A review of applications of visual inspection technology based on image processing in the railway industry

Scarlett Liu¹,², Quandong Wang¹ and Yiping Luo¹,*

¹Central South University, No. 68, South Shaoshan Road, Tianxin District, Changsha, Hunan, China and
²UNSW Sydney, Library Rd, Kensington NSW 2052, Australia

*Corresponding author. E-mail address: ypluo@csu.edu.cn

Abstract

In order to ensure the safety of railway transportation, it is necessary to regularly check for faults and defects in the railway system. Visual inspection technology is conducive to improving the low efficiency, poor economy and inaccurate detection results of traditional detection methods. This paper introduces the research and contribution of various scholars in the field of visual inspection, summarizes the application and development of visual inspection technology in the railway industry, and finally forecasts the future research direction of visual inspection technology.

Keywords: visual inspection; image-processing computer vision; railway; high-speed train

1. Introduction: background to the application of image-processing technology in the railway industry

Rail transit has become the most popular mode of transportation in China, and is favoured by the public because of its large transportation volume, fast transportation speed and low consumption. According to statistics, China’s current railway network is the largest in the world, with 127,000 km of track in operation. With such a long railway mileage, routine inspection and maintenance of the railway system is a daunting task.

The inspection of railway systems is mainly done manually by technicians. Taking rail fastener detection as an example, the technician needs to check the presence or absence of each fastener and the condition of the fastener along the railway line. It takes a long time to check the entire line, and detection accuracy is generally not satisfactory. In recent years, with the development of...
computer technology and monitoring technology, it has become possible to use visual inspection technology for railway systems instead of relying on manual troubleshooting.

A railway visual inspection system is a fast, accurate and economical transportation monitoring system that effectively integrates data acquisition technology, monitoring technology and image-processing technology into railway system fault detection. The system is divided into two main components: hardware and software. The overall hardware composition is shown in Fig. 1.

The detection target refers to those parts of the railway system that need to be periodically checked. The image information is collected by the data collection equipment and sent to the computer for processing. The software component comprises a computer program that includes an image-processing algorithm. The hardware and software work in coordination to complete the visual inspection of the rail system components and to feed back the results to the staff.

Image-processing technology plays an important role in the visual inspection system, and the image-processing algorithm directly affects the detection accuracy and detection speed. This paper describes in detail the application and development of image-processing technology in the railway industry.

Digital image processing as a discipline arose in the early 1960s. The purpose of early image processing was to improve the quality of images. It was aimed at human beings to improve the visual recognition effect of people. In image processing, the input is a low-quality image, and the output is an image with improved quality. Common image-processing methods include image enhancement, restoration, encoding and compression. The first successful application was by the American Jet Propulsion Laboratory (JPL) [1]. The JPL used image-processing techniques such as geometric correction, gradation transformation and noise removal on the thousands of lunar photographs sent back by the Space Detector Ranger 7 in 1964, taking into account the position of the sun and the environment of the moon. The impact of the successful mapping of the moon surface by computer was immense. Since then, the scientific community has begun research on image-processing technology.

Since the adoption of image-processing technology by the railway industry, visual inspection technology based on image processing has been greatly developed and applied to different parts of the railway system. Fig. 2 shows statistical data for the research conducted in the literature on visual inspection examined for this study. Based on the different test objects, we can classify the fault-detection methods used in the railway industry into four categories: the railway, the pantograph–catenary network system, the train body and infrastructure. This paper uses these four categories to review the application of image-processing technology in the railway industry.

The structure of the paper is as follows. In the following section, the application of image-processing technology in orbit detection is introduced. In Section 2, the application of image-processing technology in pantograph–catenary network detection is introduced. In Section 3, the application of image-processing technology in the detection of train body parts is discussed, while Section 4 introduces the application of image-processing technology in infrastructure inspection (such as stations, tunnels, crossings and so on).

---

**Fig. 1.** The hardware composition of the visual inspection system

**Fig. 2.** Statistical table of detection targets proposed in papers examined for this study
Finally, Section 5 summarized the problems in the existing methods and made a prospect for the application of computer vision in the railway industry.

2. Railway track

The condition of the railway track is critical to the normal operation of trains. Failure of any type of component (such as the rails, sleepers, coupling parts, anti-climbing equipment and ballasts) can cause major safety incidents, even train derailments. In order to ensure the safety of the train, it is necessary to check the railway line periodically. Traditional inspections are usually performed by trained workers who walk along the railway line to identify potential risks. However, manual inspection is time-consuming, laborious and dangerous, with great limitations in detection speed, quality, objectivity and detection range, and there is therefore an urgent need to introduce more reliable and effective orbit-detection methods. With the application of image-processing technology in the railway industry, the labour intensity of technicians is reduced, maintenance costs are reduced, and the detection efficiency of railway components is greatly improved.

At present, the hardware component of the visual inspection system of a given railway component is composed of a track component, a data-acquisition module and a data-processing module, as can be seen in Fig. 3, which shows a data-acquisition module consisting of a camera, a light source and an auxiliary sensor installed on the train (IMU, laser transmitter, and so on). The data-processing module is a computer that processes image information. During train operation, the light source emits light, illuminating the track assembly, and the camera captures the image information of the track assembly. The image information is then transmitted to the computer, and after calculation, the condition information of the track component is extracted.

2.1 Rail surface

Rail surface defects consist mainly of surface cracks and rolling contact fatigue wear, as shown in Fig. 3. Rail defects directly cause vibration during train operation, which seriously affects the comfort of the train. Excessive wear and fatigue of the rails will lead to the continuous exacerbation of rail defects, which will eventually lead to rail breakage and cause major accidents.

The early rail surface-defect detection method was based mainly on the texture characteristics of the image, and the surface defects of the rail were identified by filtering, texture analysis and classification of the rail surface image. Mandriota et al. [4] processed the orbital image by the Gabor filter bank, extracted the mean and variance using the obtained feature image to construct the feature vector, and performed the texture classification using the k-Nearest classifier. Their experiment proved the feasibility of the texture-detection method in the identification of rail surface defects. Subsequently, Mandriota et al. [5] compared the effects of filters such as the Gabor filter algorithm, wavelet-transform algorithms and the Gabor wavelet transform-algorithm on the texture detection of the track surface, analysing the track surface texture to detect the existence of ripple defects on the rail surface. They verified that the Gabor wavelet-transform algorithm had better performance in track ripple detection than other filters. Deutschl et al. [6] used the spectral image-difference algorithm to classify defect textures on the track surface. The convolution filter was used to process the surface image of the rail to obtain the texture feature map, and the decision-tree algorithm was then used to classify the feature map. The detection speed of this method reached 20 times that of the traditional rail surface-defect detection method. Jie [7] proposed a simple and fast algorithm that directly analysed the grey histogram curve (the first-order statistical texture characteristics) of the rail head surface image to identify the special rolling contact fatigue of the rail head. This method proved to be less affected by image representation and more robust than pixel-level texture methods. Huber [8] used a colour line-scan camera to perform image acquisition under
two different wavelengths of light to obtain 2D image data of the track. The Gabor filter bank was used for the 2D texture description. After Gaussian mixture modelling, the Bayesian classifier was used for texture classification. Finally, the 2D image data and 2D texture information were integrated for surface defect analysis.

