

RailGuard: Vision-Based Railway Track Inspection and Fault Detection Using a Pole-Mounted Camera

P. Anlet Pamila Suhi¹, Harsha K², Nithyashree T S², & Rakshitha Reddy V²

¹Assistant Professor, Department of Information Science and Engineering
Sri Sairam College of Engineering, Bengaluru

²Final-year B.E(Information Science and Engineering), Sri Sairam College of Engineering, Bengaluru

DOI: doi.org/10.34293/iejcsa.v4i4.162

Abstract - Railway tracks require frequent and reliable inspection because geometric irregularities, rail-surface defects, fastener failures, ballast deterioration, and objects on or near the track can affect safety, ride quality, and maintenance planning. Conventional manual patrols and specialized track-recording vehicles provide valuable measurements, but they can be labour-intensive, costly, and difficult to scale for frequent monitoring. This paper reviews contactless railway monitoring approaches based on computer vision, LiDAR, GNSS/RTK, inertial sensing, machine learning, ground-penetrating radar (GPR), and InSAR, and uses the findings to motivate an indigenous Integrated Track Monitoring System (ITMS). The proposed direction combines stereo or line-scan cameras, LiDAR, GNSS/RTK, IMU, an encoder, onboard edge computing, SSD storage, and real-time alerts. A practical system should not only detect a defect but also estimate its severity, associate it with a reliable track position, retain evidence, and communicate an actionable maintenance alert. The review highlights calibration, weather and lighting variation, vibration, limited labelled data, computing constraints, and the integration of detection with localization as key engineering challenges.

Keywords: Integrated Track Monitoring System, contactless track monitoring, railway track geometry, computer vision, LiDAR, GNSS/RTK, IMU, deep learning, YOLO, edge computing, sensor fusion

INTRODUCTION

A. Background

Railway infrastructure is a long-life safety asset, and its condition changes continuously with traffic loading, weather, ageing, and maintenance activity. Small changes in gauge, alignment, cross-level, twist, rail-head condition, fasteners, ballast, or surrounding objects may develop into larger operational problems if they are not detected in time. Track inspection is therefore not simply a periodic measurement exercise; it is a continuous information problem in which observations must be converted into reliable maintenance decisions.

B. Objective of a Contactless ITMS

A contactless ITMS can use cameras and ranging sensors to observe the track without mechanical contact with the rail. The engineering objective is broader than defect detection. A useful system should acquire synchronized sensor data, estimate geometry,

identify visible defects, determine the geographic position of an event, reject nuisance detections, preserve evidence, and generate an alert quickly enough to support maintenance action.

LITERATURE REVIEW

A. Existing Railway Track Monitoring Technologies

Traditional track geometry systems combine optical or laser measurements with inertial navigation and wheel or encoder information. Escalona described a compact Track Geometry Measurement System using an IMU, two cameras, and an encoder to estimate alignment, vertical profile, cross-level, gauge, twist, and rail-head profile (Escalona, 2020). A related multibody-kinematics formulation mapped sensor observations to geometry parameters and reported agreement with an alternative method on a scaled track (Escalona et al., 2021). These studies demonstrate that a compact sensing arrangement can recover several important parameters without direct mechanical contact.

Peng et al. (2020) used four cameras, an inclinometer, an IMU, and a laser, with Kalman filtering and point-cloud registration to support geometry and surface-defect measurement. This illustrates how complementary sensors can share a coordinate system. Instead of asking a single camera to solve every problem, the system can allow each sensor to contribute the information it measures best.

Position estimation is another essential component. GNSS/RTK can provide an absolute or map-referenced location, while an IMU supplies high-rate motion information but accumulates drift. Li et al. (2024a) combined RTK GNSS, a low-cost MEMS IMU, an odometer, and lasers, using zero-velocity updates, smoothing, nonholonomic constraints, and calibration to support irregularity inspection.

LiDAR adds metric three-dimensional information. Cameras are strong at recognizing appearance and semantic patterns, while LiDAR can provide shape and depth when surface texture is weak. Camera-LiDAR fusion has been investigated for railway-object detection (Wang et al., 2021) and real-time track detection and three-dimensional fitting (Tang et al., 2024). GPR and InSAR can reveal subsurface or deformation-related information that ordinary surface cameras cannot observe (Koohmishi et al., 2024). These technologies establish a roadmap for a modular ITMS.

