflection monitoring		
Wet-bed detection		
Ballast depth and profile		

located on the top of the rail head, a two-dimensional image can be used to detect them since the camera can focus on the top without monitoring the sides of the rail head. However, due to the rail profile geometry and conical wheel profiles, wheel/rail contact is usually not a straight line on the rail head surface. Therefore, to detect three-dimensional wear (defects can be located on the top and on the corners of the rail head simultaneously), Alippi et al. [118] proposed a laser-scanning technique in which a CCD camera takes images of the track illuminated by a laser beam. While the beam is perpendicular to the track and illuminates the rail section, the camera has a different angle that gives visibility of the intersection of the rail and the laser beam. Then the profile is reconstructed using a neural network and the images are post-processed with classical techniques to isolate the region of interest.

Alternatively, Jie et al. [75] developed a methodology to detect defects while examining the grey-scale histogram mean from images taken using a laser-scanning camera located below the vehicle, aimed at the head of the rail. To do so, the rail was divided into pixels rows perpendicular to the rail. The difference between these was used to determine wear. A more recent classical technique was developed by Zhang et al. [120] and consists of an automatic vision-based inspection system able to detect surface rail defects. The system is composed of a line-scan camera mounted on a trolley. This provides long images of the rail whose segments of interest are extracted and analysed to locate the defects. The segmentation process is primarily based on classical image analysis approaches such as grey-scale histogram monitoring, grey equalisation and filtering.

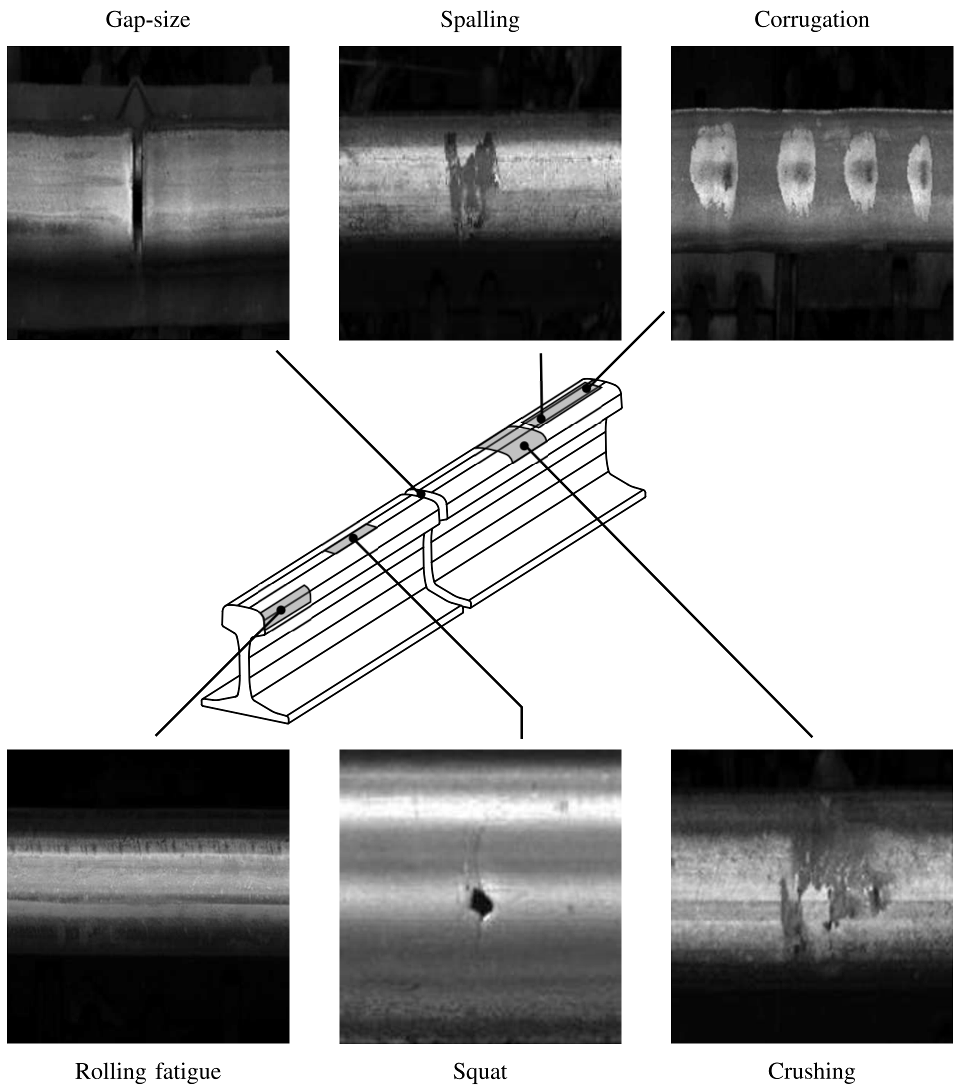


Fig. 15. Example surface rail defects.

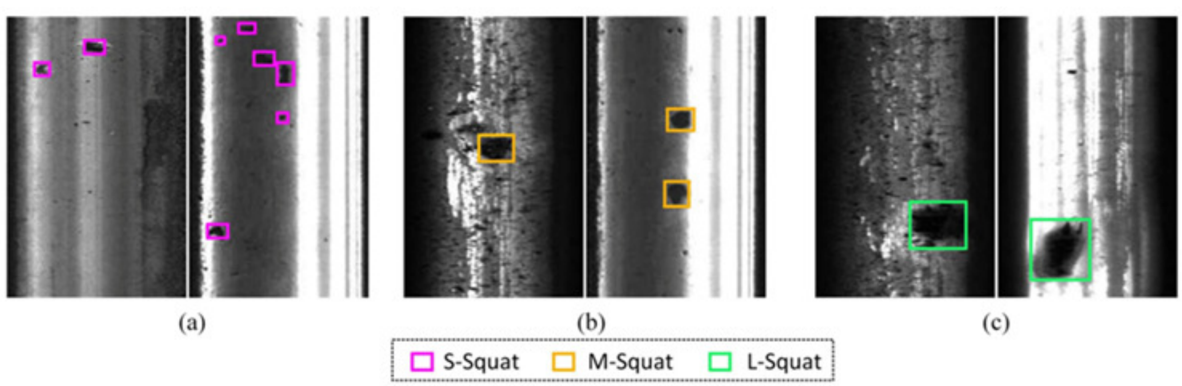


Fig. 16. Rail squat images. (a) small squats; (b) medium squats; (c) larger squats. Reproduced from [119].

A challenge with classical image processing techniques is that they tend to have a high computational cost, with some even requiring trained human input. Thus, to improve the detection efficiency, Qi et al. [101] developed the MYOLOv3-Tiny model which combined YOLOv3-tiny detection head. The backbone which can extract a feature map from the input image was designed with a linear bottleneck and inverted residuals which reversed the convolutional blocks to reduce the parameters. In this way, the computational complexity was reduced, indicating real-time inference was possible. Alternatively, Feng et al. [121] proposed detection networks using MobileNetv2 and MobileNetv3 to ensemble with YOLOv3. The MobileNetv2 bottleneck was divided into two branches while the MobileNetv3 bottleneck was embedded in the squeeze-and-excite module. Based on the test results, rail surface defects were detected and localized in real time. Further, Zhang et al. [119] established a multi-modal rail surface defect detection model named MRSDI-CNN, which included the improved single shot multibox detector and YOLOv3. Rail squats were classified into small, medium, and large based on their radii (Figure 16). Using the proposed a multi-model approach, the rail surface defect detection efficiency was largely improved. Similarly, a pyramid feature-based lightweight CNN (PFCNN) containing a pyramid feature extraction module (PFEM) was also designed for rail surface defect detection in real-time [122].

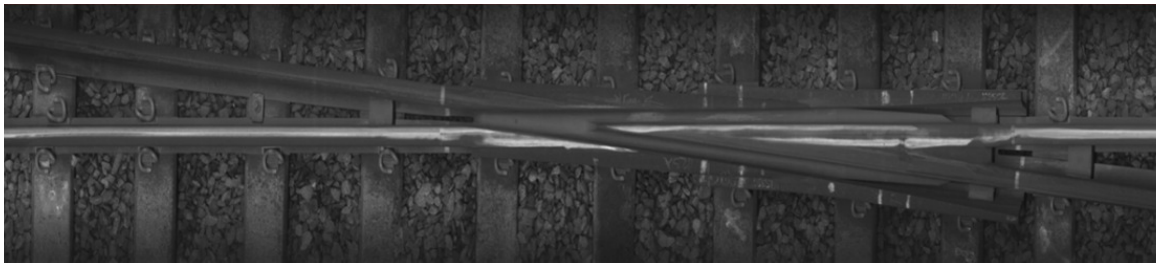
To further improve the detection accuracy of rail surface defects, multilevel or multitask learning approaches were developed. For example, Yang et al. [123] established a multilevel rail surface defect detection network. Taking advantage of differential box-counting and the GrabCut [124] algorithm, Gaussian model parameter identification and defect segmentation were executed. YOLOv2 was adopted for locating and detection of rail surface defects. However, based on the experimental results, rail surface defects were not classified, and the rail position was not accurately localised. Therefore, Meng et al. [125] developed a multitask learning architecture (MtlrNet) to detect surface defects. The characteristics of rail cracks were first analyzed, and the crack features were fused using the attention fusion module. The training loss functions used four parts which were rail defect segmentation loss, rail object detection loss, sharing loss, and omission loss. To reduce the computational cost and avoid the interruption with the other features in the image (such as rail surface background), partial residual blocks were shared and kept. Jin et al. [126] developed a deep multi-model rail inspection system (DM-RIS) to detect rail surface defects with the constrained Gaussian mixture model and Faster RCNN. In the first step, the improved Gaussian mixture model (GMM) [127] was used for defect edge segmentation. In the second step, the Faster RCNN [128] was used to learn non-defect features such as weak illumination, external noise, rust, etc. In this manner DM-RIS performed rail surface defect detection under varied environmental conditions.

Using an alternative approach, Wu et al. [129] considered the impact of image depth (referring to the distance to the camera of its corresponding RGB pixel) with the aim of improving rail surface defect detection accuracy. To do so a depth repeated-enhancement RGB network (DRER-Net) [129] was developed, where the depth and RGB information were considered using an encoder-decoder structure. With the encoder design, cross-modality fusion was conducted using the novel cross modality enhancement fusion module. In the decoder design, a multimodality complementation module was used to refine the prediction results. Further, combining the advantages of rail defect detection on a block level (i.e. processing groups of pixels organized in square blocks) and pixel level, Guo et al. [104] fine-tuned an instance segmentation mode (Mask R-CNN) for inspecting rail surface defects using a customised dataset.

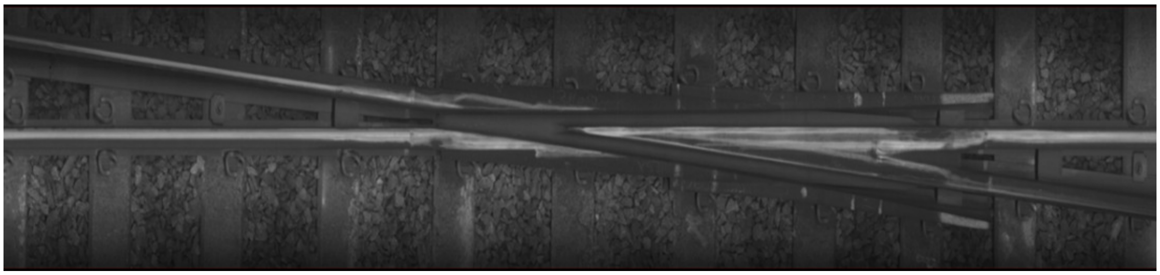
Crossing nose damage: Railway crossings (frogs) enable trains to change lines. They are characterised by a discontinuity of rail and thus the wheel is unsupported as it moves from nose to wing rail. This results in high impact forces and wear along the crossing. Thus, crossings typically have a reduced lifespan compared to plain line rail. To detect changes/defects in the running band under-carriage cameras can be used. For example, Figure 17 (upper 2 sub-figures) shows changes to a crossing nose over a multi-month period. When performing image comparison, it's important that the images being compared have been recorded by a camera at a fixed position from the rail. This can be challenging for vehicle-borne cameras attempting to precisely monitor the track because the conical wheel profiles of different trains lead to slightly different rolling paths along the rail, particularly at crossings where wheel-rail interaction is complex. However, if the positions are similar enough then change detection methods can be used. Figure 17 also shows an example of this, where a heatmap indicates the absolute square difference between the images. It's seen that the changes are located primarily on the running band and in the zone closest to the nose, which is where the wheel impact loads occur. The nose shape has changed, becoming flattened which is likely to reduce running quality.

Wheel profiles: Due to wheel-rail contact, the wheel can also develop defects that negatively affect structural integrity or reduce ride quality (e.g. wheel-flats). To study wheel defects, Zhang et al. [130] used a CCD camera as well as two laser beams fixed to the track to monitor wheel state when the vehicle passed through the measurement device. Specific attention was paid to the flange height and thickness. After calibration of the system, submillimeter precision of detection was achieved. Further, Sun et al. [131] used a camera coupled with multiple parallel laser beams to reconstruct 3D profiles of wheels. A technique was developed to work even if the lasers and the camera were not perfectly normal to the monitored wheel. Also, Emoto et al. [90] developed an algorithm combining images taken by a camera sensor with the 3D reconstruction of the wheel profile, performed with lasers, to perform wheel surface defect detection using an AI-based YOLOv5 algorithm enhanced by classical post-processing computer vision techniques.

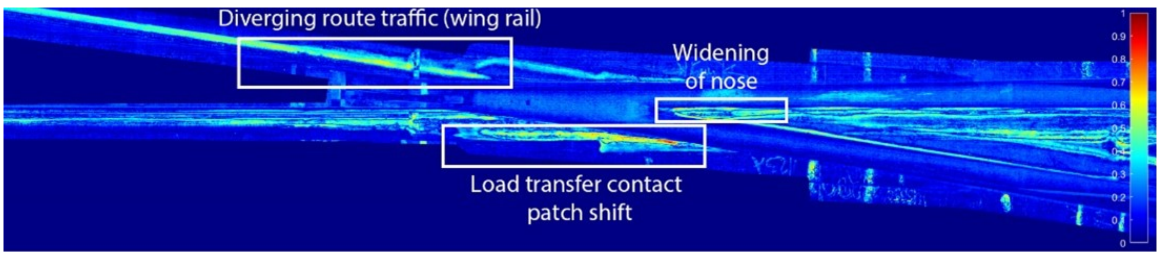
Wheel-rail contact position: Trains are self-steering due to the use of rigid axles and a wheel conicity compatible with the rail profile. This results in the wheel constantly shuffling along the rail surface, particularly in curves. The wheel flange is a safety feature and should not regularly experience contact with the rail. Monitoring the motion of the wheel on the rail and the contact point(s) is of value for monitoring ride quality, detecting hunting and measuring track geometry. To measure this, Shi et al. [132] used stereo cameras below in-service vehicles trained on the wheel-rail contact location (Figure 18). The images were processed using: 1) automatic calibration to detect the region of interest, 2) virtual point detection



(a)



(b)



(c)

Fig. 17. Changes in a crossing over time using train-borne cameras. (a) Initial crossing. (b) Crossing after six months. (c) Change detection analysis.

using a deep convolutional neural network, and 3) a domain-knowledge based rule engine to track the contact point over subsequent frames.