Due to the obvious surface defect characteristics of the rail, the image implies different semantics from the normal rail surface image. Some researchers used the image features contained in the surface defect image of the rail to detect the surface defects of the rail. Liu et al. [9], Chen et al. [10] and others achieved the automatic detection of cracks on the surface of the rail through image processing. The crack region was extracted on the image using an adaptive threshold algorithm, and a dynamic template based on the morphology of the crack was used to detect the continuous crack boundary and estimate the length of the crack. Wang et al. [11] used the learning partial differential equation of a Gaussian blurred image to perform image preprocessing for rail surface-defect detection. This method effectively removed image blur and improved the accuracy of defect detection. Marmol and Mikrut [12] obtained the head position information from the image and laser data, extracted the data from the auto-approximation mode of the track edge, fed back the position in space with high-precision laser data, and determined in the digital image the position of the rail head. This method was more general, but not suitable for low-contrast backgrounds. Li and Ren [13, 14] proposed an intelligent visual detection system (VDS) for discrete surface defects. A contrast-based image-enhancement algorithm was used to improve the discrimination between the defect and the background, and the proportional emphasis maximum entropy threshold algorithm was used to obtain the threshold for maximizing the defect entropy while keeping the defect ratio at a low level. The detection rate of discrete defects on the track surface using this VDS was able to reach 91.61%. Liu et al. [15] used an electromagnetic field to collect the defect information of the rail by calculating the sensitivity matrix using the forward problem of the electromagnetic finite-element model. Defect image reconstruction was simulated using linear projection and the Tikhonov regularization algorithm to roughly detect the shape and position of defects in the track. This experimental system was able to reconstruct defects on the track surface and internal distribution, and obtained more multidimensional projection information than the traditional measurement method. Dubey et al. [16] used a visual inspection technique based on the maximum stable extreme value region marker to extract the defects on the rail surface from the image data, through image feature calculation, three different orbital surface defects, with a detection accuracy of up to 98%. Shi et al. [17] proposed an improved Sobel algorithm to detect rail surface defects. In order to solve the problem of the Sobel algorithm’s lack of sensitivity in the X and Y directions, they added defect templates in six different directions. The defect-recognition rate of the improved Sobel algorithm was 10% higher than that of the Sobel algorithm. Ta¸stimur et al. [18, 19] processed video data in real time based on the morphological features of the rail surface-defect image. This method was more efficient and had good robustness under different illumination conditions. Yaman et al. [20] used the Otsu segmentation method to locate the surface defects of the rail in the image, and obtained the characteristic signal by calculating the variance value of the surface image to detect the fault. They then extracted the feature types from the attribute signals and used fuzzy logic to classify the faults. This method was effective at detecting cracks, erosion and undulations on the track surface. However, the detection rate of rail crust, wheel burn and rail surface collapse was very low. Gan et al. [21] used a graded extractor to detect rail surface defects. The coarse extractor was used to explore the background features and extract the rails, while the fine extractor analysed the background mode and spatial position information to distinguish the mode of the defect. This method achieved a 100% recall rate and a very low false-positive rate. Zhuang et al. [22] used Haar mode as an important feature of cracks. Three cascaded classifiers based on the LogitBoost algorithm were integrated with a voting scheme to scan the orbital image and then use the Otsu algorithm to identify cracks. It has been verified that the frame is the most effective means of detecting track surface cracks. Min et al. [23] extracted the surface defects of the orbit using the morphological algorithm, and traced the direction chain code in the LabVIEW environment to obtain the defect features. This method was able to detect speeds of up to 2 m/s in real time.

With the rapid development of machine-learning technology, researchers began to construct rail surface-defect classifiers based on machine-learning algorithms. This method requires estab-
lishing a data set of the track surface defect image, manually calibrating the track surface defect type, and using this data to train the classifier. The trained classifier is adjusted and optimized to detect rail surface defects. Wang et al. [11] proposed a method for detecting surface defects of rails based on quantum neural networks, extracting feature vectors from different spaces, and training QNN models using a sigmoid function and a multilayer activation function. The effectiveness and potential value of quantum neural networks in the detection of surface defects of rails were demonstrated by these experiments. Faghih-Roohi and Hajizadeh [24] compared the effects of deep convolutional neural networks (DCNNs) on three different scales (small, medium and large) on the detection of rail surface defects. They concluded that although the network training took a long time, the large DCNN model was better than the small and medium DCNN models for the classification task. Ma et al. [25] proposed an automatic evaluation-classification algorithm for rail surface damage in a two-layer support vector machine (SVM) network. The track surface image feature was used to train the first-level SVM to obtain the rail image area in the image. The second-order linear kernel SVM was then used as a classifier to distinguish the types of surface defects on the rail. The accuracy of the rail surface-defect detection of the device was above 90%. Santur et al. [26, 27] used a deep-learning method to process rail surface-defect images from 3D laser cameras, and introduced the Cuda library to retrieve the GPU for image classification. This method was able to achieve the real-time detection of rail surface defects at speeds of 144 km/h, with a recognition accuracy of up to 99%. Li et al. [28] used an orbital defect-localization algorithm based on weighted projection contours to segment the image and identify the rail ripple defects using the support vector machine. This method was robust to illumination changes, and the detection accuracy and recall rate were 98.47% and 96.50%, respectively. Shang et al. [29] proposed sequential processing of rail images using target location and a convolutional neural network (CNN). Canny edge detection was performed on the rail image and the pseudo-edge points were removed, and the obtained image was classified and identified using the pre-trained CNN. The detection accuracy and recall rate of this method were 92.08% and 92.54%, respectively. Chen et al. [30] proposed an adaptive weighted multi-classifier fusion-decision algorithm. An SVM was used to classify the MFL signals as single-channel and unidirectional, and the adaptive weights of different SVMs were assigned according to the entropy calculated by the posterior probability. Finally, the weighted majority-voting strategy was used to combine the classification results of different channels and different directions to make comprehensive decisions. This method had good robustness and a good recognition rate.

2.2 Track component wear and deformation detection

The degree of rail wear is an important indicator when assessing train safety. Excessive rail wear results in unstable train operation and increases the possibility of accidents. At present, rail-wear detection is mainly done manually, which is inefficient and often produces inaccurate test results. The typical response of railway managers to rail wear problems is to periodically replace the rails, which causes many rails with low wear to be decommissioned early. Although this can ensure the safety of trains, it also greatly reduces the economics of railway operations. The use of visual inspection methods to assess the amount of wear on the rails makes it easier for railway managers to control the condition of the rails.

