Publisher

Vandana Publications
www.vandanapublications.com

International Journal of Engineering and Management Research

2025 Volume 15 Number 2 April

E-ISSN:2250-0758
P-ISSN:2394-6962

Research Article

Deep Learning

Railroad Track Defect Detections using Deep Learning

Anita[1*]

DOI:10.5281/zenodo.15356939

[1*] Anita, Assistant Professor, Department of Computer Science and Engineering, Shri Ram College of Engineering and Management, Palwal, India.

However, railroad transportation continues to be one of the most important components to global terms of freight and passenger transport, and is integral to the success of economic development and logistics efficiency. With bigger demand for rail service comes ever more need to not only maintain functionality of the rail system, but also to keep it safe and reliable. With railroad tracks being the case, one of the most pressing concerns at hand is to be able to detect defects in the tracks early and accurate. Due to labor intensive, time consuming and human error prone nature of traditional inspection methods, there is an immediate requirement of approaches to automated, intelligent inspection. Using high resolution inspection images of railroad tracks, this study investigates the use of deep learning techniques for automation in detecting defects in railroad tracks. Then based on the leakage of convolutional neural networks (CNNs), transfer learning strategies and real time object detection frameworks (e.g., YOLOv5), provide the good enough track anomaly classification and locating capabilities such as cracks, broken rail, loose fasteners and vegetation interference. Experimental results show that the classification accuracy was 94% for InceptionV3 model, and the real time detection using YOLOv5 has the mAP of 0.89 and the inference time of just 23ms per image. Deep learning appears to be a robust scalable and practical method of improving railway safety and maintenance operation.

Keywords: Deep Learning, Railroad Transportation, Convolutional Neural Networks (CNNs), YOLOv5

Corresponding Author	How to Cite this Article	To Browse
Anita, Assistant Professor, Department of Computer Science and Engineering, Shri Ram College of Engineering and Management, Palwal, India. Email: anitasorout4@gmail.com	Anita, Railroad Track Defect Detections using Deep Learning. Int J Engg Mgmt Res. 2025;15(2):89-96. Available From https://ijemr.vandanapublications.com/index.php/j/article/view/1735	

Manuscript Received	Review Round 1	Review Round 2	Review Round 3	Accepted
2025-03-03	2025-03-27			2025-04-19
Conflict of Interest	**Funding**	**Ethical Approval**	**Plagiarism X-checker**	**Note**
None	Nil	Yes	6.33	

1. Introduction

1.1 Background and Motivation

Railroad infrastructure is an essential component of national and international transportation networks; it aids in fostering the development of the national and international economies and the ability to move people and goods. Railways are one of the most cost-effective and environmentally friendly modes of transportation upon which ever increasing amounts of global trade and urban development are increasingly relying. Whilst the growing dependence on rail systems also increases expectations of safety, reliability and efficiency (Akabal et al., 2017), however.

Regular inspection of track infrastructure is a critical part of railway safety. Mechanical stress, exposure to the environment, and material degradation of railroad track is always applied. Cracks, broken rails, loose fasteners, misalignments and encroachments such as vegetation can cause derailments, service delays or deadly accidents. For traditional practice, railway track inspection was mostly depending on human inspection or manual tool (Chandran et al., 2019). Although these methods are a little bit effective to an extent, but they are labor intensive and time consuming and are also prone to inconsistencies and human error by fatigue due to environmental conditions..

Figure 1: Enhancing Railway Safety and Efficiency

However, in the recent years, it has been witnessed that high technologies such as computer vision, machine learning, and artificial intelligence (AI) have offered an eventual alternative to automate the inspection. Currently it's possible to detect and classify defects at a faster and more accurate rate than previously, through the use of deep learning models which can handle (and analyze) large quantities of visual data. Such intelligent systems have the potential to be integrated into the track monitoring procedures, which will be transformative towards data driven, proactive, maintenance strategies that save on long term maintenance costs and improve operational safety (Banić et al., 2019).

