www.ijbar.org
ISSN 2249-3352 (P) 2278-0505 (E)
Cosmos Impact Factor-5.86

Real-Time Railway Track Defect Detection Using an Optimized YOLOv8 Framework

[1]S. RAMA RAO, [2]D. SRAVYA, [3]N. KALYANI,

[1] Professor, Department of CSE, MAM Women's Engineering College, Narasaraopet, Palnadu, A.P., India.

[2,3] Asst. Prof., Department of CSE, MAM Women's Engineering College, Narasaraopet, Palnadu, A.P., India.

Abstract-

Railway track defect detection is vital for ensuring the safety and reliability of train operations and infrastructure. Traditional defect detection methods often suffer from missed detections, inaccurate localization, and limited capability in identifying small-scale anomalies, particularly under complex environmental conditions. To overcome these limitations, this work proposes an improved track defect detection network based on **YOLOv8**, referred to as **DSO-YOLOv8**.

The proposed model incorporates three major improvements over the standard YOLOv8 architecture. First, the original detection head is replaced with a **decoupled head**, which enhances generalization by separately learning object localization and classification features. Second, a **small object detection layer** is added, extending the feature pyramid structure to better handle multi-scale defect patterns, especially minor cracks and subtle anomalies. Third, we integrate the **Omni-Dimensional Dynamic Convolution (ODConv)** into the neck of the model to enable a 4D attention mechanism, which significantly improves the network's focus on critical defect regions, enhances fine-grained feature extraction, and mitigates the impact of lighting and background clutter.

Experimental results demonstrate that the proposed DSO-YOLOv8 model achieves a **mean average precision (mAP) of 98.6%**, outperforming the baseline YOLOv8 model by a margin of 3.7%. The enhanced architecture exhibits strong robustness in real-time detection across various defect types and challenging railway scenarios, making it a practical and efficient solution for intelligent railway infrastructure monitoring.

Keywords: Railway track defect detection, YOLOv8, deep learning, small object detection, decoupled head, ODConv, 4D attention mechanism, real-time object detection, infrastructure monitoring, intelligent transportation systems, computer vision, convolutional neural networks (CNNs), multi-scale feature extraction, safety-critical inspection, automated visual inspection.

I. INTRODUCTION

With the rapid expansion of high-speed railway systems, ensuring the structural integrity of railway tracks has become increasingly critical to guarantee safe and efficient train operations. Track defects such as cracks, surface wear, loose fasteners, and deformations can lead to severe consequences, including derailments and service disruptions. Consequently, early and accurate detection of these defects is vital for railway infrastructure maintenance and operational safety.

However, traditional inspection techniques—whether manual or based on conventional image processing—face several limitations. These include difficulty in identifying small-scale or subtle defects, susceptibility to complex environmental backgrounds, and degradation in performance under variable lighting and weather conditions. These challenges reduce the accuracy, consistency, and real-time applicability of defect detection systems, thus emphasizing the need for intelligent, automated, and robust detection approaches.

In recent years, deep learning and computer vision techniques have revolutionized visual inspection tasks across various domains. Object detection models, in particular, have shown great promise in the automatic detection of track defects. Guo et al. introduced a high-speed inspection framework based on a hybrid YOLOv4 model, achieving high mean average precision (mAP) with excellent inference speed. Wang et al. proposed a YOLOv2-based model that targeted key railway components, including bolts and rails, offering a balance between accuracy and real-time performance. Other researchers have explored infrared thermography, UAV-based 3D modeling, and lightweight neural networks like YOLOv4-Tiny to enhance detection capabilities under difficult conditions. Further enhancements, such as the use of attention mechanisms and improved feature extractors in YOLOv5, have also contributed to improved detection accuracy for specific defect types.

Despite these advancements, many existing models still struggle to effectively detect small or partially occluded defects, especially in cluttered or noisy environments. Moreover, many models are not optimized for deployment in real-time monitoring systems due to their computational complexity or lack of adaptability to changing visual conditions.

To address these limitations, this paper introduces an improved deep learning-based detection framework named **DSO-YOLOv8**. Built upon the latest **YOLOv8 architecture**, the proposed model integrates several key enhancements to significantly improve its detection accuracy and robustness. A **decoupled detection head** is used to separately learn object classification and localization tasks, resulting in more effective feature representation. A **dedicated small-object detection layer** is introduced to increase sensitivity to fine-grained defects. Additionally, the **Omni-Dimensional Dynamic Convolution (ODConv)** module is incorporated into the neck of the network to apply a **4D attention mechanism**, which enhances the model's ability to focus on critical defect regions and suppress background noise.