B. AI and Deep Learning for Railway Defect Detection

Deep learning has become an important tool for visual railway inspection because defects and components can appear at different scales, orientations, and lighting conditions. Improved YOLO-family detectors have been applied to fastener defects, including YOLOv5-based detection (Improved YOLOv5 Study, 2022) and improved YOLOv8 detection for heavy-haul railway applications (Li et al., 2024b). Such detectors are attractive for an onboard system because they can process images rapidly and return both a defect class and a confidence score.

However, high detector accuracy on a fixed dataset does not automatically translate into safe field operation. A train-borne system experiences vibration, blur, shadows, rain, dust, changing ballast colour, partial occlusion, different fastener designs, and changes in camera height. The AI layer should therefore be treated as one part of a larger decision pipeline using confidence thresholds, temporal persistence, spatial consistency, and sensor agreement.

The availability of labelled data is another practical issue. Manual annotation of rail defects is expensive and often produces datasets concentrated around a limited set of track types and environmental conditions. Haroon et al. (2024) investigated supervised and self-supervised learning for rail-defect detection, indicating a useful direction for reducing dependence on fully labelled datasets. Locally collected images can gradually become a valuable dataset as confirmed defects and reviewed false alarms are fed back into later model versions.

Classical image-processing methods also remain useful when computing resources are limited. Embedded contour-feature processing has been used for rail-surface deformation detection (Embedded Contour-Feature Processing Study, n.d.). A hybrid strategy is therefore reasonable: lightweight screening can operate continuously at the edge, while more demanding analysis can be performed selectively or offline.

C. Contactless Track Geometry Measurement

The strongest common theme in the reviewed work is the combination of vision and inertial sensing. Camera-IMU-encoder systems can estimate multiple geometry parameters from a compact installation (Escalona, 2020; Escalona et al., 2021). Antognoli et al. (2021) evaluated stereo and thermal cameras for commercial vehicles and positioned contactless sensing as an alternative to more sophisticated diagnostic trains. Peng et al. (2020) further demonstrated that geometry and surface-defect observations can be handled within one multisensor arrangement.

Measurement accuracy depends heavily on calibration. Camera pose, lens distortion, sensor mounting height, vehicle vibration, rail visibility, encoder timing, and changes in vehicle-body motion can all influence the final result. IMU bias also accumulates when motion is integrated. These effects make it important to retain calibration parameters and sensor-health information with each measurement.

For the proposed ITMS, geometry processing can be organized as a sequence: image acquisition and synchronization, rail-region detection, geometric feature extraction, motion compensation using IMU information, distance progression using an encoder, and correction or referencing using GNSS/RTK. The result can be stored together with confidence or uncertainty indicators, allowing each stage to be tested independently.

D. Real-Time and Edge-Based Railway Monitoring

Real-time monitoring becomes useful when the system can process observations while the vehicle is moving. The reviewed camera-LiDAR track-fitting work (Tang et al., 2024) and embedded deformation-detection work (Embedded Contour-Feature Processing Study, n.d.) support local processing as a practical direction. Edge computing avoids making the complete system dependent on a continuous high-bandwidth cloud connection, which is useful on routes where network availability is inconsistent.

An onboard computer can perform frame selection, image preprocessing, defect inference, temporal confirmation, sensor fusion, and alert generation. SSD storage can retain synchronized images, point-cloud snippets, position data, and system logs. Instead of transmitting every raw frame, the system can send confirmed events and compact metadata to a remote dashboard.

An alert should communicate what was detected, where it was detected, when it was observed, how confident the system is, and what evidence is available. A maintenance engineer should be able to review the associated image or sensor record rather than receiving an unexplained numerical warning.

E. Sensor Fusion and Location-Based Fault Detection

Sensor fusion improves robustness because different sensors fail in different ways. Camera data provides appearance and semantic information; LiDAR provides three-dimensional structure; IMU provides high-rate motion information; GNSS/RTK provides geographic reference; an encoder provides distance progression; and GPR can reveal subsurface conditions.

The RTK-GNSS/MEMS-IMU/odometer/laser architecture described by Li et al. (2024a) shows how constraints and smoothing can make an economical navigation stack useful for irregularity inspection. Camera-LiDAR fusion studies (Tang et al., 2024; Wang et al., 2021) demonstrate the complementary value of semantic and metric sensing. On-board accelerometer research, including the systematic review by Bravo et al. (2025), provides another route for estimating track condition from vehicle response, although acceleration alone cannot reliably recover every geometry parameter (Antognoli et al., 2021).

