

An intelligent and highly reliable system for safe autonomous railway driving



Djamila Zamouche^{1,*}, Sarah Chabane-Mechiouri¹, Mawloud Omar², Mouna Ouazine¹ and Loubna Mohamadi¹

¹ LIMED Laboratory, Abderrahmane Mira University of Bejaia, Bejaia 06000, Algeria

² IRISA Laboratory, University of South Brittany, Vannes 56000, France

* Correspondence author; E-mail: djamila.zamouche@univ-bejaia.dz.

Highlights:

- Intelligent autonomous train framework integrating vision, AI, and decision-making.
- YOLOv5-based object detection adapted for railway environments.
- Deep neural network risk assessment aligned with Safety Integrity Levels.
- Open automata model for autonomous train control decisions.

Abstract: Railway transportation is increasingly relying on autonomous technology to improve safety and efficiency. However, conventional control systems often lack the ability to adapt to the dynamic and uncertain nature of railway environments. In this paper, we propose an intelligent autonomous train framework that integrates computer vision, artificial intelligence-based risk assessment, and adaptive decision-making. The proposed system is organized into four complementary phases, namely multi-sensor environmental perception, object detection using You Only Look Once version 5 (YOLOv5), deep neural networks aligned with safety integrity levels (SIL), and open automata-based control for real-time decision generation. All of these components work together to ensure that trains can operate independently and safely in complex situations. Experimental results demonstrate the system's ability to accurately assess risks and provide robust and reliable performance under realistic operational conditions.

Keywords: autonomous railway systems; computer vision; risk assessment; deep learning; YOLO; open automata; safety integrity levels

1. Introduction

The railway transportation sector is undergoing a profound transformation driven by technological progress and the increasing need for automation to enhance operational safety. While autonomous vehicles have received considerable attention in the road transport domain, the railway sector presents



Copyright©2026 by the authors. Published by ELSP. This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium provided the original work is properly cited.

challenges for autonomy development. Railway systems benefit from controlled environments and dedicated infrastructure; however, they must also meet stringent safety requirements due to high vehicle masses, long braking distances, and the potentially catastrophic consequences of operational failures [1]. Despite significant advancements, current railway operations still depend heavily on human supervision and decision-making, particularly in complex environments where trains interact with pedestrians, vehicles, and dynamic infrastructure elements. According to international railway safety statistics, human error accounts for approximately 40% of railway accidents [2], highlighting the potential benefits of autonomous systems that can operate with consistent precision and rapid response times.

The need for advanced automation becomes particularly critical in dense urban railway contexts, where the system must continuously interpret diverse environmental inputs, such as weather and obstacles, while maintaining operational safety. Traditional control systems often fail to handle such dynamic variability. Although existing solutions, such as train-centric communication-based train control (TC-CBTC), offer reliable inter-train coordination and movement authority management, they lack the perception and risk assessment capabilities required for fully autonomous operations [3]. Bridging this gap between infrastructure-dependent control and self autonomy remains one of the most critical challenges in next-generation railway systems.

Although recent progress in autonomous vehicle technologies has been remarkable, most developments have been confined to road transport. Computer vision models, particularly those based on You Only Look Once (YOLO) architectures, have achieved high performance in real-time object detection and classification for road vehicles [4]. However, direct adaptation of these approaches to railway environments is not straightforward. Rail conditions, such as track-constrained motion, fixed pathways, and unique obstacle types demand tailored perception and decision mechanisms. Consequently, current railway automation remains largely deterministic and infrastructure-driven, limiting its ability to handle unexpected scenarios.

Previous studies in railway autonomy have often addressed isolated aspects of the problem, such as obstacle detection or risk estimation without developing fully integrated frameworks. Vision-based detection systems [5] have shown the potential to recognize obstacles and signals, but typically operate independently of decision and control units. Similarly, traditional risk assessment methods rely on predefined hazard matrices or static probability models, which lack the adaptability required for real-time safety evaluation in dynamic environments.

Furthermore, applications of machine learning in the railway domain have focused on predictive maintenance, fault diagnosis, and traffic optimization [6]. Although these contributions have improved operational efficiency, they have not yet fully leveraged artificial intelligence for autonomous perception and decision-making. The current gap between the proven capabilities of AI in other autonomous systems and its limited deployment in railway safety applications presents a significant research opportunity. This motivates the development of an intelligent and integrated framework capable of perceiving the environment, evaluating risks in real time, and autonomously generating control actions for safe and efficient train operation.

In this paper, we propose an autonomous railway system that integrates computer vision, artificial intelligence-based risk assessment, and decision-making within a unified architecture. The proposed

approach processes environmental data through perception modules, performs real-time object detection using a YOLO version5 model, and evaluates operational risks through a deep neural network calibrated according to safety integrity levels (SIL). Control decisions are generated through an open automata mechanism designed to ensure responsive and reliable system behavior. Unlike traditional control systems, our proposal enables continuous learning and dynamic adaptation to changing environments. Experimental evaluations on railway datasets demonstrate strong object recognition performance and a high level of consistency with safety assessments, confirming the feasibility of the approach for real-world autonomous train operations.

The main contributions of this paper are as follows:

- (1) An architecture that combines perception, risk assessment, and decision-making to achieve fully autonomous and safe railway operation.
- (2) A risk assessment strategy, calibrated against SIL, capable of continuously evaluating and adapting to real-world operational uncertainties in complex railway environments.
- (3) A control strategy for generating optimal, real-time decision under dynamic and uncertain conditions.
- (4) An experimental evaluation demonstrating the robustness, accuracy and suitability of the proposed system for deployment in complex railway scenarios.

The remainder of this paper is structured as follows. In Section 2, we present an overview of the related work. In Section 3, we introduce the general idea of the proposed solution as well as present its operation in detail. The results of the performance evaluation will be provided in Section 4. Finally, we conclude the paper in Section 5.

2. Literature review

In this section, we provide a comprehensive review of existing safety mechanisms for the design of autonomous trains.

2.1. Vision-based perception and detection systems

Several works have explored computer vision and deep learning for railway environment perception. Etxeberria-García *et al.* [7] have proposed an approach based on computer vision and deep learning techniques. First, they used an object detection model to identify and locate railway signals, traffic signs, and other visual elements. Then a deep neural network is used to analyze these data and make decisions about stopping the train. The algorithm is trained on data sets that contain various examples of railway scenarios. These vision-based techniques provide high accuracy in detecting safety critical elements and offer essential perception capabilities for real-time decision-making. However, their performance remains dependent on training dataset quality and diversity, which may limit robustness in degraded or unexpected conditions.

