Ref. Ares(2023)4558196 - 30/06/2023

Deliverable D 2.4

WP2 Report on identification of future innovation needs and recommendations for improvements

Project acronym:	RAILS
Starting date:	01/12/2019
Duration (in months):	43
Call (part) identifier:	H2020-S2R-OC-IPX-01-2019
Grant agreement no:	881782
Due date of deliverable:	Month 40
Actual submission date:	June 30^{th} 2023
Responsible/Author:	Stefania Santini (CINI)
Dissemination level:	Public
Status:	Issued

Reviewed: no

Document history		
Revision	*Date*	*Description*
1	June 15^{th} 2023	First issue for internal review
2	June 30^{th} 2023	Second issue after internal review

Report contributors		
Name	Beneficiary Short Name	Details of contribution
Stefania Santini	CINI	WP Leader
Lorenzo De Donato	CINI	Contributor
Francesco Flammini	LNU	Internal Reviewer

Funding

This project has received funding from the Shift2Rail Joint Undertaking under the European Union's Horizon 2020 research and innovation programme under grant agreement n. 881782 Rails. The JU receives support from the European Union's Horizon 2020 research and innovation program and the Shift2Rail JU members other than the Union.

Disclaimer

The information and views set out in this document are those of the author(s) and do not necessarily reflect the official opinion of Shift2Rail Joint Undertaking. The JU does not guarantee the accuracy of the data included in this document. Neither the JU nor any person acting on the JU's behalf may be held responsible for the use which may be made of the information contained therein.

Contents

Executive Summary

This deliverable contains a critical examination of the work and the results obtained in WP2, also against the current state-of-the-art in railways. From the Proofs-of-Concept (PoCs) and the experience gained by developing the tasks of the project, this document reports lessons learned, weaknesses and strengths shown by each exploited technology, technical and implementation recommendations, unaddressed issues, and innovation needs. Specifically, an in-depth analysis of the methodological and experimental PoCs carried out for the two selected case studies proposed in the previous WP2 deliverables, namely, "Obstacle Detection for Collision Avoidance" and "Cooperative Driving for Virtual Coupling of Autonomous Trains", are provided. This analysis, arranged according to the SWOT structure, is going to highlight the main strengths (S) and weaknesses (W) shown by the proposed AI approaches, and to identify external opportunities (O) that could support its technical feasibility, as well as some open challenges related to the main threats (T) emerged from each PoC. Starting from this evidence, some specific recommendations are addressed regarding the effectiveness of the proposed AI approaches. Furthermore, general recommendations are drawn, including suggestions for future development, experimentation, and applications for the integration of AI in the rail sector.

Abbreviations and acronyms

Abbreviations / Acronyms	Description
AI	Artificial Intelligence
ADM	Anomaly Detection Module
ATO	Automatic Train Operation
ATP	Automatic Train Protection
CCTV	Closed Circuit Television
DDPG	Deep Deterministic Policy Gradient
DL	Deep Learning
DNN	Deep Neural Network
DRL	Deep Reinforcement Learning
DT	Digital Twin
ETCS	European Train Control System
GoA	Grade of Automation
GoI	Grade of Intelligence
IoT	Internet of Things
ITC	Intelligent Train Control
ITO	Intelligent Train Operation
ITP	Intelligent Train Protection
KPI	Key Performance Indicator
LoI	Level of Intelligence
ODM	Object Detection Module
PoC	Proof-of-Concept
RBC	Radio Block Center
RDM	Rails Detection Module
SIL	Safety Integrity Level
SOSS	Safe Operational State Space
SWOT	Strengths, Weaknesses, Opportunities, and Threats
T2I	Train-to-Infrastructure
T2T	Train-to-Train
T2X	Train-to-Everything
VC	Virtual Coupling
VBODS	Vision-Based Obstacle Detection System
VCTS	Virtually Coupled Train Set
XRL	Explainable Reinforcement Learning

1. Objective

This document aims to draw some conclusions from the work carried out in WP1 and WP2, helping to clearly identify possible future innovations, research directions, and impacts for the European railway sector. These objectives are strictly related to the work addressed in the previous WP2 deliverables, in which methodological and experimental Proofs-of-Concept (PoCs) have been carried out for two selected case studies, namely, "Obstacle Detection for Collision Avoidance"and "Cooperative Driving for Virtual Coupling of Autonomous Train".

On the basis of the results that emerged from the PoCs, the main goal of this deliverable is to provide some recommendations and innovation needs which could support the effective adoption of AI in the rail sector. To this aim, the present document focuses on the following objectives:

- a brief overview of the recent advancements in the field of the related PoCs;
- an in-depth analysis of the proposed PoCs, arranged according to the SWOT structure, to highlight the main strengths (S) and weaknesses (W) shown by the proposed AI approaches, and to identify external opportunities (O) and threats (T) which could affect their technical feasibility;
- the identification of some specific recommendations which could support the effectiveness of the proposed AI approaches;
- the drawing of general recommendations regarding the integration of AI for rail safety and automation, mainly encompassing indications for further investigation, including:
  - further development of approaches, methods, models, technologies, and tools;
  - further experimentation with additional data, case studies, pilot studies, and scenarios;
  - applications to other areas and subsystems within the railway transport sector.

The set of recommendations will be used in WP5 to identify migration strategies and roadmaps for AI integration in the rail sector.

2. Introduction

This deliverable provides a critical examination of the work and the results obtained in WP2 also against the current state-of-the-art in railways. It reports some technical/implementation recommendations and innovation needs that would support future investigations in the context of AI for rail safety and automation.

The recommendations provided by this document can be subdivided into two main macro-categories:

1. Recommendations coming from the critical examination of the Proofs-of-Concept (PoCs) developed within the RAILS's WP2 which would be potentially useful to support future development of approaches, methods, models, technologies, and tools in the specific contexts of the PoCs and related areas. Specifically, the work conducted for the Obstacle Detection scenario can be reused for similar purposes in other perception-based railway applications, and the approach taken for the Virtual Coupling scenario can inspire other control solutions.
2. General Recommendations, coming from lessons learned while both working at the PoCs and investigating the state-of-the-art of AI in railways, which aim at providing hints about practices and activities that would support the integration of AI across various railway applications.

The remainder of this deliverable is organised as follows. Chapter 3 summarises the findings of WP1 about the state-of-the-art and promising research directions in railway safety and automation, as well as the documents and results produced during the project and addressing the topics investigated in the context of WP2, including the related scientific publications stemming from the project activities. Therefore, Chapter 3 provides the context and the background of the discussion reported in the present deliverable. Chapter 4 and Chapter 5 respectively address critical examinations of the PoCs on "Obstacle Detection for Collision Avoidance" and on "Cooperative Driving for Virtual Coupling of Autonomous Trains". Chapter 4 and Chapter 5 share the same structure: Sections 4.1 and 5.1 discuss a high-level overview of the recent advancements in the context of the corresponding PoCs; Sections 4.2 and 5.2 present a bird-eye view of the investigative approaches; Sections 4.3 and 5.3 propose structured analyses of the implemented approaches in the form of SWOT (strengths, weaknesses, opportunities, and threats) analyses; lastly, Section 4.4 and 5.4 highlight the main recommendations resulted from the lessons learned while working at the PoCs. Then, Chapter 6 discusses the general recommendations and some innovation needs that would be required for the fast take-up of AI in railways. Lastly, Chapter 7 provides some concluding remarks.

3. Background

This section recalls most of the findings from the analyses carried out in the previous phases of the RAILS project with specific emphasis on AI applications for “Railway Safety and Automation”. To better set the document background, Table 3.1 reports all the manuscripts (deliverables and papers) resulting from the aforementioned research activities and specifies the main contributions/results.

Table 3.1: Published Documents discussing AI for “Railway Safety and Automation”.

Focus	Document	Type	Main contribution(s)
Taxonomy	Deliverable D1.1: Definition of a Reference Taxonomy of AI in Railways [1]	PD	1. Delineation of a definition for AI in railway 2. Establishment of a taxonomy of AI in railway 3. A first overview of regulations for AI 4. Identification of Railway Subdomains 5. Preliminary mapping of existing AI applications on Railway Subdomains
	Artificial Intelligence in Railway Transport: Taxonomy, Regulations, and Applications [2]	SP	
State of the Art	Deliverable D1.2: Summary of existing relevant projects and State-of-the-Art of AI application in railways [3]	PD	1. Review of projects conducted worldwide (with emphasis on S2R projects) dealing with AI in Railway Subdomains 2. Review of scientific papers dealing with AI in Railway Subdomains 3. Preliminary definition of future direction towards the integration of AI
	A Literature Review of Artificial Intelligence Applications in Railway Systems [4]	SP	Extended review of scientific papers dealing with AI in Railway Subdomains
	A Survey on Audio-Video Based Defect Detection Through Deep Learning in Railway Maintenance [5]	SP	In-depth review of scientific papers dealing with AI for M&I applications exploiting audio-video data
	A Systematic Review of Artificial Intelligence Public Datasets for Railway Applications [6]	SP	In-depth review of publicly available datasets for each Railway Subdomain
Application Areas	Deliverable 1.3: Application Areas [7]	PD	1. Identification of relevant railway Application Areas for AI together with the main challenges to be tackled for its effective integration basing on: i) the review of projects conducted worldwide and the scientific literature dealing with AI in railways; ii) suggestions from the Advisory Board; and iii) the results from a comprehensive Survey submitted to Researchers and Practitioners from different organisations operating worldwide 2. Delineation of basic AI usage guidelines to select the most appropriate AI approach by taking into account: i) the goal; ii) the type of available data; and iii) the required responsiveness of the AI system
Transferability	Deliverable 2.1: WP3 Report on Case Studies and Analysis of Transferability from Other Sectors [8]	PD	1. Review of AI-based emerging technologies developed in (transport and non-transport) sectors other than railways 2. Identification of AI approaches that can be transferred to or adapted for railway applications
Intelligent Train Control	A Vision of Intelligent Train Control [9]	SP	1. Definition of a vision for the introduction of AI in Train Control Systems 2. Preliminary identification of Grades of Intelligence (GoIs) 3. Preliminary identification of Levels of Intelligence (LoIs)
PoCs Development	Deliverable 2.2: WP2 Report on AI approaches and models [10]	PD	Identification of PoCs to be developed together with research questions, methodology, reference datasets, AI and ML models, and expected results
	Deliverable 2.3: WP2 Report on experimentation, analysis, and discussion of results [11]	PD	Development of identified PoCs including model/architecture description, data generation, training and validation, and evaluation and discussion of results
	Roadmap and Challenges for Reinforcement Learning Control in railway Virtual Coupling [12]	SP	1. Investigation of the main challenges related to the feasibility of Railway VC through AI techniques 2. Identification of potential AI methodologies to be transferred from automotive to railway 3. Definition of a methodological PoC based on a Reinforcement Learning control strategy and discussion on expected results
Recommendations	Deliverable 2.4: WP2 Report on identification of future innovation needs and recommendations for improvements [this document]	PD	1. Identification of sectorialized recommendations oriented at supporting AI integration in WP2 PoCs' topic and related areas 2. Identification of general recommendations aiming at supporting AI integration across different railway applications

PD: Project Deliverable; SP: Scientific Paper

Herein, we recall some of the results obtained within the first phase of the project as they impacted the subsequent investigation carried out in WP2. First, a *Taxonomy of AI in Railway* has been introduced and seven *Railway Subdomains* have been identified for the RAILS investigation [1, 2]. Then, in Deliverable D1.2, for each of the subdomains, a review of the State-of-the-Art (SOTA) of AI in railways has been developed by analysing i) research projects conducted worldwide (with a particular focus on S2R projects) and ii) the scientific literature. The latter analysis has been then extended in [4]. The main statistical findings we deducted from the SOTA are reported in Fig. 3.1.

