# AI-Driven Railway Maintenance for Fault Identification Through Object Detection and Segmentation

AKSHAR CHINTALAPALLY, AVINASH BEJJAM,
CHATURYA GANNE, VIPUL THOTA,
NAMRATHA REDDY GADDAM,
PRAFULLA KALAPATAPU
and VENKATA DILIP KUMAR PASUPULETI

## ABSTRACT

Railway inspection is an important task to ensure the safety and reliability of transportation systems. Regular inspection of key components such as sleepers, fasteners and tracks are essential to maintain the infrastructure and prevent accidents. This paper proposes a deep learning-based framework for automating railway inspection by combining object detection and fault segmentation. The experimentation was performed using the data comprising of track videos, which was further refined using image enhancement techniques where each frame of the video has been used for evaluating the models. The pipeline initiates with YOLO11 being employed in the first phase for detecting railway components due to its superior performance in limited data scenarios, with the second phase utilizing Mask R-CNN to detect and segment areas of damage and corrosion in the corresponding detected components. To benchmark our proposed pipeline, RTDETR was also used with Mask R-CNN and observed that YOLO11 combined with Mask R-CNN outperformed RTDETR in terms of accuracy and efficiency. This work highlights the potential of integrating advanced object detection and segmentation techniques to streamline railway maintenance by automating fault detection from images and video. The models were also evaluated on videos captured at varying speeds.

## INTRODUCTION

Over the past two decades, rail transit has emerged as a crucial element of transportation infrastructure globally, particularly in rapidly expanding networks such as India's. As Indian Railways transitions from traditional broad-gauge tracks to advanced metro and high-speed rail corridors, there is an increasing demand for sophisticated, automated maintenance systems, with safety becoming a priority for both public authorities and researchers. Railway tracks and their components, including sleepers and fasteners, are subjected to continuous stress from traffic loads, environmental factors, material fatigue, resulting in defects such as surface cracks, corrosion, or damaged fasteners, thereby elevating the risk of accidents.

Rail track infrastructure consists of four fundamental constituents, rails that govern the movement of trains, sleepers that transfer load to the ballast for support,

---

Akshar Chintalapally[1], Avinash Bejjam[1], Chaturya Ganne[1], Vipul Thota[1], Namratha Reddy Gaddam[1], Prafulla Kalapatapu[2], Venkata Dilip Kumar Pasupuleti[2]
[1]UG Student, [2]Faculty
Mahindra University, Hyderabad, India
se21ucse012@mahindrauniversity.edu.in,
se21ucse031@mahindrauniversity.edu.in,
se21ucse044@mahindrauniversity.edu.in,
se21ucse259@mahindrauniversity.edu.in,
se21ucse135@mahindrauniversity.edu.in,
prafulla.kalapatapu@mahindrauniversity.edu.in,
venkata.pasupuleti@mahindrauniversity.edu.in

joint bars that link rail parts and fasteners that bond sleepers to rails. Currently, track inspections are conducted manually, relying on visual assessments using basic tools under unpredictable weather conditions. This manual approach is characterized by low efficiency, inconsistent accuracy due to varying expertise, inadequate real-time monitoring, high costs and significant risks of human error. Consequently, there is a strong impetus to invest in enhancing railway inspection techniques through automation and advanced technology.

Artificial Intelligence (AI) has gained significant prominence due to its capacity to address complex problems using high-quality data. This advancement is driven by data proliferation, cost-effective storage solutions and progress in machine learning, particularly deep learning, which facilitates the creation of machines capable of emulating human abilities such as vision, auditory processing and decision-making [1]. Image-based inspection techniques utilizing computer vision have emerged as an alternative to manual inspections for crack detection.

In this context, the paper proposes an automated multi-object detection method based on a deep-learning framework, designed to detect rail surface fasteners and sleepers through a hybrid pipeline that integrates the YOLO11 (You Only Look Once) object detection model and the Mask R-CNN (Mask Region-based Convolution Neural Network) instance segmentation model.