Measuring the cross section of the rail is a prerequisite for obtaining the wear of the rail. In order to obtain the profile of the rail, researchers have tried various methods. Alippi et al. [31] used a CCD camera to capture a laser beam projected onto the rail and used a pre-filtering algorithm to limit the image area of the orbital contour while ensuring the accuracy of the contour generation through neural reconstruction. Jin et al. [32, 33] designed a single-element laser beam projector for diffusing light projected onto the surface of the rail into a uniform intensity line, and improved the direct linear transformation algorithm (DLT) to correct the nonlinear lens distortion. After the orbital profile was reconstructed, the corner algorithm was used to detect the key points in the track profile curve and match the corresponding points on the standard track section to visually determine the amount of wear. Liu et al. [34] used the centre of the large and small circles from the profile of the rail waist as the control point to record the measured track profile to the reference profile. The amount of track wear was obtained by comparing the reference profile with the standard profile. Karaduman et al. [35] used
filters and preprocessing algorithms to extract feature information from image information and depth information, and then used a fuzzy logic-based intelligent classifier to obtain rail wear from the feature information. This method was robust and efficient. Sun et al. [36] proposed a motion deviation-correction method. Track profile image data was obtained using a multiline structured-light vision sensor with multiple parallel planes, and the track profile is projected onto the auxiliary plane using a motion deviation correction method to reconstruct the rail profile to calibrate the perpendicularity between the structured light plane and the track. This method overcomes the error caused by the non-perpendicularity between the structured light plane and the track when scanning the track profile, and reduces the track wear measurement error. Zhou et al. [37] proposed an improved structured-light measurement system that solved the problem caused by the limited number of extractable calibration points during free-target calibration by extracting the spot of the laser emitter projected onto the rail surface. The measured profile was then compared to the standard profile of the rail to quantitatively evaluate the rail wear. At the same time, in order to comprehensively monitor the condition of the track, the deformation monitoring of other track components has also been favored by researchers. Labarile et al. [38] proposed a visual system capable of detecting railway anomalies. The matching tracking method was used to extract image features, and a similarity function was used to calculate the consistency of the left and right images. This approach enhanced quality control of the road system and reduced maintenance costs. Yaman et al. [39] proposed a method for detecting track surface and component defects based on particle swarm optimization (PSO). Their method was divided into two parts. The first part encompassed the detection of the surface defect on the track and the matching of the specific area of the image with the template to determine the surface-defect ratio of the rail; the second part comprised the detection of the track component and the calculation of the correlation coefficient between the image and the template. The PSO algorithm was then used to detect the track fasteners. Finally, the two test results were combined to obtain the condition of the track. The advantage of this method was the ability to simultaneously detect rail surfaces and track fasteners. Sabato and Nizerecki [40] used digital image-correlation techniques and pattern projection for track-deformation evaluation. They applied different-sized forces on a prefabricated experimental model to obtain the deformation and displacement information of the sleeper, and then used pattern drawing and pattern projection to conduct experiments, exploring the influence of damping on image resolution. The detection resolution was expected to be controlled at around 5 μs.

2.3 Rail component identification

Track fasteners are important components that prevent rail overturning and longitudinal movement. The failure of rail fasteners can directly lead to the loosening of the rails, which in turn leads to serious accidents. The track fasteners are therefore a focal point of daily railway maintenance operations. The main types of track fastener failure are missing, loose and damaged. Because loose fasteners are also loose in the image, the characterization is obvious, but detection is difficult and dependent largely on manual labour. Therefore, the visual inspection of track fasteners is used primarily to detect the absence of or damage to fasteners. Mazzeo et al. [41] proposed a vision-based railway-tight bolt detection technology to automatically detect the presence of fastening bolts that secure the rail to the sleeper. The method preprocessed images using several combinations of wavelet analysis (WT) and principal element analysis (PCA), and used a neural network to classify the extracted image information. Their final detection system was applied to long sequences of real images with high reliability and robustness. Singh et al. [42] used a series of algorithms such as Gaussian smoothing, edge detection, and short-line removal to perform image preprocessing, and extracted the rail clips by image classification and specific window recognition. The detection accuracy of the method for gray clips reaches 96.5%. Marino et al. [43] proposed a real-time detection system for railway maintenance to achieve automatic detection of the presence or absence of railway set screws. The system used two discrete wavelet transforms to simultaneously process the image data and provide the resulting information to multilayer perceptual neural classifiers (MLPNCs). These MLPNCs allowed the system to almost completely avoid false alarms, and the detection accuracy of visible and missing bolts when detecting the tightening bolts reached 99.6% and 95%, respectively. The FPGA-based architecture performed these tasks in 8.09 μs, detecting at speeds of up to 200 km/h. Sawadisavi et al. [44] proposed using a machine-vision...
system for data acquisition and analysis. The preliminary research and development focused on the inspection of spikes, anti-climbing devices, inter-pillow ballasts, transition points and ballast components. The algorithm used edge detection and texture information to provide a detection-oriented track, tie rods and tie plates. This offered a robust means of narrowing the search area. Within the search area, an algorithm was used to determine whether the spike and anti-climbing device were present. De Ruvo et al. [45] proposed a patented real-time inspection system for rail set screws. The system trained the Daubechies classifier and the Haar classifier, respectively, and implemented the detection based on the GPU. The detection accuracy of this system was 99.9%, and the rail set screws were able to be detected in real time at a speed of 187 Km/h. Compared to that of CPU-based spike detection, the detection efficiency of this method was increased by 287%. Resendiz et al. [46] developed a machine-vision system for automatically detecting specific components in a track structure. The digital image of the track was recorded using a video capture device, and the machine-vision algorithm processed the resulting video data using a global-to-local component-identification method. Edge- and texture-based detection techniques were used to narrow down areas where components might be detected. The system provided a robust approach to orbit component detection and enhanced the ability to predict the long-term condition of the orbital system. Rubinsztejn [47] proposed a track component-detection system based on image processing and pattern recognition to automatically detect the presence of railway components. The system acquired a real image of the track through a digital scanning camera installed underneath the train. Using a combination of image processing and pattern recognition, the Viola–Jones object-detection frame was applied to the detection of railway fasteners, with a detection accuracy of 98%. Yang et al. [48] proposed a method for detecting track fasteners based on image processing and pattern matching. The method extracted the direction field of the extracted track fastener as a description for pattern recognition. In addition, an appropriate matrix of weight coefficients was proposed for robust and fast pattern matching in complex environments. The experimental results showed that the algorithm had a high degree of computational efficiency (it was able to detect track fasteners at a speed of 400 km/h) and was robust enough for fastener detection in complex environments with an accuracy of 99.99%. Trinh et al. [49] proposed a real-time automatic rail-detection system for rail components such as sleepers, connecting plates and rivets. The system first efficiently collected the image information, location information and speed information of the track, and then, after integrating the acquired information, mapped the detected object to the physical track object, and performed further data integration and analysis to detect defects in the sequence level track. Finally, quantitative analysis of tests was performed under different orbital conditions to evaluate rivet conditions and compliance anomalies. The system obtained a satisfactory test result (a joint plate detection accuracy of 99.3%, a sleeper detection accuracy of 88.2%, and a rivet detection accuracy of 96.5%), and was the first system to simultaneously evaluate rivet status and track compliance level. Resendiz et al. [50] developed a railway component repetitive tracking algorithm for tracking railway components in video data. Based on the method of spectrum estimation and signal processing, the periodic changes of various parameters of the railway track are detected, and the detection and defect evaluation of the track, ballast and sleeper can be performed simultaneously. Li et al. [51] proposed a railway track detection system based on real-time computer vision. The hardware system of the research consists of multiple cameras, global positioning systems and distance measuring instruments. First, use four cameras fixed on the inspection vehicle to obtain the video, then use the global positioning system and the distance measuring instrument to synchronize the positioning and speed information with the video, start the railway component detection and optimization system to analyze the image, and finally use the abnormality detection module. Evaluate the level of association and compliance level of the two types of data. It has been verified that the average detection accuracy of the system for the three track components is 94.67%. Khan et al. [52] proposed an automatic detection technique for detecting the absence of rail fasteners. The Harris-Stephen and Shi-Tomasi feature detectors are used to extract the feature points and feature vectors of the image, and the features of the input image are matched with the features of the training image. If the matching value is less than the threshold, the rail fastener is inferred to be missing. It has been verified that the detection accuracy of the rail fasteners in the system is 83.55%. Feng et al. [2] proposed a machine-vision
system for detecting partial wear and loss of fasteners using a probabilistic subject model. This method enabled the use of unlabelled data to simultaneously simulate multiple fasteners with different orientations and lighting conditions, as well as the division of faults into three levels based on the maximum-likelihood probability method. Gibert et al. [53] proposed an SVM-based track fastener-recognition method. First, the data set used to train the SVM was adjusted; the intra-class changes were then reduced; finally, the difficult samples were bootstrapped to increase the classification bounds. Through a combination of the direction-gradient feature of the histogram and the linear SVM classifier, the accuracy of the detection of the fastener reached 98%, with a false positive rate of 1.23%. In order to improve the efficiency of rail fastener defect detection, WANG [54] et al. proposed an automatic detection method for rail fastener defects. Firstly, the background difference method is used to accurately locate the rail fasteners, and the linear features of the image are extracted based on the improved Canny operator and Hough transform method. Then, by combining the local binarization (LB) and the directional gradient direction histogram (HOG) Extracts the feature vector of the rail fastener defect; Finally, the feature vector is classified using the support vector machine (SVM). This method has higher real-time and accuracy than the traditional rail defect detection method. Aytekin et al. [55] proposed a laser-ranging camera-assisted real-time track fastener-detection system. First, extensive analysis based on pixel and histogram similarity methods was carried out on specific road segments. A fusion-phase method was then introduced, and the feasibility of this method was verified using a large database. Finally, the design experiment was implemented in LabVIEW. The system detected the hex head bolts on the track at a speed of 100 km/h.