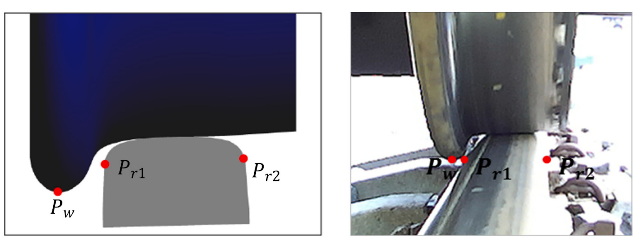


Fig. 18. Wheel-rail contact monitoring showing 3 points of interest. Left: idealised system, Right: in-situ monitoring image [132].

Clip/fastener absence: Railway clips connect the rail and sleeper (and railpad), and due to vibration of the track during train passage, can loosen and fall away from the track. This results in reduced track resistance to horizontal and vertical forces. Under-carriage cameras can be used to view the track super-structure and identify missing clips.

Due to rail industry safety requirements, track super-structure anomalies were traditionally checked by qualified engineers who determine an appropriate course of action. However, considering typical sleeper spacing is 0.6-0.7m, along an entire route this equates to many clips that require checking. Therefore, it is more practical for computer vision to either shortlist track locations for detailed review or assist with maintenance decisions. For example, early methods involved classical image analysis techniques to detect fasteners. Aytekin et al. [98] proposed a real-time fastener detection and inspection algorithm for railway tracks using a combined sheet of light technique with a high-speed, 3D and laser-range finder camera. This involved projecting a laser line on the rail while taking pictures of the laser line with an adapted camera. Pixel similarity and histogram similarity techniques were used to detect fasteners.

More recently, AI techniques have been proposed to improve speed and flexibility of detection. For example, Wei et al. [100] trained VGG16 for the detection and recognition of broken/missing fasteners, however the inference speed was slow. Alternatively, Liu et al. [102] developed a hierarchical learning approach to detect fasteners. In the first step, a multi-scale feature-based deep detection network (MSF-DDN) was constructed. In the second step, a region classification network was built to detect the type of key sub-regions including fasteners. In the last step, fastener detection was achieved using a decision tree. Another approach was used by Zhuang et al. [133] who proposed a deep learning powered two stage method for automatic inspection of railroad components. In the first stage, YOLOv3 was developed to generate the initial detection results. In the second stage, a domain-logic based hybrid model (DLHM) was proposed to improve detection performance.

Kim et al. [134] developed a deep anomaly detection system for rail fasteners. In the first step, a U-Net model, which was composed of segmentation and denoising routines, was trained. In the second step, a deep-learning based method was adopted using the feature map obtained from the U-Net encoder. Using the hierarchical design of U-Net, the features of different rail fasteners can be extracted. However, since the inference is at the pixel level, it is difficult to achieve real-time detection in the field. Wang et al. [135] developed an attention-powered deep convolutional network (AttConv-net) for the detection of the rail, clips, and bolts. Three components including a deep convolutional neural network (DCNN), cascading attention blocks (CAB) and two feed forward networks (FFN) were used to ensemble AttConv-net [135]. The effectiveness of the proposed network was validated using one real dataset and another synthesised dataset. It was found that the training process was time-consuming due to the heavy computational cost of the attention-based module.

To develop these types of deep learning-based approaches requires a large database of image data. These images then require manual labelling of each fastener, which is time consuming and introduces the possibility of human error. To overcome this pre-processing issue Liu et al. [136] developed a key component detection (KCD) method which took advantage of the key components of normal fasteners to determine the existence of abnormalities as depicted in Figure 19. Further, Liu et al. [137] proposed a vision-based fastener inspection system which applied 'few-shot' learning to acquire and label fasteners. Liu et al. [138] discussed the limitations of this method which needs high quality training procedures and can have issues with robustness. Thus, a sample generation method was proposed by producing failure samples with UNet to achieve the training data augmentation. However, since the failure modes of the generated samples were different from the real failures, detection performance was challenging.

Although these approaches are useful for reducing the need for a qualified engineer input to detect anomalies, they are challenging to use in real-time with on-board in-service vehicle cameras. This is because they use multiple steps in model design and implementation. For example, the fastener requires locating and then a decision tree or logic-based model is applied to quantify fastener state. This can require considerable computational requirements, meaning real-time defection challenging, possibly requiring off-site processing. Thus, to increase inference speed, Tu et al. [139] proposed a real-time detection of track components framework. The proposed cascade defect detection network was lightweight and therefore the defect detection inference speed allowed real-time detection. Alternatively, Guo et al. [104], [105] developed real-time rail track component detection models based on real-time object detection and instance semantic segmentation techniques. The impacts of different activation and loss functions were discussed and compared. Using a parallel network structure, the proposed instance segmentation network was able to locate rail track components and display their shapes in real-time. More recently, Tang et al. [140] developed an innovative Cascade Region-based convolutional neural network with Predefined Proposal Templates (CR-PPT) to detect missing components in railway infrastructure such as rail fasteners. One of the main advantages of this newly proposed method consists in its adaptability to novel fastening systems that may not be included in the training dataset.

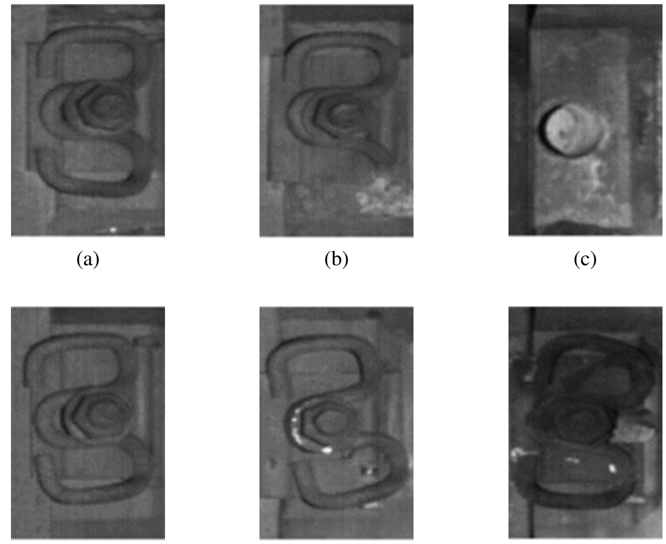

(a) (b) (c) (d) (e) (f)

Fig. 19. Abnormal states of the railway fasteners [103]. (a) Non-defective. (b) Fractured. (c) Missing. (d), (e) and (f) Rotated.

Track slab defects: Concrete slab tracks are commonly used for high-speed lines and for lower speed lines where ballasted track is undesirable (e.g. in tunnels where high track fixity is vital for maintaining clearances). A potential challenge with slab tracks compared to ballast is that they can be susceptible to defects at the slab joints and cracking (e.g. due to subgrade settlement, earthwork shrink-swelling and freeze-thaw cycles). An example of cracking is shown in Figure 20.

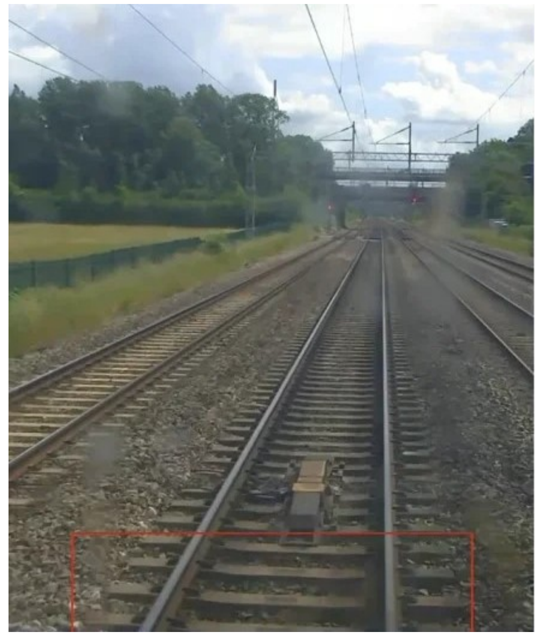
To detect slab track cracking, Ye et al. [142] used crack images from both hand-held cameras and UAVs to develop a deep learning-based network based upon dilated convolution. Compared to ResNet50 [143] and VGG16 [144], the approach achieved higher accuracy and required lower computational cost. Then, expanding upon this work, Ye et al. [145] aggregated multi-scale information to extract crack feature morphology. Alternatively, Wang et al. [146] compared the classification accuracy of different networks for slab track cracking severity. The status of the slab track was classified according to crack severity, depending upon whether the crack was larger or smaller than 0.2mm. Considering recent advances in camera technology, vehicles-borne cameras can also now detect cracking.

Leveraging the dilated convolution, STCNet I [142] was developed to detect high speed railway concrete slab cracks. The watershed algorithm was then proposed to locate the detected crack and execute crack segmentation operation. A total of 1,496 images containing concrete slab cracks with

(a)

(b)

Fig. 20. Slab track surface cracking [141]. (a) Birds eye view between rail seats. (b) Side view along concrete slab.



(a)

(b)

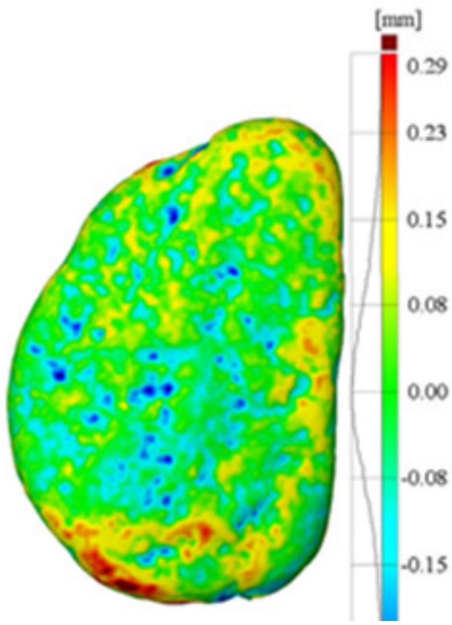
Fig. 21. Abnormal ballast conditions. (a) Inspection of ballast shoulder heights. (b) Dust abraded from unsupported concrete sleepers. Both Figures are reproduced from [58]

high resolutions were acquired to perform model training, validation, and testing.

Ballast: Railway ballast comprises a volume of individual stones which support the sleepers. The lateral resistance provided by the ballast is a key factor in preventing rail buckling, for example on warms days on continuously welded track. Therefore it is common to visually inspect that ballast shoulder heights are sufficient and that the ballast surrounding each sleeper is at a height at least level with the sleeper upper surface. These checks can be performed using computer vision from vehicle-borne cameras, as shown in Figure 21a.

Further, when first laid, the body of railway ballast must meet particle size distribution requirements, and the particles should also meet requirements related to angularity and texture. However, over time the stress cycles induced during train loading can cause ballast degradation. For example, abrasion between individual stones can cause a loss of angularity or breakage, leading to an increased number of smaller diameter particles. Similarly, cohesive particles and contaminates can infiltrate the ballast due to mud-pumping, thus effecting drainage. The ballast then no longer meets requirements, resulting in degraded performance, for example in terms of differential settlement, particularly in wet conditions. This differential settlement can result in the generation of voids between the sleeper bottom and ballast top. These unsupported sleepers hang from the rail when the track is unloaded and when excited during train passage, they cause high frequency impact on the ballast surface. This can result in accelerated abrasion of both the ballast and concrete sleepers. Figure 21b shows an example where dust abraded from unsupported concrete sleepers is present on the track surface and has been detected using computer vision.

The study of ballast particle behaviour is not just confined to in-situ computer vision. For example, to simulate the discrete behaviour of ballast particles, a common analysis technique is the Discrete Element Modelling method. This involves the



(a)

(b)

Fig. 22. Example of ballast monitoring. (a) Photogrammetry of an individual ballast particle [147]. (b) Example ballast layer image segmentation using Mask R-CNN [148].

numerical simulation of the behaviour of a body of individual ballast particles. To model realistic geometries, textures and size distributions, it is common to construct ballast particle libraries. To do so, individual ballast particles are measured under laboratory conditions and digitally reconstructed in 3D. To measure the particles, laser scanning is common, however this can be expensive and does not provide surface colour details. An alternative is to use photogrammetry, which can be more cost effective and records the surface colours. For example, Moaveni et al. [56] used a single DLSR camera and segmentation techniques to capture 3D ballast particle geometries. Further, as depicted in Figure 22, Paixao et al. [147] digitally reconstructed ballast particles using both photogrammetry and lasers and found photogrammetry to produce models of equivalent or higher quality.

Rather than imaging individual railway ballast particles for the purpose of numerical modelling, the ballast layer can also be imaged directly for the purpose of determining ballast condition. This can be done for example using a mobile phone as proposed by Zhang et al. [57], where in-situ mobile images are uploaded to a cloud server for processing. Alternatively, using under-train cameras, foreign objects on the ballast can be detected. For example, Mazzeo et al. [149] used a multilayer perceptron network with an edge histogram to detect object such as broken bottles and aerosol cans on the track.

To assess ballast condition, Kumara et al. [150] proposed the use of image analysis to quantify the level of sand fouling in ballast. It was shown possible to calculate particle gradation curves from images in a lab setting, however for in-situ ballast the technique is challenging because most of the destructive particle movement occurs below the surface, with minimal evidence except near sleepers [151]. This, coupled with the fact fouled particles that have small diameters compared to ballast and thus likely to settle towards the bottom of the ballast layer, means surface camera's struggle to capture images of fouling unless the problem is severe. One notable exception is the presence of sleeper voiding which due to ballast-sleeper abrasion can evidence itself as a white-ish powder on the ballast surface.