## BACKGROUND

Modern railway inspections increasingly utilize specialized diagnostic vehicles, such as ROGER, equipped with advanced measurement and monitoring systems for flexible and efficient assessments [2]. In image-based crack detection, feature extraction is crucial, using methods like wavelet transformation, percolation techniques, Otsu’s thresholding and morphological operations to identify defects [3-6]. After extraction, classification algorithms sort crack types; however, traditional methods demand careful technique choice due to image complexity, noise and intensity changes, often reducing reliability in real-world applications. These challenges make conventional classification techniques less robust in dynamic environments, often leading to inconsistent performance during field deployment. Furthermore, manual tuning of parameters across varied track conditions adds to operational complexity.

When it comes to localization algorithms, techniques like template matching, pixel statistics and edge detection aim to boost accuracy by focusing on specific track components [7,8]. However, pixel statistics and edge detection can face challenges with uneven lighting and complex backgrounds, while traditional template matching often struggles to accurately identify damaged track sections. Various methods have been proposed to address these issues, including double-template matching, geometric property-based localization and variance projection with wavelet transforms [9-12]. These approaches leverage the fixed positional relationships between rails, fasteners and backing plates. Classical visual detection typically relies on hand-crafted features from fastener regions, classified using shallow-learning models to assess their condition.

For railway fastener detection, traditional computer vision methods rely on manually designed shallow feature extraction techniques, including Haar-like features, Dense SIFT, direction field, edge detection and Gabor filters [13-16]. Classification is typically performed using AdaBoost classifiers, Support Vector Machines (SVM), Probabilistic Graphical Models (PGM) and Multilayer Perceptron Neural Networks [17-20]. However, these methods overlook the extraction of critical local fastener features essential for accurate defect identification.

Deep learning models like Mask R-CNN and YOLO (V5, V7, V8) have shown strong performance in instance segmentation and object detection by leveraging texture, color, brightness and distance features [21]. While Mask R-CNN integrates detection with segmentation capabilities, achieving crack detection accuracies ranging from 67.6% to 81.1% [22,23], YOLO models emphasize real-time detection through end-to-end training, trading some accuracy for speed, making them particularly applicable to railway inspection scenarios [24]. In railway applications specifically, Song et al. implemented YOLOv3 for rail defect localization without shape or size estimation, Liang et al. employed SegNet for improved segmentation precision and James et al. developed TrackNet by combining U-Net and ResNet architectures, achieving moderate segmentation accuracy [25-27].

This proposed architecture outlines a deep learning strategy for identifying railway damage, leveraging both YOLO11 for real-time detection of key features like fasteners and sleepers and Mask R-CNN for segmenting damaged areas. Our method demonstrates efficiency compared to RT-DETR (Real-Time Detection Transformer) combined with Mask R-CNN. This pairing eliminates the drawbacks of existing literature, improving railway maintenance as more secure and dependable through automation. The enhanced model synergy also facilitates easier scalability and field adaptability for diverse inspection conditions

## METHODOLOGY

### Data Collection and Preprocessing

Figure 1 shows a complete overview of the dataset collection and annotation process. The dataset comprises approximately 1500 RGB images with a resolution of 3264×2448 pixels, captured from Indian railway lines using a manually operated camera positioned at elevated vantage points. Additional images, ranging from 640×640 to 1920×1020 pixels, were sourced from the railway-track dataset [28], ProjectData-2 dataset [29], Caronrial-2 dataset [30] and Crack_Segmentation Dataset [31]. The '*makesense.ai*' tool was employed to annotate images with boundary boxes for fasteners, sleepers, missing fasteners and rails for YOLO and RT-DETR training, as well as polygons for cracks and corrosion for Mask R-CNN. Data augmentation techniques, including flipping, rotating and shearing, were applied using Albumentations [32], thereby expanding the dataset to approximately 3200 images, which were subsequently divided into 80% for training and 20% for testing.