2.4 Rail extraction and obstacle elimination

Following the adoption of visual inspection systems by the railway industry, this technology began to be used for autonomous driving and foreign object detection. A number of orbital extraction methods and foreign-matter detection methods have emerged.

Schewe et al. [56] designed a railway clearance obstacle detection software based on the LiMeZII prototype. Based on the principle of photogrammetry, the special conditions and key points of railway obstacle photography are analyzed. The software has demonstrated its efficiency through the evaluation of 18000 narrow passages over the past 3 years. Maire [57] introduced a vision-based collision-avoidance system and designed and implemented a prototype in MATLAB. The collision-avoidance system for track maintenance vehicles allowed the distance between two maintenance trains to be eliminated; the system had an original self-calibration module that took full advantage of the constrained geometry of the orbit, and also developed an algorithm for determining the initial segment of the orbit. This self-organizing technique was limited by the use of road lane marking-detection algorithms. This system exhibits good robustness under extreme lighting conditions and is at least four times more resistant to collisions than existing systems.

Kaleli and Akgul [58] proposed a dynamic programming method based on vision for railway orbit extraction. Their method was divided into three parts. First, the video data was acquired by the camera installed on the head of the train. Second, feature extraction was performed on the track to obtain the disappearance point of the train. Finally, the dynamic image-planning method was applied to the gradient image based on the vanishing point. At the same time, the rails on the left and right sides were extracted and the track section was calculated. Experiments showed that this method had good robustness at higher vehicle speeds.

Maire and Bigdeli [59] proposed a visual collision-avoidance system based on C++ (using the OpenCV computer vision library), which tracked two orbital curves by evaluating 30 m orbital parallel lines to ensure different illumination intensities. The track could be detected underneath, avoiding the problem of selecting the Canny edge threshold, directly using the relative gradient amplitude caused by the pixel intensity. Finally, the camera image was transformed into a parallel orbit using the Hough transform. The system has been shown to exhibit robustness under a wide range of lighting conditions, but tests for rain and fog have not been conducted.

Wohlfett [60] proposed a vision-based switch-detection method based on images from a digital camera to observe the space in front of (or behind) the train, by identifying all areas of an image visible in a particular detection. The track provides a wealth of information about the portion of the railway network that is currently being traversed, including the location of nearby parallel tracks and an estimate of the radius of the curve of
the track currently being traversed. Wohlfeil's method used real data from various test locations in various weather conditions and environments. It has proven to be very powerful and highly practical for rail-based selective self-positioning of railway vehicles. Train collisions can be avoided to a large extent, and the method has high robustness and reliability. In a live demonstration at Wegberg-Wildenrath, it also proved its real-time performance on mobile hardware.

Nassau and Ukai [61] described a method for extracting tracks by matching edge features to candidate orbit patterns that were simulated as parabolic segment sequences; the pattern was pre-computed in the semi-automatic offline phase for use in areas near the camera, while dynamic generation was used in more distant areas. The method aimed to solve the challenges posed by an unknown environment, without clear understanding of the train speed or camera parameter position information, and without certain hardware requirements; the running speed was therefore fast enough to play a role in practical applications.

Elberink et al. [62] proposed a method for railway rail reconstruction. The point cloud data of the scene is obtained by using a laser scanner, and the acquired point clouds are clustered according to the position, relative height and linearity of the railway track. The railway rail point cloud is fitted using an interpolated low-order Fourier curve to reconstruct the rail model. The reconstruction accuracy of this method is about 2cm.

Weichselbaum et al. [63] proposed a 3D vision-based obstacle-detection system. The system used a laser scanner to operate single and stereo cameras in the visible and infrared spectra, as well as radar and ultrasonic sensors. The system used a multi-baseline camera system, first filtering the image gap, marking adjacent pixels, independently evaluating the region of interest, and finally using layered stereo matching to achieve obstacle detection. It has been verified that the obstacle detection rate is increased from 89.4% to 97.75% using this system, while the false positive detection rate can be kept as low as 0.25%, and the overall delay of obstacle detection is much faster than 300 ms.