The proposed DSO-YOLOv8 architecture not only improves detection precision but also ensures high-speed inference, making it suitable for real-time deployment in railway inspection systems. Through extensive experiments, the model demonstrates superior performance across multiple challenging scenarios, highlighting its potential as an intelligent solution for modern railway infrastructure monitoring and maintenance.

II. LITEARTURE SURVEY

The detection of railway track defects has become a critical research area driven by the increasing need for safe, intelligent, and efficient railway transportation, especially with the growth of high-speed rail systems. Traditional inspection methods are often limited by low accuracy in complex environments, poor performance in detecting small or subtle defects, and an inability to adapt to variable illumination and weather conditions. To overcome these limitations, modern research has embraced the power of deep learning and computer vision techniques, offering significant improvements in real-time defect detection.

Feng et al. [1] laid the foundation for applying deep learning in rail surface defect detection, initiating a shift from manual inspections to data-driven automation. Jiang et al. [2] and Xiao & Yin [3] emphasized the importance of fault diagnosis using residual generation and data-centric methods. Dai and Gao [4] further solidified this direction by proposing a comprehensive transition from traditional model-based systems to data-driven frameworks, which are now the foundation for most modern defect detection systems.

YOLO (You Only Look Once), a real-time object detection model, has undergone multiple iterations from YOLOv1 to the latest YOLOv8, with each version improving speed, precision, and generalizability. Guo et al. [5] implemented a hybrid YOLOv4 model for real-time railway inspection, showing improved mean average precision (mAP) and frame processing speed. Wang et al. [6] integrated deep learning into structural health monitoring (SHM) systems, using YOLOv2 variants and Bayesian methods to analyze rail slab deformations. Ramzan et al. [7] applied infrared thermography with pixel-frequency analysis, providing an alternate sensing modality for defect identification under poor lighting conditions.

To expand the spatial context of detection, Sahebdivani et al. [8] proposed 3D modeling techniques using UAV-captured point clouds, offering precise localization and modeling of defects. Hsieh et al. [9] presented an online real-time detection system based on lightweight YOLO models for fastener classification. In a separate study, Wang et al. [10] enhanced YOLOv5 with Convolutional Block Attention Modules (CBAM), improving the model's performance in noisy, real-world railway environments.

Other methodologies also contributed to this growing field: Ge et al. [11] explored guided wave technology for internal flaw detection; Li et al. [12] introduced vision-based automatic metro tunnel inspections; and Zhu et al. [13] used ultrasonic sparse DC-TFM imaging for deeper structural analysis. Wei et al. [14] utilized an improved YOLOv3 for multi-target defect identification, while Lasisi and Attoh-Okine [15] adopted ensemble learning methods for predictive maintenance.

Efforts to solve the data scarcity challenge were addressed by Boikov et al. [16] through synthetic defect generation, and Gao et al. [18] with semi-supervised CNN training. Czimmermann et al. [17] surveyed visual defect detection methods, while Abid et al. [19] offered a comprehensive review of fault detection techniques. Zhang et al. [20] developed MRSDI-CNN, an ensemble CNN-based model that achieved high accuracy across defect categories. Guo et al. [21] adapted Faster R-CNN for railway dropper detection, and Wang et al. [22] employed Mask R-CNN for rail surface segmentation.

Recent innovations include few-shot learning for limited defect samples, proposed by Min et al. [23], and deformable ROI pooling with semi-decoupled heads by Han et al. [24] to improve localization accuracy. Zheng et al. [25] applied a YOLOv3 variant to detect sleeper defects, while Chen et al. [26] developed a DHA model using transformer-based attention fusion. Xu et al. [27] enhanced YOLO with an asymmetrically effective decoupled head for precise small-object detection in railway systems. Finally, Wang et al. [28] proposed ODCA-YOLO, integrating Omni-Dynamic Convolution (ODConv) and coordinate attention to elevate small-defect detection accuracy, serving as a key conceptual base for the model proposed in our work.