1.2 Problem Statement

Railway systems are critical infrastructure assets whose safety and reliability are of the greatest importance and constant and careful maintenance that can last for a long time. In traditional practice, human inspectors mounted specialized vehicles to ride them along tracks so that the inspector can visually observe various conditions or walk along or on top of the tracks (Edwards et al., 2009). While the historically effective degree to which these methods were applied to some extent, in the light of today's expanded and more complicate rail networks these methods have some limitations. The process is labor intensive, requires a lot of time, very prone to human error and fatigue, and is liable to generate missed or inaccurate defect identification. In addition, the need for high speed rail, as well as the increase in the volume of transport, has increased the pressure on the railways, and therefore more frequent and more accurate monitoring of the rail (Kliuiev et al., 2020). These demands are simply beyond the ability of the human eye to manually inspect in scale, and automated methods are hardly better. That is why there is a great need for an automated, scalable, and intelligent solution to reliably and in time detect and classify track defects. Deep learning, with its proven success in image recognition and anomaly detection, offers a powerful alternative (Teshaev et al., 2024). However, the advent of deep neural networks, especially convolutional architectures, has offered railways operators a mechanism to modify maintenance procedures to deliver proactive repairs, decrease operational downtime, cut costs, and foremost, increase the safety and reliability of railway systems (Yang et al., 2020).

1.3 Objectives

- Develop a deep learning model for automatic detection of railroad track defects.
- Evaluate the model performance using real-world datasets.
- Analyze the feasibility of deploying the model in real-time scenarios.

2. Related Work

In previous work, most effort has been spent in using traditional machine learning techniques and heuristic image processing techniques to detect railroad track defects.

Specific use cases which have been addressed by these methods, include the use of edge detection, morphological operations and handcrafted feature extraction, followed by support vector machines or decision trees. Yet, they often depend on a heavy use of domain expertise for the feature engineering and they are rarely able to generalize over various environmental and operational conditions (Kaparthi & Bumblauskas, 2020).

Conversely, deep learning generally convolutional neural networks, or CNNs, have intrinsically revolutionized image based classification and detection tasks. CNNs provide impressive results in medical imaging, autonomous driving and structural health monitoring, due to their inherent ability for automatic learning of hierarchical feature representations from raw input data (Τρίγκα & Δρίτσας, 2025). As CNNs offer these advantages, they are extremely well suited to complex defect detection in railroad track system problems where very subtle information and contextual information is key. Unity with deep learning has been found to be most suitable in applications such as rail infrastructure fault diagnosis in rail components, anomaly detection based on unmanned aerial vehicles (UAV) and track geometry analysis (Bai et al., 2021). While there have been these developments, there have, relatively few, studies specifically looking at visual inspection of rail tracks using deep learning models. However, among the others who have implemented similar systems, limitations are found in small datasets, lack of real time processing capabilities, as well as the difficulty of accurately recognising rare or subtle defect types. As such, deep learning based solutions hold great potential to be adapted to uniquely solve dynamic operational environments that characterize railroad track defect detection (Wu et al., 2021).

3. Methodology

3.1 Dataset Collection

To perform this study, we use a facets dataset, including more than 20,000 high-resolution images of railroad tracks that is publicly available. The dataset is presented in such a way that each image in the dataset is carefully and precisely labeled with defects such as 'crack', 'broken rail', 'loose fastener', and 'vegetation on track'. Some of the most common and critical types of structural anomalies have been put into these categories.

The dataset is also very diverse: under different environmental condition (bright sunlight, shadow, rain, fog) and so helps robustness and generalizability of the models on the dataset. Secondly, also present in the provided dataset are various geographic regions of origin, thus assuring that variations in the rail construction, material wear, and environmental impact are reflected. For training deep learning models that can reliably generalize over real world scenarios, it is critical to have this diversity. Precise annotations enable better quality of supervised learning and consequently higher model accuracy.

3.2 Data Preprocessing

Raw image data is a crucial data source and requires to be fed into the deep learning model, that’s why preprocessing is an integral step. This is the reason why each image in this study was resized to a single resolution of 224x224 pixels in order to match the input dimension that standard CNN architectures expected. Normalized pixel values to a range between 0 and 1 made the learning process faster to converge as it stabilized the learning process. Data augmentation techniques like rotating random y, horizontal flipping, vertical flipping, zooming were added to the model to further make the model more robust and less biased. Through the establishment of these simulation conditions that resemble different real world conditions the model is made to generalize to unseen data better.