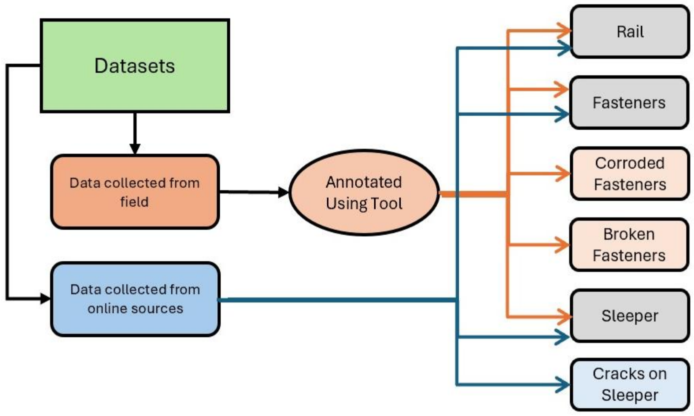



Figure 1: Illustration of Dataset Collection Process

## Model Overview

Among the various deep learning approaches, this study focuses on the use of YOLO11-object detection, RT-DETRv2 and Mask R-CNN-instance segmentation, each distinguished by its architectural design and processing workflow. The core mechanisms are detailed accordingly.

YOLO follows a single-stage approach, using grid-based analysis to process the entire image for the detection of bounding boxes. Available in various configurations, YOLO ranges from lightweight to more complex versions, depending on the specific application requirements.

In contrast, RT-DETR is a transformer-based model that combines a CNN-transformer encoder for multi-scale feature extraction with a decoder to predict bounding boxes and masks. It also incorporates deformable attention and a mask refinement module to improve accuracy.

Mask R-CNN, which builds upon Faster R-CNN, operates through a two-stage process: the first stage uses a backbone network and a Region Proposal Network (RPN), while the second stage focuses on predicting bounding boxes and class labels, with an additional mask head that generates pixel-wise segmentation masks for each detected instance. This model generates segmentation masks, class labels and confidence scores for detected objects, utilizing VGG-16 as the backbone.

## Proposed Algorithm

This research proposes a hybrid model that combines YOLO11 for real-time object detection with Mask R-CNN for pixel-level defect segmentation, as illustrated in Figure 2. Input images are first processed by YOLO11, which employs an anchor-based detection strategy using Intersection over Union (IoU) and Non-Maximum Suppression (NMS) to accurately localize railway components such as sleepers, fasteners and missing fasteners. During post-processing, predictions with an IoU exceeding a defined threshold (t) are suppressed, ensuring that only the detections with the highest confidence scores are retained.

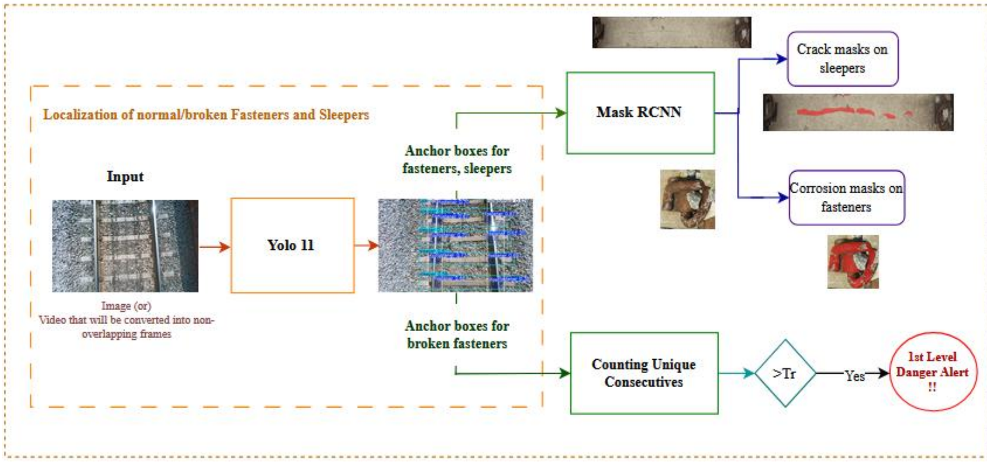



Figure 2: Overall framework of sleeper and fastener defect detection

Once the bounding boxes are generated by YOLO11, only the areas containing sleepers and fasteners are individually processed by a pre-trained Mask R-CNN model which subsequently produces pixel-level masks that delineate cracks and deformations in sleepers, as well as corrosion in fasteners. By utilizing RoI Align with bilinear interpolation and multipoint sampling, the model ensures precise spatial alignment and minimizes quantization errors, while maintaining computational efficiency through selective region processing that focuses computational resources only on relevant infrastructure components. Concentrating on localized regions rather than the entire image enhances both detection accuracy and computational efficiency, with the segmentation masks acting as visual markers for inspectors or grading machines to quantify the severity of the detected faults.

