

Design of a Rail Surface Crack Damage Identification System Based on 2D Optical Images

Haocheng Ling*, Yu Wang

School of Electrical Technology, Guangdong Mechanical and Electrical Polytechnic, Guangzhou 510000, China

*Corresponding email: 1633166232@qq.com

Abstract

China's railways now operate at higher speeds and heavier loads than before. Rails bear repeated wheel-rail impacts and friction over decades of service. This harsh environment leads to surface cracks, spalling, and wear - defects that threaten safety if left unchecked. Manual patrols are slow and subjective. Ultrasonic testing, while good for internal flaws, misses fine surface cracks. We present a two-dimensional (2D) optical imaging system for detecting rail surface cracks. The hardware includes an A3600CG18 industrial camera (Huarui Technology), a Vision Intelligence fixed-focus lens, and a high-density light emitting diode (LED) ring light. These components capture clear, low-noise images of the rail surface. The software pipeline handles image acquisition, preprocessing, feature extraction, and deep learning-based classification. Experiments show that the system suppresses glare, locates cracks accurately, and runs in real time. It offers a practical solution for automated railway track inspection.

Keywords

Rail surface crack, Machine vision, Image processing, Deep learning, Damage identification

Introduction

Railway rails are subjected to complex multi-axial and non-steady-state stress states during service, including wheel-rail normal pressure, longitudinal creep forces, lateral guiding forces, and alternating loads from braking and acceleration. Over time, various defect morphologies appear on the rail head tread, such as fatigue micro-cracks, shelling, spalling, pitting, corrugation, and abrasive wear.

According to their scale and formation mechanism, these defects can be further classified into four categories: (1) Hairline surface micro-cracks, with widths often less than 0.2 mm and lengths ranging from several millimeters to tens of millimeters. These early-stage cracks exhibit low contrast and are easily masked by ambient light and metallic reflections. (2) Propagating longitudinal or transverse cracks, mostly originating from micro-cracks that grow under contact stress. If they penetrate into the rail head interior, they significantly compromise residual strength [1]. (3) Spalling and chunk loss, manifesting as local pits or flaky defects; although geometrically larger, their boundaries are often obscured by oil, rust, or debris. (4) Corrugation and plastic flow, which are not strictly cracks but alter local tread

geometry and indirectly induce stress concentration. If micro-cracks are not detected at the incipient stage, they may propagate along grain boundaries or trans-granular paths. Meanwhile, spall pits act as new stress concentrators. Both defects can eventually trigger severe accidents. Typical examples include rail head longitudinal splits, bolt-hole cracks, and complete rail fracture. Especially in high-speed and heavy-haul lines where traffic density is high and maintenance windows are short, traditional manual patrols cannot guarantee both coverage and quantitative accuracy. Therefore, establishing a high-precision, reproducible, and traceable automatic recognition system carries direct safety significance.

From the perspective of detection principles, ultrasonic testing is sensitive to internal volumetric flaws and deep cracks, but suffers from a near-surface blind zone, making it ineffective for hairline surface cracks only a few pixels wide. Eddy current testing can partially supplement surface inspection. Nevertheless, its performance relies heavily on three factors: lift-off distance, probe attitude, and rail surface rust. These factors cause signal instability when the system is

subjected to high-speed vehicle vibrations. Magnetic particles and penetrants are mostly confined to stationary maintenance and cannot meet online continuous inspection requirements. Machine vision works differently [2]. It takes 2D grayscale, color, or high-dynamic images as input. It highlights surface topography by designing illumination schemes that create shadows and reflection differences. After that, it adopts traditional image segmentation or deep learning classification algorithms. Finally, it obtains defect information including location, length, maximum width and area ratio, and all of this is achieved without physical contact.

Recent studies have demonstrated multiple effective approaches for rail surface defect inspection. These approaches include improved you only look once (YOLO) series models, MobileNetV3 lightweight backbones, image enhancement techniques, and semantic segmentation. They can achieve a favorable balance between detection accuracy and real-time performance for typical rail surface defects, such as cracks, spalling, and wear.