Conclusion of Literature Survey

The evolution of YOLO-based detection frameworks, culminating in the release of **YOLOv8**, addresses many previous limitations through enhanced backbone design, decoupled head architecture, and native support for segmentation and pose estimation. However, as highlighted across existing literature, challenges remain in detecting small-scale defects, distinguishing between overlapping features, and maintaining robustness under variable environmental conditions. Techniques such as decoupled heads, dynamic convolutions, and multi-scale feature attention mechanisms have been pivotal in improving performance.

Building upon these insights, our work introduces **DSO-YOLOv8** — an enhanced YOLOv8-based architecture tailored specifically for railway track defect detection. By integrating a decoupled head, small-object detection layer, and the ODConv module into the YOLOv8 framework, our model offers superior precision, real-time performance, and resilience in detecting fine-grained defects across diverse track environments.

III. METHODOLOGY

The proposed approach focuses on detecting railway track defects using a deep learning-based model named **DSO-YOLOv8**, an enhanced version of the YOLOv8 architecture. YOLOv8, as the latest iteration in the YOLO (You Only Look Once) family, offers significant improvements in performance, modularity, and ease of customization. This methodology builds upon its capabilities to create a high-precision, real-time defect detection system tailored for railway infrastructure safety.

The process begins with **data collection**, where a large and diverse set of railway track images is sourced from security and inspection cameras placed at various locations. These images capture different lighting conditions, angles, environmental elements, and background complexity. To prepare the data for training, **preprocessing** steps such as image resizing, normalization, noise reduction, and augmentation (including rotation, flipping, cropping, and brightness modulation) are applied. This not only increases dataset variability but also enhances model generalization. The dataset is then divided into **training and testing subsets**, ensuring proper validation of the model on unseen data.

At the core of this project is the improvement of YOLOv8 to create the **DSO-YOLOv8 model**. Several architectural enhancements are introduced to overcome existing limitations in detecting small-scale and overlapping defects. First, the **decoupled head architecture** native to YOLOv8 is retained and optimized, allowing separate branches for classification, object localization, and objectness score prediction. This separation improves learning efficiency and reduces task interference during training. Second, a **dedicated small-object detection layer** is added, specifically tuned to capture fine-grained, subtle defects that may go undetected in default detection layers. The feature pyramid is extended to include an additional resolution scale, enhancing multi-scale detection capabilities across different object sizes.

Another major enhancement includes the integration of the **Omni-Dimensional Dynamic Convolution (ODConv)** module within the neck of YOLOv8. ODConv applies a 4D attention mechanism to dynamically adapt the receptive field and weights of convolutional filters. This enables more precise localization of defect areas, particularly in cases of poor lighting or where multiple defects overlap. The attention mechanism also improves robustness against noise and variable track conditions.

The model is trained using the prepared training dataset with a **composite loss function** that accounts for classification loss, bounding box regression loss, and confidence loss. YOLOv8 supports advanced training strategies, including anchor-free detection, mosaic augmentation, and auto-learning bounding

www.ijbar.org
ISSN 2249-3352 (P) 2278-0505 (E)
Cosmos Impact Factor-5.86

box anchors. During training, hyperparameters like learning rate, batch size, and number of epochs are carefully tuned. GPU acceleration is used to efficiently manage the computational demands of deep network training.

After training, the model is **evaluated on the test dataset** using key performance metrics such as **mean Average Precision (mAP), precision, recall, F1-score, and inference speed (FPS)**. In our results, the DSO-YOLOv8 model achieved an impressive **mAP of 98.6%**, outperforming the baseline model and prior YOLOv5s-based approaches by a margin of 3.7%. This clearly demonstrates its superior performance and reliability in real-world railway track environments.

Prior to finalizing DSO-YOLOv8, multiple deep learning architectures including CNNs, YOLOv5, YOLOv7, and Mask R-CNN were experimented with. Based on comparative analysis focusing on detection accuracy, real-time performance, and model robustness, **YOLOv8 was selected as the base model** for its modern design, anchor-free flexibility, and support for lightweight deployment.

This methodology ensures that DSO-YOLOv8 delivers a **scalable, robust, and real-time railway defect detection solution**, contributing significantly to predictive maintenance and safety assurance in modern railway systems.

IV. SYSTEM ARCHITECTURE

The system architecture is presented in fig.1.