2.2. Formal verification and control logic

Formal methods have been employed to ensure correctness and safety in railway control systems. Xie *et al.* [8] have used well-formed Petri nets (WFN) to develop logic controllers for railway automation and to verify properties such as the safety of the automated system, with the aim of modeling

control functions and generating code that could be implemented on ground infrastructure computers. Chouchani *et al.* [9] have proposed an approach to map modeling or embedded mapping based on various standards that include information on infrastructure, signage, and connections. Safety rules are integrated into the conceptual map model, with the aim of automating their integration on board and the management of safety decisions. Mohammed *et al.* [10] have developed a process for developing a prototype driverless train implemented using a peripheral interface controller (PIC) microcontroller. They have simulated the system circuits using Proteus software. The hardware circuits, built on printed circuit boards (PCBs), are interfaced with actuators and sensors. They have then assembled the hardware into a prototype train that resembles a toy. The prototype train follows a predefined trajectory in terms of appropriate starts and stops at stations, announcements to passengers, and alerts. While offering rigorous validation, these formal methods often introduce high computational demands and show limited adaptability to dynamic operating environments, raising challenges for large-scale deployment.

2.3. Multi-agent systems and coordination approaches

Agent-based architectures have been investigated for distributed railway management and coordination. Dordal *et al.* [11] have proposed an approach to develop an autonomous driving train to pass through a bidirectional crossing loop to ovoid stops. In the proposed approach, they have used an intelligent system that can simulate the railway environment using driver agents and agents with a high level of railway management. The agents coordinate with each other by exchanging messages in order to avoid stops along the route. These multi-agent approaches demonstrate adaptability and distributed coordination capabilities. However, they face challenges such as communication overhead and complexity in ensuring system-wide safety guarantees in large-scale networks.

2.4. Safety analysis and risk assessment frameworks

Several frameworks have addressed safety analysis and risk evaluation for autonomous train operations. Yan *et al.* [12] have proposed the use of the system-theoretic process analysis (STPA) method to analyze the safety of a fully automatic operation (FAO) system scenario, which represents a continuous control system. During the operation process, the internal and external elements of the system interact continuously, and then the system sends control commands to the train. Train operation scenarios describe the interaction between subsystems as a whole and ensure the safe operation of trains through the simultaneous and synchronous interaction of multiple systems. Djordjevic *et al.* [13] have developed a three-step approach for studying the determining factors and constraints for the adoption of automatic train operation (ATO) systems. The first step is based on an open Delphi questionnaire to gather expert knowledge on the challenges, risks, advantages, and critical subsystems of ATO in rail freight operations. In the second step, they have applied analytic network process (ANP) technique to develop and prioritize the relationship between key determinants in the analysis of the deployment of ATO. Finally, the authors have attempted to predict the costs and benefits of each level of automation grade of automation (GOA) in rail freight transport based on expert knowledge. Chelouati *et al.* [14] have proposed an approach to map modeling or embedded mapping based on various standards that include information on infrastructure, signage, and connections. Safety rules are integrated into the conceptual

map model, with the aim of automating their integration on board and the management of safety decisions. Naufal *et al.* [15] have proposed A2CPS, a vehicle-centric safety conceptual framework for autonomous transport systems. This approach aims to ensure the safety of autonomous vehicles by minimizing the risk of collisions. It is based on an adapted architecture from a similar transport domain. A2CPS combines the principles of cyber-physical systems (CPS) and intelligent transportation systems (ITS) in the context of road transportation systems. While these approaches provide structured safety analysis and adaptive risk assessment, they often require time-intensive expert processes, face subjectivity issues, or demand significant computational resources for real-time implementation.

2.5. Advanced control and resource management systems

Recent works have focused on intelligent control architectures for autonomous train operations. Song *et al.* [16] have provided a perception of information about the fixed infrastructure and the internal and external environment. This system uses resources to optimize the use of ground equipment. The operating scenarios of the system can be divided into interrogation, locking, occupation, and release of resources, which are similar to the traditional scenario. The autonomous train control system (ATCS) will be equipped with an on-board locking unit that analyzes data on train position, track occupancy, and movement authority. These control-oriented architectures demonstrate potential for resource optimization but require strict standardization and strong interoperability with heterogeneous data sources.

2.6. Synthesis

Our analysis of the reviewed works shows that a wide range of approaches have been investigated to enhance safety in autonomous train operations. Vision-based techniques combined with deep learning provide high accuracy in detecting railway signals, track obstacles, and other safety-critical elements. These methods offer essential perception capabilities for real-time decision-making, although their performance remains dependent on the quality and diversity of training datasets, which may limit their robustness in degraded or unexpected conditions [7]. In parallel, formal verification methods, particularly those relying on Petri nets, have been employed to ensure correctness in control logic and compliance with safety properties. While offering rigorous validation, these methods often introduce high computational demands and show limited adaptability to dynamic operating environments, raising challenges for large-scale deployment [8]. Other works have focused on decision-making and operational management. Multi-agent system architectures demonstrate adaptability and coordination between distributed agents, such as avoiding unnecessary stops in bidirectional crossings. Expert-driven approaches, including Delphi and ANP, provide structured ways of identifying risks, benefits, and adoption challenges of ATO. System-theoretic approaches like STPA allow safety analysis by modeling complex system interactions. However, these techniques face drawbacks such as communication overhead, subjectivity, time-intensive processes, or scalability issues when applied to complex railway infrastructures [11–13]. Finally, several contributions have addressed infrastructure and system-level safety. Embedded mapping approaches integrate safety rules into conceptual infrastructure models, hardware-based prototypes provide proof-of-concept validations in miniature train systems, and control-oriented architectures such as ATCS aim to optimize resource usage with onboard safety units. In addition, resilience-oriented frameworks like

dynamic risk assessment (DRA) and A2CPS enhance situational awareness and adaptive risk assessment. Despite their merits, these solutions require significant computational resources, strict standardization, and strong interoperability with heterogeneous data sources [9,10,14–16].

Based on these observations, the literature reviewed demonstrates the richness and diversity of approaches addressing autonomous train safety. Yet, no single methodology fully satisfies the stringent requirements of adaptability, scalability, and operational feasibility. This gap motivates our proposal, which aims to integrate the complementary strengths of perception accuracy, formal verification reliability, and decision-making into a unified framework capable of addressing the safety-critical needs of future autonomous railway systems.

Table 1 summarizes the reviewed approaches.

Table 1. Comparative summary of safety approaches for autonomous train operation.

