

Investigation of Rail Defects Induced by Rolling Contact Fatigue Using YOLOv8 - Based Computer Vision and Spatial - Temporal Post - Processing

Olzhas Suleimenov¹, Davran Radjibaev², Yadgor Ruzmetov², Aliya Zabiyeva^{1*}, Alipbayev Zhasulan^{3*}, Nurbol Kamzanov^{3*}

¹ Faculty of Transport and Energy, Eurasian National University, Astana, Kazakhstan

² Faculty of Railway Transport Engineering, Tashkent State Transport University, Tashkent 100176, Uzbekistan

³ Department of Technological Machines and Transport at Satbayev University, Almaty, Kazakhstan

Article Info

Article history:

Received September 4, 2026

Revised September 24, 2026

Accepted October 3, 2026

Keywords:

Rail Crack Detection,
YOLOv8,
Video Analytics,
Temporal Filtering,
Railway Infrastructure
Monitoring,
Finite Element Analysis

ABSTRACT

Rail surface defects caused by rolling contact fatigue (RCF) significantly affect railway operational safety and infrastructure reliability. This study proposes an automated rail surface defect detection framework based on the YOLOv8 object detection architecture combined with a lightweight spatial-temporal post-processing module designed for video-based monitoring. We collected a dataset under real operating conditions in Kazakhstan, including snow, variable illumination, and motion-induced blur, with temperatures down to -32 °C. Four object classes were annotated: rail, crack defect, joint, and bolt_joint. The baseline YOLOv8 model achieved an average mAP@0.5 of 0.849. However, crack detection recall was limited (0.644) due to small defect size and temporal instability in video sequences. The proposed spatial-temporal confirmation mechanism increased crack recall to 0.688 without modifying the neural network architecture or increasing inference complexity. Additionally, finite element simulations of R65 rails with longitudinal cracks (2 mm and 4 mm depth) under a 120 kN contact load demonstrated stress concentration growth and a reduction in safety factor from 3.48 (intact rail) to 3.15 (4 mm crack). The results confirm the engineering relevance of early crack detection and demonstrate the feasibility of integrating YOLO-based monitoring with structural integrity assessment.

Copyright © 2026 Reports in Mechanical Engineering.
All rights reserved.

Corresponding Author:

Aliya Zabiyeva

Faculty of Transport and Energy, Eurasian National University, Astana, Kazakhstan

Email: zabiyeva_ab@enu.kz

Alipbayev Zhasulan

Department of Technological Machines and Transport at Satbayev University, Almaty, Kazakhstan

Email: alipbayev_zhr@enu.kz

Nurbol Kamzanov

Department of Technological Machines and Transport at Satbayev University, Almaty, Kazakhstan

Email: n.kamzanov@satbayev.university

1. Introduction

Rolling contact fatigue (RCF) is one of the dominant degradation mechanisms in heavily loaded railway lines. Repeated wheel-rail contact stresses initiate microcracks at or near the rail head surface, which may propagate under cyclic loading. Without timely detection, these cracks may evolve into critical defects affecting rail integrity and safety. (Izvoltova et al., 2022).

Conventional inspection methods, including ultrasonic testing and eddy-current inspection, remain standard practice. However, periodic inspection intervals and high operational costs limit them. Integrating computer vision and deep learning enables continuous or quasi-continuous rail monitoring using standard imaging hardware (Gong et al., 2022; Peng et al., 2019).

Among modern object detection models, YOLO-family detectors provide a favorable balance between inference

speed and detection accuracy. YOLOv8, in particular, represents an optimized one-stage detector suitable for real-time applications.

However, when applied to video-based inspection, standard YOLO architectures treat frames independently, which leads to temporal instability and short-lived false detections. This limitation motivates a lightweight spatial-temporal post-processing strategy.

Main contributions of this paper:

Real-condition dataset acquisition under extreme climatic conditions.

YOLOv8-based rail crack detection framework.

A computationally efficient spatial-temporal micromodification for video stabilization.

Integration of detected crack geometry into finite element stress analysis for residual life reasoning.

1.1 Modern Diagnostic Methods for Railway Rails

Comprehensive reviews of railway inspection technologies show that traditional non-destructive testing (NDT) methods remain the cornerstone of routine rail diagnostics. Among them, ultrasonic testing (UT) is the most widely used technique for detecting internal defects such as transverse cracks, longitudinal cracks, and head checks in the rail head and web. UT is highly effective at identifying subsurface discontinuities and can operate at relatively high inspection speeds when mounted on diagnostic vehicles.

Eddy current testing (ECT) is commonly used to detect surface and near-surface defects. This method is particularly sensitive to small cracks and suits high-speed inspection because it operates without contact. In many countries, including those in Europe, the United States, Japan, and China, these technologies are integrated into dedicated inspection trains and provide high diagnostic reliability for main railway corridors. (Barlybayev et al., 2024; Gong et al., 2022).

Despite their effectiveness, conventional NDT methods have several limitations. First, inspections are typically periodic rather than continuous, creating time gaps between measurements and increasing the risk of undetected crack propagation. Second, they require specialized equipment and trained personnel, leading to high operational costs. Third, inspection quality may be influenced by calibration accuracy, environmental conditions, and human factors. These limitations motivate exploring complementary diagnostic approaches based on automated data processing and artificial intelligence.

1.2 Diagnostic Practice in the Republic of Kazakhstan

In the Republic of Kazakhstan, ultrasonic testing remains the primary and most widely used method for detecting rail defects (Figure 1). It is routinely applied on main railway lines and heavily loaded track sections, particularly in freight corridors with high axle loads and harsh climatic conditions.

Given the railway network's large geographical extent and extreme environmental factors, including low winter temperatures and significant seasonal variability, maintaining a high inspection frequency can be operationally challenging. Although ultrasonic inspection reliably detects internal defects, its periodic nature limits continuous condition monitoring.

These factors create a need for complementary technologies that can increase inspection frequency without significantly raising operational costs. In this context, computer vision-based systems integrated with moving rolling stock offer a promising way to improve the diagnostic coverage of railway infrastructure.

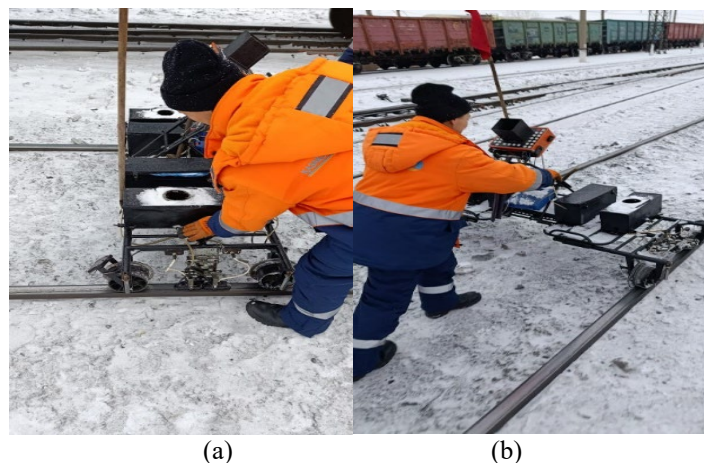


Figure 1: An Example of Working with a flaw detector in Kazakhstan

1.3 Deep Learning-Based Approaches for Rail Defect Detection

In recent years, artificial intelligence and deep learning techniques have gained significant attention as complementary tools for railway inspection. Among object detection architectures, the YOLO (You Only Look Once) framework is a convolutional neural network that performs real-time object detection and classification in a single forward pass. Unlike traditional region-proposal-based detectors, YOLO treats detection as a regression problem, enabling high inference speed while maintaining competitive accuracy.

YOLO-based models have demonstrated strong performance in detecting small objects and complex visual patterns under varying environmental conditions. Studies conducted by researchers in Kazakhstan have successfully applied YOLO architectures for small-object recognition tasks in challenging natural environments, reporting high detection accuracy and efficient data processing speeds. (Alimov & Meiramkhanov, 2024; Bai et al., 2021; Barlybayev et al., 2024; Zhexebay et al., 2025). These findings highlight YOLO's suitability for scenarios with variable illumination, occlusion, and complex backgrounds.

