CETRA2026 27-30 May 2026, Bol, Island Brač, Croatia
9th International Conference on Road and Rail Infrastructure

MACHINE LEARNING TECHNIQUES FOR LONG-TERM CONDITION MONITORING OF TRAMWAY TRACKS

Lana Miličević[1], Krešimir Burnać[1], Ivo Haladin[1], Nenad Trifunović[2], Katarina Vranešić[1], Franka Meštrović[1]
[1]*University of Zagreb, Faculty of Civil Engineering, Zagreb, Croatia*
[2]*University of Rijeka, RITEH, Rijeka, Croatia*

Abstract

Railway and tramway systems are subject to continuous mechanical stress due to high operational frequency and load conditions, with switches and crossings (S&C) being the most vulnerable parts of rail infrastructure. In urban tramway networks, where S&C density is particularly high, their maintenance alone accounts for approximately 30–50% of total tramway maintenance resources. Traditional maintenance strategies, such as corrective and time-based preventive maintenance, often lead to inefficient resource utilization, unexpected failures, and unnecessary service interruptions. In tramway systems, where safety and continuous operation are critical, adopting condition-based and predictive maintenance strategies is being increasingly recognized as essential for ensuring functionality and reducing maintenance costs. This paper provides an overview of state-of-the-art machine learning methods for track condition monitoring and predictive maintenance, with a focus on tramway systems. Furthermore, it presents a self-supervised embedding-based pipeline for detecting infrastructure-induced vibration patterns tailored to the unique operational characteristics of the Zagreb tram network designed as part of the project URITMIS – Urban railway infrastructure predictive maintenance system based on monitoring of vibro-acoustic track properties. Using monitoring data acquired from an in-service tram vehicle, the suggested approach uses machine learning algorithms to detect track irregularities associated with S&C degradation, with the overall objective of predicting maintenance cycles for efficient resource management.

Keywords: predictive maintenance, condition-based monitoring, signal processing, urban railway systems, IoT

1 Introduction

Traditionally, track maintenance operations have relied on scheduled interventions or addressed failures post-occurrence, leading to inefficiencies, increased costs, and potential service disruptions [1]. These shortcomings can significantly impact service reliability and maintenance costs, especially in tramway systems, where operational continuity and passenger safety have shown to be critical. Condition-based monitoring and predictive maintenance are becoming increasingly acknowledged as the most effective maintenance strategy in various industries, including rail transport. Through optimized response time and resource utilization, predictive maintenance can greatly reduce the effort intensity of conducting rail diagnostics, minimize the disruption of the regular operation schedule, and ultimately minimize the maintenance costs [2]. The railway sector necessitates constant improvements and novelties in track maintenance techniques to provide a contemporary service.

DOI: https://doi.org/10.5592/CO/CETRA.2026.2104

As a part of their research, authors in [3] have tried to characterize the motivation behind the need for such improvements into three categories of concerns: economical, safety, and technical. Aiming to balance technical quality and safety with economic costs, predictive techniques have proven to be very important process that allows railway operators to assess track conditions through different degradation components, such as track load, duration of operation, vehicle speed, and other parameters [4, 5]. Due to the complexity of the interaction of these parameters, machine learning (ML) techniques have shown to be particularly well-suited to mathematically model track wear, predict required maintenance operations, and adapt to various areas of focus [6, 7].

In this paper, an overview of current predictive maintenance techniques based on machine learning approaches is presented. Similar ML approaches are being used as a basis for the development of a novel, improved machine learning model as a part of project URITMIS conducted in Zagreb, Croatia. The model is being developed based on the acquired in-service vehicle vibration data on the Zagreb tram network to detect track irregularities and defects (with focus S&C locations), making the foundation for predictive maintenance track condition monitoring approach as the long-term goal for its utilization.

2 Machine learning approaches

The rapid increase in data acquired through sensors and onboard systems has enabled machine learning models that are able to characterize degradation beyond traditional threshold-based systems [7-9]. Rather than relying on predefined physical and statistical assumptions, ML models can learn directly from data. This allows them to capture nonlinear relationships, use large data volumes, and handle heterogeneous inputs [10, 11]. The challenges this approach tackles correspond perfectly to the specific type of data in railway systems - complex interactions of track geometry, vehicle dynamics and operational conditions.