In parallel, the system also tracks missing fasteners using the anchor boxes generated by YOLO11, as shown in Figure 2. This module tracks the number of distinct missing fastener locations, triggering a high-level alert if the count exceeds a predefined threshold (Tr), set here to 4 in accordance with Indian Railway Authority guidelines, thereby indicating a critical safety concern. This alert is generated alongside the identification of other defects through the pixel-level masks produced by Mask RCNN. Together, these outputs provide a comprehensive evaluation of structural integrity and surface defects.

In this study, two hybrid pipeline comparisons were conducted, namely, YOLO11 with Mask R-CNN and RTDETRv2 with Mask R-CNN. In contrast to YOLO's NMS, RTDETR employs an anchor-free self-attention mechanism to better capture long-range dependencies and intricate spatial interdependencies. Once RTDETR localizes the parts of railways, the second step would be same as before passing only those areas into Mask-RCNN. Figure 3 illustrates the accuracy of both models over 150 epochs, showing that RTDETRv2 plateaued around 55%, while YOLO11 consistently improved from 50% to 90% by epoch 45. Given the dataset's specific needs and YOLO11's continuous improvement in its superior fault-detection accuracy, lead to its selection as the primary model.

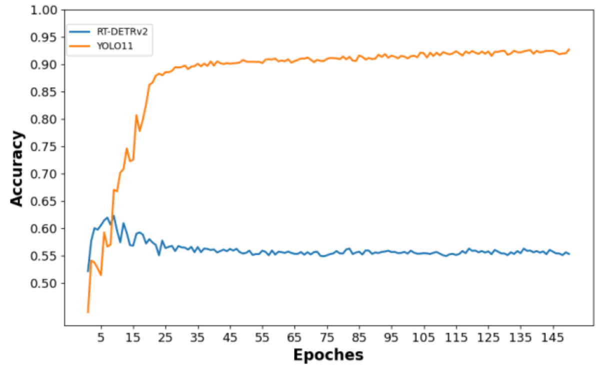



Figure 3. Accuracy Graphs Over Epochs

Mathematical Representation of the proposed Pipeline: Let the input image or video frame be denoted as I. The objective is to detect and segment n components $\{c_1, c_2, \ldots, c_n\}$, where each component has:

- Bounding box $b_{\{i\}} = (x,y,w,h)$, representing the top-left coordinates, width and height.
- Segmentation mask $m_i$, indicating the pixel-wise area of the component.

For a detection model f, the output is:

$$f^{(I)} \rightarrow \{b_1, b_2, \ldots, b_n\} \quad (1)$$

For a segmentation model g, the output is:

$$g^{(I)} \rightarrow \{m_1, m_2, \ldots, m_n\} \quad (2)$$

$$h^{(I)} = g(f^{(I)}) \rightarrow \{(b_1, m_1), (b_2, m_2), \ldots, (b_n, m_n)\} \quad (3)$$

## IMPLEMENTATION

### Training and Testing

Preprocessing techniques, including image enhancements and denoising, were applied to improve input quality and model performance under noisy environments. The system was validated through extensive experiments, including video analysis, demonstrating its scalability and efficiency in detecting, counting and analysing railway components in real time.

During training, various enhancement techniques were evaluated to improve defect visibility by enhancing contrast, sharpening edges and reducing noise. Global histogram equalization revealed faint cracks by redistributing intensities, while CLAHE and adaptive histogram equalization adjusted contrast locally. Unsharp masking emphasized high-frequency details and median and bilateral filters suppressed noise while preserving edges. These preprocessing steps proved critical in handling the challenging lighting conditions and environmental variations typically encountered in railway track inspection scenarios. Based on the results in Table I, global histogram equalization was selected for its balance of effectiveness and computational efficiency.