Similarly, research conducted by Chinese scholars has shown that improved YOLO variants can effectively detect rail surface defects, including cracks, spalling, and fastener-related anomalies, under real railway operating conditions. (Luo et al., 2023; Mi et al., 2023; Redmon et al., 2016; Xing et al., 2022; Zhang et al., 2023). Various modifications, such as feature pyramid enhancement, attention mechanisms, and multi-scale convolutional modules, have been proposed to improve sensitivity to small and low-contrast defects.

Despite these advances, applying YOLO to video-based rail inspection still faces challenges related to temporal instability, small defect scale, and motion-induced noise. Most implementations process each frame independently, without exploiting temporal consistency in video sequences. This limitation motivates the development of spatiotemporal post-processing strategies to improve defect persistence detection and reduce false positives in real operational environments.

2. Objects and Methods of Research

2.1 Environmental and Operational Conditions in the Republic of Kazakhstan

The railway network of the Republic of Kazakhstan spans regions characterized by diverse climatic and environmental conditions. These factors significantly influence both rail surface degradation mechanisms and the effectiveness of visual inspection systems.

Extreme seasonal temperature variations, snowfall, shadows, dust contamination, and variable illumination conditions affect image contrast and clarity. For convolutional neural networks such as YOLO, these environmental factors represent critical considerations during dataset preparation and model training. The detection model must recognize defects under limited visibility, non-uniform lighting, and motion-induced blur to ensure reliable, accurate defect identification in real operating environments.

2.2 YOLO Model Selection and Training Strategy

The YOLO (You Only Look Once) architecture requires supervised training on annotated image datasets (Bochkovskiy et al., 2020). To achieve high detection accuracy, hundreds or thousands of labeled images are typically required, covering variations in shooting angles, illumination conditions, object scales, and environmental backgrounds. The broader and more representative the training dataset, the more robust the model becomes against overfitting and performance degradation under real-world conditions (Redmon & Farhadi, 2018).

Effective training of YOLO involves several key aspects:

2.2.1 Loss Function Selection.

The loss function determines how accurately the model predicts bounding box coordinates, object confidence scores, and class labels. Properly configuring localization and classification losses is essential for stable convergence and small-object detection.

2.2.2 Data Augmentation

Image transformations such as rotation, flipping, brightness adjustment, scaling, and mosaic augmentation are applied to improve generalization under diverse environmental conditions. (Huang et al., 2020).

2.2.3 Architecture Micro-modifications

Previous studies suggest that modifications such as Dense Connection Spatial Pyramid Pooling (DC-SPP) or enhanced feature extraction modules can improve detection performance for small objects and complex backgrounds.

(Fang et al., 2024; Wang et al., 2024). Figures 2 and 3 show the standard YOLOv8 architecture and modified variants.

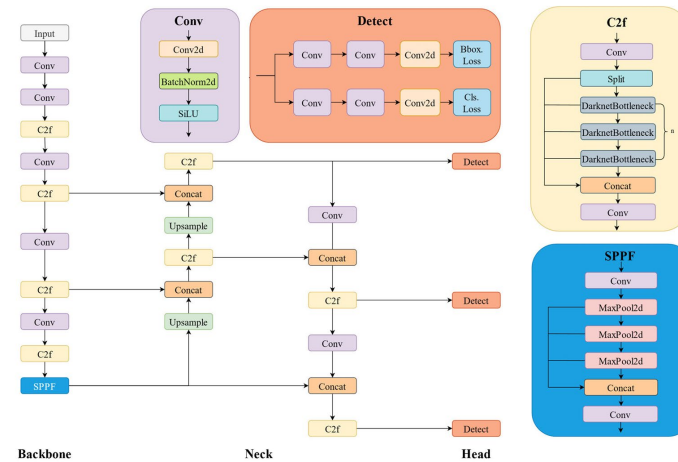


Figure 2: YOLOv8 Network Architecture (Source: (Chen et al., 2025)).

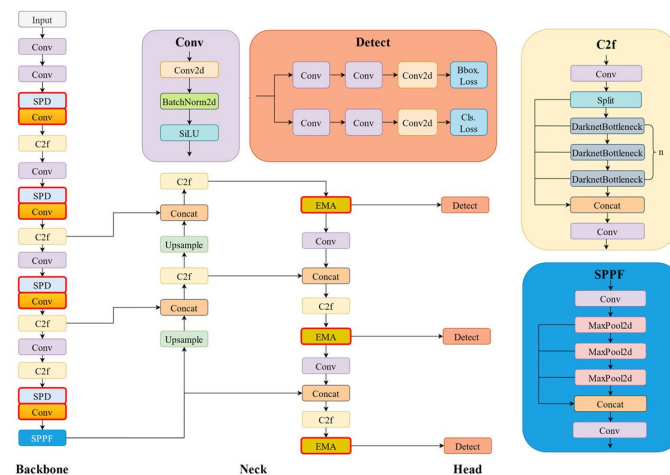


Figure 3: Improvements in the YOLOv8 Network Structure (Source: (Chen et al., 2025)).

2.3 Video Data Acquisition and Dataset Preparation

We collected video data for testing and image data for training under real railway operating conditions. We recorded the data using a Redmi 13 mobile device mounted on a flaw detection cart at a fixed height relative to the rail head. The cart moved at speeds up to 15 km/h, enabling realistic monitoring data.

Video recording was conducted under natural daylight conditions. Ambient temperature during data collection ranged from -30°C to -32°C . Additional environmental factors included light snowfall, shadows, and variable illumination, which increased dataset variability and improved the model's generalization capability. The experiments were conducted on a 5-7 km section of railway track. However, we also evaluated a 1 km segment, which had the highest concentration of rail surface defects.

This study collected a dataset to represent railway inspection under cold-climate conditions, including snowy backgrounds and adverse weather, which can increase the complexity of visual defect detection. We used 400 images to develop and evaluate the YOLOv8-based defect detection model. The dataset was divided into 320 training images (80%) and 80 validation images (20%). The training and validation subsets contained mutually exclusive images, with no identical images included in both subsets. The validation subset included challenging visual conditions and images containing a relatively high number of rail defects. However, the present dataset is suitable for experimental evaluation of the proposed approach rather than full-scale practical deployment. For implementation on the railway network of Kazakhstan, a substantially larger and more diverse dataset would be required. Such a dataset should cover different railway sections, rail conditions, defect types and sizes, seasons, weather conditions, illumination levels, camera

positions, and background variations. Therefore, developing a large-scale dataset containing thousands, or potentially tens of thousands, of representative images is an important direction for future work to improve the robustness and generalizability of the proposed model under real operating conditions.

Extracted frames were manually annotated using the LabelImg tool in YOLO format. The dataset included four object classes: rail, defect (crack), joint, bolt_joint.

All images were annotated with four classes: rail, defect (crack), joint, and bolt_joint. The training emphasized the defect (crack) class, as the primary goal of the proposed approach is reliable detection of rail surface defects. Joint and bolt-connection detection provides reference points along the railway track, which can approximate the location of detected defects. Because rail joints occur at known intervals, they serve as spatial markers for estimating defect kilometer positions when GPS data is unavailable or inaccurate.

2.4 Evaluation Metrics

We evaluated detection performance using Precision, Recall, Average Precision (AP), and mean Average Precision (mAP). Let TP denote true positives, FP false positives, and FN false negatives, and let C be the number of object classes. The evaluation metrics are defined as follows. (Chen et al., 2025; Du et al., 2024; Kang et al., 2017; Lv et al., 2025; Mao et al., 2024; Wu & Yu, 2025):

$$Precision = \frac{TP}{TP+FP} \quad (1)$$

$$Recall = \frac{TP}{TP+FN} \quad (2)$$

$$AP = \int_0^1 P(R)dR \quad (3)$$

$$mAP = \frac{1}{C} \sum_{i=1}^C AP_i \quad (4)$$

2.5 Hardware and Software Environment

The experiments to train and test the neural network model were conducted on a personal computer with the following specifications: Intel Core i5-13600KF processor, 32 GB DDR4 RAM, and an NVIDIA RTX 4070 (12 GB) graphics card with 5,888 CUDA cores.

The software environment included Windows 11 Pro, Python 3.10, PyTorch, Ultralytics YOLOv8, CUDA 12, and the Conda dependency management environment.

3. Results

3.1 Detection Performance on Validation Dataset

Table 1 presents the YOLOv8 model's object detection performance on the validation dataset.