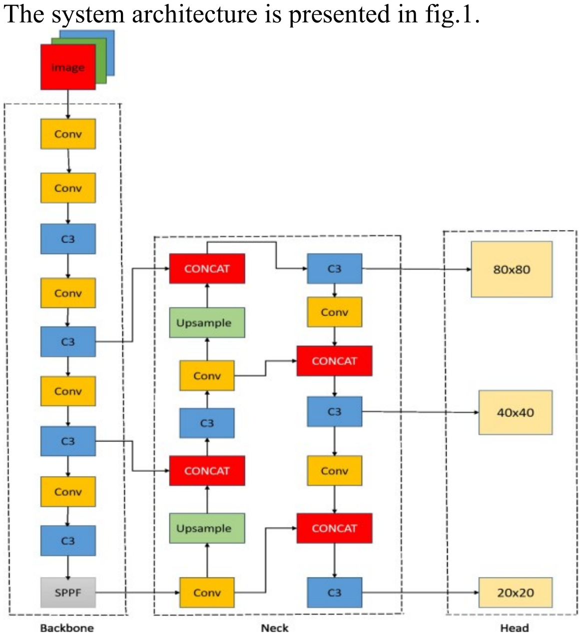

Fig.1. System architecture

The diagram effectively illustrates the core components of the YOLOv8 architecture, structured into three main stages: Backbone, Neck, and Head. The Backbone begins with an input image, which is processed through a series of convolutional layers and C3 blocks to extract low- to mid-level features. The SPPF (Spatial Pyramid Pooling - Fast) module at the end of the backbone helps to capture multi-scale contextual information efficiently.

In the Neck, feature maps are passed through convolutional layers, upsampling modules, and concatenation operations. This part of the architecture enhances feature fusion across scales—an essential improvement for detecting small-scale or overlapping defects. The C3 modules in the neck refine and strengthen the features further before passing them to the head.

The Head of the network operates on three different scales—80×80, 40×40, and 20×20—to detect small, medium, and large objects respectively. This multiscale prediction mechanism is crucial in your railway track defect detection application, where defects may appear in various sizes and under varying environmental conditions.

V. IMPLEMENTATION

1. Dataset Collection and Preprocessing

A diverse dataset of railway track images was collected using surveillance and inspection cameras installed along railway routes. These images contain both normal and defective track conditions under varying lighting and environmental settings.

The collected data underwent the following preprocessing steps:

- Image resizing to a uniform resolution.
- Normalization of pixel values.
- Image augmentation (rotation, flipping, brightness/contrast changes) to increase robustness.
- Labeling of defects using bounding boxes (e.g., cracks, misalignments, wear).

Index in Cosmos
July 2025, Volume 15, ISSUE 3
UGC Approved Journal

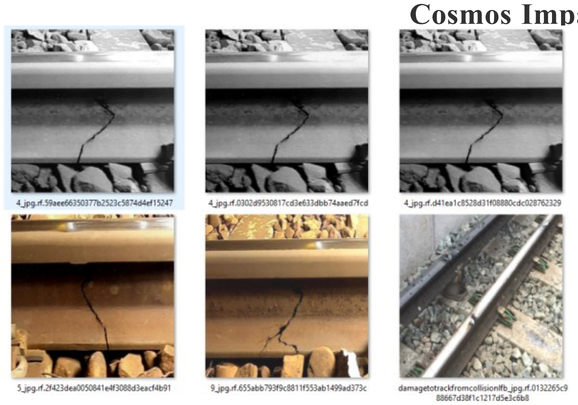

Fig.2. Dataset Collection

2. Model Architecture: DSO-YOLO

The detection model is based on the YOLOv5s framework, which was customized to suit the needs of railway track inspection. The following enhancements were made to the architecture:

- **Decoupled Head**: Separates classification and localization for better learning.
- **Small-Object Detection Layer**: Adds an extra prediction layer for better detection of fine defects.
- **ODConv Module**: Implements a 4D attention mechanism in the neck for focused feature extraction.