Vibration-based input dominates in the current ML research in railway monitoring. Axle-box acceleration (ABA) measured from in-service vehicles has been commonly used form of input data due to its interpretability, ease of acquisition, and low volume in comparison to its common counterpart, image data [6, 10]. ABA data is typically collected via accelerometers mounted on the axle box or bogie frame of in-service vehicles, with sampling rates depending on the defect types of interest [10].

Early machine learning implementations demonstrated feasibility for engineered features [7, 12], while more recent studies apply deep learning directly to raw signals or time–frequency representations [9, 13]. Furthermore, research is moving toward hybrid knowledge-fusion models and digital-twin-integrated reinforcement learning frameworks. The most advanced systems do not stop at defect detection, but move towards closed-loop maintenance, creating fully integrated predictive maintenance ecosystems. However, despite the breadth of approaches, urban tramway networks remain largely unaddressed in literature. Current research employs both supervised and unsupervised learning paradigms, with the choice between the two being largely dictated by data availability. Supervised approaches achieve strong performance when labeled defect datasets are available, but their applicability is limited by the cost and scarcity of annotations in operational settings. Unsupervised and self-supervised methods on the other hand have gained advantage as practical alternatives, particularly for anomaly detection in scenarios where labeled data cannot be assumed [11].

The approaches within the realm of machine learning can be generalized into the subcategories of: (1) classical machine learning, (2) deep learning, and (3) online and reinforcement machine learning. Each represents a distinct level of model complexity, data requirement, and adaptability to the dynamic conditions of railway operations (table 1).

Table 1 Overview of machine learning approaches used in railway engineering

Approach	Core idea and typical methods	Railway applications	Reference
Classical ML	Standard algorithms (SVM, k-NN, decision trees)	Baseline fault detection, early predictive models	[7-9]
Deep learning	Multi-layer neural networks (CNN, RNN, LSTM) - automatic feature extraction	Defect detection, forecasting, anomaly detection	[8, 9]
Online and reinforcement learning	Learning from data streams or digital twin feedback	Adaptive condition monitoring, digital twin integration	[9, 11]

2.1 Classical machine learning

The earliest efforts in data-driven railway condition monitoring and predictive maintenance were conducted using classical machine learning methods. These methods typically operate on manually engineered features derived directly from the data. Commonly used techniques include support vector machines (SVM), k-nearest neighbors (k-NN), decision trees, random forests, and gradient boosting models [7, 8].

A key advantage of classical ML methods lies in their relatively low computational complexity and moderate data requirements, making them suitable for scenarios where labelled datasets are limited or where there are real-time processing constraints. However, the effectiveness of classical ML approaches is strongly dependent on the quality of feature engineering. The manual design of features requires substantial domain expertise and yet may still fail to fully capture the complex, nonlinear interactions governing track degradation processes [9].

2.2 Deep learning

Deep learning continues the evolution of machine-based modelling and is characterized by the use of artificial neural networks with multiple hidden layers. It does not rely on manually engineered features, but rather performs automatic feature extraction [8, 9]. Common deep learning architectures include convolutional neural networks (CNNs), recurrent neural networks (RNNs), long short-term memory (LSTM) networks, transformers, and autoencoders [8, 9]. CNNs are frequently applied to vibration signals and image-based inspection data for defect detection and localization. Transformers, RNN- and LSTM-based models are well suited for time-series analysis and degradation forecasting. Autoencoders, particularly in unsupervised settings, have proven effective for anomaly detection in scenarios where labeled data are scarce [9]. Deep learning is increasingly applied in railway predictive maintenance for its ability to model complex nonlinear and spatiotemporal patterns from large datasets, though it requires substantial computational resources and extensive training data. Deep learning models typically require larger training datasets and higher processing power, both during training and deployment. Furthermore, their black-box nature raises concerns regarding interpretability and trustworthiness [8]. However, the advantages of deep learning for urban tramway condition-based monitoring and predictive maintenance heavily outweigh the challenges. Frequent switches and crossings, heterogeneous track and load structure, and high volume of monitoring data generate complex data patterns that exceed the modelling capacity of more traditional approaches. Consequently, deep learning forms the methodological foundation of the proposed approach developed within the URITMIS project.