Table 1: Quality Assessment Results

Class	Precision	Recall	mAP@0.5	mAP@0.5:0.95
rail	0.893	1.000	0.942	0.813
defect (crack)	0.867	0.644	0.648	0.309
joint	0.897	1.000	0.995	0.766
bolt_joint	0.882	0.781	0.812	0.548
Average	0.885	0.856	0.849	0.609

The model demonstrates high detection performance for structural components such as rails and joints. However, crack detection has lower recall because defects are small, low-contrast, and irregularly shaped.

The dataset consists of four object classes: rail, defect (crack), joint, and bolt_joint. The study used a total of 401 annotated instances. The object distribution across classes is imbalanced, reflecting real-world railway conditions where structural components appear more frequently than defects. Specifically, the dataset contains 192 rail instances, 114 defect (crack) instances, 29 joint instances, and 66 bolt_joint instances. The relatively small number of joint samples may lead to optimistic recall values, while the limited number of defect samples contributes to reduced detection performance for this class. We considered this imbalance when interpreting evaluation metrics.

3.2 Rail Defect Detection on Images

We selected the YOLOv8 architecture, a one-stage object detector, to address the rail defect detection task.

Experimental results show that the trained model reliably detects elongated structural elements such as rails and rail joints, even under changes in illumination and perspective distortion.

However, crack detection is more challenging due to:

- small object size,
- blurred boundaries,
- visual similarity to background textures.

Examples of rail defect detection during motion are shown in Figure 4.

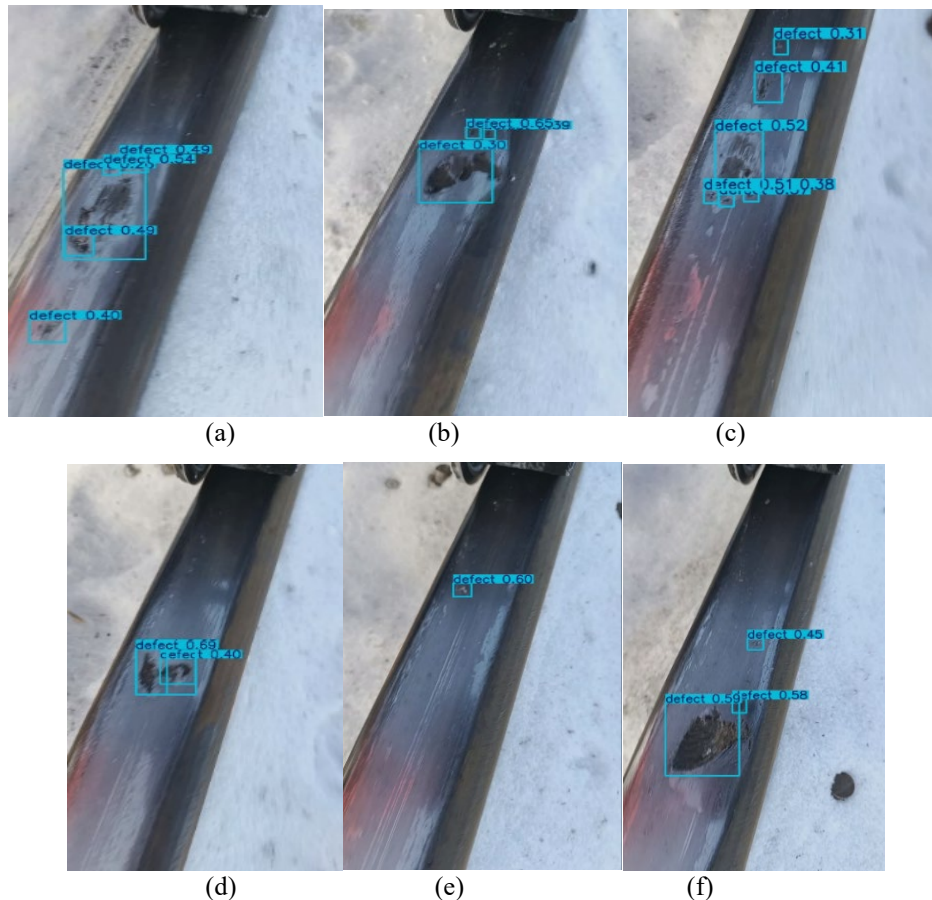


Figure 4: The Result of Detecting Rail Defects with the YOLOv8 Model in Motion.

Despite high detection accuracy for rails and joints, the model occasionally produced false positives in crack detection. These errors are primarily caused by surface heterogeneity of the rail head. The effect is particularly noticeable in video sequences and motivated the development of a spatial-temporal post-processing method described in the following section.

3.3 Rail Defect Detection in Video Sequences

When applying the standard YOLOv8 model independently to each video frame, large structural components (rails and joints) are detected consistently. However, small crack defects show short-term false detections and temporal instability, appearing in only one or two adjacent frames.

This effect occurs because YOLOv8 processes each frame independently without considering temporal continuity. As a result, transient visual noise, camera vibrations, or lighting artifacts may be misinterpreted as defects.

Figure 5 presents the proposed rail defect detection framework for video data, which combines YOLOv8 with a spatial-temporal post-processing module. Unlike the baseline YOLO algorithm, which analyzes frames independently, the proposed approach incorporates temporal persistence and spatial localization consistency across consecutive frames.

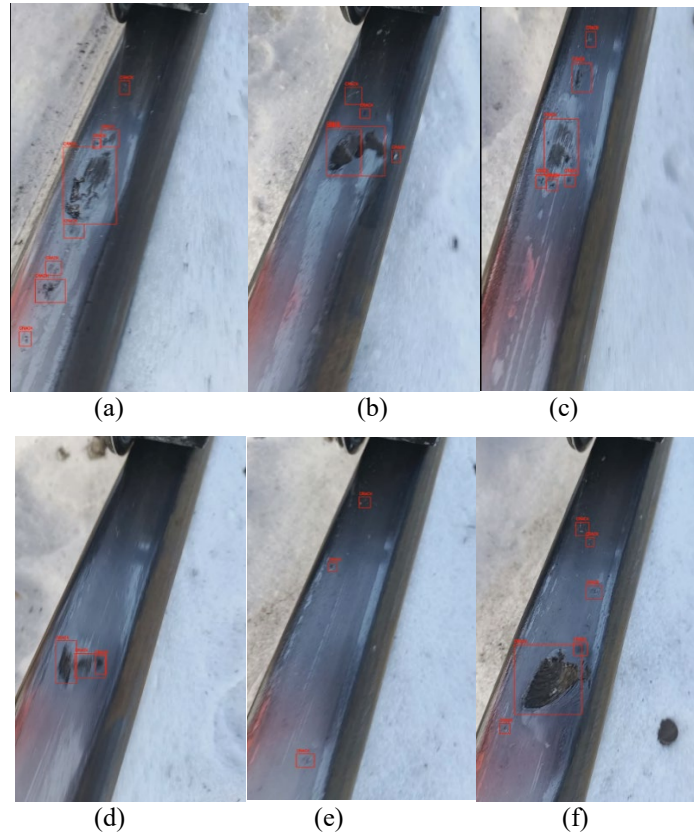


Figure 5: Schematic Diagram of Rail Defect Detection in Video Sequences Using YOLOv8 and Spatio-Temporal Filtering of Detections.

The proposed micro-modification adds a post-processing stage that dynamically analyzes the detection results produced by the YOLOv8 neural network within the video stream.

To improve the robustness of decision-making, a probabilistic detection confirmation model is applied, based on the conditional probability of object detection in adjacent frames:

$$P(D_t | D_{t-1}) > \theta \quad (5)$$

where:

$P(D_t | D_{t-1}) > \theta$ - the conditional probability of detection at frame t given that a detection was present in the previous frame;

θ - the confirmation threshold.

Additionally, a confidence smoothing mechanism is applied to reduce the influence of short-term noise and camera vibrations:

$$C_t = \alpha \times C_t + (1 - \alpha) \times C_{t-1} \quad (6)$$

where:

C_t - the adjusted detection confidence at frame t ;

α - the smoothing coefficient;

C_{t-1} - the confidence score at the previous frame.

The decision criterion for confirming the presence of a defect is based on a majority voting principle within a window of k consecutive frames:

$$\sum_{i=t}^{t+k} I(D_i) \geq m \quad (7)$$

where:

$I(D_i)$ - detection indicator (1 - defect detected, 0 - no defect);

m - the minimum number of confirmations within the window.

