Railway Track Defect Detection

Augustin Crespin[0009−0007−9238−8063], Ioannis Kostis[0000−0002−1875−8587], and Pierre Schaus[0000−0002−3153−8941]

ICTEAM/INGI, UCLouvain, Belgium
{firstname.lastname}@uclouvain.be

The Track Sentry Vision (TSV) project, funded by the Walloon Region, aims to develop a low-cost, autonomous preventative maintenance system tailored for railway sectors in developing countries, with a focus on Africa. Current maintenance practices in these regions are often reactive and costly, relying heavily on corrective interventions. TSV seeks to transform this approach by introducing an autonomous detection system for track defects, designed to integrate seamlessly with commercial trains and operate within the constraints of limited resources. The system's core objective is to predict defects through continuous track condition analysis, thereby reducing unplanned downtime and prioritizing maintenance interventions based on real-time data.

The project addresses two critical data-science challenges. First, *data signal alignment* is required to compare multimodal sensor and video streams captured during different runs, where variations in train speed and environmental conditions introduce temporal misalignments. Second, *self-supervised anomaly detection* must identify defects and obstacles without relying on pre-labeled datasets, which are typically unavailable or impractical to collect in real-world railway environments. To tackle these challenges, we designed a proof-of-concept laboratory setup using a remote-controlled LEGO train equipped with a Raspberry Pi 5, an accelerometer, and a camera. This scalable platform allowed us to test and refine our algorithms in a controlled yet representative setting.

For signal alignment, we applied Dynamic Time Warping (DTW) [1] to accelerometer data collected during runs at varying speeds. DTW successfully aligned the time-series signals by stretching and compressing segments, enabling accurate comparison of track conditions across runs despite speed differences.

For anomaly detection, we explored two complementary approaches: self-supervised feature extraction using DINOv3 [2], and fine-tuning a lightweight YOLOv11n model. The YOLO [3] model, trained on manually annotated video frames, achieved an accuracy of 0.93, precision of 0.95, recall of 0.98, and an F1-score of 0.96, demonstrating robust performance in identifying artificial obstacles with minimal false positives. Meanwhile, DINOv3's self-supervised features enabled effective clustering of anomalous frames, such as those containing unexpected objects or track irregularities.

The detection pipeline is designed to integrate with a broader maintenance management framework using a *Model Context Protocol (MCP)* [5] architecture. An *Agentic RAG* [4] backend will handle the retrieval of technical documentation and maintenance records and allow interaction between these and the end-user, while detection outputs will be directed to a dedicated analytics module for

visualization and explainable reporting. This modular design ensures scalability and decouples the detection system from documentation management, allowing for flexible deployment.

Preliminary results from our lab experiments confirm the feasibility of DTW for time-series alignment and the effectiveness of self-supervised and vision-based methods for anomaly detection. However, further validation is required to assess the robustness and scalability of these approaches in real-world operational contexts. The use of a LEGO train as a proof of concept underscores the adaptability of our solution to low-resource settings, where full-scale testing may initially be challenging. Future work will focus on field validation, algorithm optimization for edge deployment, and expanding the Agentic RAG subsystem to support data-driven decision-making. Ultimately, TSV aims to provide a scalable, cost-effective solution that enhances railway safety and operational efficiency in developing countries.

Acknowledgments. This project is funded by Logistics in Wallonia.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. P. Xu, R. Liu, Q. Sun and L. Jiang: Dynamic-Time-Warping-Based Measurement Data Alignment Model for Condition-Based Railroad Track Maintenance, in IEEE Transactions on Intelligent Transportation Systems, vol. 16, no. 2, pp. 799-812, April 2015, https://doi.org/10.1109/TITS.2014.2342235
2. Oriane Siméoni and Huy V. Vo and Maximilian Seitzer and Federico Baldassarre and Maxime Oquab and Cijo Jose and Vasil Khalidov and Marc Szafraniec and Seungeun Yi and Michaël Ramamonjisoa and Francisco Massa and Daniel Haziza and Luca Wehrstedt and Jianyuan Wang and Timothée Darcet and Théo Moutakanni and Leonel Sentana and Claire Roberts and Andrea Vedaldi and Jamie Tolan and John Brandt and Camille Couprie and Julien Mairal and Hervé Jégou and Patrick Labatut and Piotr Bojanowski: DINOv3, https://doi.org/10.48550/arXiv.2508.10104 https://arxiv.org/abs/2508.10104,
3. Joseph Redmon and Santosh Divvala and Ross Girshick and Ali Farhadi: You Only Look Once: Unified, Real-Time Object Detection, 2016, https://doi.org/10.48550/arXiv.1506.02640, https://arxiv.org/abs/1506.02640
4. Singh, Aditi and Ehtesham, Abul and Kumar, Saket and Talaei Khoei, Tala: Agentic Retrieval-Augmented Generation: A Survey on Agentic RAG, arXiv e-prints, arXiv–2501, 2025, https://doi.org/10.48550/arXiv.2501.09136
5. Anthropic, "Introducing the Model Context Protocol," Anthropic Blog, 2023. [Online]. Available: https://www.anthropic.com/news/model-context-protocol