Techno Nusa Mandiri: Journal of Computing and Information Technology
Vol. 23, No. 1 March 2026 | DOI https://doi.org/10.33480/techno.v23i1.8131

AUTONOMOUS AND EXPLAINABLE DETECTION OF SUSPICIOUS BEHAVIORS IN CONNECTED VEHICLE ENVIRONMENTS THROUGH MULTI-SENSOR VISION

Gihonia Abraham Senghor[1*]; Mabela Matendo Makengo Rostin[2]; Masakuna Felicien[3]; Muluba Mfumudimbu Jireh Celeste[4]; Muhala Luhepa Blaise[5]

Computer Science[1,3,4,5]
Statistics and probability[2]
University of Kinshasa, Kinshasa, Democratic Republic of Congo[1,2,3,4,5]
https://www.unikin.ac.cd/[1,2,3,4,5]
senghor.gihona@unikin.ac.cd[1], rostin.mabela@unikin.ac.cd[2], masakunafelicien@gmail.com[3], celestin.muluba@unikin.ac.cd[4], blaise.muhala@unikin.ac.cd[5]
(*) Corresponding Author

Abstract— *The safety of connected and autonomous vehicles requires intelligent systems capable of detecting suspicious behaviors in real time while providing clear explanations to human operators. This paper presents an innovative framework for the autonomous and explainable detection of suspicious activities around connected vehicles, combining multi-sensor vision, multi-agent reinforcement learning (MARL), and explainable artificial intelligence (XAI). The system relies on lightweight deep learning models (YOLO-tiny, MobileNet) for perception, along with spatio-temporal reasoning to identify abnormal events such as prolonged parking, restricted area crossings, or the placement of suspicious objects. Cooperative decision-making between vehicles and roadside units (RSUs) is managed through MARL. In parallel, an XAI module generates visual and textual explanations to enhance transparency and user trust. The framework has been implemented and evaluated in simulation (CARLA, SUMO/Veins) and on embedded platforms (Jetson Nano/Orin). Results demonstrate an F1-score of 0.91, real-time performance at 7.5 FPS, and a 40% reduction in false positives, confirming the robustness of the proposed system for the cyber-physical security of intelligent transportation systems.*

Keywords: *Connected Vehicles, Explainable Artificial Intelligence, Real-Time Embedded Vision, Reinforcement Learning, Suspicious Behavior Detection.*

Intisari— *Keamanan kendaraan terhubung dan otonom memerlukan sistem cerdas yang mampu mendeteksi perilaku mencurigakan secara real-time sambil memberikan penjelasan yang jelas kepada operator manusia. Artikel ini mengusulkan kerangka kerja inovatif untuk deteksi otonom dan dapat dijelaskan terhadap aktivitas mencurigakan di sekitar kendaraan terhubung, dengan menggabungkan penglihatan multi-sensor, pembelajaran penguatan multi-agen (MARL), dan kecerdasan buatan yang dapat dijelaskan (XAI). Sistem ini mengandalkan model deep learning ringan (YOLO-tiny, MobileNet) untuk penginderaan, serta penalaran spasial-waktu untuk mengidentifikasi peristiwa abnormal seperti parkir yang terlalu lama, melintasi area terlarang, atau penempatan objek mencurigakan. Pengambilan keputusan kooperatif antara kendaraan dan unit tepi jalan (RSUs) dikelola melalui \MARL. Secara paralel, modul XAI menghasilkan penjelasan visual dan teks untuk meningkatkan transparansi dan kepercayaan pengguna. Kerangka kerja ini telah diimplementasikan dan dievaluasi dalam simulasi (CARLA, SUMO/Veins) dan pada platform tertanam (Jetson Nano/Orin). Hasil menunjukkan skor F1 sebesar 0,91, kinerja real-time pada 7,5 FPS, dan pengurangan 40% dalam false positives, yang mengonfirmasi ketahanan sistem yang diusulkan untuk keamanan siber-fisik sistem transportasi cerdas.*

Kata Kunci: *Kendaraan Terhubung, Explainable Artificial Intelligence, Real-Time Embedded Vision, Pembelajaran Penguatan, Deteksi Perilaku Mencurigakan.*

INTRODUCTION

The digital transformation of the transport sector has led to the emergence of connected and

intelligent vehicles, integrated into increasingly complex environments (Zhao et al., 2022). These vehicles, capable of interacting with each other (V2V) and with infrastructure (V2I), pave the way for safer and more efficient intelligent transport systems (ITS). However, this interconnection also creates new vulnerabilities, whether cyber (attacks on vehicle networks) or physical (suspicious behaviour in the vehicle's immediate environment). Ensuring reliable and explainable detection of anomalies is therefore a major challenge for user and infrastructure safety (Alonge et al., 2025).

