



D29.2

Demonstration and Evaluation of standardised European Railway Checkpoints at borders or other operational stop points

Project acronym:	FP5-TRANS4M-R
Starting date:	2022-07-01
Duration (in months):	54
Call (part) identifier:	HORIZON-ER-JU-2022-01 (Topic: HORIZON-ER-JU-2022-FA5-01)
Grant agreement no:	101102009
Due date of deliverable:	Month 42
Actual submission date:	2026-03-30
Responsible/Author:	Behzad Kordnejad (KTH)
Dissemination level:	PU - Public
Deliverable Type:	DEM – Demonstrator, pilot, prototype
Doc Version & Status:	V2.0 Submitted

Reviewed: (Yes)



"Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the Europe's Rail Joint Undertaking. Neither the European Union nor the granting authority can be held responsible for them."

"The project FP5-TRANS4M-R is supported by the Europe's Rail Joint Undertaking and its members."

Document history		
<i>Revision</i>	<i>Date</i>	<i>Description</i>
V0.1	2025-09-09	First issue
V0.2	2026-02-24	Version sent for internal review
V0.3	2026-03-30	Version submitted to PC/PMO
V1.0	2026-03-30	Final quality check and submission
V2.0	2026-05-27	Contributions corrected

Report contributors		
<i>Name</i>	<i>Beneficiary Short Name</i>	<i>Details of contribution</i>
Behzad Kordnejad	KTH	Lead author, Editor, chapters 1, 5, 6, 9
Ingrid Nordmark	KTH	Editor, chapters 2, 3, 4, 5, 6, 9
Martin Kjellin	RISE	Chapters 7.2, 8.4, 9
Mats Åkerfeldt	RISE	Chapter 7.2
Florian Thiery	LTU	Chapter 7.2
Elisa Sengle	DBC	Chapter 7.1
Margarite Klass	DBC	Chapter 7.1, 9
Alexander Ungefucht	DBC	Chapter 7.1, 9
Vitali Rjasanzew-Frey	DBC	Chapter 8.1, 8.2, 8.3
Alexis Vidaurre Celis	ADIF	Chapters 7.3, 9
Markos Losada	CEIT	Chapters 7.3, 9
Iñigo Adin	CEIT	Chapters 7.3, 9
Teresa Hernando Bravo Manuel Blanco Castillo Francisco Parrilla Ayuso	INDRA	Chapters 7.3
Luuk Spreeuwiers	TWENTE/PRORAIL	Chapter 7.4

Reviewers		
<i>Name</i>	<i>Beneficiary Short Name</i>	<i>Details of contribution</i>
Arjen Zoeteman	ProRail	Official Review and Feedback
Francesco Fasano	Hitachi	Official Review and Feedback

Disclaimer

The information in this document is provided "as is", and no guarantee or warranty is given that the information is fit for any particular purpose. The content of this document reflects only the author's view – the Joint Undertaking is not responsible for any use that may be made of the information it contains. The users use the information at their sole risk and liability.

The content of this report does not reflect the official opinion of the Europe's Rail Joint Undertaking (EU-Rail JU). Responsibility for the information and views expressed in the report lies entirely with the author(s).

Table of Contents

1	Executive Summary	10
2	Abbreviations & Acronyms	11
3	Background	13
4	Objective/Aim.....	14
4.1	Task Description.....	14
5	European Railway Checkpoint concept	16
6	Demonstration webinar	18
7	Evaluation of ERC modules	19
7.1	Module 1: AI-powered condition monitoring (Germany)	19
7.1.1	Railway Checkpoints in Germany	19
7.1.2	Use case: Condition monitoring - Bogie Analyzer	20
7.1.3	Description of the recognized components	22
7.1.4	Reasons for selecting the components.....	24
7.1.5	Technical description of the process.....	26
7.1.6	Image Requirements for the Bogie Analyzer AI Model.....	33
7.1.7	Evaluation of the system - KPIs within the framework of ERC	33
7.1.8	Analysis of the usability of the AI model on data other checkpoints	47
7.1.9	Validation of usage on data from ProRail	49
7.1.10	Requirements for bogies.....	49
7.1.11	Further recommendations.....	49
7.2	Module 2: Checkpoints for improving sharing of logistics information (Sweden) 51	
7.2.1	Use Case: Rolling stock and ILU identification.....	51
7.2.2	Locations and roll-out plan	52
7.2.3	Evaluation	53
7.2.4	Checkpoints for improving monitoring of wheelsets via data quality checks	55
7.3	Module 3: Checkpoints deployment in Can Tunis for capturing yard movements (Spain)	60
7.3.1	Introduction to the Spanish demonstrator.....	60
7.3.2	Checkpoints deployment in Can Tunis.....	62
7.3.3	Use case: Integrating checkpoints with On-board systems	84

7.3.4	Evaluation and Preliminary Results.....	86
7.3.5	Issues and challenges identified.....	88
7.3.6	Conclusions and recommendations.....	90
7.4	Module 4: Recognition of Hazardous Materials Plates/Placards and UIC Code (the Netherlands)	91
7.4.1	Use case: Identification of dangerous goods and UIC Code	91
7.4.2	Main objectives Recognition of Hazardous Materials Plates/Placards.....	92
7.4.3	Main objectives Reading Wagon Codes	92
7.4.4	Results Recognition of Hazardous Materials Plates/Placards.....	93
7.4.5	Results Reading Wagon Codes	95
7.4.6	Issues and challenges identified.....	99
8	Data management.....	102
8.1	Data Sources	102
8.2	Conceptual Data Model (CDM)	103
8.3	Examples of Exchanged Data.....	105
8.3.1	AI-detected damage.....	105
8.3.2	Human annotation.....	106
8.3.3	Logistics-related information	106
8.4	Data Usage	107
9	Conclusions.....	109
	References.....	112

Table of Figures

Figure 7.1 Locations of DB Cargo's IVGs installed in Germany	19
Figure 7.2 Standardized IVG of DB Cargo installed at 13 locations in Germany	20
Figure 7.3 Analysed components (wheel set with tare springs and 4 cover screws)	21
Figure 7.4 Analysed components (wheel set with parabolic springs and 3 cover screws)	22
Figure 7.5 Object detection of the wheelset areas	26
Figure 7.6 Left: Segmentation (AI). Right: Smoothing of the segment edges	27
Figure 7.7 Reference geometries for determining the pixel to millimetre ratio	27
Figure 7.8 Angle of Left: Lifting T. Right: Damper ring	28
Figure 7.9 Wheelset with non-matching tare springs	30
Figure 7.10 Distance of Left: Centre point of cover screws to lower edge of lifting brackets	31
Figure 7.11 Weight distribution of an open flat wagon	32
Figure 7.12 Example of load distribution results generated by the Bogie Analyzer	32
Figure 7.13 Excerpt from the loading guidelines	33
Figure 7.14 Wagon number recognition rate 2025	35
Figure 7.15 Percentage of damaged vehicles captured by a IVG in preceding 30 days at least once	36
Figure 7.16 Latency from IVG to S3 bucket	37
Figure 7.17 Latency from S3 bucket to get the AI results	37
Figure 7.18 Distribution of touchpoints of the IVG	39
Figure 7.19 Bogie Analyzer coverage	39
Figure 7.20 Bogie Analyzer detection per country	40
Figure 7.21 Bogie detections	40
Figure 7.22 Valid wheelset detections	41
Figure 7.23 Analysis evaluation results (Lifting T)	42
Figure 7.24 Left: damping image quality. Middle: damping invalid. Right: damping valid	43
Figure 7.25 Vertical distance between the spring cap and the bogie frame (<2 mm)	43
Figure 7.26 Analysis evaluation results (damper ring)	44
Figure 7.27 Validation approach for brake pad thickness	45
Figure 7.28 A wagon image that has been produced by a checkpoint, along with a vehicle number (left) and two ILU codes (right) that can be extracted from it	51
Figure 7.29 The location of the checkpoint in Gothenburg (adapted map from OpenRailwayMap)	52
Figure 7.30 The location of the checkpoint in Malmö (adapted map from OpenRailwayMap)	53
Figure 7.31 A sketch of the deployment in Malmö and the actual installed checkpoint.	53
Figure 7.32 (a) Evolution of the median value of the detectors right side as function of time (b) Evolution of the degradation of the right wheels before rescaling/removal (c) Evolution of the degradation of the right wheels after scaling/removal	56

Figure 7.33 Temporal evolution of Mean Absolute Deviation of the reference detector by wagon type over the period 2022-2025	58
Figure 7.34 Geographical location.	60
Figure 7.35 Architecture of the Spanish Demonstrator.....	62
Figure 7.36 Location of the Llobregat and Morrot sites inside Can Tunis.	63
Figure 7.37 Detail: Llobregat site.	63
Figure 7.38 Detail: Morrot site.	64
Figure 7.39 Photograph of the topographic study.	66
Figure 7.40 Checkpoint for double-track railway installation at Morrot location.....	67
Figure 7.41 Checkpoint for single-track railway installation at Morrot location.....	67
Figure 7.42 Checkpoint for double track railway installation at Llobregat location.....	67
Figure 7.43 Checkpoint for single-track railway installation at Llobregat location.....	68
Figure 7.44 Checkpoint Llobregat section – RFIG.	68
Figure 7.45 Checkpoint Morrot section – RFIG.	69
Figure 7.46 Checkpoint Llobregat section – Port.....	69
Figure 7.47 Checkpoint Morrot section – Port.....	70
Figure 7.48 Detection area.	71
Figure 7.49 Acquisition area.....	72
Figure 7.50 Integration test Train Detection Module environment.	73
Figure 7.51 Integration test Train Detection Module detection camera.	73
Figure 7.52 Detection and Tracking example of a train in movement.	74
Figure 7.53 Homography transformation example.	75
Figure 7.54 Bounding box in transformed image for speed estimation.	75
Figure 7.55 Keypoint tracking and prediction.....	76
Figure 7.56 Train Detection Module Unitary Tests Sonar.....	76
Figure 7.57 Train Detection Module, video output result.....	77
Figure 7.58 Laboratory setup for illumination testing.	78
Figure 7.59 Captured images for different illumination conditions.	78
Figure 7.60 Acquisition camera calibration.	78
Figure 7.61 Short range camera without illuminator (distance = 300 cm).....	79
Figure 7.62 Long range camera without illuminator (distance = 750 cm).....	79
Figure 7.63 Image Acquisition Module - Unitary Tests Sonar report.....	80
Figure 7.64 Pre-classifier Module - Unitary Tests Sonar report.	81
Figure 7.65 Keypoint matching between consecutive frames of a wagon from stitching module.....	82
Figure 7.66 Stitching process: Panorama image of a complete wagon generated by merging multiple frames captured.....	82
Figure 7.67 Capture of the labelling platform.	83
Figure 7.68 Example of checkpoint dashboard.	83

Figure 7.69 a) Field layer hardware architecture, b) CEIT's modular device for WOBUM (Wireless On-Board Unit Module), WOBUC (Wireless On-Board Unit Central), or WSEN (Wireless Sensor Elements) operation, c) Installation of an on-board unit	85
Figure 7.70 Examples of Hazmat signs, data provided by DB Cargo.....	92
Figure 7.71 Examples of UIC codes. All images are provided by Garcia-Bordils et al. (2023).	93
Figure 8.1 Data transfer between German and Swedish Checkpoints	102
Figure 8.2 Dataflow from German IVG to Swedish API endpoint	102
Figure 8.3 JSON example- data exchange of a wagon.	103
Figure 8.4 Location of the damage and bounding box size in the images.	105
Figure 8.5 Damage in a metal floor with location in the image.	106
Figure 8.6 Orange-coloured plate with UN number on a wagon.....	107
Figure 8.7 Intermodal Loading Unit codes on two containers on a single wagon.	107
Figure 8.8 The user interface Tåginfo showing (using blue background colour) information about a wagon that has passed checkpoints in Seddin and Mannheim. ...	108

Table of Tables

Table 1 Task descriptions and related output in the report	15
Table 2 Overview of European Railway Checkpoints demonstrated and evaluated	16
Table 3 Topics presented at webinar and related deliverable chapter	18
Table 4 Excerpts from the GCU.....	25
Table 5 Summary of overall KPIs	38
Table 6 Summary of bogie detection and validation	46
Table 7 Validation criteria.....	87
Table 8 Performance metrics for YOLO11x (Jocher, 2025) and Faster RCNN (Ren et al., 2015) under different training strategies on the two test sets. The following training strategies were used: Early Stopping (E.S), Lower Learning rate (L.L), and Data Augmentation (D.A.).....	94
Table 9 Performance comparison of Tesseract [TA] (Hoffstaetter and Lee, 2025), EasyOCR [E-OCR] (Kittinaradorn, 2025), and Idefics2 [IDF2] (Laurençon et al., 2024) models on ProRail dataset, based on the Character Error Rate (CER) and Word Error Rate (WER). Lower CER and WER mean better performance.....	95
Table 10 Accuracy of our proposed pipeline on the ProRail dataset at different IoU thresholds. The pipeline uses the YOLO11x [YOLO] (Jocher, 2025) and Idefics2 [IDF2] (Laurençon et al., 2024) models.....	95
Table 11 Comparison of OCR model performance on the STEP Image Dataset [STEP] (Garcia-Bordils et al, 2023), demonstrating that our proposed method achieves substantially higher accuracy, recall, and precision than existing baselines.....	97
Table 12 Performance evaluation of our proposed model on video dataset subsets. Each row compares the model's performance on frames filtered by the condition in the leftmost column (subset) versus its complement (all other frames). Results show that occlusions and low-contrast conditions lead to notable drops in accuracy and recall. ...	97
Table 13 Comparison of the STEP [STEP] (Jocher, 2025) and our proposed model on line-scan data. Our model demonstrates substantially higher accuracy, recall, and precision across both individual camera views and combined camera views.....	98
Table 14 Performance Evaluation of our proposed model on top 5 most common challenges affecting line-scan data and their complement subsets. Each row compares the model's performance on frames filtered by the condition in the leftmost column (subset) versus its complement (all other frames). Results show that occlusions and low-contrast conditions lead to notable drops in accuracy and recall.....	99
Table 15 High-Level object description of data types.	104
Table 16 Description of DamageObject class.	104
Table 17 Description of BoundingBox object class.	105

1 Executive Summary

The main objective of work package (WP) 29 *Standardised European Railway Checkpoints at Borders and Other Operational Stop Points* has been to develop European Railway Checkpoints (ERC) that will partially automate Freight Train Transfer Inspections at borders or other operational stop points. Thus, this WP focuses on:

- Digitalisation and automation of manual processes through innovative sensors, video gates and handheld devices, based on a process analysis (TRL 7)
- Interoperable IT-systems for data management and processing (TRL 7)
- Harmonized procedures and regulations

This document addresses *Task 29.2: Demonstration and evaluation of the installed checkpoints* within WP29 and has been prepared to provide the output from the demonstration and evaluation. The results from the work within task 29.2 was demonstrated during a webinar, arranged on the 27th of January 2026, 11:30-13:00. In this webinar, developments within this working group were demonstrated by showcasing four modules from Germany, Sweden, Spain and the Netherlands, each focusing on a specific use case:

- Condition monitoring,
- Improving sharing of logistics information,
- Capturing yard movements and
- Recognition of Hazardous Materials Plates/Placards and UIC Codes

This report documents the demonstration within Task 29.2 and presents the evaluation of European Railway Checkpoint concept and the proposed four modules, as well as suggested methodologies for data exchange and possible paths for implementation and roll-out. In conclusion, the concept of European Railway Checkpoints is believed to represent a step forward in making European railways smarter, more efficient, and more customer focused. Further development is proposed to take place in the second call of FP5 where the aim is to reach Technology Readiness Level (TRL) 8 and have the actual system completed.

Keywords: European Railway Checkpoints (ERC), Intelligent Video Gate (IVG), Demonstrator, Image analysis, Data exchange, Rail Freight Transportation

2 Abbreviations & Acronyms

Abbreviation / Acronym	Description
ADIF	Administrador de Infraestructuras Ferroviarias
ADR	European Agreement Concerning the International Carriage of Dangerous Goods by Road
AI	Artificial Intelligence
API	Application Programming Interface
AWS	Amazon Web Services
BIC	Bureau International des Containers et du Transport Intermodal
CDF	Cumulative Distribution Function
CDM	Conceptual Data Model
CEIT	Centro de Estudios e Investigaciones Técnicas de Gipuzkoa
CER	Character Error Rate
CI/CD	Continuous Integration and Continuous Delivery (or Deployment)
CO ₂	Carbon dioxide
CVAT	Computer Vision Annotation Tool
DB	Deutsche Bahn
DBC	DB Cargo Aktiengesellschaft
ERC	European Railway Checkpoints
EU-Rail JU	Europe's Rail Joint Undertakings
EVN	European Vehicle Number
FAT	Factory Acceptance Tests
FDS	Federated Data Space
FP	Flagship Project (within Europe's Rail)
FP5 - TRANS4M-R	Europe's Rail Flagship Project 5 - Transforming Europe's Rail Freight
GA	Grant Agreement
GCU	General Contract of Use for Wagons
GNSS	Global Navigation Satellite Systems
GPS	Global Positioning System
GT	Ground-Truth
HERD	Harmonised European Railway Diagnostics (project within the System Pillar of the European Railway Joint Undertaking)
IDS	International Data Spaces
ILU	Intermodal Loading Unit; containers, swap-bodies and semi-trailers
INDRA	Indra Sistemas S.A.

IoU	Intersection over Union
IVG	Intelligent Video Gate
KPI	Key Performance Indicators
KTH	KTH Royal Institute of Technology
LoRA	Long Range
LSP	Lindholmen Science Park AB
LTE-M	Long-Term Evolution for Machines
MAD	Mean Absolute Deviation
mAP	mean Average Precision
MAWP	Europe's Rail Joint Undertaking Multi-Annual Work Programme
NB-IoT	Narrow Band Internet of Things
OCR	Optical Character Recognition
Renfe	Renfe Operadora
RFID	Radio Frequency Identification
RFIG	Red Ferroviaria de Interés General (Spanish national rail network)
RID	Regulation concerning the International Carriage of Dangerous Goods by Rail
RISE	RISE Research institutes of Sweden
S3	Amazon simple storage service
STEP	Structured Text sPotter
TAF/TAP TSI	Technical Specification for Interoperability relating to Telematics Applications for Freight/Passenger Services
TMS	Traffic Management System
TRL	Technology Readiness Level
UIC	International union of railways
UN number	United Nations number for hazardous goods
UWB	Ultra-Wideband
WBB	Wagenmeister Bildbasierte Befundung
WER	Word Error Rate
WILD	Wheel Impact Load Detector
WMS	Wayside Monitoring System
WOBU	Wireless On-board Unit
WOBUC	Wireless On-Board Unit Central
WOBUM	Wireless On-Board Unit Module
WP	Work Package
WSEN	Wireless Sensor Elements

3 Background

The present document constitutes the Deliverable D29.2 “Demonstration and Evaluation of standardised European Railway Checkpoints at borders or other operational stop points” which is part of Task 29.2: *Demonstration and evaluation of the installed checkpoints* in the framework of the Europe’s Rail Flagship Project 5 - Transforming Europe’s Rail Freight (FP5 - TRANS4M-R) as described in the Europe’s Rail Joint Undertaking Multi-Annual Work Programme (EU-RAIL MAWP) and contributes as well to the Flagship Projects FP1 – Motional and FP3 - IAM4RAIL.

European Railway Checkpoints rely on wayside sensing systems capable of automatically capturing information about trains and intermodal loading units during operation. These systems typically include imaging technologies, triggering sensors and data processing infrastructure enabling automated inspection and information extraction.

The work with European Railway Checkpoints had three phases, and took place in Work Package 29 (WP29) in FP5- TRANS4M-R. In the first phase, basic functions and technology were specified. This is described in Deliverable D25.1 *Report on the basic functional and technical specifications for the realisation of the technical enablers of Seamless Freight, also including the final specification input for FP1* (chapter 8, Hildebrandt et al., 2024).

In the second phase, the technical definition of the gates, also called Intelligent Video Gate, IVG, and related demonstrators was developed. Deliverable D29.1 *Technical definition of the standard IVG for checkpoints and related demonstrators* (Kordnejad et al., 2024) covers these topics.

The third phase covered testing and demonstration. Reports describing the installations and test plans as well as execution of tests have been produced within the project:

- *D29.3 Installation and test plan for Gates in Spain*
- *D29.4 Installation and test plan for Gates in the Netherlands*
- *D29.5 Execution of test for gates installed in Sweden*
- *D29.6 Execution of test for gates installed in Spain*
- *D29.7 Execution of test for gates installed in the Netherlands*

Two demonstrations have taken place within WP29, where this document constitutes of the final demonstration, in form of a webinar 2026-01-27. The other demonstration took place live in Nuremberg 2025-09-16 and are further described in the document *D29.8 Live-Demo of Video Gates showing process optimization in a German yard*.

4 Objective/Aim

The main objective of this work package (WP29) has been to develop European Railway Checkpoints (ERC) that will partially automate Freight Train Transfer Inspections at borders or other operational stop points. Thus, this WP has focus on:

- Digitalisation and partial automation of manual processes through innovative sensors, video gates and handheld devices, based on a process analysis (TRL 7)
- Interoperable IT-systems for data management and processing (TRL 7)
- Harmonized procedures and regulations

These objectives are addressed through the implementation and evaluation of several ERC demonstrators deployed across different European railway environments.

This document has been prepared to provide the output from the demonstration and evaluation of standardised European Railway Checkpoints at borders or other operational stop points. This document presents the demonstration carried out and covers the following topics: demonstration and evaluation of installed checkpoints, data management, harmonized procedures and implementation and roll-out plan for standardised European checkpoints.

4.1 Task Description

Task 29.2 Demonstration and evaluation of the installed checkpoints started in M31 (January 2025). Table 1 gives the match of the task described in the Grant Agreement (GA) with the output and a link to the section where more details can be found. Published deliverables can be found at FP5-TRANS4M-R's website: <https://rail-research.europa.eu/pages/fp5-trans4m-r/deliverables>.

Table 1 Task descriptions and related output in the report

Task definition from GA (Task 29.2)	Output of deliverable
Demonstrate and evaluate the development of IVG checkpoints based on the specifications in task 29.1 for cross border traffic, on main lines, at yards and terminals at suggested sites. Final selection and demonstration of use cases. Coordination of demonstration within the showcase Destination FA5 Seamless corridor together with other WPs within the Seamless cluster. Demonstrate and evaluate interoperable IT-systems for data management and processing using the Conceptual Data Model developed in task 29.1. Demonstrate and evaluate the mitigation strategies for the identified regulatory barriers for harmonized procedures and regulations and as specified in task 29.1.4.	Chapters 6, 7, and 8
This task will result in an installation report and test execution for each of the considered countries (D29.2-29.7). In addition, DBC will have a Live-Demo of Video Gates showing process optimization in a yard (D29.8). In total this will result in a common deliverable (D29.2).	D29.3-D29.4 Installation reports. D29.5-D29.7 Execution of test reports D29.8 Report of live-demo

5 European Railway Checkpoint concept

The concept of European Railway Checkpoints (ERC), also called Intelligent Video Gates (IVG), is a rather simple yet powerful way of intelligent monitoring of seamless rail freight transports. The ERC concept typically combines wayside sensing infrastructure, automated analysis modules and data exchange platforms enabling the sharing of inspection and logistics information between railway stakeholders. Based on new and emerging technologies, a wayside gate equipped with cameras as well as RFID (Radio Frequency Identification) readers and sensors. Images from the cameras undergo image analysis and OCR (Optical Character Recognition). The whole procedure enables capture of critical logistics and maintenance data in real time. These gates can favourably be positioned at strategic locations such as:

1. cross-border traffic,
2. on main lines, and
3. at yards and terminals

Thus enabling:

1. Digitalization and automation of operational processes
2. Stakeholders to improve operational efficiency and maintenance and enhance their offer to the market
3. Less administrative burdens and enhanced customer information.

Gate functionalities and technologies are described in more detail in D25.1 (Hildebrandt et al., 2024) and D29.1 (Kordnejad et al., 2024).

The ERC concept is defined a *Modular System with harmonised procedures* based on certain minimal functional requirements with possibilities for add-on functions depending on the gate owners' operational focus. The four countries where the concept has been implemented are presented in this deliverable: Germany, Sweden, Spain, and the Netherlands. An overview of the gates demonstrated and evaluated in the project and their functionalities are presented in Table 2.

Table 2 Overview of European Railway Checkpoints demonstrated and evaluated

Gate location	Country	Further reading
Marshalling yard (13)	Germany	D29.8
Marshalling yard (6)	The Netherlands	D29.7
Marshalling yard at border (2)	Spain	D29.6
Terminal (1)	Sweden	D29.5
Border (1)	Sweden	D29.5
Main line (6)	The Netherlands	D29.7

Moreover, as the concept constitutes a very good source for operational data, important alignments have been carried out regarding the aspect of data sharing, further elaborated in chapter 8. This alignment has been made with 1) FP1 Motional regarding alignment for the conceptual data model (see report D25.2 by Deiterding et al., (2023)) as well as 2) System Pillar Project HERD (Harmonised European Railway Diagnostics) regarding standardisation of diagnostics data.

Alignment has also been carried out with FP3 IAM4RAIL, as checkpoints are also being developed there although on main lines also focusing on passenger trains. This was covered by deliverable D25.3 (Kordnejad et al., 2023), dedicated for facilitating and reporting on this alignment.

6 Demonstration webinar

The results from the work with European Railway Checkpoints in WP29 was demonstrated during a webinar, arranged on the 27th of January 2026, 11:30-13:00. In this webinar, developments within this working group were demonstrated by showcasing four modules from Germany, Sweden, Spain and the Netherlands focusing on the following use cases:

- Condition monitoring: Bogie Analyzer
- Rolling stock and ILU identification
- Integrating checkpoints with On-board systems
- Identification of dangerous goods

Each module presented for 15-18 minutes and addressed a specific operational aspect of the ERC concept. The agenda for the webinar was as follows:

- *Welcome and Introduction* – KTH Royal Institute of Technology
- *Artificial Intelligence (AI) powered condition monitoring* – DB Cargo
- *Checkpoints for improving sharing of logistics information* – Trafikverket and RISE
- *Checkpoints deployment in Can Tunis for capturing yard movements* – ADIF, CEIT, and INDRA
- *Recognition of Hazardous Materials Plates/Placards and UIC Code* – ProRail and University of Twente
- *Closing remark* – KTH Royal Institute of Technology

The maximum number of viewers during the webinar was 109. The video recording of the webinar was published by Lindholmen Science Park (LSP) 2026-02-05 at <https://railsweden.lindholmen.se/en/news/webinar-european-railway-checkpoints-intelligent-monitoring-seamless-rail-freight-transport>. Table 3 presents the demonstrated topics in the webinar for each module as well as reference to which sections in this report the modules are further described and evaluated.

Table 3 Topics presented at webinar and related deliverable chapter

Module focus	Chapter	Lead partner
AI-powered condition monitoring	7.1 Module 1: AI-powered condition monitoring (Germany)	DB Cargo
Checkpoints for improving sharing of logistics information	7.2 Module 2: Checkpoints for improving sharing of logistics information (Sweden)	Trafikverket
Checkpoints deployment in Can Tunis for capturing yard movements	7.3 Module 3: Checkpoints deployment in Can Tunis for capturing yard movements (Spain)	Adif
Recognition of Hazardous Materials Plates/Placards and UIC Code	7.4 Module 4: Recognition of Hazardous Materials Plates/Placards and UIC Code (the Netherlands)	ProRail

7 Evaluation of ERC modules

This chapter presents the four modules, each focusing on a specific use case, for the *Modular System with harmonised procedures* of the European Railway Checkpoint (ERC) as described in chapter 5.

7.1 Module 1: AI-powered condition monitoring (Germany)

As part of the ongoing digitalisation of freight wagon condition monitoring, camera-based inspection systems are increasingly being deployed to automatically detect and assess (safety-)relevant components. The primary objectives of these systems are to reduce manual inspection effort, improve the objectivity of assessments, and enable the early identification of potential defects.

This chapter presents the results of multiple test and validation campaigns conducted to evaluate AI-based analysis methods for various freight wagon components. The scope includes, among others, wagon number recognition, brake pad thickness assessment, and the evaluation of additional running gear-related components. The analyses are based on real operational data and combine manual reference measurements, expert assessments, and automated model predictions.

7.1.1 Railway Checkpoints in Germany

Germany plays a key role in the deployment of Intelligent Video Gates (IVGs) within the European digitalization of rail freight. In cooperation with DB InfraGO, DB Cargo has installed 13 IVGs at eight major marshalling yards (see locations in Figure 7.1), achieving approximately 95% coverage of its wagon fleet (Kordnejad et al. 2024).



Figure 7.1 Locations of DB Cargo's IVGs installed in Germany

IVGs are designed to digitally support inspections by enabling high-resolution, automated data acquisition, thereby reducing processing times and improving operational efficiency in marshalling yards and terminals. In Germany, IVGs are installed at the hump in marshalling yards to ensure that every wagon passes the system prior to maintenance inspections, making them particularly suitable for network bottlenecks.

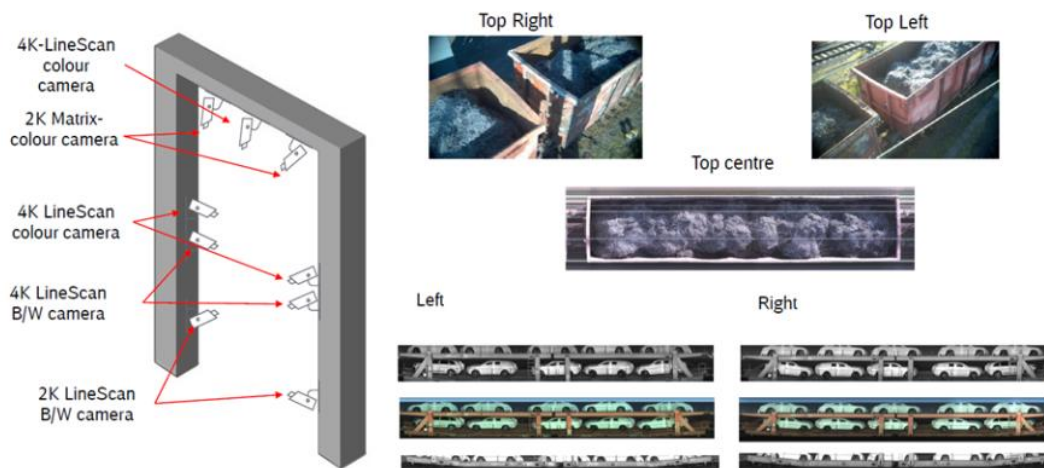


Figure 7.2 Standardized IVG of DB Cargo installed at 13 locations in Germany

A standard IVG consists of a steel frame equipped with multiple matrix and line-scan cameras that capture wagons from various perspectives, including side, top, and bogie views (see Figure 7.2). Line-scan images are stitched to form complete wagon representations for automated analysis. Wagon identification is achieved through a combination of OCR-based reading of UIC vehicle numbers and RFID scanning, ensuring reliable identification even under challenging conditions.

7.1.2 Use case: Condition monitoring - Bogie Analyzer

7.1.2.1 The basics of the Bogie Analyzer

The Bogie Analyzer employs a three-stage processing pipeline to analyse bogie components. The first stage implements a two-step object detection approach using artificial intelligence. Initially, the system detects wheelsets and buffers within the image. Subsequently, within the identified wheelset areas, the system performs detailed detection of specific objects of interest, including elements such as cover screw heads of axle caps, lifting T anchors and brackets, brake pads and tare springs.

The second stage focuses on scale and shear correction, utilizing the positions of detected screw heads to estimate scale and linear distortion parameters. This correction process compensates for scale and shear effects present in the images and provides a reliable approximation of wheel centres. The successful completion of this stage is critical for all subsequent use cases, as it enables proper undistorted angle measurements and conversion from pixel coordinates to metric units. The system requires a minimum of

three screw heads arranged in either a triangular configuration or three to four screw heads in a square configuration to proceed with processing.