A location-based fault record is valuable for maintenance. If a detected fastener problem, surface defect, or geometry irregularity is associated with a reliable chainage or GNSS position, repeated runs can be compared over time. This creates the possibility of moving from isolated alerts toward condition history and trend analysis.

F. Comparative Literature Survey

The studies reviewed in this work cover different parts of the railway-monitoring problem. Table I summarizes their sensing methods, learning or processing approaches, major findings, and limitations. The comparison indicates that no single paper provides a complete low-cost train-borne ITMS covering geometry, visible defects, localization, edge

inference, evidence storage, and alerts. This integration gap is the main motivation for the proposed indigenous architecture.

Table 1 Summary of Reviewed Studies

Study	Application	Method / Sensors	Finding	Gap
Escalona (2020)	Contactless TGMS	IMU, 2 cameras, encoder	Multi-parameter geometry	Calibration; rail visibility
Peng et al. (2020)	Geometry + defects	4 cameras, IMU, laser; Kalman/ICP	Multi-parameter measurement	Field robustness
Escalona et al. (2021)	Vision-inertial geometry	Multibody kinematics	Scaled-track agreement	Field generalization
Antognoli et al. (2021)	Wheel-to-rail sensing	Stereo/thermal cameras	Candidate contactless sensing	Alignment remains difficult
Wang et al. (2021)	Railway objects	Camera + LiDAR	Semantic + 3-D fusion	Calibration/compute
Improved YOLOv5 Study (2022)	Fastener defects	YOLOv5; images	Small-defect detection	Dataset shift
Li et al. (2024a)	Irregularity inspection	RTK GNSS, MEMS IMU, odometer, lasers	Low-cost precision approach	Train-scale validation
Li et al. (2024b)	Fastener defects	Improved YOLOv8; images	Automated detection	Weather/occlusion
Haroon et al. (2024)	Rail defects	Supervised + self-supervised	Less labelling dependence	Operational validation
Koohmishi et al. (2024)	Track health	GPR, InSAR, ML	Surface + subsurface view	Not full moving ITMS
Tang et al. (2024)	Track + 3-D fitting	Camera + LiDAR	Real-time representation	Weather/speed tests
Bravo et al. (2025)	Track condition	On-board accelerometers	Vehicle-response monitoring	Indirect measurement

RESEARCH GAP

Eight practical gaps emerge from the reviewed literature. First, high-end optical, inertial, and navigation systems can be expensive, motivating MEMS and commodity-camera alternatives (Antognoli et al., 2021; Li et al., 2024a). Second, many studies focus on one sensing function rather than a complete inspection chain. Third, real-time edge processing is less mature than offline analysis. Fourth, lighting, weather, occlusion, vibration, and changing track assets can degrade image models.

Fifth, small defects such as fastener damage are difficult to detect reliably, especially when components are partially hidden (Improved YOLOv5 Study, 2022; Li et al., 2024b). Sixth, labelled railway datasets are costly and frequently narrow; self-supervised learning is promising but still requires operational validation (Haroon et al., 2024). Seventh, detection and accurate GNSS/RTK localization are not always designed together. Eighth, GPR/InSAR,

surface vision, inertial response, and LiDAR are rarely unified in a cost-optimized indigenous platform (Bravo et al., 2025; Koohmishi et al., 2024; Tang et al., 2024; Wang et al., 2021).

Evaluation should therefore be system-oriented. Useful measures include geometry error against a reference system, defect precision and recall, localization error, processing latency, storage rate, false-alert frequency, and robustness across speed, lighting, weather, and track conditions.

PROBLEM STATEMENT

Manual patrols remain useful because experienced personnel can recognize contextual conditions that are difficult for a sensor alone to interpret. However, manual inspection is labour-intensive and does not naturally produce a dense digital record. High-end track-recording cars can measure geometry with high precision, but their deployment and operating cost can restrict how frequently they are used. These limitations have encouraged research into compact sensing units mounted on service vehicles so that inspection can take place during normal movement.