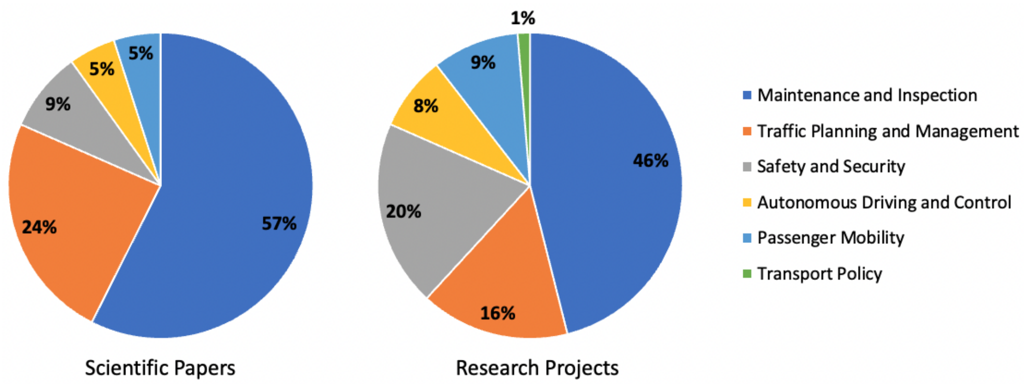

Fig. 3.1. Distribution of Scientific Papers and Research Projects per Railway Subdomain.

Notably, most of the contributions were related to the “Maintenance and Inspection” subdomain, while those referring to the “Railway Safety and Automation” macro area – i.e., “Safety and Security” (S&S) and “Autonomous Driving and Control” (ADC) – presented a net subdivision. As for scientific papers, the two subdomains of S&S and ADC resulted to receive limited attention; conversely, research projects seemed to be quite sensitive to the topic of S&S, but still limited contributions were found for ADC. Among the most investigated topics, we had: i) *Risk Assessment*, which encompasses all the contributions dealing with “the process of planning, preparing, performing, and reporting a risk analysis, and evaluating the results against risk acceptance criteria” [13]; ii) *Disruption and Anomaly Detection*, including the approaches oriented at identifying or predicting possible disruption and anomalies that could harm the safety of railway lines; iii) *Railway Accidents*, involving the studies related to the prediction and the discovery of hidden patterns among hazardous events; and iv) *Energy-Efficient Driving*, encompassing the contributions oriented at optimising energy consumption.

In addition to this analysis, we also conducted further investigations in Deliverable D1.3 with the aim of identifying the **Application Areas** *that could benefit the most from the implementation of AI techniques.* Among others, some of the Application Areas we identified for “Railway Safety and Automation” are reported herein:

- **Environment Monitoring**, including solutions giving trains the capability of analysing the surroundings;
- **Anomaly Detection**, encompassing approaches oriented at detecting anomalies in railway components (e.g., trains’ on-board components) in order to promptly adopt countermeasures and avoid potential threats;
- **Obstacle Detection and Avoidance**, involving systems aiming at detecting obstacles (e.g., on rail tracks) that could interfere with the safe operability of trains;

- **Energy Optimisation**, including approaches oriented at identifying adequate trains' speed profiles in order to reduce energy consumption as much as possible;
- **Smart Signalling** aiming to control railway traffic safely and prevent trains from colliding by leveraging modern technologies, such as cloud computing, wireless communication, location-based services, and computer vision.
- **Adaptive Automatic Train Operations**, dealing with the introduction of adaptive behaviours (based on learning capabilities) within the current ATO systems;
- **Verification and Validation (*V&V*)**, encompassing procedures for the safety assessment of AI-aided railway systems. Due to the possibly unpredictable nature of some AI approaches, their use creates concerns that need to be faced using appropriate *V&V* processes to guarantee trustworthy AI and safe autonomy.

These and other Application Areas, together with related research directions, are discussed in Deliverable D1.3. Deliverable D2.1 reports approaches and results from other sectors, including challenges in assessing ML systems. Then, the research activities in WP2 focused on autonomous train driving and control, and in particular addressed intelligent control, obstacle detection, and cooperative driving. Details about the specific PoCs we developed are given in Deliverables D2.2 and D2.3.

The investigations conducted within the aforementioned deliverables and all the other documents reported in Table 3.1 converged into the definition of the practical recommendations that are discussed in the following chapters of this deliverable.

4. A Critical Examination of the Obstacle Detection Proof-of-Concept

As described in our previous Deliverable [11], in order to move towards *full autonomy*, trains should be equipped with adequate sets of sensors and systems giving them the capability of achieving *full situation awareness* in relation to the health status of the on-board components and external threats or signals. With this PoC, we mainly focused on the latter aspect and, to be specific, we investigated the possibility of adopting Deep Learning algorithms in combination with data from a RGB camera to detect obstacles on rail tracks. Important to underline, the aim of this PoC was not to propose the most suitable and effective solution for this specific task, instead, we focused on investigating the opportunities that DL approaches could introduce when realising Vision-Based Obstacle Detection Systems (VBODSs) to detect obstacles both on rail tracks and in other railway scenarios.

4.1. Recent Advancements on Obstacle Detection

This section provides a brief overview of the latest developments concerning *Obstacle Detection on Rail Tracks* and concisely highlights the main shortcoming(s) of the VBODSs discussed so far within the literature (to the best of our knowledge).

In the last few years, different obstacle detection systems have been investigated which typically leverage multiple sensors including, among others, LiDARs, radars, and cameras to be able to detect obstacles at different distances. In addition, as discussed within the SMART [14] and SMART2 [15] projects, such systems may also involve multiple sub-systems – involving onboard, trackside, and airborne sub-systems – to monitor the environment from different perspectives. In these contexts, and if AI is involved, the usage of cameras is mainly oriented (yet not limited) to the classification of the type of obstacles, while their localisation within the environment is demanded to the other sensors. As widely analysed within, in our opinion, one of the most comprehensive literature reviews on vision-based obstacle detection applications [16], camera-based systems can be used to both locate and identify (i.e., detect) objects, however, to the best of our knowledge, they mainly rely on *Supervised approaches*. Briefly, these kinds of AI models are trained to detect a set of elements that are specified in advance; for example, a Deep Neural Network (DNN) can be trained to detect some pre-specified obstacles such as cars, pedestrians, some species of animals, and so on. The main problem with these approaches is that if an unknown obstacle[1] has to be detected, the DNN will most likely fail.

4.2. Bird-eye View of the PoC Approach

With this PoC, we aimed at understanding to what extent it would have been possible to overcome this *coverage* issue by leveraging *a single RGB camera*, in order to discuss opportunities and shortcomings introduced by the simplest (and potentially cheapest) system that is possible to implement.

Therefore, we studied the possibility of implementing a *cheaper (or supportive) alternative*, compared to the systems introduced above, *which exploits AI, artificial vision, and data com-*

[1] An obstacle that has not been taken into account when training the DNN.

ing from a single camera mounted in front of the train, with the specific aim of investigating *Unsupervised* DL approaches that would potentially allow VBODSs to detect any obstacle (and not just those specified *a priori*). To that aim, as deeply described in [11], we studied a multi-modular architecture (reported in Fig. 4.1) which includes the following modules:

- **Rails Detection Module (RDM):** this module is oriented at extracting rail tracks from the input frame. We first implemented a semi-automatic labelling approach based on Self-Training and Transfer Learning to label data we collected, then, we implemented a semantic segmentation approach based on U-Net [17] to achieve the aforementioned goal.
- **Object Detection Module (ODM):** the purpose of this module is to detect objects (i.e., obstacles known a-priori) laying on the rail tracks. This module would most likely involve a DNN trained in Supervised mode as described above. Given the large attention this topic has received in the last years, also specifically within the rail sector [16], we did not implement this module.
- **Anomaly Detection Module (ADM):** this module is oriented at detecting potentially any kind of obstacle whether known or unknown a-priori. We refer to obstacles which cannot be pre-specified as anomalies. In order to realise this module, we adopted an anomaly detection approach. Practically, inspired by [18] and [19], we build a DNN (named SSIM-VQ-VAE) and we trained it in an unsupervised manner on obstacles-free data. Eventually, we obtained a DNN capable of producing an anomaly map which highlights the anomaly (if present).
- **Obstacle Detection Module:** once the ODM will be implemented, this module will simply merge the outputs from the ODM and the ADM, then, it will produce a map that identifies the anomalous regions or identifies and classifies the objects if known a-priori. Most likely, this module would not involve any AI functionality.
- **Distance Estimation Module:** the purpose of this module would be to estimate the distance of the obstacles (i.e., both objects and anomalies) from the train. Worth mentioning, this task has recently been analysed within the SMART project (DisNet [20]) which seems to work properly when it comes to detecting the distance of the objects (for which it is possible to know the dimensions in different circumstances). Differently, as far as we know, further investigations would be required to estimate anomalies' distances.

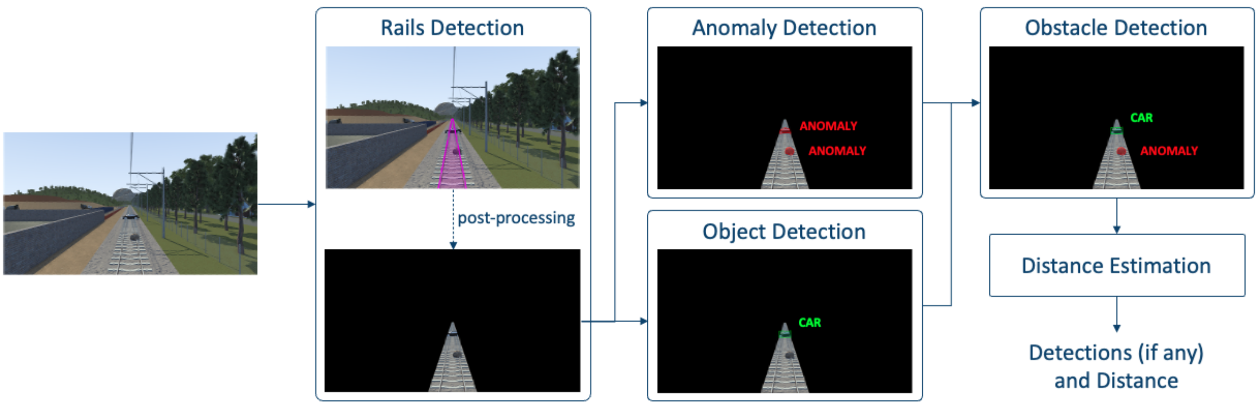

Fig. 4.1. An Architecture for Vision-Based Obstacle Detection.

4.3. A SWOT Analysis of the PoC

In this section, we highlight the inner Strengths (S) and the Weaknesses (W) of the approach we investigated and identified some external Opportunities (O) and Threats (T) that cloud respectively support or challenge its technical implementation. Important to mention, this SWOT analysis is not oriented at defining the market viability of the solution we investigated, instead, it has been used to evaluate the approach in a structured manner. Fig. 4.2 reports all the aspects that we found to be relevant, based on our investigation and tests, arranged according to the SWOT structure.

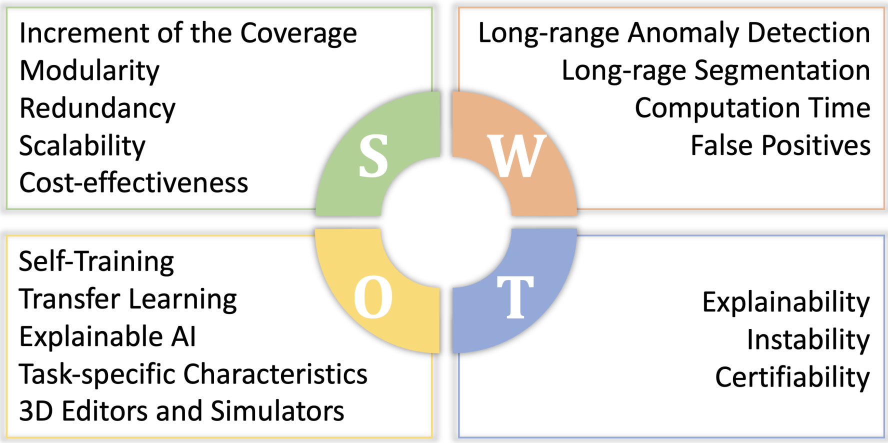

Fig. 4.2. Relevant aspects from the Obstacle Detection PoC arranged according to the SWOT Structure.