Traditionally, to gain understanding of ballast condition with depth required techniques such as visual inspection, sieve analysis or automated ballast sampling. Ground penetrating radar has more recently been introduced; however, accuracy can be affected by fluctuations in moisture content. Alternative image techniques have also been proposed, including the work of Clark et al. [152] and Tan et al. [153] who used thermal and infrared imaging to investigate ballast fouling. It was found clean and fouled ballast had different thermal properties thus evidencing the potential of the approach. Invasive camera methods have also been proposed, for example by inserting endoscope cameras into the ballast and monitoring condition. Alternatively, if environmental dust can be minimised, imaging can be performed during periods when the ballast is exposed, for example shoulder cleaning, skimming and undercutting. Alternatively, trial holes can be dug. The advantage of these approaches is that a potentially wide area and thus representative sample of the ballast layer can be imaged. This approach was used by Luo et al. [154] who recorded ballast layer images in the presence of a spherical target to determine scale. It was then used in a vision transformer-based segmentation framework to approximate ballast fouling index. The technique was further extended to create a dedicated ballast scanning vehicle [17].

Track deflection: The elastic deflection of the rail during train passage must be kept within strict tolerances for safety purposes. If too high, the track stiffness is not sufficient for the axle loads. It is challenging to measure using vehicle-borne sensors due to a lack of datum, and thus can be performed using way-side sensing. Recording sleeper bending and shape during track loading, with the aim of detecting sleeper failures has been attempted by Sabato et al. [155]. To do so, a 3D DIC and pattern projection solution was proposed, and laboratory tests showed the potential to detect bending. However, for research purposes when attempting to study track stiffness, rail deflection is commonly a primary variable of interest. Although rail-mounted accelerometer and geophone solutions have frequently been attempted, the mathematical

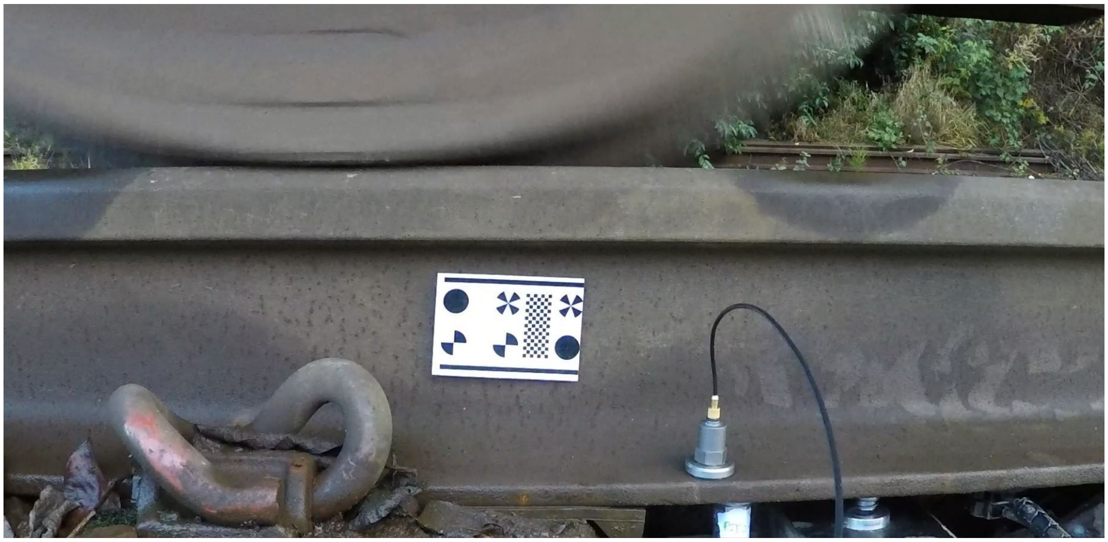


Fig. 23. Monitoring of rail deflection using targets during train passage.

integration can introduce errors, thus making it difficult to obtain accurate readings particularly when the train speed is low. This can be solved using camera-based monitoring with DIC, which usually involves placing, as depicted in Figure 23, a camera several meters from the track, trained on a reflector/target glued to the rail web [60], [61]. The fixed frame of reference then makes it more straightforward to choose a datum. The camera needs to be carefully chosen to have sufficient frame rate to capture the high-speed track deflection, an adequate pixel count and suitable lens characteristics. The pixel count and lens specification can be relaxed closer to those of a consumer grade camera under favourable conditions, for example if the site conditions allow the camera to be placed close to the rail.

The camera's vertical height of placement should be perpendicular with the rail to maximise accuracy. However, any support platform/tripod must also be robust and insensitive to the ground and air motion induced during train passage. Therefore, it's usually challenging to place it any closer than the toe of the ballast. This means some sites (e.g. deep ballast constructions) can require the construction of a temporary platform to elevate the camera. Some approaches have also used lasers encased in tubes to minimise lighting errors [156]. The complex setups and high-power requirements for this type of way-side visual monitoring campaigns mean they are typically performed over short time periods, usually for research purposes, in the presence of trained operators and the captured data is processed off-site.

B. Overhead line applications

Pantograph-catenary systems (PCS) provide power to the rolling-stock on electrified lines (Figure 24). If the system is defective, then train operation is put at risk. For example, loose components can cause short circuits which cut train power, while abnormal pantograph movement or a sagging catenary can pull-down the catenary wires. Considering these challenges, it has become common to put multiple camera technologies on train roofs to continuously monitor both the pantograph and overhead system. Further, way-side camera systems are used to monitor the passes of rolling stock over an individual location, thus allowing for the analysis of many train pantographs on a network.

Pantograph: The pantograph draws power from the catenary into the train. One aspect that has been proposed for monitoring is arcing due to a loss of contact between the pantograph and the catenary. This can occur for example when both electrical elements are separated by a thin layer of air. Then, due to the high voltage in the wires, a short polarisation of the air arises, and the air briefly becomes an electrical conductor. This creates an arc that can be spotted by monitoring the contact zone. To monitor it, Aydin et al. [157] proposed an algorithm to control the contact force applied by the pantograph on the contact line to avoid contact loss or arc-bursting that may damage the line by scraping matter from the wire. To do so, they monitored the height of the contact point by detecting the pantograph on images taken by a camera located at the top of the vehicle. Taking advantage of the pantograph being composed of several straight elements, they used the Hough transform, Canny edge detection and other filters to detect straight lines in the images. Having located the pantograph main elements, they were able to retrieve the height of the top element of the pantograph.

Another aspect is the monitoring of pantograph stagger. The pantograph contacts the catenary via a replaceable contact strip, however, to prevent rapid wear at a single location on the strip, the system uses lateral stagger to shift the contact location during train movement. This contact strip may be monitored as well to prevent any damage and plan its preventive replacement. Cho et al. [158] proposed a methodology able to assess the reliability of railway overhead power lines through the measurement of the dynamic stagger of contact wires using computer vision techniques. The considered images are taken from a camera placed on the top of the train and they use the Hough transform to detect the contact strip. Moreover, they

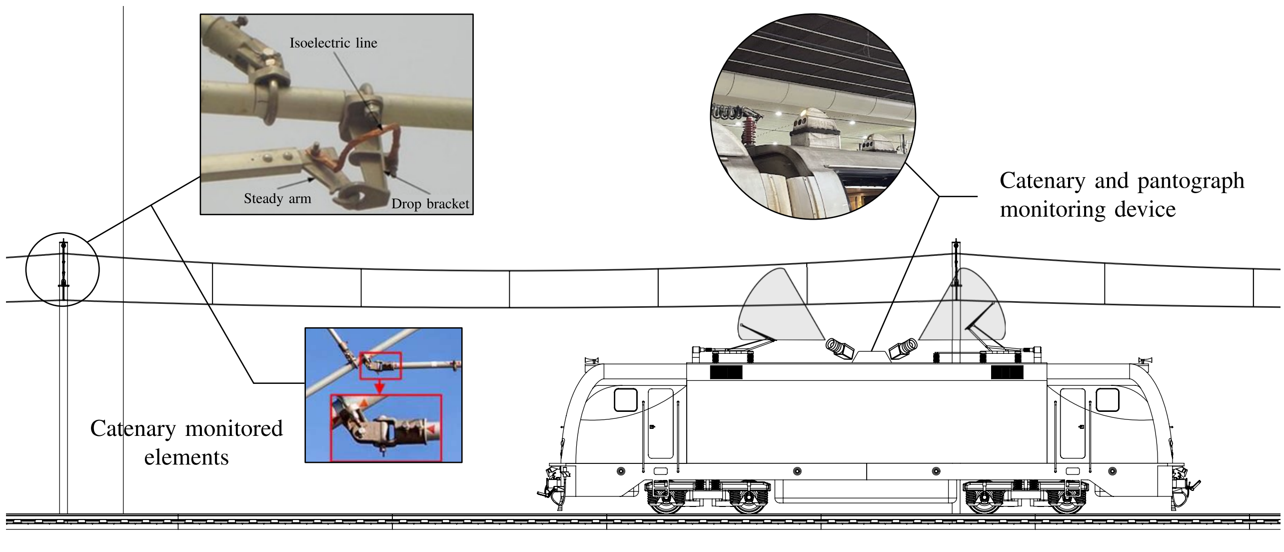


Fig. 24. High-speed railway system: Monitoring of catenary and vehicle pantograph. The part depicting the isoelectric line is reproduced from [160] and the one depicting the swivel clevis from [161].

considered the tilting motion of the train, which leads to a tilting motion of the pantograph, to enhance their detection algorithm.

Overhead line: As an alternative to cameras located on the rolling-stock roof, UAVs can also be used. For example, to monitor the static height and stagger of a high-speed railway contact wire, a LiDAR-based detection framework was developed by Geng et al. [159] using UAV collected point clouds. A self-adaptive method was used to extract the features associated with the contact wires, masts and other suspensions. The height and stagger were computed based on the specific geometric characteristics.

Object detection models have been developed to provide fast and accurate localisation of catenary components. For example, to detect joint components, Wang et al. [162] developed a method using improved YOLOv3 and deeplabv3+ algorithms to localise joint components and split pins (a fastening and protective part of the catenary). To do so, a deblur module was proposed to ensure the accuracy of semantic segmentation and classification. To address the issue of localisation of the catenary insulator which is responsible for keeping the insulation between the catenary and earth, Zhong et al. [163] developed a novel two stage defect detection network. A regression network and an external postprocess network were developed in the TOL-Framework, while an adversarial reconstruction model was adopted to reduce training data requirements. Alternatively, Chen et al. [164] developed a RetinaNet model by integrating a spatial attention map and channel weight map to detect defects in the catenary carrying ring which has a small size making it difficult to inspect manually. The advantage of RetinaNet was that it can deal with severe image data imbalances.

With respect to the inspection of swivel clevis components, Gao et al. [161] proposed an adaptive deep learning-based network for defect detection. A semantic segmentation network and local operators were incorporated, while an unreliability index was defined for monitoring the stability of the proposed network. Regarding the issue of catenary clevis fracture, Han et al. [165] developed a deep learning-based visual ensemble method to detect it. The edge map of the clevis was produced based on a region-scalable fitting model. Wavelet entropy and morphological filtering were employed to complete the fracture detection. Even though there were still several false positives, the detection performance under different conditions was stable. Additionally, Liu et al. [166] developed a novel method to detect the loosening of catenary bracing wire. Firstly, the raw images were enhanced using a deep CNN. Secondly, dynamic anchor learning was developed to localise the bracing line through an angle-based CNN. Finally, the looseness of the catenary bracing was detected using the peak distribution of Hough Transformation.

Treating catenary failure as an anomaly case, Lyu et al. [167] developed a method combing deep CNN and generative adversarial networks (GANs) to predict defects. DCNN was responsible for localisation and GANs accounted for image mapping between the image space and the high-dimensional feature space. An anomaly criterion was used to score the image. To detect the loose strands of the isoelectric line (see Figure 24), Liu et al. [160] proposed an automatic fault detection system including three stages. In the first stage, the isoelectric lines were detected with the Faster RCNN. In the second stage, the Markov random field model was adopted to conduct segmentation of the isoelectric line. In the final stage, the fault state was determined via comparison with a line without defects. Finally, other parts of the overhead line may also be inspected such as the dropper that guarantee a stable contact between the pantograph and the catenary wire [168].

C. Other applications

Earthworks: Railway earthworks can fail due to a variety of causes. For example, scour and washout in the presence of standing and moving water, collapse due to changes in soil moisture deficit (earthworks with clay content). UAVs can be advantageous for monitoring earthworks because they can assess site conditions, for example the present of standing water, the blockage of drainage channels and evidence of animal burrowing. They can also be used to detect slope movement indicators such as curved tree-trunks, quantify vegetation coverage and detect changes in images over time from more flexible camera angles compared to satellite imagery. However, UAVs are typically suited for short surveys where pilot access is available close to the site. For some applications this is challenging, for example in hilly or mountainous terrain. This is particularly true when attempting to survey sites for damage immediately after extreme weather events because access may prove particularly difficult. In these cases, helicopters can be used to collect imagery, including thermal images.

Byrraju et al. [46] explored the use of Sentinel-1 satellite and aerial imagery to survey railroad rights-of-way for conditions that may act as precursors to landslides and other geotechnical hazards. The research incorporated two Interferometric Synthetic Aperture Radar (InSAR) methodologies: Differential Interferometric Synthetic Aperture Radar (DInSAR) and Persistent Scatterer Interferometric Synthetic Aperture Radar (PS-InSAR). Case studies highlighted the capability of current satellite technologies to detect both large and minor ground shifts and variations in soil moisture with sufficient resolution. Alternatively, Chang et al. [49] used Radarsat-2 data to perform nationwide railway earthwork monitoring in the Netherlands. A persistent scatter approach was used considering a 50m wide buffer along the track and the deformations plotted to create a risk map. More recently, Azadnejad et al. [55] used multi-temporal InSAR analysis of Sentinel-1 data combined with Sentinel-2 images. In this study, the small baseline approach was used rather than persistent scatter to monitor, as seen in Figure 25, long-term settlement of the peat earthworks supporting the track. Alternatively, Kim et al. [169] used higher resolution TerraSAR-X data to study track settlement. The persistent scatter method was again used and to improve reflection intensity and coherence corner reflectors were installed. The method was validated using on-site surveys and shown to be useful for trend monitoring.