2.3 Online and reinforcement learning

Online learning extends ML approaches by enabling models to be updated continuously as new data becomes available, rather than being trained on a fixed dataset. For railway applications, this is particularly relevant in monitoring how track conditions evolve over time. Machine learning results can be further improved by integrating ML techniques with digital twins of the systems they model, using reinforcement learning techniques. The digital twin's outputs are used as feedback to the artificial neural network to learn more effectively [9]. Sresakoolchai and Kaewunruen [11] argue that future maintenance strategies should therefore incorporate digital twin frameworks and reinforcement learning paradigms to achieve optimal predictive maintenance performance.

3 Challenges in railway applications

Challenges may arise when utilizing the model outputs by human operators due to the black-box nature of machine learning methods, especially when they are based on artificial neural networks. Without the guarantee for the model output semantics, its interoperability and reusability are put to question. This is especially important in scenarios when these are to be used for operations with low or no error tolerance, such as risk management [8]. Enhancing machine learning pipelines with explainability techniques allows railway experts to more thoroughly understand the implications of model outputs for real-life applications, and to mitigate any subjectivity the model might be prone to is highlighted in [8]. This subjectivity is in most cases a result of the nature of the data, such as discrepancies in human and computer semantics, biases in data acquisition, data preprocessing oversights, and other. Other challenges arise from the shortcomings of publicly available knowledge and data. Nakhaee et al. [12] argue that it is the responsibility of the scientific and industrial community to address these shortcomings in terms of data quality and availability. Common points of failure include unbalanced classes in ML pipelines, the lack of labeled datasets, and the lack of public benchmark datasets.

3.1 Urban tramway networks

As previously mentioned, there are particular challenges surrounding the implementation of condition-based monitoring used in predictive maintenance techniques with its application to tramway networks. These challenges arise from both the physical characteristics of urban tramway infrastructure, the constraints of available data, and the lack of research in comparison to mainline rail applications. Tramway S&C employ grooved rail designs, in particular shallow groove crossings, which fundamentally differ from conventional, ballasted tracks. In shallow groove S&C, wheel guidance depends on tight groove tolerances, and progressive wear causes a qualitative change in wheel-rail contact mode. As wear progresses, contact becomes improper, transitioning from rim-groove contact to flange-running surface contact. This introduces a corresponding increase in dynamic forces and vibration levels [14]. On a vibration level, this introduces vibration response signatures specific to the system. Additional challenge is to characterize S&C from the vibroacoustic signal with the absence of a reliable single descriptor for their detection. Analysis of Zagreb tramway data carried out in an earlier phase of the URITMIS project suggests that the signal variability is governed by structural characteristics of track elements rather than operational parameters such as speed [2, 14]. These characteristics include element type, geometry, and wear state.
Consequently, the signal carries latent structural information about rail elements that cannot be decoded by simple threshold-based approaches, naturally imposing the employment of more refined statistical and machine learning methods.

This is further evidenced by the behavior of time-domain features across wear categories. Rather than increasing monotonically with degradation, RMS acceleration values showed inconsistent trends between standard and worn S&C profiles [14]. Only severely worn S&C, where the contact mode has already transitioned, produced clearly elevated acceleration levels. Even though a feature-based approach in the previous research [14] yielded results hinting at the possibility of tramway infrastructure evaluation using ML techniques, early and intermediate wear stages cannot be detected without using more advanced modeling techniques.

4 The URITMIS system for predictive maintenance

The URITMIS project - Urban railway infrastructure predictive maintenance system based on vibroacoustic track characteristics - is developing an integrated system for condition-based monitoring and predictive maintenance specifically for the Zagreb tram network. The system developed in this project integrates onboard sensing of an operating tram with a self-supervised deep learning pipeline for event detection. The Zagreb network is a case representative of the challenges described in the section 3.1. It is a dense urban tramway network with a high concentration of rail elements and tight curves, and heterogeneous traffic conditions in terms of operating vehicles, types of ballastless track superstructure, and passenger load. The URITMIS system directly addresses the limitations of manual track inspection and the challenges of employing condition-based monitoring in an urban tramway environment.