TABLE I: COMPARISON OF DIFFERENT IMAGE ENHANCEMENT TECHNIQUES

| Techniques | Avg Preprocessing Time (s) | Avg Prediction Time (s) | Avg Prediction Confidence (All Classes) |
|---|---|---|---|
| Original | 0.0000 | 0.0152 | 0.7135 |
| Unsharp Masking | 0.0214 | 0.0123 | 0.7770 |
| Histogram Equalization | 0.0031 | 0.0117 | 0.7509 |
| Adaptive Histogram Equalization | 0.0096 | 0.0117 | 0.7307 |
| Median Filter | 0.0121 | 0.0115 | 0.7249 |
| Laplacian Sharpening | 0.0611 | 0.0170 | 0.7196 |
| Gaussian Blur | 0.0052 | 0.0137 | 0.7166 |
| Bilateral Filter | 0.2918 | 0.0133 | 0.7126 |
| CLAHE | 0.0126 | 0.0136 | 0.6870 |

To enhance robustness against noisy inputs, a denoising step was also added before inference. Non-local means (NLM) and wavelet-based filtering were compared, with NLM averaging similar patches to suppress noise while preserving critical features, such as fasteners and crack edges. In contrast, wavelet filtering reduces high-frequency noise across multiple scales while maintaining edge sharpness. NLM was selected as the denoising algorithm due to its superior performance in handling a broader range of noise types, particularly in challenging outdoor environments where dust, vibrations and varying illumination conditions significantly impact image quality

## Evaluation Metrics

To assess the effectiveness of various localization models for real-time railway fault inspection in the initial stage of the analysis pipeline, several variants of the YOLO architecture, including RT-DETR, were thoroughly trained and evaluated. Specifically, four versions of YOLO (YOLOv8, YOLOv9, YOLOv10 and YOLO11) were implemented, each with different size configurations, to identify the most suitable model for the specific datasets used in this study.

The evaluation of object detection systems was based on key performance metrics, including Precision, Recall, F1 Score and mean Average Precision (mAP) [33]. Recall indicates the percentage of true positives correctly identified by the model, while Precision measures the ratio of true positives to all predicted positives. The F1 Score provides a combined measure of Precision and Recall. On the other hand, mAP offers a comprehensive assessment of model performance, with mAP50 evaluating detection accuracy at an intersection over union (IoU) threshold of 0.50 and mAP50-95 expanding evaluation across IoU thresholds from 0.50 to 0.95.

TABLE II: COMPARISON OF DIFFERENT YOLO MODELS BY PRECISION, RECALL, F1 SCORE, MAP-50, MAP50-95, INFERENCE AND POST-PROCESSING SPEED

| Models | | Error Metrics | | | | | Speed | |
|---|---|---|---|---|---|---|---|---|
| | | P | R | F1 Score | mAP 50 | mAP 50-95 | Inference (ms) | Postprocess (ms) |
| Yolo11 | s | 0.856 | 0.742 | 0.794 | 0.826 | 0.373 | 7.2 | 4.2 |
| Object | l | 0.825 | 0.707 | 0.761 | 0.798 | 0.355 | 19.8 | 3.3 |
| Detection | x | 0.742 | 0.784 | 0.762 | 0.794 | 0.353 | 33.5 | 2.2 |
| Yolo v10 | s | 0.75 | 0.757 | 0.753 | 0.764 | 0.359 | 9.1 | 0.5 |
| Object | l | 0.797 | 0.732 | 0.763 | 0.757 | 0.343 | 23.9 | 0.2 |
| Detection | x | 0.676 | 0.771 | 0.720 | 0.743 | 0.343 | 37.6 | 0.2 |
| Yolo v9 | s | 0.774 | 0.781 | 0.777 | 0.808 | 0.362 | 11.9 | 4.2 |
| Object | c | 0.844 | 0.693 | 0.761 | 0.756 | 0.346 | 21.5 | 2.1 |
| Detection | | | | | | | | |
| Yolo v8 | s | 0.691 | 0.805 | 0.743 | 0.791 | 0.36 | 7.4 | 4.9 |
| Object | l | 0.857 | 0.701 | 0.771 | 0.74 | 0.358 | 23.7 | 2 |
| Detection | | | | | | | | |