Vegetation management: On newer high speed railway lines vegetation is typically well controlled. However, on older intercity lines it can be common for heavy vegetation such as tress to be located near the track. This is typically undesirable because trees can pose risks to train operation. These include grass/weeds growing through the trackbed, leaves falling on the track or even an entire tree collapsing on the line. Therefore, computer vision can be used to monitor vegetation growth and schedule vegetation maintenance. For example, the Normalized Difference Vegetation Index (NVDI) can be derived from multispectral satellite images to study the growth of vegetation near the track over time [42]. Alternatively, from a UAV perspective, drones can be used to identify the species

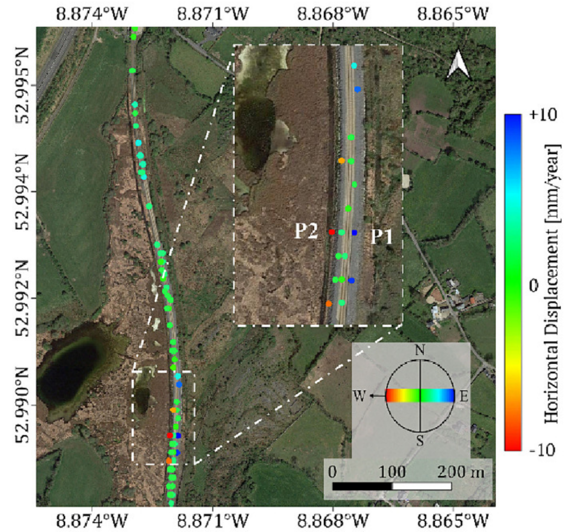


Fig. 25. Horizontal displacement rate map of the an embankment. Reproduced from [55].

of different wayside trees [170]. Alternatively, as illustrated in Figure 26, Ilic et al. [171] developed a dedicated herbicide spraying vehicle with front facing vehicle cameras to capture the presence of vegetation. In real-time, when certain vegetation limits were exceeded, herbicide was deployed depending upon the location it was detected.

Tunnel subsidence: During and after the construction of a tunnel, the hydrological regime around the excavated area may be perturbed [172]. Depending on the local geology, this can lead to a tunnel subsidence which creates ground settlement. Therefore Roccheggiani et al. [172] used Satellite SAR interferometry to detect this motion of the heavily urbanised ground surface above the excavated tunnel of Genoa in Italy. Further, Wang et al. [173], used PSInSAR analysis to detect and identify new subway tunnel lines in Shanghai. To do so, the settlement of the ground located above the tunnels was monitored over a period of time to detect changes. Ge et al. [174] used the same PSInSAR technique to monitor subway and high-speed railway tunnels in Beijing. They established that the ground settlement depended upon tunnel construction technique. Another study provided by Perissin et al. [175] investigated ground subsidence created by recently excavated tunnels in Shanghai using PSInSAR techniques.

Alternatively, Farahani et al. [89] proposed a DIC-based technique to determine the deformation of the surface of a shallow railway tunnel whose construction consisted of drilling and blasting rocks. To develop their methodology, they used a previously 3D-scanned version of the actual tunnel with a terrestrial laser scanner to reconstruct a 3D-printed and scaled version of the tunnel that can be studied in a laboratory. This first scan provided a point-cloud on which a surface was reconstructed to design the 3D model to be printed. On this scaled model a laser scanning technique was used based on a camera and a circular laser to scan and reconstruct the entire surface of the tunnel. Two other cameras were used to analyse the deformation field of the tunnel using DIC analysis

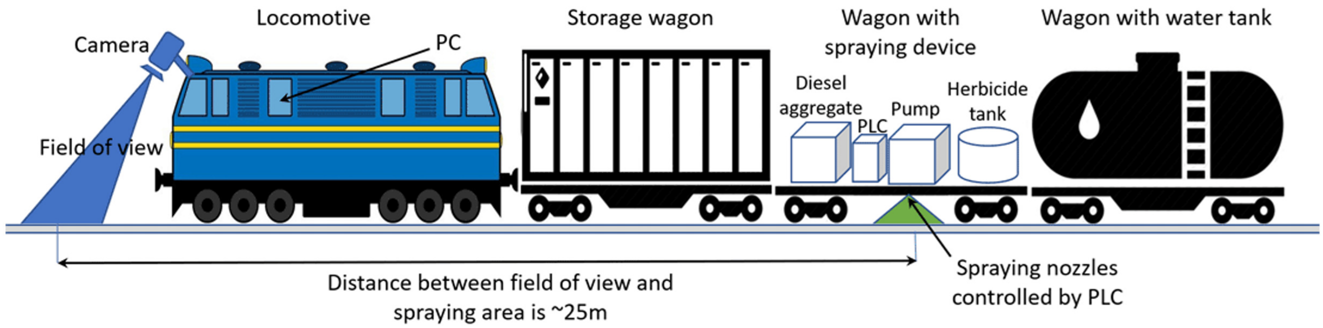


Fig. 26. Herbicide spraying vehicle commanded using computer vision techniques. Reproduced from [171].

on the acquired images of the deformed scaled tunnel when subjected to a known constraint. The final objective was to locate, with the camera/laser measurement system, possible geometrical change of one part of a real tunnel.

Tunnel inspection: Zhou et al. [176] proposed a laser scanning based inspection method for tunnel clearance inspection. As for other laser techniques, a mobile laser scanning system placed on the top of a trolley has been used to monitor the surface of tunnels. These tunnels are then reconstructed using point cloud segments that are realigned using the alignment of the rails. Therefore, a dynamic clearance coordinate system is computed considering clearance restrictions.

Otherwise, for tunnels being used past their initially intended lifetime, it is important to perform inspections to prevent failure that could either reduce the safety of passengers or delay traffic. Jenkins et al. [177] developed both hardware and software to help perform the monitoring of railway tunnels to detect structural degradation. Their system consisted of an array of camera and lighting placed on a trolley. The aim was to monitor and detect defaults of the tunnel lining such as missing bricks, open joints, surface corrosion, cracks as well as water ingress. The software reconstructed a 3D point cloud and a mesh of the tunnel including the texture by stitching the images taken by the array of cameras. A similar technique was also proposed by Gavilan et al. [178] but with arrays of laser cameras mounted on a track-modified truck that scanned the tunnel at a relatively high speed to perform long tunnel inspection on a regular basis.

Sometimes, UAVs are also used to perform tunnel inspection such as in the PLUTO Project [179] in which a UAV carried out the inspection of confined spaces such as tunnels for which communication with the outside was not possible.

On-track safety hazards: Potential safety hazards (PSHs) on the tracks, such as falling trees and other objects that may foul the track, present risks to both railroad safety and operational reliability, especially for high-speed railroads. They require prompt attention to prevent the escalation of these issues into more severe incidents or accidents. For example, in China during 2021, foreign objects such as plastic tarps obstructed HSR tracks on more than ten occasions, leading to numerous service disruptions, cancellations, traffic delays and economic losses.

As a result, regular inspections of track surroundings are important. Traditional methods of PSH inspection, which typically involve visual checks conducted by inspectors on foot, are subjective and largely reliant on the personal expertise and discretion of the inspectors. To address these challenges, Wu et al. [180] developed YOLARC (You Only Look at Railroad Coefficients), an automated PSH detection framework. YOLARC utilises Unmanned Aerial Vehicle (UAV) imagery combined with AI-powered image processing to monitor HSR tracks. Moreover, the system is equipped with a hazard level evaluation (HLE) methodology that assesses the proximity of detected objects to the tracks, thereby quantifying the level of risk. Testing on UAV-acquired high-speed rail datasets has demonstrated YOLARC can process UAV imagery into actionable information, achieving high detection rates and fast processing speeds.

Level-crossings: At-grade crossings (aka level crossings) are junctions where railways and highways intersect. They are a common source of accidents due to collisions. To access traffic information at the crossing, Guo et al. [72] proposed an improved YOLOv3 model to detect multiple traffic instances at the grade crossing both during the day and night. The proposed feature fusion module was validated to improve local feature representation.

Trespass: Railway track trespass, either at a level-crossing or elsewhere is one of the greatest sources of injury in the industry. To improve safety, Zhang et al. [181] developed an AI-aided analytic platform to automatically detect trespassing based on surveillance video footage of a level-crossing monitored by a fixed camera. Experiments demonstrated trespassing can be detected during both the day and night. However, addressing pedestrian behaviour at railway crossings presents significant challenges due to the nuanced differences between normal and potentially hazardous actions. This was studied by Jiang et al. [182] who proposed a deep learning framework capable of detecting unusual pedestrian behaviours through video analysis and skeleton tracking. Further progress was made by Song et al. [183], who developed a GAN-based framework for analysing pedestrian behaviour without the need for location-specific adjustments, thereby enhancing its applicability across various settings. However, dangers at railway crossings extend beyond intentional or unintentional pedestrian actions. Thus, responding to the need for broader detection capabilities, Tang et al. [184] introduced the RC-SAFE Network, a system designed to identify any foreign object at crossings, extending beyond pedestrian monitoring alone.

Alternatively, Zhang et al. [181] presented a computer vision algorithm to detect trespass near misses using surveillance video footage of railway-road grade crossings. This algorithm was designed to be robust under changing lighting conditions throughout the day-night cycle and performed well under varying weather conditions. The same author also introduced an improved AI-powered framework [12] to automatically detect trespassing events on railroads. The deep learning tool identified trespassing incidents, classified violators, created video clips, and consolidated event data into a centralized dataset. Further, Qin et al. [185] used a semantic segmentation model (DeepLab) and an object detection model (YOLOv5) to dynamically identify regions of interest, trespassers, and obstacles, trained on over 10,000 images.

VII. Future Research Directions

Data fusion: Combining imaging with additional datasets, for example using a data-fusion approach, has the potential to achieve more powerful insights. One example relates to combining imaging and laser scanning to create 3D scans with texture, which then has the potential to provide even more detailed datasets for more advanced AI operations. For example, Cui et al. [186] studied point cloud acquisition to detect rail fasteners along a ballasted track. They used a laser mounted on a trolley to scan the track and deep learning techniques to detect the fasteners that can, in some cases, be covered by the ballast. Farahani et al. [89] developed a coupled system using 3D laser scanning and digital image correlation to monitor the deformation of railway tunnels. Hackel et al. [187] developed a technique to perform rail and crossing detection along railway tracks using vehicle-borne point clouds and associated images.

Hyperspectral/Multispectral imaging: While the human eye sees colour in three bands (red, green and blue), hyperspectral imaging collects information across bands in the electromagnetic spectrum, many of which are beyond the visible ranges. Rather than 2D images, this results in 3D data 'cubes', with each location having its own spectral signature. These signatures can be used to identify the materials that the imaged object is comprised from. Although it is starting to gain interest in fields such as highways [188], so far hyperspectral imaging has received limited attention in the railway field [189]. This is perhaps due to the high cost of hyperspectral camera systems and also the large datasets generated which require processing.

Affordable and user-friendly monitoring solutions: The goal is to craft compact, yet accurate, computer vision technologies that fit into budget-friendly edge-computing environments. Even though the market offers a variety of inspection systems, uptake can be curtailed by steep prices and the need for specialist training. A direction of new research may aim to produce systems that are not only economical but also easy to use, enhancing their appeal and functionality for those managing railways. For example, Tang et al. [184] introduced a streamlined light-weight approach for inspecting tracks, and Tang and Qian [116] demonstrated the integration of this approach into a portable, power-efficient edge-computing device. Additionally, wearable technology and augmented reality (e.g. Apple Vision Pro and Microsoft HoloLens) continues to evolve and leveraging these devices for inspection purposes may become a key area for researchers in the future, particularly for training and education.

Autonomous solutions: Track inspections can be demanding as they require personnel to manage several responsibilities and analyse various data simultaneously. By creating technologies capable of independently performing specific inspection activities, it may be possible to reduce the burden on human inspectors. For example, using drones for the real-time detection of rails and track centreline tracking to navigate without relying on GPS, lays the groundwork for future self-operating inspection systems [190]. The fusion of lightweight inspection models, edge-computing technology, and autonomous platforms presents an exciting direction for upcoming research efforts.

Emerging technologies: Potential future technologies that may benefit railroad computer vision are edge AI, generative models, and quantum computing [191], [192]. Edge AI enables on-site, real-time detection of trespassers and obstacles by processing data directly on devices placed on trains or trackside. Generative models, such as GANs or diffusion models, can contribute by generating realistic synthetic training data to cover a broad range of environmental scenarios (e.g. various weather and lighting conditions), which can improve model robustness. Meanwhile, quantum computing, though still emerging, promises a step-change in speeding up data-intensive tasks, such as processing large volumes of video data to detect trespassing patterns in real time or optimizing model parameters efficiently. The convergence of these technologies could ultimately produce a resilient, intelligent railroad industry, where edge AI systems, continually enhanced by generative models and quantum-accelerated data processing, operate as a self-optimizing system for monitoring and responding to safety risks with increased precision and adaptability.

Remaining challenges: Although the application of CV for railways has advanced quickly in recent years, challenges remain, particularly related to the railway environment [193]. For example, lighting and weather changes, from day-night cycles to rain, snow, and fog, all significantly affect image quality, requiring robust models to handle noise while maintaining accuracy. Also, camera positioning further complicates detection: fixed cameras may lack comprehensive angles, while mobile cameras on trains face rapid perspective changes and field-of-view limitations. For fixed wayside cameras there is also the risk of theft which can limit the long-term deployment of high-end camera technology. Further, the brief appearance of trespassers or obstacles, combined with occlusions from environmental features, demands high sensitivity to prevent missed detections. Additionally, railroad operations produce track vibrations and signal lights, which create image instability and visual noise that can confuse algorithms. Real-time processing requirements, constrained by the limited computational power of edge devices, add pressure for optimized models that balance speed and accuracy. These algorithms must also generalize across varied rail settings, yet data scarcity hampers training, as high-quality, labelled datasets of rail-specific scenarios remain limited. Moreover, regulatory standards mandate high system reliability, which requires CV and AI models to minimize false positives that could disrupt operations.

VIII. Conclusions

Computer vision in the railway field has the potential to provide new and enhanced methods for asset inspection compared to traditional manual techniques. It offers the potential of improved quality control, safety, time and cost, and therefore it is a highly active research area. This paper has performed a state-of-the-art review of the field, attempting to capture the current trends and future opportunities. To do so, first camera hardware requirements were studied. This focused on the unique challenges associated with operating optical equipment in a railway environment, such as contamination, power sources and lighting. Data requirements were also discussed, delving into the challenges associated with storage, transmission and security. In this section, rules of thumb as well as practical orders of magnitudes were specified. Next, the different types of camera mounting device were explored, investigating the differences between handheld, trolley, vehicle-borne, drone and satellite technologies. Image processing was then explored by dividing the main approaches into 'classical' and 'AI' based. Classical approaches such as edge detection, Hough transforms, and Gabor filters were presented before extending to machine learning algorithms such as YOLO. Considering the hardware requirements and image processing algorithms, the most common applications for computer vision in the rail industry were studied. This included detection of rail surface defects, missing fasteners, concrete slab track cracks, ballast particle condition, earthworks movement, level crossing blockage and overhead line equipment defects. Finally, opportunities for future research direction were discussed.