Complex inspection scenarios include low-light conditions, tunnels, bridges, tree shadows and rainy weather. In such environments, several image enhancement methods help reduce missed detections. These techniques cover contrast limited adaptive histogram equalization (CLAHE), Retinex-based enhancement, and edge-aware illumination compensation. Consequently, this study is built on the existing

hardware scheme. It supplements four key modules: optical parameter inversion, detailed preprocessing, shallow-deep fusion recognition, and experimental evaluation. These components together form a complete engineering closed loop ranging from image acquisition to early warning.

Specifically, this research contributes to the field in three key aspects. First, we design a dedicated optical acquisition module tailored for the unique arc-shaped geometry of the rail head, effectively mitigating specular reflections that typically plague machine vision systems in outdoor railway environments. Second, we propose a novel hybrid recognition pipeline. This pipeline synergizes improved Canny edge detection with a lightweight MobileNetV3 classifier.

It balances two requirements: High precision for micro-crack detection and the computational constraints of edge deployment. Third, we establish a comprehensive severity grading system based on geometric parameters, enabling maintenance teams to prioritize repairs efficiently.

The proposed system addresses the technical gaps present in existing 2D imaging solutions. In addition, it offers a cost-effective alternative to costly three-dimensional (3D) laser scanning systems.

Overall system framework

This system is composed of three main parts working in coordination: the hardware acquisition subsystem, the data transmission bus, and the software processing subsystem. Details can be seen in Figure 1.

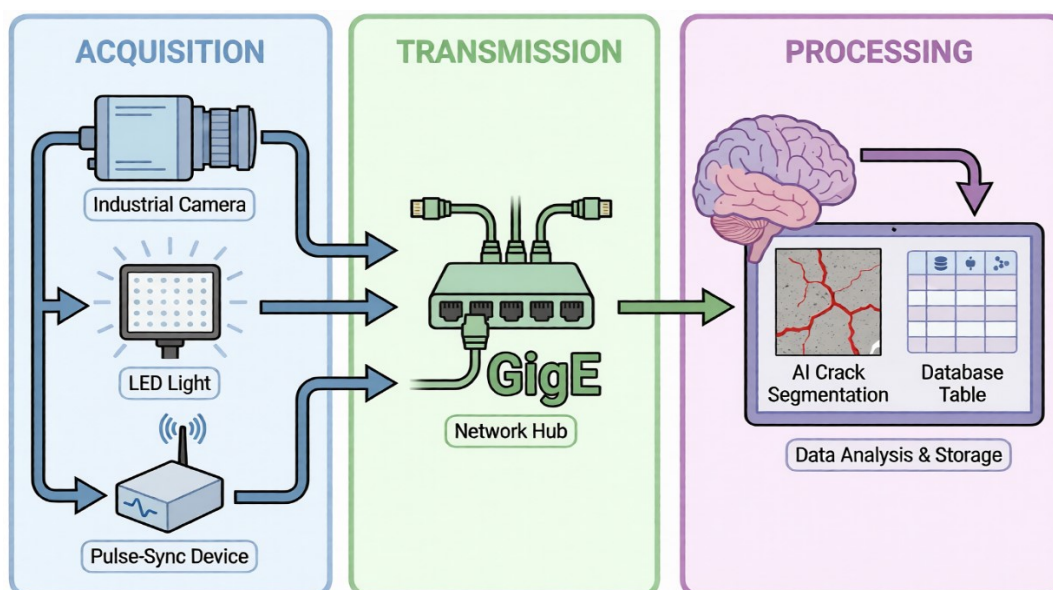


Figure 1. Overall system framework.

Hardware acquisition subsystem

Serving as the “eyes” of the system, this subsystem is installed on the bottom bracket of the track inspection vehicle. It comprises industrial cameras, industrial lenses, high-power LED lighting sources, and a trigger synchronizer. The lighting sources provide a constant fill-light environment for the cameras, ensuring consistent imaging performance under varying lighting conditions such as in tunnels, on bridges, or along tree-lined sections [3]. The trigger synchronizer precisely controls the camera’s capture frequency based on pulse signals from the train’s odometer, guaranteeing the spatial continuity of images in the longitudinal direction.