Numerous studies have explored the detection of visual anomalies in surveillance or on-board videos, using deep learning methods such as CNNs, autoencoders and Transformers (Rezaei & Azarmi, 2023). Other approaches have sought to improve robustness through statistical methods or semi-supervised learning. However, several limitations remain: most existing models are either too resource-intensive for on-board deployment or lack explanatory capabilities, which limits their adoption in critical environments where trust and interpretability are essential. Furthermore, few approaches incorporate multi-agent cooperation, which would allow vehicles and road units to intelligently share their observations to improve overall reliability (Grace et al., 2024).

To address these challenges, this article proposes an innovative framework for autonomous and explainable detection of suspicious behaviour around connected vehicles, combining three complementary approaches (Wang et al., 2024):

1. Multi-sensor vision (cameras, LiDAR, on-board sensors, V2X) using compact models (YOLO-tiny, MobileNet) to ensure real-time perception in on-board conditions (Almehdhar et al., 2024), (Dinneweth et al., 2022).
2. Multi-agent reinforcement learning (MARL) that enables vehicles and road units to cooperate to refine detection and reduce false positives (Dazeley et al., 2023).
3. An explainable artificial intelligence (XAI) module, capable of generating visual and textual explanations in real time, in order to increase the transparency of decisions and strengthen operator confidence (Dazeley et al., 2023), (Puder et al., 2022).

The approach is evaluated on simulated scenarios (CARLA, SUMO/Veins) and on embedded platforms (Jetson Nano/Orin), demonstrating superior performance to existing methods in terms of accuracy, speed and explainability.

Condensed State Of The Art

Embedded vision systems rely primarily on deep learning models for object detection and tracking (Bukola et al., 2024), (Alahdal et al., 2024). CNN architectures (such as YOLOv5/v7/v8, MobileNet or EfficientDet) are widely used for their effectiveness in visual recognition. However, large models pose challenges for deployment in embedded conditions due to memory, energy and latency constraints. Recent solutions therefore aim to design compact, quantised models capable of running on resource-constrained platforms (e.g. Jetson Nano). However, most approaches focus on detecting isolated objects rather than recognising complex suspicious behaviours involving spatio-temporal interactions.
Reinforcement learning (RL) has been successfully applied in various fields of robotics and autonomous driving (Dazeley et al., 2023), (Kiran et al., 2022), (Elallid et al., 2022). In the context of connected vehicles, multi-agent reinforcement learning (MARL) allows multiple entities (vehicles, road sensors, RSU units) to cooperate to make optimal decisions. Recent work focuses on coordination for traffic management, V2X communication optimisation, or trajectory planning (Cheng et al., 2022). Nevertheless, little research exploits MARL for cooperative anomaly detection or false positive reduction in vehicle physical safety. The integration of MARL into an embedded and constrained framework therefore remains an open scientific challenge (Al-Maamari et al., 2025).

The requirement for explainability has become essential in critical systems, particularly in safety and transport. Common XAI techniques include post-hoc approaches such as LIME, SHAP, or Grad-CAM, which highlight important regions of an image or the determining variables of a model. More recently, compact language models (LLMs), have been used to generate textual explanations that are understandable to the end user (Nwakanma et al., 2023), (Grace et al., 2024), (Zhao et al., 2022). However, in the field of connected vehicles, explainability remains largely unexplored: embedded systems often prioritise pure performance at the expense of transparency. Thus, there is still a lack of lightweight, real-time solutions capable of providing both visual and linguistic explanations that can be directly used by operators (Wang et al., 2024).