IX. Acknowledgements

The authors wish to acknowledge the Belgian funding association (Fonds de la Recherche Scientifique – FNRS) as well as the University of Mons (Fonds Franeau) for funding the academic visit of Bryan OLIVIER at the University of Leeds that lead to the conception of this paper. This work was also supported by the Natural Science Foundation of China (Grant Number: 52308457), Natural Science Foundation of Shandong Province (Grant Number: ZR2023QE220), China Postdoctoral Science Foundation (Grant Number: 2024M761811).

References

[1] J. Blake, "Fogs. Uncommon Sense. Schenectady Gazette.," 1932. [Online; accessed 08-November-2024].

[2] N. Rail, "Inspection and Maintenance of Permanent Way: Manual," 2020.

[3] K. Oh, M. Yoo, N. Jin, J. Ko, J. Seo, H. Joo, and M. Ko, "A Review of Deep Learning Applications for Railway Safety," *Applied Sciences*, vol. 12, p. 10572, Oct. 2022. doi: 10.3390/app122010572.

[4] P. Aela, H.-L. Chi, A. Fares, T. Zayed, and M. Kim, "UAV-based studies in railway infrastructure monitoring," *Automation in Construction*, vol. 167, p. 105714, Nov. 2024. doi: 10.1016/j.autcon.2024.105714.

[5] P. Aela, J. Cai, G. Jing, and H.-L. Chi, "Vision-based monitoring of railway superstructure: A review," *Construction and Building Materials*, vol. 442, p. 137385, Sept. 2024. doi: 10.1016/j.conbuildmat.2024.137385.

[6] S. Chen, G. T. Frøseth, S. Derosa, A. Lau, and A. Rönnquist, "Railway Catenary Condition Monitoring: A Systematic Mapping of Recent Research," *Sensors*, vol. 24, p. 1023, Feb. 2024. doi: 10.3390/s24031023.

[7] M. J. Pappaterra, M. L. Pappaterra, and F. Flammini, "A study on the application of convolutional neural networks for the maintenance of railway tracks," *Discover Artificial Intelligence*, vol. 4, p. 30, May 2024. doi: 10.1007/s44163-024-00127-2.

[8] Z. Cao, Y. Qin, L. Jia, Z. Xie, Y. Gao, Y. Wang, P. Li, and Z. Yu, "Railway Intrusion Detection Based on Machine Vision: A Survey, Challenges, and Perspectives," *IEEE Transactions on Intelligent Transportation Systems*, vol. 25, pp. 6427–6448, July 2024. doi: 10.1109/TITS.2024.3412170.

[9] L. Chapman, J. E. Thornes, and S. P. White, "Thermal imaging of railways to identify track sections prone to buckling," *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, vol. 220, pp. 317–327, May 2006. doi: 10.1243/09544097JRRT73.

[10] M.-S. Kim, S.-C. Oh, G.-Y. Kim, and S.-J. Kwon, "Underbody component monitoring system of railway vehicles using the infrared thermal images," in *2014 International SoC Design Conference (ISOCC)*, (Jeju), pp. 222–223, IEEE, Nov. 2014. doi: 10.1109/ISOCC.2014.7087616.

[11] J. Kim, "Non-Destructive Characterization of Railway Materials and Components with Infrared Thermography Technique," *Materials*, vol. 12, p. 4077, Dec. 2019. doi: 10.3390/ma12244077.

[12] Z. Zhang, P. Chen, Y. Huang, L. Dai, F. Xu, and H. Hu, "Railway obstacle intrusion warning mechanism integrating YOLO-based detection and risk assessment," *Journal of Industrial Information Integration*, vol. 38, p. 100571, Mar. 2024. doi: 10.1016/j.jii.2024.100571.

[13] J. Cámara-Molina, E. Moliner, M. Martínez-Rodrigo, D. Connolly, D. Yurchenko, P. Galvín, and A. Romero, "3D printed energy harvesters for railway bridges-Design optimisation," *Mechanical Systems and Signal Processing*, vol. 190, p. 110133, May 2023. doi: 10.1016/j.ymssp.2023.110133.

[14] J. Cámara-Molina, A. Romero, E. Moliner, D. Connolly, M. Martínez-Rodrigo, D. Yurchenko, and P. Galvín, "Design, tuning and in-field validation of energy harvesters for railway bridges," *Mechanical Systems and Signal Processing*, vol. 208, p. 111012, Feb. 2024. doi: 10.1016/j.ymssp.2023.111012.

[15] Railvue, "Locomotive Camera System." https://www.railvue.com/, 2024. [Online; accessed 20-March-2024].

[16] H. Trinh, N. Haas, Y. Li, C. Otto, and S. Pankanti, "Enhanced rail component detection and consolidation for rail track inspection," in *2012 IEEE Workshop on the Applications of Computer Vision (WACV)*, pp. 289–295, IEEE, 2012.

[17] J. Luo, K. Ding, H. Huang, J. M. Hart, I. I. A. Qamhia, E. Tutumluer, H. Thompson, and T. R. Sussmann, "Toward Automated Field Ballast Condition Evaluation: Development of a Ballast Scanning Vehicle," *Transportation Research Record: Journal of the Transportation Research Board*, p. 03611981231178302, June 2023. doi: 10.1177/03611981231178302.

[18] H. Wang, J. Berkers, N. Van Den Hurk, and N. F. Layegh, "Study of loaded versus unloaded measurements in railway track inspection," *Measurement*, vol. 169, p. 108556, Feb. 2021. doi: 10.1016/j.measurement.2020.108556.

[19] X. Jiang and S. Wang, "Railway Panorama: A Fast Inspection Method for High-Speed Railway Infrastructure Monitoring," *IEEE Access*, vol. 9, pp. 150889–150902, 2021. doi: 10.1109/ACCESS.2021.3125645.

[20] Y. Wu, Y. Qin, Y. Qian, and F. Guo, "Automatic detection of arbitrarily oriented fastener defect in high-speed railway," *Automation in Construction*, vol. 131, p. 103913, Nov. 2021. doi: 10.1016/j.autcon.2021.103913.

[21] A. Eriksen, J. Gascoyne, and W. Al-Nuaimy, "Improved productivity and reliability of ballast inspection using road-rail multi-channel gpr," in *Railway Engineering*, pp. 6–7th, Commonwealth Institute London, 2004.

[22] X. Gibert-Serra, A. Berry, C. Diaz, W. Jordan, B. Nejikovsky, and A. Tajaddini, "A Machine Vision System for Automated Joint Bar Inspection From a Moving Rail Vehicle," in *ASME/IEEE 2007 Joint Rail*

Conference and Internal Combustion Engine Division Spring Technical Conference, (Pueblo, Colorado, USA), pp. 289–296, ASMEDC, Jan. 2007. doi: 10.1115/JRC/ICE2007-40039.

[23] Ying Li, Hoang Trinh, N. Haas, C. Otto, and S. Pankanti, “Rail Component Detection, Optimization, and Assessment for Automatic Rail Track Inspection,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, pp. 760–770, Apr. 2014. doi: 10.1109/TITS.2013.2287155.

[24] E. Resendiz, J. Hart, and N. Ahuja, “Automated Visual Inspection of Railroad Tracks,” *Intelligent Transportation Systems, IEEE Transactions on*, vol. 14, pp. 751–760, May 2013. doi: 10.1109/TITS.2012.2236555.

[25] Z. Liu, W. Liu, and Z. Han, “A High-Precision Detection Approach for Catenary Geometry Parameters of Electrical Railway,” *IEEE Transactions on Instrumentation and Measurement*, vol. 66, pp. 1798–1808, July 2017. doi: 10.1109/TIM.2017.2666358.

[26] M. Karakose, O. Yamanand, K. Murat, and E. Akin, “A New Approach for Condition Monitoring and Detection of Rail Components and Rail Track in Railway*:,” *International Journal of Computational Intelligence Systems*, vol. 11, no. 1, p. 830, 2018. doi: 10.2991/ijcis.11.1.63.

[27] Y. Min, B. Xiao, J. Dang, B. Yue, and T. Cheng, “Real time detection system for rail surface defects based on machine vision,” *EURASIP Journal on Image and Video Processing*, vol. 2018, p. 3, Dec. 2018. doi: 10.1186/s13640-017-0241-y.

[28] G. Jing, X. Qin, H. Wang, and C. Deng, “Developments, challenges, and perspectives of railway inspection robots,” *Automation in Construction*, vol. 138, p. 104242, June 2022. doi: 10.1016/j.autcon.2022.104242.

[29] M. Circelli, N. Kaviani, R. Licciardello, S. Ricci, L. Rizzetto, S. S. Arabani, and D. Shi, “Track geometry monitoring by an on-board computer-vision-based sensor system,” *Transportation Research Procedia*, vol. 69, pp. 257–264, 2023. doi: 10.1016/j.trpro.2023.02.170.

[30] SnakeGrid, “Low-distortion coordinate systems for large engineering projects.” https://snakegrid.org/, 2024. [Online; accessed 21-March-2024].

[31] Network Rail, “How we use drones.” https://www.networkrail.co.uk/running-the-railway/looking-after-the -railway/our-fleet-machines-and-vehicles/air-operations/drones-or-unmanned-aircraft-systems-uas/, 2024. [Online; accessed 20-March-2024].

[32] L. Tong, L. Jia, Y. Geng, K. Liu, Y. Qin, and Z. Wang, “Anchor-adaptive railway track detection from unmanned aerial vehicle images,” *Computer-Aided Civil and Infrastructure Engineering*, vol. 38, pp. 2666–2684, Dec. 2023. doi: 10.1111/mice.13004.

[33] Y. Wu, Y. Qin, Z. Wang, X. Ma, and Z. Cao, “Densely pyramidal residual network for UAV-based railway images dehazing,” *Neurocomputing*, vol. 371, pp. 124–136, Jan. 2020. doi: 10.1016/j.neucom.2019.06.076.

[34] L. Tong, Z. Wang, L. Jia, Y. Qin, Y. Wei, H. Yang, and Y. Geng, “Fully Decoupled Residual ConvNet for Real-Time Railway Scene Parsing of UAV Aerial Images,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, pp. 14806–14819, Sept. 2022. doi: 10.1109/TITS.2021.3134318.

[35] H. Huang, G. Zhao, Y. Bo, J. Yu, L. Liang, Y. Yang, and K. Ou, “Railway intrusion detection based on refined spatial and temporal features for UAV surveillance scene,” *Measurement*, vol. 211, p. 112602, Apr. 2023. doi: 10.1016/j.measurement.2023.112602.

[36] N. Yague-Martinez, P. Prats-Iraola, F. Rodriguez Gonzalez, R. Brcic, R. Shau, D. Geudtner, M. Eineder, and R. Bamler, “Interferometric Processing of Sentinel-1 TOPS Data,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, pp. 2220–2234, Apr. 2016. doi: 10.1109/TGRS.2015.2497902.

[37] M. E. D. Chaves, M. C. A. Picoli, and I. D. Sanches, “Recent Applications of Landsat 8/OLI and Sentinel-2/MSI for Land Use and Land Cover Mapping: A Systematic Review,” *Remote Sensing*, vol. 12, p. 3062, Sept. 2020. doi: 10.3390/rs12183062.

[38] A. Stumpf, J.-P. Malet, P. Allemand, and P. Ulrich, “Surface reconstruction and landslide displacement measurements with Pléiades satellite images,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 95, pp. 1–12, Sept. 2014. doi: 10.1016/j.isprsjprs.2014.05.008.

[39] A. Javed, K. A. Qazi, M. Maqsood, and K. A. Shah, “Efficient Algorithm for Railway Tracks Detection Using Satellite Imagery,” *International Journal of Image, Graphics and Signal Processing*, vol. 4, pp. 34–40, Oct. 2012. doi: 10.5815/ijigsp.2012.11.05.

[40] D. C. Wickramasinghe, T. T. Vu, and T. Maul, “Satellite remote-sensing monitoring of a railway construction project,” *International Journal of Remote Sensing*, vol. 39, pp. 1754–1769, Mar. 2018. doi: 10.1080/01431161.2017.1415481.

[41] K. Nogueira, C. Cesar, P. H. T. Gama, G. L. S. Machado, and J. A. Dos Santos, “A Tool for Bridge Detection in Major Infrastructure Works Using Satellite Images,” in *2019 XV Workshop de Visão Computacional (WVC)*, (São Bernardo do Campo, Brazil), pp. 72–77, IEEE, Sept. 2019. doi: 10.1109/WVC.2019.8876942.

[42] M. Kučera and Z. Dobesova, “Analysis of the Degree of Threat to Railway Infrastructure by Falling Tree Vegetation,” *ISPRS International Journal of Geo-Information*, vol. 10, p. 292, May 2021. doi: 10.3390/ijgi10050292.

[43] G. Dial, H. Bowen, F. Gerlach, J. Grodecki, and R. Oleszczuk, “IKONOS satellite, imagery, and products,” *Remote Sensing of Environment*, vol. 88, pp. 23–36, Nov. 2003. doi: 10.1016/j.rse.2003.08.014.

[44] S. Jérôme, “Shaping the future of earth observation with pléiades neo,” in *2019 9th International Conference on Recent Advances in Space Technologies (RAST)*, pp. 399–401, IEEE, 2019.

[45] P. Bernhard, D. Haener, and O. Frey, “Detection of Railway Track Anomalies Using Interferometric Time Series of TerraSAR-X Satellite Radar Data,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, pp. 1–11, 2024. doi: 10.1109/JSTARS.2024.3405019.

[46] S. V. Byrraju, D. C. Rizos, and Y. Qian, “Satellite Radar Imagery for Detection and Monitoring of Geohazards,” *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2674, pp. 283–292, Mar. 2020. doi: 10.1177/0361198120910746.

[47] V. Gagliardi, F. Tosti, L. Bianchini Ciampoli, M. L. Battagliere, L. D’Amato, A. M. Alani, and A. Benedetto, “Satellite Remote Sensing and Non-Destructive Testing Methods for Transport Infrastructure Monitoring: Advances, Challenges and Perspectives,” *Remote Sensing*, vol. 15, p. 418, Jan. 2023. doi: 10.3390/rs15020418.

[48] L. Chang, R. Dollevoet, and R. Hanssen, “Railway Infrastructure Monitoring using Satellite Radar Data,” *International Journal of Railway Technology*, vol. 3, no. 2, pp. 79–91, 2014. doi: 10.4203/ijrt.3.2.5.