## RESULTS AND DISCUSSION

After model training, the test data set of track detection image is fed into the training model, to obtain the crack detection result. Figure 4 illustrates the YOLO11 localization outputs for sleepers, fasteners, missing fasteners and railway track. The detections accurately enclose each object across varied orientations, streamlining the subsequent crack-detection stage. All bounding boxes for sleepers and fasteners are then forwarded to Mask R-CNN for detailed segmentation. The Figure 4 also shows that the model remained robust under noise and occlusions, demonstrating strong generalization supported by the applied preprocessing.

Figure 5 illustrates closer views of the defects detected by Mask R-CNN in the pipeline's second stage, representing the final output. Figure 5(a) corresponds to corroded regions on fasteners that are masked to highlight rust-induced degradation on it. Figure 5(b) shows mask on cracks of sleeper, demonstrating the model's ability to analyze the defect and localize it precisely with pixel-level accuracy that enables quantitative assessment of defect severity and extent.

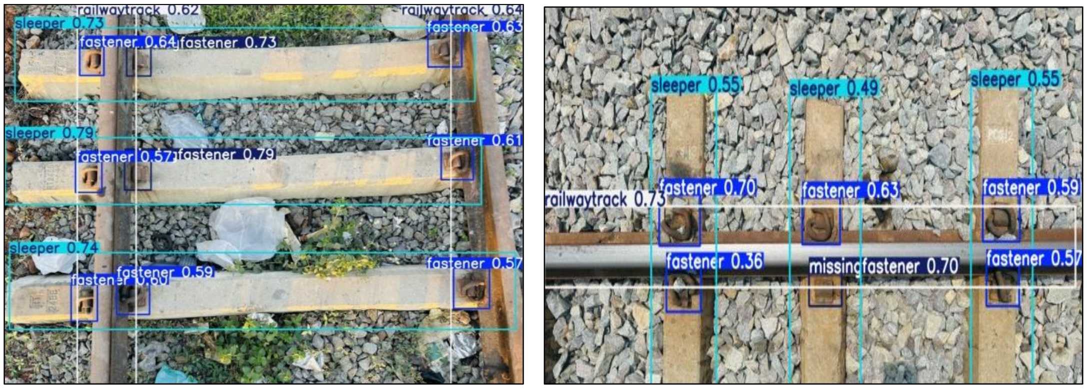
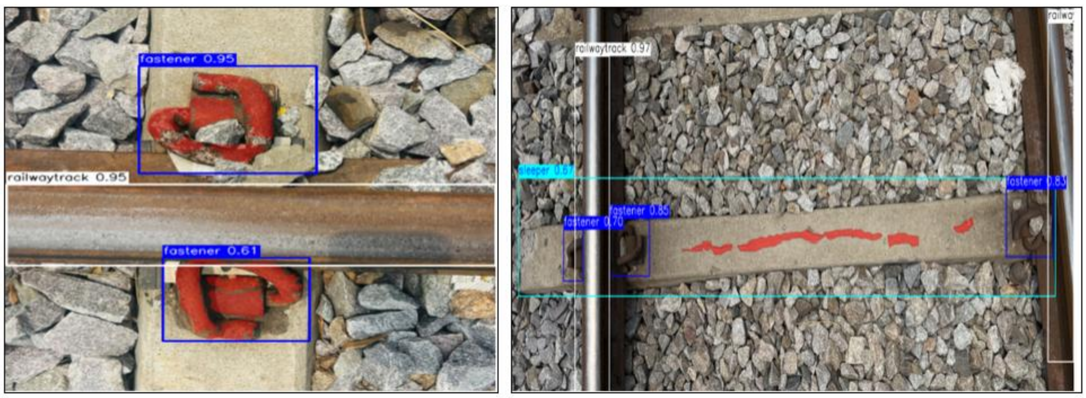



Figure 4. Some Class Localization Results of Rail Near a Station




(a) (b)

Figure 5. Zoomed In View of Defect Detection Results

## CONCLUSION AND FUTUREWORK

This paper introduced an automated framework for railway maintenance and inspection. The proposed pipeline effectively identifies structural defects of fasteners and sleepers, enhancing railway safety and efficiency. Furthermore, the automated approach can be integrated into mobile monitoring inspection devices, such as drones or railway-mounted inspection vehicles with onboard cameras, facilitating real-time defect detection improving emergency management and structural resilience.