The condition of set intersection formalizes the final confirmation of detection:

$$D_t \cap D_{t+1} \cap \dots \cap D_{t+k} \cap \neq \emptyset \quad (8)$$

where:

D_t - the set of detections at frame t ;

k - the minimum number of consecutive frames;

$\cap$ - the set intersection operator.

In this study, we set the temporal window size to $k = 5$ consecutive frames. We set the detection confidence threshold to $\theta = 0.25$ and the smoothing coefficient to $\alpha = 0.5$. A defect was considered temporally confirmed when it was detected in at least $m = 2$ frames within the five-frame observation window. For spatial association between consecutive frames, we grouped the centers of the detected bounding boxes into 50×50 -pixel spatial cells. These parameters were kept fixed during evaluation of the proposed spatial-temporal post-processing method.

This approach significantly reduces false positives caused by vibration, illumination changes, and visual noise typical for railway infrastructure monitoring. (Baur & Nitsche, 2025; Bewley et al., 2016; Wojke et al., 2017; Zhuang et al., 2021).

3.4 Comparative Analysis of Detection Performance

Table 2 presents performance metrics for crack detection using the baseline YOLOv8 model and the proposed spatial-temporal micro modification.

Table 2: Comparison of methods

Method	Precision	Recall	mAP@0.5
Yolov8	0.867	0.644	0.648
YOLOv8 + Our Micro-Modification	0.867	0.688	0.648

The average metric values (Average), calculated across all classes, are presented for the baseline YOLOv8 model. The temporal micro-modification is applied selectively to improve the defect class recall; therefore, its impact on the averaged performance indicators is not evaluated.

As shown in the table, the baseline YOLOv8 model achieves high Precision and recall for large, structurally stable objects such as rails and rail joints. However, for the defect class, a decrease in Recall is observed, which is associated with the small object size and the influence of camera motion. (CanYang et al., 2025; Guan et al., 2025; Jamali et al., 2025; Tarkasis et al., 2025; Wang et al., 2026).

To improve defect detection recall, we proposed a temporal micro-modification based on confirming object detections across several consecutive frames. This approach increased recall for the defect class from 0.644 to 0.688 without modifying the neural network architecture or degrading Precision or mAP@0.5.

Despite the overall improvement in defect detection recall, the experiments identified rare false positives, accounting for no more than 20% of the total frames. Such errors predominantly occur under unfavorable external conditions, such as snow cover, contamination, glare, or repair work on the rail surface (Figure 6).

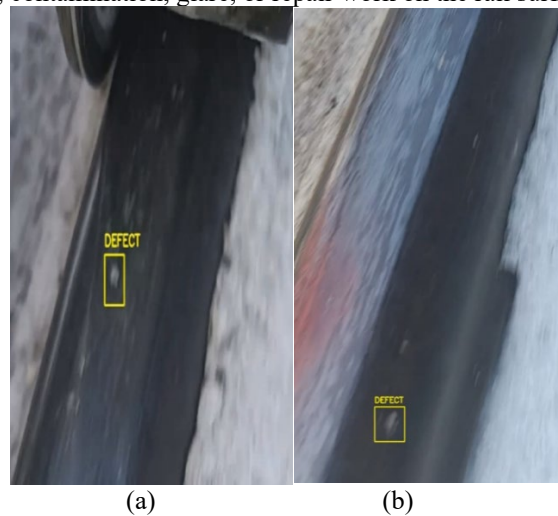


Figure 6: False Alarms

The analysis showed that the model's architectural limitations do not cause these false detections, but rather the

visual similarity of certain formations to actual rail defects. It should be noted that such cases are typical for most computer vision methods and are widely reported in the scientific literature (Ghodrati et al., 2018; Liu et al., 2024; Nezhad et al., 2024; Wang et al., 2022).

At the same time, applying the temporal micro-modification partially mitigates the influence of such factors by enforcing temporal consistency in detections, since random visual artifacts typically do not persist across consecutive frames.

4. Rail Strength Analysis Using Numerical Simulation

Using the YOLO neural network with the proposed micro-modification, it became possible to determine crack shape and location. In the future, the system can also be adapted to estimate defect dimensions. These data enable the construction of realistic CAD/CAE models of defects for subsequent engineering analysis of rail strength and residual service life. (Abdurasulov et al., 2025; Jamshiddoust, 2025; Khromova et al., 2017; Liu et al., 2024; Tulekov et al., 2023; Wang et al., 2023).

To quantitatively assess the influence of defects on R65 rails with a length of 1000 mm under a standard contact load of 120 kN, a series of simulations was carried out in SolidWorks Simulation 2023.

Table 3: Physical and Mechanical Properties of R65 Rail Material Used in SolidWorks Simulation

Parameter	Meaning	Units
Modulus of Elasticity	2.1×10^{11}	Pa
Poisson's Ratio	0.28	-
The Shear Modulus	7.9×10^{10}	Pa
Material Density	7700	kg/m ³
Ultimate Tensile Strength	7.24×10^8	Pa
Compressive Strength	7.24×10^8	Pa
Yield Strength	6.2×10^8	Pa
Coefficient of Linear Thermal Expansion	1.3×10^{-5}	1/K
Thermal Conductivity	50	-
Specific Heat Capacity	460	-
Structural Damping Coefficient	0.01 -0.03	-

The main objectives of the study were to determine the stress distribution, evaluate local displacements, and investigate deformations in the crack zone on the rail surface.

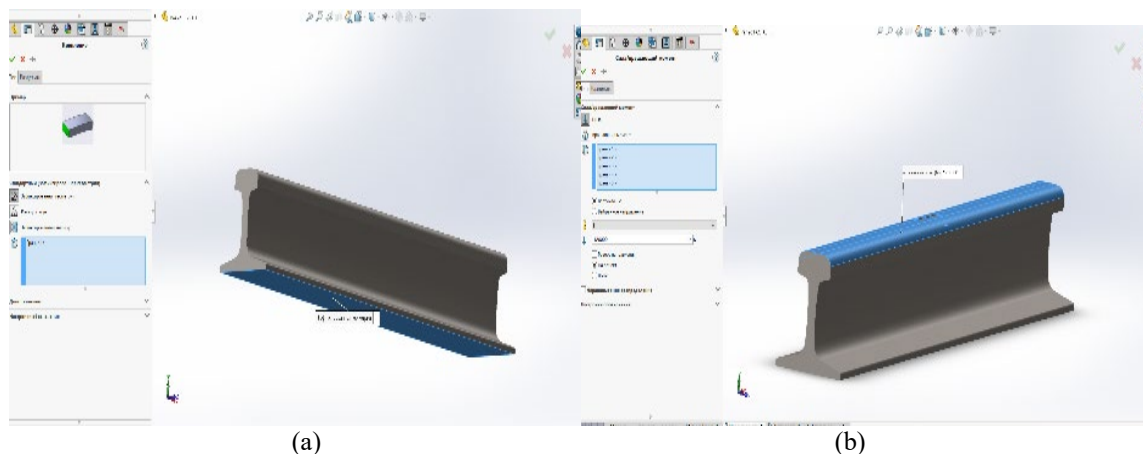


Figure 7: Boundary Conditions: (a) - Fixed Geometry; (b) - Load Application;

The lower surface of the rail foot was fully constrained using a fixed-geometry boundary condition, restricting translational displacement in all directions. A vertical load of 120 kN was applied to the upper surface of the rail head, perpendicular to the selected surface. The applied force was distributed over the selected surface of the rail head to represent the vertical loading acting on the rail during operation.