Paper	Method	Objective	Main Contribution	Performance
Etxeberria-Garci <i>et al.</i> [7]	Computer vision + Deep Learning	Reliable detection of railway signals and decision-making for train stopping	Trained neural networks on diverse railway scenarios	Detection accuracy, decision latency, dependency on training data
Xie <i>et al.</i> [8]	WFN	Safety verification of railway automation logic for autonomous train operation	Formal modeling of control functions with code generation	Formal verifiability, safety assurance, limited adaptability to unexpected cases
Dordal <i>et al.</i> [11]	Multi-agent system simulation	Avoiding unnecessary stops at bidirectional crossings in autonomous train routes	Intelligent coordination among agents via message exchange	Reduced delays, improved traffic flow, scalability with multiple agents
Djordjevic <i>et al.</i> [13]	Delphi method + ANP + cost-benefit analysis	Assessing determinants and constraints for adoption of autonomous train operation (ATO) systems	Prioritization of key factors and economic feasibility evaluation	Cost efficiency, operational feasibility, dependency on expert judgment
Yan <i>et al.</i> [12]	STPA	Safety analysis of fully automated train operation scenarios	Identification of interactions among subsystems for risk prevention	Comprehensive hazard coverage, proactive failure detection
Chouchani <i>et al.</i> [9]	Embedded cartographic modeling	Integration of safety rules into on-board decision-making for autonomous trains	Standard-based map model including infrastructure and signals	Robustness in complex environments, automated safety enforcement
Mohammed <i>et al.</i> [10]	Hardware prototype (PIC microcontroller, Proteus, PCB)	Demonstration of driverless train prototype with basic automation	Miniature autonomous train with predefined trajectory, stops, and passenger alerts	Low-cost validation, proof-of-concept, limited scalability
Song <i>et al.</i> [16]	Autonomous Train Control System (ATCS)	On-board decision-making to optimize infrastructure and environment management	On-board locking unit analyzing position, track occupancy, and movement authority	Resource optimization, compatibility with traditional operation modes
Chelouati <i>et al.</i> [14]	DRA framework	Real-time adaptation and anticipation of risks for autonomous train driving	Perception, prediction, and decision-making modules with online DRA layer	Resilience, adaptability to unforeseen scenarios, responsiveness
Naufal <i>et al.</i> [15]	A2CPS (Vehicle-Centric Safety Framework)	Ensuring collision-free autonomous train operation by adapting concepts from intelligent transport systems	Cross-domain CPS and ITS integration applied to railway context	Collision risk minimization, cross-domain applicability, integration potential

3. The proposed solution

In this section, we present the proposed solution and provide a detailed description of each phase.

3.1. Overview

The proposed autonomous train architecture is designed to address the challenges of urban railway transportation. It integrates perception capabilities, artificial intelligence, and autonomous decision making capabilities. Unlike conventional systems, this approach can function autonomously, independently of communications with the infrastructure. This allows the system to dynamically adapt to complex environmental conditions.

An overview of our architecture is shown in Figure 1. This architecture is structured in four distinct phases. The first phase is the intelligent perception, in which multisensor acquisition and fusion are used to understand the environment comprehensively. The second phase is the specialized detection, in which artificial vision is adapted to precisely identify critical entities. The third phase is the risk assessment, in which intelligent threat quantification is achieved using deep learning. The fourth and final phase is the autonomous decision, in which the generation of control commands is achieved through automata-based modeling. Mutual influence and bidirectional integration among the three components (computer vision, AI-based risk assessment, and adaptive decision-making) allow a high level of safety and reactivity to be achieved. Table 2 summarizes the interaction mechanisms.

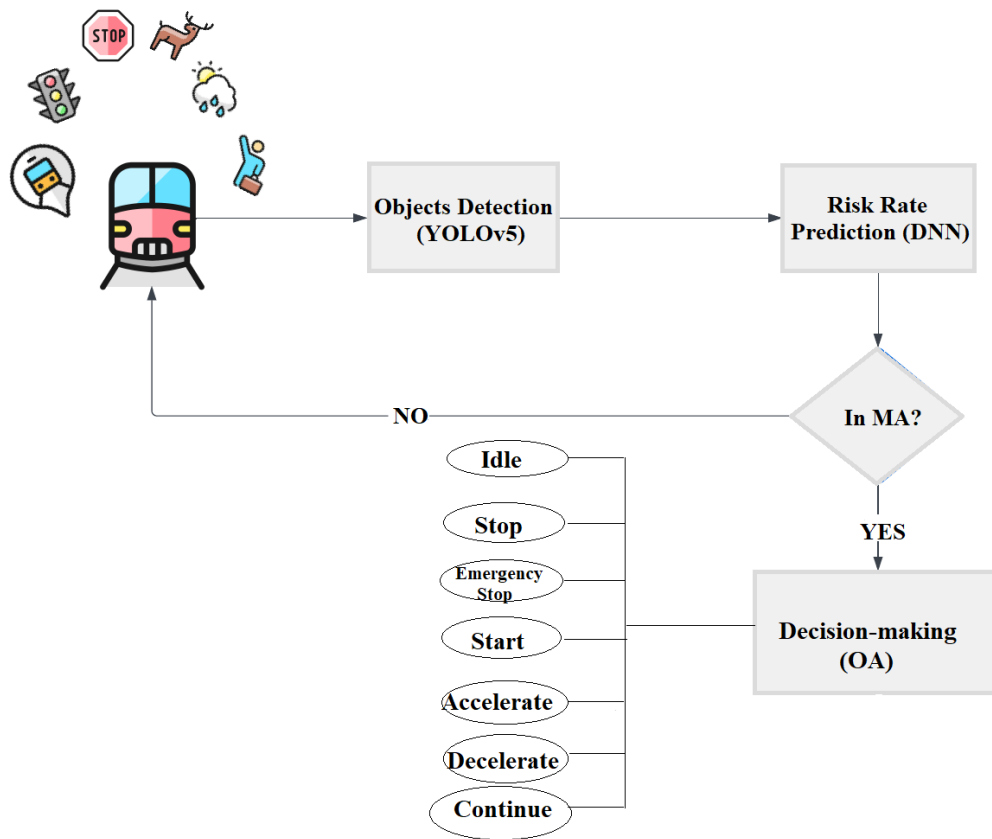


Figure 1. Overview of the proposed architecture.

Table 2. Mutual influence mechanisms among system components.

Influence Path	Information Flow	System Impact
Vision → Risk Assessment	Object type & position	Determines base risk level
	Detection confidence	Adjusts risk uncertainty factor
	Multi-object count	Triggers combinatorial risk
Risk Assessment → Decision Making	SIL-1/2	Normal/Cautious operation
	SIL-3	Alert mode, speed reduction
	SIL-4	Emergency stop activation
Decision Making → Vision (Feedback)	Risk state elevation	Lower confidence threshold
	Emergency mode	Expand detection
	Alert state	Increase processing frequency
	Speed reduction	Activate temporal tracking
Decision Making → Risk Assessment	Current action state	Contextualizes risk evaluation
	Speed & trajectory	Adjusts proximity thresholds

3.2. Environment perception phase

The perception subsystem enables the complete real-time understanding of the train's environment, rather than relying on infrastructure-to-train communications, as is used in conventional control systems. Cameras have been used to detect critical elements, such as pedestrians at level crossings, vehicles, road and rail signals, and obstacles. In addition, a rotating light detection and ranging (LiDAR) system has been implemented to provide precise three-dimensional mapping of the immediate environment and long-range detection of moving objects. For autonomous localization, a hybrid localization system that combines high-precision global navigation satellite system (GNSS), an inertial measurement unit, and odometers have been utilized to ensure reliable localization, despite adverse conditions such as, tunnels and bridges.