The third stage encompasses use case processing, where the system analyses detected contours using scale information from the previous stage. The detected contours undergo processing through geometric approximations to obtain mathematical representations of the objects. Each use case implements custom business logic to derive key indicators regarding the state of bogie and wheelset components.

Overall, the Bogie Analyser is able to recognize nine different components across two types of wheelsets. As can be seen in Figure 7.3, the following elements are recognized in wheelsets with tare springs:

- brake pads
- tare springs
- four cover screws
- damper ring
- lifting T consisting of:
 - o lifting anchor
 - o lifting brackets
 - o lifting screw

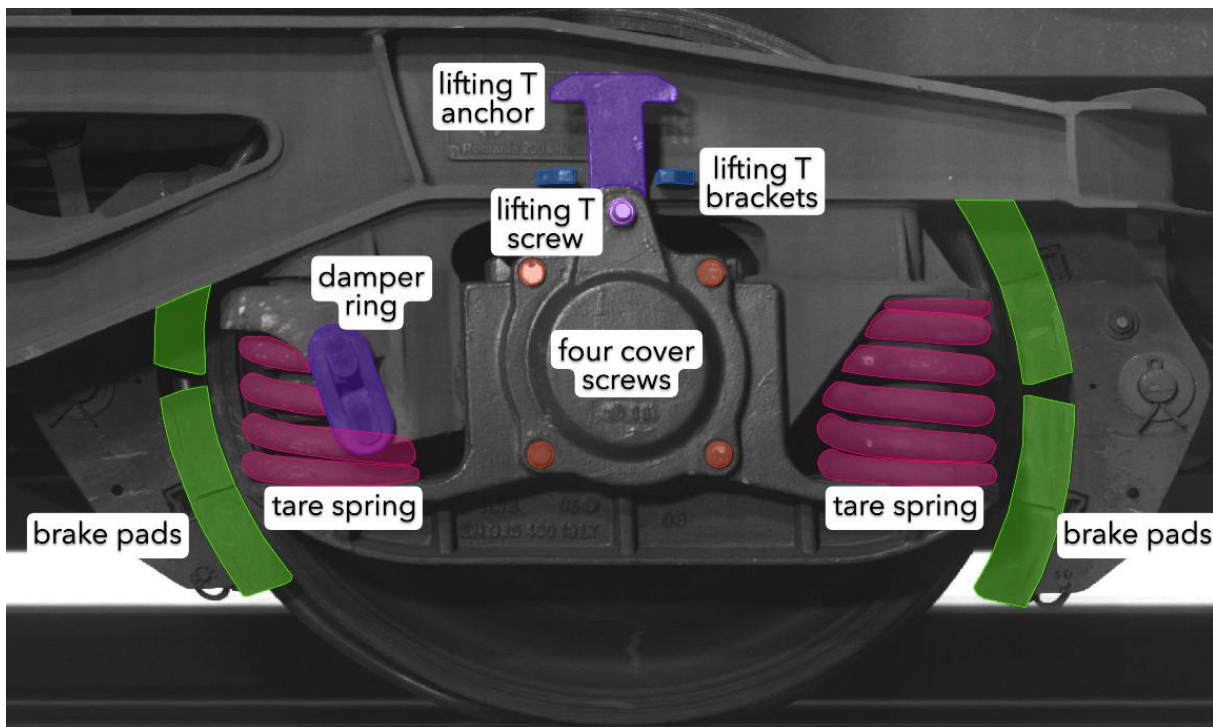


Figure 7.3 Analysed components (wheel set with tare springs and 4 cover screws)

For wheelsets with parabolic springs, these are (Figure 7.4):

- brake pads
- three cover screws
- spring bucket
- spring stop

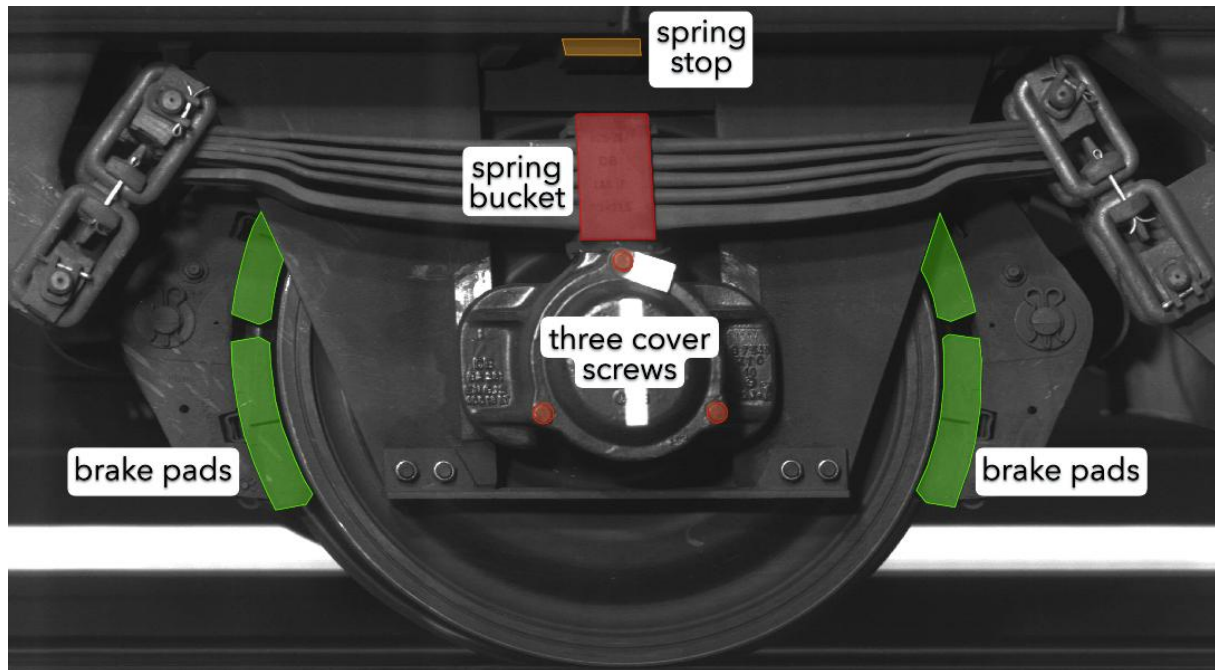


Figure 7.4 Analysed components (wheel set with parabolic springs and 3 cover screws)

7.1.3 Description of the recognized components

7.1.3.1 Brake pad

A brake pad on freight wagons is an essential component of the braking system, which generates friction by pressing against the wheel tread to reduce speed. Today, it is mainly made of modern composite materials to reduce noise emissions and wheel wear compared to old cast-iron brake pads.

7.1.3.2 Tare spring

Some bogies are equipped with coil spring sets. The outer spring (tare spring) has a second spring inside it. This second spring (load spring) has a smaller diameter and has a slightly lower height. This achieves a progressive spring effect: when unloaded, only the outer spring is active, the suspension is relatively soft. Above a certain load, the inner spring is also compressed.

Depending on their winding direction, coil springs can be divided into right-hand and left-hand (ascending) springs. On bogies for 20 t axle loads, the outer springs are right-hand (ascending), while on bogies for 22.5 t axle loads, they are left-hand (ascending). The winding direction of the inner springs is always opposite.

7.1.3.3 Cover screw

The cover of the axle-box on freight wagons is a component that protects the wheelset bearing from contamination, moisture and lubricant loss. The screws that secure the cover are used as reference geometries. More on this in Chapter 2.3.

7.1.3.4 Damper ring

Damper rings are important components on freight wagons that connect the spring assembly to the frame, compensate for longitudinal expansion and dampen vibrations through friction. They improve running smoothness by absorbing vertical and transverse vibrations via friction plates, especially at high axle loads.

7.1.3.5 Lifting T

A lifting T is a technical device that prevents the wagon body from detaching from the bogies or lifting off in an impermissible manner. It is relevant to safety to prevent derailments in the event of strong vibrations, cornering or accidents, especially in the case of unloaded or lightly loaded freight wagons.

7.1.3.6 Spring bucket and spring stop

A spring bucket is the central connecting element that firmly encloses and holds together the individual layers of a leaf spring in the middle. It ensures the transmission of force between the spring and the frame/axle holder and prevents the spring leaves from falling apart. The spring stop is the element that meets the spring collar if the freight wagon bounces too much, for example because it is overloaded.

The distance between the lifting T and the centre of the cover screws, as well as the distance between the spring bucket and the spring stop, can be used to determine the load status. The smaller the distances, the higher the load on a wagon. By looking at the distances on each wheelset of the freight wagon together, it is also possible to determine the load distribution within a freight wagon.

7.1.4 Reasons for selecting the components

During the selection of components for the Bogie Analyzer, extensive technical exchange was conducted with professionals who work with freight wagons daily. The objective was to identify which features and information would provide the greatest support for their operational tasks. The implementation of components was carried out in an incremental manner, integrating one component at a time into the Bogie Analyzer. The first component to be implemented was the brake pad, with the cover bolts included to enable scaling. Following initial tests, which confirmed reliable interaction of all components, additional elements were integrated. These components included the damper ring, the lifting T with its three distinct elements, and the tare springs. Shortly thereafter, the idea emerged to also detect the load condition. This led to the implementation of the detection of the components of the spring assembly as well as the spring retaining stop. As all these ideas emerged within a very short time frame, they were trained and implemented jointly within a single model of the Bogie Analyzer.

A key selection criterion was the visibility of the individual components. The components needed to be clearly identifiable by their contours and distinguishable from the background. Moreover, the aim was to generate value for the employees. Additional elements can be incorporated into the further development of the Bogie Analyzer.

Every component that is analysed is included in the GCU (General Contract of Use for wagons). This ensures that only important and relevant components are considered. Table 4 shows excerpts from the GCU.

Table 4 Excerpts from the GCU

Component	Code no.	Irregularities/Criteria/Notes	Action to be taken	Irregularity class
Leaf spring	2.1.5	Insufficient spring clearance: Vertical distance between buckle and fixed parts of body, underframe or bogie frame less than 15 mm • signs of recent contact between buckle and fixed parts of the underframe or bogie frame • signs of recent contact between wheel and underframe or wagon floor/body	Detach wagon	5
	2.5.3	Damper ring(s) missing or broken • contact marks	K	5
	2.5.3.1	- one ring per bogie		3
	2.5.3.2	- more than one ring per bogie	Detach wagon	5
Suspension system of Y 25 bogies or derived systems	2.5.5	Lifting T (safety catch) loose or missing	M	3
Brake block	3.2 3.2.1	Cast-iron brake block – missing – broken, cracked right through, even if still held together by its metal insert – worn so that thickness X near brake block is less than 10 mm	Replace. If not possible, K + R1 (isolate brake)	3
	3.2.2	Composite brake block – missing – radial crack from friction surface through to plate edge (except at the designated expansion joint) Friction material: – visible shelling of the friction material over more than one quarter of the block length, or metal inclusions – detached from back plate by more than 25 mm – cracked over more than 25 mm in direction of wheel circumference – lowest thickness X less than 10 mm	Replace. If not possible, K + R1 (isolate brake)	3

In short, the Bogie Analyzer enables the early identification of overloads or irregularities, reduces the need for manual inspections, and contributes to increased operational safety, efficiency, and maintenance quality in freight transport.

Additionally, the workload of trackside personnel is reduced as visual inspections become increasingly automated, minimizing the need for staff to work outdoors on freight wagons in adverse weather conditions.

7.1.5 Technical description of the process

7.1.5.1 Object detection

For instance, when analysing brake pads, the initial step involves detecting the wheel set areas, as illustrated in Figure 7.5. The first AI model is responsible for identifying them and returning the coordinates of the recognized areas. In a second step, these areas are analysed in more detail by a segmentation model.



Figure 7.5 Object detection of the wheelset areas

7.1.5.2 Segmentation

A second model provides a more detailed analysis of the cut-out wheelset area. This is necessary because the segmentation models cannot work with the original resolution of the whole wagon. Most AI models have a limited input size of 1024x1024 or 2048x2048 pixels and prefer square images. It is important to note that higher resolutions require more hardware resources to run these models efficiently. The output is presented as contours surrounding the objects to be detected (see Figure 3 and 4 in chapter 1). The thickness of the brake pads (green) can be calculated from these contours using an undisclosed algorithm.

7.1.5.3 Analysis of geometries

The components analysed by the Bogie Analyzer are examined using a variety of geometric methods to obtain the most accurate information. For some components, precise measurements of distances and dimensions are performed, while for others, angles and relative orientations are determined. Based on these geometric parameters, reliable conclusions can be drawn regarding function, wear condition, and load distribution. This differentiated analysis enables the Bogie Analyzer to reliably capture even complex interactions within the bogie and prepare the data for subsequent evaluation.

7.1.5.4 Using the example of the brake pad / cover screws

Geometric analysis is necessary because the segmentation model does not provide information about the condition of the brake pads. Therefore, the AI cannot decide whether the pads are sufficiently thick or damaged. Thickness is determined by analysing the geometry of the segments. This analysis includes various plausibility checks and the smoothing of the inner and outer contours of the pads (see Figure 7.6).

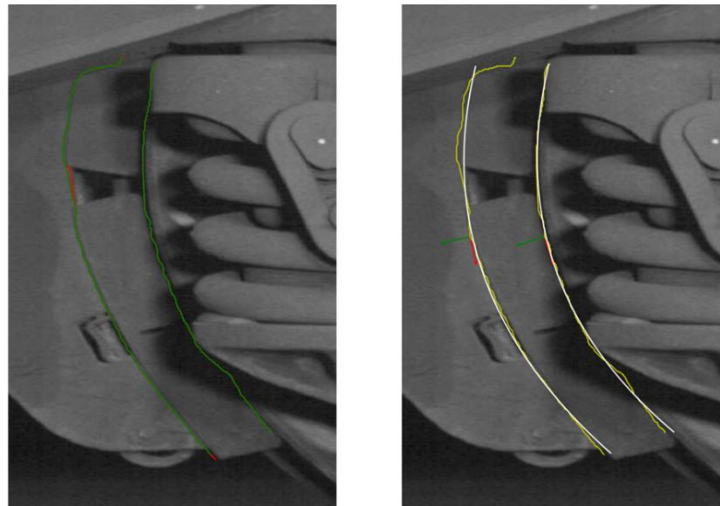


Figure 7.6 Left: Segmentation (AI). Right: Smoothing of the segment edges

As line scan images may be subject to stretching and compression along the X-axis, reference geometries such as bolts (see Figure 7.7) are used to provide a more precise representation of the pixel-to-millimetre ratio. The cover screws are detected using the same segmentation model as the brake pads, and they were chosen as a reference because their spacing is standardized across different bogies.

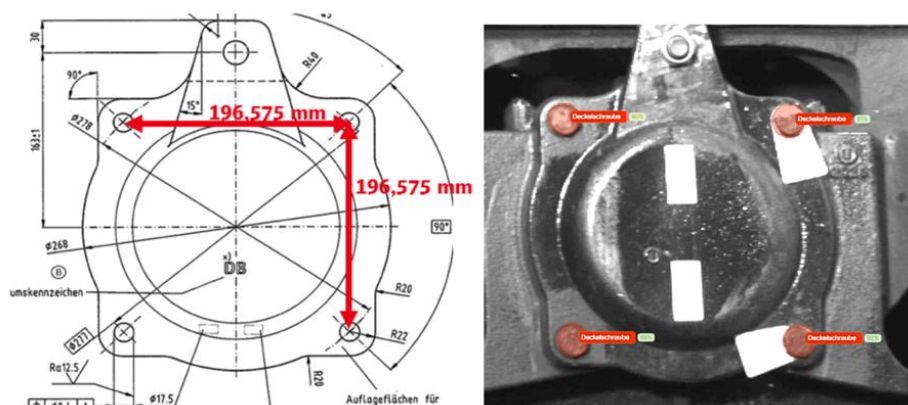


Figure 7.7 Reference geometries for determining the pixel to millimetre ratio

7.1.5.5 Using the example of the lifting T / damper ring

The example of the lifting T and the damper ring illustrates how the Bogie Analyzer utilizes angular measurements to analyse components. To determine the angle of the lifting T the

lifting T screw serves as a reference point, allowing precise calculation of the component's orientation. For the damper ring, a diagonal is drawn across the component and compared to a vertical reference line to accurately measure its angle. The vertical reference is derived from the alignment of the cover screws, while potential shear effects introduced during the stitching of line-scan camera images are compensated for to ensure robust and reliable angle measurement. In this way, even subtle deviations in component alignment can be detected, providing insights into their condition, functionality, and potential damages within the bogie. Examples of the measured angles can be seen in Figure 7.8. When installing a damper ring, a mounting angle of 27° must be maintained.

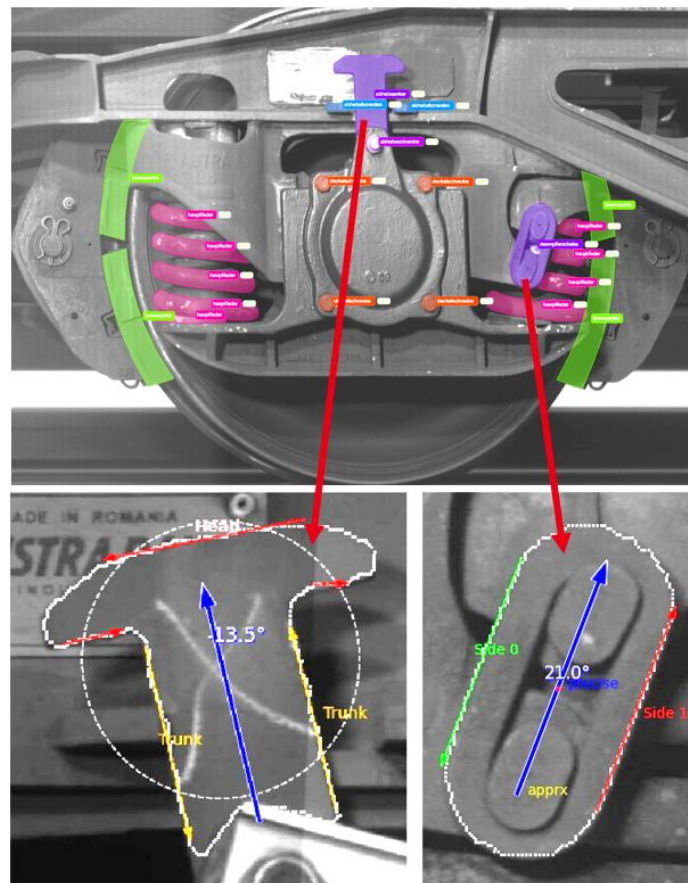
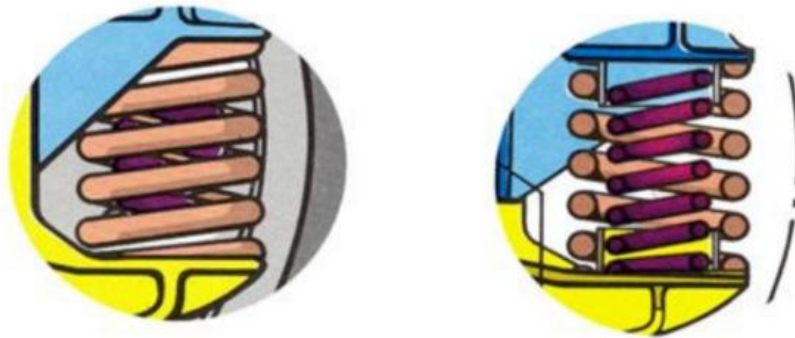


Figure 7.8 Angle of Left: Lifting T. Right: Damper ring

7.1.5.6 The tare springs

As translated from the German "Handbook 93612 – Technical Rules for the Operation and Handling of Freight Wagons" (DB Fahrwegdienste GmbH., 2021), the following can be stated:



Main spring (tare spring) and auxiliary spring (load spring)

The winding direction of the coil springs is of critical importance.

For wheelsets with a load of 20 t, the outer compression spring is right-hand wound, while the inner compression spring is left-hand wound.

For coil spring assemblies used in Y25-type bogies with a wheelset load of 22.5 t, the outer springs are left-hand wound, while the inner compression springs are right-hand wound.

In this context, it must be ensured that:

- all spring sets of a bogie within a wagon have the same winding direction,
- within a single spring set, the inner and outer coil springs are installed with opposite winding directions.

Incorrect installation of coil springs may result in entanglement of the springs, which negatively affects the spring characteristics and thus the running behaviour of the wagon. If the above-described defect is detected, the wagon must be withdrawn from service.

For this reason, the Bogie Analyzer compares the pitch of the individual tare spring windings on the left and right sides of a wheelset. The calculated pitches are evaluated against each other to determine whether they match or exhibit deviations. Based on this comparison, conclusions can be drawn regarding the correct design and installation of the tare springs.

Figure 7.9 illustrates an example of a wheelset with non-matching tare spring windings.

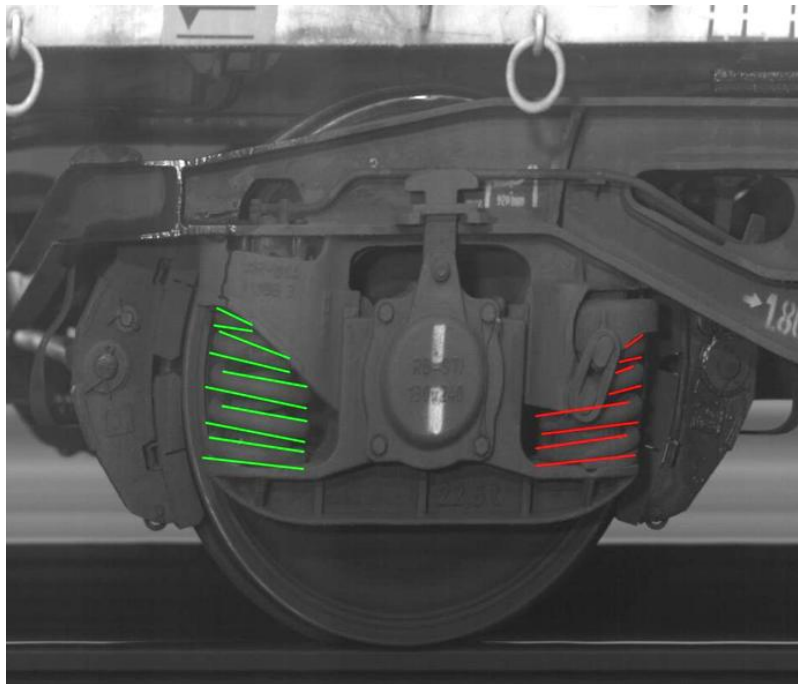


Figure 7.9 Wheelset with non-matching tare springs

7.1.5.7 Using the example of load detection

For wheelsets equipped with tare springs, the components required for load detection are the cover screws and the lifting T brackets. In this case, the distance between the centre point of the cover screws and the lower edge of a reference line of the lifting T brackets is determined (see Figure 7.10).

For wheelsets equipped with parabolic or leaf springs, the distance between the upper edge of the spring bucket and the lower edge of the spring stop is measured (see Figure 10). Both distances are specified in pixels as well as in millimetres.

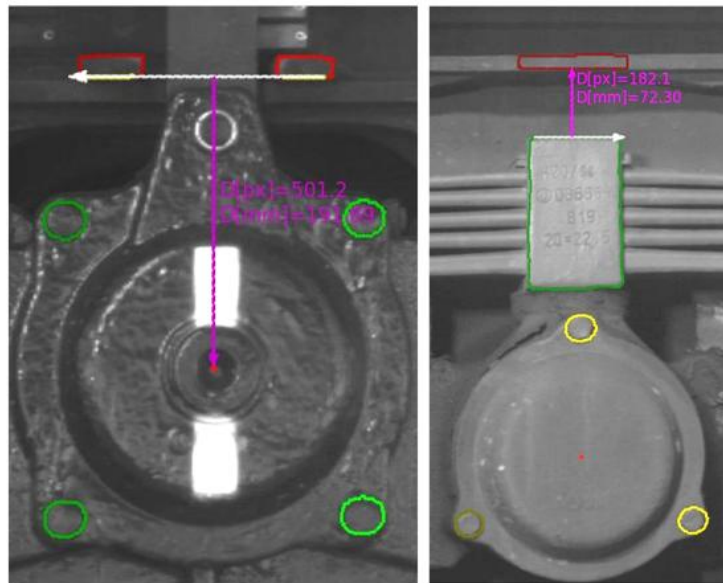
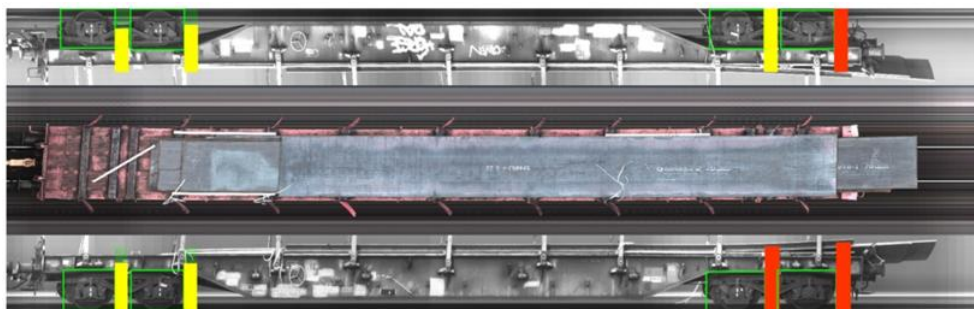


Figure 7.10 Distance of Left: Centre point of cover screws to lower edge of lifting brackets

By combining the measured distances of all wheelsets of a freight wagon, conclusions can be drawn regarding the total load weight and the load distribution. In particular, for closed freight wagons or wagons carrying bulk goods, it is often not possible to visually assess whether the load is evenly distributed. In such cases, the Bogie Analyzer enables a reliable assessment of load distribution. An example of load distribution determination is shown in Figure 7.11. The image illustrates that the load is significantly shifted to the right and even protrudes beyond the wagon boundary. This may be caused either by incorrect loading or by load displacement during operation.

The load is identified as faulty in the image, which is additionally indicated by the bars displayed. The filled bars next to the wheelsets provide information on the distances between the lift-off bracket and the wheelset centre, and thus on the degree of suspension compression. With increasing compression, the load on the wheelset increases, which is reflected by an increase in the bar size. The lower part of the Figure 7.11 visualizes the resulting load distribution, clearly showing that the centre of gravity is shifted far to the right.



Weight distribution

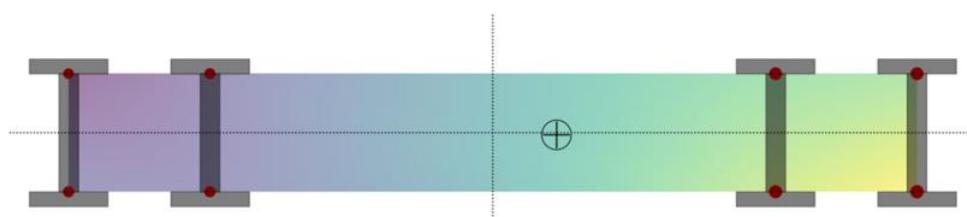


Figure 7.11 Weight distribution of an open flat wagon

Another example of the results provided by the Bogie Analyzer is shown in Figure 7.12. The Figure illustrates the load distribution between the front and rear of the wagon as well as between the left and right sides. In addition, the weight ratios between front and rear and between left and right are displayed.

UIC: - Date: - Location:
URL:

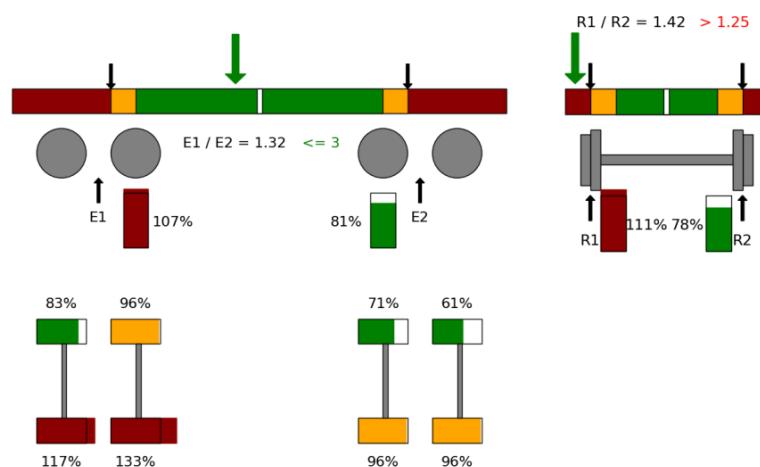


Figure 7.12 Example of load distribution results generated by the Bogie Analyzer

According to the applicable loading guidelines, bogie wagons shall not exceed a mass ratio of 3:1 between the individual bogies. The corresponding permissible limits are presented in Figure 7.13.

3.3 Distribution of the load

The load should be distributed as evenly as possible on the wagon, without exceeding the maximum mass per axle.

The load distribution should be such that the following ratios are not exceeded:

- for 2-axle wagons: ratio of 2:1 between the masses per axle,
- for bogie wagons: ratio of 3:1 between the masses per bogie,
- ratio of masses per wheel: 1.25:1¹⁾

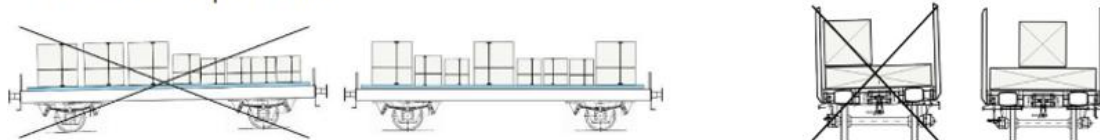


Figure 7.13 Excerpt from the loading guidelines

7.1.6 Image Requirements for the Bogie Analyzer AI Model

To recognize the contours and see where one component ends and the next begins, the images need a certain brightness and contrast between the components. For this reason, the number of pixels should not be too low. However, it should not be too high either, since most AI models have a maximum input size of 1024x1024 or 2048x2048 pixels and prefer square images.

The camera perspective (vertical or horizontal orientation, and lateral positioning) and the viewing angle have a noticeable influence on detection performance. The results indicate that optimal accuracy is achieved when the camera is aligned horizontally. Ideally, the camera should be positioned at the height of the cover screws up to the lifting T screw of the wheelset.

7.1.7 Evaluation of the system - KPIs within the framework of ERC

7.1.7.1 Achieving KPIs

A series of workshops was conducted to establish a common set of key performance indicators (KPIs) for the ERC framework. These workshops enabled alignment among all partners and ensured a shared understanding of the overarching evaluation criteria. In addition to the common KPIs, each partner and system module defined supplementary KPIs that specifically reflect the objectives and functional scope of their respective ERC contributions. The KPIs relevant to our system will be described and analysed in greater detail in the following sections of this report.

7.1.7.2 Common KPIs

We have agreed on KPIs that each ERC can fulfil in its own way. The following three KPIs show how well our system works and how effective optimisation and utilisation by the system is.

This chapter describes the selected KPIs used to evaluate the performance and data quality of the ERC-based wagon monitoring process. The KPIs focus on:

- the confirmation of wagon sets,
- data coverage, and
- end-to-end processing time.

Together, they provide insight into both operational reliability and data usability within the current system architecture.

7.1.7.2 Time to confirm wagon set

The time to confirm wagon set measures the duration how long an employee needs to walk along the train, passes every wagon and check the wagon number and wagon order with the one in the system.

The IVGs are located at a point where individual wagons pass the IVG sequentially. Consequently, the IVG records wagon-level data from every passing wagon before the final train composition is physically established. What is not a problem as all wagons which passes the gates are recorded by the IVG, with image, wagon number, timestamp and location. The sequence in which wagons pass the detection gates is recorded in the IVG dataset. This means that the needed wagons in the train can be compared with the recorded wagon order of the gate as their order is already the one, they're passing.