4.3.1. Strengths and Weaknesses

Increment of the Coverage. The architecture we investigated is oriented at *increase the coverage of VBODSs* which typically rely on supervised approaches and are then capable of detecting only a pre-specified set of objects.

Modularity, Redundancy, and Scalability. The *modularity* of the architecture allows the application of this idea also to already existing Object Detection systems without compromising their functioning. Indeed, assuming that another Object Detection system exists (and is, to some extent, “reliable”) which detects potential obstacles within the whole video frame and then applies post-processing to understand whether they are on the track or not, the pipeline “RDM $\rightarrow$ ADM” can be simply applied in parallel without interfering with the aforementioned system. The modularity also introduces a sort of *redundancy*. Indeed, for each known obstacle, we would potentially have two detections: one from the ODM and the other from the ADM. Therefore, given that the two modules adopt different DL models (which are either trained on different datasets or process data in two different ways) it is quite unlikely that both would fail at the same time. Redundancy can also be improved by *scaling* the sensors. In our analyses, we considered data coming from a *single camera*, however, it would be possible to use *multiple cameras* and always the same architecture and DL models. For example, assuming a sensing system with three cameras (each of which monitors the rail track from a different angle), three VBODSs could be implemented, each of which will be trained on data coming from only one of the aforementioned cameras. Therefore, at each instant, we could have three predictions that can be merged to i) filter out possible prediction noises and ii) avoid possible miss-detections.

Cost-effectiveness and Potential Applications. As pointed out at the beginning of this chapter, the main intent of this PoC was to analyse the potential of DL for vision-based obstacle detection, and not to provide the best possible solution for any kind of trains or railway lines (e.g., freight/passenger trains, high-speed railway lines, etc.). Although, based on our tests, *anomalies can be detected only within a range of about 70 meters*, which makes the ADM not suitable for high-speed railway lines or heavy freight trains since their braking distance is higher than that values, there are other scenarios where the investigated approach can potentially bring benefits and help to *improve safety or automation in a cost-effective manner*. For example, the *one-camera* solution can be adopted to assist the driver or to automatise those railway lines where trains run at reduced speed (e.g., regional or urban lines). In addition, it could also be used in main lines as low-speed complement and/or to assist the driver during procedures such as the "Track Ahead Free" or to integrate a sort of "return to home" functionality [9]; i.e., in case of malfunctions would not allow trains to safely operate at their maximum efficiency, a one-camera system could support the train to autonomously and safely reach the nearest station by proceeding at reduced speed. In addition, such an architecture can also be exploited to improve safety at Level Crossings (LCs), which is one of the most sensitive railway assets; some LCs are already equipped with CCTV cameras for security purposes, and these could be potentially exploited also to implement the obstacle detection approach we propose so that it would potentially possible to save costs due to sensors installation.

Anomaly Detection Distance and Computation Time. Diving into more technical aspects, the pipeline we implemented in Deliverable [11] encompasses only two of the models that should compose the whole VBODS: the RDM and the ADM. According to our results, as mentioned above, *anomalies can be detected within a range of about 70 meters* and the *computation time required to process a single frame is about 476 ms*. The former may not be a shortcoming, depending on the specific application the architecture would be adopted for, however, actions are required to reduce the computation time to meet *real-time constraints*. The pipeline encompasses two DL models and post-processing algorithms to adjust the outcome of the DL models; these algorithms are those that contribute the most in terms of computation time, especially the one that post-processes the RDM outcome which, alone, takes about 363 ms. Therefore, further investigations would be required to optimise these processing algorithms and, hopefully, improve the processing speed of the DL models.

Long-rage Rail Track Segmentation. The RDM properly segments rail tracks within a medium-long range, however, *there are cases in which the last part of the track, i.e., that one that is the most far from the train, is not properly extracted.* The causes leading to this phenomenon have already been identified in Deliverable [11] together with some possible actions to overcome this issue. However, this is a crucial aspect to investigate as, given the architecture we identified, if tracks are not properly extracted, it would not be possible to identify anomalies or objects.

False Positives. The *ADM resulted not to be enough robust to false positives*. In our tests, we considered a rail track section with a tree projecting its shadow on the rail tracks which is detected by the ADM as an anomaly. Issues of this kind may lead to unnecessary slowdowns which will reflect in a reduction of the capacity of the lines and delays; therefore, further investigations would be required to understand if such an issue can be somehow mitigated.

4.3.2. Opportunities and Threats

Domain-specific Characteristics. One of the main issues we encountered during the development of this PoC was the *lack of suitable datasets* to train and test DL algorithms. They exist some datasets depicting railway scenes (e.g., RailSem19[2], we actually used to pre-train our RDM) but, as far as we know, the lack of some *domain-specific characteristics* that could potentially help DL models to perform better. As explained in detail in Deliverable [11], the structure of railway lines does not change over time and this *invariance* can be exploited to build specific datasets (one for each railway line or one for each train path) to train different DL models each of which is specialised in working on a specific railway line. Practically, instead of building a single DL model, "overcharging" it with data coming from different railway lines so that it would (probably) be capable of working in any railway scenario, in our view, it would be better to build, for each railway line, a different dataset. Then, multiple VBODS can be implemented (without varying the architecture or the adopted DL models), each of which is trained on a specific dataset collected from a specific railway line. *The main advantage is that the generation of different datasets (one for each railway line) could help to reduce the structural aleatory the DL models should face so that the aleatory is "limited" to weather and light conditions and possible intrusions*. This aspect also leads to another advantage: if the system focuses on a specific railway line and has been trained with multiple videos depicting that line under different light and weather conditions, *it would be highly probable that, at run-time, the VBODS would analyse a scenario that is very similar (but, of course, not identical) to something it has already seen during the training phase*.

3D Editors and Simulators. Given what has been explained above, we leveraged the MathWorks' 3D editor RoadRunner[3] to build a railway scenario depicting a specific railway line from which to capture videos to train and test AI algorithms. In general, *3D editors (or simulators)* of this kind can bring huge benefit when it comes to generating data for the development of PoCs for two main reasons: first, the latest technological developments have given to these software (e.g., Unity[4] or Unreal Engine[5]) the capability of *properly approximating the reality in terms of graphics, weather/light conditions, and simulation of the behaviour of the assets*; second, they allow to *safely generate anomalous events* (e.g., obstacles on the tracks). Clearly, even though these approaches can be extremely useful to develop PoCs, experiment DL models, and define suitable reasoning to address the problem, to evaluate the real effectiveness of the implemented architecture, tests with real data should be performed.

Self-Training and Transfer Learning. Another implementability challenge that we encountered was *data labelling*. To efficiently train and test DL algorithms, data should be properly labelled. The main problem is that, in many cases, data labelling is an extremely time-consuming process. In order to address this problem, we experimented a *Self-Training* approach based on *Transfer Learning* (by pre-training the network with

[2] https://wilddash.cc/railsem19
[3] https://www.mathworks.com/products/roadrunner.html
[4] https://unity.com
[5] https://www.unrealengine.com/en-US

a customised version of the RailSem19 dataset) to implement a *semi-automatic labelling*. Results seem to be quite promising but, as deeply explained in [11], further investigation would be required to make this process as effective as possible.

Certifiability and Explainable AI. To conclude, the main threats come from the *extreme aleatory of the environment*, the *explainability* of DL models, and the *certifiability* of AI systems. As for the former, besides the railways' structure (which, as mentioned above, can be considered quite constant over time), the ADM should be deeply tested and evaluated by considering hundreds or thousands of possible combinations between weather/light conditions and possible obstacles which may vary in dimensions, colour, distance, and type (e.g., cars, rocks, trees, animals, etc.). Exhaustive tests may not be feasible, however, at the current level of development of AI systems in general and to the best of our knowledge, it would also be extremely challenging to statistically identify possible *failure modes* and the *Safety Integrity Level (SIL)* of these systems given their *instability* and poor *explainability*. Instability refers to the fact that different DL algorithms may not be robust to slight variations of the input, i.e., specific pixel patterns (which may also be imperceptible to the human eye) can completely fool the DL model inducing it to produce completely wrong outputs. This is one of the main vulnerabilities exploited by *adversarial attacks* which, for example, may apply specific pixel patterns to the input frame to mislead the functioning of the system. In addition to that, given that many AI approaches (especially DL ones) are opaque by design, it would be extremely challenging to understand the reasoning that the DL model applied to obtain a specific output given a specific input. From this perspective, *eXplainable AI (XAI)* approaches, which are oriented at unrolling and, indeed, explain the behaviour of DL models, may introduce several opportunities for the understandably and, consequently, the certifiably of AI systems in railways.

4.3.3. Recommendations from the PoC

Herein, we report all the recommendations that specifically came out from the analyses we conducted in the context of the PoC discussed in this chapter.

Exploit 3D Editors and Scenario Simulators. As mentioned above, 3D Editors and Scenario Simulators offer different benefits when it comes to easily and safely collecting data to develop proofs-of-concept. The latest technology developments give this software the capability of approximating reality both in terms of the physics of the elements and graphic characteristics. Therefore, we think their potential is worth to be taken into account also in future research. In addition, ad-hoc railway software (as RoadRunner is for automotive scenes) facilitates further the replication of railway lines or railway assets (e.g., level crossings, bridges, etc.) and video data collection.

Exploit Task-Specific Characteristics. Our suggestion is to facilitate the tasks that DL models should face by exploiting some characteristics which are peculiar to the problem that is intended to be addressed. For example, we adopted this strategy when implementing the Rails Detection Module (RDM) in the context of our Obstacle Detection PoC. It would be possible to build a "general-purpose" RDM which is capable of extracting the rail tracks of potentially any railway line, however, in our view, it would be better to build a RDM for each rail line (or rail section) so that the DL model would not face the structure aleatory introduced by multiple railway lines. By training a DL model with data coming from the same railway line (or section) it will operate on, we are practically feeding it with almost all the possible structural data so that, at run-time, it will most likely process a scenario which structure has already been seen during the training phase. Therefore, the model should be “only” capable of generalising in terms of light and weather conditions, and possible intrusions. Important to mention, in this context, by “structural data” we mean the physical structure of the railway line, e.g., the number of tracks, curves angles, the structure of the trackside equipment, and so on.

Modular Approaches for Redundancy and Robustness. At the current stage of development, as far as we know, DL models, especially those oriented at processing images, may not be so robust to specific changes in the input data. For example, some particular pixel patterns, which may not be known a-priori and can be both generated through adversarial attacks or simply due to the aleatory of the environment, may fool the DL models, i.e., induce them to produce a wrong output. Therefore, especially in safety-critical scenarios, it would be advisable to adopt modular approaches so that each model can potentially compensate possible malfunctions of the other modules or extend their functionalities. In our Obstacle Detection PoC, the ADM is intended to extend the functionality of the ODM which is typically implemented for these kinds of tasks; however, the ADM could also act as a supportive module to confirm the output of the ODM (e.g., by highlighting an anomaly in the same position of the object detected by the ODM).

The redundancy can be scaled further by adding more sensors. Using different kinds of sensors may be the best choice, however, it may also be the most costly in terms of the cost of the sensors themselves and time to develop modules capable of processing different types of data. A first step that could be taken is to introduce multiple cameras monitoring the rail tracks from different angles. The object detection architecture would always be the same, but it will be replicated for each camera and trained on data coming from a particular camera only (i.e., Architecture *“n”* will be trained with images coming from Camera *“n”*). From a theoretical perspective, if the same architecture is trained on two different datasets, the two resulting models (i.e., the combination of a particular architecture and a specific dataset it is trained on) would most likely be different in terms of parameters (i.e., they are two different systems); therefore, it would be more unlikely that the same pixel pattern would fool both of them or, in general, that both the models will miss-detect the same obstacle. Since this could ensure, to some extent, more robustness to the whole detection system, we think that it would be worth performing further investigations and extensive evaluations in this direction.