Multi-sensor fusion represents a major improvement over traditional systems that process information sequentially. In our approach, we propose to simultaneously integrate all sensory modalities to construct a coherent representation of the environment. The fusion process weights each sensor's contributions according to their contextual reliability, precision, and availability.

3.3. Object detection phase

After collecting data on its environment, the system then proceeds to the second phase, which consists of object detection. This process is based on the YOLOv5 algorithm [17], which was previously trained in offline mode.

The key parameters required for object detection are produced in this step by using a convolutional neural network (CNN). The CNN has been optimized for railway applications using transfer learning. It starts with pre-trained weights and adapts the final layers to railway-specific object classes. For each region of the image, the CNN predicts the probability that an object is present for each class, as well as the location coordinates of potential bounding boxes (b) surrounding these objects in each grid cell.

Let $\{c_0, \dots, c_{11}\}$ the set of objects classes to detect, where $c_0 = \text{car}$, $c_1 = \text{truck}$, $c_2 = \text{cyclist}$, $c_3 = \text{pedestrian}$, $c_4 = \text{traffic light}$, $c_5 = \text{green signal}$, $c_6 = \text{green signal left}$, $c_7 = \text{red signal}$, $c_8 = \text{red signal left}$, $c_9 = \text{yellow signal}$, $c_{10} = \text{yellow signal left}$, $c_{11} = \text{train}$.

For each predicted bounding box b , confidence scores $Trust(c_i|b)$ are estimated for each of the 12 classes:

$$Trsut(c_i|b) = Pr(c_i|Obj) * Pr(Obj) * IOU(b, real Obj) \quad (1)$$

where,

- $Pr(c_i|Obj)$ represents the conditional probability that the detected object belongs to class i given that an object is present.
- $Pr(Obj)$ indicates the probability that an object is actually present in bounding box b .
- $IOU(b, real Obj)$ is the Intersection over Union ratio between the predicted bounding box b and the target bounding box.

After performing the inference on an image, a redundancy elimination step is applied to remove boxes that detect the same object. Only boxes with the highest confidence scores are maintained. These boxes constitute the final object detections, along with their predicted class and precise location.

3.4. Risk rate prediction phase

In this phase, we use a deep neural network to calculate the risk estimation, given its ability to process complex and nonlinear data. The network has been trained using labeled dataset. The inputs are features extracted from YOLOv5, and the outputs are risk estimations. The network architecture consists of an input layer that receives object detection information from the previous step. This is followed by two hidden layers that perform internal processing and an output layer that produces the final prediction in the form of an estimated risk rate.

The network produces for each detected obstacle an estimated risk score ranging from 0 (no risk) to 1 (critical risk). This quantification is based on an adaptation of SIL [18] used in railway safety standards (see Algorithm 1). The risk score strategy follows the thresholds inspired by SIL:

Algorithm 1 Risk level prediction

Input: Detection set $D = \{d_1, d_2, \dots, d_{11}\}$

Output: Risk level $R = \{r_1, r_2, \dots, r_{11}\}$

foreach detection $d_i \in D$ **do**

$r_i \leftarrow DNN_RiskPrediction(d_i);$

if $r_i \in [0.0, 0.25]$ **then**

$risk_level \leftarrow 1;$

else

if $r_i \in [0.25, 0.50]$ **then**

$risk_level \leftarrow 2;$

else

if $r_i \in [0.50, 0.75]$ **then**

$risk_level \leftarrow 3;$

else

$risk_level \leftarrow 4;$

end

end

end

return $risk_level;$

end

- Negligible risk (0.0–0.25): Corresponds to SIL-1 scenarios where basic safety functions are sufficient. Objects are at safe distances with low probability of collision under normal operating conditions.
- Moderate risk (0.25–0.50): Aligns with SIL-2 requirements for enhanced safety monitoring. Situations include moving objects at intermediate distances or mildly degraded operational conditions that require increased vigilance.

- High Risk (0.50–0.75): Matches SIL-3 safety standards for critical safety functions. Scenarios involve potential collision trajectories, restrictive signaling, or environmental conditions that significantly affect braking performance.
- Critical Risk (0.75–1): complies with the SIL-4 emergency safety requirements with the highest integrity levels. These include imminent collision threats, signal violations, or obstacles directly on the train path that require immediate emergency intervention.

3.5. Decision-making phase

After the risk rate prediction phase, we model decision-making using Open Automata (OA) [19]. These are automata with partial specifications, deriving their partiality from “holes” that act as interface ports. These ports enable the systematic integration of additional OA elements to complete the system specification, while variables enable symbolic representation and modular definition of partial behaviors in complex systems.

Definition 1 (Open Automaton): An OA is a tuple with S a set of states, $s_0 \in S$ the initial state, V the finite set of variable names, $\sigma_0 : V \rightarrow$ the initial valuation of variables, J the set of hole names and T the set of open transitions.

An open transition is a structure made of several composing entities, equivalent to a tuple. In an open transition $s, s' \in S$ are the source and target states, $\alpha \in$ is the resulting action that can be observed from the outside, $J' \subseteq J$ are the holes involved in the transition, $g \in$ is the guard that may constraint the transition, and $\psi : V \rightarrow$ are the variable assignments that have an effect on the state of the automaton. Each $\beta_j \in$ is an action of the holes j .

Modelling the decision making phase: We now define the OA for our decision-making system. First, we specify a function for calculating train speed deceleration based on the risk level (see Algorithm 2) and the variables involved in the automaton. The variables are listed below.

- (1) $risk_level \in \{SIL_1, SIL_2, SIL_3, SIL_4\}$ that specifies the risk of collision and is updated by the previous phase.
- (2) $light \in \{\text{“red”}, \text{“yellow”}, \text{“green”}\}$ that specifies the values of the traffic lights.
- (3) $position \in \{\text{“InStation”}, \text{“OnTrack”}\}$ that specifies if the train is in the station or on track.

Algorithm 2 Adjusted_speed process

Input: risk_level (SIL0, SIL1, SIL2, SIL3, SIL4), speed

Output: target_speed

```

switch risk_level do
  case SIL0 do
    | return Speed;
  end
  case SIL1 do
    | return Speed × 0.8;
  end
  case SIL2 do
    | return Speed × 0.6;
  end
  case SIL3 do
    | return Speed × 0.3;
  end
  case SIL4 do
    | return 0;
  end
end
end

```

- Open automata of risk-aware predictive system: The decision-making phase relates to the risk prediction through an Open Automaton (*cf.*, Figure 2) that acts as a bridge between the neural network’s risk assessment and the train’s control system. This automaton operates with a two-state structure: it begins in an idle state (I) where the system waits for incoming data, performing a continuous monitoring loop (θ action) until new information arrives. When risk prediction data becomes available, the automaton transitions to the decision state (D) via the ReceiveData() action, which occurs unconditionally (*True*). The system receives three essential parameters from the risk prediction phase: the current light signal status (*lt*), the calculated risk level (*risk_l*), and the train’s position (*pos*). These inputs are processed through the Update(*lt*, *risk_l*, *pos*) function, which updates the automaton’s internal variables to reflect the current operational context. This data will be the inputs for the train’s decision-making process.