4.1 Data acquisition

Data is being acquired using an inertial measurement unit (IMU) mounted on the central bogie of an in-service tram vehicle, with a sampling frequency of 25 kHz. The signal has been downsampled to 5 kHz since the empirically discovered range of important frequencies for this use-case spans from 0 to 2.5 kHz. The points of interest for recording FLAC files were determined based on the condition of the tram network and areas prone to wear. These points of interest include 38 geofence zones, 12 of which are for monitoring the corrugated track sections, and the remaining 26 are for the locations with switches and crossings, covering a total of 215 S&C elements. The instrumented vehicle operates on a rotating schedule across multiple lines of the Zagreb tram network, covering all designated points of interest. When the tram enters a geofence zone, the sensor captures raw vibroacoustic signals across three modalities: a broadband signal – noise levels (geo), and the vibration response – acceleration of the left and right rail separately (geol, geor). Aside from the raw signals, positional and operational data including GPS coordinates, speed, bearing, equivalent sound level, maximum sound level, and equivalent peak vibroacoustic velocity and acceleration are recorded continuously along the entire track with 1 Hz sampling rate.

4.2 Data processing and machine learning

The proposed approach adopts self-supervised representation learning on raw vibroacoustic windows of the left and right rail signals (geol, geor). The FLAC file undergoes transformation to a CWT input, split into 2-meter windows with 50% overlap; windows and overlaps being chosen empirically. For the training that yielded the results mentioned in this paper, we observed 75 ground truth S&C elements, in 583 unique passages through 20 different element series (i.e. ways to pass through a crossroad). The model that was used was an autoencoder trained under a multi-objective curriculum, integrating three losses: a reconstruction loss to ensure the encoder captures signal structure faithfully, a delta prediction loss to encourage sensitivity to change in track geometry along the route, and a temporal contrastive loss to enforce consistency between embeddings from similar track positions (figure 1).

Loss weights were ramped gradually across training epochs, allowing the model to first establish good signal reconstructions before being asked to impose geometric and temporal structure on the latent space.

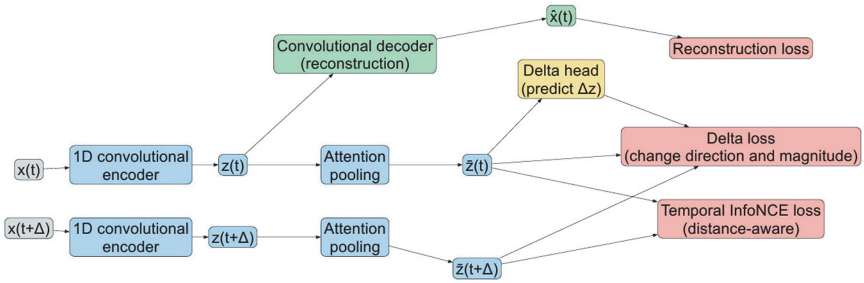

Figure 1 Architecture of the temporal 1D convolutional autoencoder

The resulting embeddings encode both local signal characteristics and positional context, forming the basis for downstream event detection and, in a later phase, condition and degradation monitoring. The downstream analysis pipeline identifies events by integrating two embedding interpretation approaches: (1) regime analysis and (2) PELT change-point segmentation (figure 2). Regime analysis applies KMeans clustering to the standardized embedding space, partitioning the track into three distinct regimes that reflect structural conditions in terms of track events (figure 2).

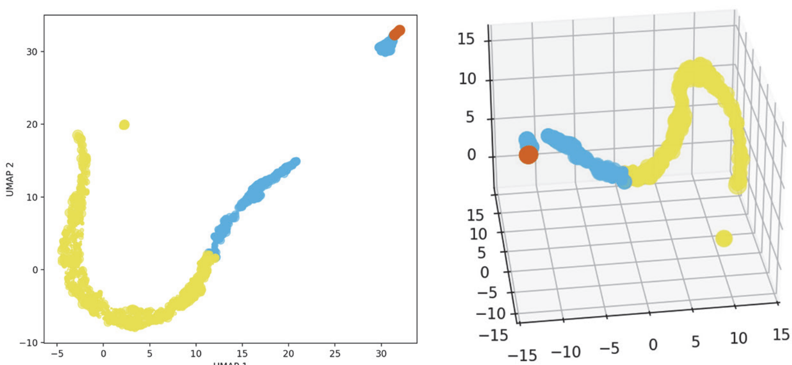

Figure 2 UMAP projection of latent embeddings z' (t), colored by KMeans regime cluster

To characterize transitions, temporal delta was computed: the frame-to-frame Euclidean distance of consecutive embeddings, which captured the change of vibroacoustic response along the track, with high values indicating changes in structure. Event detection turned the temporal delta into a signal using PELT change-point segmentation, a method to identify segments where the embedding deviates from the local baseline. The two approaches were combined such that regime labels from the first phase modulate PELT detection confidence scores – a regime change within a segment serve as a bonus, while a segment falling entirely into the dominant regime will be penalized. The output of the pipeline was a set of candidate event locations corresponding to infrastructure-induced transitions in the track (figure 3). A consensus of these event locations over multiple passages was calculated and shown on a map in order to compare it with real S&C locations (figure 4).