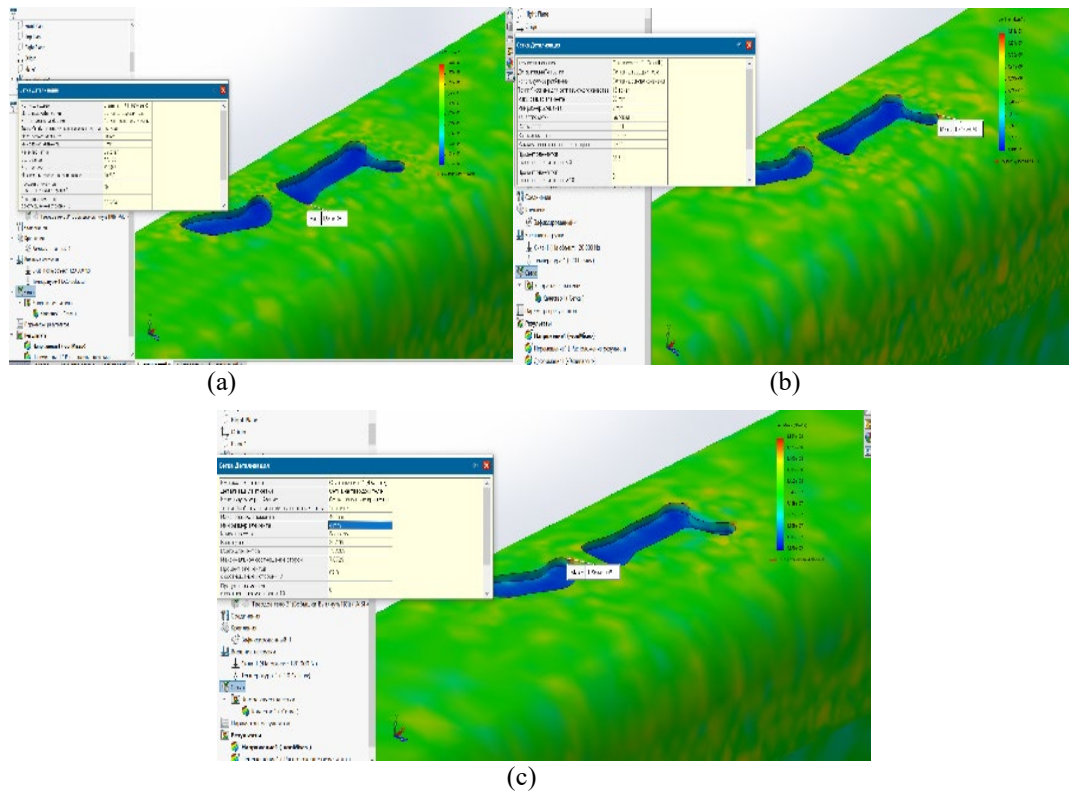


Figure 8: Mesh: (a) - 10mm/1mm; (b) - 30mm/2mm; (c) - 40mm/4mm

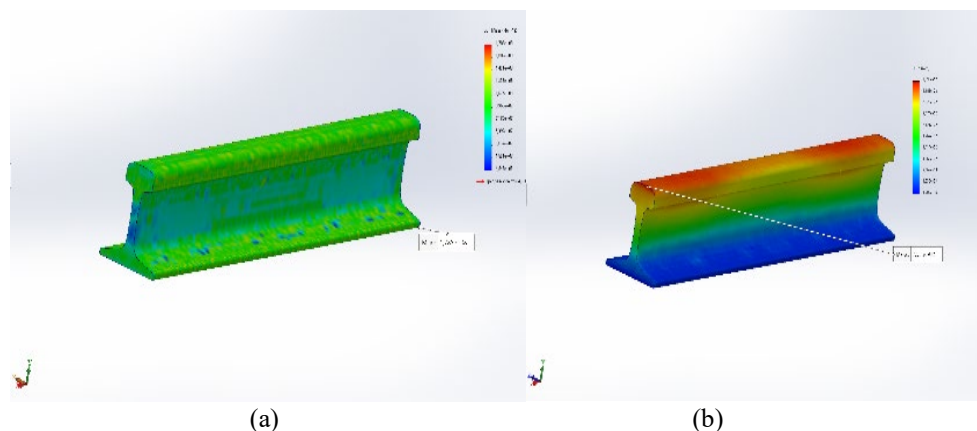
Table 4: Mesh-Independence Assessment of the Finite-Element Model

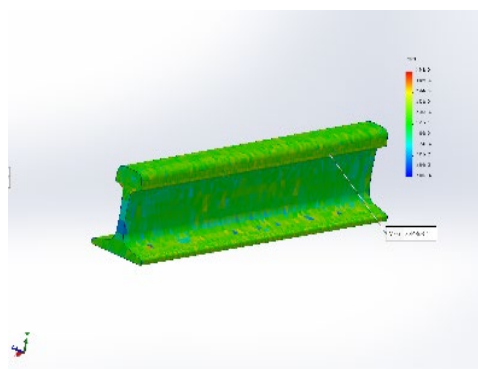
Max / Min Element	Max von Mises, MPa	Error, %
10mm / 1mm	188	3.8%
30mm / 2mm	181	0%
40mm / 4mm	186	2.7%

The medium mesh with maximum and minimum element sizes of 30 mm and 2 mm, respectively, was selected for the subsequent finite-element analysis. Compared with the selected mesh, the maximum von Mises stress varied by approximately 3.8% for the fine mesh and 2.7% for the coarse mesh. The relatively small variation in the calculated stress shows that the numerical results are not highly sensitive to further mesh refinement.

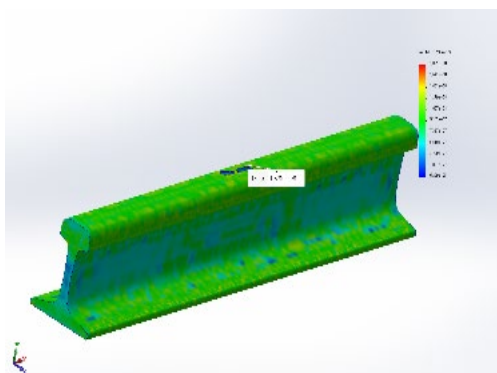
Three rail configurations were considered. (Khromova et al., 2017):

- Intact rail (without defects) (Figure 9);
- Rail with a longitudinal crack 2 mm deep (Figure 10);
- Rail with a longitudinal crack 4 mm deep (Figure 11).

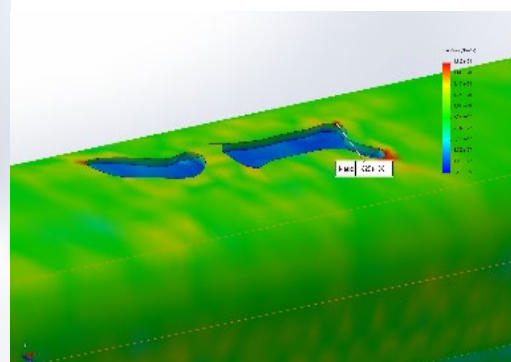




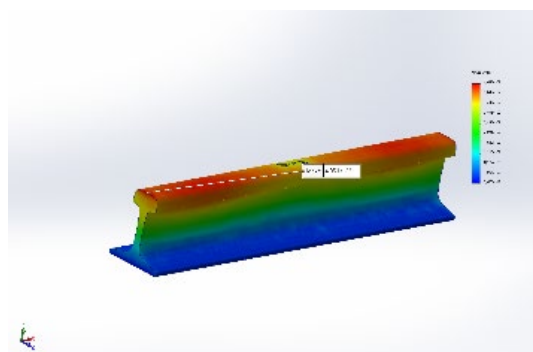
(c)