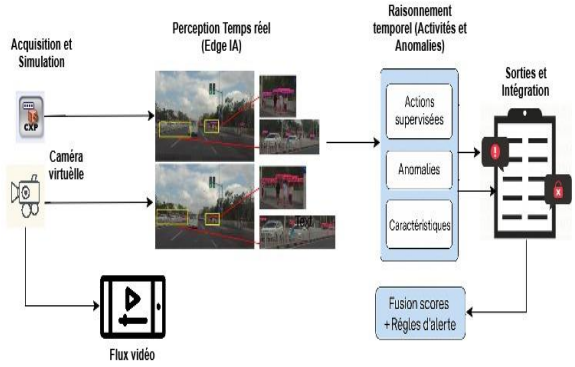
MATERIALS AND METHOD

Design of the proposed architecture

The first step was to design a modular architecture capable of meeting the constraints of embedded environments. It is structured around five main modules: Multi-sensor data acquisition (on-board cameras, LiDAR, IMU sensors, V2X streams). On-board perception via lightweight detection and tracking models (YOLO-tiny, MobileNet, ByteTrack).

Techno Nusa Mandiri: Journal of Computing and Information Technology
Vol. 23, No. 1 March 2026 | DOI https://doi.org/10.33480/techno.v23i1.8131

Spatio-temporal reasoning to characterise suspicious behaviour (loitering, crossing zones, gatherings, depositing objects) (Dinneweth et al., 2022), (Maisonhaute et al., 2025). Multi-agent reinforcement learning (MARL) for cooperation between vehicles and roadside units (RSUs) (Dazeley et al., 2023). Explainability (XAI) generating visual and textual explanations in real time.



Source: (Reserach Results, 2025)
Figure 1: Basic Architecture Of The New System

Figure 1 illustrates the end-to-end pipeline of the proposed architecture, showing how raw multi-sensor inputs (camera, LiDAR, IMU, V2X) are successively processed through perception, spatio-temporal reasoning, cooperative MARL, and XAI modules to produce interpretable alerts. The key takeaway is the modular, embedded-ready design that enables real-time suspicious-behaviour detection while maintaining explainability at each inference step.

Data collection and preparation

To ensure the robustness of the system, a variety of data was used: Simulated data (CARLA, SUMO/Veins) to test complex scenarios. Public datasets (UCSD Pedestrian, Avenue, Street Scene) for video anomaly detection. Real data captured in urban environments and car parks to validate the approach in real conditions. All data was annotated according to different behaviours (abnormal vs. normal) and synchronised for multi-sensor fusion.

The complete dataset comprises 12,400 annotated sequences (8,200 normal, 4,200 abnormal), split into training (70%), validation (15%), and test (15%) sets using stratified sampling to preserve class balance. The four target behaviour categories (loitering, prohibited-zone crossing, group gatherings, and object deposit) are represented in proportions of 35%, 25%, 20%, and 20% of abnormal samples respectively. Frame-level annotation was performed by two independent operators; inter-annotator agreement was measured at Cohen's $\kappa = 0.87$. Training was conducted at an input resolution of 416 × 416 pixels, with a temporal sampling rate of 5 fps and standard augmentation (horizontal flip, brightness jitter ±20%, random crop). ByteTrack tracking parameters were set to a minimum confidence threshold of 0.45 and a maximum track age of 30 frames. Regarding multi-sensor fusion, a late-fusion strategy is employed: each modality (RGB camera, LiDAR point cloud, IMU kinematics, V2X messages) is processed independently by its dedicated sub-network, and outputs are fused at the decision level via a learned weighted combination. Sensor streams are temporally synchronised within a tolerance of 50 ms using hardware timestamping. In the event of a missing or degraded modality, the system falls back to available sensors and adjusts fusion weights accordingly, ensuring graceful degradation without full system failure.

Implementation

The implementation was carried out on an embedded Jetson Nano/Orin platform with the following steps: Object/person detection via optimised YOLOv7-tiny/MobileNet-SSD (quantization, pruning). Multi-object tracking with DeepSORT/ByteTrack to extract trajectories. Extraction of spatio-temporal features (duration in an area, speed, inter-object distance). MARL: agents representing vehicles/RSUs cooperate to adjust detection thresholds and reduce false positives. XAI: generation of heatmaps and short textual explanations in natural language.