3.3 Model Architecture

In order to evaluate the utility of deep learning in railroad track defect detection, we used Convolutional neural network (CNN) based architectures including VGG16, ResNet50, InceptionV3 and YOLOv5 to test their performance. They chose these models based on their past record for image classification and object detection tasks in different domains. As a very strong baseline, we used the deep yet very simple architecture of VGG16. The residual connections are introduced by ResNet50 to mitigate the vanishing gradient problem such that deeper networks are able to learn effectively. As an optimization between the accuracy and computational efficiency, inceptionV3 used the asymmetric convolutions and the factorized filters. As a real time detection model, it was decided to use YOLOv5, a state of the art object detection model.

However, all classification models were fine tuned from ImageNet pre trained weights using transfer learning. To produce predictions for four specific defect categories, cracks, broken rails, loose fasteners, and vegetation obstructions, the final classification layers were modified.

3.4 Training Configuration

Two prominent frameworks TensorFlow and PyTorch were used to train the deep learning models so they can stay compatible and for comparison in performance. The two frameworks are in constant use in the research community due to their wide support of advanced neural networks and flexibility in constructing custom architectures. It took a batch size of 32 to enable fast training speed while useful memory utilization. With this batch size, the training does not overwhelm the GPU memory during processing, but is also efficient so training performance remains stable when changing architectures. The learning rate was chosen to be 0.0001, a value that is used to fine tune a pre-trained model while preventing a large gradient from falling, lest it result in training instability. We chose the Adam optimizer as it does facilitate a adaptive learning rate which helps speed up convergence during training. For classification tasks Categorical Crossentropy was used as the loss function and for object detection tasks where you want to localize track defects a more appropriate loss function is Binary Crossentropy so that you can better model presence or absence of the object in the image. For the models to learn from the data and avoid over fitting, training was done for 50 epochs.

3.5 Evaluation Metrics

For evaluating the performance of deep learning models for the detection of railroad track defects, a comprehensive set of evaluation metrics were used in this study. An accuracy is a fundamental one that determines the number of correctly classified instances among all instances forming a general overview of the model's efficiency. However, one must take great caution when dealing with imbalanced datasets where one class is orders of magnitude less than others as accuracy is not indicative enough. Therefore, precision was also employed, as precision represents the ratio of true positive predictions to all positive predictions.

In defect detection it is particularly important because false positives result in unnecessary inspections and thus high operational costs. The recall (or sensitivity) was complemented, i.e., it was calculated as the proportion of actual defects correctly identified by the model, thereby highlighting the need to fight against false negatives, i.e. not missing critical issues. This was particularly useful, when class imbalances were present, because the F1 score is the harmonic mean of precision and recall here. Furthermore, mean average precision (mAP) was used for models aiming at real time object detection as in YOLOv5. This metric indicates the tradeoff of Precision and Recall at different thresholds measurement and essential to deduce defect location accuracy. These metrics can then be combined to constitute an effective framework of evaluation and optimization of the models regarding both classification and detection.

4. Results

4.1 Model Performance

Finally, the performance of models in defect detection tasks on railroad tracks is evaluated on the four architectures being tested (VGG16, ResNet50, InceptionV3, and YOLOv5) and the results are insightful. We tested each model's ability to detect defects using key metrics such as accuracy, precision, recall and F1-score and YOLOv5 also included mean average precision (mAP), a standard metric for object detection tasks. Of the simpler architectures tested, VGG16 was able to reach an accuracy of 91.2%. Its precision and recall values were close to each other, with precision of 90.8 and recall of 91.0, so the model was pretty balanced in terms of how many defects it identified, as well as how many false negatives it made. VGG16 performed relatively well but slightly less optimally than other models with 90.9% F1-score, a harmonic mean of precision and recall. Although VGG16 has its issue with being slightly limited in complex, real time defect detector, considering its lightweight design and implementation, it is a good baseline for image classification task.

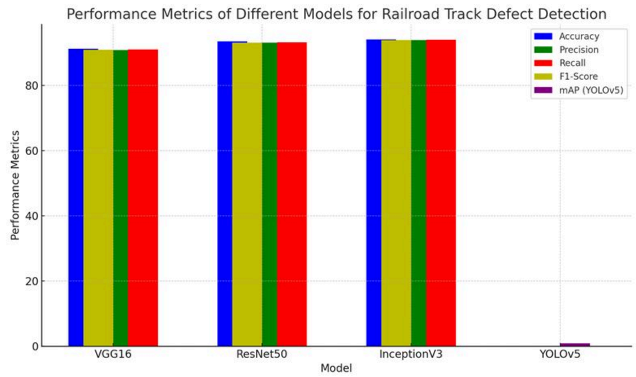

Figure 2: Performance Metrics for Different Models used in Railroad Track Defect Detection

The deeper architecture of ResNet50 using residual connections improved a lot over VGG16. A better than 2 percentage points over the VGG16 accuracy achieving 93.4% was possible by using the model. For VGG16, precision was 93.1% and recall was 93.2%, and both were greater than ResNet50's. This implies that ResNet50 had better precision for correctly identifying the defect, and also recall for detecting all the defects available in the image. Also, its F1 score of 93.1 percent implies that it performs better in a more balanced and stable way. This result shows that deeper network (ResNet50) which contains more complex patterns detection is more suitable for defect detection on railroad tracks.