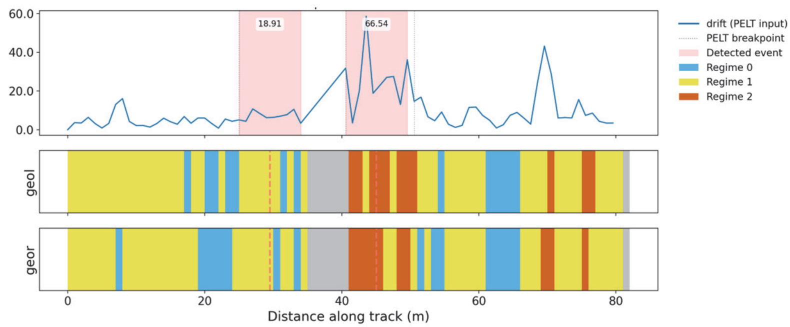

Figure 3 Example regime timeline for a track segment showing (top) PELT segmentation with confidence scores, and (bottom) KMeans regime assignment for signals

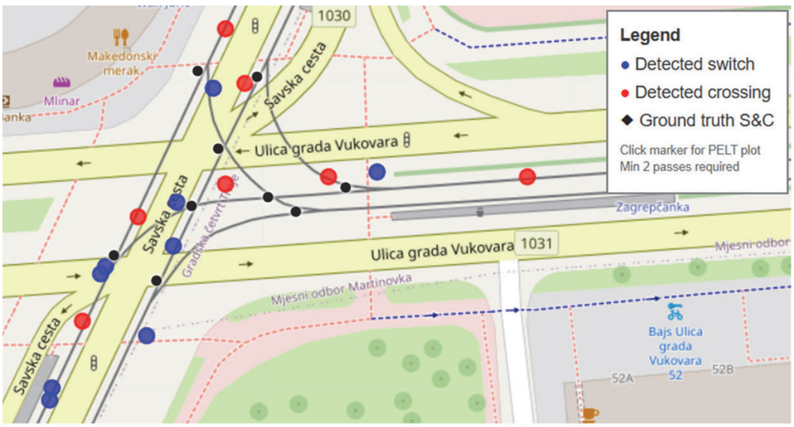

Figure 4 Detected events and ground truth locations on an example of one crossing, though most of the elements do get detected, they are often of the wrong type, the events detected on straight track likely symbolize defects

4.3 Challenges and limitations

The current implementation has several limitations that affect the direction of future work. While event detection proves effective in capturing structural transitions, type discrimination between switches and crossings remains unreliable due to the overlap in vibroacoustic signatures of those elements. That is further complicated by the track geometry differing at the selected sampling points. Additionally, false positive events are present in event detection, most likely due to corrugated track sections. Overall, quantitative performance of the pipeline remains limited, and the long-term objective of condition assessment will be addressed in the continuation of our work. Therefore, this stage of research served primarily as proof of concept that S&C elements can embody a meaningful representation in the embedding space created by an autoencoder. In the next stage of the research, focus will move from element identification towards wear characterization, aiming to qualify and quantify the characteristics of different wear stages and model its progression.

5 Conclusion

This paper presents an overview of machine learning approaches for railway predictive maintenance, with emphasis on the challenges of urban tramway networks. The URITMIS project is introduced as a case study addressing those challenges for the Zagreb tramway network, as well as the lack of tramway network representation in such research. Additionally, the suitability of deep learning and unsupervised methods to the tramway monitoring problem was addressed, as the signal signature is governed by structural factors of the network, and labeled data was not available. The URITMIS pipeline demonstrated that a multi-objective autoencoder trained on raw vibration responses can produce embeddings that encode structural information. The analysis of these embeddings allowed for the detection of S&C events along the track. Current limitations include false positive events and the inability of event semantic segmentation, which are going to be addressed in future research. The main direction for future work is obtaining historic data and used them as indicators predicting condition and degradation, fulfilling the objective of predictive maintenance.