This research also explores data augmentation through deliberate noise introduction during training, improving model robustness and enabling effective generalization to real world conditions. The proposed approach demonstrates adaptability and scalability and offers a foundation for extension to other rail components or sub-parts needing error analysis in the future. Future research could larger dataset with additional types of railway faults and various fastener types, enhancing model training, accuracy and defect outlining precision.

## REFERENCES


1. S. Rawas, "AI: the future of humanity," *Discover Artificial Intelligence*, vol. 4, art. 25, Mar. 2024. DOI: 10.1007/s44163-024-00118-3.
2. MERMEC, "Roger Inspection Vehicle Web Page"
3. R. Nigam and S. K. Singh, "Crack detection in a beam using wavelet transform and photographic measurements," *Structures*, vol. 25, pp. 436–447, 2020.
4. Z. Qu, Y.-X. Chen, L. Liu, Y. Xie and Q. Zhou, "The Algorithm of Concrete Surface Crack Detection Based on the Genetic Programming and Percolation Model," *IEEE Access*, vol. 7, pp. 57592–57603, 2019.
5. B. Chen, X. Zhang, R. Wang, Z. Li and W. Deng, "Detect concrete cracks based on Otsu algorithm with Differential Image," *Journal of Engineering*, vol. 2019, pp. 9088–9091, 2019.
6. H. Hou and W. Lin, "A new approach for the detection of concrete cracks based on adaptive morphological filtering," *Fuzzy Systems and Data Mining*, vol. 6, pp. 312–319, 2020.
7. C. Aytekin, Y. Rezaeitabar, S. Dogru and I. Ulusoy, "Railway fastener inspection by real-time machine vision," *IEEE Transactions on Systems, Man and Cybernetics: Systems*, vol. 45, no. 7, pp. 1101–1107, 2015.
8. L. Jiajia, X. Ying and L. Bailin et al., "Research on automatic inspection algorithm for railway fastener defects based on computer vision," *Journal of the China Railway Society*, vol. 38, pp. 73–80, 2016.
9. Y. Qiu, X. Chen and Z. Lv, "Rail fastener positioning based on double template matching," *Complexity*, vol. 2020, Article ID 8316969, 2020.
10. Y. Li, C. Otto, N. Haas et al., "Component-based track inspection using machine-vision technology," in *Proc. ACM Int. Conf. Multimedia Retrieval*, Trento, Italy, 2011.
11. Y. Li, H. H. Trinh, N. Haas, C. Otto and S. Pankanti, "Rail component detection, optimization and assessment for automatic rail track inspection," *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 2, pp. 760–770, 2014.
12. X. Wei, Z. Yang, Y. Liu et al., "Railway track fastener defect detection based on image processing and deep learning techniques: a comparative study," *Engineering Applications of Artificial Intelligence*, vol. 80, pp. 61–81, 2019.
13. H. Feng, Z. Jiang, F. Xie, P. Yang, J. Shi and L. Chen, "Automatic fastener classification and defect detection in vision-based railway inspection systems," *IEEE Transactions on Instrumentation and Measurement*, vol. 63, no. 4, pp. 877–888, 2014.
14. J. Yang, W. Tao, M. Liu, Y. Zhang, H. Zhang and H. Zhao, "An efficient direction field-based method for detection of fasteners on high-speed railways," *Sensors*, vol. 11, no. 8, pp. 7364–7381, 2011.
15. P. Dollar and C. L. Zitnick, "Fast edge detection using structured forests," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, no. 8, pp. 1558–1570, 2015.
16. X. Gibert, V. M. Patel and R. Chellappa, "Robust fastener detection for autonomous visual railway track inspection," in *Proc. IEEE Winter Conf. Applications of Computer Vision*, Waikoloa, HI, USA, 2015, pp. 694–701.
17. H. Trinh, N. Haas, Y. Li, C. Otto and S. Pankanti, "Enhanced rail component detection and consolidation for rail track inspection," in *Proc. IEEE Workshop Applications of Computer Vision*, Breckenridge, CO, USA, 2012, pp. 289–295.
18. P. D. Ruvo, A. Distante, E. Stella and F. Marino, "A GPU-based vision system for real-time detection of fastening elements in railway inspection," in *Proc. IEEE Int. Conf. Image Processing*, Cairo, Egypt, 2009, pp. 2309–2312.
19. F. Marino, A. Distante, P. L. Mazzeo and E. Stella, "A real-time visual inspection system for railway maintenance: automatic hexagonal-headed bolts detection," *IEEE Trans. Systems, Man, Cybernetics Part C*, vol. 37, no. 7, pp. 418–428, 2007.
20. X. Wei, Z. Yang, Y. Liu et al., "Railway track fastener defect detection based on image processing and deep learning techniques: a comparative study," *Engineering Applications of Artificial Intelligence*, vol. 80, pp. 61–81, 2019.
21. A. M. Hafiz and G. M. Bhat, "A survey on instance segmentation: state of the art," *International Journal of Multimedia Information Retrieval*, vol. 9, no. 3, pp. 171–189, 2020.
22. K. He, G. Gkioxari, P. Dollár and R. Girshick, "Mask R-CNN," in *Proc. IEEE Int. Conf. Computer Vision*, 2017, pp. 2961–2969.