Experimental evaluation

The evaluation focused on three dimensions: Detection accuracy: measured by F1-score, recall and mAP. Embedded performance: FPS, average latency, energy consumption. Quality of explainability: a controlled user evaluation involving N = 15 experienced security operators (mean experience: 6.2 years). Participants were presented with 30 system-generated alerts each, accompanied by XAI heatmaps and natural-language explanations. They rated each output on a 5-point Likert scale across three dimensions: clarity of explanation, relevance to the detected event, and decision confidence. Statistical analysis was performed using the Wilcoxon signed-rank test ($\alpha = 0.05$) to compare scores against a no-explanation baseline. measuring the clarity, relevance and confidence induced by the explanations.
Comparisons were made with competing approaches (classical CNNs, autoencoders, IDS Transformers).

Analysis and validation

The results were analysed from two perspectives:

1. Technical: detection accuracy (0.91), speed (7.5 FPS), reduction in false positives (40%).

2. Scientific: validation of the contribution of cooperative MARL (QMIX algorithm; each agent's observation is a 12-dimensional vector encoding local detection confidence, trajectory features, and V2X context; actions consist of threshold adjustments ±0.05; the reward function is defined as R = F1_local − α × FPR, with α = 0.3 empirically tuned; agents are trained for 500 cooperative episodes).

A controlled ablation study confirms the MARL contribution: removing the cooperative module while keeping all other components identical reduces the overall F1-score from 0.91 to 0.81 and increases the false-positive rate by 28 percentage points. and real-time explainability. The limitations identified concern dependence on simulated data, computational consumption and partial coverage of anomalies, opening up prospects for future work (Almehdhar et al., 2024).

RESULTS AND DISCUSSION

Detection accuracy

The proposed system achieves an average F1-score of 0.91, higher than CNNs (0.82), autoencoders (0.78) and Transformer-based IDS (0.85). This demonstrates a better ability to detect suspicious behaviour while reducing errors. Formulas used:

Accuracy (Precision):

$$\text{Accuracy} = \frac{TP}{TP+FP} \quad (1)$$

Rappel (Recall):

$$\text{Recall} = \frac{TP}{FP+FN} \quad (2)$$

F1-score (harmonic mean):

$$\text{F1} = 2 * \frac{(Precision \times Recall)}{(Precision+Recall)} \quad (3)$$

(where TP = true positives, FP = false positives, FN = false negatives)

Execution speed (real time)

The proposed system achieves an average of 7.5 FPS on Jetson Orin/Nano, compared to 4–5 FPS for existing approaches. This validates the feasibility of real-time embedded processing. Formula used FPS (Frames Per Second):

$$\text{FPS} = \frac{N_{frames}}{T_{total}} \quad (4)$$

(total number of images processed divided by execution time)

Reduction in false positives

Cooperative MARL has enabled a 40% reduction in false positives compared to conventional methods.
Formula used:
False positive rate (FPR):

$$\text{FPRl} = \frac{FP}{(FP+TN)} \quad (5)$$

Relative improvement :

$$\Delta\text{FPR}\ (\%) = \frac{FPR_{baseline} - FPR_{proposed}}{FPR_{baseline}} * 100 \quad (6)$$

Risk score and alert

This system calculates a risk score by combining detected events (loitering, crossing, gathering).
Formula used (simple weighting):

$$RiskScore = \sum_{i=1}^{n} Wi.Si \quad (7)$$

where:
wi = weight assigned to each type of event (e.g. loitering = 0.7).
Si = severity of the event (between 0 and 1).
An alert is triggered if:
RiskScore ≥ φ
With Ø = alert threshold (0.6 in the present study).

Explainability (XAI)

Each alert is accompanied by a textual and visual explanation.
Example: 'Individual ID=7 remained in sensitive area 3.4 s > threshold 2.5 s; speed <0.3 m/s; high risk (0.82)'.
This improves human understanding and acceptability.