5. A Critical Examination of the Virtual Coupling Proof-of-Concept

In recent years, a new paradigm emerged in the railway field based on the concept of Virtual Coupling (VC). It aims both to overcome the limitation of current fixed blocks railway systems, and to go beyond the concept of moving block introduced by ETCS Level 3, by considering the concept of a relative braking distance among trains. The idea, strictly related to the platooning concept in the automotive field, is to virtually couple two or more trains via Train-to-Train (T2T) communication, so that they can travel in formation with the same velocity while maintaining a desired inter-train safety distance among them. Being railway VC a fairly new and visionary paradigm, much still needs to be done to give an exhaustive answer to its numerous unaddressed issues. In the following, the most recent advancements are summarised, with a particular focus on the emerging AI-based solutions in VC control applications.

5.1. Recent Advancements on Virtual Coupling

Railway VC is currently a subject of intensive research. The interest in this field is indeed exponentially increasing in the last years: Fig. 5.1 shows the evolution of publication regarding railway VC applications over time. A significant research sub-field is emerging which aims at understanding the adoption of AI techniques in VC from a control viewpoint. Interesting results investigate the use of ML techniques in VC control applications, whose main goals are optimising the inter-train distances while preserving safe operations. [21, 22] show a recent comprehensive review of control techniques in railway VC, including AI-based emerging applications. We emphasise that these recent results update the overview of the state of the art addressed in RAILS Deliverable D2.2 [10]. In this perspective, the proposed PoC analysis is meant to be a further step towards the identification of some operational and technical issues related to the potential adoption of AI techniques in railway VC.

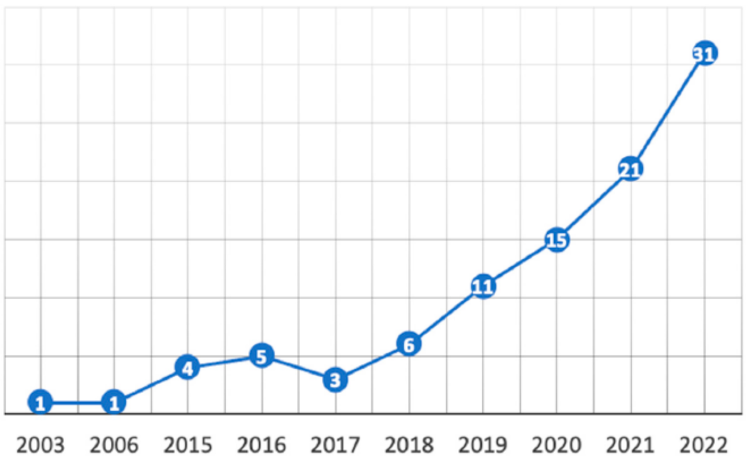

Fig. 5.1. Publications per year of VC applications in the railway field [21]

5.2. Bird-eye View of the PoC Approach

In RAILS Deliverables D2.2 [10] and D2.3 [11], we addressed a methodological and experimental PoC to explore the potential adoption of AI techniques from a control-based perspective. Specifically, we proposed a Deep Deterministic Policy Gradient (DDPG) control strategy, which belongs to the Deep Reinforcement Learning (DRL) methods, to investigate the effectiveness of the VC tactical layer functionalities, according to the conceptual VC view proposed in [23]. The model considers a Virtually Coupled Train Set (VCTS) of heterogeneous autonomous consists able to share information on their current position and speed with the other communicating trains via a T2T wireless communication network. The desired reference behaviour for the convoy is imposed through a Train-to-Infrastructure (T2I) communication network by the Radio Block Center (RBC), which acts as a virtual leader (see Fig. 5.2 for a conceptual architecture of railway VC). The proposed DDPG controller had

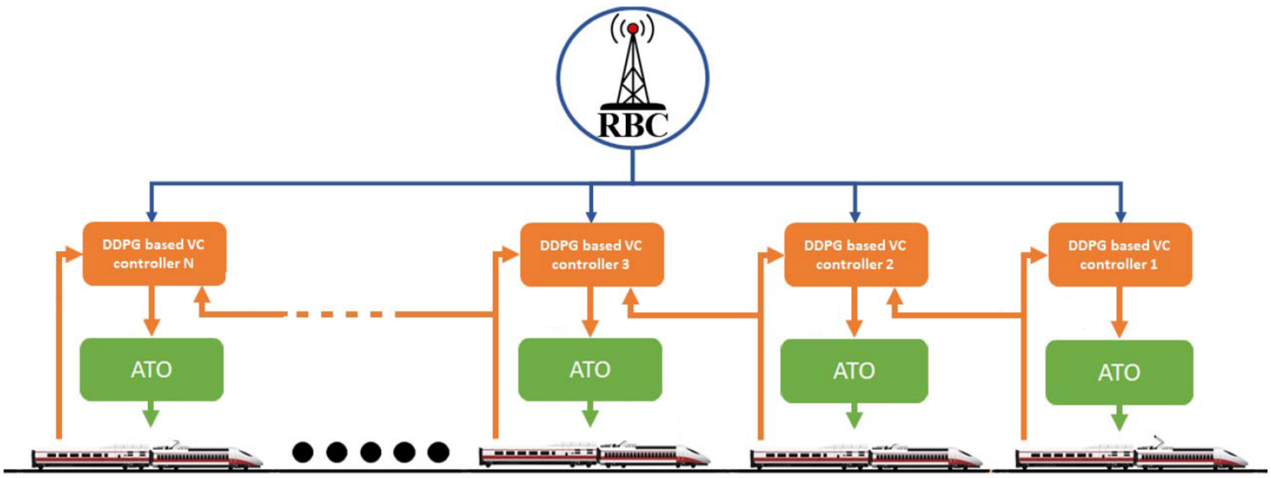

Fig. 5.2. Conceptual architecture of VC: VCTS of autonomous consists receiving information from RBC through the T2I communication network (blue lines), and sharing information among each other through the T2T communication network (orange lines).

the final aim to guarantee VC even in complex scenarios, despite the presence of: uncertainties in train dynamic parameters; heterogeneity in trains dynamics (e.g., different braking capabilities, different speed categories); uncertainties in track conditions (e.g., adhesion factors, gradients profile); external disturbances (e.g., wind speed) and unknown exogenous forces due to the curvature and the slope; uncertainties in reaction delay when performing braking manoeuvres, which can influence the inter-train distance; uncertainties in train location information; on-board speed error measurements. Therefore, it has been tested and validated through an ad hoc simulation platform in different operational scenarios and considering some basic VC manoeuvres: VCTS forming, VCTS splitting, and leader tracking. Preliminary results showed its effectiveness in guaranteeing the achievements of the VC requirements in the considered operational scenarios. Furthermore, a comparison analysis with a model-based MPC control strategy highlighted some potential advantages that the use of DRL techniques could provide in this field with respect to traditional models.

5.3. A SWOT Analysis of the PoC

In the following, an in-depth analysis of the methodological and experimental PoC proposed in the previous deliverables is carried out, arranged according to the SWOT structure. We highlight that a SWOT analysis which explores demand trends and operational scenarios in railway VC has been provided in [24], followed by the definition of scenario-based roadmaps as a useful tool for stakeholders to identify potential risks and criticalities in the deployment of VC [25]. In this context, the focus of the following analysis to report the major strengths (S) and weaknesses (W) shown by the proposed AI technology, and to identify external opportunities (O) that could support its technical feasibility, as well as some open challenges related to the main threats (T) emerged from the PoC (see Fig. 5.3 for an overview of the proposed SWOT analysis).

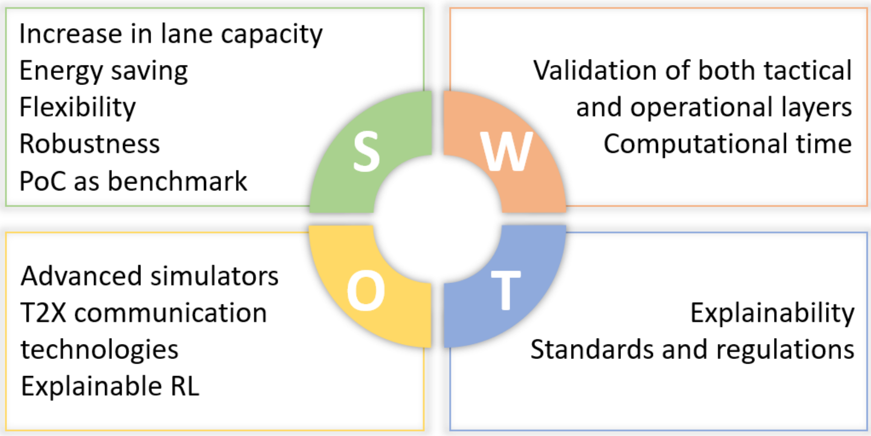

Fig. 5.3. Relevant aspects from the Virtual Coupling PoC arranged according to the SWOT Structure.

5.3.1. Strengths and Weaknesses

Increase in lane capacity and energy saving. In the proposed PoC, some evidences have arisen from the comparison analysis with a traditional MPC approach which highlighted some potential strengths of the proposed AI control strategy. The main evidence is represented by an increase in lane capacity: the proposed AI approach showed a better reference tracking, thus optimising the lane exploitation, while at the same time guaranteeing a relevant reduction in energy consumption. It is worth highlighting that capacity optimisation and energy saving are considered among the main objectives pursued by the VC paradigm with respect to the currently applied and in development fixed and moving block railway systems; thus, these strengths could have a significant weight in the evaluation of the proposed AI methodology.

Flexibility and robustness. Another advantage is represented by the inherent nature of the DDPG approach itself. Namely, being a model-free and a learning-based approach, it is suitable to face with heterogeneous train features and environments without having any prior knowledge of the train dynamics and the travelling scenarios. This guarantees more flexibility compared to model-based approaches, which require detailed modelling of the train dynamics and the surrounded environment, and thus must be adapted to the different train characteristics and the encountered operational scenarios. Furthermore, the proposed approach can better address uncertainties and inaccuracies due to unknown track conditions and external disturbances, thus improving robustness with respect to traditional models. As a consequence, the proposed strategy could better cope with the complex and uncertain railway real-life scenarios.

PoC as benchmarck. As discussed in Section 5.1, the interest of the research community in understanding feasibility of AI technologies in railway VC is increasing in the last years. Towards this direction, the proposed PoC could act as a benchmark for future research. Namely, additional Key Performance Indicators (KPIs), operational scenarios, and approaches could be considered to improve knowledge on the potential usage of RL control strategies in the field of VCTS paradigm.

Validation of tactical and operational layers. The proposed approach is focused on the tactical layer functionalities: it coordinates the actual platoon movements and manoeuvres from the instance a joining request is received until the platoon is dissolved. To do this, our DDPG control strategy computes in real-time the trajectory that each consist should follow in order to allow the VC. In the validation phase carried out in Deliverable D2.3 [11], we assumed that the operational layer located on each train (which should embed an Automatic Train Operation system with Grade of Automation GoA4) is able to automatically drive each consist guaranteeing the execution of the computed trajectory without any deviation from the latter. This assumption may be considered a limitation of the proposed approach and should be better investigated. Therefore, an extensive validation phase could be conducted by considering an architecture encompassing both the tactical and the operational layers to assess the effectiveness of the local train controllers to perform the tactical layer commands and to guarantee VC in a safe manner.

Computational time. Eventually, computational time should be better examined, since it could affect the performance of the proposed control strategy in real world scenarios.