While combining the wagon data set the order can be confirmed. This automated matching process would allow the wagon order and thus the complete wagon set to be confirmed without the intervention of a human. So, this results into a reduction of the confirmation time through a human down to 0 s¹. Just in the significantly reduced 1.3%² passed wagons, the correct wagon number couldn't be obtained. This means that in 1.3% cases where the wagon number is missing the employees would still need to check the wagon set manually in the tracks.

In total this means a reduction from 30 sec/wagon to 0 s/wagon in 98.7% of the times the wagon set needs to be confirmed. In 1.3% of the time the needed time would still be 30 sec/wagon. To fully automate the confirmation, we need to increase the wagon number recognition of the IVG up to 100%. The wagon number recognition rate remained at a consistently high level throughout 2025 (see Figure 7.14). The target value of 95% was clearly exceeded in all months. Monthly recognition rates ranged between 97.4% and 99.3%, with peak performance observed during spring and summer. A slight decline can be seen towards the end of the year; however, overall performance remains stable and well above the target threshold.

¹ In theory as it is not implemented

² 37,197 wagon numbers of 2,765,725 (=1.3%) are not identified in 2025

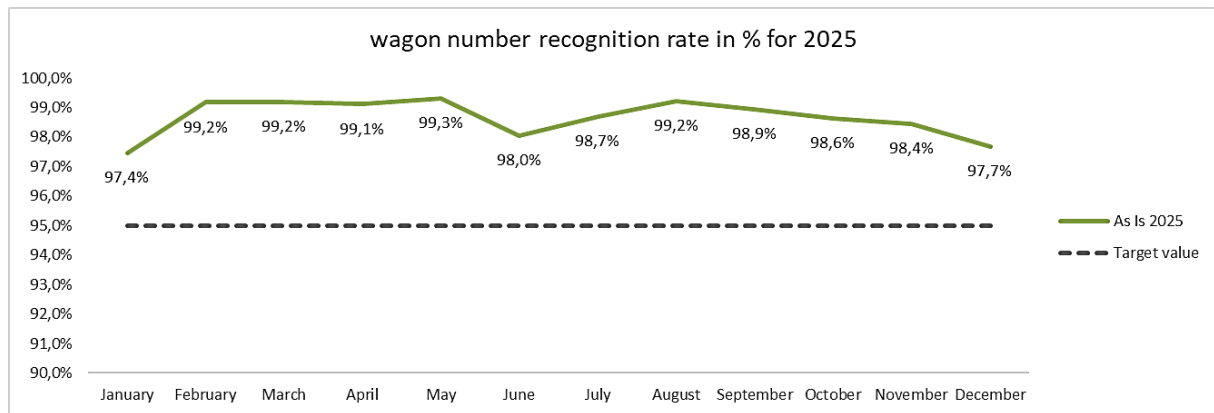


Figure 7.14 Wagon number recognition rate 2025

In addition to the increase of the wagon number recognition in the future could the system more reliable with more operational information matched, as the GPS data of our wagons or at least the track number.

7.1.7.2 Wagon coverage

Wagon recording for maintenance purposes is particularly important in the case of damaged wagons. For this reason, this module focuses on the recording of damaged wagons by IVG in Germany. To this end, damaged wagons from a period of 12 months were collected and compared with the IVG database. The result shows a very high level of recording and therefore confirms the initial use case for the IVG. Figure 7.15 illustrates the monthly share of wagons that entered the workshop within 30 days after passing through an IVG, comparing the years 2024 and 2025. In 2025, this share is consistently higher than in 2024 and generally ranges between approximately 70% and 76%. Higher coverage levels are particularly evident during the spring and autumn months.

Across the entire year 2025, 114,985 wagons passed through the IVG and entered the workshop within the subsequent 30 days, out of a total of 156,870 workshop cases. This corresponds to an overall coverage rate of 73.3%. The underlying workshop data therefore indicate an improved level of system coverage in 2025. However, the results also show that not all workshop entries are preceded by an IVG. The target was to capture 75% of the damaged wagons which go to a maintenance workshop with the built IVG.

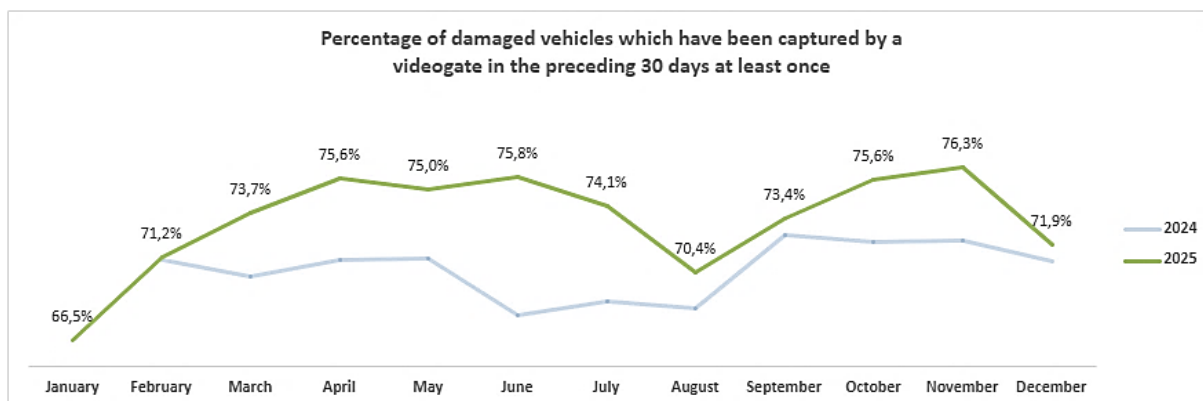


Figure 7.15 Percentage of damaged vehicles captured by a IVG in preceding 30 days at least once

7.1.7.2 Processing Time

We consider the process time in two parts: firstly, the process time between the wagon passing through and thus being captured, up to and including storage and availability, is relevant (Figure 7.16) and secondly, in the next and currently important step, the processing time between image input and AI processing is relevant (Figure 7.17).

This KPI encompasses the entire data pipeline, including:

- initial data capture by the IVG,
- data transmission and storage,
- data validation and enrichment,
- and final availability for use in downstream systems.

Processing time reflects the system's responsiveness and its suitability for near-real-time or time-critical applications. Long processing times may indicate bottlenecks in data handling, system integration, or manual intervention steps. Conversely, short and stable processing times suggest a mature and well-integrated data pipeline. In our case, most of our uses aren't time-sensitive, so we don't need real-time data. But it would be technically possible if we used a higher service level.

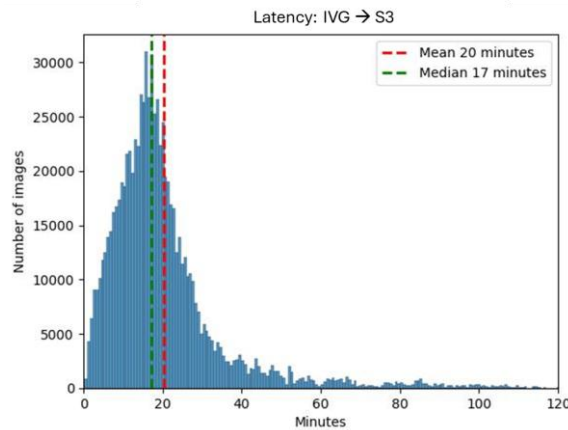


Figure 7.16 Latency from IVG to S3 bucket

The average processing time from wagon registration to availability in the object storage service S3 (Amazon simple storage service) bucket is 20 minutes, with a median of 17 minutes. The objective was to maintain the average below 30 minutes, as real-time processing is not required, as previously noted.

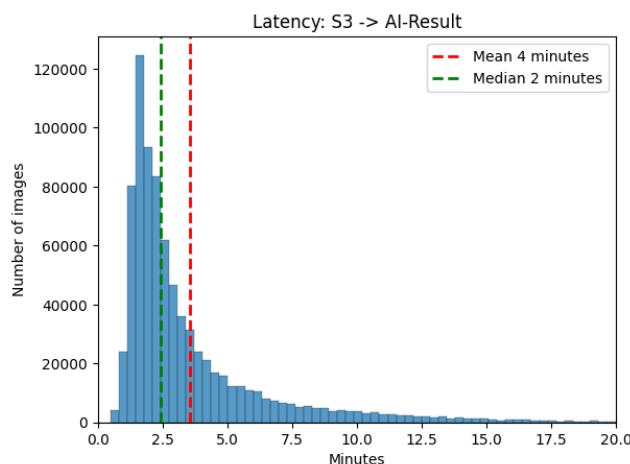


Figure 7.17 Latency from S3 bucket to get the AI results

The processing time, from retrieving data from the S3 bucket to the AI results, averages 4 minutes, with a median of 2 minutes. This ensures secure and rapid damage assessment, which can then be further processed and checked by humans. The target of the ai processing was 10 minutes.

So, to conclude the KPI of the processing time the mean latency of the image transfer is 20 minutes and the latency of the ai processing is 4 minutes. The target of the image transfer is an average of 30 min and the target of the ai processing is 10 minutes.

To summarise the overall KPIs can be see together in Table 5.

Table 5 Summary of overall KPIs

KPI	Evaluation Basis	Measured Results	Target Results (Current Phase)	Target Met
Damaged wagon coverage (IVG match)	Damaged wagons from period 2025 matched against IVG records in Germany; focus on maintenance-relevant wagons	73.3% of workshop cases were covered by a prior IVG, with 114,985 out of 156,870 wagons entering the workshop within 30 days after being recorded	75%	Not always, in this period not
Processing time	Latency from wagon registration/capture to data availability in S3	Mean 20 min; median 17 min	Average < 30 min (real-time not required)	Yes
	Latency from data retrieval in S3 to AI output availability	Mean 4 min; median 2 min	≤ 10 min	Yes
Time to confirm wagon set	Reduction in confirmation time versus the manual, employee driven process (baseline)	When a wagon passes the IVG and is correctly captured (including wagon number recognition), confirmation is fully automated; human action required: 0 seconds	0 seconds (automatic confirmation; manual fallback only if number recognition is missing)	Yes
Coverage of wagon variations	Share of wagon by country successfully processed with valid, usable analysis results	81.2% German operators and 18.8% international Operators (1867862 German wagons vs. international wagons 432164 in 2025)	Share of international wagons ≥ 15%	Yes
	Share of wagon by operator successfully processed with valid, usable analysis results	79.7% of DB Wagons and 93.6% of other operators' wagons	Share of analysed wagons ≥ 70%	Yes
Detection rate of wheelsets	Share of validly detected wheelsets over all wheelsets in the dataset	99.9% (77,525 of 77,599)	99%	Yes

7.1.7.3 KPIs within the framework of the module (Bogie Analyzer)

To evaluate the performance and reliability of the AI based wagon and bogie analysis system, several KPIs are defined. These indicators are intended to assess system coverage, detection capability, and analytical accuracy across different wagon configurations and components.

7.1.7.3 Bogie Analyzer coverage of wagon variations

The coverage of wagon variations describes the extent to which the AI model can successfully analyse different wagon types and country specific configurations. And it should also show what expansion opportunities there are for the Bogie Analyzer across Europe. There is more than one step to describe how the coverage can be measured and what its use is.

First, for the comprehension the distribution of the touchpoints captured is shown in the Figure 7.18. The variation of the different wagons based on the touchpoints is a distribution of 60.8% DB wagons and 39.2% other wagons. The other wagon can be ones from other companies, rented ones and from other countries.

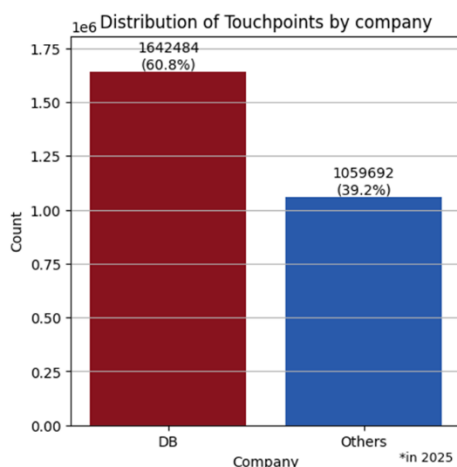


Figure 7.18 Distribution of touchpoints of the IVG

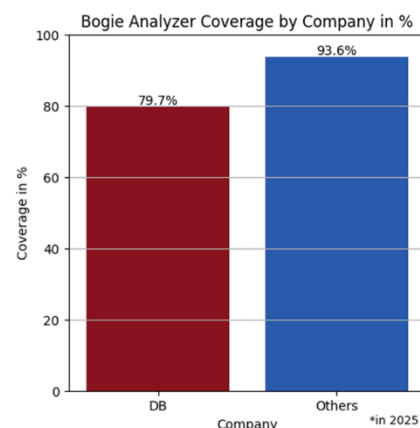


Figure 7.19 Bogie Analyzer coverage

Based on the distribution of the captured images the coverage on them for the AI model of the Bogie Analyzer covers 79.7% of the DB wagons and 93.6% of the other wagons, as illustrated by Figure 7.19.

In addition, in Figure 7.20 you can see the distribution of the bogie detections for the top 10 countries. There are countries covered as Netherlands and Sweden.

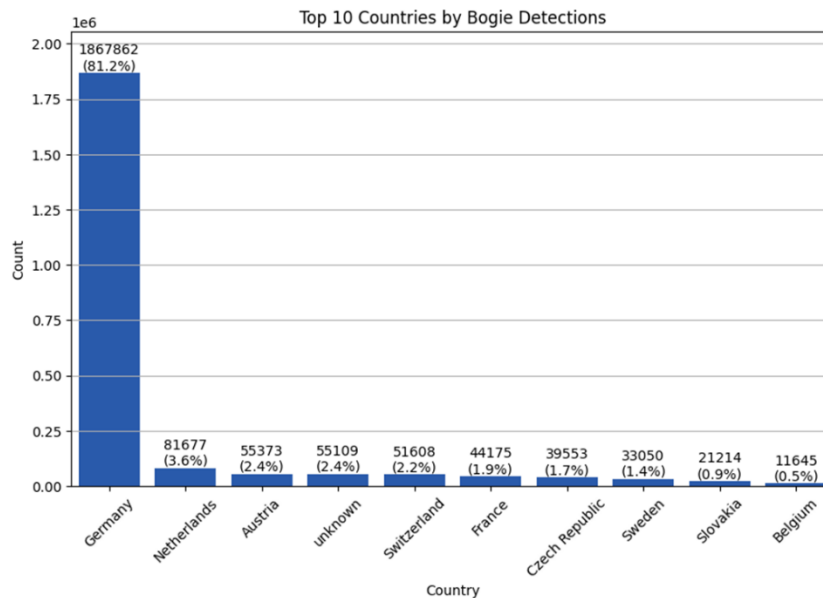


Figure 7.20 Bogie Analyzer detection per country

In Figure 7.21 the overall distribution of the bogie detection is shown, 81.2% are detected in the DB wagons and 18.8% on the others.

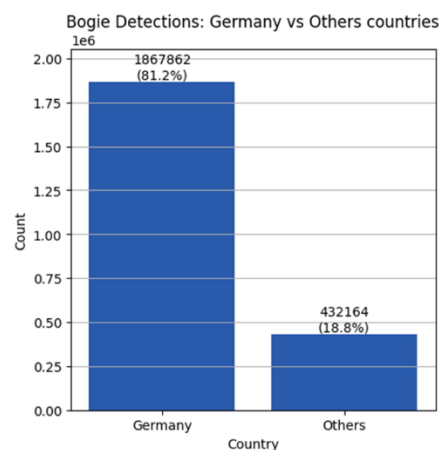


Figure 7.21 Bogie detections

7.1.7.3 Detection rate of wheelsets

The detection rate of wheelsets is measured as the number of correctly identified wheelsets per 1,000 recorded wagons. This indicator reflects the capability of the system to detect the region of interest under operational conditions. Only detections that are confirmed using verified ground truth data are counted as correct. Normalization per 1,000 wagons enables comparison across datasets of different sizes. To keep things simple and reduce the complexity of the analysis options (due to different bogie designs,

wagon types, number of wheel sets required, etc.), the following section looks at a selection of 1,000 wagons from three types and the performance of wheel set detection on them. Figure 7.22 shows that 99.9% of the wheelsets of these carriages and types were successfully recognised. This was compared with the ground truth. The target was to detect 99% of 1,000 wheelsets.

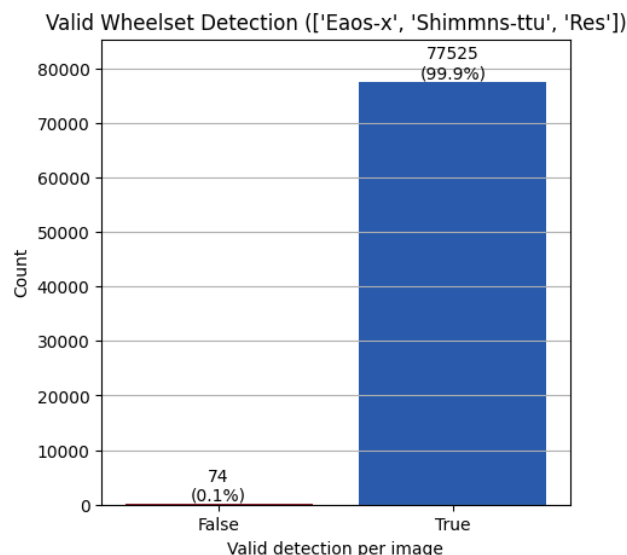


Figure 7.22 Valid wheelset detections

7.1.7.3 False positive rate per component for the Bogie Analyzer

As a complementary performance indicator, the false positive rate is used to evaluate the Bogie Analyzer. To assess this, the false positives for the components were tested with the following tests for the different components with the experts evaluating. In the following passages, you will read about the tests and the corresponding false positive rates. This indicator is particularly relevant for operational deployment, as an elevated false positive rate may lead to unnecessary inspections or maintenance actions.

7.1.7.4 Conduction of Tests and Evaluation: Detection of loose lifting T

The lifting T anchor and lifting T brackets are two of the components detected by the Bogie Analyzer model. As loose lifting Ts are a possible safety hazard, one of the damages that are detected is a loose lifting T. This is achieved by measuring the position of the device in relation to the brackets. If the two components touch, it is very likely that the lifting T is loose.

Detection results were evaluated by three wagon masters specialized on evaluating wagon images ("Wagenmeister Bildbasierte Befundung", WBB) in a test that used 421 distinct images with 122 images containing detected loose lifting Ts and 313 images without loose lifting Ts (105 of which contained detected missing damper rings).

WBBs evaluated wholly unmarked images and were asked to mark any loose lifting Ts. These results were then evaluated matches between the loose lifting Ts marked by the Bogie Analyzer and by the WBBs. The findings have been classified into 4 categories:

- Correct: Both the Bogie Analyzer and all WBBs marked the device as loose
- Wrong: The Bogie Analyzer marked the device as loose, but all WBBs didn't
- Missing: The Bogie Analyzer didn't mark the device as loose, but all WBBs did
- Uncertain: Any result where the WBBs differed in opinion were marked as uncertain and excluded from the results as the veracity of the result cannot be verified

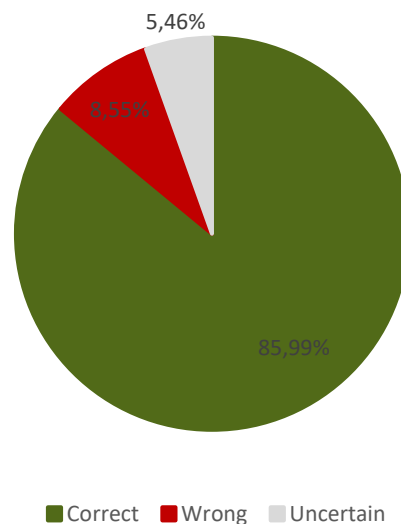


Figure 7.23 Analysis evaluation results (Lifting T)

Figure 7.23 shows the evaluation results: about 85.99% of the images have been classified correctly by the Bogie Analyzer (*correct* detection results and images that do not contain *missing* lifting T 8.55% of the images have been classified incorrectly (which contain *wrong* detection results and *missing* lifting Ts). 5.46% of the results could not be ascertained due to inconsistent classification by the WBBs.

7.1.7.5 Conductions of tests and Evaluation: Detection damper rings and tare springs

In a test conducted the damper ring, 139 images marked as defective by the Bogie Analyzer were checked by a human. The following categories were categorized:

- insufficient image quality
- invalid
- valid

The first category (insufficient image quality) describes cases in which insufficient image quality was the decisive factor for classifying an image of the damper shackle as faulty. This includes images that were either blurred or significantly stretched or compressed due to the characteristics of the line-scan camera used.

Category 2 (invalid) includes images that were incorrectly classified as faulty despite the absence of an actual defect.

Category 3 (valid) refers to images that were correctly classified as faulty. The defect was primarily detected based on an impermissible angle of the damper ring. In addition, the defect can also be identified by evaluating the measured distance between the spring cap and the bogie frame, which serves as a supplementary criterion to validate the angular measurement. The relevant distance is indicated by a red arrow in Figure 7.24. According to the applicable specifications, the distance between the spring cap and the bogie frame must not be less than 2 mm (Figure 7.25).

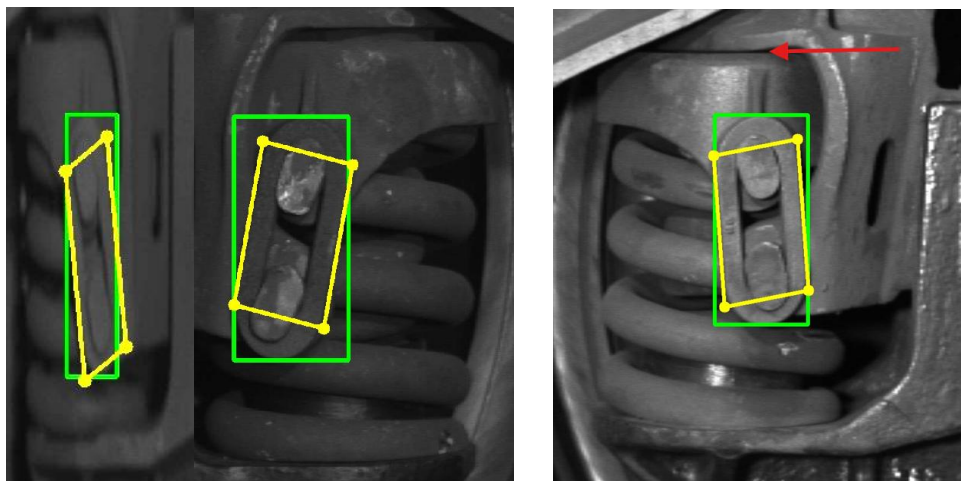


Figure 7.24 Left: damping image quality. Middle: damping invalid. Right: damping valid

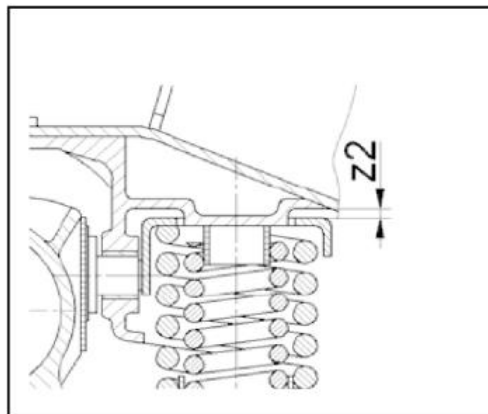


Figure 7.25 Vertical distance between the spring cap and the bogie frame (<2 mm)

The evaluation results of the damper ring classification are presented in Figure 7.26. A total of 56% of the images were classified as *valid*, indicating correct identification of defective damper rings by the Bogie Analyzer. 9% of the images were classified as *invalid*, representing cases in which the system falsely detected a defect despite the absence of an actual fault. For 35% of the images, a reliable classification was not possible due to insufficient image quality, which limited accurate evaluation.

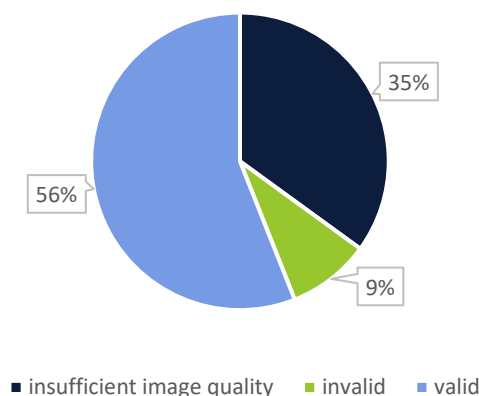


Figure 7.26 Analysis evaluation results (damper ring)

A similar approach was used with the results of the tare spring tests. It emerged that the model sometimes delivered incorrect results for fully loaded wagons with heavily compressed springs, as the individual spring coils were sometimes so close together that it was almost impossible to detect any incline. However, to obtain statistically meaningful results, additional data points are required. The defect under investigation occurs very rarely but is severe in nature, as it can lead to the withdrawal of the freight wagon from service. Consequently, an insufficient number of image samples is currently available to fully validate the results. For this reason, the test campaign is still ongoing.

7.1.7.6 Conducting of tests for the brake pad

To avoid assessing the Bogie Analyzer exclusively based on the false positive rate, several complementary evaluation approaches were applied to obtain a more comprehensive and differentiated performance assessment.

As part of a larger test campaign (see Figure 7.27) focusing on brake pads thickness assessment, a survey was conducted involving technically qualified staff. The participants were shown brake pads of various thicknesses and were asked to estimate the thickness in millimetres. The actual brake pad thickness was determined through manual measurements using a measuring device. Measurements were taken at the upper, central, and lower positions of each brake pad to identify potential wedge-shaped wear

patterns. Furthermore, the thickness values estimated by the AI-based model were included in the evaluation.

The results indicate that for thicker brake pads, the estimations provided by the participants show a larger spread and tend to overestimate the actual measured brake pad thickness. In contrast, for thinner brake pads, the estimations are more consistent and tend to underestimate the true thickness.

Overall, the average deviation of the human estimations from the measured brake pad thickness amounts to 5.4 mm. The uncertainty range stated by the participants themselves is, on average, 5.5 mm. By comparison, the AI model demonstrates a significantly higher accuracy, with an average deviation of 1.8 mm, referring to the thinnest measured point of the brake pad. Based on these tests, the AI based approach achieves a higher precision and less variation compared to the estimation by experts.



Figure 7.27 Validation approach for brake pad thickness

7.1.7.7 Planned Tests and Validation of load detection

Further tests are required to comprehensively validate the load detection functionality of the Bogie Analyzer. Tests with freight wagons loaded under different conditions shall be conducted to verify the accuracy of distance measurements across varying load states. Both evenly and unevenly distributed loads need to be considered.

In addition, tests under real operational conditions are planned to evaluate the influence of environmental factors such as contamination, lighting, and weather on detection performance. The objective of these investigations is to further improve system robustness and ensure reliable load detection in practical operation. For this purpose, weighing records from track scales have been collected over several months to ensure sufficient variation in measurement values. In addition, different types of freight wagons are included to cover a wide range of designs and loading scenarios.

7.1.7.8 Summary of KPIs for bogie component detection and validation

The results of the tests and illustrates the envisioned KPIs as well as their status of target achievement are summarised in Table 6.

Table 6 Summary of bogie detection and validation

Component	Evaluation Basis	Measured Results	Target Results (Current Phase)	Target Met
Lifting T	421 images evaluated by three trained wagon inspectors (WBB). 122 images with loose lifting T and 313 without	Approx. 86 % correctly classified. Around 9 % incorrect. About 5 % inconclusive due to differing expert assessments	≥ 85 % correct classification on evaluable images, with inconclusive cases below 10 %	Yes
Damper ring	139 images flagged as defective by the Bogie Analyzer and reviewed by a human expert	56 % confirmed as valid defects. 9 % false positives. 35 % not evaluable due to insufficient image quality	≥ 50 % confirmed valid detections on evaluable images, with false positives below 15 %	Yes
Brake pad	Survey-based human estimation compared with measuring device measurements and AI-based thickness detection	AI model shows mean deviation of 1.8 mm ; human estimations show mean deviation of 5.4 mm	AI-based thickness estimation with deviation ≤ 2.0 mm	Yes
Tare spring	Initial tests with limited number of defect cases. Expert review ongoing	No statistically robust accuracy value available	Qualitative validation of detection feasibility and identification of limiting conditions	Not yet
Load detection	Planned validation using operational data and track scale records	Quantitative results not yet available. Tests ongoing	Completion of controlled test covering different load states and wagon types	Not yet

7.1.8 Analysis of the usability of the AI model on data other checkpoints

7.1.8.1 Processing of the dataset of the checkpoint in Gothenburg

This evaluation was conducted on a dataset from a camera gate in Sweden, consisting of side-view recordings of undercarriages and wagons. The images have a resolution of 2048 pixels in height, with approximately 20% (400 pixels) of the image height dedicated to capturing bogies and wheelsets. The width of the images varies between 4000 and 15000 pixels.

Detection Performance and Limitations

The evaluation dataset comprised 1,132 images containing 4,212 wheelsets. The object detection system successfully identified 72.5% (3,055) of the wheelsets present in the images. However, only 17% (512) of the detected wheelsets could be successfully processed through the scale and shear correction stage and advanced to the use case processing stage. This significant reduction in processable data stems primarily from the failure to detect screw heads, which is predominantly attributed to false negatives resulting from the AI system not being trained specifically for this camera gate configuration and insufficient image quality. Additionally, some screw heads remain undetectable due to obstruction by carriage structures.

The dataset revealed a complete absence of wheelsets with parabolic leaf springs after the scale and shear correction stage, rendering the spring collar and catch distance use case un-processable. Among the wheelsets that successfully passed preprocessing, 82.5% exhibited coil spring suspension, while the remaining 17.5% remained undetermined, with neither tare spring nor parabolic leaf spring configurations identified.

Use Case Analysis Results

The spring bucket to spring stop distance analysis could not be performed due to the absence of relevant data after preprocessing. The lifting T anchor angle analysis achieved a 71% success rate among available processed wheelsets, with image resolution proving just adequate for recognizing and evaluating the geometric properties of the anchors.

The lifting T brackets collision detection faced significant challenges due to limited bracket detection rates. Only 9% of processed cases had either left or right bracket detected, while a mere 0.5% had brackets detected on both sides. Since reliable collision prediction requires bilateral bracket detection, the limited data prevented comprehensive analysis. The average dimensions of bracket bounding boxes measured only 12 by 6 pixels, indicating resolution constraints that limit the system's potential even with retraining efforts.

The lifting T bracket distance measurement suffers from similar detection limitations but requires only unilateral bracket detection. While the system demonstrates functional capability with limited accuracy, camera perspective introduces measurement bias, causing actual vertical distances to appear shorter than ground truth values. This

systematic error could potentially be addressed through correction factor implementation.

Tare spring direction analysis proved particularly challenging, with only 30% of cases providing sufficient data for algorithmic evaluation. The remaining 70% lacked adequate coil detections on one side, as the evaluation requires at least five detected visible coil segments on both left and right sides. Even successful processing cases yield statistically questionable results due to the camera's downward viewing angle on coils, which invalidates the line-based approximation of coil windings that the algorithm relies upon.

Damper ring analysis demonstrated more promising results, with 86% of available detections successfully analysed. The primary failure modes trace back to poor contour quality rather than fundamental algorithmic limitations. Similarly, brake pad analysis achieved an 80% processing rate among available wheelset datasets with brake pad detections, though only 50% of these (40% of total) produced valid and reliable results. The main impediments include noisy contours and partial obstructions of the brake pad components.

Conclusions and Recommendations

The current Bogie Analyzer implementation cannot be deployed at Swedish camera gates without prior adjustments. Targeted adaptations and additional AI retraining will be required to address the detection rate deviations and quality-related challenges identified during this evaluation. Among the evaluated use cases, lifting T analysis and damper ring evaluation currently show the highest potential for successful implementation.