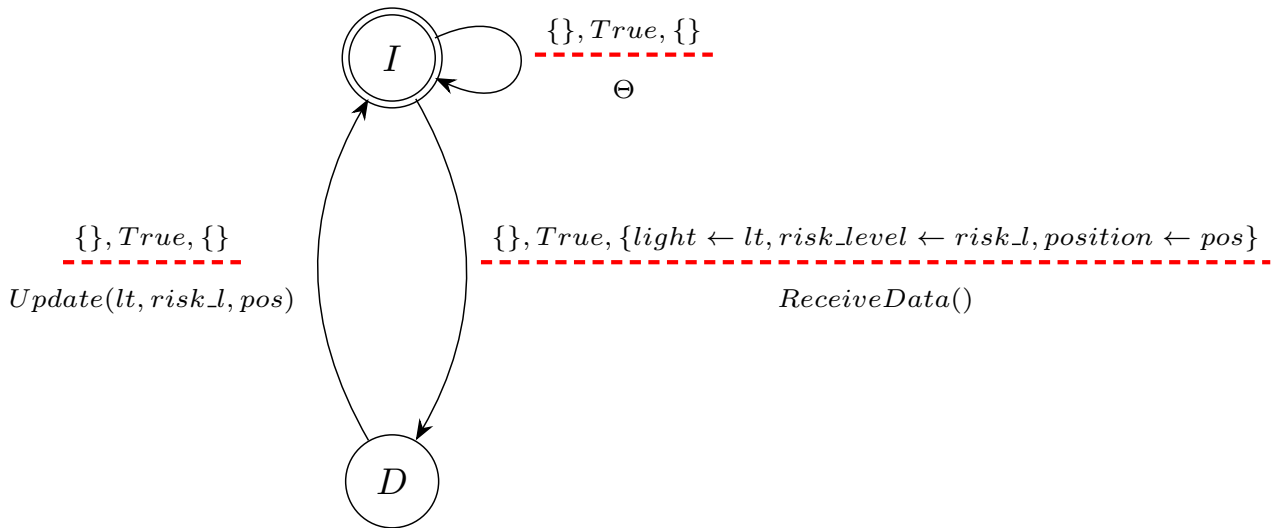


Figure 2. Open automata of the risk-aware predictive system.

- Open automata of the train’s control system: The OA depicted in Figure 3 represents the train’s control system. This system has two states: the stopped state (*S*) where the train remains stationary, and the moving state (*M*), where the train is in motion. We can notice on each transition the presence of a systematic hole (*pred*) receiving updates of the risk parameters. This hole serves as a dedicated interface that connects the train’s control system to the risk predictive system (Figure 2), allowing the control automaton to receive the essential risk parameters necessary for decision-making.

From the initial state (*S*), we can see that the speed is initialized to zero(0) the system can perform two actions:

- (a) The action *Idle* means that the train remains in the same stopped state when the lights are not green or the risk level is high (SIL₂ or higher).
- (b) The action *Start* leads to the state *M* (moving) when the lights are green and there is no risk or minimal risk. Hence, the speed is increased and the position of the train is updated to “On Track”.

From the Moving state (*M*) there are five possible actions. The first three transitions keep the train in movement.

- (a) The *Continue* action maintains the train at constant speed, executed when the traffic lights are green, the train has reached cruising speed, and the risk level is either absent or minimal.
- (b) The *Accelerate* action increases the train’s speed by a value “*i*”, executed when the lights are green, the train has not reached the cruising speed and the risk level is either absent or minimal.
- (c) The *Decelerate* action decreases the train’s speed based on the risk level ($speed \leftarrow Adjust_speed(speed, risk_level)$). The train decelerates when the lights are not red and the risk level is medium or high.

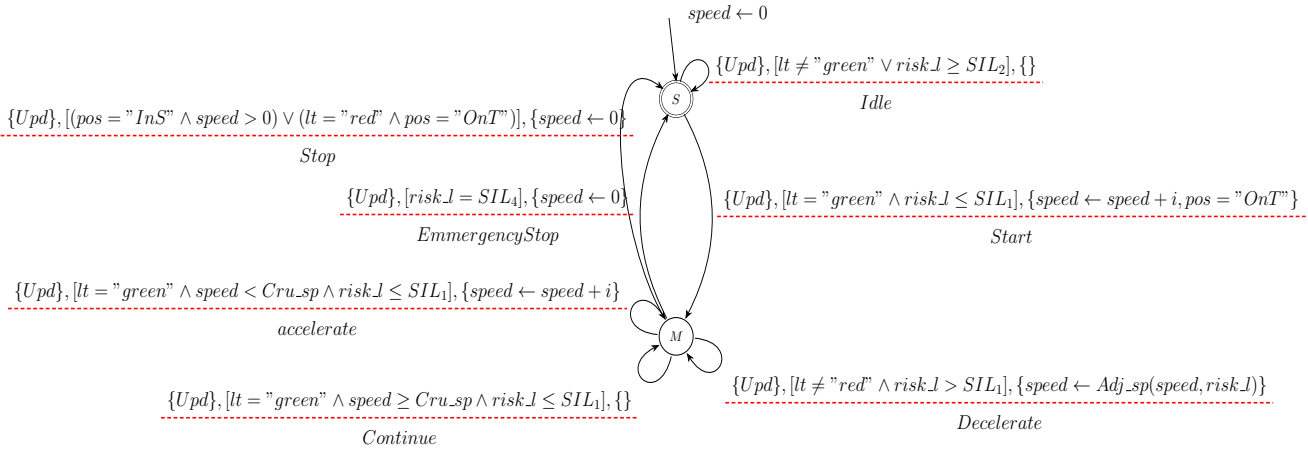


Figure 3. Open automata of the train’s control system.

Finally, two actions lead from the state *M* to *S* :

- (a) The *Stop* action sets the speed to zero. This action is performed when the train reaches the station or when the lights are red and it is on track.
- (b) The *EmergencyStop* action also sets the speed to zero. It is performed when there is a high risk (SIL_4) of collision.

Table 3 summarizes the abbreviations used.

Table 3. Abbreviations.

Abbreviation	Meaning
Upd	prd=Update(lt, risk_level, pos)
OnT	OnTrack
InS	InStation
risk_l	risk_level
Cru_sp	Cruising_speed
Adj_sp	Adjusted_speed
lt	light (traffic signal)

4. Performances evaluation

To demonstrate the effectiveness of the proposed system, we conduct a comprehensive experimental evaluation focusing on both individual component performance and overall system behavior.