Figure 9: Intact Rail: (a) – Voltage; (b) - Moving; (c) - Deformation

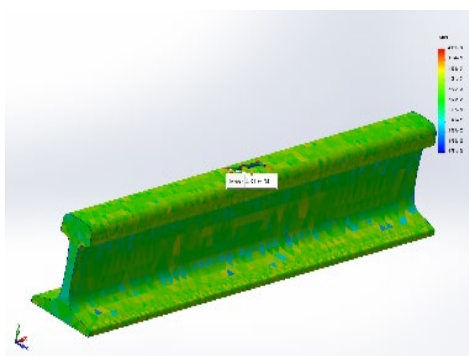
(a) - Voltage



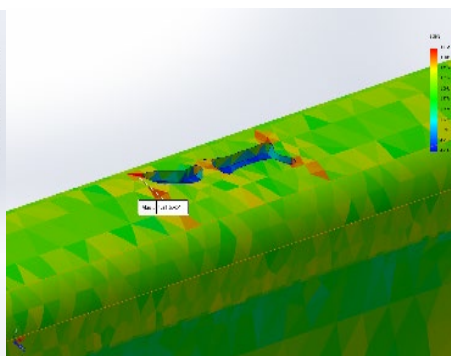
(a) - Voltage



(b) - moving



(c) - deformation



(c) - deformation

Figure 10: Rail with a Longitudinal Crack 2 mm deep

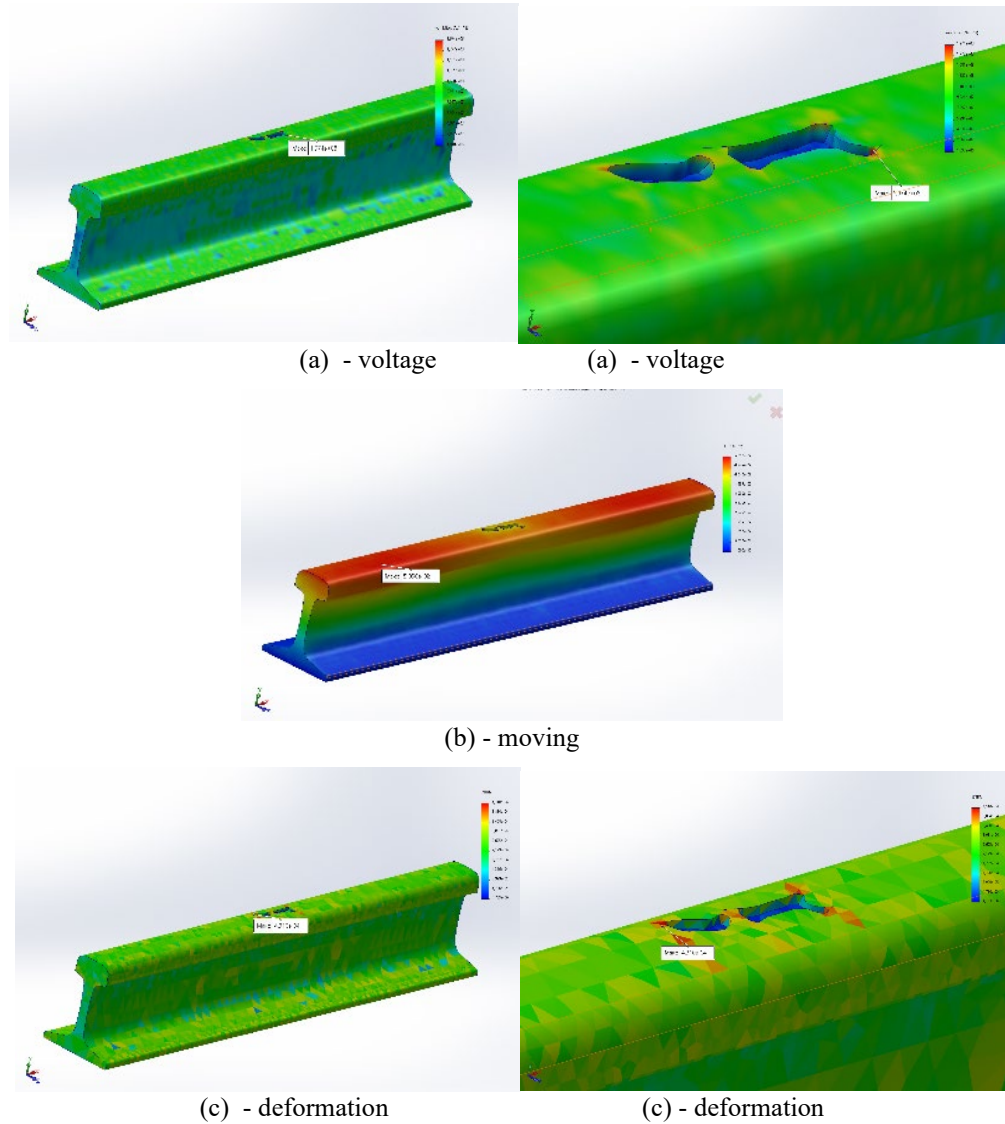


Figure 11: Rail with a Longitudinal Crack 4 mm deep: (a) - Voltage; (b) - Moving; (c) - Deformation

Equivalent von Mises stress was used as the criterion for evaluating the material's stress state. This parameter was automatically calculated in SolidWorks Simulation based on the stress tensor components. The theoretical expression is (Dumitriu & Gheți, 2018):

$$\sigma_{vM} = \sqrt{\frac{(\sigma_x - \sigma_y)^2 + (\sigma_y - \sigma_z)^2 + (\sigma_z - \sigma_x)^2 + 6(\tau_{xy}^2 + \tau_{yz}^2 + \tau_{zx}^2)}{2}} \quad (9)$$

where

$\sigma_x, \sigma_y, \sigma_z$ - normal stresses;

$\tau_{xy}, \tau_{yz}, \tau_{zx}$ - shear stresses.

The residual service life of the rail was evaluated based on the safety factor, defined as the ratio of the material's yield strength to the equivalent von Mises stress. (Dumitriu & Gheți, 2018; Jamshiddoust, 2025; Khromova et al., 2017; Khromova et al., 2025):

$$n = \frac{\sigma_y}{\sigma_{vM}} \quad (10)$$

For the intact rail, the safety factor was 3.48, indicating a significant residual service life.

With a longitudinal crack 2 mm deep, the safety factor decreased to 3.41, which reflects a moderate reduction in

residual life.

For a 4 mm deep crack, the safety factor further decreased to 3.15, indicating a noticeable reduction in the rail's residual service life.

Thus, the higher the local stress and deformation, the lower the rail's residual service life.

Table 5: Experimental Results

Configuration	σ_{max} (MPa)	δ_{max} (mm)	$\varepsilon_{max} \times 10^{-4}$
Intact rail	148	0,035	3,4
2 mm crack	182	0,04	5,6
2 mm crack	182	0,04	5,6

The calculation at the rail crack tip used fracture mechanics to evaluate the stress concentration. In particular, we determined the stress intensity factor K_I , which characterizes the potential risk of further defect propagation.

The stress intensity factor was calculated using the following formula. (Tarkasis et al., 2025; Trollé et al., 2014):

$$K_I = Y \times \sigma \times \sqrt{\pi a} \quad (11)$$

where

Y - geometric factor (≈ 1.12 for a surface crack);

σ - local stress at the crack tip (MPa);

a - crack depth (m).

However, increasing crack depth increases stress concentration and reduces residual service life. These results allow prediction of the dynamic strength and residual life of rails on Kazakhstan's main railway lines. Even cracks of 2-4 mm require monitoring and may influence maintenance scheduling or rail replacement decisions. (Trollé et al., 2014).

The increase in local stresses and deformations is explained by stress concentration at the crack tip. Even a small surface crack leads to a noticeable increase in σ_{max} and local displacements, thereby reducing the residual service life of the rail.

To more accurately assess the risk of defect propagation, we calculated the stress intensity factor K_I at the crack tips. The calculations showed that for a crack depth of 2 mm, $K_I \approx 16 \text{ MPa}\sqrt{\text{m}}$ and for a crack depth of 4 mm, $K_I \approx 25 \text{ MPa}\sqrt{\text{m}}$, which remains below the critical fracture toughness value $K_{IC} \approx 60 \text{ MPa}\sqrt{\text{m}}$ for rail steel.

However, increasing crack depth and number of cracks raises K_{IKI} and strengthens stress concentration, further reducing the rail's residual service life.

These results enable prediction of the dynamic strength and residual life of rails on main railway lines in Kazakhstan. Even cracks with a depth of 2-4 mm require monitoring and may influence the frequency of scheduled maintenance or rail replacement decisions. (Ning et al., 2025; Togizbayeva et al., 2023).

As a promising direction for developing rail track diagnostic systems, we propose using passenger rolling stock to carry video monitoring equipment. The concept involves installing a specialized measuring beam under passenger trains, equipped with a high-speed video recording system and a train speed sensor. Rail surface imaging is performed during normal train operation at speeds of 60–100 km/h, without interfering with the operational timetable (Figure 12).

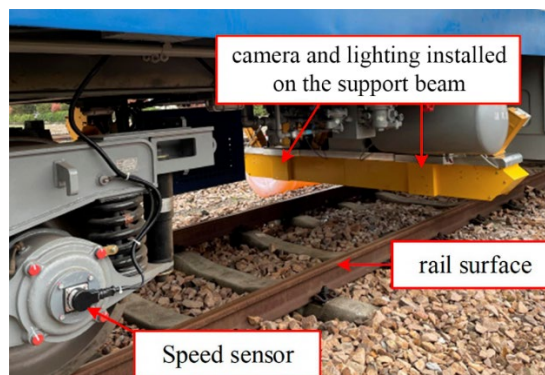


Figure 12: Schematic Diagram of a Rail Track Video Monitoring System based on Chen et al. (2025).