Selver et al. [66] designed a camera-based driver assistance system. The image acquired by the camera is divided into four different sub-regions according to the distance. Then different filters are designed based on the 2-D Gabor wavelet decomposition method to extract the railway rails in the corresponding sub-regions. Finally, based on the morphological analysis, the rails extracted from different sub-regions are spliced to obtain a complete railway rail. This method is robust and can be used in complex environments.

Kishore et al. [67] proposed a method based on CV for monitoring the orbital state. The method captured the track opened by the train and the adjacent track image by a camera placed on top of the train. On these images, edge and feature extraction methods were applied to determine the orbit. The faults generated between the railways were investigated to determine whether there was a fault. The experimental results showed that the method was tested and the observed results were very effective.

Wu et al. [68] proposed a real-time tram-detection method. First, the local cumulative histogram was used to estimate the adaptive parameters, after which the adaptive multilevel threshold was used to segment the ROI of the trolley track. The regional growth method was then used to reduce noise, extract the best tram track, and predict the trend of the tram. It has been verified that the method exhibits high accuracy and real-time performance under different environments.
Singh et al. [69] proposed the use of drones in computer-vision systems for monitoring. The study first recorded video and images through 4K cameras and Sony sensors, and used MATLAB to perform different preprocessing tasks such as image-size adjustment. After Gaussian smoothing for noise reduction, HSV colour extraction converted the image to greyscale, located the region of interest using the Canny edge detector for edge detection, then performed orbit detection and meter measurement for defect detection. It has been verified that the method can perform orbit detection under different weather conditions.

Gabara and Sawicki [70] proposed an image-based point-cloud measurement method based on the configuration of digital image and reference control networks, using the Nikon D5100 DSLR camera to obtain the railway track test portion. Six digital images, dense point clouds and 3D mesh models were generated using two software systems, Reality Capture and PhotoScan, which implemented different matching and 3D model reconstruction techniques (multi-view stereo and semi-global matching), all of which were able to generate suitable 3D models. The MeshLab program was used to filter the final mesh of the 3D model. The Cloud Compare application was then used to determine the cross section of the track gauge and tilt definition, and the results obtained from the point cloud by dense image-matching techniques were compared to the results of direct geodetic measurements. The RMS difference obtained in the horizontal (specification) and vertical (tilted) planes was RMS $\Delta_1$ 0.45 mm. The accuracy achieved was in accordance with the accuracy conditions (error $m_1$ 1 mm) of the orbital measurements and inspections specified in the Polish branch railway and the European technical specifications.

3. Pantograph–catenary system

The pantograph–contact network system is an energy-supply facility for electric locomotives. Its failure results in a disruption in the energy supply to the locomotive, which leads to serious accidents. In order to ensure the normal operation of the train, the pantograph–contact network system must be periodically checked for faults. The pantograph–contact net contains a large number of components, and failure of any component may cause loss of function of the net system. Researchers have conducted targeted research on different parts of this system.

In order to reduce installation and maintenance costs, Chen et al. [71] proposed a new automatic onboard neutral section passing control device based on image recognition. Its hardware mainly includes high-speed digital signal processor (DSP) and industrial smart camera. The industrial smart camera is used to collect the images of break/close markers along railways, and the edge information and edge refinement of the image are used to extract the feature information such as the size and shape of the markers. The break/close time is optimized by calculation of image size of break/close markers. The results of the experiment show that with the device is not only the distance between the train and break/close markers can be determined but also the train's speed can be calculated. Aydın et al. [72] proposed a pantograph–contact network-fault diagnosis technique based on image processing for detecting arc faults between the pantograph and the catenary. The technique extracted the edge of the pantograph using an edge detection algorithm, and detected the arcing phenomenon between the arch nets by checking the position of the contact line at the edge of the pantograph. Experiments have shown that this technology has broad potential in diagnosing train arch failure. Aydın et al. [73] proposed a pantograph–contact network-fault diagnosis technique based on image processing for detecting arc faults between the pantograph and the catenary. The technique extracted the edge of the pantograph using an edge detection algorithm, and detected the arcing phenomenon between the arch nets by checking the position of the contact line at the edge of the pantograph. Experiments have shown that this technology has broad potential in diagnosing train arch failure.

Researchers have conducted targeted research on different parts of this system.

In order to reduce installation and maintenance costs, Chen et al. [71] proposed a new automatic onboard neutral section passing control device based on image recognition. Its hardware mainly includes high-speed digital signal processor (DSP) and industrial smart camera. The industrial smart camera is used to collect the images of break/close markers along railways, and the edge information and edge refinement of the image are used to extract the feature information such as the size and shape of the markers. The break/close time is optimized by calculation of image size of break/close markers. The results of the experiment show that with the device is not only the distance between the train and break/close markers can be determined but also the train's speed can be calculated. Aydın et al. [72] proposed a pantograph–contact network-fault diagnosis technique based on image processing for detecting arc faults between the pantograph and the catenary. The technique extracted the edge of the pantograph using an edge detection algorithm, and detected the arcing phenomenon between the arch nets by checking the position of the contact line at the edge of the pantograph. Experiments have shown that this technology has broad potential in diagnosing train arch failure. Aydın et al. [73] proposed a pantograph–contact network-fault diagnosis technique based on image processing for detecting arc faults between the pantograph and the catenary. The technique extracted the edge of the pantograph using an edge detection algorithm, and detected the arcing phenomenon between the arch nets by checking the position of the contact line at the edge of the pantograph. Experiments have shown that this technology has broad potential in diagnosing train arch failure.
power line and the support arm were detected to generate the region of interest for detailed processing. The hypotheses of potential anomaly patterns were then collected from the endpoints extracted from the local curves and line segments to enhance the perceived difference. Following this, the anomaly model and background model were implemented to classify the candidates in the hypothesis. The experimental results showed that the scene-structure-enhanced Bayesian detector was superior to the classical object detector in the presence of a large amount of interference and variance. Arastounia et al. [76] proposed a method for identifying railway infrastructure based on 3D lidar data. Data acquisition is performed using the Optech Lynx mobile mapping system fixed on the train. The railway components are identified by physical properties, geometric characteristics and topological relationships between the components, and the railways key components such as the catenary and the catenary support columns are identified by image processing algorithms. The point cloud horizontal accuracy of this method is above 96.4%.

Aydin [77] proposed an automatic detection system based on the firefly optimization algorithm for detecting the pantograph and the arcing that occurs. The method located the pantograph within a rectangular frame and used the Otsu method based on the firefly algorithm to detect the arcing occurring in the rectangular frame. Experimental results showed that the average detection accuracy in the four true pantograph videos was above 96%, while the detection time per frame was 0.031 s.