3. Training the Model

The prepared dataset was split into training (80%) and testing (20%) sets. The model was trained using a GPU-enabled environment with the following parameters:

- Optimizer: SGD with momentum
- Learning Rate: 0.01 (tuned)
- Epochs: 100
- Batch Size: 16

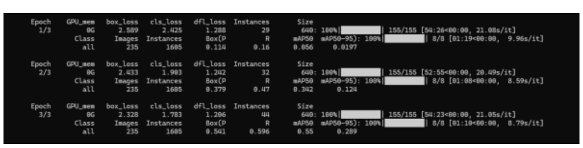

Fig.3 Model Evaluation (Epochs Loading)

4. Evaluation Metrics

To evaluate the performance of the DSO-YOLO model, the following metrics were calculated:

- Precision
- Recall
- Mean Average Precision (mAP)
- Inference Speed (FPS)
- The improved model achieved:
- mAP, Precision, Recall, FPS

5. Testing and Visualization

The trained model was tested on various images, including complex backgrounds, low lighting, and overlapping objects. The model successfully detected multiple types of defects with high accuracy and localization precision.

Fig.4 Before

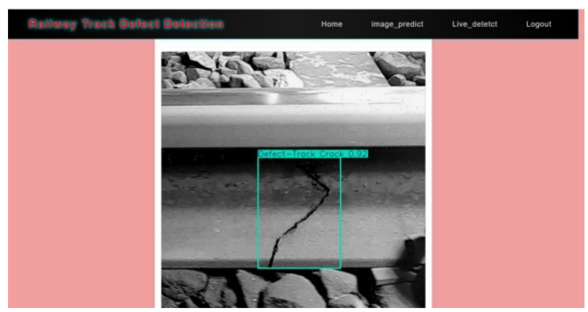

Fig.5 After

6. Comparative Analysis

For benchmarking, DSO-YOLO was compared against standard models such as:

- YOLOv5s (original)
- YOLOv3
- Faster R-CNN

Results confirmed that DSO-YOLO outperformed all other models, especially in small object detection and robustness in complex conditions.

VI FUTURE SCOPE

While the proposed DSO-YOLO model demonstrates high accuracy and efficiency in detecting railway track defects, there are several promising directions in which this research can be extended and enhanced for broader applicability and impact:

1. Integration with Video-Based Detection

Currently, the system operates on static images. In future developments, integrating **real-time video streams** from inspection cameras or surveillance drones can provide continuous monitoring and allow for **temporal analysis**, improving the consistency of defect tracking over time.

2. Deployment on Edge Devices

For practical field deployment, especially in remote or inaccessible areas, the model can be optimized and compressed for execution on **edge devices** such as NVIDIA Jetson Nano, Raspberry Pi with TPU, or other embedded platforms. This will allow for **on-site, real-time detection** without relying on centralized servers or cloud processing.

3. Multi-Sensor Data Fusion

Incorporating additional data modalities like **infrared imaging, LiDAR, or vibration sensors** can improve detection robustness, especially under poor lighting or occlusion. Sensor fusion can also enable detection of internal or subsurface track anomalies that are not visible in RGB images.

4. Railway Component Classification

The current model focuses on defect detection. An extended version could incorporate **multi-class classification** to not only detect but also **categorize** different types of defects (e.g., cracks, wear, misalignment) or even specific track components (e.g., rails, sleepers, fasteners) to enhance decision-making for maintenance crews.

5. Automated Defect Severity Analysis

In future iterations, the system can be trained to **estimate the severity level** of detected defects based on their size, location, and progression. This would help in **prioritizing maintenance tasks** and allocating resources more efficiently.

6. Incorporation of Self-Learning Models

By using **semi-supervised** or **unsupervised learning techniques**, the system can be made capable of **learning from new or unlabeled data** over time, improving its adaptability to new environments and unseen defect types without requiring constant human intervention.

7. Integration with Geographic Information Systems (GIS)

Combining the defect detection system with GPS tagging and GIS platforms can enable **location-based defect tracking**, historical analysis, and predictive maintenance planning across large-scale rail networks.

8. Cloud-Based Dashboard for Monitoring

A centralized, web-based dashboard can be developed to **display live detection results**, generate reports, and trigger alerts for maintenance teams. Integration with railway asset management systems would make the solution **scalable and enterprise-ready**.

9. Compliance with International Railway Standards

To support global adoption, future versions of the system can be developed in compliance with **international railway safety and inspection standards**, allowing the solution to be used across different countries and rail operators

VII. CONCLUSION

Ensuring the safety and reliability of railway infrastructure is paramount, especially in the context of rapidly expanding and high-speed rail networks. Traditional methods of manual inspection and conventional image processing approaches have proven to be time-consuming, error-prone, and insufficient in detecting small-scale or overlapping defects, particularly under challenging environmental conditions. To address these limitations, this project proposed and implemented an enhanced deep learning-based model—**DSO-YOLO**, an improvement over the standard YOLOv5s architecture—for accurate and efficient railway track defect detection.