PROPOSED SYSTEM

A. Relevance to the Proposed ITMS

The proposed ITMS addresses the identified gaps through a train-borne, contactless, modular architecture. Stereo or line-scan cameras can support visual inspection and geometry estimation; LiDAR can provide three-dimensional structure and assist track fitting; GNSS/RTK and an encoder can support event localization; an IMU can provide high-rate motion compensation; an edge computer can perform immediate inference; SSD storage can preserve synchronized evidence; and selective cloud or remote transmission can support maintenance dashboards.

The indigenous aspect is not simply the use of inexpensive sensors. It is the integration of commercially obtainable components into a system that can be calibrated, maintained, upgraded, and adapted to local railway conditions. A modular design also allows future sensors or improved AI models to be added without replacing the complete platform.

The reviewed literature (Table I) supports the feasibility of the individual sensing modalities, but end-to-end performance across local rolling stock, track classes, operating speeds, and weather conditions remains a project hypothesis that must be demonstrated experimentally. Field calibration, diverse data collection, uncertainty reporting, independent validation, and repeat-run testing should therefore be core development activities.

B. Expected Contribution and Practical Significance

The expected contribution of the proposed work is an indigenous, modular approach that brings several established sensing ideas into one practical monitoring chain. The emphasis is on integration: a defect should be detected, measured where appropriate,

localized, supported by evidence, and communicated in a form that a maintenance user can understand.

From an engineering perspective, the use of commodity or comparatively low-cost components can make experimentation and deployment more accessible. The architecture also creates a path for future upgrades. Improved detection models, additional sensors, or better localization algorithms can be incorporated as modules without redesigning the complete system.

The practical value of such a platform is the creation of a repeatable digital inspection history. Over time, this history can support condition-based maintenance and help teams focus attention on locations that show repeated anomalies. The proposed system should therefore be viewed as an inspection and information platform, not merely as an AI defect detector.

SYSTEM ARCHITECTURE

A. Proposed Indigenous ITMS Architecture

The proposed architecture is intentionally modular. At the sensing layer, stereo or line-scan cameras provide visual information, LiDAR supplies three-dimensional structure, GNSS/RTK provides absolute location, the IMU provides motion information, and the encoder provides distance progression. An optional GPR or additional condition sensor can be added in a later version when the project scope and vehicle installation permit it.

The edge-computing layer acts as the local decision centre. It receives synchronized streams, performs preprocessing, runs the selected AI models, estimates geometry, and generates events. Keeping the first level of inference onboard has two practical advantages. It reduces the amount of data that must be transmitted and allows the system to continue operating when a network connection is weak or temporarily unavailable.

The storage layer should use SSD storage because inspection produces a large volume of images and sensor measurements. Continuous raw storage may not be necessary for every channel; the system can use configurable compression and event-based retention. For example, confirmed defects can retain full-resolution evidence, while normal sections can be stored using a lower-rate summary. This approach can extend storage capacity without losing the information needed for investigation.

The communication and dashboard layer can receive selected events rather than every raw sensor sample. Each event can be displayed using a simple maintenance-oriented record: location, defect class, confidence, date and time, severity category, and evidence. Repeated observations of the same location can then be compared, allowing engineers to see whether a condition is stable, improving, or deteriorating.

An important design principle is graceful degradation. If RTK becomes unavailable, the system should continue recording using IMU and encoder information and mark the position uncertainty accordingly. If the LiDAR stream fails, camera-based inspection can continue. If the AI model produces uncertain results, the system should store the evidence

as a review candidate rather than silently discarding it. This makes the platform more practical for real railway operation.

METHODOLOGY / WORKFLOW

The proposed development can be organized as a sequence of sensing, synchronization, preprocessing, inference, fusion, localization, storage, and reporting stages. An overview of the seven-stage pipeline, from multi-sensor data acquisition to track-health and maintenance decision, is shown in Fig. 1. During a monitoring run, the camera, LiDAR, IMU, GNSS/RTK receiver, and encoder should share a common time reference. Synchronization is important because even a small timing mismatch can produce an apparent spatial error when the vehicle is moving quickly.