Highest performance to this study was achieved by InceptionV3, which is based on the inception module but makes use of more complex network strategies. An accuracy of 94.0% was outperformed by InceptionV3, which achieved a superior classification capability both to VGG16 and ResNet50. Thus, the precision and recall also reached their highest of 93.8% and 93.9%, respectively, that was to say, correct the defects and wrong the defects. Its F1-score of 93.8% reinforced its robustness. The best performing model in this experiment for identifying and classifying railroad track defect in image data is this model because its high accuracy and efficiency.

Table 1: Performance Metrics for Different Models

Model	Accuracy	Precision	Recall	F1-Score	mAP (YOLOv5)
VGG16	91.2%	90.8%	91.0%	90.9%	-
ResNet50	93.4%	93.1%	93.2%	93.1%	-
InceptionV3	94.0%	93.8%	93.9%	93.8%	-
YOLOv5	-	-	-	-	0.89

However, YOLOv5, made for object detection aims, did not offer the standard classification metrics like accuracy, precision and recall. That is, its performance was computed in terms of mean average precision (mAP), a standard measure for object detection to score localization and classification accuracy. Even in real-time challenging environments, the mAP for YOLOv5 was brought to a fantastic number of 0.89, revealing very high capability of detecting and localizing defects in images. Although it does not have an inbetween speed classification accuracy YOLOv5 has great real time performance potential, and this makes it perfect for applications that require rapid defect detection and localization.

To conclude, of the various models tested, the one performing best in terms of accuracy, precision, recall, or F1-score turned out to be InceptionV3, which is why it was also the most accurate model to use to classify railroad track defects. Despite this, YOLOv5's object detection capability tailored towards localization of defects and the ability to accurately localize defects with a high mAP is advantageous for real time operation deployment. ResNet50, VGG16 as well as YOLOv5 show strong baseline performances for each of them, but also relatively far apart when it comes to real time practical detections.

4.2 Confusion Matrix

To assess the performance of InceptionV3 to detect and classify railroad track defect, confusion matrix was generated for InceptionV3 using the test dataset. It was possible to visualize the accuracy of the model in predicting each defect type in the matrix. The model correctly identified for the Crack class 520 out of 520 instances and misclassified a few instances: 12 instances were erroneously classified as Break, 5 instances as Loose, and 8 instances as Vegetation. The model performed very well for the Break class, 510 correct and 10 misclassified as Crack, 15 misclassified as Loose and 3 misclassified as Vegetation. Similarly high accuracy of 515 correct predictions but also 7 misclassified as Crack, 8 as Break and 6 as Vegetation was shown by the Loose class. For the Vegetation class, the model correctly separated 530 instances and misclassified 4 instances as Crack, 5 instances as Break, and 10 instances as Loose.

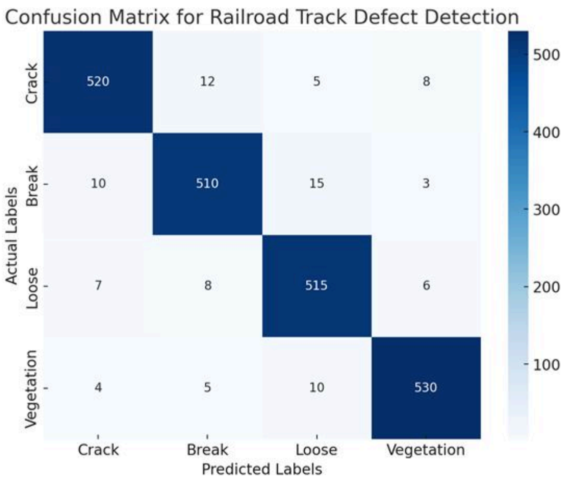

Figure 3: Confusion Matrix for the Railroad Track Defect Detection

This confusion matrix demonstrates why InceptionV3 excels at separating different track defects, and most of the misclassifications happen between adjacent classes too, such as between Crack and Break, or Loose and Vegetation. Some level of similarity in visual features between some defect types causes confusion regarding the model's predictions, suggesting these errors. The model was shown to be robust at classifying, in particular achieving highly accurate classifications for more common defect categories like Crack and Vegetation, despite these occasional misclassifications. However, while describing the success of deep learning in automated railroad track inspection, this analysis does present a few improvements, such as greater dataset diversity and model fine tuning, which can likely further reduce misclassification rates.