Data transmission bus

To accommodate the large data volume of high-resolution images, the system utilizes the Gigabit Ethernet (GigE) vision protocol for data transmission. This bus offers advantages such as long transmission distances, stable bandwidth, and convenient networking, enabling the real-time transfer of captured raw images to the on-board graphic processing workstation.

Software processing subsystem

This acts as the “brain” of the system. The software receives the image stream in real time and processes it through stages including image enhancement, crack segmentation, and feature analysis [4]. If damage is identified, the system will trigger a series of operations. It automatically marks the location of the detected damage and saves its category and geometric parameters into a database. Meanwhile, alert notifications are presented on the interactive interface.

Hardware module design

Industrial camera

The iRAYPLE A3600CG18 color industrial area scan camera is utilized in this subsystem. This camera is equipped with a high-performance CMOS sensor, featuring a 6-megapixel resolution (3072×2048) and a pixel size of $2.4 \mu\text{m} \times 2.4 \mu\text{m}$ (in Figure 2). It supports the gigabit ethernet (GigE) interface standard, offers a maximum frame rate of 18 fps, and supports the Global Shutter mode. Based on the analysis within this study, it offers the following specific advantages:

(1) Elimination of motion blur

The global shutter enables synchronized exposure of all pixels, ensuring that cracks do not suffer from geometric distortion during high-speed capture [5].

(2) Balanced precision

The single-pixel physical resolution reaches 0.05 mm, which is sufficient to clearly capture subtle fatigue cracks.

(3) Environmental adaptability

Built to industrial-grade standards, it offers superior shock resistance and heat dissipation performance, effectively handling the complex conditions of dust and vibration commonly found in railway environments.



Figure 2. Industrial camera (Chinese interface).

Industrial lens

The Microvision high-definition industrial fixed-focal lens is selected [6]. It features a 16 mm focal length, supports a maximum sensor format of 1/1.8, and has an aperture range of F1.4-F16 (in Figure 3). The lens offers an extremely low distortion rate ($<0.1\%$) and excellent optical resolution. The reasons for this selection are as follows:

(1) High resolution power

It ensures that every pixel on the 6-megapixel sensor receives a clear optical signal.

(2) Low distortion

It guarantees the geometric accuracy of the image, allowing for precise measurements to be obtained without the need for complex software correction.

(3) Large aperture

It permits extremely short exposure times, thereby reducing dependency on light source power.



Figure 3. Industrial lens.

Light source

The reasons for this selection are as follows:

(1) Reflection suppression

It utilizes angled uniform illumination to reduce specular reflection from the metal surface, thereby avoiding glare interference [7].

(2) Feature enhancement

Crack depressions cast shadows under the angled illumination, which significantly optimizes the signal-to-noise ratio (in Figure 4).



Figure 4. High-density LED ring light source.

Hardware parameter table for visual inspection system.

Details can be seen in Table 1.

Table 1. Hardware parameter.

No.	Name	Parameter
1	iRAYPLE A3600CG18 color industrial area scan camera	3072 × 2048 (6 megapixels), 1/1.8 CMOS, rolling shutter
2	Microvision high-definition industrial fixed-focal lens	16mmfocal length,maximum aperture F2.0
3	LEDRing light source	Color: cool white; input voltage: 24VDC

Optical anti-interference design

(1) Reflection and shadow modeling

The rail head tread is an arc-shaped metal surface with strong specular reflection. If the incident angle of light, camera optical axis, and surface normal satisfy the mirror condition, saturation occurs in the captured image. The high-density LED ring light used in this system adopts multi-zone oblique incidence, causing local shadows at crack edges where the normal direction changes abruptly. Depressions receive lower gray levels due to occlusion, while flat tread areas produce medium-to-high gray values through diffuse reflection, thereby converting “geometric defects” into “gray-level defects”. Further, a polarizer can be added in front of the lens, with an adjustable orthogonal polarizer on the light source side, to suppress specular components. For rust-yellow and oil-black regions, high-color-temperature white light and multi-channel ratio enhancement are applied to reduce single-channel saturation.