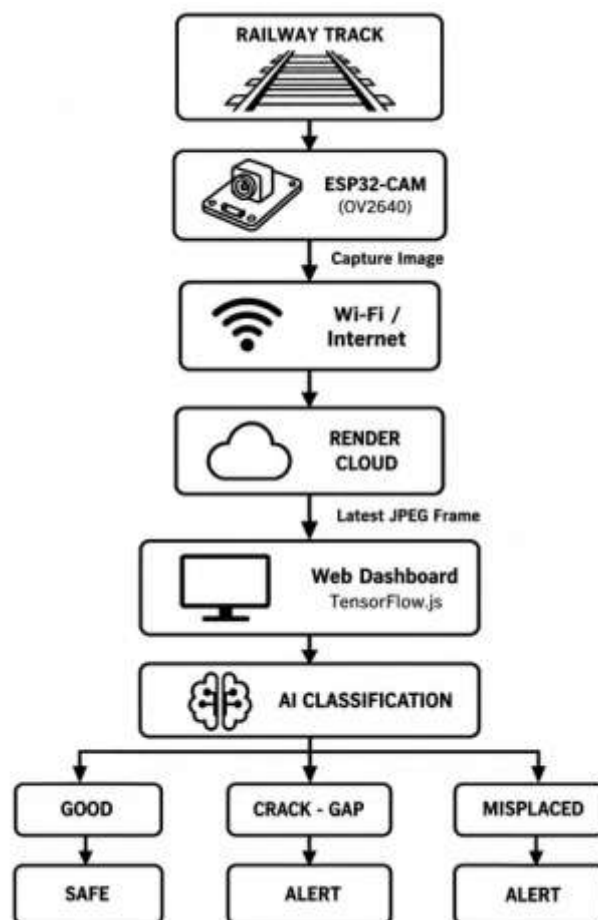


Figure 1 Seven-Stage Methodology Pipeline of the Proposed RAIL-GUARD System

Data acquisition. The first stage is continuous data acquisition. Cameras capture the rail corridor and nearby components, LiDAR records depth information, the IMU records acceleration and angular motion, GNSS/RTK provides position, and the encoder estimates distance travelled. The acquisition layer should record sensor health indicators such as dropped frames, GNSS quality, IMU saturation, and storage status. These indicators help distinguish a real track problem from a problem caused by the monitoring equipment itself.

Preprocessing. The second stage is preprocessing. Images can be corrected for lens distortion and normalized for exposure where necessary. LiDAR points can be filtered to remove obvious outliers. IMU signals can be filtered for noise, while GNSS observations can be checked for quality before they are used as absolute constraints. The aim is not to over-process the data, but to produce stable inputs for the later geometry and AI modules.

Detection and measurement. The third stage is detection and measurement. A visual model can identify rail components, fasteners, surface anomalies, obstacles, and other visible conditions. In parallel, geometry algorithms can estimate gauge, alignment, cross-level, twist, and profile-related parameters. LiDAR can assist with three-dimensional structure and track fitting, while the IMU helps compensate for vehicle motion. The results are then fused and assigned a position.

Decision support. The final stage is decision support. A candidate event can be accepted when its confidence, persistence, spatial consistency, and sensor agreement meet defined thresholds. The system should save an evidence package containing the event type, timestamp, position or chainage, confidence, relevant image, and selected sensor measurements. This creates a traceable record that can be reviewed by maintenance personnel.

A. Operational Workflow

A typical inspection run begins with a pre-run equipment check. Camera lenses, LiDAR windows, sensor mounts, cables, storage capacity, GNSS reception, and system time are checked before the vehicle starts. Calibration parameters are loaded and the system records the vehicle and configuration information for that run. A short initialization period can be used to verify that all streams are synchronized.

During movement, the system continuously acquires sensor data. The encoder provides distance progression, while GNSS/RTK and IMU information establish motion and position. The camera pipeline identifies candidate regions and defects, and the LiDAR pipeline supplies depth or track geometry information. The fusion stage combines these observations and attaches a location to each accepted event.

After a run, the system produces a compact inspection report and retains the evidence needed for review. False alarms and confirmed defects can be labelled by a human reviewer. These reviewed records become useful training material for future model updates. In this way, the system can improve gradually instead of requiring a perfect dataset before the first field deployment.

The workflow also supports repeat inspection. When the same route is monitored again, the system can compare the new measurements with earlier records. This temporal view is more informative than a single snapshot because maintenance teams can prioritize locations showing persistent or worsening abnormalities.

IMPLEMENTATION AND EVALUATION

A. Implementation and Validation Plan

The development should proceed incrementally rather than attempting a full railway-ready system in the first build. A laboratory prototype can first validate sensor synchronization, image acquisition, LiDAR capture, GNSS/RTK logging, IMU integration, and encoder distance measurement. A controlled test section can then be used to validate geometry calculations and event localization.