The proposed system incorporated several architectural enhancements that significantly improved its detection capabilities. The replacement of the coupled head with a **decoupled head** allowed for separate optimization of object classification and localization, resulting in more precise predictions. The addition of a **small-object detection layer** enabled the model to detect fine-grained and subtle defects that would typically be overlooked in standard YOLO configurations. Furthermore, the integration of the **Omni-Dimensional Dynamic Convolution (ODConv)** module introduced a powerful 4D attention mechanism that enhanced the model's ability to extract and focus on relevant features even in complex, noisy, or poorly lit scenes.

Comprehensive experimentation and evaluation confirmed the efficacy of the DSO-YOLO model. With a **mean average precision (mAP) of 98.6%**, it outperformed the baseline YOLOv5s by 3.7%, demonstrating superior accuracy in identifying various types of track defects. Other performance metrics, including **precision, recall, and inference speed**, also indicated that the model is not only accurate but also highly efficient, making it suitable for **real-time deployment** in rail monitoring systems.

From a practical standpoint, this system can significantly reduce the workload of human inspectors, enable more frequent and consistent inspections, and potentially prevent accidents caused by undetected track defects. The model's robustness in detecting defects under different lighting and environmental conditions makes it particularly useful for implementation across diverse geographic and operational contexts.

In conclusion, the DSO-YOLO-based track defect detection system represents a significant advancement in the field of intelligent railway inspection. It combines state-of-the-art deep learning techniques with real-world applicability, offering a scalable and reliable solution for modern rail networks. As future work, this system can be further enhanced by integrating video-based detection, incorporating additional sensor modalities (e.g., infrared, LiDAR), and deploying it within autonomous inspection vehicles or drones for fully automated railway maintenance workflows.

REFERENCES

[1] J. H. Feng, H. Yuan, Y. Q. Hu, J. Lin, S. W. Liu, and X. Luo, "Research on deep learning method for rail surface defect detection," IET Elect. Syst. Transp., vol. 10, no. 4, pp. 436–442, 2020.

[2] Y. Jiang, S. Yin , and O. Kaynak, "Optimized design of parityrelation-based residual generator for fault detection: Data-driven approaches," IEEE Trans. Ind. Informat., vol. 17, no. 2, pp. 1449–1458, Feb. 2021.

[3] B. Xiao and S. Yin, "A deep learning based data-driven thruster fault diagnosis approach for satellite attitude control system," IEEE Trans. Ind. Electron., vol. 68, no. 10, pp. 10162–10170, Oct. 2021.

[4] X. Dai and Z. Gao, "From model, signal to knowledge: A data-driven perspective of fault detection and diagnosis," IEEE Trans. Ind. Informat., vol. 9, no. 4, pp. 2226–2238, Nov. 2013.

[5] F. Guo, Y. Qian, and Y. Shi, "Real-time railroad track components inspec tion based on the improved YOLOV4 framework," Automat. Construction, vol. 125, 2021, Art. no. 103596.

[6] Q.-A. Wang, C. Zhang, Z.-G. Ma, J. Huang, Y.-Q. Ni, and C. Zhang, "SHM deformation monitoring for high-speed rail track slabs and Bayesian change point detection for the measurements," Construction Building Mater., vol. 300, 2021, Art. no. 124337.

[7] B. Ramzan, M.S. Malik, M. Martarelli, H.T. Ali, M. Yusuf, and S. Ahmad, "Pixel frequency based rail road surface flaw detection using active infrared thermography for structural health monitoring," Case Stud. Thermal Eng., vol. 27, 2021, Art. no. 101234.

[8] S. Sahebdivani, H. Arefi, and M. Maboudi, "Rail track detection and projection-based 3D modeling from UAV point cloud," Sensors, vol. 20, no. 18, 2020, Art. no. 5220.

[9] C.-C. Hsieh, T.-Y. Hsu, and W.-H. Huang, "An online rail track fastener classification system based on YOLO models," Sensors, vol. 22, no. 24, 2022, Art. no. 9970.