5.3.2. Opportunities and Threats

Advanced railway simulation platforms. As already mentioned in Deliverable D2.2 [10], one of the main issues of the DRL methods is represented by the need of ad-hoc simulators for both the training and validation phases. Namely, these methods work in high-dimensional and continuous action spaces, which are difficult to explore efficiently; therefore, any datasets would be inadequate for the purpose. To the best of our knowledge, even though some simulators are currently available for railway virtual testing and RL training (see for instance SUMO [26] and Anylogic [27]), they do not allow the possibility of considering the VCTS paradigm, which also requires to take into account T2T communications and the involved dynamical phenomena. In view of this, an ad-hoc simulation platform has been provided in the PoC for the training and validation of the proposed DDPG approach. However, although this simulator takes into account different operative scenarios, train characteristics, track conditions, and exogenous factors, this is just a first attempt along this direction: future works will include the enrichment of the proposed simulation platform by encompassing other crucial aspects of trains dynamics and T2T communication, such as the multi-body modelling, the wheel-rail contact effects, suspensions effects and communication losses, etc., so to obtain a more accurate model of train dynamics able to emulate as much as possible the real systems. The development of advanced railway simulators, indeed, could represent a great opportunity for the effectiveness of our approach, and basically of DRL methods, which require extensive training and validation activities.

Emerging T2X communication technologies. Train-to-Everything (T2X) communication is one of the enabling technologies for the exploitation of the proposed VC control strategy, since its performances strongly depends on the communication latency. The current communication systems that support data rates up to 100 Mbps, operating at frequencies lower than 6 GHz, may not be sufficient to handle the future communication requirements of the railway industry. One potential solution to this challenge is represented by 5G technology, which is expected to play a fundamental role in Internet of Things (IoT) applications that require more reliable and low-latency communication. Namely, one of the key benefits of 5G technology is the reduction in latency thanks to the use of millimeter waves, which operate at higher frequencies than the existing technologies used in 4G. [28]. However, the implementation of 5G technology in the railway industry requires careful consideration of various factors, including signal propagation, network architecture, and cybersecurity. Additionally, the integration of 5G technology with existing railway systems and infrastructure may require significant investments and modifications.

Explainability and Explainable RL. Explainability may be considered a critical aspect for the proposed DRL algorithm, and more generally for the widespread acceptance of AI systems. It refers to the ability of the algorithm to provide transparent and interpretable justifications for its decisions and actions in order to enable humans to understand the reasoning behind the model's behaviour. An explanation for the automated decisions is especially crucial in safety critical applications, where understanding the causes of actions or decisions is a natural human requirement. To date, although RL models can be considered reliable, the effectiveness of these systems is limited by their current inability to explain their decisions and actions to human users. Namely, because of their inherently probabilistic nature, their specific processes of decision-making are currently almost impossible to understand. In this field, Explainable Reinforcement Learning (XRL) [29], a relatively new subfield of XAI, is emerging with the aim of supporting the spreading of RL-based applications with diverse audiences, requiring ethical, responsible and trustable algorithms. Constructing explainable intelligent driving systems is a viable promise for the trustworthy use of RL control algorithms for autonomous trains and VC.

Standards and regulations. The above concept of explainability should be accompanied by the compliance of the proposed approach with safety standards and principles established by transportation regulators. However, this expectation does not coincide with the current requirements of the European railway regulations for the functional safety of railway systems and their verification procedures, which do not account for AI-based applications. Therefore, the certification of the proposed DRL autonomous and cooperative train control system is an open challenge: while traditional train control sub-systems can be certified on the basis of today's CENELEC standards, this is not the case for all AI-based algorithms. Thus, the certification of such AI systems in railways will require extensions or modifications of the current standards, or the development of additional ad-hoc standards.

5.3.3. Recommendations from the PoC

In the following, the main recommendations emerged from the above SWOT analysis are summarised.

Development of advanced railway simulators for virtual testing. The development and availability of advanced railway simulation platforms for virtual testing could contribute to the actual exploitation of DRL approaches in the railway field, since the performances of these methods is strictly related to exhaustive training and validation phases. Emerging paradigms such as railway VC and its related enabling technologies should be taken into account in order to allow the virtual testing and validation of RL-based control strategies for VCTS.

Validation of the VC architecture. The assessment of the proposed DDPG controller shall include the verification and validation of the VC architecture according to the vertical layer structure proposed in [23]. As a recommendation emerged from the PoC, a first step forward could be that of considering the validation of both the tactical and operational layers, in order to assess the effectiveness of the operational layers located on each train to safely perform the VCTS manoeuvres computed by the proposed DDPG control strategy.

Exploring the potential of XRL for VC applications. As stated before, the proposed DDPG approach is affected by explainability issues related to its inherently probabilistic nature. In an attempt to overcome explainability limitations, the effectiveness of XRL methods could be investigated for the future deployment of VC control strategies. Indeed, some promising approaches to enhance the explainability of RL models are emerging [29]. Although XRL methods need to be further assessed, they could be crucial for the future exploitation of RL techniques in the real world applications. Namely, they could allow the development of explainable, trustable and responsible algorithms that could be deployed in practice and can be more understandable by the general public.

Extension of the current standards and regulations. The certification of the proposed control system is an open challenge, since, as highlighted in the previous section, current CENELEC railway standards do not explicitly account for the certification of safety critical AI-based railway systems. This is indeed a general issue which goes beyond the specific PoC, since it affects the actual exploitation of AI in railways applications. Certifications of AI-based algorithms in the railway field will indeed require the introduction of new concepts and requirements as well as the extension of the current standards. General recommendations and innovation needs for safety-critical railway functionalities can be found in Chapter 6, in which new concepts are introduced that could support the identification of specific requirements in the field of autonomous trains applications.

6. Recommendations and Innovation Needs

6.1. Deal with the Data Challenge

The problem of data availability (and labelling) is a cross-area, which means that we had to face it when addressing the majority of the PoCs developed within the various work packages of the RAILS project. This problem is described more in detail in Deliverable D3.4, which summarises our recommendations on AI for railway maintenance and inspection applications, where tons of data are typically required to properly set up defect detection/failure prediction systems.

However, as highlighted by the PoCs discussed in this document, this problem also affects the integration and development of AI approaches for railway safety and automation applications. In the following, we discuss the main directions we think could be useful to promote the development of AI approaches in this context, while we remand to Deliverable D3.4 for additional insights on the topic of data challenges.

Premised that data collected in real scenarios should always be preferred, there are some alternatives that can be exploited to generate data to be used to set up a suitable strategy for the resolution of a railway problem through AI approaches. In this context, our recommendations are:

Exploit 3D Editors and Scenario Simulators. As widely discussed in Chapter 4, the latest technological developments have given to these software the capability of approximating reality with a high degree of fidelity, both in terms of the physics of the objects and graphics. The advantage compared to real data collection is that these software gives the possibility of re-creating potentially any kind of event and thus allows the collection of various kinds of data in a reduced amount of time. In addition, it would be possible to: i) safely simulate harmful situations; and ii) avoid any problem related to data privacy.

Exploit Data Augmentation and Transfer Learning. In case real data are available, but only in a limited amount, it would be possible to exploit some techniques (e.g., those recently reviewed in [30]) to i) generate synthetic data starting from the real ones or ii) simply apply transformations (e.g., cropping, rotation, stretching) to increase the dimension of the datasets. Generative Adversarial Networks (GANs) are the most common example when it comes to generating images; they allow for the generation of real-like synthetic data which are, indeed, generated starting from the distribution of data collected in real scenarios.

Exploit Automatic / Assisted Labelling. Once data have been collected, whether they are real, synthetic, or a mix of these, they must be properly labelled to train and test AI models. The labelling process is typically performed manually by human operators, which implies costs in terms of time and human resources especially if hundreds of thousands of data should be labelled.

However, there are three main options that can be exploited to reduce labelling efforts. First, 3D Editors may include functionalities to automatically label objects in the generated scene (e.g., Unity's Perception Package[1]), in these case, in addition to generating data, the software would also automatically generate labels. Second, most of the labelling software and tools that have been developed in the last years, especially for image/video labelling, include functionalities to assist human operators in the labelling processes[2,3,4,5]; if properly used, these can extremely facilitate this task. Lastly, it would be possible to exploit AI approaches to create semi-automatic labelling algorithms based, for example, on self-training in order to manually label only a subset of data and then automatically generate labels for the remaining data. We adopted the last strategy to develop an algorithm to semi-automatically label data in the context of the Obstacle Detection PoC (see Deliverable D2.3 [11] for additional details); it resulted to be viable but further analyses would be required to understand its real effectiveness when it comes to small details in the images.

To conclude, we think that synthetic datasets can be used as benchmarks for the instigation of AI approaches and the definition of a suitable strategy oriented at improving safety and automation in railways. However, these kinds of data could be extremely useful only at the proof-of-concept level; once a strategy has been defined and a potential AI approach identified, their effectiveness should, obviously, be tested with real data.

6.2. Promote public challenges based on safety cases

Strictly related to data availability, the definition of public challenges (to be intended as competitions) may contribute to the fast take-up of AI in railways. In the past years, challenges like those proposed on the MS COCO[6], ImageNet[7], and RailSem19[8] datasets have promoted the technological advancement of AI applications on some specific topics with some solutions achieving state-of-the-art performances. The same can be done in railways, by establishing challenges for those railway problems that are intended to be tackled so that all the interested researchers and practitioners could participate in their resolution.

Important to mention, in order to propose proper challenges, there are some factors that should be taken into account:

- First, a challenge requires a shared dataset that all the participants should use to train and validate their AI models. In our view, this can be mitigated by generating synthetic datasets as we have done for some of the PoCs developed within the whole RAILS project. In this way, it would also be avoided any problem related to data privacy.
- Second, each challenge should encompass a series of requirements and constraints that should be adequately established by experts and the data for the challenge should be properly collected or generated so that they reflect these characteristics. The main problem in this case is that, in our view, this should be done in a centralised way. An authoritative entity, which could be an alliance between railway stakeholders and/or

[1] https://docs.unity3d.com/Packages/com.unity.perception@1.0/manual/index.html
[2] https://it.mathworks.com/help/vision/ug/get-started-with-the-image-labeler.html
[3] https://www.v7labs.com
[4] https://keymakr.com/annotation-tool.php
[5] https://segment-anything.com
[6] https://cocodataset.org/#home
[7] https://www.image-net.org/challenges/LSVRC/index.php
[8] https://wilddash.cc/benchmark/summary_tbl?hc=semantic_rob

railway experts, should be in charge of: i) defining the problems to be addressed; ii) establishing proper criteria, requirements, and constraints so that the challenge would reflect as much as possible the real problem; iii) identifying guidelines for the realisation of the dataset(s); iv) generating/collecting the required data; and v) managing the results of the competitions.

Despite these tough issues, we think that competitions of this kind, i.e., oriented at defining benchmarks for the resolution of specific railway problems, could effectively promote technology advancements in the rail sector.

6.3. Introduce new concepts in risk assessment

As a matter of fact, when dealing with safety-critical functionalities, the systems' certifiablity is of paramount importance. However, as also emerged from the aforementioned PoCs, the current state of development of AI applications does not allow, in most cases, the correct definition of failure modes and the precise identification of the SIL of AI systems. Therefore, new concepts should be taken into account in order to promote the integration of AI in safety-critical railway applications.

6.3.1. Safety Envelope as a Barrier for Intelligent Train Control

An opportunity to mitigate the certifiability issue and start introducing AI in Autonomous Train Driving functionalities is given, as described in [9], by the **Safety Envelope** paradigm [31–33]. The main idea is that *autonomous trains should move within an area (the Safe Operational State Space (SOSS) defined by the Safety Envelope) that is free from any hazard and is continuously computed and updated; any action that would bring the system outside the safety envelope should be detected and mitigated or aborted.*

An example can be made by taking into account two systems that have already been widely discussed in the rail sector, i.e., the Automatic Train Protection (ATP) and the Automatic Train Operation (ATO). The separation of concerns between ATP and ATO can be framed in the concept of Safety Envelope as follows. From a high-level perspective:

- The ATP is responsible for ensuring that the correct dynamic speed profile is applied in order to avoid derailments, collisions, and keep a safe distance between trains. Therefore, in the context of the Safety Envelope paradigm, the ATP would be the *"checker"* and would be responsible for *computing* the *SOSS* within which trains can operate safely.
- On the other hand, the ATO, which is oriented at automatically taking actions to move trains from one station to another, would be the *"doer"* and its actions would be *checked* by the ATP. For example, in case the ATO would accelerate over the maximum speed computed by the ATP, the ATP itself would perform some countermeasures to ensure safety (e.g., application of emergency braking).