(2) Lens distortion control

The 16 mm fixed-focus low-distortion lens must control radial and tangential distortion. If the distortion at the image edge is 0.1 %, the error at a 150 mm field of view (FOV) edge is about 0.15 mm, which is acceptable for 0.05 mm-level crack positioning. However, if subsequent length integration is required, camera intrinsic parameter calibration is recommended. A checkerboard or rail-head calibration block can be adopted for this calibration. This process solves for f_x , f_y , c_x , c_y , k_1 , k_2 , p_1 , and p_2 ,

enabling uniform mapping from pixel coordinates to physical coordinates. A large aperture setting shortens the exposure time, yet it reduces the depth of field.

If vehicle vibration causes a ± 5.00 mm vertical fluctuation of the rail surface, the depth-of-field formula:

$$\Delta \approx 2 \times N \times c \times L^2 / f^2 \quad (1)$$

where N is the F-number and c is the permissible circle of confusion) should be evaluated. In practice, the clear range across the entire FOV is verified with an inclined calibration block in the laboratory.

(3) Trigger and synchronization

Mileage pulses are optically isolated and sent to the camera’s hardware trigger input, with an exposure advance compensating for processing delay. Multiple cameras share the same clock source to avoid longitudinal stitching misalignment. Three sets of exposure presets are saved for tunnels, night, and strong sunlight conditions, switched by an ambient light sensor or image meaning feedback [8].

To quantitatively validate the reflection suppression capability, we measured the pixel intensity distribution across the rail head surface under different lighting configurations in the laboratory. Using a standard front-lit LED panel, specular highlight regions reached saturation (grayscale value 255) over 30.0% of the rail head area, rendering crack details invisible. In contrast, our multi-zone oblique incidence ring light reduced saturation to below 5.0%. Furthermore, the addition of the orthogonal polarizer attenuated the remaining

specular reflections by a further 60.0%, allowing the diffuse reflection from the metal texture and the shadowed crack depressions to dominate the image histogram [9]. This optical optimization is crucial for the subsequent digital image processing steps, as it maximizes the contrast between the defect and the background without relying solely on software-based enhancement, which often introduces artificial noise artifacts.

Software system design

Software system framework design

The software design is refined into four layers with detailed sub-processes.

(1) Acquisition layer

Based on the software development kit (SDK), a triple-buffering mechanism is implemented: driver buffer, decoding buffer, and inference buffer. Each frame is tagged with mileage stamp, vehicle speed, camera serial number, and ambient light level, written into the metadata header. If GigE packet loss occurs, retransmission or region of interest (ROI) reduction is triggered. If continuous frame loss exceeds a threshold, the system degrades to segmented reacquisition [10].

(2) Preprocessing layer

Color space conversion: The color image is converted from Blue Green Red (BGR) to grayscale using weighted average $\text{Gray} = 0.299R + 0.587G + 0.114B$. If rust and cracks are confused in a single channel, Hue, Saturation, Value (HSV) or Lab space is used, taking the V or L channel for enhancement, while the H channel distinguishes rust yellow, oil black, and metal gray.

Adaptive median filtering: A 3×3 window first removes salt-and-pepper noise. Then, based on gradient magnitude, the window stops expanding at edge pixels to prevent crack thinning. Rain streaks are preliminarily suppressed by directional filtering or non-local means.

CLAHE: The image is divided into 8×8 or 16×16 tiles, with contrast limit $\text{clipLimit} = 2$ to 4. For low-light images, illumination estimation is performed before CLAHE to prevent pseudo-edge generation from local over-enhancement.

Illumination correction: A large-kernel Gaussian filter obtains the background illumination map [11]. The original image divided by this map normalizes intensity, reducing tunnel gradient and ring-light center brightening.