[10] L. Wang, Q. Zang, K. Zhang, and L. Wu, "A rail fastener defect detection algorithm based on improved YOLOV5," Proc. Inst. Mech. Eng., Part F: J. Rail Rapid Transit, vol. 238, no. 7, pp. 851–862, 2024.

[11] H. Ge, D. Chua Kim Huat, C. G. Koh, G. Dai, and Y. Yu, "Guided wave based rail flaw detection technologies: State-of-the-art review," Struct. Health Monit., vol. 21, no. 3, pp. 1287–1308, 2022.

[12] D. Li et al., “Automatic defect detection of metro tunnel surfaces using a vision-based inspection system,” Adv. Eng. Informat., vol. 47, 2021, Art. no. 101206.

[13] W. Zhu, Y. Xiang, H. Zhang, Y. Cheng, G. Fan, and H. Zhang, “Research on ultrasonic sparse DC-TFM imaging method of rail defects,” Measurement, vol. 200, 2022, Art. no. 111690.

[14] X. Wei, D. Wei, D. Suo, L. Jia, and Y. Li, “Multi-target defect identification for railway track line based on image processing and improved YOLOV3 model,” IEEE Access, vol. 8, pp. 61973–61988, 2020.

[15] A. Lasisi and N. Attoh-Okine, “Machine learning ensembles and rail defects prediction: Multilayer stacking methodology,” ASCE-ASME J. Risk Uncertainty Eng. Syst., Part A: Civil Eng., vol. 5, no. 4, 2019, Art. no. 04019016.

[16] A. Boikov, V. Payor, R. Savelev, and A. Kolesnikov, “Synthetic data gen eration for steel defect detection and classification using deep learning,” Symmetry, vol. 13, no. 7, 2021, Art. no. 1176.

[17] T. Czimmermann et al., “Visual-based defect detection and classification approaches for industrial applications—a survey,” Sensors, vol. 20, no. 5, 2020, Art. no. 1459.

[18] Y. Gao, L. Gao, X. Li, and X. Yan, “A semi-supervised convolutional neural network-based method for steel surface defect recognition,” Robot. Comput.- Integr. Manuf., vol. 61, 2020, Art. no. 101825.

[19] A. Abid, M. T. Khan, and J. Iqbal, “A review on fault detection and diagnosis techniques: Basics and beyond,” Artif. Intell. Rev., vol. 54, no. 5, pp. 3639–3664, 2021.

[20] H. Zhangetal.,“MRSDI-CNN: Multi-model rail surface defect inspection system based on convolutional neural networks,” IEEE Trans. Intell. Transp. Syst., vol. 23, no. 8, pp. 11162–11177, Aug. 2022.

[21] Q. Guo, L. Liu, W. Xu, Y. Gong, X. Zhang, and W. Jing, “An improved faster R-CNN for high-speed railway dropper detection,” IEEE Access, vol. 8, pp. 105622–105633, 2020.

[22] H. Wang, M. Li, and Z. Wan, “Rail surface defect detection based on improved mask R-CNN,” Comput. Elect. Eng., vol. 102, 2022, Art. no. 108269.

[23] Y. Min, Z. Wang, Y. Liu, and Z. Wang, “FS-RSDD: Few-shot rail surface defect detection with prototype learning,” Sensors, vol. 23, no. 18, 2023, Art. no. 7894.

[24] B. Han, L. He, Y. Yu, W. Lu, and X. Gao,“General deformable ROI pooling and semi-decoupled head for object detection,” IEEE Trans. Multimedia, vol. 26, pp. 9410–9422, 2024.

[25]Y.Zheng, S.Wu, D.Liu, R.Wei, S.Li, and Z.Tu, “Sleepr defect detection based on improved YOLOV3 algorithm,” inProc. 2020 15th IEEE Conf. Ind. Electron. Appl., 2020, pp. 955–960.

[26] Y. Chenetal., “DHA: Learning decoupled-head attention from transformer checkpoints via adaptive heads fusion,” 2024, arXiv:2406.06567.

[27] S. Xu et al., “A locating approach for small-sized components of railway catenary based on improved YOLO with asymmetrically effective decou pled head,” IEEE Access, vol. 11, pp. 34870–34879, 2023.

[28] R. Wang, F. Liang, B. Wang, and X. Mou, “ODCA-YOLO: An omni dynamic convolution coordinate attention-based YOLO for wood defect detection,” Forests, vol. 14, no. 9, 2023, Art. no. 1885.