Acknowledgments

This paper was written as a part of the "URITMIS – Urban railway infrastructure predictive maintenance system based on monitoring of vibroacoustic track properties" project NPOO. C3.2.R2-I1.06.0001 that is funded as a part of Development and research grants from National plan of recovery and resilience, financed by EU Commission.

References

[1] Oudshoorn, M., Koppenberg, T., Yorke-Smith, N.: Optimization of annual planned rail maintenance, Computer-Aided Civil and Infrastructure Engineering, 37 (2022) 6, pp. 669-687, DOI: 10.1111/mice.12764

[2] Haladin, I., Burnać, K., Baniček, M., Vranešić, K., Trifunović, N.: Machine learning for network-wide predictive maintenance on urban railway tracks: URITMIS project overview, 8th International Conference on Road and Rail Infrastructure - CETRA 2024, pp. 1063–1070, Cavtat, Croatia, 15-17 May 2024, DOI: 10.5592/CO/cetra.2024.1510

[3] Binder, M., Mezhuyev, V., Tschandl, M.: Predictive maintenance for railway domain: A systematic literature review, IEEE Engineering Management Review, 51 (203) 2, pp. 120-140, DOI: 10.1109/EMR.2023.3262282

[4] Elkhoury, N., Hitihamillage, L., Moridpour, S., Robert, D.: Degradation prediction of rail tracks: A review of the existing literature, The Open Transportation Journal, 12 (2018) 1, pp. 88-104, DOI: 10.2174/1874447801812010088

[5] Hoelzl, C., Arcieri, G., Ancu, L., Banaszak, S., Kollros, A., Dertimanis, V., Chatzi, E.: Fusing expert knowledge with monitoring data for condition assessment of railway welds, Sensors, 23 (2023) 5, DOI: 10.3390/s23052672

[6] Falamarzi, A., Moridpour, S., Nazem, M.: Development of a tram track degradation prediction model based on the acceleration data, Structure and Infrastructure Engineering, 15 (2019) 10, pp. 1308-1318, DOI: 10.1080/15732479.2019.1615963

[7] Liao, Y., Han, L., Wang, H., Zhang, H.: Prediction models for railway track geometry degradation using machine learning methods: A review, Sensors, 22, (2022) 19, 7275, DOI: 10.3390/s22197275

[8] Hadj-Mabrouk, H.: Contributions and limitations of AI and machine learning in railway operations, maintenance, and safety: A literature review, Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability, 240 (2025) 1, pp. 3-36, DOI: 10.1177/1748006X251364324

[9] Di Summa, M., Griseta, M.E., Mosca, N., Patruno, C., Nitti, M., Renò, V., Ettore, S.: A review on deep learning techniques for railway infrastructure monitoring, IEEE Access, 11 (2023), pp. 114638–114661, DOI: 10.1109/ACCESS.2023.3309814

[10] Tsunashima, H.: Condition monitoring of railway tracks from car-body vibration using a machine learning technique, Applied Sciences, 9 (2019) 13, 2734, DOI: 10.3390/app9132734

[11] Sresakoolchai, J. Kaewunruen, S.: Railway defect detection based on track geometry using supervised and unsupervised machine learning, Structural Health Monitoring, 21 (2022) 4, pp. 1757-1767, DOI: 10.1177/14759217211044492

[12] Nakhaee, M.C., Hiemstra, D., Stoelinga, M. van Noort, M.: The Recent Applications of Machine Learning in Rail Track Maintenance: A Survey, International Conference on Reliability, Safety, and Security of Railway Systems, pp. 91-105, Lille, France, 4-6 June 2019., DOI: 10.1007/978-3-030-18744-6_6

[13] Zuo, Y., Lundberg, J., Chandran, P., Rantatalo, M.: Squat detection and estimation for railway switches and crossings utilising unsupervised machine learning, Applied Sciences, 13 (2023) 9, 5376, DOI: 10.3390/app13095376

[14] Haladin, I. Burnać, K., Vranešić, K., Trifunović, N.: Detection and evaluation of switches & crossings on tramway infrastructure from vibration signals using machine learning techniques, 31st International Congress on Sound and Vibration, Incheon, South Korea, 6-11 July 2025.