The video acquisition system includes high-speed cameras with enhanced spatial and temporal resolution, along

with a dedicated illumination system that ensures stable imaging under varying natural lighting conditions. The cameras and lighting elements are mounted on a support beam with an anti-vibration design to reduce the impact of dynamic loads and vibrations that occur during train movement.

Using a speed sensor enables spatial referencing of the acquired video data and synchronizes frames with the traveled distance, a necessary condition for subsequent automated analysis and localization of rail defects. The recorded video data can serve as input for intelligent analysis algorithms based on deep neural network object-detection models, enabling automatic identification of rail surface defects under real operating conditions.

5. Conclusion

Within the framework of this study, we developed and investigated an approach for automatically detecting railway rail surface defects using video data analysis with deep neural object detection models from the YOLO family. The study demonstrated that these models can detect defective rail sections under real operating conditions in the presence of noise.

We proposed a micro-modification to the baseline YOLO architecture to improve defect detection efficiency in video data from moving rolling stock. The modification does not significantly increase the model's computational complexity and preserves its applicability in near real-time data processing systems.

Experimental results show that the proposed micro-modification improves Recall compared to the baseline YOLO architecture: Recall reached 0.688, whereas the unmodified model did not exceed 0.644. These results indicate that the model can detect defective rail sections more effectively while maintaining acceptable computational costs. At the same time, experiments identified isolated false positives, accounting for no more than 20% of the analyzed frames. The analysis showed that these false detections mainly result from challenging imaging conditions, including rail surface contamination, uneven illumination, and foreign objects in the inspection area.

The structural analysis showed that longitudinal cracks in rails increase local stresses and deformations due to stress concentration at the defect tip. Increasing crack depth reduces the safety factor and residual rail life, as confirmed by stress-strain state calculations. Although the calculated stress intensity factors remain below the critical fracture threshold, they increase steadily with growing defect size. This indicates a potential risk of further crack propagation under long-term operation and cyclic loading. The results confirm the need for regular monitoring, even for small surface defects. The proposed approach enables more accurate prediction of residual rail life and justification of maintenance or replacement intervals.

Based on the results, we proposed a concept for implementing the developed approach within a перспективная rail video diagnostic system. The system includes installing a specialized anti-vibration measuring beam under passenger rolling stock, equipped with speed cameras, artificial lighting, and a speed sensor.

Such an implementation enables monitoring of rail track condition during regular train operation at speeds of 60-100 km/h without disrupting the operational schedule. For practical deployment and improved robustness, we recommend expanding the training dataset by at least 10-15 times the volume used in the experimental study. This will allow consideration of a wider range of operating conditions and reduce the influence of rare imaging scenarios.

6. Author Contributions

Author Contributions: Conceptualization, O.S. and A.Z.; methodology, O.S.; software, O.S.; validation, O.S., D.R. and A.Z.; formal analysis, O.S.; investigation; resources, O.S.; data curation, O.S.; writing-original draft preparation, Z.A. and Y.R.; writing-review and editing, N.K. and Y.R.; visualization, O.S.; supervision, O.S.; project administration, O.S.; funding acquisition, A.Z. All authors have read and agreed to the published version of the manuscript.”

Olzhas Suleimenov, Aliya Zabiyeveva, Davran Rajibayev, Yadgor Ruzmetov, Alipbayev Zhasulan, Nurbol Kamzanov

All authors participated in drafting the manuscript and critically revising it for important intellectual content.

All authors approved the final version of the manuscript and agree to be accountable for all aspects of the work, ensuring the accuracy and integrity of the research.

All authors have read and agreed to the published version of the manuscript.

7. Data Availability Statement

The data supporting the reported results are available from the corresponding author upon reasonable request.

Because of the large size of the video dataset and operational restrictions related to railway infrastructure

monitoring, we cannot make the raw video data publicly available.

References

- Abdurasulov, S., Zayniddinov, N., Khamidov, O., Yusuf, A., & Jamilov, S. (2025). Stress-strain state analysis of the crossbeam of the main frame of industrial electric locomotives PE2M and PE2U. In *AIP Conference Proceedings* (Vol. 3256, pp. 060011). AIP Publishing LLC. <https://doi.org/10.1063/5.0266927>
- Alimov, M., & Meiramkhanov, T. (2024). Domain generalization in autonomous driving: Evaluating yolov8s, rt-detr, and yolo-nas with the road-almaty dataset. *arXiv preprint arXiv:2412.12349*. <https://doi.org/10.48550/arXiv.2412.12349>
- Bai, T., Gao, J., Yang, J., & Yao, D. (2021). A study on railway surface defects detection based on machine vision. *Entropy*, 23(11), 1437. <https://doi.org/10.3390/e23111437>
- Barlybayev, A., Amangeldy, N., Kurmetbek, B., Krak, I., Razakhova, B., Tursynova, N., & Turebayeva, R. (2024). Personal protective equipment detection using YOLOv8 architecture on object detection benchmark datasets: a comparative study. *Cogent Engineering*, 11(1), 2333209. <https://doi.org/10.1080/23311916.2024.2333209>
- Baur, J., & Nitsche, F. O. (2025). A False-Positive-Centric Framework for Object Detection Disambiguation. *Remote Sensing*, 17(14), 2429. <https://doi.org/10.3390/rs17142429>
- Bewley, A., Ge, Z., Ott, L., Ramos, F., & Upcroft, B. (2016). Simple online and realtime tracking. In *2016 IEEE international conference on image processing (ICIP)* (pp. 3464-3468). Ieee. <https://doi.org/10.1109/ICIP.2016.7533003>
- Bochkovski, A., Wang, C.-Y., & Liao, H.-Y. M. (2020). Yolov4: Optimal speed and accuracy of object detection. *arXiv preprint arXiv:2004.10934*. <https://doi.org/10.48550/arXiv.2004.10934>
- CanYang, X., Yingying, L., Yongqiang, L., Runliang, T., & Tao, G. (2025). Lightweight rail surface defect detection algorithm based on an improved YOLOv8. *Measurement*, 242, 115922. <https://doi.org/10.1016/j.measurement.2024.115922>
- Chen, L., Sun, Q., Han, Z., & Zhai, F. (2025). DP-YOLO: a lightweight real-time detection algorithm for rail fastener defects. *Sensors*, 25(7), 2139. <https://doi.org/10.3390/s25072139>
- Du, J., Zhang, R., Gao, R., Nan, L., & Bao, Y. (2024). RSDNet: A new multiscale rail surface defect detection model. *Sensors*, 24(11), 3579. <https://doi.org/10.3390/s24113579>
- Dumitriu, M., & Gheț, M. (2018). Influence of the interference of bounce and pitch vibrations upon the dynamic behaviour in the bogie of a railway vehicle. In *IOP Conference Series: Materials Science and Engineering* (Vol. 400, pp. 042020). IOP Publishing. <https://doi.org/10.1088/1757-899X/400/4/042020>
- Fang, Z., Li, L., Peng, L., Zheng, S., Zhong, Q., & Zhu, T. (2024). YOLOv8n-RSDD: a high-performance low-complexity rail surface defect detection network. *Ieee Access*, 12, 196249-196265. <https://doi.org/10.1109/ACCESS.2024.3466559>
- Ghodrat, M., Ahmadian, M., & Mirzaeifar, R. (2018). Modeling of rolling contact fatigue in rails at the microstructural level. *Wear*, 406, 205-217. <https://doi.org/10.1016/j.wear.2018.04.016>
- Gong, W., Akbar, M. F., Jawad, G. N., Mohamed, M. F. P., & Wahab, M. N. A. (2022). Nondestructive testing technologies for rail inspection: A review. *Coatings*, 12(11), 1790. <https://doi.org/10.3390/coatings12111790>
- Guan, Z., Wang, Z., Zhang, G., Li, L., Zhang, M., Shi, Z., & Jiang, N. (2025). Multi-object tracking review: retrospective and emerging trend. *Artificial Intelligence Review*, 58(8), 235. <https://doi.org/10.1007/s10462-025-11212-y>
- Huang, Z., Wang, J., Fu, X., Yu, T., Guo, Y., & Wang, R. (2020). DC-SPP-YOLO: Dense connection and spatial pyramid pooling based YOLO for object detection. *Information Sciences*, 522, 241-258. <https://doi.org/10.1016/j.ins.2020.02.067>
- Izvoltova, J., Izvolt, L., & Sestakova, J. (2022). *Analysis of methods used to diagnostics of railway lines*. IntechOpen. <https://doi.org/10.5772/intechopen.100835>