23. M. Bai and H. Sezen, "Detecting cracks and spalling automatically in extreme events by end-to-end deep learning frameworks," in *ISPRS Annals Photogrammetry, Remote Sensing and Spatial Info. Sciences*, 2021.
24. J. Redmon, S. Divvala, R. Girshick and A. Farhadi, "You Only Look Once: Unified, real-time object detection," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2016, pp. 779–788.
25. Y. Song, H. Zhang, L. Liu and H. Zhong, "Rail surface defect detection method basedonYOLOv3deeplearningnetworks," in Proceedings of the 2018 Chinese Automation Congress, pp. 1563–1568, Xi'an, China, December 2018.
26. Z. Liang, H. Zhang, L. Liu, Z. He and K. Zheng, "Defect detection of rail surface with deep convolutional neural networks," in Proceedings of the 2018 13th World Congress on Intelligent Control and Automation, pp.1317–1322, Changsha, China, July 2018.
27. A. James, J. Wang, X. Yang et al., "TrackNet- a deep learning based fault detection for railway track inspection," in Pro ceedings of 2018 International Conference on Intelligent Rail Transportation, pp. 1–5, Singapore, Singapore, December 2018.
28. Jaisingh Bhandari, "railway-track Dataset," Roboflow Universe, Roboflow, Sep 2021. Available: https://universe.roboflow.com/jaisingh-bhandari/railway-track.
29. Railway, "ProjectData 2 Dataset," Roboflow Universe, Roboflow, Dec 2023. Available: https://universe.roboflow.com/railway/projectdata-2
30. project-v9fbe, "caronrial 2 Dataset," Roboflow Universe, Roboflow, Dec 2023. Available: https://universe.roboflow.com/project-v9fbe/caronrial-2
31. IIIT Naya Raipur, "Crack_Segmentation Dataset," *Roboflow Universe*, Roboflow, Oct. 2023. Available: https://universe.roboflow.com/iiit-naya-raipur-ajytz/crack_segmentation-kxa7r
32. A. Buslaev, A. Parinov, E. Khvedchenya, V.I. Iglovikov and A.A. Kalinin, "Albumentations: fast and flexible image augmentations," arXiv preprint arXiv:1809.06839, 2018
33. Goutte, C., & Gaussier, E. (2005, March). A probabilistic interpretation of precision, recall and F-score, with implication for evaluation. In *European conference on information retrieval* (pp.345-359). Springer, Berlin, Heidelberg