The tare spring direction analysis is not recommended for future deployment under the present camera configuration, as the required orthogonal viewing perspective is not consistently provided. Certain technical constraints may continue to influence system performance even after optimization measures. In particular, camera perspective, image resolution, and overall image quality define practical performance boundaries. Furthermore, the algorithms were originally developed for German wagon configurations and may require adaptation to reliably process foreign wagon types with differing wheelset and bogie geometries.

Some challenges are expected to remain difficult to fully resolve, especially in cases where critical components (e.g., cover screw heads or coil windings) are partially obstructed. Even with extended AI retraining and parameter optimization, the achievable recognition rate for certain use case-relevant components is expected to remain around 80%. Considering additional quality-related influencing factors, the overall system performance under the current framework is projected to reach approximately 50%, which should be considered when evaluating deployment scenarios.

7.1.9 Validation of usage on data from ProRail

In additional tests using images from the Netherlands, the results were comparable to the ones with the Swedish data. In these cases, as well, several wheelsets could not be analysed due to the failure to detect the cover screws, caused by insufficient image resolution or lack of contrast. All other results were also very similar to the results of the Swedish image test. A further factor limiting the significance of the tests based on the image data is related to the camera installation geometry. The cameras in the Netherlands are mounted at a higher position compared to the German camera systems, resulting in a different viewing angle of the wheelset area. Due to this altered perspective, the model is currently not able to reliably detect all relevant components, which negatively affects the evaluation accuracy.

7.1.10 Requirements for bogies

For the Bogie Analyzer to analyse a wheelset, certain requirements must be met. A reference value is required to determine the widths and heights, for which at least three visible cover screws are necessary. In addition, all evaluable components must be visible, i.e. they must not be covered by other elements. In particular, the brake shoe is partially or completely covered in some types of wheelsets.

Depending on the type of wheelset – with coil springs or leaf or parabolic springs – only the corresponding elements are evaluated. It should be noted that for wheelsets with coil springs, the lifting T and the "consoles" must be visible, and for wheelsets with leaf or parabolic springs, the spring buckle and the spring stop must be visible in order to be able to recognize the distances between them. This can then be used to calculate the load condition and the distribution of the load.

7.1.11 Further recommendations

The Bogie Analyzer provides a scalable foundation for the continued digitalization and automation of freight wagon inspection processes. In the future, the system can be extended to include additional bogie components, enabling a more comprehensive assessment of wagon condition, load state, and load distribution. By continuously integrating new data and operational feedback, the Bogie Analyzer has the potential to support predictive maintenance strategies and data-driven decision-making in freight transport operations.

A major benefit of the system is the support it provides to trackside personnel. By automating routine inspections and visual assessments, employees are relieved from frequent manual work outdoors, improving both efficiency and safety.

Furthermore, the Bogie Analyzer generates large volumes of structured data during operation. These data can be shared with other teams, departments, or stakeholders, enabling collaborative improvements, knowledge transfer, and broader support for maintenance and operational decision-making.

A key focus of further development will be the continued improvement of detection performance, particularly for components and areas that are currently poorly recognized. This includes optimizing training data, refining annotation quality, and adapting model architectures to improve robustness under varying environmental conditions such as lighting, weather, and contamination. Through iterative model updates and validation in real-world scenarios, the reliability and accuracy of the Bogie Analyzer can be systematically enhanced over time.

In summary, the results of the lifting T assessment demonstrate a generally high level of detection accuracy. The majority of cases are classified as correct (approximately 86%). A smaller proportion is identified as incorrect (around 9%), while a limited share of cases (approximately 5%) is rated as uncertain.

The evaluation of the damper ring condition presents a more differentiated distribution. Approximately 56% of the cases are assessed as valid, whereas about 9% are classified as invalid. A significant proportion of roughly 35% cannot be reliably evaluated due to insufficient image quality. This comparatively high share highlights the strong dependency on the quality of the acquired images.

Overall, the results indicate that both assessment functions provide fundamentally robust performance. However, the proportion of uncertain and non-evaluable cases underlines the need for additional data and improved image acquisition conditions to further enhance the reliability and validity of the results. In this context, the model is currently being continuously improved and trained with additional data. Consequently, further enhancements in performance and result quality are expected in the near future as the available data set increases and the model optimization progresses.

7.2 Module 2: Checkpoints for improving sharing of logistics information (Sweden)

The checkpoints of the Swedish infrastructure manager, Trafikverket, are intended to capture logistics-related information concerning rolling stock and intermodal loading units (ILUs), such as containers and semi-trailers. In other words, they are designed to handle the use case “rolling stock and ILU identification”. The main method for capturing this information that has been explored in the project is to perform automatic analysis of the images that are produced by the checkpoints (Figure 7.28). This way, the European vehicle numbers (EVN, also called UIC numbers) of wagons and locomotives can be extracted and the composition of trains can be determined. Likewise, the ILU codes of intermodal loading units can be extracted from the images.

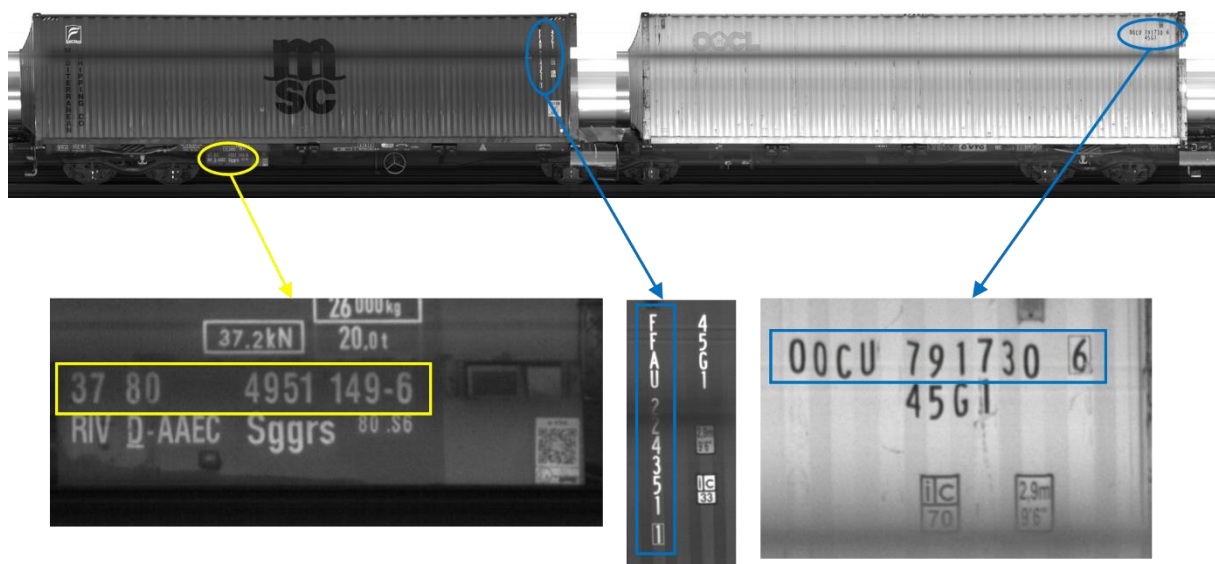


Figure 7.28 A wagon image that has been produced by a checkpoint, along with a vehicle number (left) and two ILU codes (right) that can be extracted from it

7.2.1 Use Case: Rolling stock and ILU identification

The purpose with this use case is that information captured by the checkpoints can be shared with stakeholders in the transport chain, such as other infrastructure managers, railway undertakings, intermodal terminals, railway yards, ports and their customers. There are several possible advantages of sharing the information, for example increased transparency when all authorised organisations in the transport chain have access to the same information about the location of trains, wagons and ILUs. Moreover, improved information sharing enables improved operational efficiency, increased customer satisfaction and productivity. As an example, it is easier to plan the transshipment of ILUs from a train when the composition of the train is confirmed before the train arrives to its destination.

7.2.2 Locations and roll-out plan

A goal of the project has been to have two checkpoints deployed in Sweden. The first checkpoint was built near the port of Gothenburg (the largest port in Scandinavia), already in 2022, during the Shift2Rail project FR8RAIL III (Figure 7.29). The installation includes an Intelligent Video Gate (IVG) system designed and deployed by Hitachi Rail, integrating cameras, illumination units, triggering sensors and processing infrastructure for automated image acquisition. Image analysis is currently performed using systems developed by Hitachi and Indra, outside Trafikverket's IT environment, based on image data captured by the IVG installation. Indra's system delivers the output from the image analysis to the research platform Deplide (see Section 8.4), which has been developed by RISE. The Gothenburg checkpoint represents an operational deployment of the ERC concept in a real logistics environment.

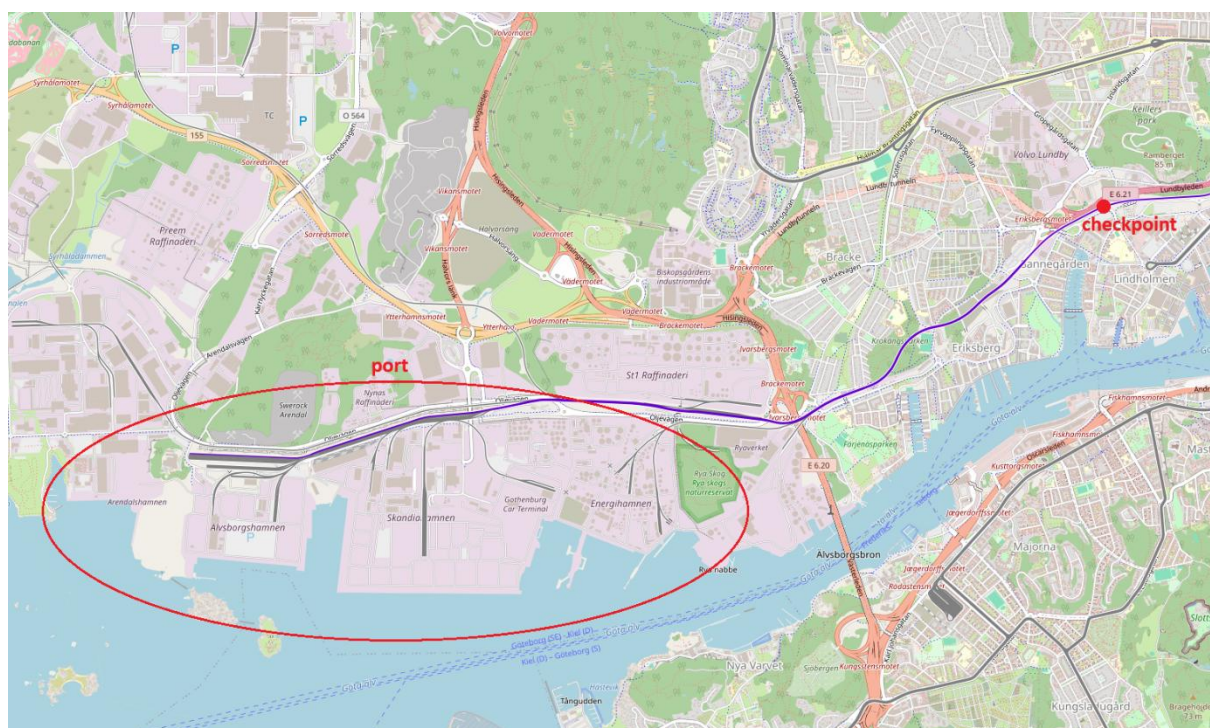


Figure 7.29 The location of the checkpoint in Gothenburg (adapted map from OpenRailwayMap)

The second checkpoint is located in Malmö, near the Öresund bridge and the border to Denmark (Figure 7.30). Cable ducts and poles for the cameras were in place in December 2025. The deployment of the checkpoint including image analysis capabilities was executed in March 2026 (Figure 7.31). Trafikverket was responsible for the physical installations, and the image analysis is performed using Indra's system, within Trafikverket's IT environment. The resulting data will be shared with stakeholders using a system that is already in use at Trafikverket.

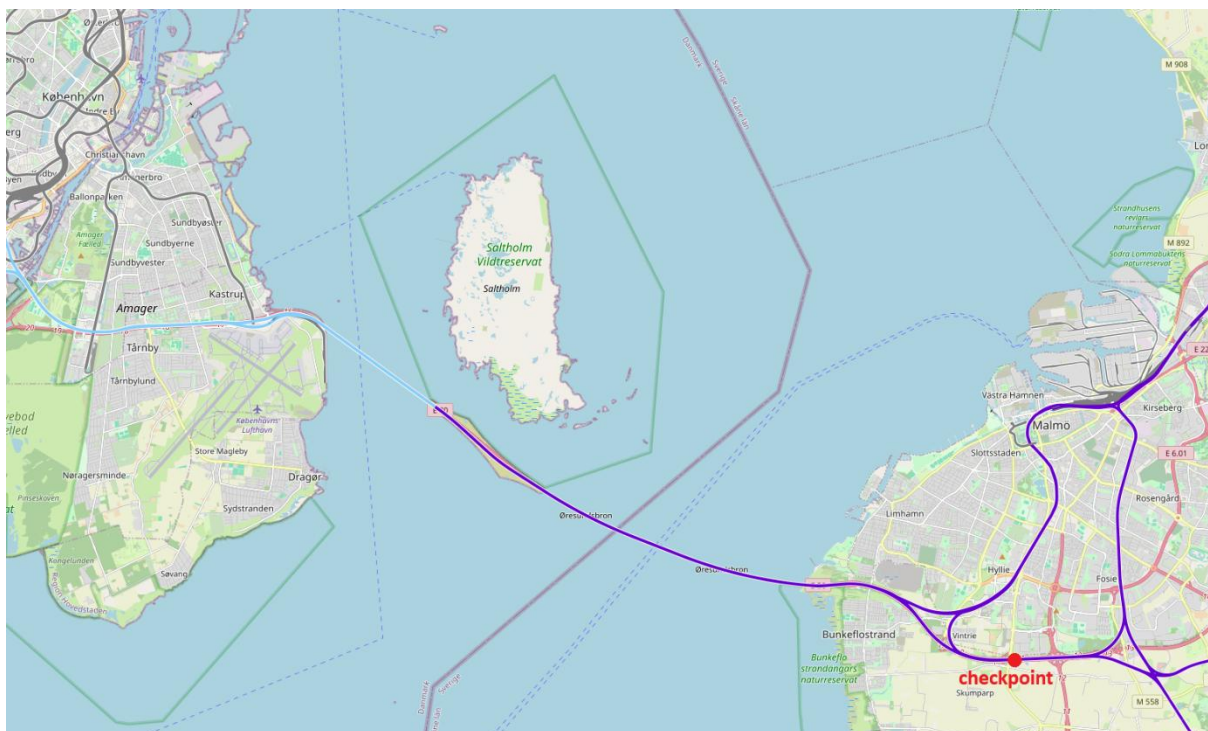


Figure 7.30 The location of the checkpoint in Malmö (adapted map from OpenRailwayMap)

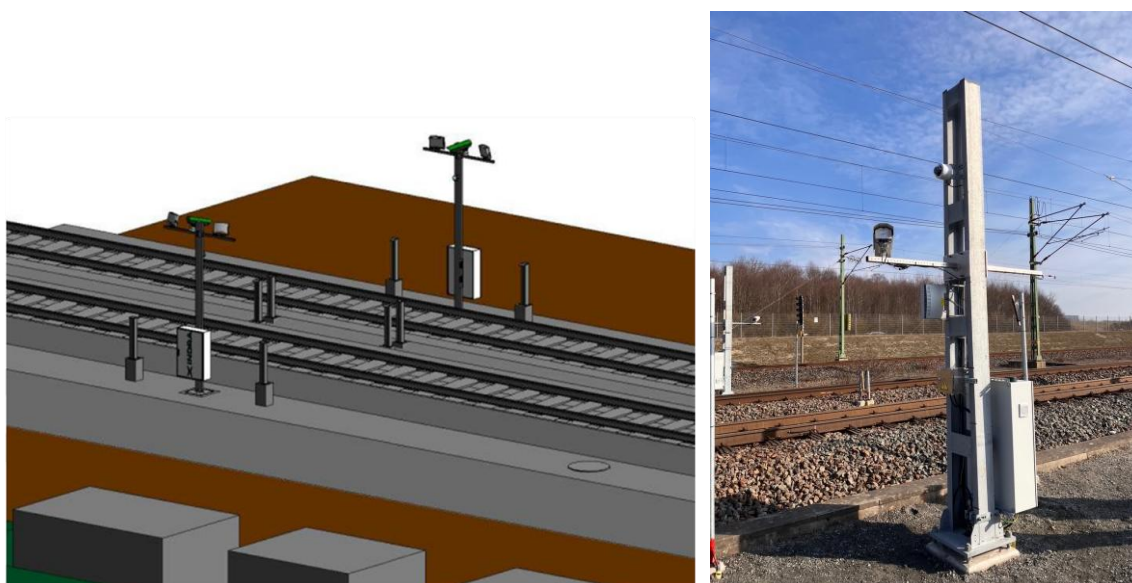


Figure 7.31 A sketch of the deployment in Malmö and the actual installed checkpoint.

7.2.3 Evaluation

During the project, the image analysis functionality of two previously implemented systems (one from Indra and one from Hitachi) has been further developed. The ability of these systems to extract logistics-related information has been evaluated through manual comparisons between images and extracted information. Here, the results of this

evaluation will be summarised. Further details can be found in Kordnejad et al. (2025). The image analysis system from Indra has been evaluated using 190 pairs of images (each pair consisting of one image for each side of a wagon), while the system from Hitachi has been evaluated using 200 pairs of images. The two sets of images are non-overlapping. All images have been produced by the checkpoint in Gothenburg. According to the previously established requirements (Hildebrandt et al., 2024), the output of the systems must be correct in at least 95 % of the cases, so that it can generally be trusted without manual checks.

In the evaluation of the ability of the systems to extract vehicle numbers, a number is considered correctly extracted if (a) all twelve digits in the number are visible in at least one of the two images of the wagon and (b) exactly twelve digits are reported by the system, and all of these are correct. (For Hitachi's system, which reports separate analyses for the two images, it is enough that the latter condition is true for one of the images.) Otherwise, the outcome is considered incorrect. Indra's system produces a correct vehicle number for 64.7 % of the image pairs, while Hitachi's system succeeds in 63.0 % of the cases. The results for the two systems are thus similar, but the reasons behind failures to correctly extract numbers are partly different. For Indra's system, 22.4 % of the failures are caused by damaged or occluded vehicle numbers (which are impossible to read also for humans), while 77.6 % are caused by the system. The corresponding figures for Hitachi's system are 5.4 % and 94.6 %, respectively. Even if the wagons with unreadable numbers are excluded, we must conclude that neither system fulfils the requirement to correctly extract 95 % of the vehicle numbers.

The image analysis systems should also be able to count the number of ILUs visible in each image pair, as well as to report their ILU codes. For this part of evaluation, however, only output from Indra's system was available. The system succeeds in counting the number of ILUs for 103 of the image pairs (54.2 %) included in the test set. The requirement of correctly counting the ILUs in more than 95 % of the cases is thus not fulfilled.

The evaluation of the ability of Indra's image analysis system to correctly extract ILU codes includes the ILUs that are both visible in the images and included in the system's ILU count. In total, the evaluation includes 222 ILUs. The system reports the correct ILU code for 204 of these. This corresponds to 91.9 % of the ILUs, which is close to the requirement of producing correct ILU codes in more than 95 % of the cases but doesn't completely fulfil it. (Note, however, that the system misses many ILUs altogether, due to the counting problems described in the previous paragraph.) All the incorrect ILU codes are caused by the image analysis system, since the code is clearly visible in at least one image in every pair.

7.2.4 Checkpoints for improving monitoring of wheelsets via data quality checks

As mentioned in previous section, the Intelligent Video Gates installed on the Swedish checkpoints are mainly used for sharing information related to logistics. However, the checkpoints related to classical wayside monitoring systems (also managed by the Infrastructure Manager) which are distributed along the Swedish railway network, are first designed to capture safety critical events, such as wheel flats and alarms related to hot-box/hot wheel that indicates imminent failures. However, information from these detectors can be also used to extract information related to non-critical defects. However, sharing information related to these non-critical defects necessitates a harmonization procedure of the different wayside monitoring systems as the outputs from the impact loads can vary based on the detector quality (Kordnejad et al., 2025). It is therefore essential to clarify which factors influence detector readings and to establish appropriate data processing or normalization procedures before sending new condition indicators to operators. This also applies when integrating data with other monitoring technologies, such as Intelligent Video Gates dedicated to wagon monitoring, to either complement or compare automatically extracted indicators with visually identified defects

7.2.4.1 Purpose

The purpose of this case study is to demonstrate a practical and scalable method for harmonizing outputs from different Wayside Monitoring Systems (WMS), with a specific focus on Wheel Impact Load Detectors (WILD). Although WMS are essential for railway safety and predictive maintenance, measurements collected from different detector sites often show systematic discrepancies. These differences arise from variations in sensor technologies, installation conditions, calibration procedures, signal processing algorithms, and local track characteristics. As a result, identical vehicles may generate significantly different readings depending on where they are measured, which complicates data interpretation, prevents the use of common alarm thresholds, and reduces the reliability of predictive models (increase of uncertainty in Remaining Useful Life prediction).

To address this issue, the study introduces cumulative distribution function (CDF) rescaling based on quantile mapping. Instead of correcting each detector through hardware adjustments, the method statistically aligns the empirical distributions of multiple detectors to a selected reference detector. By mapping the quantiles of one detector onto the quantiles of another, the approach preserves the ranking of measurements while reducing systematic bias. The objective is to improve comparability across sites, enhance early fault detection, and enable lower and more reliable lower alarm thresholds across the network.

7.2.4.2 Case study

The methodology is simply demonstrated using operational data from Wheel Impact Load Detectors installed along the Iron ore line (Malmbanan) freight corridor in northern Sweden. The case study focuses on a specific freight wagon running across multiple sections of the line and passing six different detectors that rely on different sensing technologies, including strain gauges and fibre optics depending on the manufacturer.

Initial analysis revealed clear discrepancies between detectors. Some showed multimodal distributions, higher spreads, or unstable median values over time, indicating potential calibration issues or local track influences. Others displayed stable and consistent statistical behaviour. Based on distribution stability and noise characteristics, one detector was selected as the reference, while detectors showing too many structural inconsistencies were excluded from the rescaling process.

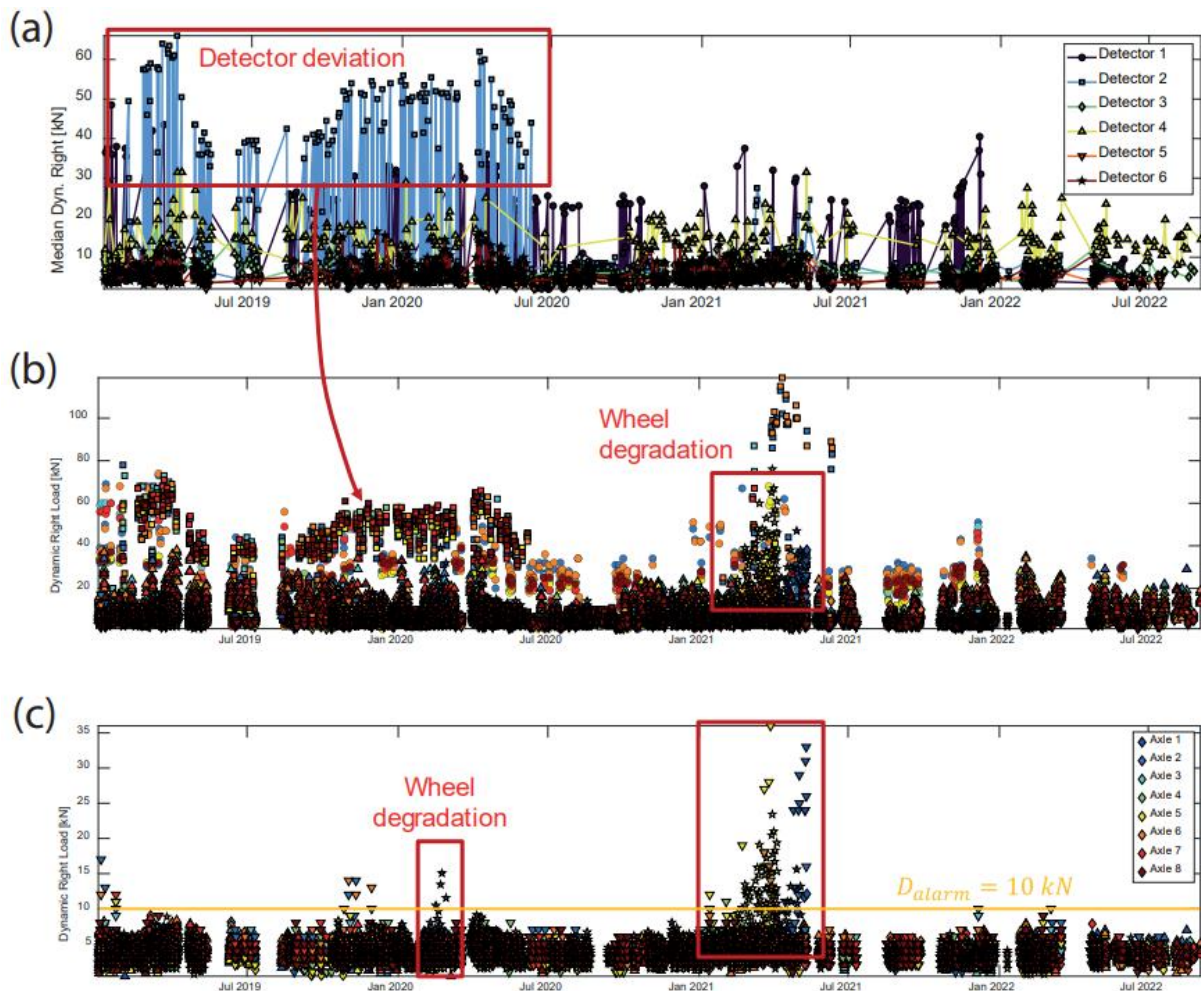


Figure 7.32 (a) Evolution of the median value of the detectors right side as function of time (b) Evolution of the degradation of the right wheels before rescaling/removal (c) Evolution of the degradation of the right wheels after scaling/removal

After applying CDF rescaling to the remaining detectors, the variability between sites was reduced (Figure 7.32). Degradation patterns that were partially masked in the raw data became more visible, including early-stage wheel degradation events. The harmonized data allowed the use of lower alarm thresholds with fewer false alarms, demonstrating the benefit of statistical normalization for predictive maintenance.

The demonstrator also highlights the importance of continuous detector quality assessment and the need for standardized metadata to ensure transparency in data processing. Overall, the case study confirms that CDF rescaling can serve as an effective network-level harmonization strategy for WMS, improving comparability, and early fault detection without requiring major hardware modifications.

7.2.4.3 Evaluation of detector quality

The previous example showed that it is possible to improve the detection of smaller defects by simply rescaling the CDF statically with a reference. However, it should be mentioned that even the reference detector can show variation as function of time (seasonal variation), wagon types, load and speed, meaning that the rescaling should be performed dynamically (i.e. per passage, day or week) to improve the detection of minor defects.

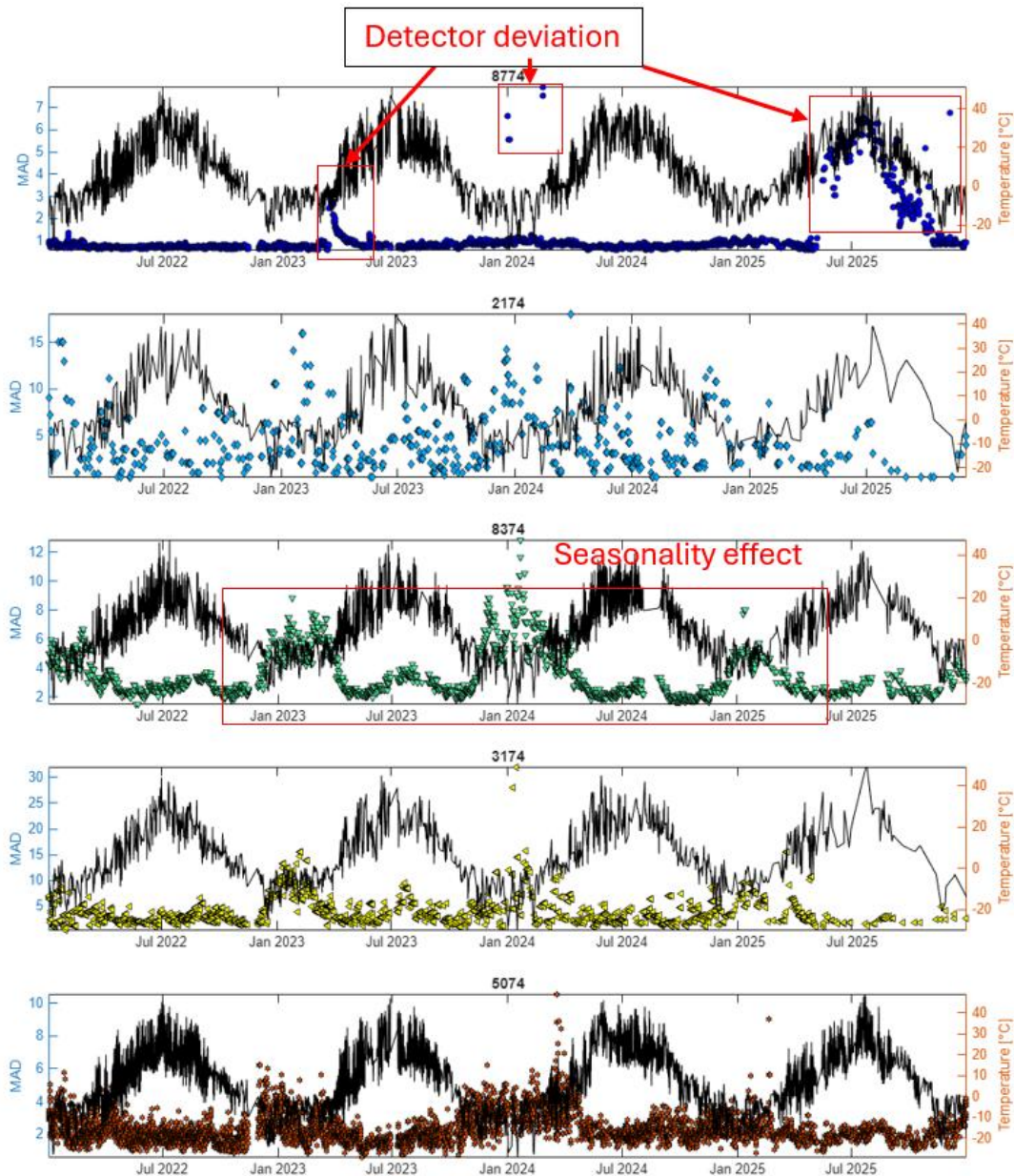


Figure 7.33 Temporal evolution of Mean Absolute Deviation of the reference detector by wagon type over the period 2022-2025

As seen in Figure 7.33, an example of the moving mean of the time series of Mean Absolute Deviation (MAD) values, separated by wagon types, exhibit clear differences in both baseline level and variability. Some wagon types (e.g., 8774, which is the example used in the previous rescaling process) show relatively low and stable MAD values for most of the observation period, whereas others (e.g., 2174, 8374 and 3174) display substantially higher dispersion and more frequent extreme values. In addition, a recurring seasonal pattern is visible in several wagon types, particularly 8374, 3174 and 5074, with elevated MAD values during winter periods across multiple years. This suggests a consistent environmental or operational influence, probably temperature related.

In addition to seasonal effects, irregular increases are observed. For instance, early 2023 shows elevated MAD levels across multiple wagon types, indicating a possible system-wide environmental or operational effect during that period. Furthermore, wagon type 8774 exhibits a sustained increase in MAD during mid-to-late 2025, distinct from previous seasonal fluctuations, which may indicate a structural or operational change specific to that wagon type or fleet. It should be noted that this behaviour only occurred for loaded wagons and correlates also with outside temperatures.