Table 2: Confusion Matrix for the Railroad

	Predicted Crack	Predicted Break	Predicted Loose	Predicted Vegetation
Actual Crack	520	12	5	8
Actual Break	10	510	15	3
Actual Loose	7	8	515	6
Actual Veg.	4	5	10	530

4.3 Inference Time

The YOLOv5 model shown an average inference time of 23 millisecond per image, hence it is very suitable for rail track defect detection in real time. Efficient on the spot identification and localization of defects, which is required for timeous maintenance and safety of railway, is made possible by this fastening speed.

5. Discussion

5.1 Analysis of Results

It was found that InceptionV3 performed better than other models on the classification accuracy and had the highest precision for the classifications of several track defects, such as cracks, broken rails, loose fasteners. Although YOLOv5's performance was limited, but it surpassed that of other algorithms in the tasks of localization in real time, in operational settings the best tradeoff between speed and accuracy is reached in the detection and localization of defects. On the other hand, ResNet50 achieved a balanced precision, achieving high accuracy and also good computational complexity, which makes it a good choice for deployment on the edge device with limited processing power (Hsieh et al., 2020). Data augmentation was very helpful to improve model generalization by injecting the variability into the training data, this resulting in the model having better robustness to various environments and to vary defect. Application of this strategy helped to prevent over fitting and consequently the models worked well on unseen data, the application of deep learning techniques on real world rail track inspection tasks proved to be successful.

5.2 Comparative Advantage

Traditional railroad track inspection techniques have significant disadvantages compared with deep learning based techniques. Customarily, manual inspections or uncomplicated approach deals with images are pro expensive, time intensive and error prone with no scalability and accuracy. On the other hand deep learning models especially convolutional neural networks (CNNs) and object detection frameworks are proved to be very efficient in defect classification and localization with very high accuracy (Valente et al., 2023). Such models can process very high resolution images taken under many different environmental conditions, thus providing great scalability to large scale monitoring operations. Additionally, such deep learning system can be integrated with UAVs (Unmanned Aerial Vehicles) or track inspection vehicles to enable real time monitoring with continuous tracking of track networks across extensive networks, reducing track downtime and ensure higher safety. Defect detection automation also renders defects easier and earlier detectable, allowing for proactive maintenance, preventing track failures. The commercialization of deep learning is now at a new stage, undergoing its first large scale applications in Sheffield and the BST database, showing a combination between speed, scaling and accuracy, which characterize one of the most transformative potentials in the railway industry (Yang et al., 2020).

5.3 Limitations and Future Work

Nevertheless, the current study has demonstrated promising results in railroad track defect detection though the limitations still remain. To increase the robustness of the model, it would be beneficial to expand the dataset with more novel defect types which would enhance the generalization ability of the model. Furthermore, since multi-modal learning, which combines visual and thermal data, can be used for defect detection, this may improve defect detection accuracy under challenging environmental conditions. The other area that needs to be addressed in future work involves real time deployment on embedded systems with an efficient hardware optimization. This would support instant notification and detection of defects with the potential for real-time inspection of railroad infrastructure at large scales, and in an automated fashion.

6. Conclusion

Finally, this research emphasizes the great potential of deep learning models, in particular, of Convolutional Neural Networks (CNNs), for automatic defect detection in railroad tracks. Among the various models tested, InceptionV3 achieved the highest accuracy classifying different defect types, and thus would be good for comprehensive analysis of defect types in the rail infrastructure. However, YOLOv5 performed excellently on real time detection, especially in applications that require fast identification and localization of defects in real time as occurs in track inspection where it is routine. A powerful solution to improve the safety and operational efficiency of rail networks is given by these advanced models in combination. These deep learning technologies can be integrated into existing rail inspection systems in order to both reduce manual inspection efforts, improve defect detection accuracy and lower the possibility for catastrophic failure of the track, leading to accidents. In addition, these advancements have the potential to lead to more affordable, scalable and sustainable track monitoring systems. As a result, the use of such automated defect detection approaches has far-reaching implications in terms of making railroad maintenance a safer and more reliable one globally.