Fig. 6.1 graphically shows this relationship by means of a flowchart that integrates some notions typical of the Activity Diagrams. The *start* (black circle) would cyclically happen any time the environment changes that much to require the computation of a new SOSS or when the ATO initially *"Takes an Action"*. The *end* (white circle with inner black circle), instead, is only related to the operations that the ATO would perform given that the ATP is continuously operating on the railway line. Lastly, the decision nodes are related to the following questions: "Safe?" stands for *"Does the action keep the system within the SOSS?"*; "Stop?" is

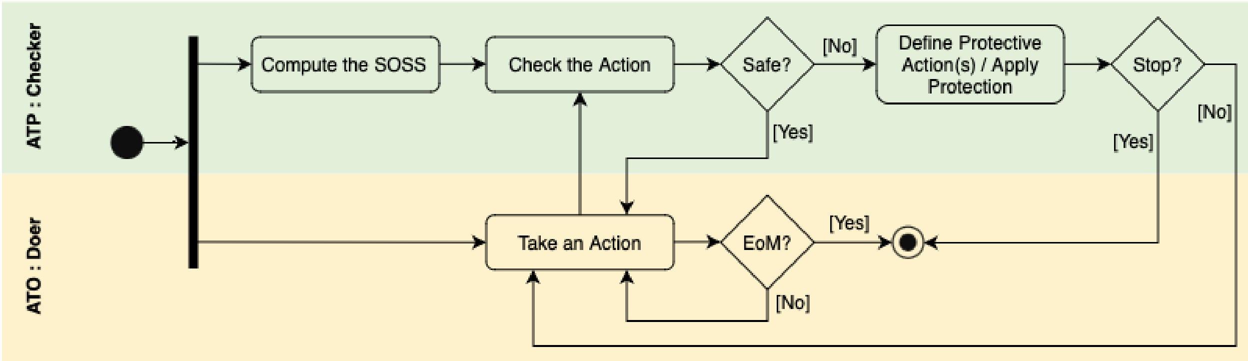

Fig. 6.1. ATP and ATO Concerns According to the Safety Envelope Paradigm.

related to *"Is the problem too critical to require a full-stop of the system (e.g., activate the emergency brake to avoid collision)?"*; lastly, "EoM" stands for End-of-Mission, i.e., *"Did the mission of the system end?"*. In this scenario, only the *"Computation of the SOSS"* and the operation of *"Checking the Action"* must be rated at SIL4, leaving the decision of how to optimally run the train (i.e., the actions taken by the ATO) to non safety-related and possibly complex software [9]. We report herein the **requirements** that, in our view, must be met in order to make possible the application of this paradigm in railways:

1. *Safety Envelope computation* must be based on sound and certified principles including not only information like the received Movement Authorities as in ETCS, but also any sensor input that contributes to the delineation of the safety envelope (e.g., the distance from leading train in VC).
2. *Safety Envelope computation* must consider all uncertainties on measures from sensors (position, distance, etc.) that impact the safety of train actions (e.g., braking distance) on a probabilistic basis. Then, the probability of producing a too permissive safety envelope (which can potentially lead to unpleasant events) must be kept under the limits given by SIL4.
3. *Safety Envelope computation software* must be formally provable through proper formal verification techniques in order to ensure that they meet the above principles.
4. *Safety Envelope checking operations* must focus only on verifying that driving commands to actuators do not bring the train outside the current SOSS. Such operations must be available at any time.
5. *Safety Envelope checking software* must be formally provable through model-checking.

Basing on this idea, it would be possible to move, step-by-step, towards Intelligent Train Control as discussed below.

6.3.2. Introducing AI in Train Driving: Grades of Intelligence.

Railway lines are typically classified according to four Grades of Automation (GoA), summarised in Table 6.1, depending on which kind of Automatic Train systems are implemented on that line. However, current GoAs do not allow to include in the classification *Autonomous* and/or *Intelligent* functionalities that railway systems could implement. In order to consider these aspects, a new classification should be introduced.

First, it would be important to underline the relation that holds among Automatic (Automation), Autonomous (Autonomy), and the role of AI [9]:

Table 6.1: Grade of Automation (GoA) levels.

Level	Description
GoA 0	Train operations are manually supervised by the driver, no automation.
GoA 1	Train operations are manually supervised by the driver supported by ATP.
GoA 2	Semi-automatic train operation. ATO and ATP systems automatically manage train operations and protection while supervised by the driver.
GoA 3	Driverless train operation with on-board staff handling possible emergencies.
GoA 4	Unattended train operation, neither the driver nor the staff is required.

Table 6.2: Automatic Train Management Systems.

System	Description
Automatic Train Operation (ATO)	Used to automatically drive the train and stop at stations when needed.
Automatic Train Protection (ATP)	Used to automatically protect the train by applying brakes when needed.
Automatic Train Control (ATC)	Both ATP and ATO are in place to ensure full control of the train.
Automatic Train Supervision (ATS)	Used to manage train schedules and coordinate routes along whole tracks.

- **Automation** refers to the ability of a system to automatically or semi-automatically (i.e., without or with partial human intervention) perform a given task based on *pre-specified rules*.
- **Autonomy** extends the concept of Automation and refers to the ability of a system to dynamically adapt to unexpected scenarios by taking independent decisions.
- In this context, **AI has the role of** making vehicles capable of learning from experience and taking autonomous decisions to adapt to changes in the environment.

According to these definitions, current driverless trains (e.g., those running in GoA3/4 metro lines), which are often considered autonomous rather than automatic, cannot be considered "intelligent" since they miss any learning and adaptation capabilities.

As the second and last point to consider, given the consideration above, new concepts for Intelligent Train Management Systems should be introduced that build upon and extend Automatic Train Management Systems; these are summarised in Tables 6.2 and 6.3.

All these considerations converge in the definition of *new* **Grades of Intelligence (GoIs)** which, extending in a complementary manner the already existing GoAs, could help to classify AI-supported railway lines and introduce AI in railways towards full-autonomous trains [9]:

- **GoI1**: this level represents **limited or no autonomy**. It includes ATC implementations where AI is not used or it is used for limited functions such as optimisation within ATS (i.e., *ITS*).
- **GoI2**: this level supports **partial autonomy** by including *only ITO (GoI2.1)* as an adaptive ATO with energy, capacity and/or comfort optimisation capabilities, or *only ITP (GoI2.2)* for driving assistance and/or as a low-speed backup system in case of ATP unavailability or limited supervision.
- **GoI3**: this level represent **full autonomy** by including *both ITO and ITP*, although no

Table 6.3: Intelligent Train Management Systems.

System	**Description**
Intelligent Train Operation (ITO)	Extends ATO functionalities by applying intelligent and adaptive behaviour in order to optimise passenger comfort, energy consumption, and line capacity.
Intelligent Train Protection (ITP)	Supports or potentially replaces ATP. If the *ATP is not available* (e.g., in old railway lines or when ATP failures oblige to partial supervision), ITP could replace ATP by, e.g., automatically recognising signals and/or obstacles. If the *ATP is available*, ITP can be a useful complement to detect events that are not managed by ATP.
Intelligent Train Control (ITC)	As for the ATC, ITC expects to have both ITO and ITP systems operating in a fully-connected environment.
Intelligent Train Supervision (ITS)	Extends ATS functionalities by exploiting AI to efficiently optimise (or maximise) railway line utilisation and average throughput by providing appropriate train routing solutions (e.g., to promptly respond to disruptions).

advanced learning and adaptation capabilities is considered. For instance, at GoI3, the artificial vision algorithms of ITP can be trained only once, e.g. to detect on-track obstacles, and never updated.

- **GoI4**: as GoI3, this level supports **full autonomy** by including *both ITO and ITP*, but it also requires **advanced learning and adaption capabilities** (e.g., online learning) and a **fully-connected and dynamically updated ecosystem** (e.g., based on extremely reliable train-to-train and train-to-infrastructure connections). GoI4 should be supported by higher levels of fog/cloud intelligence (introduced below) by using external AI models for big data analytics, such as those enabled by Digital Twins.

GoI levels are graphically shown in Fig. 6.2.

6.3.3. Bridging the Gaps towards GoIs

In order to move towards GoI1 and GoI2, the Safety Envelope paradigm, as explained above, can be exploited to establish safety barriers for the safe operability of the ITS and ITO systems.

Towards GoI1. As mentioned, GoI1 does not introduce any autonomy in relation to train driving; instead, it could involve some advanced ATS functionalities based on AI (i.e., the ITS). Despite the fact that at the current level of development it would be challenging to estimate the SIL of the ITS, it would be possible to *exploit the Interlocking Subsystem (IXL), which is rated at SIL4, as the Safety Envelope for the ITS*.

Towards GoI2.1. Similarly to what has been discussed above, *the ATP, which is rated at SIL4, could act as the Safety Envelope for the ITO*. As an example, the ITO could implement a functionality based on RL to intelligently compute the optimal train speed to save energy and ensure passenger comfort; in case the computed speed would exceed the breaking curve computed by the ATP, this would take the proper action to protect the network.

Towards GoI2.2 and GoI3. GoI2.2 expects the ITP to be implemented, but not the ITO; differently, GoI3 expects that both systems are operating. In these cases, AI would be directly involved in the computation of safety functionality, in case no human operators would be involved to manually check the correctness of ITP decisions. To the best of our knowledge, although theoretically possible, *such levels seem not to be practically*

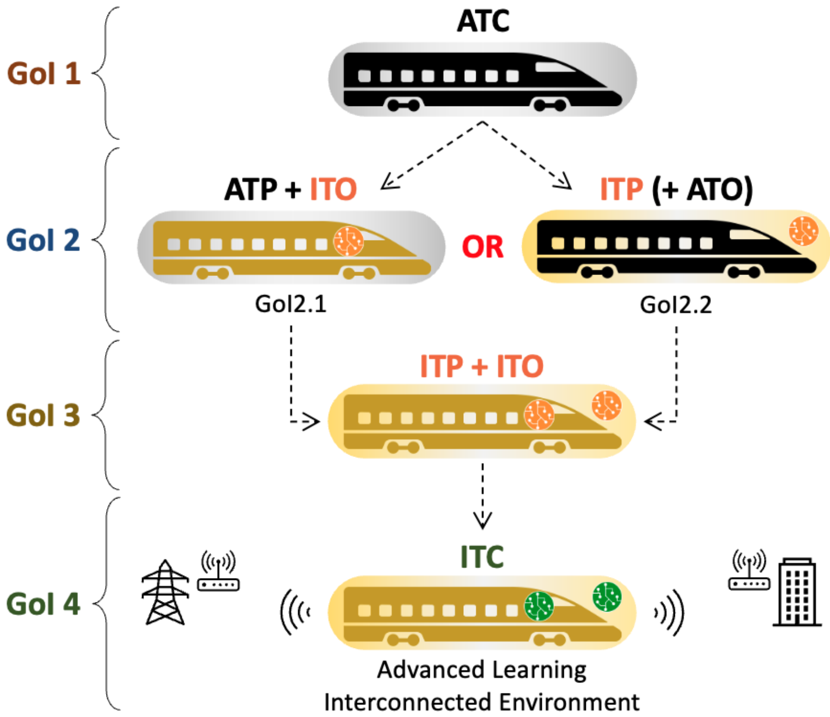

Fig. 6.2. GoI Levels.

reachable yet as ad-hoc standards and regulations would be required to quantitatively assess the trustworthiness of AI systems, together with legal, ethical, robustness, and explainability implications.