[49] L. Chang, R. P. B. J. Dollevoet, and R. F. Hanssen, “Nationwide Railway Monitoring Using Satellite SAR Interferometry,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 10, pp. 596–604, Feb. 2017. doi: 10.1109/JSTARS.2016.2584783.

[50] L. Chang, R. P. B. J. Dollevoet, and R. F. Hanssen, “Monitoring Line-Infrastructure With Multisensor SAR Interferometry: Products and Performance Assessment Metrics,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 11, pp. 1593–1605, May 2018. doi: 10.1109/JSTARS.2018.2803074.

[51] L. Chang, N. P. Sakpal, S. O. Elberink, and H. Wang, “Railway Infrastructure Classification and Instability Identification Using Sentinel-1 SAR and Laser Scanning Data,” *Sensors*, vol. 20, p. 7108, Dec. 2020. doi: 10.3390/s20247108.

[52] J. Zhang, W. Zhu, Y. Cheng, and Z. Li, “Landslide Detection in the Linzhi–Ya’an Section along the Sichuan–Tibet Railway Based on InSAR and Hot Spot Analysis Methods,” *Remote Sensing*, vol. 13, p. 3566, Sept. 2021. doi: 10.3390/rs13183566.

[53] Y. Wassie, Q. Gao, O. Monserrat, A. Barra, B. Crippa, and M. Crosetto, “DIFFERENTIAL SAR INTERFEROMETRY FOR THE MONITORING OF LAND SUBSIDENCE ALONG RAILWAY INFRASTRUCTURES,” *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XLIII-B3-2022, pp. 361–366, May 2022. doi: 10.5194/isprs-archives-XLIII-B3-2022-361-2022.

[54] S. Selvakumaran, S. Plank, C. Geiß, C. Rossi, and C. Middleton, “Remote monitoring to predict bridge scour failure using Interferometric Synthetic Aperture Radar (InSAR) stacking techniques,” *International Journal of Applied Earth Observation and Geoinformation*, vol. 73, pp. 463–470, Dec. 2018. doi: 10.1016/j.jag.2018.07.004.

[55] S. Azadnejad, A. Hrysiewicz, A. Trafford, F. O’Loughlin, E. Holohan, F. Kelly, and S. Donohue, “InSAR supported by geophysical and geotechnical information constrains two-dimensional motion of a railway embankment constructed on peat,” *Engineering Geology*, vol. 333, p. 107493, May 2024. doi: 10.1016/j.enggeo.2024.107493.

[56] M. Moaveni, S. Wang, J. M. Hart, E. Tutumluer, and N. Ahuja, “Evaluation of Aggregate Size and Shape by Means of Segmentation Techniques and Aggregate Image Processing Algorithms,” *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2335, pp. 50–59, Jan. 2013. doi: 10.3141/2335-06.

[57] B. Zhang, S. J. Lee, Y. Qian, E. Tutumluer, and S. Bhattacharya, “A Smartphone-Based Image Analysis Technique for Ballast Aggregates,” in *International Conference on Transportation and Development 2016*,

(Houston, Texas), pp. 623–630, American Society of Civil Engineers, June 2016. doi: 10.1061/9780784479926.057.
[58] OneBigCircle, "AIVR Product Suite." https://onebigcircle.co.uk/aivr/, 2024. [Online; accessed 21-March-2024].
[59] D. Bowness, A. Lock, D. Richards, and W. Powrie, "Innovative Remote Video Monitoring of Railway Track Displacements," *Applied Mechanics and Materials*, vol. 3-4, pp. 417–422, Aug. 2006. doi: 10.4028/www.scientific.net/AMM.3-4.417.
[60] D. Bowness, A. C. Lock, W. Powrie, J. A. Priest, and D. J. Richards, "Monitoring the dynamic displacements of railway track," *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, vol. 221, pp. 13–22, Jan. 2007. doi: 10.1243/0954409JRRT51.
[61] L. Le Pen, G. Watson, W. Powrie, G. Yeo, P. Weston, and C. Roberts, "The behaviour of railway level crossings: Insights through field monitoring," *Transportation Geotechnics*, vol. 1, pp. 201–213, Dec. 2014. doi: 10.1016/j.trgeo.2014.05.002.
[62] J. Wilson, G. Tian, I. Mukriz, and D. Almond, "PEC thermography for imaging multiple cracks from rolling contact fatigue," *NDT & E International*, vol. 44, pp. 505–512, Oct. 2011. doi: 10.1016/j.ndteint.2011.05.004.
[63] C. Gartrell, J. Reagan, and C. Carter, "Modernizing Rail Inspections for the U.S. Army Corps of Engineers," in *ASME 2012 Rail Transportation Division Fall Technical Conference*, (Omaha, Nebraska, USA), pp. 189–197, American Society of Mechanical Engineers, Oct. 2012. doi: 10.1115/RTDF2012-9404.
[64] D. Mishra, J. Gallier, E. Tutumluer, Y. Haddani, and R. Gourves, "Application of PANDA and Geo-Endoscopy Techniques for Improved Assessment of Track Substructure Conditions at Railroad Bridge Approaches," in *2015 Joint Rail Conference*, (San Jose, California, USA), p. V001T01A027, American Society of Mechanical Engineers, Mar. 2015. doi: 10.1115/JRC2015-5744.
[65] F. Lamas-Lopez, Y. Cui, S. Costa D'Aguiar, and N. Calon, "Geotechnical auscultation of a French conventional railway track-bed for maintenance purposes," *Soils and Foundations*, vol. 56, pp. 240–250, Apr. 2016. doi: 10.1016/j.sandf.2016.02.007.
[66] Y.-J. Cui, "Unsaturated railway track-bed materials," *E3S Web of Conferences*, vol. 9, p. 01001, 2016. doi: 10.1051/e3sconf/20160901001.
[67] Y. Haddani, P. Breul, G. Saussine, M. A. B. Navarrete, F. Ranvier, and R. Gourvès, "Trackbed Mechanical and Physical Characterization using PANDA®/Geoendoscopy Coupling," *Procedia Engineering*, vol. 143, pp. 1201–1209, 2016. doi: 10.1016/j.proeng.2016.06.118.
[68] S. Barbier, Y. Haddani, C. Sastre-Jurado, and M. Benz-Navarrete, "Development of an automatic method to determine gothechnical soil properties through pandoscope," in *Journées Nationales de Géotechnique et de Géologie de l'Ingénieur - Champs-sur-Marne 2018*, 2018.
[69] L. Rui, E. Zappa, and A. Collina, "Vision-based measurement of crack generation and evolution during static testing of concrete sleepers," *Engineering Fracture Mechanics*, vol. 224, p. 106715, Feb. 2020. doi: 10.1016/j.engfracmech.2019.106715.
[70] CamlinGroup, "Train Monitoring Solutions." https://camlingroup.com/, 2024. [Online; accessed 14-Mai-2024].
[71] H. Deilamsalehy, T. C. Havens, P. Lautala, E. Medici, and J. Davis, "An automatic method for detecting sliding railway wheels and hot bearings using thermal imagery," *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, vol. 231, pp. 690–700, July 2017. doi: 10.1177/0954409716638703.
[72] F. Guo, Y. Wang, and Y. Qian, "Computer vision-based approach for smart traffic condition assessment at the railroad grade crossing," *Advanced Engineering Informatics*, vol. 51, p. 101456, Jan. 2022. doi: 10.1016/j.aei.2021.101456.
[73] F. Guo, Y. Wang, and Y. Qian, "Real-time dense traffic detection using lightweight backbone and improved path aggregation feature pyramid network," *Journal of Industrial Information Integration*, vol. 31, p. 100427, Feb. 2023. doi: 10.1016/j.jii.2022.100427.
[74] Y. Qian, Y. Wang, D. C. Rizos, *et al.*, "Intelligent camera aided railway emergency system (i-cares)," *Center for Connected Multimodal Mobility, Clemson University*, 2021.
[75] Lin Jie, Luo Siwei, Li Qingyong, Zhang Hanqing, and Ren Shengwei, "Real-time rail head surface defect detection: A geometrical approach," in *2009 IEEE International Symposium on Industrial Electronics*, (Seoul, South Korea), pp. 769–774, IEEE, July 2009. doi: 10.1109/ISIE.2009.5214088.
[76] N. Otsu *et al.*, "A threshold selection method from gray-level histograms," *Automatica*, vol. 11, no. 285-296, pp. 23–27, 1975.
[77] Q. Li and S. Ren, "A Real-Time Visual Inspection System for Discrete Surface Defects of Rail Heads," *IEEE Transactions on Instrumentation and Measurement*, vol. 61, pp. 2189–2199, Aug. 2012. doi: 10.1109/TIM.2012.2184959.
[78] I. Aydin, "A new approach based on firefly algorithm for vision-based railway overhead inspection system," *Measurement*, vol. 74, pp. 43–55, Oct. 2015. doi: 10.1016/j.measurement.2015.07.022.
[79] T. Shi, J.-y. Kong, X.-d. Wang, Z. Liu, and G. Zheng, "Improved Sobel algorithm for defect detection of rail surfaces with enhanced efficiency and accuracy," *Journal of Central South University*, vol. 23, pp. 2867–2875, Nov. 2016. doi: 10.1007/s11771-016-3350-3.
[80] A. Landi, L. Menconi, and L. Sani, "Hough transform and thermovision for monitoring pantograph-catenary system," *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, vol. 220, pp. 435–447, July 2006. doi: 10.1243/0954409JRRT41.
[81] C. Mandriota, E. Stella, M. Nitti, N. Ancona, and A. Distante, "Rail corrugation detection by Gabor filtering," in *Proceedings 2001 International Conference on Image Processing (Cat. No.01CH37205)*, vol. 2, (Thessaloniki, Greece), pp. 626–628, IEEE, 2001. doi: 10.1109/ICIP.2001.958571.
[82] R. Huber-Mörk, M. Nölle, A. Oberhauser, and E. Fischmeister, "Statistical Rail Surface Classification Based on 2D and 21/2D Image Analysis," in *Advanced Concepts for Intelligent Vision Systems* (J. Blanc-Talon, D. Bone, W. Philips, D. Popescu, and P. Scheunders, eds.), vol. 6474, pp. 50–61, Berlin, Heidelberg: Springer Berlin Heidelberg, 2010. doi: 10.1007/978-3-642-17688-3_6.
[83] E. Guclu, I. Aydin, and E. Akin, "Development of Vision-Based Autonomous UAV for Railway Tracking," in *2021 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT)*, (Zallaq, Bahrain), pp. 120–125, IEEE, Sept. 2021. doi: 10.1109/3ICT53449.2021.9581919.
[84] R. Marmo, L. Lombardi, and N. Gagliardi, "Railway sign detection and classification," in *2006 IEEE Intelligent Transportation Systems Conference*, (Toronto, ON, Canada), pp. 1358–1363, IEEE, 2006. doi: 10.1109/ITSC.2006.1707412.
[85] C. Tastimur, M. Karakose, and E. Akın, "Image Processing Based Level Crossing Detection and Foreign Objects Recognition Approach in Railways," *International Journal of Applied Mathematics, Electronics and Computers*, vol. 1, pp. 19–23, Aug. 2017. doi: 10.18100/ijamec.2017SpecialIssue30465.
[86] C. Sacchi and C. Regazzoni, "A distributed surveillance system for detection of abandoned objects in unmanned railway environments," *IEEE Transactions on Vehicular Technology*, vol. 49, pp. 2013–2026, Sept. 2000. doi: 10.1109/25.892603.
[87] S. Zheng, X. Chai, X. An, and L. Li, "Railway track gauge inspection method based on computer vision," in *2012 IEEE International Conference on Mechatronics and Automation*, (Chengdu, China), pp. 1292–1296, IEEE, Aug. 2012. doi: 10.1109/ICMA.2012.6284322.
[88] M. Niu, K. Song, L. Huang, Q. Wang, Y. Yan, and Q. Meng, "Unsupervised Saliency Detection of Rail Surface Defects using Stereoscopic Images," *IEEE Transactions on Industrial Informatics*, pp. 1–1, 2020. doi: 10.1109/TII.2020.3004397.
[89] B. V. Farahani, F. Barros, P. J. Sousa, P. P. Cacciari, P. J. Tavares, M. M. Futai, and P. Moreira, "A coupled 3D laser scanning and digital image correlation system for geometry acquisition and deformation monitoring of a railway tunnel," *Tunnelling and Underground Space Technology*, vol. 91, p. 102995, Sept. 2019. doi: 10.1016/j.tust.2019.102995.
[90] T. Emoto, A. A. Ravankar, A. Ravankar, T. Emaru, and Y. Kobayashi, "Automatic inspection of wheel surface defects using a combination of laser sensors and machine vision," *SICE Journal of Control, Measurement, and System Integration*, vol. 17, pp. 57–66, Dec. 2024. doi: 10.1080/18824889.2024.2314800.
[91] D. He, R. Ren, K. Li, Z. Zou, R. Ma, Y. Qin, and W. Yang, "Urban rail transit obstacle detection based on Improved R-CNN," *Measurement*, vol. 196, p. 111277, June 2022. doi: 10.1016/j.measurement.2022.111277.
[92] S. Xu, Q. Feng, J. Fei, G. Zhao, X. Liu, H. Li, C. Lu, and Q. Yang, "A Locating Approach for Small-Sized Components of Railway Catenary Based on Improved YOLO With Asymmetrically Effective Decoupled Head," *IEEE Access*, vol. 11, pp. 34870–34879, 2023. doi: 10.1109/ACCESS.2023.3264441.
[93] Y. Wang, J. Jiang, Y. Hong, S. Gao, M. Hu, C. Li, M. Li, and D. Liu, "Railway Structure Diagnosis Based on Swin-Transformer Backbone with Mask R-CNN," in *2022 IEEE 5th International Conference on Electronics Technology (ICET)*, (Chengdu, China), pp. 1126–1130, IEEE, May 2022. doi: 10.1109/ICET55676.2022.9824154.
[94] L. Gao, J. Zhang, C. Yang, and Y. Zhou, "Cas-VSwin transformer: A variant swin transformer for surface-defect detection,"

Computers in Industry, vol. 140, p. 103689, Sept. 2022. doi: 10.1016/j.compind.2022.103689.

[95] J. R. Edwards, J. M. Hart, S. Sawadisavi, E. Resendiz, C. Barkan, and N. Ahuja, "Advancements in railroad track inspection using machine-vision technology," in *AREMA Conference Proceedings on American Railway and Maintenance of Way Association*, vol. 290, 2009.