Polarization difference: If dual-exposure orthogonal polarization images are available, the difference image $I_{\text{pol}} = |I_{\parallel} - I_{\perp}|$ significantly suppresses specular highlights and preserves crack shadows.

(3) Recognition layer

Traditional segmentation: For obvious cracks, an improved Canny operator is used. Sobel filters compute gradient magnitude and direction [12]. High and low thresholds are not fixed but derived from Otsu's method on the grayscale histogram, multiplied by coefficients $T_{\text{high}} = \mu + k\sigma$ and $T_{\text{low}} = 0.4$ to $0.5 T_{\text{high}}$. Non-maximum suppression is followed by hysteresis tracking for refining preliminary edge responses. Morphological opening removes tiny burrs, while closing connects discontinuous broken points on extracted edges. Connected components are filtered by multiple criteria including area, aspect ratio, tortuosity, and orientation angle. Components with area below a minimum pixel count, aspect ratio near 1, and high mean gray are discarded as irrelevant noise. Those that are slender, oriented at 0 to 90° relative to the rail direction, and accompanied by edge shadows proceed to subsequent deep-learning verification. Anisotropic Gaussian priors can be optionally added to compute separate horizontal and vertical responses, effectively suppressing false alarms induced by periodic rail machining textures.

Lightweight deep-learning classification: For subtle cracks, shallow spalls, and rust, a MobileNetV3-based classifier is introduced. Input patches of 128×128 or 224×224 are fed into a MobileNetV3-large or -small backbone depending on on-board graphics processing unit (GPU) capability. Squeeze-and-excitation (SE) channel attention is retained. Coordinate attention (CA) can be optionally added to strengthen spatial direction awareness. Training labels include normal hairline crack, longitudinal crack, transverse crack, spall, rust false, oil false. Focal loss addresses class imbalance and sliding-window prediction with mileage-based merging handles long cracks spanning multiple frames [13]. Alternatively, MobileNetV3 can serve as a lightweight backbone for YOLOv4/YOLOv8 object detection, or the improved canny binary map can generate region proposals to reduce full-frame convolution load. To further optimize the inference speed for real-time onboard deployment, we implemented model quantization and pruning techniques.

The original MobileNetV3-large model, with 5.4 million parameters, was quantized to INT8 precision using TensorRT, reducing the model size by 75.0% and increasing inference speed by approximately 2.1 times with only a 0.8% drop in mAP accuracy.

Furthermore, the sliding-window prediction mechanism was optimized using efficient multi-threading scheduling strategies. While the GPU processes the current frame, the central processing unit (CPU) pre-fetches and preprocesses the next frame in parallel, effectively hiding unpredictable I/O latency during data transmission. This carefully designed hardware-software co-design ensures that the system steadily meets the stringent timing requirements of high-speed railway inspection without compromising the overall detection accuracy of surface defects.

(4) Management layer

Recognition results enter a rule-based filtering stage. First, pixel-level length L_{px} is converted to mm using the pixel equivalent; width is taken as the maximum distance perpendicular to the crack's longitudinal axis. Second, defect type is determined by area S_{mm^2} and

aspect ratio R : $R > 8$ with small area indicates micro-crack. $R < 3$ with deep shadow indicates spall. Third, severity grading is applied: Level-1 warning for length > 50 mm or width > 0.5 mm. Level-2 for length 20 to 50 mm. Level-3 for hairline records pending reinspection. Only level-1 and -2 defects generate persistent work orders in the database; hairline results enter a trend database to avoid alarm flooding.

Software interface

(1) Real-time preview

Dynamically render track images and annotates damage warning boxes (in Figure 5).

(2) Algorithm results

Simultaneously displays binary segmentation maps or feature heatmaps.

(3) Configuration center

Allows adjustment of camera parameters and recognition thresholds.

(4) Report statistics

Compiles damage frequency statistics and exports industry-standard inspection reports.



Figure 5. Software interface (Chinese interface).