Overall, the data demonstrates:

- clear differences in typical MAD levels for various wagon fleet
- recurring seasonal variation for specific wagon types
- occasional regime-like shifts, supporting the need to explicitly account for wagon type, load, time, and environmental factors in subsequent statistical modelling.

7.2.4.4 Conclusion

The observed variability in MAD values across wagon types and over time demonstrates that raw WILD measurements alone are insufficient for reliable assessment of detector quality. Clear seasonal patterns indicate a strong environmental influence, likely temperature-related, while substantial differences in baseline level and dispersion between wagon types show that fleet composition significantly affects measured impacts. Additionally, episodic shifts suggest that operational changes or maintenance events may alter measurement behaviour, but only to specific wagon fleet or specific load ranges. Consequently, meaningful evaluation of detector performance requires that WILD data be accompanied by contextual metadata, including temperature, wagon type, speed, load, and precise timestamps (and ideally maintenance or calibration records). Without such information, variations arising from environmental or operational factors may be incorrectly interpreted as detector degradation or instability. To ensure that the operator can trust and take decisions based solely on the data, quality metrics related to the detectors should be shared during each train passage to know if rescaling can be applied or not, or to internally perform a rescaling process before sending the new processed data. This new data can be then shared to other systems such as IVG and compared with information related to wheelset condition extracted from both systems.

7.3 Module 3: Checkpoints deployment in Can Tunis for capturing yard movements (Spain)

With the goal of supporting the development of European Railway Checkpoints within the framework of FP5-TRANS4MR WP29 project, the Spanish demonstrator has been developed to accurately capture and digitalise yard movements.

This solution operates as an advanced railway monitoring system designed to collect, process, and analyse essential data concerning train operations and their composition. By delivering precise insights into yard activity, it enhances operational awareness, safety, and decision-making within railway environments.

7.3.1 Introduction to the Spanish demonstrator

The Spanish demonstrator is being deployed and validated in Can Tunis, a key freight terminal in Barcelona. The site is strategically connected to the national rail network (Red Ferroviaria de Interés General, RFIG) and directly linked to both the Port of Barcelona and the nearby Morrot terminal. This location provides an ideal real-world environment in which to test and refine the system, forming the basis for future enhancements (Figure 7.34).

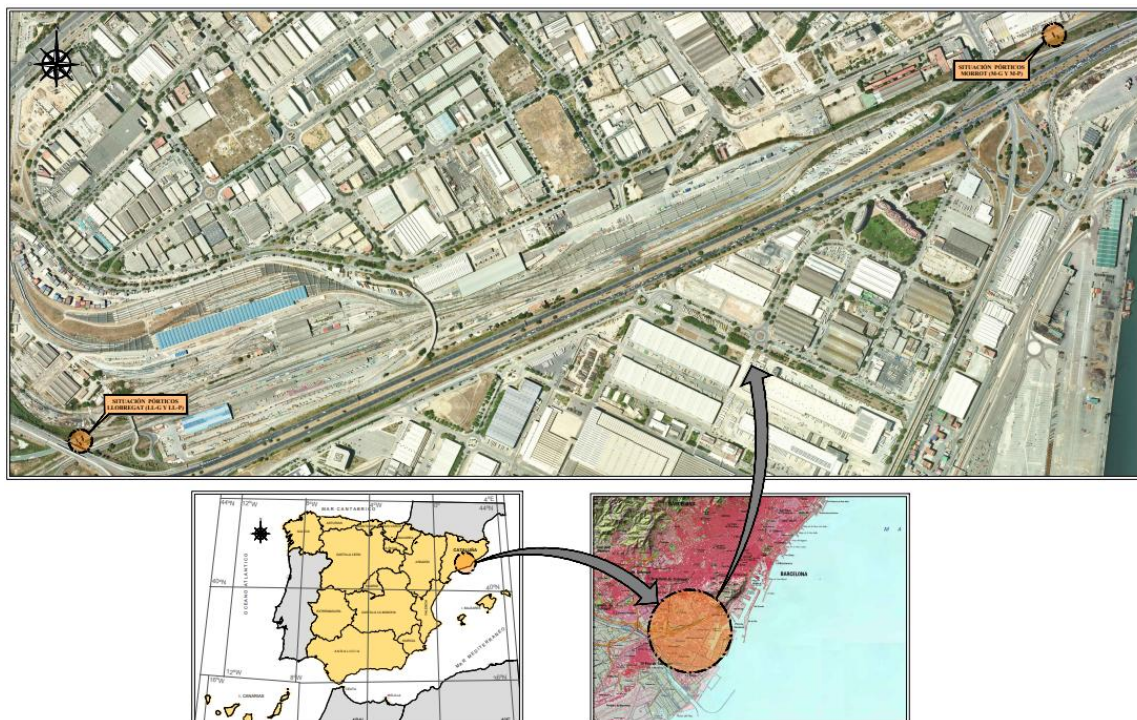


Figure 7.34 Geographical location.

The Spanish solution is structured around three coordinated components, as can be seen in Figure 7.35:

- **European Railway Checkpoints:** These systems consist of video gates equipped with sensors, high-resolution cameras, and AI-based processing modules that enable the automatic identification of trains and wagons, the detection of dangerous goods, and the acquisition of key operational data. ERCs are strategically installed at the entrance and exit of the Can Tunis terminal, ensuring comprehensive monitoring of all rolling stock circulating within the facility.
- **On-board sensors:** This technology provides continuous information about the operations and events occurring on the trains where it is installed. It comprises a set of sensors placed at strategic locations on the wagon to collect relevant data relating to logistics, transported cargo, and train composition.

Integrating on-board sensors offers a more comprehensive view of the condition and behavior of rolling stock, enhancing traceability, optimizing logistics processes, and complementing the analysis of data captured by the video gates. By bridging the information gap between static checkpoints, the system ensures that the status of the assets is monitored not just at specific nodes, but continuously throughout the operational flow.

- **Conceptual Data Model (CDM):** To enable seamless and standardised data exchange between the European Railway Checkpoint system, on-board sensors and other railway information systems, data sharing must follow a structured, interoperable, and secure approach.

For further details on system requirements, technical specifications, and deployment aspects of the Spanish demonstrator, please refer to deliverables D29.1 (Kordnejad et al., 2024) and D29.3 (Pelegri n Preixens et al., 2025).

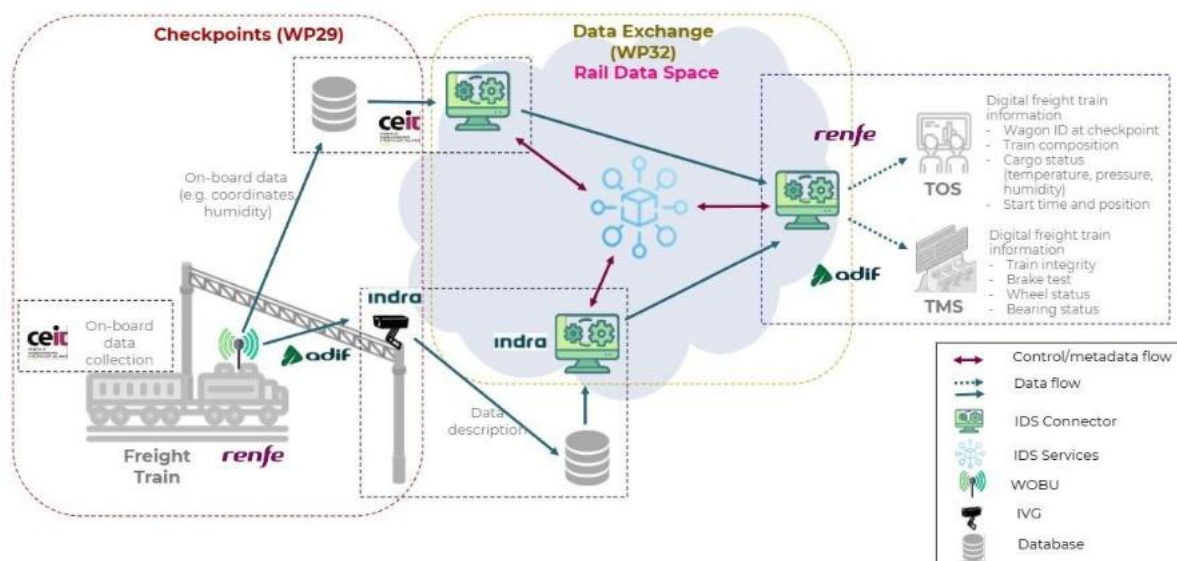


Figure 7.35 Architecture of the Spanish Demonstrator.

The installation of the ERC in Can Tunis is currently in progress and is expected to be finalized and tested in the near term. Meanwhile, the on-board sensor system has already been fully implemented and successfully validated.

For a comprehensive overview of the installation progress and testing activities for both the Railway Checkpoints and the on-board sensors, please consult deliverable D29.6.

7.3.2 Checkpoints deployment in Can Tunis

The Railway Checkpoints will be installed in two different locations at Can Tunis: to the west of the terminal, known as the *Llobregat* location, and to the east of the terminal, known as the *Morrot* Location. Both locations are situated at the ends of the terminal, as shown in Figure 7.36. In this way, all rolling stock circulating through the terminal can be identified recording all entering and exiting times.



Figure 7.36 Location of the Llobregat and Morrot sites inside Can Tunis.

The European Railway Checkpoints developed by Indra are being deployed in Can Tunis as intelligent trackside gantries forming part of the Intelligent Video Gate (IVG) system. These checkpoints are designed to automatically detect trains approaching and passing through the interface between the Spanish national rail network (RFIG) and the railway infrastructure of the Port of Barcelona. Each checkpoint integrates triggering sensors and high-resolution image capture systems capable of recording the passage of locomotives and wagons, generating precise timestamps and digital records of the train composition.

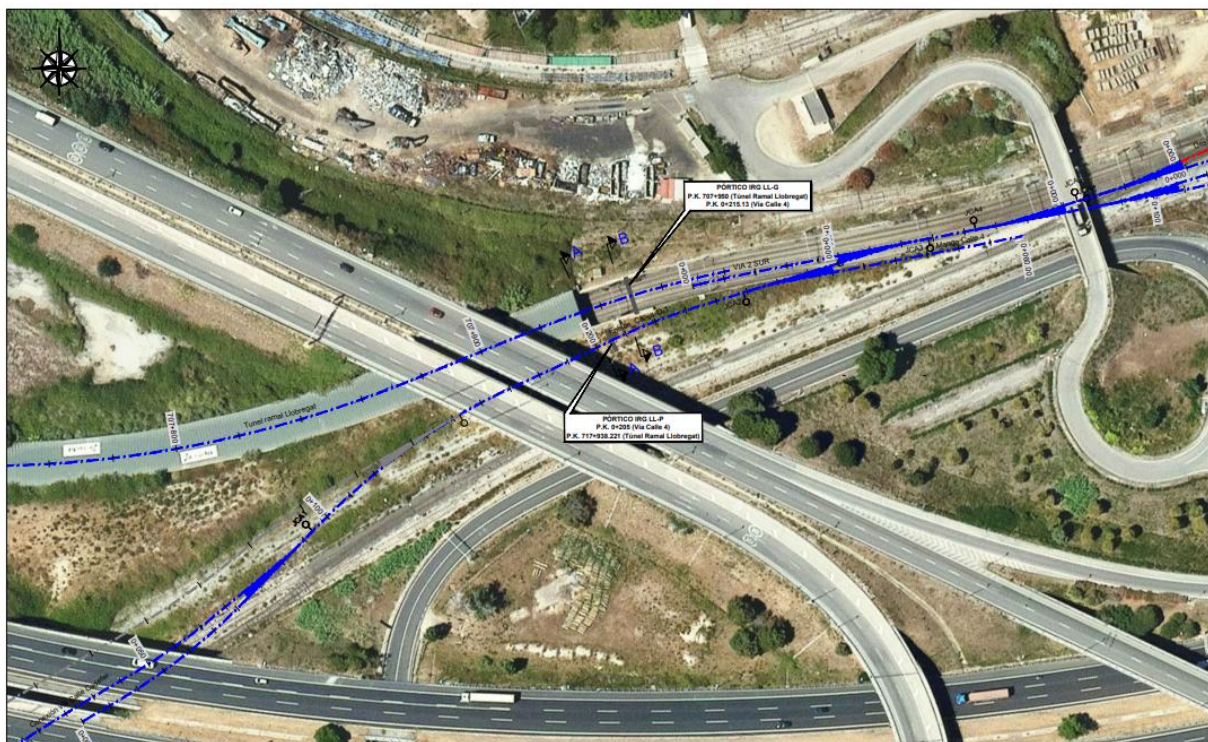


Figure 7.37 Detail: Llobregat site.

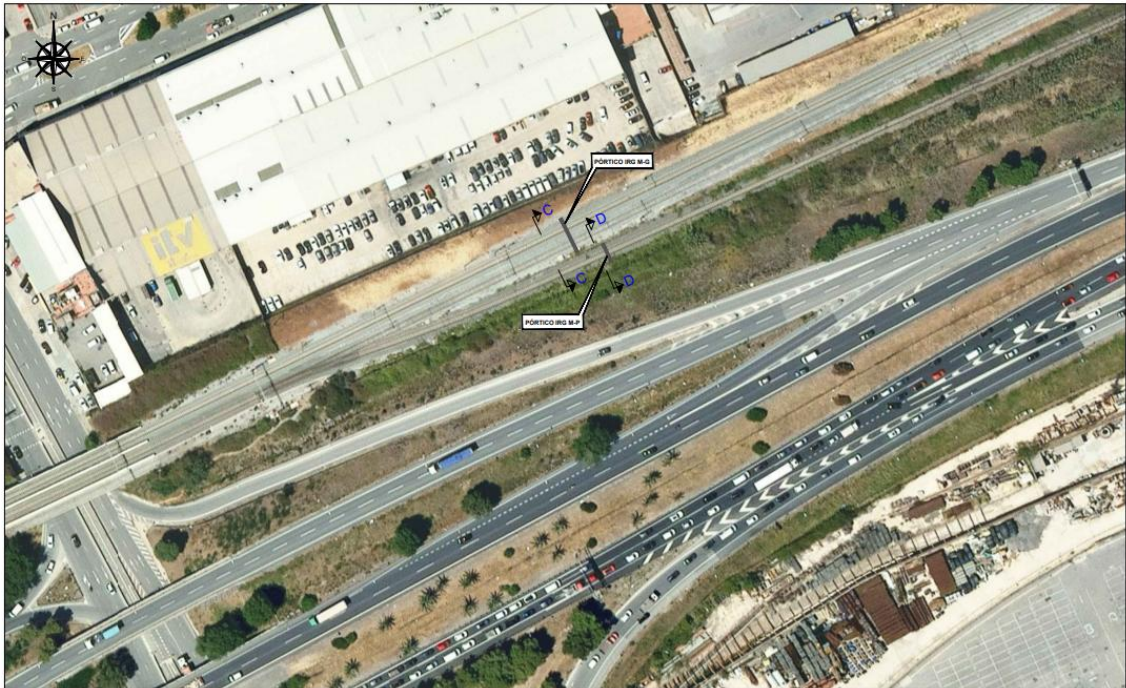


Figure 7.38 Detail: Morrot site.

Currently, information on the composition of freight trains is collected manually by rail and logistics operators. This includes recording containers, wagon numbers and signals, types of goods, gauge validation, and border inspection, as well as detecting damage, discrepancies in cargo, and signalling errors. The automation of these processes through ERC seeks to digitize and optimize these tasks, responding to the need to transform rail transport into a key element of the logistics chain by digitizing and improving the flow of information.

The monitored railway section is being equipped with four intelligent gantries installed at the interface between the RFIG network and the railway infrastructure of the Port of Barcelona. These gantries are distributed across the access points that connect the national railway network with the port rail infrastructure, enabling the monitoring of both incoming and outgoing freight trains.

The gantry is equipped with a set of detection cameras and sensors that continuously monitor the track area in order to identify approaching trains. When a train is detected approaching the checkpoint, the system automatically activates the image acquisition sequence. A preliminary classification stage is performed to confirm the presence of a train and trigger the appropriate capture workflow. Once the passage is confirmed, the system captures a sequence of high-resolution images that are subsequently processed by computer vision algorithms.

Through this processing stage, the system extracts key operational information from the images, including the identification of the train, the recognition of freight identification plates and pictograms, the detection and counting of wagons and the determination of the beginning and end of the train composition. This automated analysis enables the generation of a structured digital record of the convoy as it passes through the checkpoint. The information obtained also makes it possible to monitor the dwell time of freight convoys on the tracks owned by the Port of Barcelona, supporting a more efficient management of railway and logistics operations within the terminal.

In addition to image capture, the system incorporates speed estimation capabilities. As the train approaches the gantry, the sensors allow the estimation of its speed in real time. This information is used to dynamically adjust the internal configuration of the cameras, regulating parameters such as the capture frequency (frames per second, FPS), exposure time and triggering intervals. In this way, the system automatically adapts image acquisition to the real operating conditions, ensuring clear identification of wagons even at higher speeds while avoiding unnecessary data generation when trains move at low speeds. This functionality optimizes both detection accuracy and the efficiency of data processing and storage.

From an operational perspective, the ERC solution enables comprehensive supervision of rolling stock movements within the marshalling yard of the railway logistics node. The system automatically identifies trains entering and leaving the terminal, providing objective control of freight flows circulating between the Spanish national rail network (RFIG) and the railway infrastructure of the Port of Barcelona.

Furthermore, the data and images captured by the gantries enable advanced monitoring of railway operations and safety:

- Freight monitoring: detection of potential defects in the transported goods, verification of parameters such as temperature, volume and size of the cargo, ensuring that loading gauge limits are respected and preventing operational risks.
- Automatic train and cargo identification: through the captured images and subsequent image processing, the system is able to obtain the train identifier, identify the freight wagons and recognise cargo-related information, including the pictograms and identification plates associated with the transported goods.
- Train integrity monitoring: supervision of number of wagons of a train, possible mechanical failures in wagons, and general monitoring of the condition of the rolling stock.

The automated recording of train entries and exits makes it possible to monitor the dwell times of trains and goods on the terminal tracks, which is essential for economic and

operational management. The information generated will allow the railway operator to objectively allocate infrastructure usage times and the corresponding tariffs related to the presence of freight on the tracks, providing traceability, transparency and reliability in terminal management.

Following the definition of the installation points at Llobregat and Morrot, the deployment of the ERC infrastructure is currently in progress. INDRA leads the civil Works, engineering, development, implementation, deployment and validation activities of the IVG system, ensuring its adaptation to the specific geometric, electrical and operational conditions of the Can Tunis terminal. The deployment represents a large-scale railway infrastructure project that involves significant engineering, integration and installation activities. Due to the operational nature of the terminal and the coexistence with ongoing civil works, the implementation is being carried out progressively. Although the system has not yet reached full deployment, the procurement of equipment, preparation of infrastructure and installation works are already in progress.

Figure 7.39 Photograph of the topographic study, shows an example photograph taken by the team responsible for the topography:

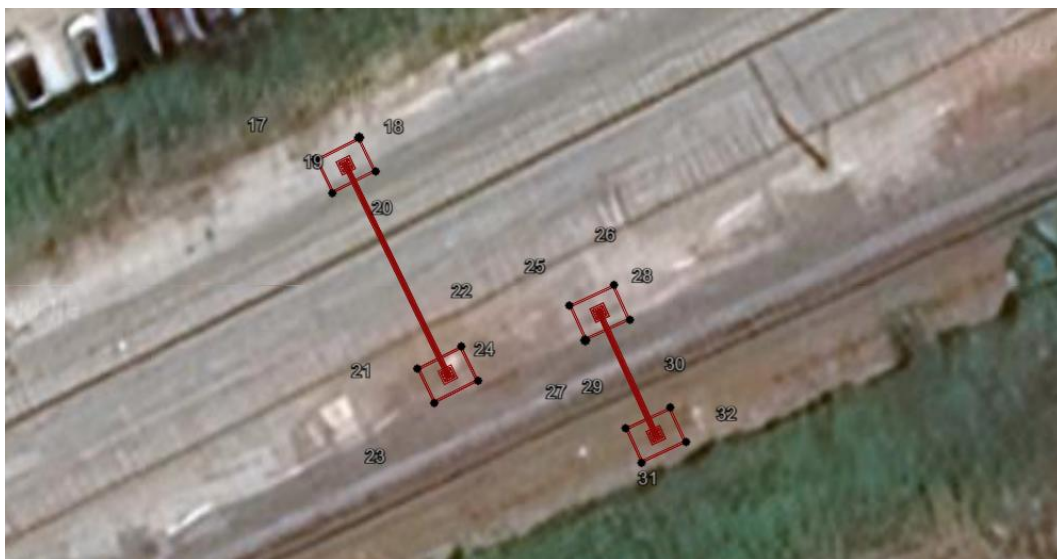


Figure 7.39 Photograph of the topographic study.

The following figures (Figure 7.40, Figure 7.41, Figure 7.42 and Figure 7.43) show the exact planned location of the checkpoints in each of the corresponding track sections.



Figure 7.40 Checkpoint for double-track railway installation at Morrot location.

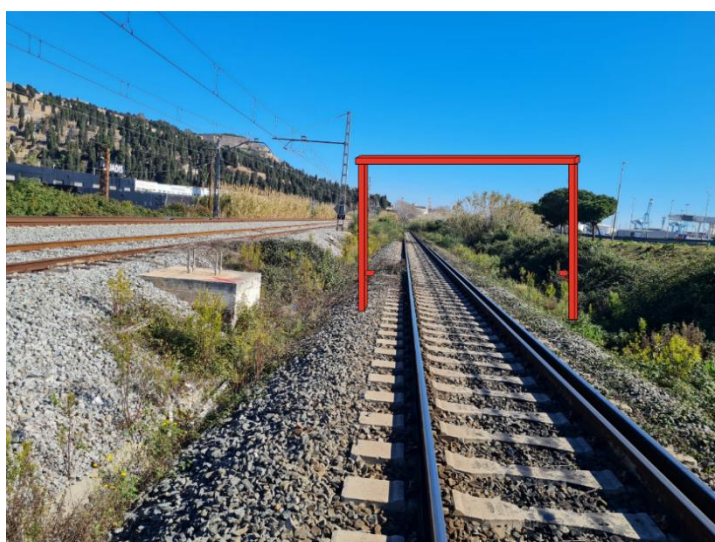


Figure 7.41 Checkpoint for single-track railway installation at Morrot location.



Figure 7.42 Checkpoint for double track railway installation at Llobregat location.



Figure 7.43 Checkpoint for single-track railway installation at Llobregat location.

The project represents a technically complex railway infrastructure deployment, requiring phased implementation due to the operational nature of the terminal, the coexistence with ongoing civil works, and the regulatory requirements typical of the railway environment. For this reason, the full deployment of the system is planned as a medium-term objective, progressing gradually and in coordination with the infrastructure managers.

The following two figures shows the detailed design of the checkpoints, including the structural layout and the main installation elements. Figure 7.44, shows the cross-section of the checkpoint located in Llobregat for the RFIG tracks, while Figure 7.45, shows the corresponding section for the Morrot area.

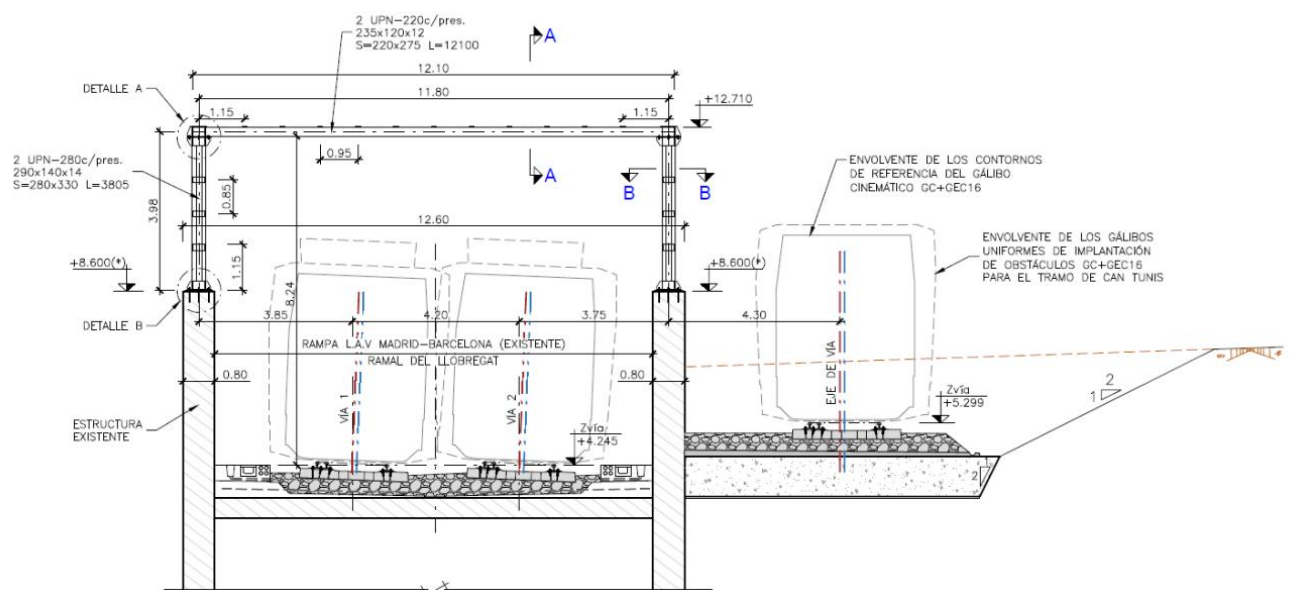


Figure 7.44 Checkpoint Llobregat section – RFIG.

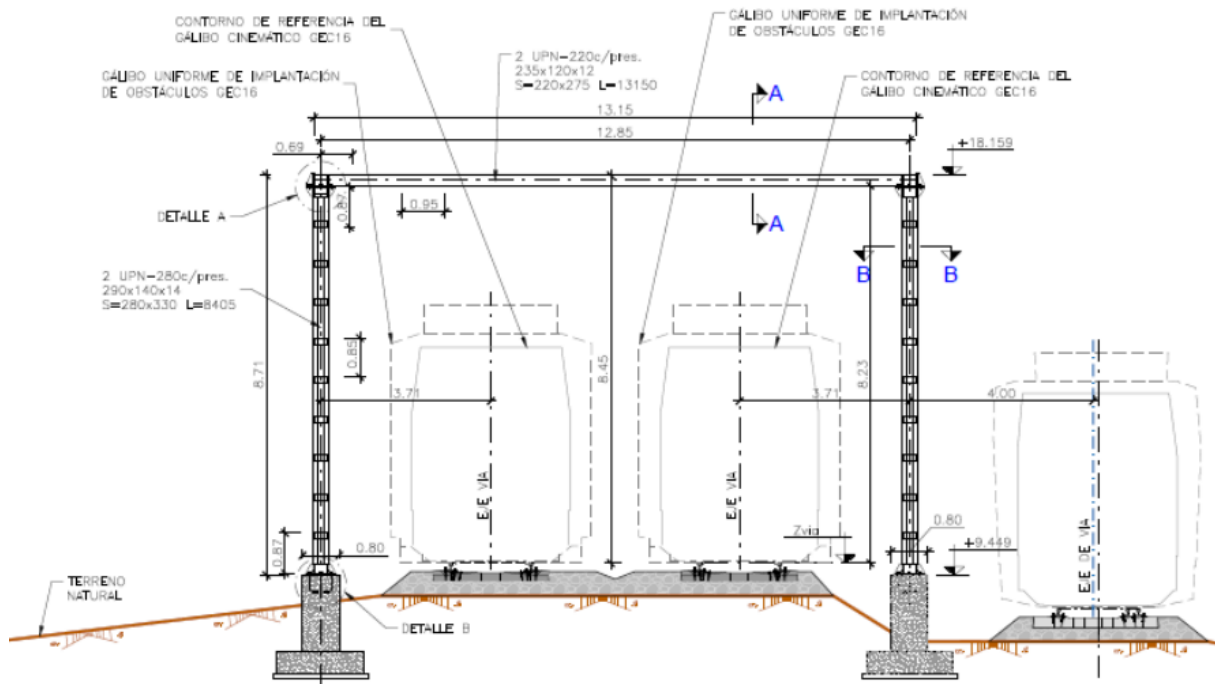


Figure 7.45 Checkpoint Morrot section – RFIG.

Analogously to the two previous figures (Figure 7.44 and Figure 7.45), two additional figures (Figure 7.46 and Figure 7.47) are presented, corresponding to the checkpoints located on the railway tracks of the Port of Barcelona.

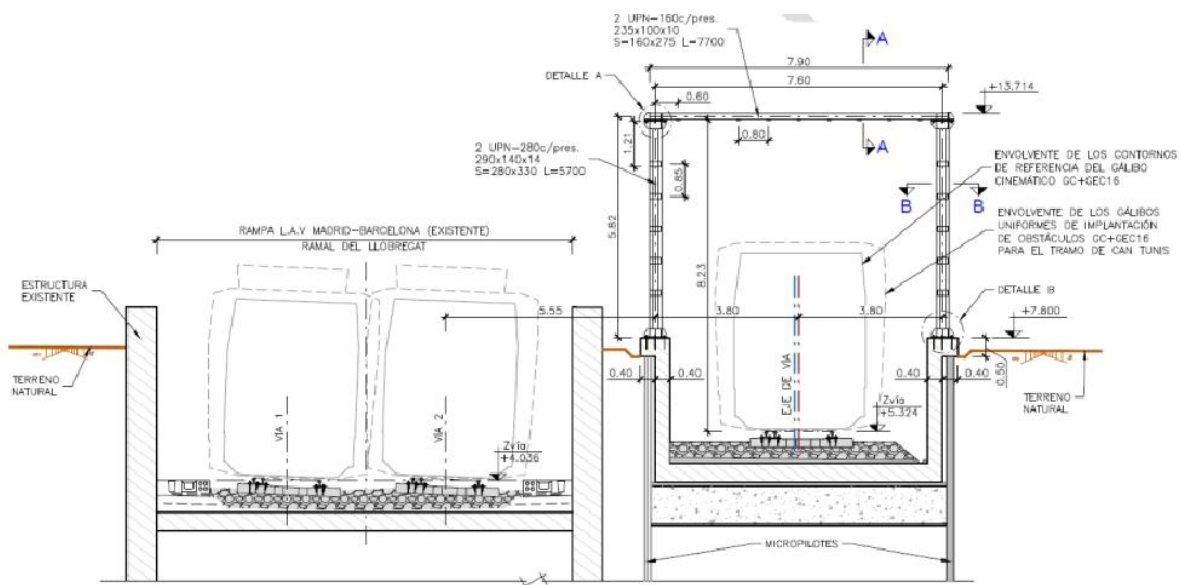


Figure 7.46 Checkpoint Llobregat section – Port

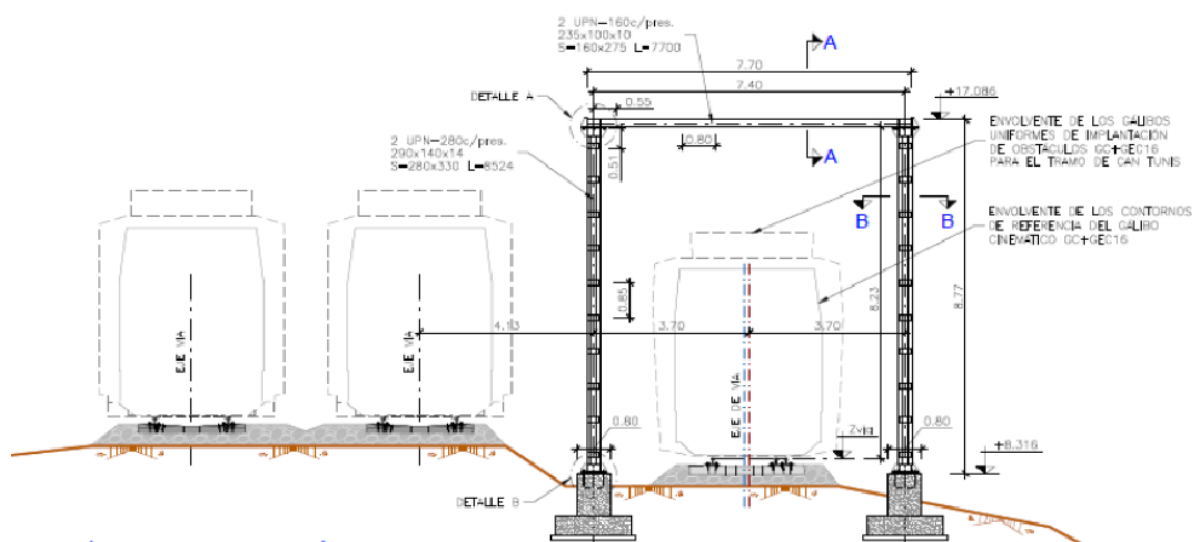


Figure 7.47 Checkpoint Morrot section – Port.