[96] H. Feng, Z. Jiang, F. Xie, P. Yang, J. Shi, and L. Chen, "Automatic Fastener Classification and Defect Detection in Vision-Based Railway Inspection Systems," *IEEE Transactions on Instrumentation and Measurement*, vol. 63, pp. 877–888, Apr. 2014. doi: 10.1109/TIM.2013.2283741.

[97] X. Gibert, V. M. Patel, and R. Chellappa, "Robust Fastener Detection for Autonomous Visual Railway Track Inspection," in *2015 IEEE Winter Conference on Applications of Computer Vision*, (Waikoloa, HI, USA), pp. 694–701, IEEE, Jan. 2015. doi: 10.1109/WACV.2015.98.

[98] C. Aytekin, Y. Rezaeitabar, S. Dogru, and I. Ulusoy, "Railway Fastener Inspection by Real-Time Machine Vision," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 45, pp. 1101–1107, July 2015. doi: 10.1109/TSMC.2014.2388435.

[99] X. Gibert, V. M. Patel, and R. Chellappa, "Deep Multitask Learning for Railway Track Inspection," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, pp. 153–164, Jan. 2017. doi: 10.1109/TITS.2016.2568758.

[100] X. Wei, Z. Yang, Y. Liu, D. Wei, L. Jia, and Y. Li, "Railway track fastener defect detection based on image processing and deep learning techniques: A comparative study," *Engineering Applications of Artificial Intelligence*, vol. 80, pp. 66–81, Apr. 2019. doi: 10.1016/j.engappai.2019.01.008.

[101] H. Qi, T. Xu, G. Wang, Y. Cheng, and C. Chen, "MYOLOv3-Tiny: A new convolutional neural network architecture for real-time detection of track fasteners," *Computers in Industry*, vol. 123, p. 103303, Dec. 2020. doi: 10.1016/j.compind.2020.103303.

[102] J. Liu, Y. Teng, B. Shi, X. Ni, W. Xiao, C. Wang, and H. Liu, "A hierarchical learning approach for railway fastener detection using imbalanced samples," *Measurement*, vol. 186, p. 110240, Dec. 2021. doi: 10.1016/j.measurement.2021.110240.

[103] T. Bai, J. Yang, G. Xu, and D. Yao, "An optimized railway fastener detection method based on modified Faster R-CNN," *Measurement*, vol. 182, p. 109742, Sept. 2021. doi: 10.1016/j.measurement.2021.109742.

[104] F. Guo, Y. Qian, D. Rizos, Z. Suo, and X. Chen, "Automatic Rail Surface Defects Inspection Based on Mask R-CNN," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2675, pp. 655–668, Nov. 2021. doi: 10.1177/03611981211019034.

[105] F. Guo, Y. Qian, and Y. Shi, "Real-time railroad track components inspection based on the improved YOLOv4 framework," *Automation in Construction*, vol. 125, p. 103596, May 2021. doi: 10.1016/j.autcon.2021.103596.

[106] F. Guo, Y. Qian, and H. Yu, "Automatic Rail Surface Defect Inspection Using the Pixelwise Semantic Segmentation Model," *IEEE Sensors Journal*, vol. 23, pp. 15010–15018, July 2023. doi: 10.1109/JSEN.2023.3280117.

[107] Y. Yan, X. Jia, K. Song, W. Cui, Y. Zhao, C. Liu, and J. Guo, "Specificity autocorrelation integration network for surface defect detection of no-service rail," *Optics and Lasers in Engineering*, vol. 172, p. 107862, Jan. 2024. doi: 10.1016/j.optlaseng.2023.107862.

[108] W. Zhou, J. Hong, X. Ran, W. Yan, and Q. Jiang, "DSANet-KD: Dual Semantic Approximation Network via Knowledge Distillation for Rail Surface Defect Detection," *IEEE Transactions on Intelligent Transportation Systems*, vol. 25, pp. 13849–13862, Oct. 2024. doi: 10.1109/TITS.2024.3385744.

[109] F. Guo, J. Liu, Y. Qian, and Q. Xie, "Rail surface defect detection using a transformer-based network," *Journal of Industrial Information Integration*, vol. 38, p. 100584, Mar. 2024. doi: 10.1016/j.jii.2024.100584.

[110] Z. Chen, J. Yang, L. Chen, Z. Feng, and L. Jia, "Efficient railway track region segmentation algorithm based on lightweight neural network and cross-fusion decoder," *Automation in Construction*, vol. 155, p. 105069, Nov. 2023. doi: 10.1016/j.autcon.2023.105069.

[111] Z. Feng, J. Yang, Z. Chen, and Z. Kang, "LRseg: An efficient railway region extraction method based on lightweight encoder and self-correcting decoder," *Expert Systems with Applications*, vol. 238, p. 122386, Mar. 2024. doi: 10.1016/j.eswa.2023.122386.

[112] C. Wan, S. Ma, and K. Song, "TSSTNet: A Two-Stream Swin Transformer Network for Salient Object Detection of No-Service Rail Surface Defects," *Coatings*, vol. 12, p. 1730, Nov. 2022. doi: 10.3390/coatings12111730.

[113] C. Tastimur, O. Yaman, M. Karakose, and E. Akin, "A real time interface for vision inspection of rail components and surface in railways," in *2017 International Artificial Intelligence and Data Processing Symposium (IDAP)*, (Malatya), pp. 1–6, IEEE, Sept. 2017. doi: 10.1109/IDAP.2017.8090267.

[114] F. Guo, Y. Qian, Y. Wu, Z. Leng, and H. Yu, "Automatic railroad track components inspection using real-time instance segmentation," *Computer-Aided Civil and Infrastructure Engineering*, vol. 36, pp. 362–377, Mar. 2020. doi: 10.1111/mice.12625.

[115] J. Guo, S. Zhang, Y. Qian, and Y. Wang, "An adaptively weighted loss-enabled lightweight teacher–student model for real-time railroad inspection on edge devices," *Neural Computing and Applications*, vol. 35, pp. 24455–24472, Dec. 2023. doi: 10.1007/s00521-023-09038-2.

[116] Y. Tang, Y. Wang, and Y. Qian, "Edge-Computing Oriented Real-Time Missing Track Components Detection," *Transportation Research Record: Journal of the Transportation Research Board*, p. 03611981241230546, Mar. 2024. doi: 10.1177/03611981241230546.

[117] S. Liu, Q. Wang, and Y. Luo, "A review of applications of visual inspection technology based on image processing in the railway industry," *Transportation Safety and Environment*, vol. 1, pp. 185–204, Dec. 2019. doi: 10.1093/tse/tdz007.

[118] C. Alippi, E. Casagrande, F. Scotti, and V. Piuri, "Composite real-time image processing for railways track profile measurement," *IEEE Transactions on Instrumentation and Measurement*, vol. 49, pp. 559–564, June 2000. doi: 10.1109/19.850395.

[119] H. Zhang, Y. Song, Y. Chen, H. Zhong, L. Liu, Y. Wang, T. Akilan, and Q. M. J. Wu, "MRSDI-CNN: Multi-Model Rail Surface Defect Inspection System Based on Convolutional Neural Networks," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, pp. 11162–11177, Aug. 2022. doi: 10.1109/TITS.2021.3101053.

[120] H. Zhang, X. Jin, Q. M. J. Wu, Y. Wang, Z. He, and Y. Yang, "Automatic Visual Detection System of Railway Surface Defects With Curvature Filter and Improved Gaussian Mixture Model," *IEEE Transactions on Instrumentation and Measurement*, vol. 67, pp. 1593–1608, July 2018. doi: 10.1109/TIM.2018.2803830.

[121] J. H. Feng, H. Yuan, Y. Q. Hu, J. Lin, S. W. Liu, and X. Luo, "Research on deep learning method for rail surface defect detection," *IET Electrical Systems in Transportation*, vol. 10, pp. 436–442, Dec. 2020. doi: 10.1049/iet-est.2020.0041.

[122] Y. Liu, H. Xiao, J. Xu, and J. Zhao, "A Rail Surface Defect Detection Method Based on Pyramid Feature and Lightweight Convolutional Neural Network," *IEEE Transactions on Instrumentation and Measurement*, vol. 71, pp. 1–10, 2022. doi: 10.1109/TIM.2022.3165287.

[123] H. Yang, Y. Wang, J. Hu, J. He, Z. Yao, and Q. Bi, "Deep Learning and Machine Vision-Based Inspection of Rail Surface Defects," *IEEE Transactions on Instrumentation and Measurement*, vol. 71, pp. 1–14, 2022. doi: 10.1109/TIM.2021.3138498.

[124] M. Tang, L. Gorelick, O. Veksler, and Y. Boykov, "Grabcut in one cut," in *Proceedings of the IEEE international conference on computer vision*, pp. 1769–1776, 2013.

[125] S. Meng, S. Kuang, Z. Ma, and Y. Wu, "MtlrNet: An Effective Deep Multitask Learning Architecture for Rail Crack Detection," *IEEE Transactions on Instrumentation and Measurement*, vol. 71, pp. 1–10, 2022. doi: 10.1109/TIM.2022.3181940.

[126] X. Jin, Y. Wang, H. Zhang, H. Zhong, L. Liu, Q. M. J. Wu, and Y. Yang, "DM-RIS: Deep Multimodel Rail Inspection System With Improved MRF-GMM and CNN," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, pp. 1051–1065, Apr. 2019. doi: 10.1109/TIM.2019.2909940.

[127] F. Pernkopf and D. Bouchaffra, "Genetic-based EM algorithm for learning Gaussian mixture models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, pp. 1344–1348, Aug. 2005. doi: 10.1109/TPAMI.2005.162.

[128] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, pp. 1137–1149, June 2017. doi: 10.1109/TPAMI.2016.2577031.

[129] J. Wu, W. Zhou, W. Qiu, and L. Yu, "Depth Repeated-Enhancement RGB Network for Rail Surface Defect Inspection," *IEEE Signal Processing Letters*, vol. 29, pp. 2053–2057, 2022. doi: 10.1109/LSP.2022.3211199.

[130] Z.-F. Zhang, Z. Gao, Y.-Y. Liu, F.-C. Jiang, Y.-L. Yang, Y.-F. Ren, H.-J. Yang, K. Yang, and X.-D. Zhang, "Computer Vision Based Method and System for Online Measurement of Geometric Parameters of Train Wheel Sets," *Sensors*, vol. 12, pp. 334–346, Dec. 2011. doi: 10.3390/s120100334.

[131] J. Sun, Z. Zhang, and J. Zhang, "Reconstructing Normal Section Profiles of 3-D Revolving Structures via Pose-Unconstrained Multiline Structured-Light Vision," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–12, 2021. doi: 10.1109/TIM.2020.3024339.

[132] D. Shi, E. Šabanovič, L. Rizzetto, V. Skrickij, R. Oliverio, N. Kaviani, Y. Ye, G. Bureika, S. Ricci, and M. Hecht, "Deep learning based virtual point tracking for real-time target-less dynamic displacement measurement in railway applications," *Mechanical Systems and Signal Processing*, vol. 166, p. 108482, Mar. 2022. doi: 10.1016/j.ymssp.2021.108482.

[133] L. Zhuang, H. Qi, T. Wang, and Z. Zhang, "A Deep-Learning-Powered Near-Real-Time Detection of Railway Track Major Components: A Two-Stage Computer-Vision-Based Method," *IEEE Internet of Things Journal*, vol. 9, pp. 18806–18816, Oct. 2022. doi: 10.1109/JIOT.2022.3162295.

[134] B. Kim, Y. Jeon, J.-W. Kang, and J. Gwak, "Multi-task Transfer Learning Facilitated by Segmentation and Denoising for Anomaly Detection of Rail Fasteners," *Journal of Electrical Engineering & Technology*, vol. 18, pp. 2383–2394, May 2023. doi: 10.1007/s42835-022-01347-1.

[135] T. Wang, Z. Zhang, F. Yang, and K.-L. Tsui, "Automatic Rail Component Detection Based on AttnConv-Net," *IEEE Sensors Journal*, vol. 22, pp. 2379–2388, Feb. 2022. doi: 10.1109/JSEN.2021.3132460.

[136] J. Liu, H. Liu, C. Chakraborty, K. Yu, X. Shao, and Z. Ma, "Cascade Learning Embedded Vision Inspection of Rail Fastener by Using a Fault Detection IoT Vehicle," *IEEE Internet of Things Journal*, vol. 10, pp. 3006–3017, Feb. 2023. doi: 10.1109/JIOT.2021.3126875.

[137] J. Liu, Y. Huang, Q. Zou, M. Tian, S. Wang, X. Zhao, P. Dai, and S. Ren, "Learning Visual Similarity for Inspecting Defective Railway Fasteners," *IEEE Sensors Journal*, vol. 19, pp. 6844–6857, Aug. 2019. doi: 10.1109/JSEN.2019.2911015.

[138] J. Liu, Y. Teng, X. Ni, and H. Liu, "A Fastener Inspection Method Based on Defective Sample Generation and Deep Convolutional Neural Network," *IEEE Sensors Journal*, vol. 21, pp. 12179–12188, May 2021. doi: 10.1109/JSEN.2021.3062021.

[139] Z. Tu, S. Wu, G. Kang, and J. Lin, "Real-Time Defect Detection of Track Components: Considering Class Imbalance and Subtle Difference Between Classes," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–12, 2021. doi: 10.1109/TIM.2021.3117357.

[140] Y. Tang, Y. Wang, and Y. Qian, "Railroad missing components detection via cascade region-based convolutional neural network with predefined proposal templates," *Computer-Aided Civil and Infrastructure Engineering*, vol. 39, pp. 3083–3102, Oct. 2024. doi: 10.1111/mice.13279.

[141] J.-W. Lee, S.-J. Lee, and S.-H. Kee, "Evaluation of a Concrete Slab Track with Debonding at the Interface between Track Concrete Layer and Hydraulically Stabilized Base Course Using Multi-Channel Impact-Echo Testing," *Sensors*, vol. 21, p. 7091, Oct. 2021. doi: 10.3390/s21217091.

[142] W. Ye, S. Deng, J. Ren, X. Xu, K. Zhang, and W. Du, "Deep learning-based fast detection of apparent concrete crack in slab tracks with dilated convolution," *Construction and Building Materials*, vol. 329, p. 127157, Apr. 2022. doi: 10.1016/j.conbuildmat.2022.127157.

[143] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778, 2016.

[144] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[145] W. Ye, J. Ren, A. A. Zhang, and C. Lu, "Automatic pixel-level crack detection with multi-scale feature fusion for slab tracks," *Computer-Aided Civil and Infrastructure Engineering*, vol. 38, pp. 2648–2665, Dec. 2023. doi: 10.1111/mice.12984.