Wu et al. [78] proposed a new bird’s nest detection framework for SVMs. The main processing flow of this framework was: image binarization, branch detection, hanging-point detection, stripe extraction and pattern learning. The Canny edge detector was used to detect the backbone of the contact network; the Hough point detection method was used to detect the suspension point of the bird’s nest; the stripe direction histogram and the stripe length histogram were used to assist with pattern recognition of the nest; and the SVM was used to implement pattern learning. The detection rate of this method reached 91.06%. Ge et al. [79] proposed a detection method for a pantograph-detection system based on Faster R-CNN. The method used a regional proposal network (RPN) to generate a bounding box proposal, using a bounding-box suggestion method to generate a bounding box for the image, after which the image was passed to the CNN. A region of interest (ROI) was then applied to the final layer Conv feature to create a fixed-size vector. The ROI vector was then passed to the classification network and the regression network. In terms of system processing speed, some convolutional layers and local response normalization (LRN) layers were removed to improve processing speed. It has been verified that the detection accuracy of this method exceeds 94.9%. The system can work with subway trains in a variety of environments. Capece et al. [80] proposed a Pantobot-3D self-checking system that used a self-checking algorithm to enable 3D representation and modelling of pantographs, classification of pantograph models, and integrity assessment of pantograph assemblies. This system has been functionally tested on the main axis of the Italian high-speed rail network. In the study of Karaduman et al. [81], an approach using deep learning is proposed for the detection of arcs in pantograph-catenary systems. Arc detection is performed using CNN (Convolutional Neural Network). Compared with similar methods, the detection effect of this method is competitive. In order to achieve railway tracking and railway fault diagnosis, Karakose et al. [83] proposed a computer vision based solution. The railway image in front of the train is acquired using a camera mounted on the top of the train, the railway rail profile in the image is extracted based on Canny edge detection, and the fault of the railway rail is diagnosed based on the morphological processing method. This method can accurately identify the pitch fault, dressage fault and expansion fault of railway rails.

Based on the principle of laser ranging, Liu et al. [84] proposed a high-precision detection approach for catenary geometry parameters of electrical railway. The image coordinates of the spot formed on the contact line by a laser emitter are mapped to the world system to obtain the conductor height and stagger. The genetic particle filter algorithm based on particle swarm optimization (PSO-GAPF) is presented to track and locate The addition in the image. In addition, Kalman filter is used for correcting the detected value of catenary geometry parameters. This method has a high detection speed, and can meet the requirements of real-time catenary geometry parameters detection system.

Chen et al. [84] proposed a contact network support device fastener-detection system based on a DCNN, including three coarse-to-fine detection stages based on DCNNs. Two of the detectors were used to position the cantilever hinge and its fasteners, and a classifier was used to diagnose fastener defects. The system had a high detection rate and
good adaptability and robustness in complex environments. Liu et al. [85] used a variety of DCNNs to evaluate the detection of contact network support components. A series of experiments were performed in a unified test environment for detection of components using neural network architectures such as Faster-CNN, R-FCN, SSD and YOLOv2. R-FCN proved more suitable for detecting contact net support components by comparing different evaluation indicators such as accuracy, recall rate and average accuracy.

4. Train body components

4.1 Wheel-set defect detection

The wheel pair is the part of the train that is in contact with the rail, receives all static and dynamic loads from the rolling stock, and transmits the load generated by the irregularity of the line to the rolling stock. Wheel defects are a major source of damage to railway installations and vehicles, increasing vibration during driving and generating a great deal of noise. It is therefore important to regularly check the condition of train wheels for the safe operation of the railway system.

Zhang et al. [86] proposed an online non-contact method for measuring wheel-set parameters based on machine-vision technology. A CCD camera with a selected optical lens and frame grabber was used to capture an image of the light profile of the wheel set illuminated by the linear laser. The analogue signal of the image was converted to a corresponding digital grey value. The image pixel coordinates were then transformed into spatial coordinates using the mapping-function method. After the image of the wheel set was captured, the internal thickness of the rim and the thickness of the flange were measured and analysed. Theoretical and experimental results showed that a CV measurement system could meet the requirements of online wheel-set measurement. Gao et al. [87] used a line-structured light vision sensor to dynamically measure the wheel diameter. The sensor is used to collect the contour data of the wheel, and then the wheel contour diameter edge point set is selected from the contour data, and finally the selected point set is fitted into a spatial circle to obtain the wheel diameter. Experimental results have proved that this method is reliable and efficient. Krummenacher et al. [88] proposed two machine-learning methods for detecting wheel defects based on wheel vertical forces. One approach was based on new features for classifying time series data, and was used for classification using SVMs. In addition, the measured vertical force time series of the wheel was represented in 2D and used for artificial neural-network training. The multi-sensor structure of the measurement system was clearly modelled by multi-instance learning and shift-invariant networks, and the neural network method improved the performance of the wheel in cases of wheel flat and out-of-roundness. Gong et al. [89] introduced the rotation centre as the centre of the circle, used the corresponding algorithm to fit the coordinates of the points obtained on the wheel pair, and calculated the wheel-diameter data in actual operation. Their program used eight sensors to detect a set of wheel pairs. Each wheel corresponded to four sensors, one of which was responsible for measuring the coordinates of the centre axis of rotation, while the other three were responsible for measuring the coordinate data of three points on the wheel. The image was processed using 3D modelling to obtain the required data. Static and dynamic tests showed that the method met the requirements of railway vehicles for wheel-pair detection in terms of accuracy and stability.

4.2 Small train components inspection

The train body contains a lot of parts, some of which (such as the angled door, bogie combination key, bogie block key (BBK), and so on) are very important for the safe operation of the train, and their absence leads to serious accidents. These parts are usually small, but large in number, and traditional inspections are time-consuming and labour-intensive. In response to this situation, scholars have conducted the following research.