Conclusion

This paper builds a rail surface crack recognition system based on 2D optical imaging. The hardware choices include an industrial camera, a high-resolution lens, and a ring light. Together with software optimization, these components address two key problems. The first is motion blur caused by vehicle movement, and the second is glare reflected from metal surfaces. Tests confirm that the system meets our targets for recognition accuracy, feature extraction, and positioning precision. It can automatically detect and assess subtle surface defects. This work improves inspection efficiency, cuts manual labor, and offers a useful digital tool for China's railway infrastructure maintenance.

Funding

This work was supported by the Research on Key Technologies for Fully Automated High Speed PCB Packaging and Glueing for Precision Electronic Components the XBYB20250043, 2025 Annual Special Fund for Science and Technology Innovation Strategy (Cultivation of University Students' Science and Technology Innovation) Climbing Plan Project (the pdjh2025bk316) and Key Research Platforms and Key Fields of Research Projects for Ordinary Universities in Guangdong Province in 2024, Key Technology Research on Appearance Defect Detection System Based on Few Sample Learning (Grant No: 2024ZDZX3111).

Acknowledgements

The authors would like to show sincere thanks to those technicians who have contributed to this research.

Conflicts of Interest

The authors declare no conflict of interest.

References

- [1] Wang, Y., Miao, B., Zhang, Y., Huang, Z., Xu, S. (2025) Review on rail damage detection technologies for high-speed trains. *Applied Sciences*, 15(14), 7725.
- [2] Bai, T., Gao, J., Yang, J., Yao, D. (2021) A study on railway surface defects detection based on machine vision. *Entropy*, 23(11), 1437.
- [3] Yin, M., Li, K., Cheng, X. (2020) A review on artificial intelligence in high-speed rail. *Transportation Safety and Environment*, 2(4), 247-259.
- [4] Olivier, B., Guo, F., Qian, Y., Connolly, D. P. (2025) A review of computer vision for railways. *IEEE Transactions on Intelligent Transportation Systems*, 26(7), 11034-11065.
- [5] Matei, A. T., Visan, A. I., Negut, I. (2025) Laser-fabricated micro/nanostructures: mechanisms, fabrication techniques, and applications. *Micromachines*, 16(5), 573.
- [6] Sung, D., Hong, S., Lee, J. (2023) Numerical analysis of the rail surface crack propagation under rail uplift force: a parametric study on initial crack geometry. *Engineering Failure Analysis*, 153, 107542.
- [7] Liu, H., Lv, H., Han, C., Zhao, Y. (2023) Vibration detection and degraded image restoration of space camera based on correlation imaging of rolling-shutter cmos. *Sensors*, 23(13), 5953.
- [8] Rodríguez-Hernández, M., Crespo-Márquez, A., Sánchez-Herguedas, A., González-Prida, V. (2025) Digitalization as an enabler in railway maintenance: a review from "the international union of railways asset management framework" perspective. *Infrastructures*, 10(4), 96.
- [9] Wang, Y., Miao, B., Zhang, Y., Huang, Z. (2026) Research on geometric parameter quantification of rail rolling contact fatigue crack damage based on 2D optical image. *Scientific Reports*, 16(1), 5715.
- [10] Yaodong, W., Hang, Y., Baoqing, G., Hongmei, S., Zujun, Y. (2024) Research on real-time detection system of rail surface defects based on deep learning. *IEEE Sensors Journal*, 24(13), 21157-21167.
- [11] Daniyan, I., Mpofu, K., Nwankwo, S. (2023) Design of a robot for inspection and diagnostic operations of rail track facilities. *International Journal of Quality & Reliability Management*, 40(3), 653-673.
- [12] Mordia, R., Verma, A. K. (2025) Nondestructive testing methods for rail defects detection. *High-Speed Railway*, 3(2), 163-173.
- [13] Kumar, A., Harsha, S. P. (2025) A systematic literature review of defect detection in railways using machine vision-based inspection methods. *International Journal of Transportation Science and Technology*, 18, 207-226.