INDRA is responsible for the complete lifecycle of the wayside system, including the definition of the architecture, detailed engineering, structural and mechanical design of the gantries, as well as the integration of camera systems, detection sensors, lighting modules and edge computing processing units.

Additionally, INDRA has developed and validated in a controlled environment the artificial intelligence-based detection and identification algorithms that will enable the automatic recognition of rolling stock. The system has been prepared for interoperability with the Conceptual Data Model (CDM) and the Federated Data Space (FDS), ensuring its future integration with onboard systems and other railway management platforms. Before the start of trackside installation, a laboratory validation phase was carried out including factory acceptance tests (FAT), verification of the algorithms through simulations and recordings of train passages, time synchronization tests, communication robustness testing and cybersecurity validations. This phase helped reduce risks and ensured a high level of technological maturity before deployment in a real operational environment. Currently, the project has entered the physical installation phase. The engineering design was updated to adapt to modifications derived from ongoing civil works, which affect track geometry and electrification systems. Technical inspections have been completed including the structural validation of the installation locations, the manufacturing and pre-configuration of hardware components, and the coordinated planning of field activities. The installation of the gantries and the preparation of power supply and communication infrastructures are currently being initiated. As part of the civil engineering project, continuous topographical surveys have been carried out on site to precisely define the location of the gates and plan the installation work.

Within this deployment phase, a wide range of engineering, civil works, installation, and systems integration activities are being carried out. These include civil engineering works such as earthworks and the construction of foundations for the portal structures, the preparation and storage of portal components, and the installation of the structures themselves. Other infrastructure activities include the deployment of communication and power connections, the installation of direct current (DC) power supply units, and the integration of communication equipment such as 5G routers, input/output modules, and PoE switches. The track-side sensor infrastructure includes the installation of train detection cameras, high-resolution analysis cameras, and photoelectric sensors for train activation and monitoring. At the same time, systems engineering and integration tasks are being carried out under the leadership of INDRA, including detailed engineering, calibration of cameras and sensor equipment, on-site installation and validation of the system, development and implementation of the Intelligent Video Gate (IVG) recognition solution, and deployment of the cloud support infrastructure necessary for data processing and system operation.

Figure 7.48 and Figure 7.49, shows respectively, in a schematic manner, the train detection area and the image acquisition area, from which images will be captured for data processing.

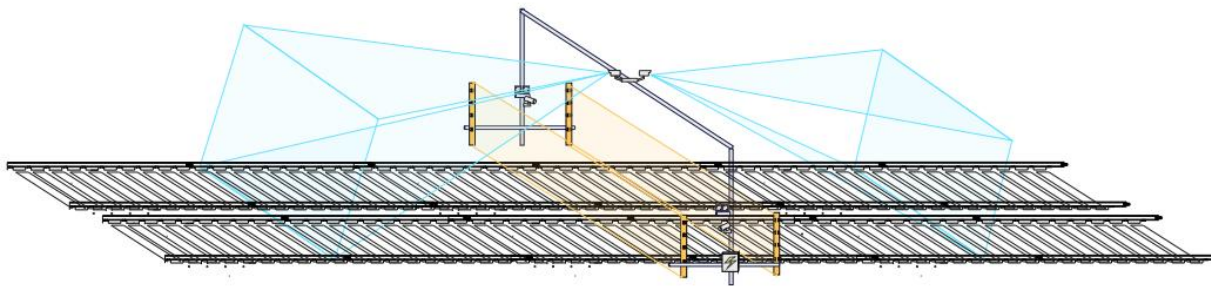


Figure 7.48 Detection area.

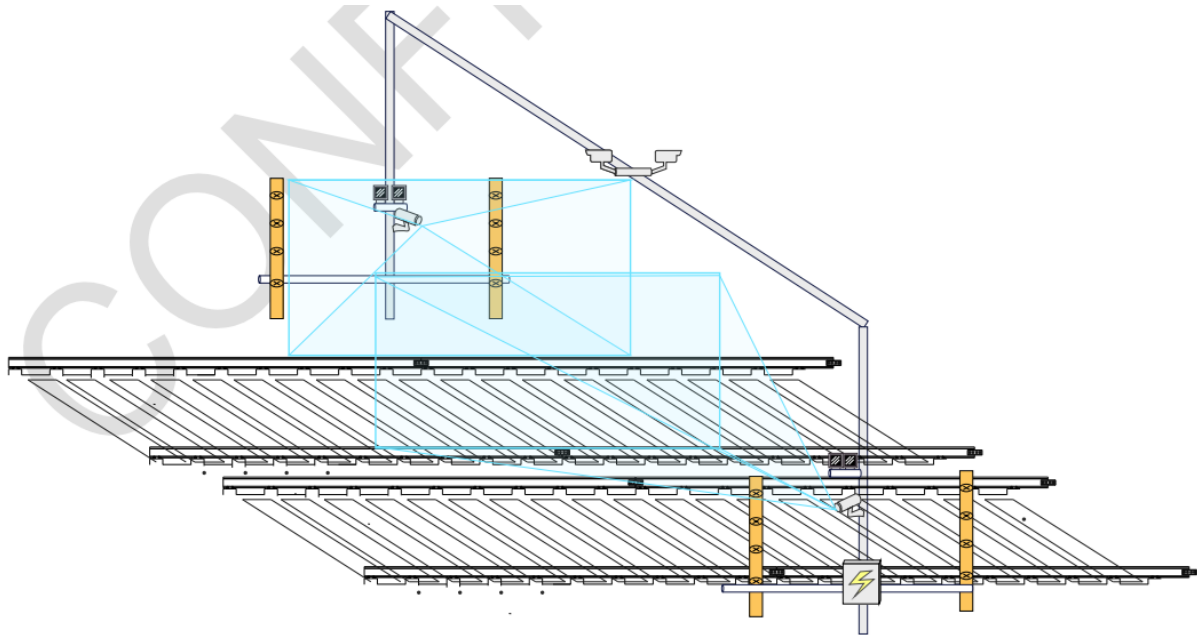


Figure 7.49 Acquisition area.

In the following phases, the project will address the full deployment of the installation, including the completion of mechanical assembly, electrical and communication integration, and the progressive commissioning of the system under real operating conditions.

Once the physical installation is completed, field tests will be conducted under real train circulation conditions. These tests will allow validation of the detection performance, the accuracy of speed estimation, the quality of captured images and the correct generation of timestamped operational events. During this stage, adjustments and iterations will be carried out to address potential issues identified, with the aim of optimizing system performance and ensuring stable and reliable operation. Subsequently, an integrated validation phase with the onboard systems and the CDM/FDS ecosystem will be performed, ensuring the consistency and synchronization of the generated data. After successfully completing all technical and functional tests, the system will undergo final approval and enter nominal operational use. The system will be used by the freight railway operator within the logistics node as a tool for access control, freight supervision, safety monitoring, defect detection and management of track occupancy times, enabling an objective and fully digitalized management of the railway infrastructure.

Validation in laboratory

The different modules of the checkpoints have been validated in a laboratory environment to verify the different functionalities and interaction under controlled conditions and ensure the proper operation of the checkpoints installed in Can Tunis, to fulfil the project objectives:

1. **Train Detection Module** This module, in charge of detecting the presence of a train on a specific track, running at a specific speed and in a specific direction. Making use of software designed and developed by INDRA, images of the rolling stock circulating through the operational control points installed in Can Tunis are obtained and processed so this rolling stock can be 'geolocated'.

Testing has been performed in different environments: unitary, modular and integration test environments.

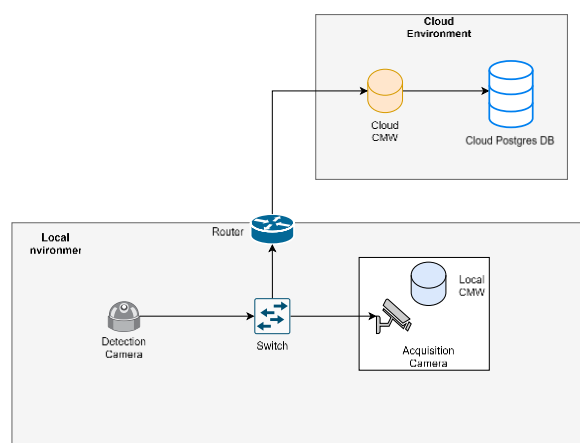


Figure 7.50 Integration test Train Detection Module environment.



Figure 7.51 Integration test Train Detection Module detection camera.

Figure 7.50 and Figure 7.51, shows the laboratory environment and setup used during the testing phase. In this configuration the camera is positioned facing a screen displaying video footages of trains (previously recorded on a relevant railway line), allowing to simulate the presence and movement of a real train in a controlled environment. This approach enabled the validation of the detection module under conditions that approximate the real Can Tunis scenario, while maintaining the flexibility and repeatability of laboratory testing.

Once images are captured and analysed, tracking is applied to the detections to avoid outliers and frames without detections. The following images (Figure 7.52) show an example of the tracking process applied to a detected train in movement.



Figure 7.52 Detection and Tracking example of a train in movement.

Once the train is identified by the system, tracking is applied to each bounding box detected in the frame.

The speed of the train is estimated from the displacement observed between consecutive frames, which initially provides a measurement expressed in pixels per second (px/s). However, due to the perspective of the camera with respect to the track, this value cannot be directly interpreted in real-world units. To address this, a homography transformation is applied to transform pixels to meters in order to map the image plane to the ground plane of the track. An example of this process is illustrated by Figure 7.53.



Figure 7.53 Homography transformation example.

This transformation allows both the correction of the perspective distortion and the conversion of the pixel-based speed measurements into real-world units.

Once this transformation is performed, bounding box tracking is applied to estimate the train's speed. Figure 7.54 illustrates the bounding box placed around the train in the transformed view, which allows to estimate the speed using the velocity of the bottom right corner of the bounding box (v_{x2} , v_{y2}), which has been found to provide more stable measurements than the top side of the box.

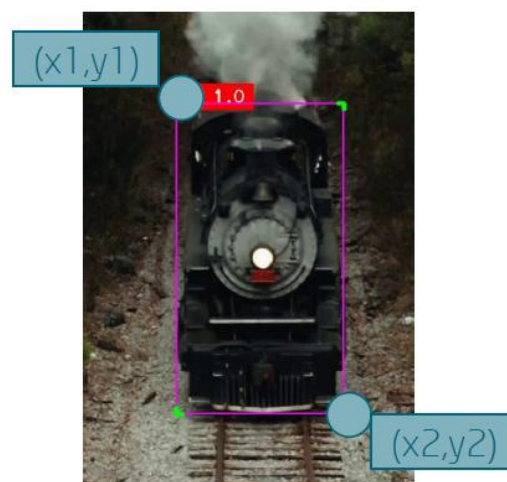


Figure 7.54 Bounding box in transformed image for speed estimation.

An example of the keypoint tracking and prediction results is shown in Figure 7.55. Successful detections are indicated by green points, while missed detections are shown as red points.

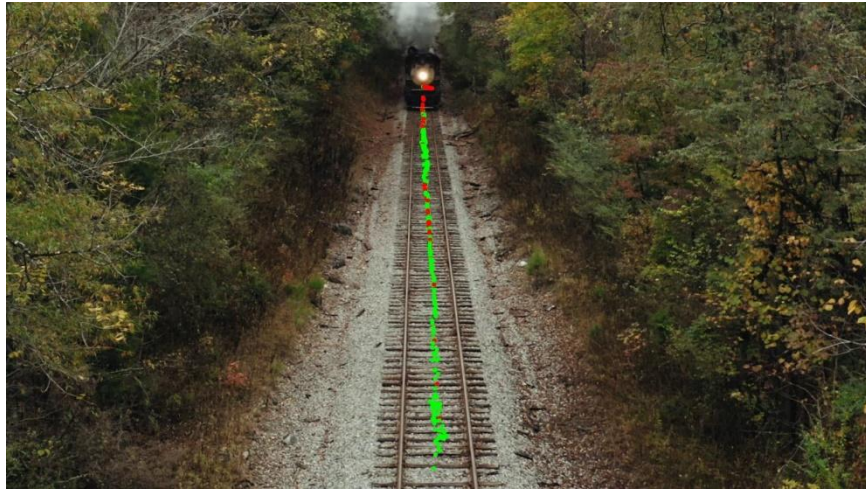


Figure 7.55 Keypoint tracking and prediction

Unitary Tests

Unitary tests were performed to verify the correct behaviour of individual functions and to ensure their stability against future changes. These tests were automatically executed as part of the CI/CD pipeline and their result were monitored through Sonar metrics, as it can be seen in Figure 7.56. The Unitary Tests have passed with a coverage of 80.4%.

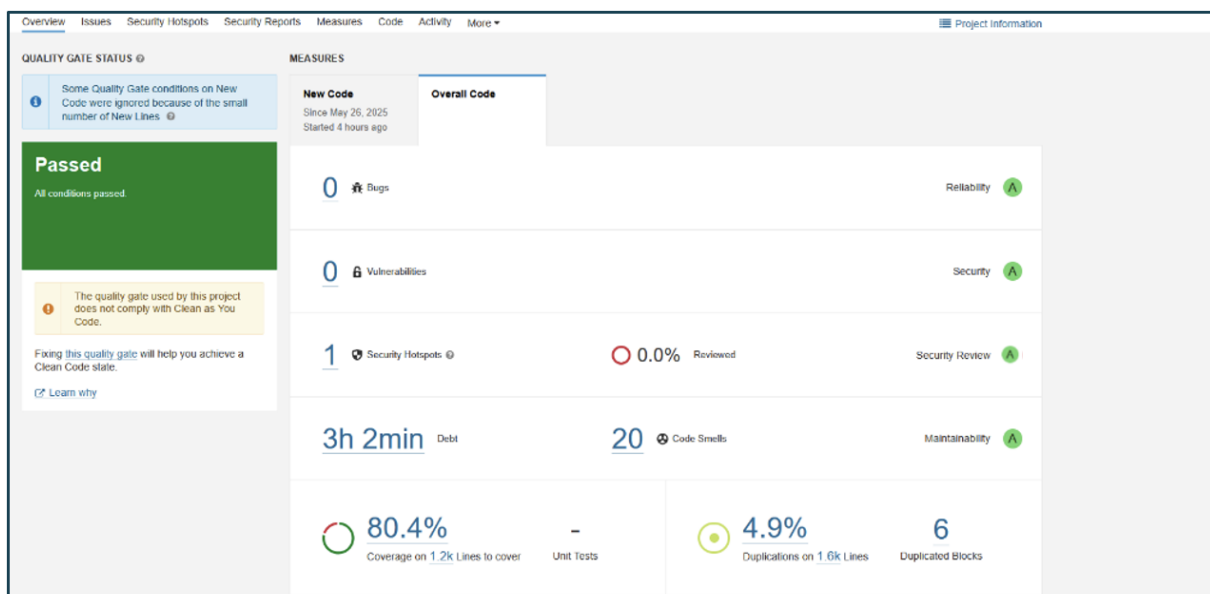


Figure 7.56 Train Detection Module Unitary Tests Sonar.

Modular Tests

Several modular tests were conducted to validate the different functionalities of the Train Detection Module. These tests verified that the module operated correctly under controlled conditions.

To verify the train detections have been correctly identified, an output video is generated containing all the processed frames. In this video, the bounding boxes, the estimated speed, and the detection confidence are displayed. This video is generated at the end of the Train Detection Module, once all frames from the input video have been processed. An example of the output video is shown in Figure 7.57.



Figure 7.57 Train Detection Module, video output result.

Acquisition Module

The acquisition module is responsible for initiating the frame capture and storage processes when a train is detected crossing the checkpoints. The objective of this process is to ensure that the acquired frames are valid and available for subsequent processing and analysis by the different modules of the system.

To guarantee the correct image acquisition, several laboratory tests have been carried out to validate the acquisition conditions before deployment in the Can Tunis scenario. These tests focus on evaluating the influence of illumination conditions and performing the adequate camera calibration to ensure stable and consistent frame capture. Examples of the laboratory setup and test performed to assess illumination conditions and camera calibration for both short and long-range scenarios are represented in Figure 7.58 - Figure 7.62.



Figure 7.58 Laboratory setup for illumination testing.



Figure 7.59 Captured images for different illumination conditions.



Figure 7.60 Acquisition camera calibration.



Figure 7.61 Short range camera without illuminator (distance = 300 cm).

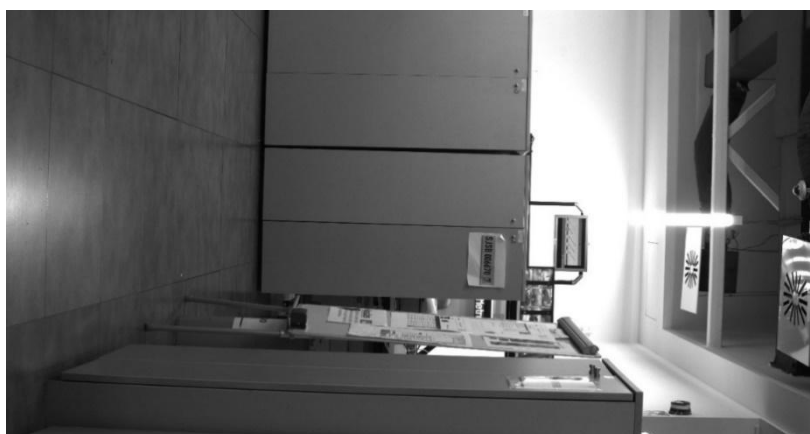


Figure 7.62 Long range camera without illuminator (distance = 750 cm).

Unitary Tests

Unitary tests were performed to verify the correct behaviour of the Acquisition Module, which have passed with a coverage of 90.7%, as it can be seen in the Figure 7.63 representing the Sonar report:

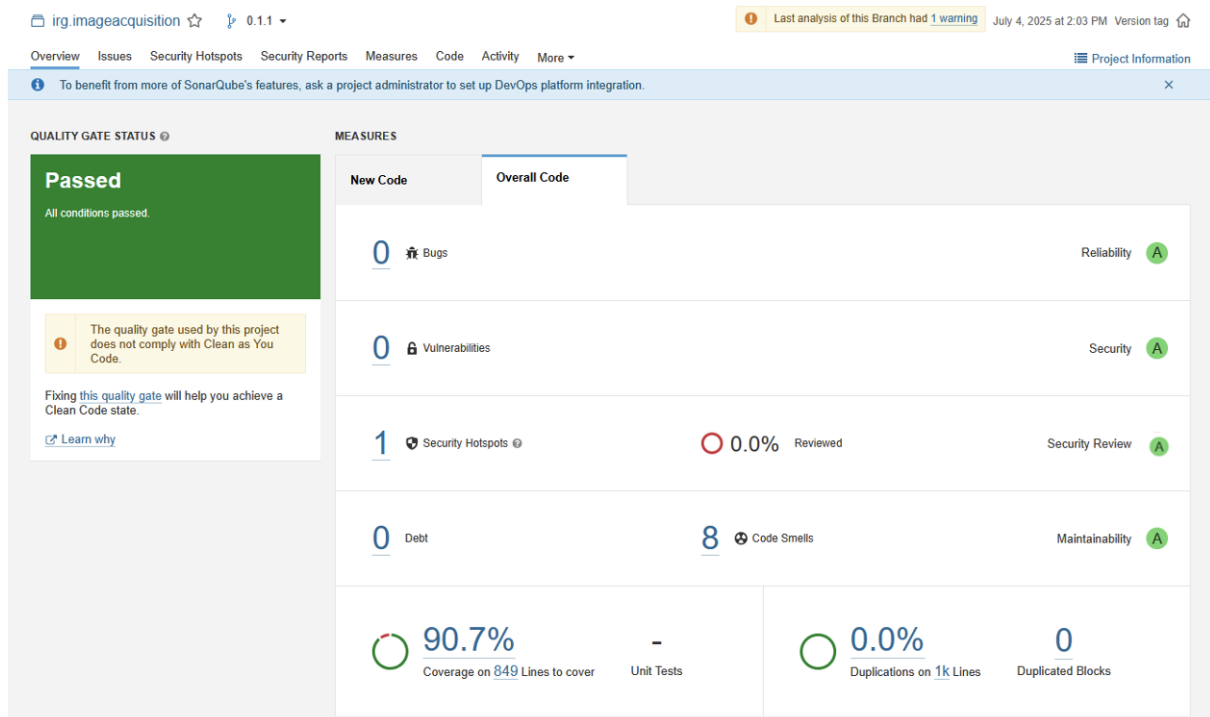


Figure 7.63 Image Acquisition Module - Unitary Tests Sonar report.

2. Classification and Processing Modules

Once the images of the train are acquired, the system performs a classification step followed by several processing and post-processing stages to analyse these frames, extract relevant visual information and identify the different elements of interest and train composition.

Through this sequence of classification, processing and post-processing steps, the system is able to reliably identify and characterise the different elements of the train, providing structured information that can be used by the subsequent modules of the system.

Unitary Tests

Unitary tests were performed to verify the correct behaviour of the different system modules. The Pre-classifier Module unitary tests have passed with a coverage of 83.6% as illustrated by Figure 7.64.

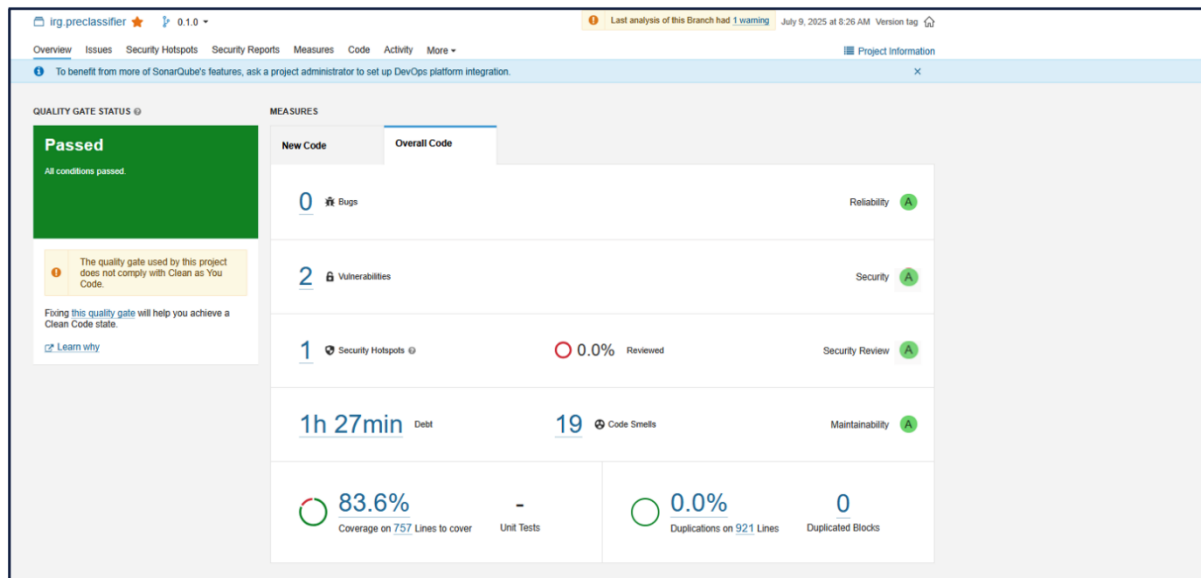


Figure 7.64 Pre-classifier Module - Unitary Tests Sonar report.

3. Stitching Module

The stitching module is responsible for generating complete images of each wagon from the sequence of captured frames. Since the acquisition module captures multiple partial views as the train moves across the camera field of view, the system must reconstruct a continuous representation of each wagon from these individual images.

As a result of the processes performed in the previous classification and processing modules, the images are properly grouped and organised according to the detected train and corresponding wagons. The reconstruction process relies on keypoint matching between consecutive images, allowing the system to identify common visual features and estimate the spatial relationship between frames:

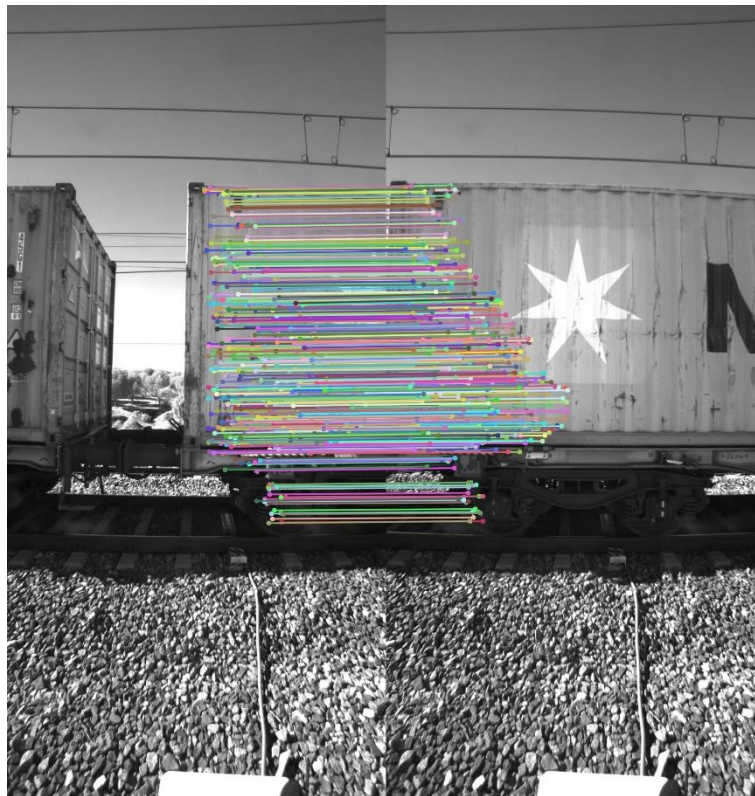


Figure 7.65 Keypoint matching between consecutive frames of a wagon from stitching module.



Figure 7.66 Stitching process: Panorama image of a complete wagon generated by merging multiple frames captured.

Figure 7.65 illustrates an example of the keypoint matching process between consecutive frames used to estimate the alignment. Figure 7.66, shows the resulting composed image obtained after combining the different frames belonging to the same wagon.

4. Labelling Platform for Training and Testing

The Labelling platform is the essential tool that we use to transform raw visual data into a structured, understandable format for machines. It creates the high-quality datasets required to train robust AI models and the reliable benchmarks needed to test them,

ultimately giving the ERC system deployed by Indra, the real-time visual intelligence to drive the different functions. During the labelling of images captured by on real environment, several processes are carried out to ensure the images are correctly labelled. Figure 7.67 shows a capture of the labelling platform. These processes take place across different systems and must follow a specific format that needs to be strictly adhered to.

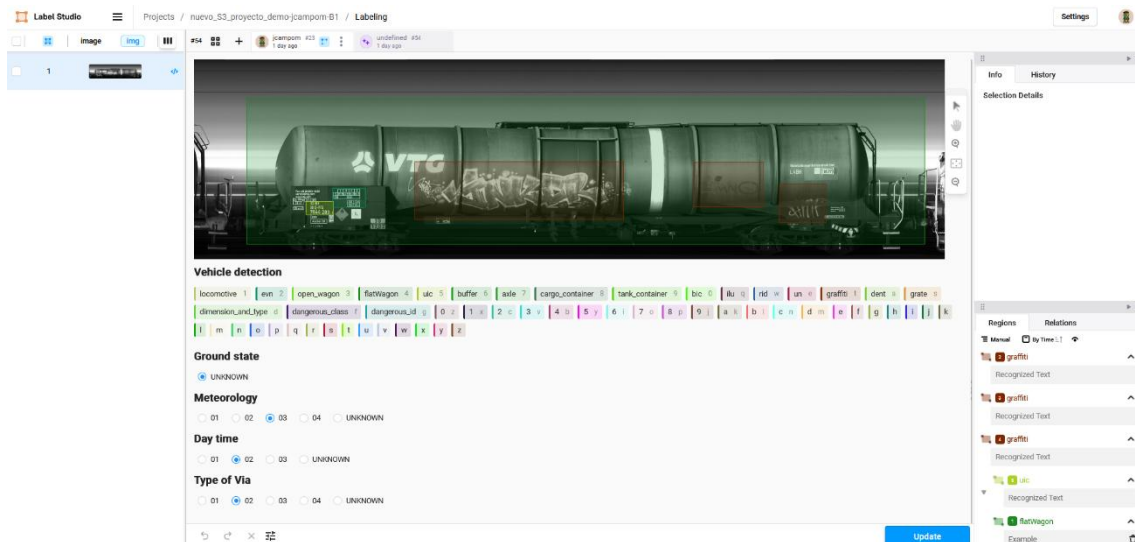


Figure 7.67 Capture of the labelling platform.

5. Dashboards

The ERC system integrates specific dashboards for each instantiation of the system that allow the visualisation of the system data. Figure 7.68 is a capture of the Dashboard component.

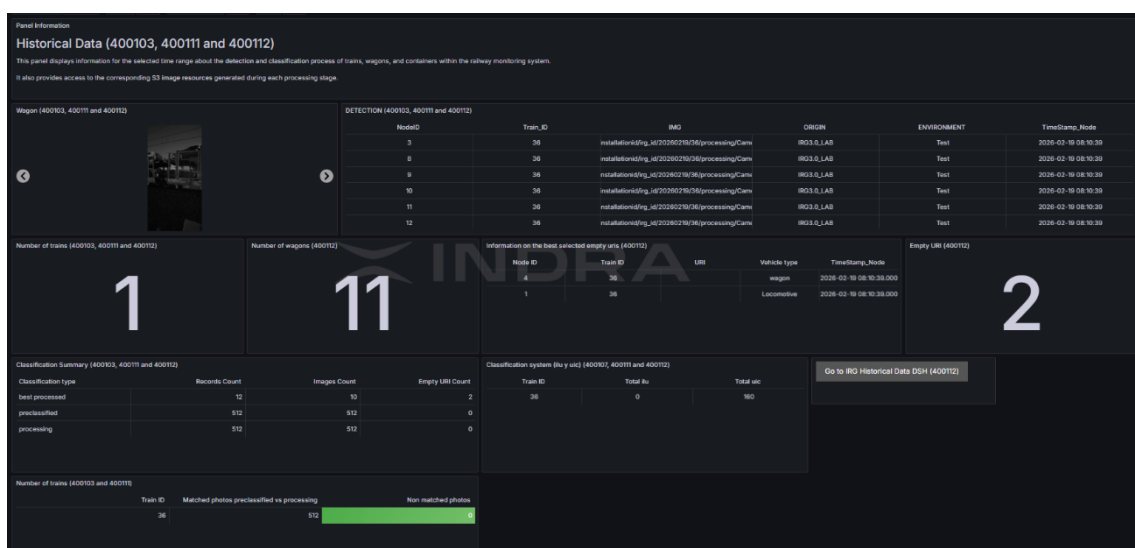


Figure 7.68 Example of checkpoint dashboard.