4.1. Experiment setting

To evaluate our approach, we have constructed a composite dataset by merging two distinct sources, specifically tailored to our research focus on railway safety. The first dataset is the Udacity Self-Driving Car dataset [20], which contains 15,000 images with 97,942 labels across 11 classes. This multimodal dataset includes images (1920×1200 pixels), videos, LiDAR sensor readings, global positioning system (GPS) information, and bounding box annotations for object detection. The second dataset, Railway Track Fault Detection [21], consists of 383 images (191 defective and 192 non-defective tracks), annotated to indicate faults and anomalies. These datasets were preprocessed through the following steps: (1) normalization of pixel values, (2) annotation format standardization to You Only Look Once (YOLO) format, and (3) consolidation into a unified directory structure with separate folders for images and corresponding annotation files. The combined dataset, totaling 15,383 images, was subsequently divided into training (70%, 10,768 images), validation (20%, 3077 images), and test (10%, 1538 images) subsets using stratified random sampling to ensure robust evaluation and prevent data leakage.

The experiments were implemented in Python language. The evaluation process relies on performance metrics including real-time response, accuracy, precision, recall, and F1-score. Table 4 summarizes the dataset characteristics, and the analysis parameters.

Table 4. Dataset characteristics and experimental parameters.

Parameter	Value
Dataset Composition	
Total Images	15,383
Training Set	10,768 (70%)
Validation Set	3077 (20%)
Test Set	1538 (10%)
Image Resolution (preprocessed)	640×640 pixels
Experimental Parameters	
Threshold Range Analyzed	0.1–0.5
Number of Thresholds Tested	5
Optimal Threshold Range	0.3–0.4
Model Variant	YOLOv5
Number of Epochs	10

4.2. Results analysis

Figure 4 represents the precision-recall curve for the YOLOv5 model training phase, demonstrating the trade-off relationship between precision and recall metrics across different detection thresholds. The curve exhibits a positive correlation with precision values ranging from 0.85 to 0.9 corresponding to recall values between 0.74 and 0.87 cross five measurement points. This upward trajectory indicates that the adapted YOLOv5 model successfully balances the detection of true positive railway objects while maintaining low false positive rates. The concentrated data points in the high-precision, high-recall region (precision > 0.87 , recall > 0.80) demonstrate the model's effectiveness in railway-specific object detection, suggesting robust performance for safety-critical applications where both missed detections and false alarms must be minimized.

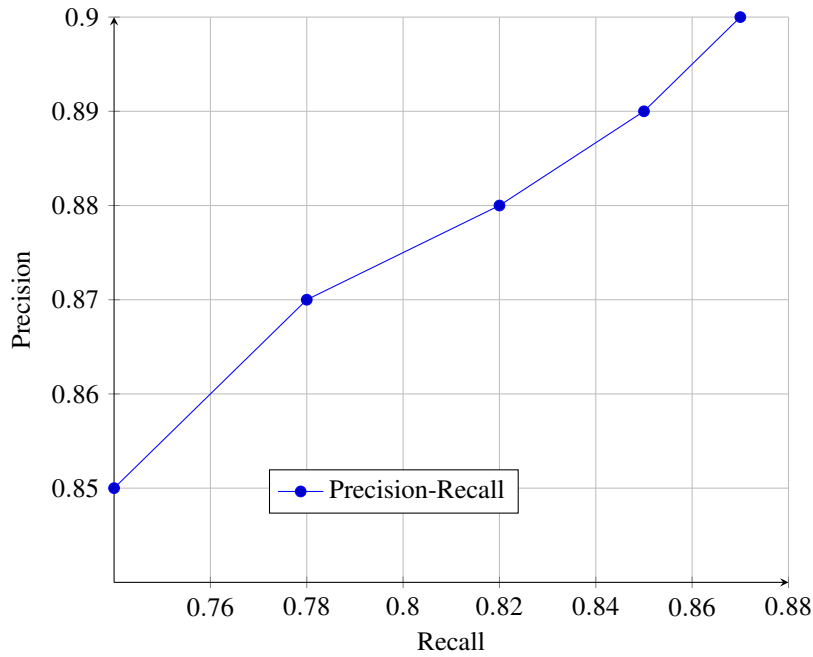


Figure 4. Results of the YOLOv5 model training.

Figure 5 illustrates the performance of the YOLOv5 model during the testing phase on the test set comprising 1538 images (10% of the total dataset of 15,383 images). The relatively high accuracy of 89% demonstrates the model’s overall correctness in classification decisions across all railway object categories. The balanced performance between precision 80% and recall 87% indicates that the model effectively detects railway objects while maintaining acceptable false positive rates. The F1-score of 87% confirms the harmonic balance between precision and recall, validating the model’s reliability for the combined railway safety dataset containing both self-driving car scenarios and railway track fault detection cases.

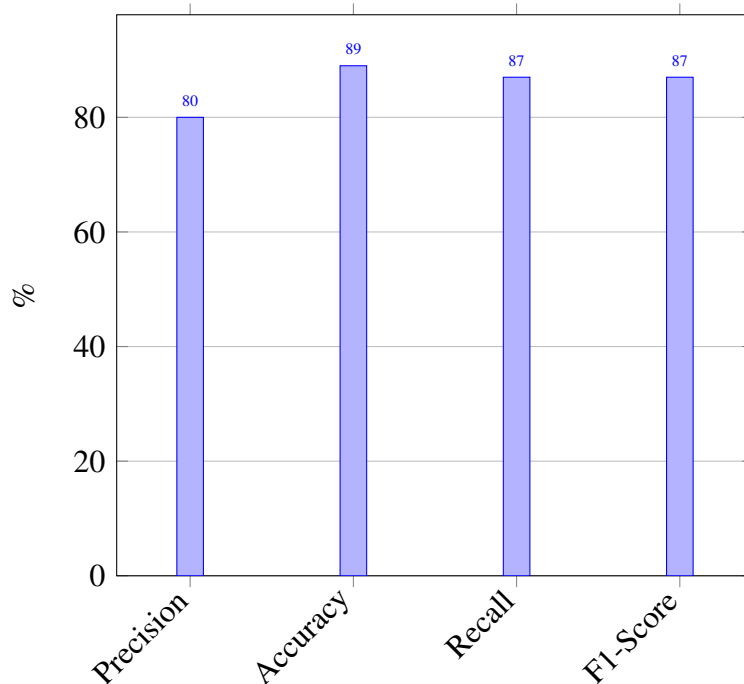


Figure 5. Results of the YOLOv5 model testing.