In order to overcome the noise and safety problems of active steering bogie, Kim et al. [90] proposed an effective scheme for measuring the displacement between the wheels and rail of active steering bogie using a computer vision system. The wheel is detected using a template matching algorithm and projection. and the labeling algorithm is used to get rail candidates from a lot of edge components in the rail region. Finally, the wheel-rail relative displacement is calculated by the adaptive Hough transform algorithm in the world coordinate system. This method can effectively prevent accidents in the curve section and intersection section. Schlake et al. [91] proposed an image-acquisition system and
a machine-vision detection algorithm. The testing of the image-acquisition system was carried out at several railway-maintenance facilities. The images collected were used to develop several types of machine-vision algorithm to analyse the image of the railway vehicle chassis and evaluate the condition of certain structural components. The results showed that this method was able to improve the utilization of inspection and maintenance resources, and improve security and network efficiency. Yang et al. [92] proposed a train center plate bolts fault identification algorithm based on the train fault detecting system (TFDS). According to the physical characteristics of the train bogie image positioning, Gray mapping used in conjunction with the gradient mapping and transformation. First, segment the area where the center plate bolts are located in the image, and then use the adaptive threshold method to filter the irrelevant features. Finally, the Hough transform is used to extract the main features of the application line and identify the failure of the center plate bolts. This method has good detection speed, accuracy and reliability. Zhou et al. [93] proposed a visual inspection method for the angled plug door of a freight train, which was used to detect the missing handle of the angled plug door during the train’s operation. The method obtained the angled gate image of the freight train using the TFDS installed on the ground, extracted the photometric level-independent features through the gradient-coded histogram (GEH), and used an SVM to classify the features. The experimental results showed that the system achieved a fault-detection rate of 99.8%. In order to keep wheel sets from separating out of bogies, and to prevent terrible security incidents, Liu et al. [94] design a vision-based system to inspect the missing of bogie block key (BBK) automatically. First, a hierarchical detection framework consisting of fault Area extraction and object detection is proposed to divide image regions which contain the inspected component from the complex background. Subsequently, a component detector based on the sparse histograms of oriented gradients and support vector machine(SVM) is proposed to verify the candidate image regions to check whether the BBK is missing or not. Experiments show that the accuracy of this method for BBK fault detection is above 99.36%. In order to improve the efficiency of railway maintenance, Sun et al. [95] proposed an automatic fault recognition system (AFRS) based on TFDS. AFRS is a two-stage system: In the first stage, a coarse-to-fine scheme based on CNN model is adopted to detect the target regions of side frame keys (SFKs) and shaft bolts (SBs) simultaneously. In the second stage, another CNN model for multi-fault determination is established to identify four typical faults that occur in the target areas of the SFKs and SBs. The experimental results show that for SFKs and SBs, the positioning accuracy of this method is above 97.7%, the recognition rate of missing faults is 100%, and the recognition rate of loose faults is above 92.5%. Kishore et al. [96] proposed an intelligent monitoring method for train rolling stock. Based on the Chan Vese active contour model (CV-AC), a shape prior seeds Chan Vese active contour model (SP-CV-AC) is established for segmenting the train rolling stock, and the spatial distances are used to propel the initial contour towards final shape contour. This method enables the bogie image segmentation of a train with a speed of 30 km/h.

4.3 Sign and limit detection

In order to ensure that the safety state and load condition of a train are within the control range, it is often necessary to record the train identification and detect train overrun. This work is generally done manually by technicians, and is therefore relatively rigid and susceptible to the subjective judgement of the technicians, which can lead to misjudgement.

Balsi et al. [97] designed an automatic identification system for train tail signs. A high speed camera is aimed at the railway track to obtain a digital image data of the train tail signs and the data is stored in a buffer. CNN is used to identify the train tail signs and the results are transmitted to the station management system to assist train scheduling. Kumar et al. [98] designed an imaging system for generating landmarks from a rail vehicle including a camera mounted proximate an end of a rail vehicle, the camera transmitting imaging data indicating of images acquired, and a communication system for transmitting the Image data from the camera to at least one of a user onboard the rail vehicle, a storage device, and to a user located remote the rail vehicle. The system can be used to assist railway staff in diagnosing railway failures. Xie Fei et al. [99] proposed a method for detecting the over-limit of railway freight trains, combining binocular stereo-vision measurement technology and digital image-processing technology to achieve non-contact over-limit measurement of cargo train
contours. In this method, a special target was designed, and the parameters of the vertical plane of the track were transmitted to the detecting device through the image form, thereby realizing the non-contact indirect measurement of the vertical plane of the track centre, and solving the problem that the rail was blocked in the visual measurement. It is difficult to get the problem of the datum. Experimental results showed that the measurement error of the reference plane under the actual outdoor large-scale measurement condition was less than 10 mm, which was able to meet the requirements of the actual over-limit detection. Zhang et al. [99] proposed a method for freight train gauge exceeding detection based on 3D stereo vision measurement. To reach high measurement accuracy under large-scale situation, the factors which influence the 3D measurement error are analyzed in detail. Algorithm to accurately extract the laser stripe feature projected by the measurement system is described. With the obtained stripe features, the 3D structure of the freight train can be reconstructed with nonlinear optimization procedure. Specially designed targets are used to identify the global coordinate system for gauge-exceeding detection. The method has a detection error of less than 3.29 mm for freight train. Lisanti et al. [100] proposed a machine-vision system that used a multi-camera gantry to acquire data. Through the use of a variety of cameras and algorithms, the train identification, train-temperature monitoring and pantograph positioning and detection were able to be completed simultaneously. The results were analysed and presented to the driver using a multi-touch user interface. The data set recorded by the system on the test bench facilitated system evaluation.

5. Train system infrastructure

5.1 Tunnel crack detection

When a train is running, it spends a large proportion of its journey in tunnels. Tunnel cracks can occur due to factors such as material, load and erosion. If timely measures are not taken to stop these cracks, the consequences may be fatal. Therefore, it is necessary to periodically inspect the inner walls of tunnels in order to find cracks in good time. At present, the artificial detection of tunnel cracks is inefficient, and scholars have conducted a number of studies to improve efficiency and accuracy.

Qi et al. [101] proposed an algorithm to detect the crack based on the image processing. This algorithm simplifies the image preprocessing procedure, and use the block binaryzation to replace the traditional method of binaryzation. Then extraneous noisy pixels are removed from the image according to the differences of characteristics between noise and crack. Finally, the tunnel crack is automatically detected by comparing the crack width with a preset threshold. This method has good noise immunity. Soni et al. [102] described the application of Terrestrial Laser Scanning (TLS) and Close Range Photogrammetry (CRP) to the monitoring of a set of masonry arches during a major station refurbishment. Firstly, the ability to use TLS compared to traditional survey Methods was investigated. Inter-epoch comparison demonstrates a capability to detect change but highlights a requirement to understand the structure and data quality in making valid interpretations. Secondly, TLS and CRP techniques as monitoring tools for creating point cloud data on the same set of masonry arches were compared. These investigations generate significant volumes of data helping the engineers to make informed decisions. Han et al. [103] have developed an in-vehicle 3D laser-scanning system. The system consists of a 360-degree 3D laser scanner, a GPS, an IMU, a distance-measuring instrument and a number of cameras. On the carriage, synchronous data is used to collect data in real time. At present, 1.5 km of railway 3D measurement data acquisition and modelling has been completed, and the measurement time has been shortened. At the same time, the system is able to increase the control point to improve the point-cloud accuracy without a GPS signal (the optimal distance of the control point should be approximately 50 m), and the point-cloud result can also be used to construct the 3D scene model to improve the follow-up data support for path-transfer operations and management.

5.2 Railway crossings

Level crossings of railways and highways have become one of the most common places for railway injuries and fatal accidents due to the intrusion of foreign objects or abnormal fork conditions. Therefore, it is necessary to monitor abnormal situations at level crossings in real time to avoid the occurrence of tragedy.