The next stage should focus on data collection under varied conditions. Images should include changes in daylight, shadows, weather, track cleanliness, component appearance, and vehicle speed. The dataset should also include normal examples because a robust inspection system must learn what healthy track components look like. Where possible, the same locations should be captured repeatedly to study temporal consistency.

AI performance should be measured using precision, recall, F1-score, and false-alert rate rather than accuracy alone. Geometry performance should be compared with an accepted reference measurement system. Localization should be evaluated in metres or chainage error. Real-time performance should include processing latency and the percentage of data that can be handled at the target vehicle speed.

System-level validation should finally test failure cases. Examples include poor GNSS reception, vibration, camera blur, partial occlusion, low illumination, rain, temporary sensor dropouts, and limited network availability. A useful prototype is one that reports uncertainty and continues operating safely when one data source becomes unreliable, rather than one that performs well only under ideal conditions.

B. Dashboard Interface

Figs. 2 to 4 show the web dashboard. Fig. 2 shows track alignment, live sensor monitoring, and the alert center; Fig. 3 shows a track alignment fault with the live ESP32-CAM feed and inspection location status; and Fig. 4 shows track alignment fault detection and AI inspection status.

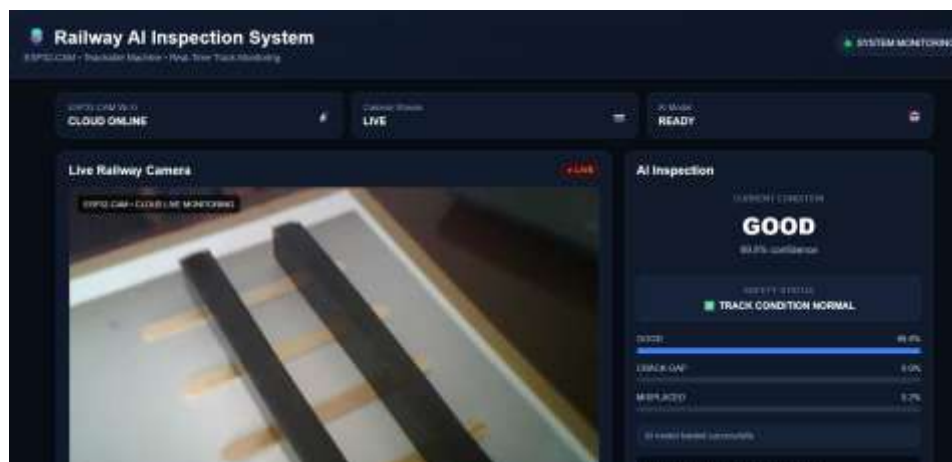


Figure 2 Web Dashboard: Track Alignment, Live Sensor Monitoring, and Alert Center

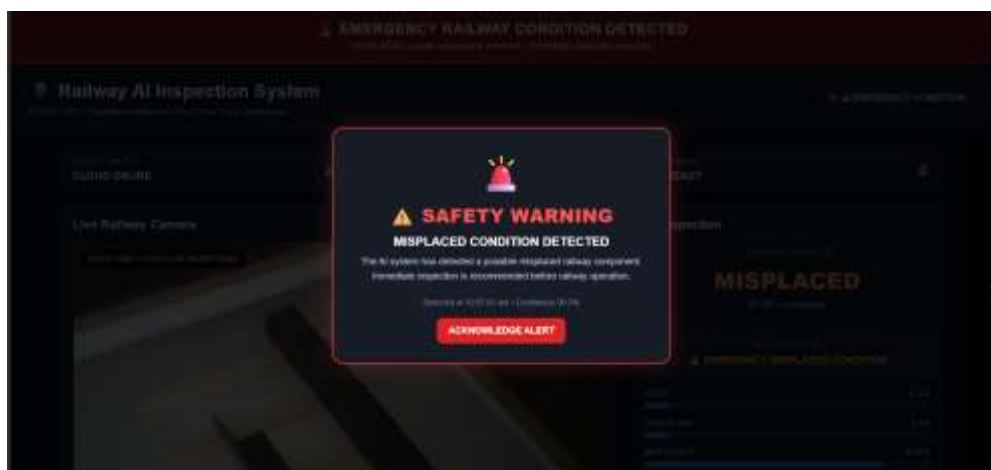


Figure 3 Web Dashboard: Track Alignment Fault, Live ESP32-CAM Feed, and Inspection Location Status