[146] W. Wang, W. Hu, W. Wang, X. Xu, M. Wang, Y. Shi, S. Qiu, and E. Tutumluer, "Automated crack severity level detection and classification for ballastless track slab using deep convolutional neural network," *Automation in Construction*, vol. 124, p. 103484, Apr. 2021. doi: 10.1016/j.autcon.2020.103484.

[147] A. Paixão, R. Resende, and E. Fortunato, "Photogrammetry for digital reconstruction of railway ballast particles – A cost-efficient method," *Construction and Building Materials*, vol. 191, pp. 963–976, Dec. 2018. doi: 10.1016/j.conbuildmat.2018.10.048.

[148] J. Luo, H. Huang, I. I. A. Qamhia, J. M. Hart, and E. Tutumluer, "Deep Learning-Based Segmentation for Field Evaluation of Riprap and Large-Sized Aggregates," in *Geo-Congress 2023*, (Los Angeles, California), pp. 424–434, American Society of Civil Engineers, Mar. 2023. doi: 10.1061/9780784484678.043.

[149] P. L. Mazzeo, E. Stella, M. Nitti, and A. Distante, "Potential dangerous object detection on railway ballast using digital image processing," in *Computers in Railways X*, vol. 1, (Prague, Czech Republic), pp. 157–166, WIT Press, June 2006. doi: 10.2495/CR060161.

[150] J. J. Kumara, K. Ogiwara, S. Hoshi, and K. Hayano, "Evaluation on grain size distribution of railway ballast with and without sands by image analyses," in *3rd International conference on new developments in soil mechanics and geotechnical engineering, North Cyprus*, 2012.

[151] X. Bian, W. Cai, Z. Luo, C. Zhao, and Y. Chen, "Image-Aided Analysis of Ballast Particle Movement Along a High-Speed Railway," *Engineering*, p. S2095809922006300, Sept. 2022. doi: 10.1016/j.eng.2022.08.006.

[152] M. Clark, D. McCann, and M. Forde, "Infrared thermographic investigation of railway track ballast," *NDT & E International*, vol. 35, pp. 83–94, Mar. 2002. doi: 10.1016/S0963-8695(01)00032-9.

[153] Y. Tan, Y. Chen, A. W. Peterson, and M. Ahmadian, "Monitoring and Detecting Fouled Ballast Using Forward-Looking Infrared Radiometer (FLIR) Aerial Technology: Possibilities and Limitations," in *2019 Joint Rail Conference*, (Snowbird, Utah, USA), p. V001T01A019, American Society of Mechanical Engineers, Apr. 2019. doi: 10.1115/JRC2019-1327.

[154] J. Luo, H. Huang, K. Ding, I. I. A. Qamhia, E. Tutumluer, J. M. Hart, H. Thompson, and T. R. Sussmann, "Toward Automated Field Ballast Condition Evaluation: Algorithm Development Using a Vision Transformer Framework," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2677, pp. 423–437, Oct. 2023. doi: 10.1177/03611981231161350.

[155] A. Sabato and C. Niezrecki, "Feasibility of digital image correlation for railroad tie inspection and ballast support assessment," *Measurement*, vol. 103, pp. 93–105, June 2017. doi: 10.1016/j.measurement.2017.02.024.

[156] P. Alves Costa, R. Calçada, and A. Silva Cardoso, "Track–ground vibrations induced by railway traffic: In-situ measurements and validation of a 2.5D FEM-BEM model," *Soil Dynamics and Earthquake Engineering*, vol. 32, pp. 111–128, Jan. 2012. doi: 10.1016/j.soildyn.2011.09.002.

[157] I. Aydin, E. Karakose, M. Karakose, M. T. Gencoglu, and E. Akin, "A new computer vision approach for active pantograph control," in *2013 IEEE INISTA*, (Albena, Bulgaria), pp. 1–5, IEEE, June 2013. doi: 10.1109/INISTA.2013.6577665.

[158] C. J. Cho and H. Ko, "Video-Based Dynamic Stagger Measurement of Railway Overhead Power Lines Using Rotation-Invariant Feature Matching," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, pp. 1294–1304, June 2015. doi: 10.1109/TITS.2014.2361647.

[159] Y. Geng, F. Pan, L. Jia, Z. Wang, Y. Qin, L. Tong, and S. Li, "UAV-LiDAR-Based Measuring Framework for Height and Stagger of High-Speed Railway Contact Wire," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, pp. 7587–7600, July 2022. doi: 10.1109/TITS.2021.3071445.

[160] Z. Liu, L. Wang, C. Li, and Z. Han, "A High-Precision Loose Strands Diagnosis Approach for Isoelectric Line in High-Speed Railway," *IEEE Transactions on Industrial Informatics*, vol. 14, pp. 1067–1077, Mar. 2018. doi: 10.1109/TII.2017.2774242.

[161] S. Gao, G. Kang, L. Yu, D. Zhang, X. Wei, and D. Zhan, "Adaptive Deep Learning for High-Speed Railway Catenary Swivel Clevis Defects Detection," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, pp. 1299–1310, Feb. 2020. doi: 10.1109/TITS.2020.3024216.

[162] J. Wang, L. Luo, W. Ye, and S. Zhu, "A Defect-Detection Method of Split Pins in the Catenary Fastening Devices of High-Speed Railway Based on Deep Learning," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, pp. 9517–9525, Dec. 2020. doi: 10.1109/TIM.2020.3006324.

[163] J. Zhong, Z. Liu, C. Yang, H. Wang, S. Gao, and A. Nunez, "Adversarial Reconstruction Based on Tighter Oriented Localization for Catenary Insulator Defect Detection in High-Speed Railways," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, pp. 1109–1120, Feb. 2022. doi: 10.1109/TITS.2020.3020287.

[164] Y. Chen, B. Song, Y. Zeng, X. Du, and M. Guizani, "Fault diagnosis based on deep learning for current-carrying ring of catenary system in sustainable railway transportation," *Applied Soft Computing*, vol. 100, p. 106907, Mar. 2021. doi: 10.1016/j.asoc.2020.106907.

[165] Y. Han, Z. Liu, Y. Lyu, K. Liu, C. Li, and W. Zhang, "Deep learning-based visual ensemble method for high-speed railway catenary clevis fracture detection," *Neurocomputing*, vol. 396, pp. 556–568, July 2020. doi: 10.1016/j.neucom.2018.10.107.

[166] W. Liu, Z. Liu, Y. Li, H. Wang, C. Yang, D. Wang, and D. Zhai, "An Automatic Loose Defect Detection Method for Catenary Bracing Wire Components Using Deep Convolutional Neural Networks and Image Processing," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–14, 2021. doi: 10.1109/TIM.2021.3113121.

[167] Y. Lyu, Z. Han, J. Zhong, C. Li, and Z. Liu, "A Generic Anomaly Detection of Catenary Support Components Based on Generative Adversarial Networks," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, pp. 2439–2448, May 2019. doi: 10.1109/TIM.2019.2954757.

[168] S. Liu, L. Yu, and D. Zhang, "An Efficient Method for High-Speed Railway Dropper Fault Detection Based on Depthwise Separable Convolution," *IEEE Access*, vol. 7, pp. 135678–135688, 2019. doi: 10.1109/ACCESS.2019.2942079.

[169] B.-K. Kim, W. Kim, C. Lee, M. Yoo, and I. Lee, "Validating Railway Infrastructure Deformation Monitoring: A Comparative Analysis of Field Data and TerraSAR-X PS-InSAR Results," *KSCE Journal of Civil Engineering*, vol. 28, pp. 1777–1786, May 2024. doi: 10.1007/s12205-024-1676-1.

[170] M. M. Rob Mahid, A. F. Clark, J. C. Woods, X. Zhai, and J. Brennan, "UAV and AI-Driven Approaches for Accurate Species Classification in Railway Trackside Vegetation Management," in *2024 29th International Conference on Automation and Computing (ICAC)*, (Sunderland, United Kingdom), pp. 1–5, IEEE, Aug. 2024. doi: 10.1109/ICAC61394.2024.10718773.

[171] V. Ilic, M. Romic, and S. Ilic, "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), pp. 251–254, IEEE, May 2022. doi: 10.1109/ZINC55034.2022.9840689.

[172] M. Roccheggiani, D. Piacentini, E. Tirincanti, D. Perissin, and M. Menichetti, "Detection and Monitoring of Tunneling Induced Ground Movements Using Sentinel-1 SAR Interferometry," *Remote Sensing*, vol. 11, p. 639, Mar. 2019. doi: 10.3390/rs11060639.

[173] Z. Wang, D. Perissin, and H. Lin, "Subway tunnels identification through Cosmo-SkyMed PSInSAR analysis in Shanghai," in *2011 IEEE International Geoscience and Remote Sensing Symposium*, (Vancouver, BC, Canada), pp. 1267–1270, IEEE, July 2011. doi: 10.1109/IGARSS.2011.6049430.

[174] D. Ge, L. Zhang, M. Li, B. Liu, and Y. Wang, "Beijing subway tunnelings and high-speed railway subsidence monitoring with PSInSAR and TerraSAR-X data," in *2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, (Beijing, China), pp. 6883–6886, IEEE, July 2016. doi: 10.1109/IGARSS.2016.7730796.

[175] D. Perissin, Z. Wang, and H. Lin, "Shanghai subway tunnels and highways monitoring through Cosmo-SkyMed Persistent Scatterers," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 73, pp. 58–67, Sept. 2012. doi: 10.1016/j.isprsjprs.2012.07.002.

[176] Y. Zhou, S. Wang, X. Mei, W. Yin, C. Lin, Q. Hu, and Q. Mao, "Railway Tunnel Clearance Inspection Method Based on 3D Point Cloud from Mobile Laser Scanning," *Sensors*, vol. 17, p. 2055, Sept. 2017. doi: 10.3390/s17092055.

[177] M. David Jenkins, T. Buggy, and G. Morison, "An imaging system for visual inspection and structural condition monitoring of railway tunnels," in *2017 IEEE Workshop on Environmental, Energy, and Structural Monitoring Systems (EESMS)*, (Milan, Italy), pp. 1–6, IEEE, July 2017. doi: 10.1109/EESMS.2017.8052679.

[178] M. Gavilán, F. Sánchez, J. Ramos, and O. Marcos, "Mobile inspection system for high-resolution assessment of tunnels," in *The 6th International Conference on Structural Health Monitoring of Intelligent Infrastructure*, p. 10, 2013.

[179] A. Falcone, D. Miccone, and G. Vaccarino, "Primary level uav for tunnel inspection: the pluto project," in *Second Italian Conference on Robotics and Intelligent Machines*, 2020.

[180] Y. Wu, F. Meng, Y. Qin, Y. Qian, F. Xu, and L. Jia, "UAV imagery based potential safety hazard evaluation for high-speed railroad using Real-time instance segmentation," *Advanced Engineering Informatics*, vol. 55, p. 101819, Jan. 2023. doi: 10.1016/j.aei.2022.101819.

[181] Z. Zhang, C. Trivedi, and X. Liu, "Automated detection of grade-crossing-trespassing near misses based on computer vision analysis of surveillance video data," *Safety Science*, vol. 110, pp. 276–285, Dec. 2018. doi: 10.1016/j.ssci.2017.11.023.

[182] Z. Jiang, F. Guo, Y. Qian, Y. Wang, and W. D. Pan, "A deep learning-assisted mathematical model for decongestion time prediction at railroad grade crossings," *Neural Computing and Applications*, vol. 34, pp. 4715–4732, Mar. 2022. doi: 10.1007/s00521-021-06625-z.

[183] G. Song, Y. Qian, and Y. Wang, "Analysis of abnormal pedestrian behaviors at grade crossings based on semi-supervised generative adversarial networks," *Applied Intelligence*, vol. 53, pp. 21676–21691, Oct. 2023. doi: 10.1007/s10489-023-04639-9.

[184] Y. Tang, Y. Wang, and Y. Qian, "Railroad Crossing Surveillance and Foreground Extraction Network: Weakly Supervised Artificial-Intelligence Approach," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2677, pp. 525–538, Sept. 2023. doi: 10.1177/03611981231159406.

[185] H. Qin, A. Zaman, and X. Liu, "Artificial intelligence-aided intelligent obstacle and trespasser detection based on locomotive-mounted forward-facing camera data," *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, vol. 237, pp. 1230–1241, Oct. 2023. doi: 10.1177/09544097231156312.

[186] H. Cui, J. Li, Q. Hu, and Q. Mao, "Real-Time Inspection System for Ballast Railway Fasteners Based on Point Cloud Deep Learning," *IEEE Access*, vol. 8, pp. 61604–61614, 2020. doi: 10.1109/ACCESS.2019.2961686.

[187] T. Hackel, D. Stein, I. Maindorfer, M. Lauer, and A. Reiterer, "Track detection in 3D laser scanning data of railway infrastructure," in *2015 IEEE International Instrumentation and Measurement Technology Conference (I2MTC) Proceedings*, (Pisa, Italy), pp. 693–698, IEEE, May 2015. doi: 10.1109/I2MTC.2015.7151352.

[188] M. Abdellatif, H. Peel, A. G. Cohn, and R. Fuentes, "Pavement Crack Detection from Hyperspectral Images Using a Novel Asphalt Crack Index," *Remote Sensing*, vol. 12, p. 3084, Sept. 2020. doi: 10.3390/rs12183084.

[189] E. Ichi and S. Dorafshan, "Spectral characterization of fouled railroad ballast using hyperspectral imaging," *Construction and Building Materials*, vol. 394, p. 132076, Aug. 2023. doi: 10.1016/j.conbuildmat.2023.132076.

[190] Federal Railroad Administration, "Automated Track Centerline Following for Drone Flight Automation." https://railroads.dot.gov/elibrary/automated-track-centerline-following-drone-flight-automation, 2024. [Online; accessed 28-March-2024].

[191] L. Yu, S. Gao, D. Zhang, G. Kang, D. Zhan, and C. Roberts, "A Survey on Automatic Inspections of Overhead Contact Lines by Computer Vision," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, pp. 10104–10125, Aug. 2022. doi: 10.1109/TITS.2021.3119023.

[192] A. Kumar and S. Harsha, "A systematic literature review of defect detection in railways using machine vision-based inspection methods," *International Journal of Transportation Science and Technology*, p. S2046043024000716, July 2024. doi: 10.1016/j.ijtst.2024.06.006.

[193] Y. Tang, Y. Wang, and Y. Qian, "Real-time railroad track components inspection framework based on YOLO-NAS and edge computing," *IOP Conference Series: Earth and Environmental Science*, vol. 1337, p. 012017, May 2024. doi: 10.1088/1755-1315/1337/1/012017.