- Jamali, M., Davidsson, P., Khoshkangini, R., Ljungqvist, M. G., & Mihailescu, R.-C. (2025). Context in object detection: a systematic literature review. *Artificial Intelligence Review*, 58(6), 175. <https://doi.org/10.1007/s10462-025-11186-x>
- Jamshiddoust, A. (2025). Improving dynamic stability and strength investigation of a locomotive bolster-type bogie frame through fabricated design and damper implementation. *Advances in Mechanical Engineering*, 17(12), 16878132251401390. <https://doi.org/10.1177/16878132251401390>
- Kang, D., Emmons, J., Abuzaid, F., Bailis, P., & Zaharia, M. (2017). Noscope: optimizing neural network queries over video at scale. *arXiv preprint arXiv:1703.02529*. <https://doi.org/10.48550/arXiv.1703.02529>
- Khromova, G., Mukhamedova, Z., & Yutkina, I. (2017). Mathematical model of oscillations of bearing body frame of emergency and repair railcars. *Transport problems*, 12(1), 93-102. <https://doi.org/10.20858/tp.2017.12.1.9>
- Khromova, G., Radjibaev, D., Zabiyeveva, A., Kenesbek, A., & Mavlanov, A. (2025). Calculation of the main parameters of the two-line helical traction transmission of an electric locomotive based on diagnostic parameters. *Applied Sciences*, 15(4), 1730. <https://doi.org/10.3390/app15041730>
- Liu, Y., Hu, T., Wang, H., Zhuang, J., & Sheng, J. (2024). Dynamic implementation and verification of the gauge changing function of high-speed variable gauge bogies. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, 238(24), 11366-11378. <https://doi.org/10.1177/09544062241280688>
- Luo, H., Cai, L., & Li, C. (2023). Rail surface defect detection based on an improved YOLOv5s. *Applied Sciences*, 13(12), 7330. <https://doi.org/10.3390/app13127330>
- Lv, D., Meng, J., Meng, G., & Shen, Y. (2025). Railway Fastener Defect Detection Model Based on Dual Attention and MobileNetv3. *World Electric Vehicle Journal*, 16(9), 513. <https://doi.org/10.3390/wevj16090513>
- Mao, Y., Zheng, S., Li, L., Shi, R., & An, X. (2024). Research on rail surface defect detection based on improved CenterNet. *Electronics*, 13(17), 3580. <https://doi.org/10.3390/electronics13173580>
- Mi, Z., Chen, R., & Zhao, S. (2023). Research on steel rail surface defects detection based on improved YOLOv4 network. *Frontiers in neurorobotics*, 17, 1119896. <https://doi.org/10.3389/fnbot.2023.1119896>
- Nezhad, M. S., Larsson, F., Kabo, E., & Ekberg, A. (2024). Finite element analyses of rail head cracks: Predicting direction and rate of rolling contact fatigue crack growth. *Engineering Fracture Mechanics*, 310, 110503. <https://doi.org/10.1016/j.engfracmech.2024.110503>
- Ning, Y., Wang, L., Qiu, S., Zhang, X., Xing, B., Fu, X., Li, R., Zhang, J., Deng, K., & Wang, H. (2025). Shear strain-assisted cross-twin orientation regulation enhances the plasticity of ZM21 magnesium alloy. *Journal of Alloys and Compounds*, 183040. <https://doi.org/10.1016/j.jallcom.2025.183040>
- Peng, Z., Wang, C., Ma, Z., & Liu, H. (2019). A multifeature hierarchical locating algorithm for hexagon nut of railway fasteners. *IEEE Transactions on Instrumentation and Measurement*, 69(3), 693-699. <https://doi.org/10.1109/TIM.2019.2905309>
- Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 779-788). <https://doi.org/10.1109/CVPR.2016.91>
- Redmon, J., & Farhadi, A. (2018). Yolov3: An incremental improvement. *arXiv preprint arXiv:1804.02767*. <https://doi.org/10.48550/arXiv.1804.02767>
- Tarkasis, K., Kaparis, K., & Georgiou, A. C. (2025). Enhancing trustworthiness in real time single object detection. *Information Systems Frontiers*, 1-12. <https://doi.org/10.1007/s10796-025-10584-4>
- Togizbayeva, B., Zabiyeveva, A., Balabekova, K., & Kenesbek, A. (2023). Development of a Calculation Method for Determining the Permissible Dimensions of the Defect 11. *Acta Polytechnica Hungarica*, 20(4). <https://tinyurl.com/mpj8258a>
- Trollé, B., Baietto, M.-C., Gravouil, A., Mai, S. H., & Prabel, B. (2014). 2D fatigue crack propagation in rails taking into account actual plastic stresses. *Engineering Fracture Mechanics*, 123, 163-181. <https://doi.org/10.1016/j.engfracmech.2014.03.020>

- Tulekov, A., Togizbayeva, B., Kenesbek, I., Kenesbek, A., & Zabiyeveva, A. (2023). The Use of Composite Materials in the Production of Tower Cranes. <https://doi.org/10.12700/APH.20.9.2023.9.16>
- Wang, J., Guo, X., Jing, L., & WANG, K. (2022). Finite element simulations of wheel-rail impact response induced by wheel tread spalling of high-speed trains. *Explosion and Shock Waves*, 42(4), 045103-045101-045103-045115. <https://www.sciengine.com/ESW/doi/10.11883/bzycj-2021-0374>
- Wang, L., Zhang, Q., & Wang, J. (2023). Microstructural effects on overall dynamics of composites: an analytical method via spatiotemporal nonlocal model: L. Wang et al. *Archive of Applied Mechanics*, 93(1), 73-91. <https://doi.org/10.1007/s00419-022-02206-1>
- Wang, M., Su, H., Yan, Z., Chen, C., Zheng, C., Lv, B., & Zhang, F. (2026). A Review of Experimental and Simulation Methods for Rolling Contact Fatigue. *Metals*, 16(2), 168. <https://doi.org/10.3390/met16020168>
- Wang, Y., Zhang, K., Wang, L., & Wu, L. (2024). An improved YOLOv8 algorithm for rail surface defect detection. *Ieee Access*, 12, 44984-44997. <https://doi.org/10.1109/ACCESS.2024.3380009>
- Wojke, N., Bewley, A., & Paulus, D. (2017). Simple online and realtime tracking with a deep association metric. In *2017 IEEE international conference on image processing (ICIP)* (pp. 3645-3649). IEEE. <https://doi.org/10.1109/ICIP.2017.8296962>
- Wu, X., & Yu, S. (2025). Research on the lightweight detection method of rail internal damage based on improved YOLOv8. *Journal of Engineering and Applied Science*, 72(1), 14. <https://doi.org/10.1186/s44147-025-00584-1>
- Xing, Z., Zhang, Z., Yao, X., Qin, Y., & Jia, L. (2022). Rail wheel tread defect detection using improved YOLOv3. *Measurement*, 203, 111959. <https://doi.org/10.1016/j.measurement.2022.111959>
- Zhang, C., Xu, D., Zhang, L., & Deng, W. (2023). Rail surface defect detection based on image enhancement and improved YOLOX. *Electronics*, 12(12), 2672. <https://doi.org/10.3390/electronics12122672>
- Zhexebay, D., Skabylov, A., Turlykozhaeva, D., Bolysbay, A., Turmaganbet, U., Namazbayev, T., & Mingliang, T. (2025). Real-time small object detection with YOLOv8n/8s and YOLOv11n/11s models in complex natural landscapes. *Physical Sciences and Technology*, 12(1-2), 142-153. <https://doi.org/10.26577/phst202512114>
- Zhuang, L., Qi, H., & Zhang, Z. (2021). The automatic rail surface multi-flaw identification based on a deep learning powered framework. *IEEE Transactions on Intelligent Transportation Systems*, 23(8), 12133-12143. <https://doi.org/10.1109/TITS.2021.3109949>