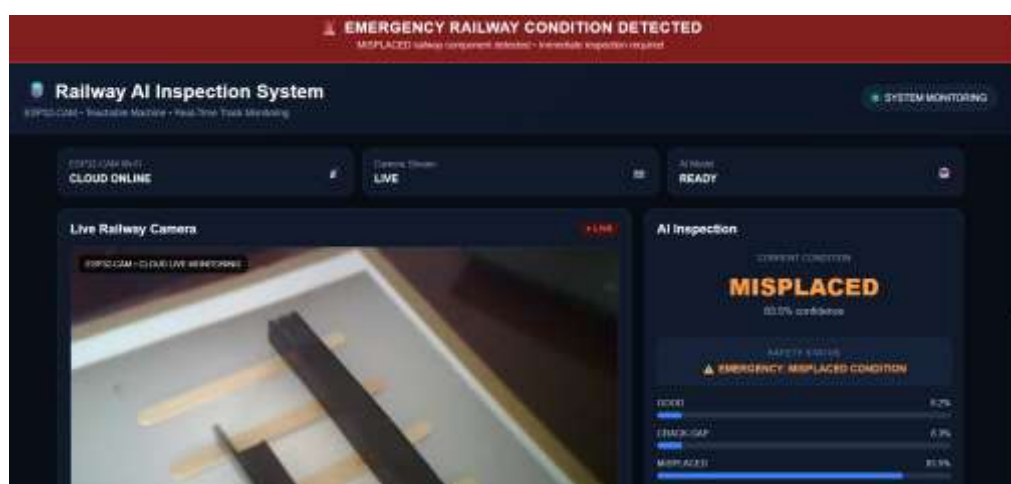


Figure 4 Web Dashboard: Track Alignment Fault Detection and AI Inspection Status

CONCLUSION

Railway monitoring is gradually moving from periodic inspection toward continuous, contactless, and data-driven observation. Camera-IMU geometry measurement provides a foundation for multi-parameter inspection (Escalona, 2020; Escalona et al., 2021), while LiDAR fusion, RTK localization, accelerometers, GPR/InSAR, and deep learning extend the range of conditions that can be observed (Bravo et al., 2025; Haroon et al., 2024; Improved YOLOv5 Study, 2022; Koohmishi et al., 2024; Li et al., 2024a, 2024b; Tang et al., 2024; Wang et al., 2021).

The main unresolved engineering challenge is integration. A useful system must detect heterogeneous defects, measure geometry, locate events, operate in real time, retain evidence, and remain affordable enough for repeated deployment. An indigenous ITMS that combines these functions through sensor fusion and edge computing is therefore a justified research direction. Its success should ultimately be measured not only by model

accuracy but by reliable field performance, clear maintenance information, and repeatable measurements under real railway conditions.

REFERENCES

- [1] Escalona, "Contactless track geometry measurement system (TGMS) using IMU, cameras and encoder," 2020.
- [2] Peng et al., "Geometry and surface-defect measurement using cameras, inclinometer, IMU and laser with Kalman filtering and ICP registration," 2020.
- [3] Antognoli et al., "Comparative assessment of stereo and thermal cameras for contactless wheel-to-rail sensing," 2021.
- [4] Escalona et al., "Vision-inertial geometry measurement using multibody kinematics," 2021.
- [5] Wang et al., "Camera-LiDAR fusion for railway object detection," 2021.
- [6] "Improved YOLOv5 study: Automated detection of railway fastener defects," 2022.
- [7] Haroon et al., "Supervised and self-supervised learning for rail-defect detection," 2024.
- [8] Koohmishi et al., "Machine learning with GPR and InSAR for track-health and asset-management applications," 2024.
- [9] Li et al., "Low-cost irregularity inspection using RTK GNSS, MEMS IMU, odometer and lasers with ZUPT, RTS and NHC constraints," 2024a.
- [10] Li et al., "Improved YOLOv8-based fastener-defect detection for heavy-haul railways," 2024b.
- [11] Tang et al., "Real-time camera-LiDAR track detection and 3-D fitting," 2024.
- [12] Bravo et al., "Systematic review of data-driven on-board accelerometer methods for track-condition monitoring," 2025.
- [13] "Embedded contour-feature processing study for rail-surface deformation detection," n.d. As cited in the supplied manuscript.