Figure 6 represents the performance evolution of the YOLOv5 algorithm across different thresholds ranging from 0.1 to 0.5, demonstrating the expected trade-off behavior between precision and recall metrics evaluated on the 1538 test images. The precision curve shows a steady improvement from 0.76 to 0.87 as thresholds increase, reflecting the model's enhanced ability to reduce false positive detections with stricter confidence requirements. In contrast, the recall metric shows the anticipated decrease from 0.89 to 0.82, indicating that higher thresholds result in filtered out some true positive detections. The F1-score maintains relatively stable values between 0.82-0.84, with a slight peak at threshold 0.3-0.4, suggesting optimal balance between precision and recall in this range. The accuracy metric shows consistent improvement from 0.86 to 0.89, demonstrating enhanced overall classification performance with higher confidence requirements. Five threshold values (0.1, 0.2, 0.3, 0.4, 0.5) were selected to analyze the model's behavior across a representative range, enabling identification of the optimal operating point for railway safety applications.

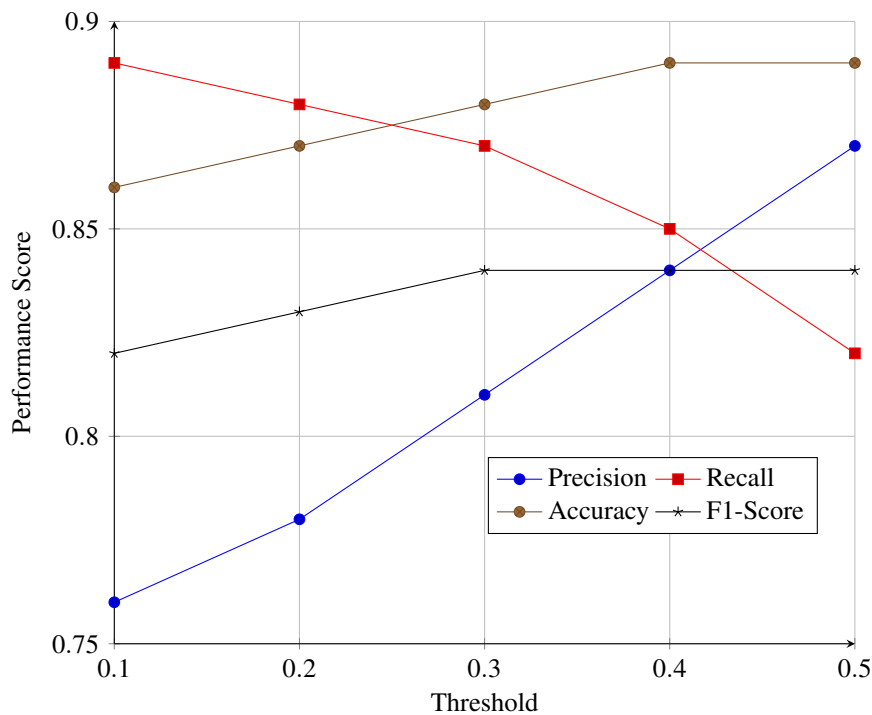


Figure 6. YOLOv5 Performance evolution across confidence thresholds.

Figure 7 illustrates the performance of the proposed system in function of key performance metrics evaluated on the complete test set of 1538 images. The system achieves 89% detection accuracy, demonstrating reliable object recognition capabilities for railway environments. Classification precision reaches 80%, indicating effective discrimination between different object classes critical for safety assessment. Real-time processing achieves 100% compliance with the 50 ms latency requirement essential for railway safety applications. The overall system performance of 87% validates the effectiveness of the system as an integrated solution for intelligent autonomous railway operation, combining perception, risk assessment, and decision making. These metrics were measured across the entire test set, which includes diverse railway scenarios from both the Udacity Self-Driving Car dataset and the Railway Track Fault Detection dataset, ensuring comprehensive evaluation under varied conditions.

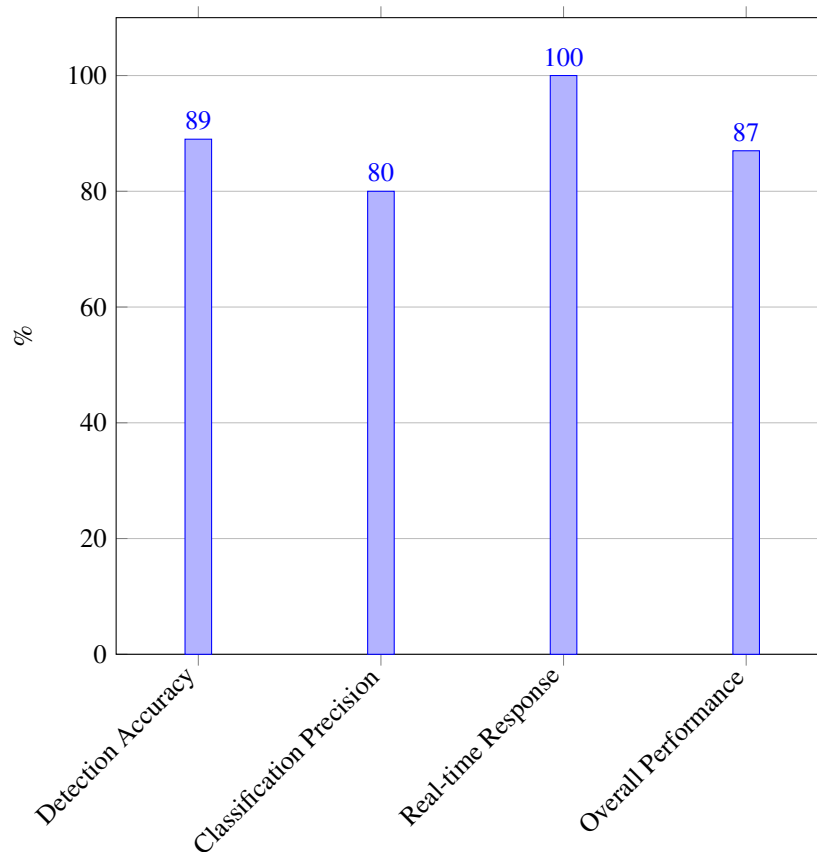


Figure 7. Performance of the proposed system.

5. Conclusion

This paper have presented an intelligent system for autonomous railway operation that addresses critical limitations of existing railway automation systems. The primary contributions include the development of railway-specific YOLOv5 adaptation optimized for railway environments, creation of an intelligent risk assessment framework calibrated against Safety Integrity Levels, and design of an integrated architecture combining autonomous perception with adaptive decision-making while maintaining compatibility with existing railway infrastructure. Experimental results confirm the system's effectiveness, achieving expert-level risk assessment accuracy while maintaining rapid response times essential for safe autonomous operation.

Although the proposed system demonstrates promising performance under standard conditions, several limitations must be acknowledged. First, the system's performance under adverse weather conditions (heavy rain, fog, snow) has not been extensively evaluated, as these scenarios may significantly degrade visual perception quality and affect object detection accuracy. Second, reliance on camera-based detection makes the system vulnerable to occlusion scenarios, such as when objects are partially hidden by infrastructure elements, vegetation, or other trains. Third, the current implementation does not include redundancy mechanisms for sensor failure scenarios, which could compromise system reliability in safety-critical railway operations.