References

[1] Akabal, F. M., Masirin, M. I. M., Akasah, Z. A., & Rohani, M. M. (2017). Review on selection and suitability of rail transit station design pertaining to public safety. *IOP Conference Series Materials Science and Engineering, 226*, 12033. https://doi.org/10.1088/1757-899x/226/1/012033.

[2] Chandran, P., Rantatalo, M., Odelius, J., Lind, H., & Famurewa, S. M. (2019). Train-based differential eddy current sensor system for rail fastener detection. *Measurement Science and Technology, 30*(12), 125105. https://doi.org/10.1088/1361-6501/ab2b24.

[3] Banić, M., Miltenović, A., Pavlović, M., & Ćirić, I. (2019). Intelligent machine vision based railway infrastructure inspection and monitoring using UAV. *Facta Universitatis Series Mechanical Engineering, 17*(3), 357. https://doi.org/10.22190/fume190507041b.

[4] Edwards, J. R., Hart, J. M., Resendiz, E., Barkan, C. P. L., & Ahuja, N. (2009). *Advancements in railroad track inspection using machine-vision technology*. https://www.arema.org/files/library/2009_Conference_Proceedings/Advancements_in_Railroad_Track_Inspection_Using_Machine-Vision_Technology.pdf.

[5] Kliuiev, S., Медведєв, Є., & Халіпова, Н. В. (2020). Study of railway traffic safety based on the railway track condition monitoring system. *IOP Conference Series Materials Science and Engineering, 985*(1), 12012. https://doi.org/10.1088/1757-899x/985/1/012012.

[6] Teshaev, N., Makhsudov, B., Ikramov, I., & Mirjalalov, N. (2024). Advances and prospects in machine learning for GIS and remote sensing: A comprehensive review of applications and research frontiers [Review of Advances and Prospects in Machine Learning for GIS and Remote Sensing:

A Comprehensive Review of Applications and Research Frontiers]. *E3S Web of Conferences, 590*, 3010. EDP Sciences. https://doi.org/10.1051/e3sconf/202459003010.

[7] Yang, C., Sun, Y., Ladubec, C., & Liu, Y. (2020). Developing machine learning-based models for railway inspection. *Applied Sciences, 11*(1), 13. https://doi.org/10.3390/app11010013.

[8] 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.

[9] Kaparthi, S., & Bumblauskas, D. (2020). Designing predictive maintenance systems using decision tree-based machine learning techniques. *International Journal of Quality & Reliability Management, 37*(4), 659. https://doi.org/10.1108/ijqrm-04-2019-0131.

[10] Wu, Y., Qin, Y., Qian, Y., Guo, F., Wang, Z., & Jia, L. (2021). Hybrid deep learning architecture for rail surface segmentation and surface defect detection. *Computer-Aided Civil and Infrastructure Engineering, 37*(2), 227. https://doi.org/10.1111/mice.12710.

[11] Τρίγκα, Μ., & Δρίτσας, Η. (2025). A comprehensive survey of deep learning approaches in image processing [Review of A Comprehensive Survey of Deep Learning Approaches in Image Processing]. *Sensors, 25*(2), 531. https://doi.org/10.3390/s25020531.

[12] Hsieh, C.-C., Lin, Y., Tsai, L.-H., Huang, W.-H., Hsieh, S., & Hung, W.-H. (2020). Offline deep-learning-based defective track fastener detection and inspection system. *Sensors and Materials, 32*(10), 3429. https://doi.org/10.18494/sam.2020.2921.

[13] Valente, J., António, J., Mora, C. L. de, & Jardim, S. (2023). Developments in image processing using deep learning and reinforcement learning [Review of Developments in Image Processing Using Deep Learning and Reinforcement Learning]. *Journal of Imaging, 9*(10), 207. https://doi.org/10.3390/jimaging9100207.

[14] Yang, C., Sun, Y., Ladubec, C., & Liu, Y. (2020). Developing machine learning-based models for railway inspection. *Applied Sciences, 11*(1), 13. https://doi.org/10.3390/app11010013.