Aminmansour et al. [104] proposed a novel method for detecting near-miss occurrences at railway level crossings from video data of trains. Firstly, the image was converted from the cabin
view to the bird’s eye view, and the railway centerline was extracted therefrom. Then track the railway from frame to frame and the railway is segmented from the frame. Finally, the Pedro Felzenszwalb detector is used to detect the vehicle in the image and the distance of the vehicle from the railway is calculated based on the homography between cabin view and bird’s eye view. This system is able to work in wide variety of conditions from daytime to nighttime, without having to change any parameter settings. Taştimur et al. [105] performed a switch and horizontal cross-detection using vision-based non-contact image-processing techniques. Switch detection was performed using ROI splitting, the Hough transform and a hierarchical SVM classifier. The proposed method was written in MATLAB 2014b. The average elapsed time of the proposed switch-detection method was 0.41 s, and the accuracy was 90.3%. Zhang et al. [106] proposed a Computer Vision (CV) algorithm to automatically detect trespassing near-misses based on surveillance video footage of railway-road grade crossings. The red signal in the video frame is used as a trigger for detecting moving objects, and a background model based on the video data was built. Finally, based on moving objects detected in the form of a set of moving pixels, the near-miss events are identified during that red signal. This method has good robustness and can operate normally under different lighting conditions.

5.3 Railway stations

Railway stations usually have a considerable amount of traffic, are difficult to manage, and are often the site of accidents. In order to improve the monitoring of station traffic density in real time and the detection of passenger behaviour abnormalities (overcrowding, falling to orbit, destroying infrastructure, and so on), scholars have conducted the following studies:

Marana et al. [107] proposed a method for automatically estimating human-flow density based on image textures of human flow. Their method used a grey-level dependence matrix for texture analysis of the acquired image to obtain the population-density feature vector, and used a self-organizing map neural network to identify the feature vector and estimate the human-flow density (divided into five levels). The method was able to automatically detect the pedestrian-flow density of the railway station, and the overall accuracy of the human-flow density detection reached 81.88%. Sacchi et al. [108] proposed a neural network-based video-based surveillance system for detecting vandalism in unmanned railway environments. Using a morphological noise filter and background update to preprocess the image acquired by the camera to obtain the binary image of the differences between The updated background and the current frame. The convolutional neural network is then used as a classifier to identify abnormal behavior of the person in the image. Through experiments, the detection accuracy of the system in the complex background environment can reach 84.21%. Oh et al. [109] proposed an image security technology for implementing a safety-monitoring system for train stations, and specifically for detecting passengers falling from the platform. The system was used in combination with multiple cameras to monitor the security of the entire track line and to determine in real time whether people or dangerous obstacles were in the preset surveillance area using image-processing techniques. Experiments showed that the system was able to robustly detect train status and objects, and provide video information and alarm information for local stations, central control rooms and trains. Lai et al. [110] developed a roadside machine vision MV system for automatically scanning a passing train and evaluating the aerodynamic efficiency of the loading mode, using advanced cameras to image each container or trailer as the train passed by. A digital video recording system was used to record passing trains, and the MV algorithm analyzed images to detect and identify loads and measure gaps. After each train was recorded, the video was processed and a histogram of the gap was generated to represent the loading mode of the train; using the resulting data, a scoring system was then proposed for comparing the actual configuration with the ideal configuration, so that the terminal manager was provided with feedback on train loading conditions. Combined with AEI’s automotive information, an index was developed based on the aerodynamic effects of IM loads and vehicle combinations to evaluate tank efficiency. The system demonstrated that at the macro level, a data collection and analysis system could be deployed to monitor system-wide IM traffic efficiency. At the micro level, it could provide feedback from specific trains at specific terminals to help managers create more efficient loading trains. Song et al. [112] proposed a vision-based method for estimating train position and railway monitored motion,
using a fuzzy classifier to determine train status. The proposed method used frame difference and background subtraction to estimate train motion and presence, respectively. These features were used as linguistic variables for the fuzzy classifier. Experimental results showed that the method was able to correctly estimate train position and motion. The method could therefore be used to estimate population density or to develop a safe railway-monitoring system. Delgado et al. [113] proposed a method for automatically detecting people jumping or descending from a train platform. The method detected moving objects and analysed their positions relative to the edge of the orbital bed. Train presence was studied by analysing motion vectors in the orbital bed area. Experimental evaluations were performed using video data sets recorded at local train stations, and the accuracy and recall rates were found to be 90%.

6. Problems and prospects

Compared with traditional detection methods, image-processing technology has the following advantages:

(i) The detection efficiency is high: some detection methods work effectively at speeds above 200 km/h.
(ii) There is no risk of subjective misjudgement, and tests are carried out strictly according to procedure.
(iii) The accuracy is stable. This type of inspection is non-contact detection. There is no contact friction during the inspection.
(iv) Costs are low; equipment can be reused.
(v) The scalability is strong. If the test results are not satisfactory, the adjustment algorithm can be optimized.
(vi) Test data is stored in a database for convenient review.
(vii) Since it is non-contact detection, there is no damage to sensors or detection targets.
(viii) Although visual inspection technology has many advantages, there are still many problems that need to be solved. A visual inspection system for railway components that can be put to practical use needs to have the characteristics of a short processing time, a low calculation amount and high reliability. Moreover, the method used in such a system must be sufficiently robust to interferences such as the installation-position offset of data-acquisition equipment, image noise caused by vehicle motion and image blur caused by excessive train speed. In addition, the visual inspection equipment is susceptible to light intensity, and detection accuracy is poor in the morning and evening hours; changes in the weather are also an important factor influencing detection. In cases of rain and snow, the visual inspection system usually cannot operate normally; the characteristics of the image data are obvious, and this is usually the cause of false detection. Whether the shadow problem can be solved depends on the reliability of the detection result; the detection object is usually contaminated with dust and oil stains, which changes the original colour and appearance, and may cause false detection.

(ix) The above problems limit the development of visual inspection technology, and should be the focus of future research in this area. Paying attention to accuracy and integrating multiple detection methods is a major development trend of future visual inspection technology. Reviewing the development of visual inspection technology for railways in recent years, early research was affected by computer performance. Early visual inspection methods were relatively simple in order to meet the requirements of algorithms with lower computing power. As computer CPUs have developed, detection methods have become increasingly complex, and multiple data preprocessing algorithms may be run at the same time to improve visual detection accuracy. Recently, due to the development of GPU technology, image-processing speeds have further increased. Under the premise of meeting the timeliness, the computer vision system needs to perform more precise processing on the image data to obtain more accurate detection results. In addition, multi-sensor information fusion of visual inspection technology is also the focus of future research. The image data acquired by the camera is susceptible to illumination, weather, and obstructing objects, resulting in unsatisfactory processing results. A visual inspection system that combines sensors such as IMU, laser radar, mechanical sensors, and temperature sensors will have more accurate data processing capabilities.

(x) The intelligence of visual inspection technology is also undergoing considerable develop-
The increasing interest in AI provides a new model for railway visual inspection technology. Enhancing the self-determination and early warning capability of visual inspection systems is of great significance to increasing the productivity of the railway industry and improving detection efficiency.

Conflict of interest statement. None declared.

References


57. Maire FD. Vision based anti-collision system for rail track maintenance vehicles.