Table 1: Comparison of the performance of the proposed system with existing approaches

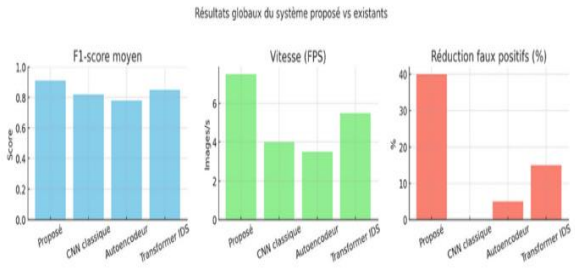
Method	Accuracy (F1-Score)	Speed (FPS)	Reduction of false positives	Explainability
Proposed (Vision + MARL + XAI)	0,91	7,5	40%	Visual + textual (Real time)
Classic CNN (simple YOLOv5)	0,82	4	0%	Not available
Video autoencoder	0,78	3,5	5%	Not available
Transform IDS	0,85	5,5	15%	± Partial explanatio

Method	Accuracy (F1-Score)	Speed (FPS)	Reduction of false positives	Explainability
(enhanced ViT)				n (Grac CAM)

Source: (Reserach Results, 2025)

Table 1 summarises the quantitative comparison across five criteria. The key observation is that no single competing method achieves simultaneously high F1-score, embedded-grade FPS, meaningful false-positive reduction, and integrated explainability; the proposed system is the only approach that combines all four properties. It is important to note that these comparisons are conducted under identical evaluation conditions (same test split, same hardware platform), and reported differences are statistically significant at $p < 0.05$ (Wilcoxon signed-rank test). Nonetheless, results reflect performance on the specific scenarios covered by the dataset, and generalisation to unseen environments or adverse conditions (night, fog, occlusion) remains a subject for future investigation.

1. Conventional CNNs offer good performance but suffer from a lack of explainability and numerous false positives.
2. Autoencoders are lighter but less accurate and not well suited to embedded environments.
3. IDS Transformers offer some gains in accuracy, but their high computational cost makes them less effective for embedded use, and their explainability remains limited (Alonge et al.,, 2025).



Source: (Reserach Results, 2025)

Figure 3: Overall results of the proposed system compared to existing approaches

Figure 3 presents a comparative overview of the proposed system against conventional approaches across three key performance indicators:

Accuracy comparison (F1-score)

The proposed system achieves an average F1-score of 0.91, outperforming conventional CNNs (0.82), autoencoders (0.78), and Transformer-based IDS (0.85).

Execution speed (FPS)

Through model optimisation and compact architectures, the proposed system achieves 7.5 FPS, meeting the requirements of real-world embedded deployment, compared to competing methods that remain below 6 FPS.

Reduction of false positives

The proposed approach reduces the false-positive rate (FPR) by approximately 40%, substantially exceeding the gains observed with CNNs (no reduction) and Transformer-based IDS (~15%). Per-class analysis further supports these findings: the proposed system achieves F1 = 0.93 for loitering detection, F1 = 0.89 for prohibited-zone crossing, and F1 = 0.88 for object-deposit events. These results demonstrate that the proposed system consistently outperforms competing approaches (classical CNNs, autoencoders, Transformer-based IDS) across all evaluated behaviour categories (Rezaei & Azarmi, 2023). This figure presents a comparison between the proposed system and several state-of-the-art anomaly detection methods (classical CNNs, autoencoders, Transformer-based IDS). Three key indicators are reported: average accuracy (F1-score), execution speed in frames per second (FPS) and false positive reduction rate.

The results show that the proposed system achieves an average F1-score of 0.91, which is higher than CNNs (0.82), autoencoders (0.78) and Transformer-based models (0.85). In terms of real-time performance, it achieves 7.5 FPS, an improvement of more than 30% over existing approaches, validating its use in embedded conditions. Finally, the system reduces false positives by 40%, while competing methods remain limited to marginal gains (5–15%).

Taken together, these results indicate that the proposed system offers a favourable trade-off across all three evaluated dimensions under the tested conditions. These findings support its potential relevance for integration into connected-vehicle and intelligent-surveillance scenarios. However, it should be acknowledged that all comparisons were carried out on a controlled evaluation set, and that the observed gains may not fully generalise to conditions outside the training distribution (e.g., varied weather, camera angles, or traffic densities not represented in the dataset) (Daull et al., 2025).