Towards GoI4. The same considerations as above for the combination of ITO and ITP hold also in this case, however, GoI4 also expects to have a fully-connected environment from which trains can receive/collect information at run-time to promptly take autonomous decisions. In order to achieve that, it would be necessary to establish a new view for the railway environment which specifically identifies how AI would be distributed at different *Levels of Intelligence (LoIs)* (i.e., edge, fog, and cloud) and the role that the AI systems deployed at the different LoIs would have in the context of intelligent fully-connected railways (or ITC).

6.3.4. Distribute Computations over different Levels of Intelligence (LoIs)

Fig. 6.3 shows an overview of the distribution of AI at different levels which are necessary to implement autonomous railway control and supervision. The Levels of Intelligence include:

- **Edge Intelligence**: AI systems mounted on-board the assets (e.g., trains) would be in charge of implementing *local autonomy* (e.g., on-board obstacle detection). They may be characterised by possible limitations in terms of computing power, due to constrained devices, but they would also be advantaged in terms of response times and data security, due to shorter communication links.
- **Fog Intelligence**: AI systems at this level would monitor a cluster of assets and manage their interconnections in order to achieve a comprehensive optimisation of a railway sub-system. For example, fog intelligence may represent trackside control where capacity optimisations (e.g., VC) can be orchestrated based on a larger knowledge of what is happening within a whole railway line.

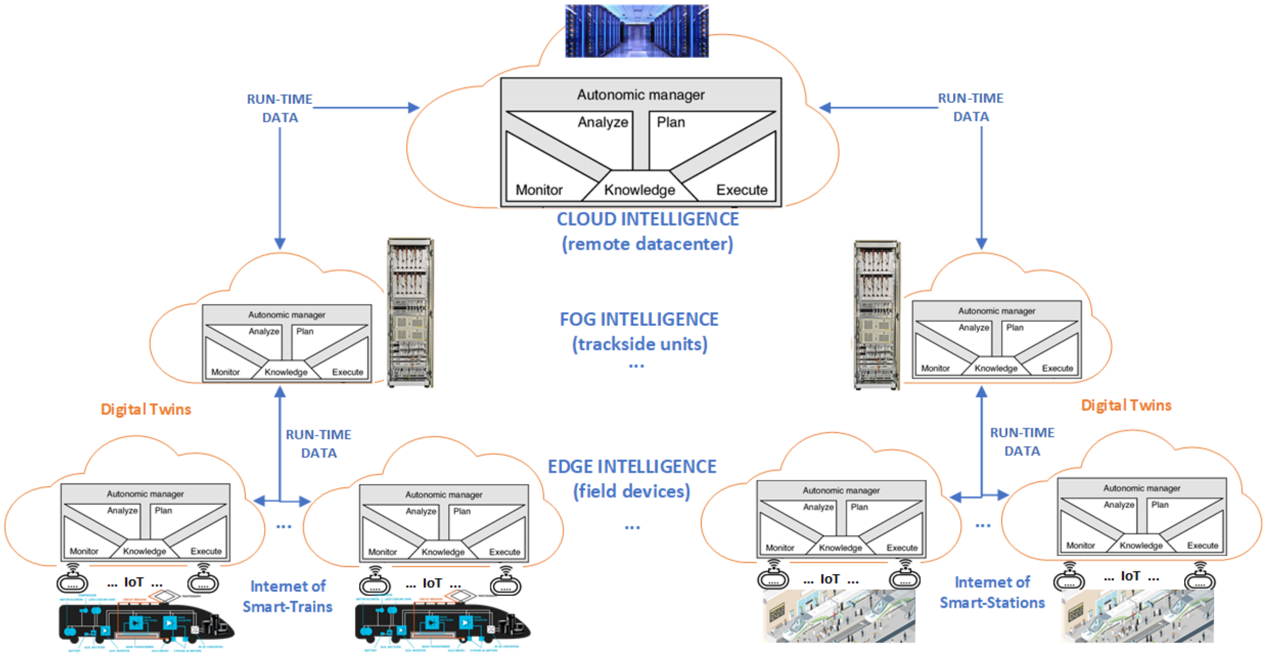

Fig. 6.3. Levels of Intelligence in Railway Control and Supervision.

- **Cloud Intelligence**: AI systems deployed at this level would aim at elaborating big amount of data, possibly coming from multiple assets installed worldwide, in order to collect information and knowledge to be either shared with the AI systems operating at lower LoIs or exploited to implement more comprehensive systems. For example, a failure prediction system can be trained with data coming from multiple installations of the same asset. These would most likely include more information about potential failures compared to data obtained by monitoring a single asset only.

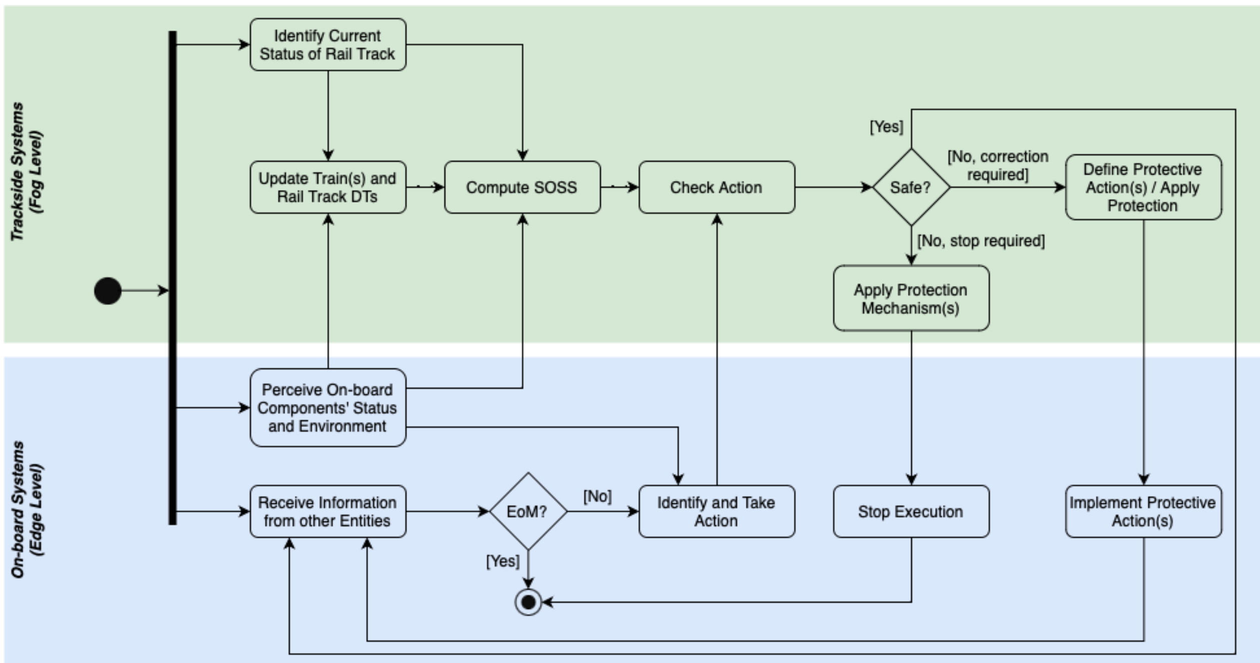

Fig. 6.4. Integrating LoIs and Safety Envelope.

An example of how Autonomous Trains functionalities and the concept of Safety Envelope can be distributed according to these LoIs is given in Fig. 6.4, which extends the example given in Fig. 6.1. Basically, an Autonomous Train receives information from other vehicles or the trackside equipment, then, if the mission has not ended, it integrates this information with those retrieved from the onboard sensors to autonomously define the action to take. Then, the action is processed by the Trackside Systems which act as the safety Envelope as discussed in Section 6.3.1. Hence, if the action is considered to be safe, the process restarts, otherwise, it depends on the severity of the problem that occurred. If drastic action is required to ensure safety, the Autonomous Train would be notified to stop or activate some emergency procedure, otherwise, it will be notified to process some corrective actions. In this extended example, besides the computation of SOSS and the check operation, also the various communication systems and the "Stop Execution" and "Implement Protective Action(s)" operations must be rated SIL4, as they would be extremely critical to avoid unpleasant consequences. This means that, at the current stage of development, such functionalities should not involve any AI and must have full control over AI-aided functions or must have higher priority.

6.4. Exploit the potential of Digital Twins

Among the various definitions that have been given for Digital Twins (DTs) [34], we rely on the concept formulated in [35]: *"DTs concentrate on bilateral interdependency between physical and virtual assets"*. Therefore, a DT is not a *simple* digital model of a real asset, instead, *"a DT is an accurate model of a physical entity which is kept alive at run-time and updated with real-time data collected from (IoT) monitoring devices"* [36]. In other words, at each time instant, a DT is a digital replica of the corresponding physical asset and represents its current state of evolution. DTs of this kind could introduce several opportunities towards the automation of railway lines.

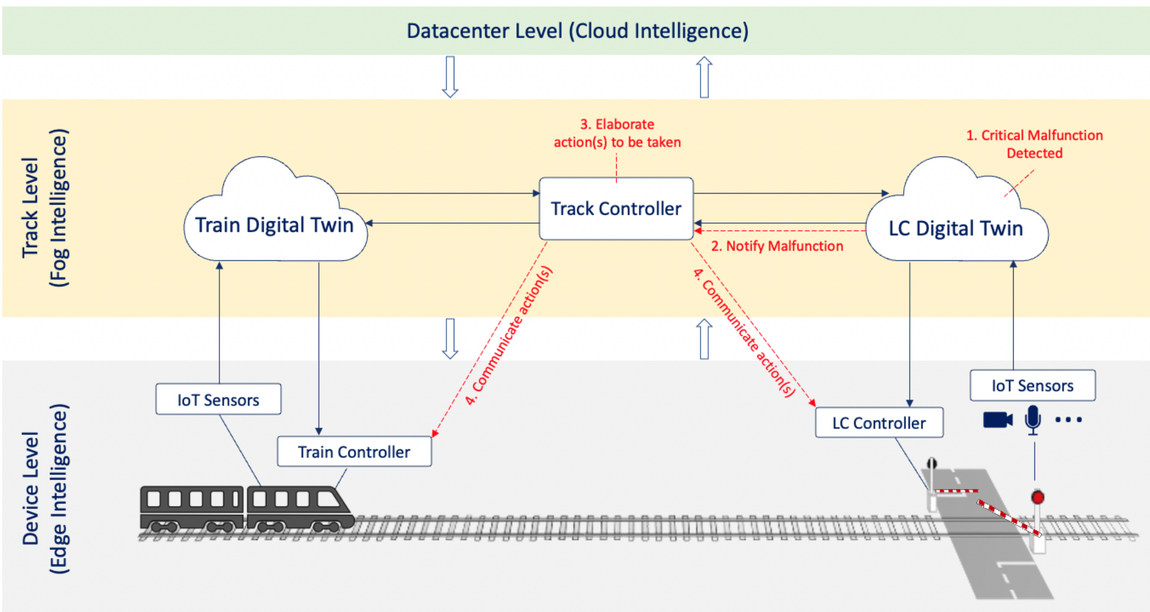

Fig. 6.5. Integrating DTs and LoIs to Ensure Protection in Case of Assets' Failure (excerpted from [37]).

DTs and LoIs. The architecture in Fig. 6.3 would build upon the DT paradigm: each asset, whether it is a single entity (e.g., a train) or an ensemble of entities (e.g., a railway line), would be managed through the corresponding DT. Therefore, in addition to direct communications between physical assets at the Edge Level (e.g., T2T and T2I), it would also be possible to exploit communications between DTs at the Fog or Cloud levels to take decisions that would optimise operations on a railway line and, at the same time, increase the level of safety. An example is given in [37] and, even though it is strictly related to maintenance at level crossings (a WP3 PoC), we think it could be relevant also in this case to understand how DTs and LoIs could be integrated to ensure protection and increase the safety of railway lines. As shown in Fig. 6.5, assuming it would be possible to generate DTs for both trains and level crossings, these can be exploited at the Fog Level to promptly adopt countermeasures in case of assets' failure. Additional details are given in Deliverable D3.2 [37].