The data generated by each module of the different ERCs combined with a speed and configured repository and the dashboard give us the different metrics to validate the different modules, obtaining up to 94% of confidence on the different functionalities implemented.

7.3.3 Use case: Integrating checkpoints with On-board systems

The core use case is the implementation of on-board sensors, aimed at improving data acquisition capabilities for the monitoring and analysis of railway operations and complement and enhance Railway Checkpoints information.

In the use case that CEIT is proposing with RENFE and ADIF, the aim is to compare the data generated by the On-board system with the data acquired by the Checkpoint. However, this system can function autonomously within the train and provide valuable information such as:

- **Train composition:** The system identifies the specific sequence and orientation of every wagon in the convoy during shunting operations, providing an exact, real-time digital inventory of the train topology.
- **Wagon Integrity & Geolocation:** It provides continuous real-time tracking of the train's position on the terminal map, eliminating the need for manual visual checks and verifying the integrity of the convoy structure.
- **Automatic Brake Test:** It supports the digitalization of the manual inspection workflow via the on-board interface, allowing brake status to be recorded directly into the digital system.
- **Condition monitoring:** The continuous acquisition of "invisible" parameters regarding wagon health, such as vibrations and accelerations, to detect potential mechanical anomalies during operation.
- **Cargo monitoring:** The monitoring of internal environmental conditions (e.g., temperature, pressure) to ensure the integrity of the goods throughout the journey.



Figure 7.69 a) Field layer hardware architecture, b) CEIT's modular device for WOBUM (Wireless On-Board Unit Module), WOBUC (Wireless On-Board Unit Central), or WSEN (Wireless Sensor Elements) operation, c) Installation of an on-board unit

All these autonomous capabilities are made possible by the On-board system's unified architecture, which was validated in the demonstration across three layers:

1. **Field Layer** (Figure 7.69): The solution is built upon a versatile platform where all devices share the same high-performance base hardware. Depending on their configuration, they fulfil two distinct roles:
 - **Wireless On-board Unit (WOBUC):** The core gateway responsible for long-range communications (LTE-M/NB-IoT/LoRa), high-precision localization (GNSS/GPS/UWB), and the coordination of the on-board Wireless Sensor Network.
 - **Sensor Nodes:** Modular nodes installed on specific points (like axle-boxes) to detect hot axles, blocked brakes, or flat wheels.

Regardless of their role, all devices feature an Advanced Processing Core capable of processing complex data (vibration, temperature), Intelligent Connectivity to combine protocols and prevent signal loss, and Ultra-Efficient Power Management designed for years of autonomy.

2. **Data Layer:** The integration allows the transformation of a "silent" freight wagon into a connected asset. By feeding this data into a Federated Data Space (FDS) through IDS Connectors, the On-board system provides the necessary telemetry to complement the visual information captured by the wayside Intelligent Video Gates (IVG). This fusion fulfils the "Beyond the Checkpoint" vision, creating a seamless flow of information between the rolling stock and the infrastructure.
3. **Client Layer:** The operational core is supported by a Client Layer, which acts as the primary dashboard for value-driven data consumption. During the Can Tunis demonstration, this tool provided:
 - **Remote Command and Reconfiguration:** Allowing operators to adjust parameters and communication intervals without physical intervention.
 - **Operational Intelligence:** Providing real-time visualization of the train composition, asset positioning on the terminal map, and the status of critical components.
 - **Data Integration:** Serving as the visual gateway for all collected data, ensuring information is actionable for traffic management.

7.3.4 Evaluation and Preliminary Results

Further results of the Spanish demonstrator can be found in deliverable D29.6. The following is a summary of the results obtained.

7.3.4.1 European Railway Checkpoint

As previously mentioned, the Railway Checkpoints have not yet been deployed at Can Tunis at the time of writing this report. Consequently, it is not currently possible to perform an evaluation or carry out an analysis of the results on the real environment, on laboratory as we mentioned and based on synthetic datasets we have reached a preliminary confident near 94% for the overall functionalities developed, nevertheless, the final assessments will be carried out once the installation and testing phases are completed, and the future results will be available in deliverable D29.6.

The full deployment of the Intelligent Video Gate (IVG) system will have a major impact on logistics, safety, and operational efficiency. It will enable automated monitoring of freight trains, ensuring accurate identification of rolling stock and cargo, tracking dwell times, and supporting infrastructure management. At the same time, the system will be the basis for future functionalities to be developed on FP5 Wave 2; functionalities that will help to prevent risks for the railway infrastructure, reducing the train/Cargo inspection and contributing to safer and fully digitalized rail operations.

7.3.4.2 On-board sensors

The field trials carried out at Can Tunis generated a complete dataset confirming the effectiveness of the on-board system. Key results include:

- **Installation Efficiency:** Even with wagon corrosion, the combination of adhesive plates and high-strength magnets enabled installation in under 5 minutes per asset.
- **Dynamic Train Composition Accuracy:** The system achieved 100% accuracy in detecting wagon orders. It correctly identified the initial configuration (Loco-3-2-4) and automatically updated it after shunting to (Loco-4-2-3).
- **Real-time Geolocation:** Continuous and precise convoy tracking was maintained throughout the test lap, eliminating manual checks and demonstrating stable GPS/LTE-M performance in motion.
- **Sensor Data Monitoring:** The Wireless Sensor Network consistently delivered high-quality raw data (temperature, pressure, vibrations, magnetometer, accelerations). All values were transmitted via LTE-M, displayed correctly on the dashboard, and ingested into the FDS.
- **Remote System Control:** Bidirectional communication was validated. Operators successfully sent reconfiguration commands to the WOBUs, confirming remote management capabilities without the need for physical access.

Table 7 summarizes the validation of the system against the initial requirements defined for the Can Tunis pilot.

Table 7 Validation criteria

Requirement / KPI	Target Criteria	Test Result	Status
Deployment Speed	< 10 min per wagon	< 5 min (including surface prep)	PASSED
Composition Logic	Correct sequence detection	100% Accuracy (Both configurations)	PASSED
Asset Visibility	Real-time map tracking	Continuous & Precise	PASSED
Data Integrity	Reliable Raw Data capture	Complete stream ingested to FDS	PASSED
Operational Control	Remote command capability	Verified via Web Dashboard	PASSED

The integration of the on-board system with the Railway Checkpoints delivers a major improvement in terminal operations demonstrating that it is possible to transform a 45-minute technical stop into a data flow lasting only a few seconds. This shift from manual inspections to automated data synchronisation is the foundation of the *"Beyond the Checkpoint"* vision.

The benefits of this combined solution span three key transformation areas:

- **Operational Efficiency (The Corridor):**
 - KPI 5.4 – Increased Commercial Speed: Eliminating manual delays raises the effective speed of freight movements.
 - KPI 5.6 – Enhanced Reliability: Automated checks and real-time condition monitoring reduce unexpected failures and unplanned stops.
- **Environmental Sustainability:**
 - KPI 5.9 – Energy Savings and Emission Reduction: Removing the traditional 45-minute idling period cuts fuel consumption and significantly lowers CO₂ emissions per convoy.
- **Commercial Control and Predictability:**
 - KPI 5.10 – Precise Billing and Traceability: Real-time timestamps and asset status data enable more accurate billing, complete cargo traceability, and greater predictability across the supply chain.

7.3.5 Issues and challenges identified

Throughout the development and deployment phases, several technical, operational, and integration-related challenges emerged that provide valuable insights for the system's future evolution. This section summarizes the main issues encountered and outlines the lessons learned that will guide the next steps of the roadmap.

7.3.5.1 European Railway Checkpoints

Although the deployment of the ERCs has been fully designed and planned, the start of the physical installation has experienced significant delays. These delays are not the result of a single obstacle but rather a combination of different issues that have a direct impact on the project timeline and must be resolved before any on-site work can safely begin.

- **Can Tunis civil works:** The current civil works taking place at the Can Tunis terminal affect both the track alignments and the catenary posts and electrification systems. As a result, the initial design of the Railway Checkpoints had to be revised to accommodate the new track configuration and ensure proper coordination between the construction activities and the checkpoint installation.

- **Administrative requirements:** Obtaining the necessary permits to install the system on railway infrastructure entails a lengthy and complex administrative process. The approval procedure involves strict regulatory compliance, and thorough technical validation, all of which significantly extend the overall authorisation timeline.
- **Mechanical installation issues:** The high energy requirements of the equipment and the remote location of the installation points make difficult identifying a suitable and reliable supply source within the terminal proved complex.

7.3.5.2 On-board systems

The Can Tunis pilot has successfully demonstrated the viability of the On-board system as a cornerstone for the digitalization of freight rail. However, the trials also identified technical challenges and regulatory requirements that define the roadmap for future development.

- **Mechanical Installation and Surface Integrity:** A significant challenge was encountered due to the corrosion on the legacy wagons' metallic structures, which compromised the primary magnetic adhesion. To ensure stable and reliable mounting, the procedure was adapted by installing intermediate metal plates secured with industrial adhesive and fasteners. This experience highlights the need for a standardized "surface preparation kit" for older rolling stock to maintain the plug-and-play advantage.
- **Wireless Interference and Range:** It was observed that 5G interference in the terminal environment significantly affected Ultra-Wideband (UWB) performance, reducing the effective range from approximately 20 m to 10 m. This reduction directly impacts the train composition algorithm, which relies on high-precision distance and orientation. To overcome this, the deployment strategy for the Can Tunis demo was adapted by installing two WOBUs per wagon. This configuration effectively shortened the wireless link distances between WOBUs and ensured a robust, redundant network topology that maintained composition accuracy despite the challenging electromagnetic environment.
- **Regulatory Compliance:** Following Renfe's requirements, a formal modification document was developed in alignment with RUE 402 (Commission Implementing Regulation (EU) No 402/2013). This document ensures that the installation is recognized as a non-intrusive, low-complexity, and fully reversible modification, providing the necessary safety impact and risk assessments for commercial operation.

7.3.6 Conclusions and recommendations

Although the European Railway Checkpoints have not yet been installed, the work conducted so far at the Can Tunis demonstrator provides clear evidence of the potential impact of combining future ERC capabilities with the already validated on-board sensing system. The trials have demonstrated the feasibility to improve the detection, timing and reliability of yard events, supporting CDM and anticipating benefits in efficiency and situational awareness.

Next steps include:

- **European Railway Checkpoints:**
 - Complete the deployment of the ERC in Can Tunis and carry out the final system tests and functional validation.
 - Refine algorithms and conflict resolution rules.
 - Develop new automated fault-detection capabilities, including acoustic monitoring (microphones) and computer-vision-based inspection.
 - Integrate and automate data exchange with the Traffic Management System (TMS) to ensure real-time synchronization with traffic-control operations.
- **On-board systems:**
 - **Scalability and Large-scale Deployment:** Future phases will require longer-duration deployments with a significantly larger number of wagons and locomotives (full convoys) to validate system stability in massive fleet operations.
 - **Brake Test Digitalization:** While the Automatic Brake Test workflow was successfully integrated into the dashboard, the next step involves the final calibration of the detection algorithms. The goal remains to fully replace manual inspections with a digital record in the FDS, enhancing safety and drastically reducing departure bureaucracy.
 - **Advanced Data Analytics:** Having validated the reliable transmission of High-Quality Raw Data, future work will focus on moving from monitoring to "active alerting." This involves the deployment of edge-computing algorithms to detect specific anomalies (e.g., hot axle boxes or flat wheels) automatically.
 - **Wayside Integration:** The system is now fully "FDS-ready." Once the Indra IVG gate infrastructure is deployed, the final validation phase will involve the real-time synchronization of On-board data with wayside gate events, completing the *Beyond the Checkpoint vision*.

7.4 Module 4: Recognition of Hazardous Materials Plates/Placards and UIC Code (the Netherlands)

In the Netherlands, ProRail has six IVG's in use at railway yards: at *Moerdijk* (1), at *Waalhaven* (2) and at *Botlek* (3) mainly for code recognition and wagon counting to check the wagon lists. At Moerdijk Zijdelijke randweg a big portal with multiple cameras for monitoring of train wagons and brake detection is installed. Moreover, two Videogates for freight trains are operational near the Port of Rotterdam at the "Havenspoorlijn" mainline. These Videogates have been used for code recognition and wagon counting to check the wagon lists in the test reported in this section.

Further gates are installed for monitoring passenger trains and pantograph monitoring, e.g. in *Hogebrug* on the main line and at *Schiphol Airport*. In the future these may also be utilized for freight train monitoring. The Videogates for passenger trains are used for image recognition of pantograph defects, pantograph carbon strip wear and wheel defects (missing bolts, loose cables).

The locations are further described in D29.1 (Kordnejad et al., 2024). Installation and test plan is described in D29.4 (Veld and Tijink, 2025), while execution of test is presented in D29.7 (Veld and Tijink, 2026).

7.4.1 Use case: Identification of dangerous goods and UIC Code

It is a legal obligation (Network Statement 2024/2025 Article 58 of the Railway Act (Spoorwegwet), (ProRail, n.d.) for rolling stock operators to indicate where wagons carrying dangerous goods are located in the train/wagon sequence. The infrastructure manager needs to have this information in order to guard the safety of the railway system and its surroundings. The indication of dangerous goods is currently done manually and are also indicated on the wagon itself, with "hazardous materials signs".

Hazardous materials can be indicated with (hazmat) signs. These are diamond-shaped signs with information on the type of hazmat. The signs have distinctive colours and inscriptions to identify classes and subclasses. The hazmats are also displayed in an orange sign. These orange signs contain a number that contains information on the substance (UN number) and the degree of danger. The UN number is the only required information, however it can be useful to detect the hazmat signs as well. They can function as a backup or to boost confidence when predicting the UN number.

Using the IVG we can visually inspect the wagons to locate and identify all hazmat signs on a wagon.

Automation of rolling stock and ILU code identification is a topic of great interest in the railway industry. Every rolling stock and every ILU are identified with a different code that

is used to track each element of a freight train composition. The UIC codes must be read and registered every time a freight train enters or leaves a railway yard, a station, crosses a border or a control point. This code registration task is manually done by operators and represents large costs due to the time required to check the codes, the possibility of human errors and the risk to operators due to the harsh conditions of railway environments.

7.4.2 Main objectives Recognition of Hazardous Materials Plates/Placards

Trains have two types of Hazardous Materials (hazmat) signage: an orange, rectangular plate with two lines of numbers (UN-number) and diamond-shaped placards (see Figure 7.70). The orange plates indicate the precise material being transported, while diamond-shaped placards indicate the general hazards associated with the material, such as flammability or oxidization.

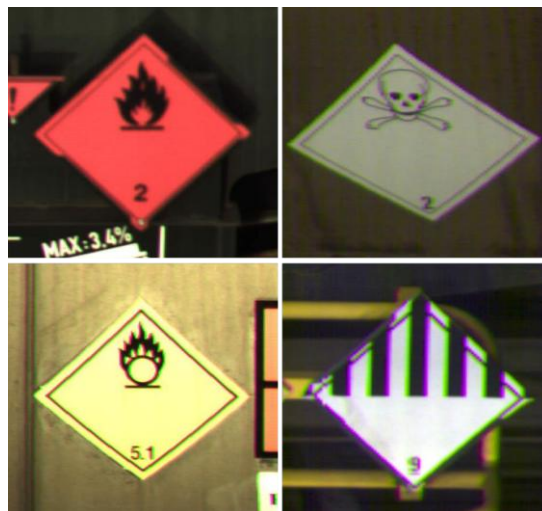


Figure 7.70 Examples of Hazmat signs, data provided by DB Cargo.

To automatically read the hazmat numbers, we designed a pipeline that detects the plates, subsequently extracts the plates and then subjects them automatic number reading, also known as Optical Character Reading (OCR).

7.4.3 Main objectives Reading Wagon Codes

Wagon numbers, or UIC numbers, are 12-digit numerical codes that uniquely identify a wagon, see examples in Figure 7.71. Automatically reading these codes from the IVG data would make it possible to automatically create a wagon list and verify said list along the whole route.

To automatically read UIC numbers, we designed a pipeline that detects and reads UIC codes from an input image or video frame. The pipeline is based on the STEP (Structured TExt sPotter) method, designed by Garcia-Bordils et al. (2023). Uniquely, this method performs localization and reading in the same step, which other methods typically do in two steps (as in the hazmat plate method described above). To guide the search, we do

not consider all text on the wagon but help the algorithm by defining the structure of the text we are looking for. Using a regular expression (regex), we can explicitly tell the model to look only for sequences such as digit-digit-space-digit-digit-space-digit-digit-digit. A drawback of the existing STEP algorithm was its suitability for one-line text only, while UIC codes often appear on multiple lines. Furthermore, code structures vary widely in reality, as they may span a large surface and may not appear in full in one easily demarcated space. By adding several modules, tapping into prior knowledge of the codes, we implement hierarchical regex matching to deal with the semi-structured nature of the UIC codes and multi-line UIC codes, instead of focusing solely on full 12-digit codes. A full description of the methodology is provided in a manuscript currently under preparation and will be published upon acceptance.



Figure 7.71 Examples of UIC codes. All images are provided by Garcia-Bordils et al. (2023).

7.4.4 Results Recognition of Hazardous Materials Plates/Placards

To train these algorithms, we created a dataset based on video data captured by a ProRail IVG. Dataset diversity is good, as video lengths vary and illumination conditions differ as a result of the time of day (day/night) and weather conditions. Only frames with a hazmat plate were considered. In total, the dataset contains 9,560 frames, extracted from 24 annotated videos, with 13 unique UN hazmat codes. The dataset was divided into training, validation, and testing subsets following an 80%-10%-10% split, resulting in 7,648 frames for training, 956 frames for validation, and 956 frames for testing. The plates were annotated using bounding boxes, and ADR/RID (European Agreement Concerning the International Carriage of Dangerous Goods by Road/Regulation concerning the International Carriage of Dangerous Goods by Rail) and UN numbers were manually annotated.

We evaluated the methodology in three steps: placard detection, OCR, and the full pipeline. Placard detection was evaluated using the Intersection over Union (IoU), see Equation (1), of the predicted bounding box and the labelled ground truth. The mean Average Precision (mAP) can be computed for different IoU thresholds. The mAP represents the average area under the precision-recall curve for all classes. The IoU threshold determines whether a prediction is considered correct or false. A high mAP means both good precision and recall, meaning a better model.

$$\text{IoU} = \frac{A \cap B}{A \cup B} = \frac{\text{Area of Overlap}}{\text{Area of Union}} \quad (1)$$

The OCR step was evaluated using the Character Error Rate (CER) and Word Error Rate (WER), shown in Equations (2) and (3), respectively. Both metrics are based on the minimum number of operations (deletions, insertions, and substitutions) needed to transform one text into another. CER measures errors at the character level, while WER evaluates accuracy at the word level. Both lower CER and WER mean better model performance.

$$CER = \frac{\text{Number of incorrect characters}}{\text{Total number of characters}} \quad (2)$$

$$WER = \frac{\text{Number of incorrect words}}{\text{Total number of words}} \quad (3)$$

The evaluation framework for the pipeline assesses performance across two key stages: detection and OCR. Placard detection is evaluated using the Intersection over Union (IoU) metric, where a detection is considered valid if its IoU with at least one ground-truth (GT) bounding box exceeds a predefined threshold. Detections are considered invalid if this threshold is not reached or if no detection is made. OCR is applied to successful detections to extract the UN numbers, which are then compared to GT values. A prediction is correct only if both numbers match the GT, while any mismatch results in an error. We evaluated overall performance by measuring accuracy, or the number of correct predictions compared to the total number of predictions. Higher accuracy means better performance of the full pipeline.

For the detection stage we considered two models: Yolo11x [YOLO] (Jocher, 2025) and Faster-RCNN [F-RCNN] (Ren et al., 2015). Details on the models can be found in the full paper. Both models are trained on detecting hazmat plates, using the annotated training portion of our labelled test set of ProRail IVG data. Quantitative results are shown in the Table 8 below. For a mAP of 50%, detection accuracy can reach values above 99%.

Table 8 Performance metrics for YOLO11x (Jocher, 2025) and Faster RCNN (Ren et al., 2015) under different training strategies on the two test sets. The following training strategies were used: Early Stopping (E.S), Lower Learning rate (L.L), and Data Augmentation (D.A.)

Model	ProRail-Test	
	mAP50-95 (%)	mAP50(%)
YOLO11x [YOLO]	57.69	99.37
YOLO11x E.S.	50.72	98.77
YOLO11x E.S.+L.L.	56.46	99.38
Faster-RCNN [F-RCNN]	56.34	98.92
Faster-RCNN D.A.	61.67	98.98
Faster-RCNN E.S.+D.A.	61.25	98.97

The results of the OCR step are shown in Table 9 below. The WER and CER scores are shown for three different models: Tesseract [TA] (Hoffstaetter and Lee, 2025), EasyOCR [E-OCR] (Kittinaradorn, 2025) and Idefics2 [IDF2] (Laurençon et al., 2024); Some qualitative examples are shown in the table below. Although the third model, Idefics2, clearly outperforms the other models, even the best performing model can fail as a result of low-resolution images or intra-plate font variations.

Table 9 Performance comparison of Tesseract [TA] (Hoffstaetter and Lee, 2025), EasyOCR [E-OCR] (Kittinaradorn, 2025), and Idefics2 [IDF2] (Laurençon et al., 2024) models on ProRail dataset, based on the Character Error Rate (CER) and Word Error Rate (WER). Lower CER and WER mean better performance.

Model	ProRail CER	ProRail WER
Tesseract [TA]	56.3	182.19
EasyOCR [E-OCR]	29.99	98.6
Idefics2 [IDF2]	4.85	15.20

The performance of the full pipeline, using the best performing models Yolo11x [YOLO] (Jocher, 2025) and Idefics2 [IDF2] (Laurençon et al., 2024), is summarized in Table 10, where we show the overall accuracy (correctly detected and read plates) for different IoU thresholds. This table shows that we can reach over 90% accuracy on the IVG data when reading Hazmat number plates. The main failed cases are low resolution, weather, or skewed surfaces.

Table 10 Accuracy of our proposed pipeline on the ProRail dataset at different IoU thresholds. The pipeline uses the YOLO11x [YOLO] (Jocher, 2025) and Idefics2 [IDF2] (Laurençon et al., 2024) models.

Threshold IoU	Accuracy ProRail-Test
40%	90.9
50%	90.3
70%	54.2

7.4.5 Results Reading Wagon Codes

The method was evaluated on three different data types: image data, video data and line-scan data. Since several camera types are still used within the project, we preferred a method that could handle all inputs rather than only one data type. We will first describe the datasets, then the evaluation metrics, and finally show the performance.

7.4.5.1 Image data

The dataset is provided by the authors of STEP (Garcia-Bordils et al, 2023) and consists of 396 pictures of stationary wagons taken different distances and angles. The wagon codes have different fonts and regular expressions, but most are single-line codes.

7.4.5.2 Video data

This dataset consists of video data captured by an IVG in the Netherlands over the course of several months at one location. The camera was triggered by a passing train and collected up to 10 minutes of data. A selection of 70 videos of varying lengths and capture times was made to label the UIC codes. Codes were labelled using the interpolation tool of the Computer Vision Annotation Tool (CVAT). During the labelling process, any anticipated challenges with the codes were also noted, such as occlusions and low contrast. The full videos were also annotated, distinguishing between day/night and the presence of rain.

7.4.5.3 Line-scan Data

The third dataset is also private and consists of grayscale line-scan images of freight wagons. There are two line-scan cameras on both sides of the train tracks, meaning that each wagon is photographed from two sides. Due to asymmetry and failed photos, the images do not fully overlap, resulting in 831 images total, of which 430 are from one camera and 401 from the other. This dataset was also labelled using the CVAT tool. In addition to the UIC codes, challenging conditions were also noted.

7.4.5.4 Evaluation metrics

To evaluate UIC code-reading performance, we have to define when the model output is correct and when not. The first criterion concerns the final digit of the 12-digit code, which is a control digit that serves to identify any mistakes in the prior 11 digits. The control digit is determined using Luhn's algorithm. The second criterion consists of verifying whether the predicted code matches the ground truth: the actual UIC code present in the image. Using the True/False Positives and True/False negatives, we can calculate Accuracy, Precision and Recall. The image dataset contains no True Negatives, as all images feature a UIC code.

The video frames were evaluated using a sliding window format, as combining information from several frames may produce stronger predictions. As we only need to identify unique wagons to make a wagon list, we do not consider individual frame predictions and only output unique predictions per video. The sliding window has two parameters: window length (how many frames are considered together) and a threshold (to determine the minimum number of frames within a window in which the same code has to be found to make a final prediction). This approach means that we only compare the predicted list of unique UIC codes to the GT list of UIC codes, ruling out any True Negatives.

As mentioned before, two-side views are available for most line-scan images. In addition to considering all individual images, we also combine the frames. We use the validity of

the code as a selection criterion to determine the final prediction for each wagon. For example, if no code or an invalid code is found on the left-side view, while a valid code is found on the right-side view, the prediction of the second view is considered final. In this dataset True Negatives are possible, as some images do not have a UIC code, and this should also be predicted.

7.4.5.5 Experiment 1 – Image data

In this experiment, we evaluated the images in Dataset 1 – Image data. The results (see Table 11) show that we can reach an accuracy of more than 79% on this data, marking a major improvement compared to the original model. There is, however, still room for improvement. The results demonstrate that the main causes of incorrect predictions were difficult fonts (e.g. open fonts), occlusions, and other 12-digits codes.

Table 11 Comparison of OCR model performance on the STEP Image Dataset [STEP] (Garcia-Bordils et al, 2023), demonstrating that our proposed method achieves substantially higher accuracy, recall, and precision than existing baselines.

Model	Accuracy (%)	Recall (%)	Precision (%)
STEP	22.5	23.1	89.0
Our model	79.3	84.4	92.9

7.4.5.6 Experiment 2 – Video Data

Table 12 below shows the results per category with a window length of 12 and a threshold of 60%.

Table 12 Performance evaluation of our proposed model on video dataset subsets. Each row compares the model's performance on frames filtered by the condition in the leftmost column (subset) versus its complement (all other frames). Results show that occlusions and low-contrast conditions lead to notable drops in accuracy and recall.

Subset	Percentage dataset	Metrics Subset			Metrics Complement Subset		
		Accuracy(%)	Precision(%)	Recall(%)	Accuracy(%)	Precision(%)	Recall(%)
All data	100	75.0	92.1	80.1	-	-	-
No issues	52.5	79.1	88.4	88.1	70.4	97.1	72.0
Night	28.9	71.3	96.1	73.5	76.5	90.7	83.0
Occlusion (all)	14.5	55.8	100.0	55.8	78.3	91.2	84.6
Rain	3.5	76.2	94.1	80.0	75.0	92.0	80.1
Low contrast	2.7	56.3	56.3	100.0	75.5	92.0	81.0
Night + occlusion	2.4	35.7	100.0	35.7	76.0	92.0	81.3
Occlusion only	11.7	62.3	100.0	62.3	76.7	91.3	82.7

The first row shows the results on the complete dataset, with accuracy peaking at 75%. The second row shows the results for images without any reported issues, with accuracy reaching 79% for the no-issue set and 70% for the complement subset (all data with issues). For UIC codes captured at night, recall is significantly lower, indicating that this

condition is more likely to cause false negatives (wrong or missed codes). We also see that occlusions and low contrast hamper model performance, as evidenced by a clear drop in performance. The combination of nighttime conditions and occlusion in particular result in very poor performance, with an accuracy score of 35.7%. The models are especially sensitive to these errors because their core underlying principle is finding regular expressions. It is an “all-or-nothing” approach, meaning models will look only for full codes or nothing. If a part of a code is occluded, the complete code will not be found. The remaining failures (since accuracy was only 79.1% at best), could be caused by unlabelled issues, missing regular expressions, or wrongly detected codes.

The results suggest that rain does not negatively impact performance, with the proviso that rain is only present in a small portion of the data (3.5%). While we can therefore hypothesize that it will not cause problems, future research is needed to analyse the full impact of rain on the results, focusing particularly on heavy rain and the combination of rain and nighttime/low-light conditions. This calls for a larger dataset, either collected by an IVG or synthesized by implementing synthetic rain on the available data. The latter method would provide full control over the severity of the rain and light conditions, but real data should always be used to verify the impact on real-world implementation.

7.4.5.7 Experiment 3 – Line-scan data

The results on the line-scan data are shown in Table 13 below. The model boasts 79.8% accuracy and even reaches an accuracy score of 86.2% on wagons captured by both cameras. Combining information from different views clearly improves performance and makes the system more robust against missing codes or failed detections.

Table 13 Comparison of the STEP [STEP] (Jocher, 2025) and our proposed model on line-scan data. Our model demonstrates substantially higher accuracy, recall, and precision across both individual camera views and combined camera views.

Model	Data Subset	Total	Accuracy (%)	Recall (%)	Precision (%)
STEP [STEP]	All	831	12.6	2.84	32.8
	Camera 1	430	13.5	4.08	44.1
	Camera 2	401	11.7	1.48	18.5
	Combined	507	19.1	3.84	30.6
Our model	All	831	79.8	84.6	90.4
	Camera 1	430	80.2	84.2	91.6
	Camera 2	401	79.3	85.0	89.1
	Combined	507	86.2	91.8	90.4

Similar to the experiment on video data, we ran an extensive analysis on possible challenges within the data. The results of this analysis, focusing only individual images (not combined views), are shown in Table 14.

Table 14 Performance Evaluation of our proposed model on top 5 most common challenges affecting line-scan data and their complement subsets. Each row compares the model's performance on frames filtered by the condition in the leftmost column (subset) versus its complement (all other frames). Results show that occlusions and low-contrast conditions lead to notable drops in accuracy and recall.

Subset	Percentage dataset	Metrics Subset			Metrics Complement Subset		
		Accuracy(%)	Precision(%)	Recall(%)	Accuracy(%)	Precision(%)	Recall(%)
Font size variation	12.39	74.8	81.1	90.6	80.5	92.0	83.7
Faded	8.66	68.1	90.6	72.7	80.6	90.3	85.8
Occlusion	7.70	65.6	85.7	73.7	81.0	90.7	85.6
Low contrast	5.42	62.2	93.3	65.1	80.8	90.2	85.9
Angle	5.29	97.8	97.8	100.0	78.8	89.8	83.5
No issues	44.89	84.7	93.0	90.5	75.3	86.5	78.5

The results show that occlusions and low contrast are particularly challenging for the model, as was also the case for the video data. Since the recall is lower in these situations, false negatives (missing codes) are more common. The last row of Table 14 shows that the accuracy is 84.7% on the subset without any labelled issues. The remaining errors are caused to mislabelled codes, missed challenges while labelling, other 12-digit codes on wagons resulting in false positives, and regular expressions missing from the algorithm's regex list. The results help pinpoint several ways to improve the overall methodology: robustness against occlusions, low contrast, and font variations. It is also strongly advised that wagon owners address these issues by adopting more legible and standardized wagon codes.

7.4.6 Issues and challenges identified

Focusing on the ProRail IVG data, the findings show that hazmat number plate recognition performs very well in real-life operating conditions: the full pipeline achieves over 90% accuracy on the ProRail test set when reading UN numbers. For hazmat placards, the main conclusion is that including synthetically damaged training data significantly improves real-world performance. Models trained only on publicly available datasets (mainly U.S. placards) perform notably worse on the DB wagon placard dataset, whereas training strategies that add EU-style synthetic damage produce the highest F1 scores. This clearly highlights a gap between public, clean training data and the worn, scratched, or partially degraded placards found in the real world, and demonstrates the value of synthetic augmentation for bridging this gap. Similarly, UIC recognition reaches 75% accuracy on ProRail video data and 79–86% on line-scan data, while combining information from both camera views improves performance. These results indicate that the proposed pipelines are already robust to typical IVG conditions such as speed variation, day/night settings, and minor visibility challenges.