Future research directions will focus on several key areas. First, we plan to evaluate and enhance the system's robustness under adverse weather conditions through the integration of multi-modal sensing approaches, combining camera data with LiDAR and radar to maintain performance in low-visibility

scenarios. Second, we will develop occlusion handling mechanisms using temporal information and predictive tracking algorithms to maintain object awareness even when targets are temporarily obscured. Third, implementing sensor fusion and redundancy strategies will be explored to ensure system reliability during sensor failures or degraded sensor performance.

Data availability statement

The datasets analyzed in this study are publicly available. The Udacity Self-Driving Car dataset is available at <https://github.com/udacity/self-driving-car>, and the Railway Track Fault Detection dataset is available at <https://www.kaggle.com/datasets/salmaneunus/railway-track-fault-detection>.

Declaration of generative AI and AI-assisted technologies

During the preparation of this manuscript, the authors used AI-assisted tools (deepL translate, grammarly) to improve language and readability. No AI tool was used for content generation, drafting, or idea development. The authors take full responsibility for the content of the manuscript.

Acknowledgments

This work was carried out in the framework of the research activities of the LIMED laboratory, university of Bejaia and IRISA laboratory, university Bretagne sud, France. This work has been sponsored by the General Directorate for Scientific Research and Technological Development, Ministry of Higher Education and Scientific Research (DGRSDT), Algeria.

Authors' contribution

Conceptualization, D.Z., Mo.O. and L.M.; methodology, Mo.O. and L.M.; software, Mo.O. and L.M.; validation, D.Z., Mo.O. and L.M.; formal analysis, D.Z., S.C.M., Mo.O. and L.M.; investigation, Mo.O. and L.M.; writing—original draft preparation, D.Z. and Ma.O.; writing—review and editing, D.Z., S.C.M. and Ma.O.; supervision, D.Z. and Ma.O.; funding acquisition, D.Z. All authors have read and agreed to the published version of the manuscript.

Conflicts of interest

Mawloud Omar holds the position of Associate Editor for *Artificial Intelligence and Autonomous Systems* and has not peer reviewed or made any editorial decisions for this paper.

References

- [1] Cui Y, Martin U, Zhao W. Review of safety and security of cyber-physical systems in transportation. *Transp. Res. Rec.* 2019, 2673(4):1–14.
- [2] Dhillon BS. *Human Reliability, Error, and Human Factors in Engineering Maintenance: with Reference to Aviation and Power Generation*, 1st ed. Boca Raton: CRC Press, 2007.

- [3] Zhao L, Xu T, Li K, Chen L. Survey on automatic train control systems. *IEEE Trans. Intell. Transp. Syst.* 2017, 18(7):1854–1871.
- [4] Redmon J, Divvala S, Girshick R, Farhadi A. You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, Las Vegas, USA, June 27–30, 2016, pp. 779–788.
- [5] Berg A, Öfjäll K, Ahlberg J, Felsberg M. A detecting rails and obstacles using a train-mounted thermal camera. In *Proceedings of 19th Scandinavian Conference*, Copenhagen, Denmark, June 15–17, 2015, pp. 430–441.
- [6] Zhang C, Yao X, Zhang J, Jin H. Tool wear prediction in milling operations using machine learning techniques. *J. Manuf. Processes.* 2019, 37:487–494.
- [7] Etxeberria GM, Labayen M, Zamalloa M, Arana AN. Application of computer vision and deep learning in the railway domain for autonomous train stop operation. In *Proceeding of 2020 IEEE/SICE International Symposium on System Integration*, Honolulu, USA, January 12–15, 2020, pp. 943–948.
- [8] Xie Y, Khlif BM, Toguyéni A. Well-formed Petri net-based pattern for modeling logic controllers for autonomous trains. In *Proceedings of the 10th International Conference on Integrated Modelling and Analysis in Applied Control and Automation*, Barcelona, Spain, September 18–20, 2017, pp. 25–34.
- [9] Chouchani N, Debbech S, Perin M. Model-based safety engineering for autonomous train map. *J. Syst. Softw.* 2022, 183:111082.
- [10] Mohammed TS, Al-Azzo WF, Akaak MA, Suroor ML. Full automation in driverless trains: a microcontroller-based prototype. *Int. J. Adv. Res. Electr. Electron. Instrum. Eng.* 2014, 3(7):10417–10422.
- [11] Dordal OB, Ávila BC. Autonomous driving trains to pass in bidirectional crossing loop preventing stops. *Procedia Comput. Sci.* 2016, 96:197–206.
- [12] Yan F, Zhang S, Tang T. Autonomous train operational safety assurance by accidental scenarios searching. In *Proceeding of 2019 IEEE Intelligent Transportation Systems Conference*, Auckland, New Zealand, October 27–30, 2019, pp. 3488–3495.
- [13] Djordjević B, Fröidh O, Krmac E. Determinants of autonomous train operation adoption in rail freight: knowledge-based assessment with Delphi-ANP approach. *Soft Comput.* 2023, 27(11):7051–7069.
- [14] Chelouati M, Boussif A, Beugin J, Koursi EE. A framework for risk-awareness and dynamic risk assessment for autonomous trains. In *Proceedings of the 32nd European Safety and Reliability Conference*, Dublin, Ireland, August 28–September 1, 2022, pp. 2128–2135.
- [15] Naufa, JK, Camargo JB, Vismari LF, de Almeida JR, Molina C, *et al.* A²CPS: a vehicle-centric safety conceptual framework for autonomous transport systems. *IEEE Trans. Intell. Transp. Syst.* 2017, 19(6):1925–1939.
- [16] Song H, Gao S, Li Y, Liu L, Dong H. Train-centric communication based autonomous train control system. *IEEE Trans. Intell. Veh.* 2022, 8(1):721–731.
- [17] Terven J, Córdova-Esparza DM, Romero-González JA. A comprehensive review of YOLO architectures in computer vision: from YOLOv1 to YOLOv8 and YOLO-NAS. *Mach. Learn. Knowl. Extr.* 2023, 5(4):1680–1716.

- [18] CENELEC. EN 50129:2018. Railway applications—communication, signalling and processing systems—safety related electronic systems for signalling. 2018. Available: <https://standards.iteh.ai/catalog/standards/clc/f6548cc3-5885-43aa-8654-9e71383b892e/en-50129-2018> (accessed on 20 June 2025).
- [19] Ameer-Boulifa R, Corradi Q, Henrio L, Madelaine E. Refinements for open automata. In *Proceedings of the 21st International Conference on Software Engineering and Formal Methods*, Eindhoven, The Netherlands, November 6–10, 2023, pp. 11–29.
- [20] Ghodsian B, Khademi M, Khodabandeh M. Self-driving car steering angle prediction: let transformer be a car again. *arXiv* 2022, arXiv:2204.12748.
- [21] Hafique R, Siddiqui HUR, Rustam F, Ullah S, Siddique MA, *et al.* A novel approach to railway track faults detection using acoustic analysis. *Sensors* 2021, 21(18):6221.