Predicting Assets' Evolution with Cognitive DTs. If DTs evolve with the corresponding assets, it would be possible to make copies of DTs (which represent the physical assets at a given time instant) and stress them with various inputs to study the potential evolution of physical systems *and/or generate data that can be used to train AI models*. Cognitive Digital Twins (CDT) [38], i.e., DTs enhanced with cognitive capabilities making them capable of learning and reasoning, would introduce several opportunities in this direction as these kinds of DTs not only evolve with the physical counterpart but would also be able of simulating their behaviour.

DT-in-the-Loop for testing and validation. In addition to what is described above, *DTs can also support the development, validation, and testing of railway systems* [39], including those leveraging AI for Autonomous Train Driving. DTs could help to overcome, at least in the first phase of development, some challenges introduced by on-field tests including elevated costs and time required, limited scenarios (as it would not be possible to evaluate scenarios different from those captured through sensors and it would be extremely time-consuming to wait for all possible evolution of the environment), and high risk in case of failures. DTs can be exploited, especially if combined with Mixed Reality (MR) and AI, to evaluate the behaviour of systems and their functionalities similar to what happens with Hardware-in-the-Loop. As spotted within the Automotive field, MR has shown great potential in linking physical vehicles with their digital counterparts making both of them susceptible to events occurring either in the real or in the virtual world [40, 41]. As a simple example, this can be exploited in the rail sector to test the physical reaction of a train (running on an obstacle-free track) and its obstacle detection system to obstacles (of any kind and in any position) which are not physically on the tracks but are simulated in the virtual world (see Fig. 6.6).

Fig. 6.6. DT-in-the-Loop and Example of Mixed Reality.

7. Conclusions

This document reported the identification of possible future innovation needs and recommendations in the railway industry to enhance safety, efficiency, and intelligence in train operations. It addressed a detailed analysis of the two proofs-of-concept (PoCs) proposed in the previous WP2 deliverables, namely, Obstacle Detection and Virtual Coupling. The report critically examined the outcomes of the PoCs, discussing recent advancements in their respective fields, and conducting SWOT analyses to identify the main Strengths, Weaknesses, Opportunities, and Threats. Some recommendations emerged that aimed to address the identified challenges in order to enhance the technical and operational feasibility of the proposed AI approaches.

Specifically, as to Obstacle Detection, recommendations are mainly oriented at highlighting the benefits that 3D editors could introduce when it comes to data collection for the fast realisation of PoCs, promoting the exploitation of domain- and task-specific characteristics which are peculiar to the rail sector, and underlining the potential of unsupervised anomaly detection approaches (integrated within a modular architecture) to detect any possible obstacle on rail tracks. Regarding Virtual Coupling, the recommendations are strictly related to the potential of Reinforcement Learning methods to develop control strategies that could guarantee the Virtual Coupling objectives in a safe manner. The main criticalities concern the lack of explainability due to the inherent probabilistic nature of the proposed approach, the need for advanced simulators for virtual testing and validation, and the possible extension of the current European railway standards and regulations to account for the certification of safety-critical AI-based applications.

Eventually, general recommendations and innovation needs for future developments on rail safety and automation have been drawn. These include dealing with data challenges related to the lack of data availability, promoting public challenges based on safety cases, introducing new concepts in risk assessment, such as safety envelopes and Grades of Intelligence (GoI), and exploiting the potential of Digital Twins and Mixed Reality to test and validate railway systems and AI functionalities. These recommendations will be exploited in WP5 to identify migration strategies and roadmaps for AI integration in the rail sector.

Bibliography

[1] RAILS, "Deliverable D1.1: Definition of a reference taxonomy of AI in railways," October 2020. [Online]. Available: https://rails-project.eu/downloads/deliverables/

[2] N. Bešinović, L. De Donato, F. Flammini, R. M. P. Goverde, Z. Lin, R. Liu, S. Marrone, R. Nardone, T. Tang, and V. Vittorini, "Artificial intelligence in railway transport: Taxonomy, regulations, and applications," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 9, pp. 14 011–14 024, 2022.

[3] RAILS, "Deliverable D1.2: Summary of existing relevant projects and state-of-the-art of AI application in railways," 2021. [Online]. Available: https://rails-project.eu/downloads/deliverables/

[4] R. Tang, L. De Donato, N. Besinović, F. Flammini, R. M. Goverde, Z. Lin, R. Liu, T. Tang, V. Vittorini, and Z. Wang, "A literature review of artificial intelligence applications in railway systems," *Transportation Research Part C: Emerging Technologies*, vol. 140, p. 103679, 2022.

[5] L. D. Donato, F. Flammini, S. Marrone, C. Mazzariello, R. Nardone, C. Sansone, and V. Vittorini, "A survey on audio-video based defect detection through deep learning in railway maintenance," *IEEE Access*, vol. 10, pp. 65 376–65 400, 2022.

[6] M. J. Pappaterra, F. Flammini, V. Vittorini, and N. Bešinović, "A systematic review of artificial intelligence public datasets for railway applications," *Infrastructures*, vol. 6, no. 10, p. 136, 2021.

[7] RAILS, "Deliverable 1.3: Application Areas," 2021. [Online]. Available: https://rails-project.eu/downloads/deliverables/

[8] RAILS, "Deliverable D2.1: WP2 Report on case studies and analysis of transferability from other sectors (safety and automation)," 2021. [Online]. Available: https://rails-project.eu/downloads/deliverables

[9] F. Flammini, L. De Donato, A. Fantechi, and V. Vittorini, "A vision of intelligent train control," *Preprint submitted to International Conference on Reliability, Safety, and Security of Railway Systems (RSSRail 2022), Lecture Notes in Computer Science, Springer*, 2022.

[10] RAILS, "Deliverable D2.2 – WP2 Report on AI approaches and models," 2022. [Online]. Available: https://rails-project.eu/downloads/deliverables

[11] ——, "Deliverable D2.3 - WP2 Report on experimentation, analysis, and discussion of results," 2023. [Online]. Available: https://rails-project.eu/downloads/deliverables

[12] G. Basile, E. Napoletano, A. Petrillo, and S. Santini, "Roadmap and challenges for reinforcement learning control in railway virtual coupling," *Discover Artificial Intelligence*, vol. 2, no. 1, p. 27, 2022.

[13] *Main Elements of Risk Assessment*. John Wiley Sons, Ltd, 2020, ch. 3, pp. 59–85. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/9781119377351.ch3

[14] SMART, "Smart Automation of Rail Transport," 2016-2019, http://smart.masfak.ni.ac.rs.

[15] SMART2, "Advanced integrated obstacle and track intrusion detection system for smart automation of rail transport," 2019-2022, https://smart2rail-project.net.

[16] D. Ristić-Durrant, M. Franke, and K. Michels, "A review of vision-based on-board obstacle detection and distance estimation in railways," *Sensors*, vol. 21, no. 10, p. 3452, 2021.

[17] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*. Springer, 2015, pp. 234–241.

[18] L. Wang, D. Zhang, J. Guo, and Y. Han, "Image anomaly detection using normal data only by latent space resampling," *Applied Sciences*, vol. 10, no. 23, 2020.

[19] P. Bergmann, S. Löwe, M. Fauser, D. Sattlegger, and C. Steger, "Improving unsupervised defect segmentation by applying structural similarity to autoencoders," in *Proceedings of the 14th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*. SCITEPRESS - Science and Technology Publications, 2019.

[20] D. Ristić-Durrant, M. Franke, K. Michels, V. Nikolić, M. Banić, and M. Simonović, "Deep learning-based obstacle detection and distance estimation using object bounding box," *Facta Universitatis. Series: Automatic Control and Robotics*, vol. 20, no. 2, pp. 075–085, 2021.

[21] J. Felez and M. A. Vaquero-Serrano, "Virtual coupling in railways: A comprehensive review," *Machines*, vol. 11, no. 5, p. 521, 2023.

[22] Q. Wu, X. Ge, Q.-L. Han, and Y. Liu, "Railway virtual coupling: A survey of emerging control techniques," *IEEE Transactions on Intelligent Vehicles*, 2023.

[23] "X2Rail-3 Deliverable 6.1 - Virtual Train Coupling System Concept and Application Conditions," 2020. [Online]. Available: https://projects.shift2rail.org/s2r_ip2_n.aspx?p=X2RAIL-3

[24] J. Aoun, E. Quaglietta, and R. M. Goverde, "Exploring demand trends and operational scenarios for virtual coupling railway signalling technology," in *2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2020, pp. 1–6.

[25] ——, "Roadmap development for the deployment of virtual coupling in railway signalling," *Technological Forecasting and Social Change*, vol. 189, p. 122263, 2023.

[26] "SUMO - Simulation of Urban MObility." [Online]. Available: https://www.eclipse.org/sumo/

[27] "Anylogic." [Online]. Available: https://www.anylogic.com/

[28] W. Gheth, K. M. Rabie, B. Adebisi, M. Ijaz, and G. Harris, "Communication systems of high-speed railway: a survey," *Transactions on Emerging Telecommunications Technologies*, vol. 32, no. 4, p. e4189, 2021.

[29] A. Heuillet, F. Couthouis, and N. Díaz-Rodríguez, "Explainability in deep reinforcement learning," *Knowledge-Based Systems*, vol. 214, p. 106685, 2021.

[30] A. Mumuni and F. Mumuni, "Data augmentation: A comprehensive survey of modern approaches," *Array*, vol. 16, p. 100258, 2022.

[31] P. Koopman and M. Wagner, "Toward a framework for highly automated vehicle safety validation," SAE Technical Paper, Tech. Rep., 2018.

[32] A. Rudolph, S. Voget, and J. Mottok, "A consistent safety case argumentation for artificial intelligence in safety related automotive systems," in *ERTS 2018*, 2018.

[33] Y. Lyu, Q. Pan, C. Zhao, Y. Zhang, and J. Hu, “Vision-based uav collision avoidance with 2d dynamic safety envelope,” *IEEE Aerospace and Electronic Systems Magazine*, vol. 31, no. 7, pp. 16–26, 2016.

[34] M. Liu, S. Fang, H. Dong, and C. Xu, “Review of digital twin about concepts, technologies, and industrial applications,” *Journal of Manufacturing Systems*, vol. 58, pp. 346–361, 2021.

[35] G. Bhatti, H. Mohan, and R. Raja Singh, “Towards the future of smart electric vehicles: Digital twin technology,” *Renewable and Sustainable Energy Reviews*, vol. 141, p. 110801, 2021.

[36] R. Dirnfeld, L. De Donato, F. Flammini, M. S. Azari, and V. Vittorini, “Railway digital twins and artificial intelligence: Challenges and design guidelines,” in *Dependable Computing–EDCC 2022 Workshops: SERENE, DREAMS, AI4RAILS, Zaragoza, Spain, September 12, 2022, Proceedings*. Springer, 2022, pp. 102–113.

[37] RAILS, “Deliverable 3.2 – WP3 Report on AI approaches and models,” 2022. [Online]. Available: https://rails-project.eu/downloads/deliverables/

[38] X. Zheng, J. Lu, and D. Kiritsis, “The emergence of cognitive digital twin: vision, challenges and opportunities,” *International Journal of Production Research*, vol. 60, no. 24, pp. 7610–7632, 2022.

[39] B. Yu, C. Chen, J. Tang, S. Liu, and J.-L. Gaudiot, “Autonomous vehicles digital twin: A practical paradigm for autonomous driving system development,” *Computer*, vol. 55, no. 9, pp. 26–34, 2022.

[40] M. F. Drechsler, J. Peintner, F. Reway, G. Seifert, A. Riener, and W. Huber, “Mire, a mixed reality environment for testing of automated driving functions,” *IEEE Transactions on Vehicular Technology*, vol. 71, no. 4, pp. 3443–3456, 2022.

[41] M. Szalai, B. Varga, T. Tettamanti, and V. Tihanyi, “Mixed reality test environment for autonomous cars using unity 3d and sumo,” in *2020 IEEE 18th World Symposium on Applied Machine Intelligence and Informatics (SAMI)*, 2020, pp. 73–78.