For hazmat plates, the evaluation was carried out on a relatively small real-world test set, meaning that important sources of variability (rain, snow, or other challenging outdoor conditions) were not fully represented. For placard classification, only cropped placard images were used rather than full-frame wagon images, so the detection stage (locating the placard on the wagon) was not fully assessed, nor were systematic outdoor effects such as weather, lighting changes, or surface damage beyond what was synthetically generated. For UIC recognition on both line-scan and video data, only limited weather effects were included in the datasets, and the results show that occlusions and low-contrast situations remain the dominant causes of failure, indicating that the system still lacks robustness in these common real-world scenarios.

In addition, all evaluations were performed in an offline setting, meaning the algorithms were tested separately from the operational IVG hardware rather than in a fully integrated, real-time system. Although real-world IVG data was used throughout, the processing itself ran on external hardware, and the full end-to-end pipeline was not implemented. As a result, processing time and system efficiency were not evaluated. Assessing real-time performance on the actual hardware platform is therefore an essential next step, as the focus of this work was on demonstrating the recognition potential of the algorithms rather than their execution speed or integration readiness.

Across all evaluated components, the most effective solution combines the YOLO11x detector with the Idefics2 OCR model for hazmat number plates, delivering over 90% accuracy on ProRail data and showing strong robustness under typical IVG conditions. For hazmat placards, the classification models trained with synthetically damaged EU and US placards achieve the highest F1 scores, demonstrating that realistic augmentation is essential for real-world performance. For UIC codes, the method with hierarchical regex matching provides a substantial improvement over existing baselines across image, video, and line-scan data. Together, these components currently form the strongest-performing configuration and constitute the most reliable foundation for an integrated wagon-inspection system.

Future work should focus on moving the system towards full operational integration. A key next step is to implement wagon counting, which requires full access to all raw sensor data so that information from multiple sensors can be combined effectively. More extensive evaluation under real-world conditions, including varied weather, lighting, and environmental effects, is essential to ensure robustness. For hazmat placards, the recognition pipeline should be tested on full-frame wagon images, which will require additional labelled data to evaluate the complete detection–classification chain. Beyond

recognition accuracy, the efficiency and processing time of the full pipeline must be assessed on the actual hardware to verify real-time feasibility. In the longer term, improving resilience to occlusions and unreadable UIC codes, for example by introducing a dedicated “code not readable” flag, will further enhance reliability. Ultimately, the goal is to combine all developed algorithms into a single integrated system capable of handling plate reading, placard classification, UIC recognition, and wagon counting within one unified framework.

8 Data management

To facilitate the sharing of rail-specific data across borders, a minimal conceptual data model (CDM) was developed and tested using real rail and freight data. More than 2,000 datasets were sent by DB Cargo to an Application Programming Interface (API) endpoint of the Deplide system provided by RISE. The data chosen for sharing included AI-detected information about brake pad thickness, damping ring and lifting T, as mentioned in Section 7.1, as well as human annotations about damages seen in images from German IVGs. Additional manual tests were performed with information about UN numbers and ILU codes for certain wagons.

8.1 Data Sources

The datasets used for the evaluation activities described in this deliverable originate from the ERC demonstrators deployed in the participating countries. To ensure an overlap, the wagons for the data exchange were chosen by comparing lists of wagon numbers from the Swedish IVG in Gothenburg with those that had passed through at least one of the 13 German IVGs (see Figure 8.1). As not all of these wagons were damaged, the datasets also included all Swedish wagons that had passed through the German IVGs in the last six months. The data source for the exchange was the AWS (Amazon Web Services) backend cloud system of the German IVGs. A scheduled service extracted the data from the PostgreSQL databases and sent it to the Swedish API on a weekly basis over multiple weeks. The path of the data exchange is simplified and can be seen in Figure 8.2.



Figure 8.1 Data transfer between German and Swedish Checkpoints

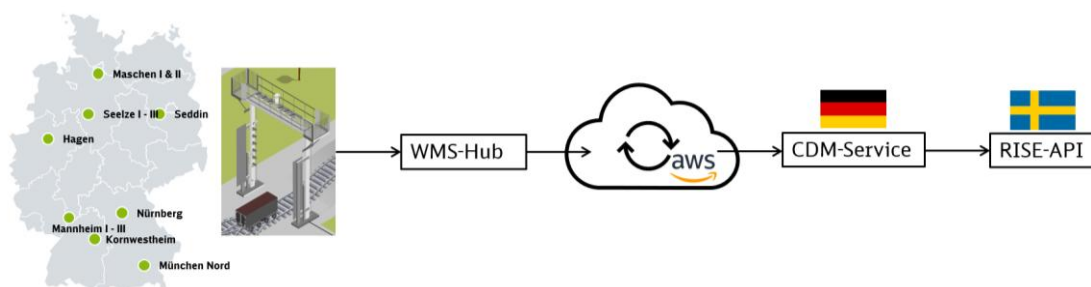


Figure 8.2 Dataflow from German IVG to Swedish API endpoint

8.2 Conceptual Data Model (CDM)

The data from the German backend system was transformed into JSON-formatted data using the developed minimal conceptual data model, as described in Table 15, 16 and 17. The following example (see Figure 8.3) illustrates an exemplary wagon, alongside the relevant information regarding AI-detected damage and human annotations. In this test, the data relates to a single wagon only. This is because the source system of the German IVGs only provides data for single wagons and the train information is not yet available when the wagon passes through the IVG.

```
{
  "wagon_number": "378049930613",
  "timestamp": "2025-08-01T00:00:32.000Z",
  "location": "urn:epc:id:sgln:7332862.00211.1",
  "train_id": "0080202508010456",
  "vehicle_position": 1,
  "wagon_side": "left",
  "view": "bogie",
  "img_url": "https://.../20250801000017_1337.jpg",
  "damages": [
    {
      "error_code": "3.2.2",
      "source": "ai",
      "confidence": 0.86,
      "bounding_box": { "x_min": 1053, "y_min": 532, "width": 323, "height": 254 }
    },
    {
      "error_code": "2.1.5",
      "source": "human",
      "bounding_box": { "x_min": 1400, "y_min": 600, "width": 440, "height": 376 }
    }
  ],
  "un_numbers": [
    "1088",
    "1089",
    "1090",
    "1091"
  ],
  "loading_units": [
    "BICU1234567",
    "BICU1234568"
  ]
}
```

Figure 8.3 JSON example- data exchange of a wagon.

Table 15 High-Level object description of data types.

Field	Type	Req.	Description
wagon_number	string	Yes	12-digit UIC wagon number.
timestamp	string (ISO 8601 UTC)	Yes	Detection time. Use Z suffix for UTC.
location	string	Yes	An ID for the gate and possibly for the sensor that registered the information (for example, a URI following the standard for Global Location Numbers in https://ref.gs1.org/standards/rail/1.1.0/application-standard-gs1-epcis-for-rail-vehicle-visibility-Release1-0-1).
train_id	string	Optional	Train ID (TAF/TAP TSI or some other identifier of the train).
vehicle_position	integer	Optional	1-based index of wagon within train consist (direction of movement).
wagon_side	string enum: left, right	Optional	Side inspected; omit if unknown / not applicable.
view	string	Optional	Camera perspective (e.g. bogie, side, top).
img_url	string (URL)	Optional	Image showing detected damage(s).
damages	array of DamageObject	Optional	List of detected / reported damages. Empty or omit if none.
loading_units	array of strings	Optional	ILU-Codes (EN 13044), identical structure to BIC-Codes. Alternatively: any other loading unit IDs, or null for each loading unit without (known) ID.
un_numbers	array of strings	Optional	4-digit UN numbers for hazardous goods.

Table 16 Description of DamageObject class.

Field	Type	Req.	Description
error_code	String	Optional	Classification code (GCU – reference 20250101_GCU_EN_FULL.pdf).
Source	string enum: ai, human	Optional	Origin of detection.
Confidence	number (0–1)	Optional	AI confidence; omit for human source.
bounding_box	BoundingBox	Optional	Pixel coordinates in image reference frame.

Table 17 Description of BoundingBox object class.

Field	Type	Req.	Description
x_min	integer	Yes	X of top-left corner (pixels).
y_min	integer	Yes	Y of top-left corner (pixels).
Width	integer	Yes	Width (pixels).
Height	integer	Yes	Height (pixels).

8.3 Examples of Exchanged Data

The following three examples demonstrate the type of data for AI, human and logistics-related information that the CDM uses.

8.3.1 AI-detected damage

The results of the Bogie Analyser were packaged into the minimal CDM format and sent to the API endpoint. This data included information on the location and code of the damage, the bounding box, and the level of confidence in the detection (see Figure 8.4).

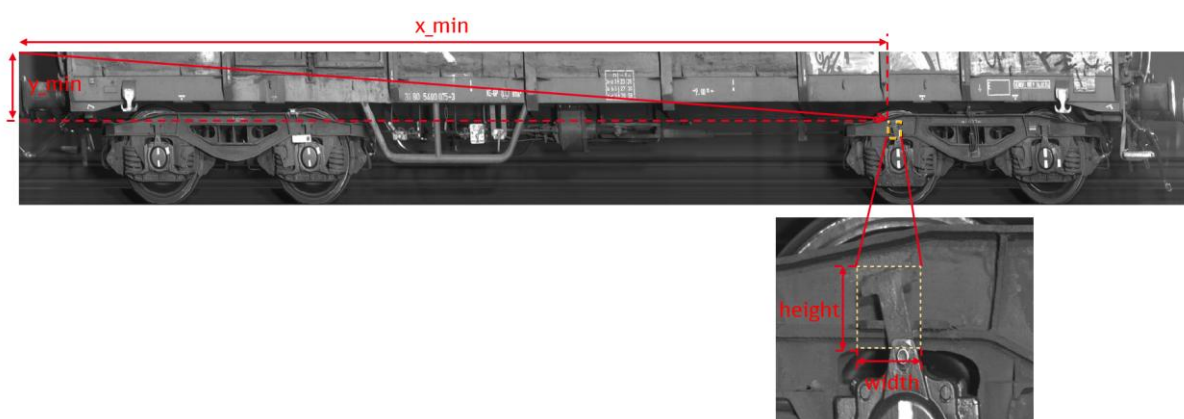


Figure 8.4 Location of the damage and bounding box size in the images.

8.3.2 Human annotation

Employees can use the DB Cargo front end (BISON) to view images of the IVGs and highlight any damage. In the example below (Figure 8.5), a human operator has marked a damaged metal floor and sent the data to the Swedish RISE API endpoint. One image can contain multiple annotations if the operator has identified more than one instance of damage. These annotations contain the damage code, as defined in the 'General Contract of Use for Wagons' (GCU, 2025), as well as the coordinates of the relevant finding and the width and height of the damage. Some of these are not critical but are still maintenance-related and can represent valuable information for owner or maintainer of the asset/wagon.

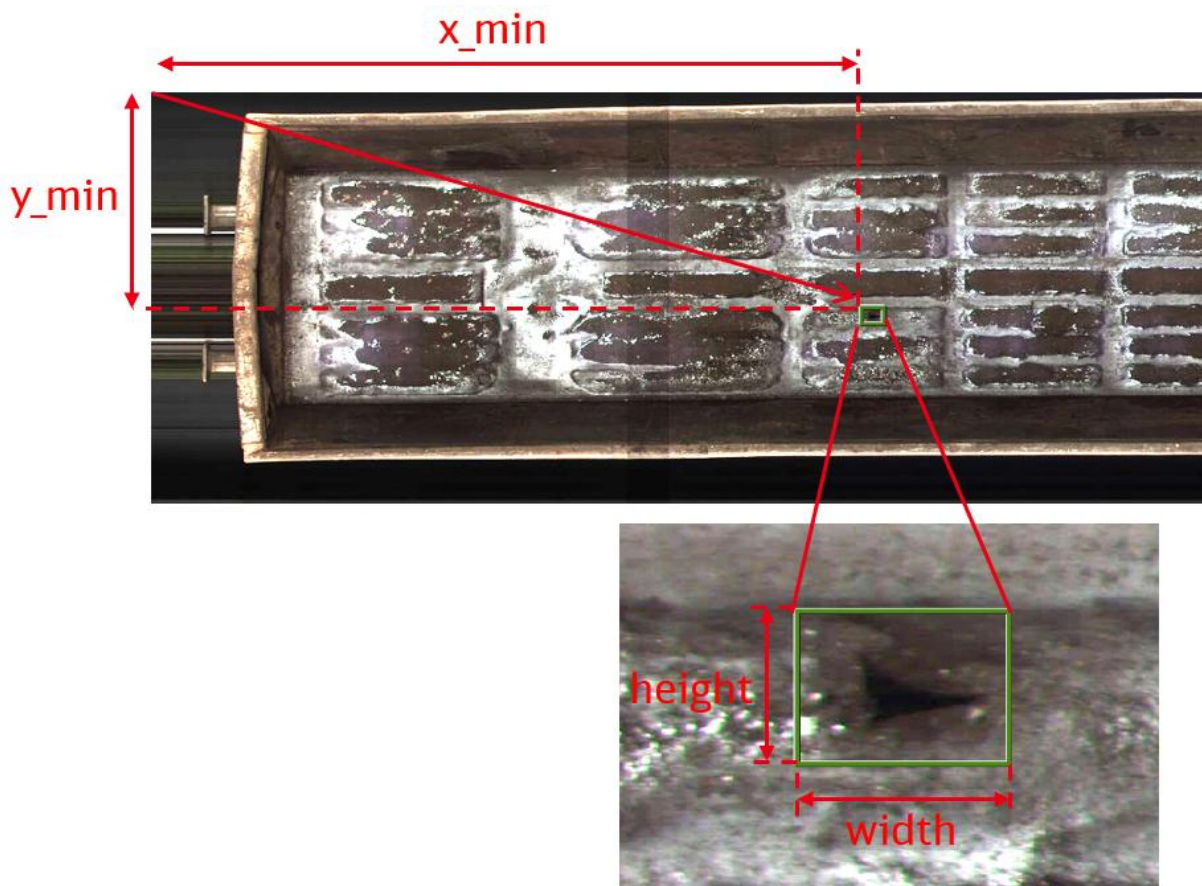


Figure 8.5 Damage in a metal floor with location in the image.

8.3.3 Logistics-related information

Logistics-related information may contain UN numbers (Figure 8.6) and/or ILU (Intermodal Loading Unit) codes. One wagon can carry multiple ILUs and therefore be associated with several ILU codes, see example in Figure 8.7. Similarly, one wagon can be associated with several UN numbers. This information is provided as lists, since position is not required.



Figure 8.6 Orange-coloured plate with UN number on a wagon.

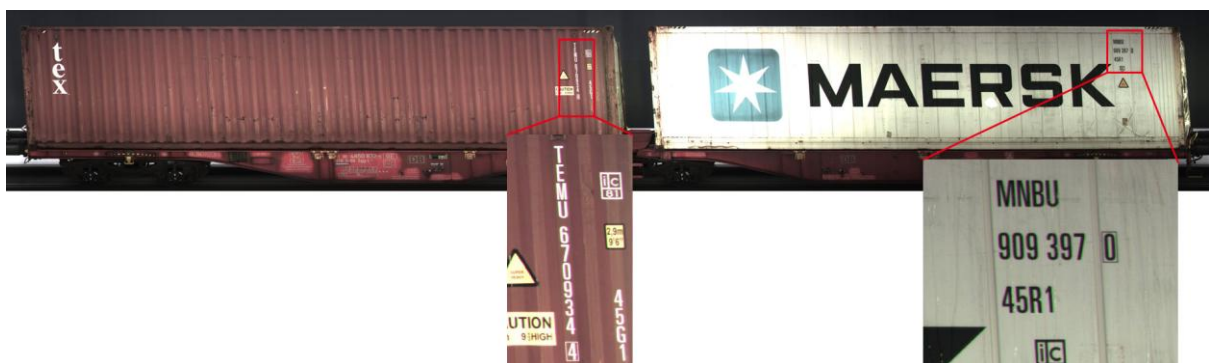


Figure 8.7 Intermodal Loading Unit codes on two containers on a single wagon.

8.4 Data Usage

The research platform Deplide can receive data captured by checkpoints from the systems of both Indra and DB. Deplide is based on the event streaming platform Apache Kafka and organises data into so-called topics, each of which can be described as a channel for a certain type of data. Apart from making the data available to stakeholders and their systems, Deplide can combine and reformat the data in different ways.

One way of making data handled by Deplide available to stakeholders is to present them in graphical user interfaces. Such an interface, called Tåginfo (Swedish for *Train Info*) is shown in Figure 8.8. There, information about wagons and trains that has been captured by checkpoints in Germany and Sweden is presented in a table, along with a map that shows the locations of the checkpoints in question.


Tåginfo

Ej inloggad
Stationsök
Visa favoriter
Hjälp


Sök tåg, spår, destination...
Anpassa

Alla länder
Alla stationer
22 januari 2026
Sök

Visar 21-40 av 100 tåg
Rader per sida: 20
< Föregående 1 2 3 4 5 Nästa >

Välj	Tågnummer	Källa	Föregående	Senaste plats	Senast	Uppdaterad
☐	338027434210	DB Cargo	Maschen2	Hagen	Detected at db_cargo_bison_hagen	2026-01-22T14:26:06
☐	378027702760	DB Cargo	Maschen2	Hagen	Detected at db_cargo_bison_hagen	2026-01-22T14:25:21
☐	238029412190	DB Cargo	Seddin	Mannheim2	Detected at db_cargo_bison_mannheim2	2026-01-22T09:17:41
☐	247442181837	DB Cargo	Seddin	Mannheim1	Detected at db_cargo_bison_mannheim1	2026-01-22T09:11:05
☐	247442202500	DB Cargo	Maschen1	Mannheim1	Detected at db_cargo_bison_mannheim1	2026-01-22T09:11:05
☐	338027430226	DB Cargo	Maschen2	Maschen1	Detected at db_cargo_bison_maschen1	2026-01-22T07:27:11
☐	838027450439	DB Cargo	Maschen2	Hagen	Detected at db_cargo_bison_hagen	2026-01-22T03:48:18
☐	338027434087	DB Cargo	Maschen2	Hagen	Detected at db_cargo_bison_hagen	2026-01-22T03:48:18
☐	338527404176	DB Cargo	Maschen2	Hagen	Detected at db_cargo_bison_hagen	2026-01-22T03:48:18
☐	338027421985	DB Cargo	Maschen2	Hagen	Detected at db_cargo_bison_hagen	2026-01-20T14:33:36
☐	247442206881	DB Cargo	Seelze3	Muenchen	Detected at db_cargo_bison_muenchen	2026-01-20T03:21:51
☐	377202601191949460	INDRA	-	Kville checkpoint	Composition updated at rfid_reader_jvg	2026-01-19T16:49:46
☐	377202601191850190	INDRA	-	Kville checkpoint	Composition updated at rfid_reader_jvg	2026-01-19T15:50:19
☐	377202601191832450	INDRA	-	Kville checkpoint	Composition updated at rfid_reader_jvg	2026-01-19T15:32:45
☐	377202601191725230	INDRA	-	Kville checkpoint	Composition updated at rfid_reader_jvg	2026-01-19T14:25:23
☐	377202601191713060	INDRA	-	Kville checkpoint	Composition updated at rfid_reader_jvg	2026-01-19T14:13:06
☐	377202601191643430	INDRA	-	Kville checkpoint	Composition updated at rfid_reader_jvg	2026-01-19T13:43:43
☐	377202601191558100	INDRA	-	Kville checkpoint	Composition updated at rfid_reader_jvg	2026-01-19T12:58:10
☐	377202601191552480	INDRA	-	Kville checkpoint	Composition updated at rfid_reader_jvg	2026-01-19T12:52:48
☐	377202601191440480	INDRA	-	Kville checkpoint	Composition updated at rfid_reader_jvg	2026-01-19T11:40:48

Visar 21-40 av 100 tåg
Rader per sida: 20
< Föregående 1 2 3 4 5 Nästa >

Tåg 238029412190

Seddin → Mannheim2

Ankomst: 2026-01-22T09:17:41

Spår: -

Operatör: DB Cargo

Land: Tyskland

Visa mer >



Kartförklaring

Tågrutt

Avgångsstation

Ankomststation


For feedback, kontakta railwaycdm.support@ri.se

© 2026 RISE Research Institutes of Sweden.
Under utveckling
Frontend: 5.0.0 • API: 5.0.0 • Consumer: 5.0.0

Figure 8.8 The user interface Tåginfo showing (using blue background colour) information about a wagon that has passed checkpoints in Seddin and Mannheim.

9 Conclusions

In conclusion, the concept of European Railway Checkpoints is believed to represent a step forward in making European railways smarter, more efficient, and more customer-focused. The results from the work with European Railway Checkpoints in WP29 were demonstrated during a webinar, arranged on the 27th of January 2026, 11:30-13:00. In this webinar, developments within this working group were demonstrated by showcasing four modules from Germany, Sweden, Spain and the Netherlands focusing on the following use cases.

- Condition monitoring: Bogie Analyzer
- Rolling stock and ILU identification
- Integrating checkpoints with On-board systems
- Identification of dangerous goods

This document aims to report the demonstration and emphasizes the evaluation part of *Task 29.2: Demonstration and evaluation of the installed checkpoints*. Although the ERC modules are focusing on showcasing different use cases, it should be noted that D25.1 has defined a minimal viable product for the ERC concept, where some functionalities were deemed necessary to include whereas others can be included as possible add-ons.

Regarding **module 1** led by DB Cargo, within this overall framework, the present section summarises the results of multiple validation and evaluation campaigns conducted in the context of camera-based condition monitoring of freight wagons. The work forms part of the ongoing effort to enhance automated inspection processes through the application of AI-based image analysis systems, in particular the Bogie Analyzer. The core challenge addressed in this work was the reliable detection and assessment of safety-relevant wagon components under varying operational and technical conditions. Particular attention was given to differences in camera configurations, image quality and wagon types. The objective was to evaluate the current system performance, identify technical limitations, and determine the feasibility of broader deployment scenarios. The methodology combined several complementary approaches. These included performance evaluations based on image data from different countries, comparison against manual reference measurements (e.g., brake pad thickness measured with measuring device), expert-based human assessments, and statistical performance indicators beyond false positive rates. In addition, coverage analyses were conducted to assess the proportion of workshop entries preceded by a videogate passage. The results demonstrate that the system delivers robust performance in selected use cases, such as brake shoe thickness estimation, lifting T detection, and damper ring evaluation. In particular, the AI-based brake shoe thickness assessment significantly outperformed human estimations, showing a markedly lower average deviation. At the same time, the evaluation revealed several limitations. System performance is influenced by camera installation geometry, viewing angles, image resolution, and overall image quality.

Differences in wagon configurations across countries also impact detection reliability. Certain components may be partially occluded, which constrains achievable recognition rates even after model optimization. Additionally, some use cases (e.g., tare spring direction analysis under non-orthogonal camera perspectives) are currently not suitable for reliable deployment under the existing hardware configuration. Overall, the planned objectives have been achieved. The evaluation provides a transparent and differentiated performance assessment and identifies realistic deployment boundaries. Future work could focus on extended AI retraining using diversified international datasets and optimization for varying camera geometries to ensure consistent cross-site comparability.

Regarding **module 2** led by Trafikverket, two image analysis systems, one developed by Indra and one by Hitachi, have been evaluated for their ability to extract logistics-related information from images produced by the checkpoint in Gothenburg, Sweden. The evaluation shows that neither system currently fulfils the requirement of correctly extracting vehicle numbers in 95 % of all cases. For ILU codes, Indra's system doesn't fulfil the 95 % requirement, while Hitachi's ILU detection capability was not evaluated within this specific assessment campaign, as the dataset used in this study only included output generated by Indra's system.

Regarding **module 3** led by ADIF, the work carried out at the Can Tunis demonstrator highlights the strong potential of integrating future European Railway Checkpoints (ERC) with on-board sensing systems to monitor and track yard movements. The core use case, "Integration of the on-board system", has been successfully tested and validated. The analysis of the results confirms that the requirements defined in earlier phases have been met and demonstrates that the system operates in a robust and reliable manner under real-world yard conditions. Although validation of data exchange with the ERC could not be completed due to ERC installation delays, the current trials demonstrate that this combined approach significantly enhances event detection, timing accuracy, and overall reliability, thereby improving operational efficiency and situational awareness. The immediate next steps focus on completing the installation of the ERCs, carrying out their validation and test their combined operation with the on-board sensors, which is to be reported in D29.6. Looking further ahead, several exploratory lines are being considered, such as refining ERC algorithms and introducing new automated inspection capabilities, including acoustic monitoring and computer-vision-based analysis. For on-board systems, upcoming phases will scale up to full-convoy deployments, finalize digital brake-test calibration, and progress toward active anomaly detection using edge computing. The ultimate milestone will be the real-time synchronization of on-board data with the forthcoming wayside IVG gate infrastructure, enabling the full implementation of the "Beyond the Checkpoint" concept.

Regarding **module 4** led by ProRail, effective solutions were found in the ProRail led tests for robust identification of hazmat number plates, especially the YOLO11x detector with the Idefics2 OCR model (as described in Section 7.4.4). Accuracy levels over 90% are well proven possible. For hazmat placards, the recognition pipeline should be tested on full-frame wagon images to evaluate the complete recognition accuracy. The efficiency and processing time of the full pipeline must be assessed as well on the actual hardware to verify real-time feasibility. Ultimately, the goal is to combine all developed algorithms into a single integrated system capable of handling plate reading, placard classification, UIC recognition, and wagon counting within one unified framework. Future work should focus on moving the system towards full operational integration. Implement wagon counting is a key next step, which requires full access to all raw sensor data so that information from multiple sensors can be combined effectively. To ensure robustness more extensive evaluation under real-world conditions, including varied weather, lighting, and environmental effects, is essential.

In conclusion this document has presented the evaluation of European Railway Checkpoint (ERC) concept and the proposed four modules, as well as suggested methodologies for data exchange and possible paths for implementation and roll-out. Overall, the demonstrators provide evidence of the feasibility of the ERC concept in operational railway environments and contribute to the technological maturation of the involved solutions. Moreover, the outcome of task 29.2 which is covered in this deliverable, is intended to give input to other deliverables in FP5 in particular D1.6 *Final KPI assessment*, D32.1 *Results from Demonstrators*, and D33.2 *Report of Seamless corridor*.

The modules demonstrated in this deliverable incorporate diverse functionalities that are able to achieve improvements from automatic recognition and deriving information, from automatic recognition of UIC codes, dangerous goods plates and ILUs to defects to be inspected. These improvements can be achieved in a rather efficient way through the potential of new scanning techniques and AI techniques. The ERC can thus still growth in importance to manage the fleet of rolling stock efficiently and also to protect the infrastructure and environment against damages caused by defecting rolling stock. Thus, further development is proposed to take place in the second call of FP5 where the aim is to reach TRL8 and have the actual system completed.

References

- Commission Implementing Regulation (EU) No 402/2013 of 30 April 2013 on the common safety method for risk evaluation and assessment and repealing Regulation (EC) No 352/2009 https://eur-lex.europa.eu/eli/reg_impl/2013/402/oj
- DB Fahrwegdienste GmbH. (2021). 93612 | Technische Regeln: Güterwagen bedienen und behandeln (Version 4.0, 01.01.2021, F. Schelzke)
- Deiterding L., Hildebrandt F., Kroth G., Sterbak Y., Hauser F., Gade K. (2023). *D25.2: High-level specification of requirements, challenges and a future target state for freight cross-border planning and operations from an operator perspective as input for FP1*. FP5-TRANS4M-R, Europe's Rail. https://rail-research.europa.eu/wp-content/uploads/2025/02/FP5-TRANS4M-R_D25.2_V1-2023-06-30.pdf
- Garcia-Bordils S., Karatzas D., and Rusinol M. (2023). "STEP – Towards Structured Scene-Text Spotting," Sep. 2023. Available: <http://arxiv.org/abs/2309.02356>.
- GCU (2025). *General Contract of Use for Wagons*. Version 1 January 2025. https://gcubureau.org/wp-content/uploads/Contract/2025/20250101_GCU_EN_FULL.pdf
- Hildebrandt, F., Deiterding, L., Walker, T., Pelegrin, M., Mendizabal Roche, A., Bueso, M., ... Bryne, B. (2024). *Deliverable 25.1 Report on the basic functional and technical specifications for the realisation of the technical enablers of Seamless Freight, also including the final specification input for FP1*. FP5-TRANS4M-R, Europe's Rail. https://rail-research.europa.eu/wp-content/uploads/2025/03/FP5-TRANS4M-R_D29.1_V3.pdf
- Hoffstaetter, S., Lee, M. (2025). Pytesseract · PyPI, <https://pypi.org/project/pytesseract/>, Accessed: June 2025
- Jocher G. (2025) Ultralytics Yolo 11, <https://github.com/ultralytics/ultralytics>, Accessed: June 2025
- Kittinaradorn, R. (2025). Easyocr · PyPI, <https://pypi.org/project/easyocr/>, Accessed: June 2025
- Kordnejad B., Nordmark I., Åkerfeldt M., Bueso M., Mendizábal A., Pelegrin M., Adin I., Losada M., Nielsen J., Klein C., Sengle E., Böhmer L-B., Klass M., Wohlrath R., Vinh-Le T., Neumann T, Salvitti G., Lengu R, Parrilla Ayuso F., de la Iglesia Sánchez R., Díaz Díaz A, Giménez Gómez J., Thiery F., Pum T., Yedema B., Koolstra E., Huergo A., Kjellin M., Tijink M. (2024). *D29.1 Technical definition of the standard IVG for checkpoints and related demonstrators*. FP5-TRANS4M-R, Europe's Rail. https://rail-research.europa.eu/wp-content/uploads/2025/03/FP5-TRANS4M-R_D29.1_V3.pdf

- Kordnejad B., Nordmark I., Thierry F., Nielsen J., Kjellin M., Åkerfeldt M. (2025). *D29.5 Execution of test for gates installed in Sweden*. FP5-TRANS4M-R, Europe's Rail. (submitted)
- Kordnejad B., Nordmark I., Wohlrath R., Klein C. C. (2023). *D29.3 Report on the basic functional and technical specifications regarding CMS as relevant input for FP3*. FP5-TRANS4M-R, Europe's Rail. https://rail-research.europa.eu/wp-content/uploads/2025/02/FP5-TRANS4M-R_D25.3_V1.0.pdf
- Laurençon, H., Tronchon, L., Cord, M., Sanh, V. (2024) *What matters when building vision-language models?* <https://doi.org/10.48550/arXiv.2405.02246>
- Pelegrín Preixens, M., Mendizábal Roche, A., Vidaurre Celis, A. A., de la Iglesia Sánchez, R., Parrilla Ayuso, F., Losada Gobantes, M., Adin Marcos, I., Huergo Luz, A. (2025). *D29.3 Installation and test plan for Gates in Spain*. FP5-TRANS4M-R, Europe's Rail. <https://rail-research.europa.eu/wp-content/uploads/2026/01/D29-3-Installation-and-test-plan-for-Gates-in-Spain.pdf>
- ProRail (2026), Network Statement 2024/2025 Article 58 of the Railway Act. <https://www.prorail.nl/samenwerken/vervoerders/network-statement>, Accessed: Feb. 2026
- Ren, S., He, K., Girshick, R., Sun, J. (2015). *Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks*. IEEE Transactions on Pattern Analysis and Machine Intelligence 39(6), 1137–1149 (6 2015)
- Sengle E., Margarite Klass M., Pi Schwarz P., Hedtke M. (2025). *D29.8 Live-Demo of Video Gates showing process optimization in a German yard*. FP5-TRANS4M-R, Europe's Rail.
- Veld D., Tijink M. (2025). *D29.4 Installation and test plan for Gates in the Netherlands*. FP5-TRANS4M-R, Europe's Rail. (submitted)
- Veld D., Tijink M. (2026). *D29.7 Execution of test for gates installed in the Netherlands*. FP5-TRANS4M-R, Europe's Rail. (submitted)