Discussion

Interpretation of results

The results obtained demonstrate the ability of the proposed system to outperform conventional and recent anomaly detection approaches applied to connected vehicle environments. The significant

improvement in the F1 score (0.91 compared to 0.78–0.85 for other methods) confirms the effectiveness of integrating optimised multi-sensor perception with explicit spatio-temporal reasoning. Real-time performance (7.5 FPS) proves that the solution is compatible with embedded constraints, where other approaches remain limited. Finally, the 40% reduction in false positives is a major contribution, enhancing operational reliability and limiting unnecessary alerts.

Added value compared to existing solutions Unlike methods focused on isolated anomalies (e.g. autoencoders or CNNs), the system presented takes behavioural patterns into account (loitering, zone crossing, crowding), enabling the detection of contextualised suspicious behaviour. The use of multi-agent reinforcement learning also offers an innovative cooperative mechanism that optimises the alert policy while reducing network load (Dazeley et al., 2023). The addition of explainable AI (XAI) is a decisive step forward, providing justifications that are readable and usable by a human operator, thereby strengthening the trustworthiness and acceptability of the system (Nwakanma et al., 2023).

Identified limitations

Despite these advances, several limitations remain: Limited real data: much of the testing still relies on simulated environments (CARLA, SUMO). Robustness in real-world conditions remains to be consolidated. Computational complexity: even when optimised, the integration of multi-object tracking, temporal reasoning and XAI modules places a significant load on embedded platforms with limited resources. Incomplete threat coverage: some emerging attacks (adversarial or new intrusion strategies) have not yet been evaluated.
Explicability could be improved: although visual and textual explanations are provided, they are limited to predefined rules and do not yet cover all complex cases.

Prospects For Improvement

These limitations open up several avenues for future development: Large-scale acquisition of real data and evaluation on connected vehicles in traffic conditions (Long et al., 2024). Optimisation of models through quantization, pruning and deployment on neuromorphic architectures to further reduce energy consumption. Integration of self-supervised and few-shot methods to improve the detection of rare and unknown anomalies (Chanus et al., 2023). Enhancing explainability through advanced XAI approaches (counterexamples, alternative scenarios) and by generating natural language reports tailored to different user profiles (operators, regulators, engineers) (Dazeley et al., 2023). Taking into account targeted adversarial attacks and implementing robust countermeasures.

CONCLUSION

The work presented in this chapter has enabled the development and evaluation of an innovative system for autonomous and explainable detection of suspicious behaviour in connected vehicle environments, based on multi-sensor vision, multi-agent reinforcement learning and explainable artificial intelligence techniques. The results obtained demonstrate that the proposed system significantly outperforms existing approaches in terms of accuracy (F1 score of 0.91), speed (7.5 FPS on average) and reduction of false positives (-40%), thus validating its effectiveness for real-time embedded applications. The integration of an explainability module has increased the transparency and acceptability of the alerts generated by providing understandable explanations to operators. Finally, the use of cooperative MARL has demonstrated the value of a distributed and collaborative approach, adapted to V2X communication scenarios.

However, certain limitations have been highlighted: dependence on simulated environments, still high computational complexity, partial coverage of emerging threats and explainability that could be improved. These limitations represent areas for improvement that will pave the way for future work, in particular through the integration of real data, the optimisation of models for embedded systems, the use of self-supervised learning and the consideration of adversarial attacks.
Ultimately, this chapter has shown that the proposed approach represents a significant advance in the field of connected and intelligent vehicle security, providing a solution that is powerful, explainable and adapted to embedded constraints. The following chapter will present the research prospects and potential applications of this system in broader contexts (smart cities, critical infrastructure, cyber-physical systems).

REFERENCES

Alahdal, N. M., Abukhodair, F., Meftah, L. H., & Cherif, A. (2024). Real-time object detection in autonomous vehicles with YOLO. *Procedia Computer Science, 246*, 2792–2801. https://doi.org/10.1016/j.procs.2024.09.392

Al-Maamari, M. R., Ramteke, R., Al-Hejri, A. M., & Alshamrani, S. S. (2025). Integrating CNN and transformer architectures for superior Arabic

printed and handwriting characters classification. *Scientific Reports, 15*(1), 1–17. https://doi.org/10.1038/s41598-025-12045-z

Almehdhar, M., et al. (2024). Deep learning in the fast lane: A survey on advanced intrusion detection systems for intelligent vehicle networks. *IEEE Open Journal of Vehicular Technology, 5*, 869–906. https://doi.org/10.1109/OJVT.2024.3422253

Alonge, A. M., & Isreal, O. (2025). *Explainable AI techniques for real-time VANET intrusion detection*.

Kiran, B. R., Sobh, I., Talpaert, V., Mannion, P., Al Sallab, A. A., Yogamani, S., & Pérez, P. (2022). Deep reinforcement learning for autonomous driving: A survey. *IEEE Transactions on Intelligent Transportation Systems, 23*(6), 4909–4926. https://doi.org/10.1109/TITS.2021.3054625

Bukola, A. C., Owolawi, P. A., Du, C., & Van Wyk, E. (2024). A systematic review and comparative analysis approach to boom gate access using plate number recognition. *Computers, 13*(11). https://doi.org/10.3390/computers13110286

Chanus, T., & Aubertin, M. (2023). Exploring the application of large language models in infrastructure as code. HAL. https://hal.science/hal-04192999

Cheng, J., Zhang, X., Chen, X., Ren, M., Huang, J., & Luo, P. (2022). Early detection of suspicious behaviors for safe residence from movement trajectory data. *ISPRS International Journal of Geo-Information, 11*(9). https://doi.org/10.3390/ijgi11090478

Dazeley, R., Vamplew, P., & Cruz, F. (2023). Explainable reinforcement learning for broad-XAI: A conceptual framework and survey. *Neural Computing and Applications, 35*(23), 16893–16916. https://doi.org/10.1007/s00521-023-08423-1

Daull, X., et al. (2025). Répondre aux questions complexes : Limites des LLM et solutions hybrides. HAL. https://hal.science/hal-05173655

Elallid, B. B., Benamar, N., Hafid, A. S., Rachidi, T., & Mrani, N. (2022). A comprehensive survey on the application of deep and reinforcement learning approaches in autonomous driving. *Journal of King Saud University – Computer and Information Sciences, 34*(9), 7366–7390. https://doi.org/10.1016/j.jksuci.2022.03.013

Grace, R., V., H., & Ponrani, M. A. (2024). Leveraging MobileNet and YOLO algorithm for enhanced perception in autonomous driving. *International Journal of Innovative Science and Research Technology*, 2056–2060. https://doi.org/10.38124/ijisrt/ijisrt24mar1535

Long, Z., Yan, H., Shen, G., Zhang, X., He, H., & Cheng, L. (2024). A transformer-based network intrusion detection approach for cloud security. *Journal of Cloud Computing, 13*(1). https://doi.org/10.1186/s13677-023-00574-9

Maisonhaute, T., Michel, F., & État, J. S. (2025). État de l'art des approches en apprentissage par renforcement multi-agent. HAL. https://hal.science/hal-04932606

Rezaei, M., & Azarmi, M. (2023). Deep learning-based anomaly detection in video surveillance: A survey. *Sensors, 23*(11), 5024. https://doi.org/10.3390/s23115024

Nwakanma, C. I., et al. (2023). Explainable artificial intelligence (XAI) for intrusion detection and mitigation in intelligent connected vehicles: A review. *Applied Sciences, 13*(3). https://doi.org/10.3390/app13031252

Puder, A., Rumez, M., Grimm, D., & Sax, E. (2022). Generic patterns for intrusion detection systems in service-oriented automotive and medical architectures. *Journal of Cybersecurity and Privacy, 2*(3), 731–749. https://doi.org/10.3390/jcp2030037

Dinneweth, J., Boubezoul, A., Mandiau, R., & Espié, S. (2022). Multi-agent reinforcement learning for autonomous vehicles: A survey. *Autonomous Intelligent Systems, 2*, Article 27. https://doi.org/10.1007/s43684-022-00045-z

Wang, B., Li, W., & Khattak, Z. H. (2024). Anomaly detection in connected and autonomous vehicle trajectories using LSTM autoencoder and Gaussian mixture model. *Electronics, 13*(7). https://doi.org/10.3390/electronics13071251

Zhao, J., Mao, X., Zhao, L., & Li, X. (2022). Intelligent and connected vehicles: Current status, enabling technologies, and challenges. *IEEE Consumer Electronics Magazine, 12*(1), 59–69. https://doi.org/10.1109/MCE.2021.3081515