UNIVERSITY OF NIS FACULTY OF MECHANICAL ENGINEERING

THE TENTH INTERNATIONAL CONFERENCE
TRANSPORT AND LOGISTICS
https://doi.org/10.46793/TIL2025.240C

til 2025

# EXPLAINABLE ARTIFICIAL INTELLIGENCE FOR TRUSTWORTHY EDGE-BASED ROAD INTERSECTION MANAGEMENT

***Žarko Ćojbašić[1], Goran Petrović[1], Dejan Rančić[2], Emilija Ćojbašić[2], Nemanja Marković[1], Bratislav Lukić[1]***

*[1)]Faculty of Mechanical Engineering, University of Niš, Serbia*
*[2)]Faculty of Electronic Engineering, University of Niš, Serbia*

ORCID iDs: Žarko Ćojbašić https://orcid.org/0000-0002-4581-1048
Goran Petrović https://orcid.org/0000-0002-3562-9576
Dejan Rančić https://orcid.org/0000-0002-9445-7700
Emilija Ćojbašić https://orcid.org/0009-0008-3843-6703
Nemanja Marković -
Bratislav Lukić https://orcid.org/0009-0004-1397-8195

***Abstract***

*Artificial Intelligence (AI) is increasingly applied to road intersection management to optimize traffic flow, reduce congestion, and improve safety. However, the opaque nature of AI-based decisions can undermine trust and limit adoption in intelligent transportation systems. This paper extends research on AI-driven intersection management by integrating Explainable Artificial Intelligence (XAI) methods ensuring the decisions of AI are interpretable and trustworthy for real-world deployment by traffic management authorities. A YOLO-based model is used for vehicle detection and traffic state recognition, while XAI techniques are applied to provide interpretable insights into system decisions. Implementation on the NVIDIA Edge AI device is considered and tested on a large dataset of real-world road scenes with multiple traffic objects, including various vehicle types and pedestrians. Results from experiments show that XAI integration enhances system comprehensibility without significantly reducing performance. These findings underline the importance of explainability for deploying trustworthy, accountable, and safe AI solutions in road intersection management.*

***Keywords***: *Explainable Artificial Intelligence (XAI); Edge AI; Deep Learning; YOLO; Intersection Management.*

## 1 INTRODUCTION

The critical need for effective road intersection management to combat congestion, mitigate rising emissions adversely affecting public health, and enhance overall urban safety is increasingly being addressed through the sophisticated and predictive modeling capabilities of recent AI and machine learning (ML) techniques to ensure sustainable and efficient urban traffic flow [1]. The motivation for research in this paper is the move from simply optimizing traffic flow with complex AI to ensuring the decisions of AI are interpretable and trustworthy for real-world deployment by traffic management authorities. Conventional traffic controls fail to mitigate congestion due to their static nature, while non-explainable AI models, despite their good performance, often suffer from a "black-box" lack of opacity that severely limits acceptance and complicates the auditing of safety-critical decisions [2], thereby establishing the crucial need for Explainable AI (XAI) in intelligent transportation systems (ITS) [3]. Furthermore, the necessity of using Edge AI for real-time traffic control, specifically leveraging NVIDIA devices, is well-justified in recent literature by the need to overcome the fundamental latency constraints of cloud-based systems for time-critical decision-making [4].

The contribution of this paper is integration of XAI with a YOLO-based perception model on an Edge AI platform to achieve portability, performance and trustworthiness in road intersection management.

The rest of the paper is organized as follows. In Section 2 related work is briefly reviewed, in Section 3 Edge Ai platform and methodology are introduced, in Section 4 results and disscusion are presented while in Section 5 conclusions and future research directions are outlined.

## 2 RELATED WORK

Contemporary AI and ML approaches for Traffic Signal Control are recently being advanced through strong development of Adaptive Traffic Signal Control (ATSC) methods, ranging from established optimization and metaheuristic algorithms to cutting-edge Deep Reinforcement Learning (DRL) techniques like Multi-Agent DRL, improving their effectiveness in managing dynamic traffic and improving real-time flow [5][6].

The growing need for Edge AI, which embeds intelligence directly into devices for real-time processing on site with improved privacy and reduced latency, has driven the recent development in the field, focusing on enabling core technologies like specialized hardware accelerators and optimized software, while exploring emerging opportunities such as neuromorphic hardware, continual learning algorithms, and enhanced edge-cloud collaboration [7]. Similarly, there is the growing need for Edge AI in Intelligent Transportation Systems which is motivated by its ability to resolve the high latency, security, and bandwidth challenges of cloud-centric systems, driving recent developments across layered Edge AI architectures and advanced application scenarios like autonomous driving and intelligent and autonomous vehicular transportation [4][8]. Recent development in deep learning vision-based traffic perception centers on the YOLO (You Only Look Once) [9] algorithm, which, as a leading single-stage model, provides the optimal balance of speed and precision necessary for real-time applications like vehicle and pedestrian counting, tracking, and general traffic state estimation [10], also providing possible connection to our previous results in other fields such as energy harvesting for transport applications [11].

Previously discussed growing need for Explainable AI (XAI), explained by the recently recognized imperative for trustworthy AI in critical domains like healthcare and finance, has led to the development of many post-hoc and intrinsic methods, with SHAP (Shapley Additive Explanations), LIME (Local Interpretable Model-agnostic Explanations), and Grad-CAM (Gradient-weighted Class Activation Mapping) being among the most prevalent techniques utilized for model interpretability [12].

Our research is aimed at identifying the research gap reflected in the lack of integrated solutions that simultaneously address real-time performance, object detection accuracy, and mandated explainability on resource-constrained edge devices in AI applications for intersection management.

## 3 EDGE SYSTEM ARCHITECTURE AND METHODOLOGY

In our research, complex traffic situations have been analyzed by deep learning on an edge device in order to identify major congestion in different directions at road intersections, where overcrowding of different groups involved in traffic were considered, such as vehicles, pedestrians, bicycles and others. The aim was to allow for automated decisions regarding traffic regulation and congestion and pollution reduction. On top of deep learning comprehensive traffic participants detection, a XAI layer has been added to provide for trustworthy and reliable detection system that can explain its decisions. The focus was on implementation with an Edge AI device.

For the approach considered in this research The Traffic Detection Project dataset was used [13], which is a traffic-camera image dataset containing 8,693 images annotated for object detection in YOLO format, split into training (7,566), validation (805) and test (322) subsets. The images depict road scenes with multiple traffic objects, including various vehicle types (cars, motorbikes, buses, trucks), pedestrians and traffic signs captured from fixed cameras. Most images are taken in Turkey, especially in cities such as Bursa, Istanbul and Konya, while still providing some geographic diversity with scenes from other countries. The dataset spans a range of environmental conditions (different weather, lighting and traffic density levels), which makes it useful for testing the robustness of detection models in realistic traffic scenarios. Thanks to its size, high-quality bounding-box annotations and variability, it is well suited for training and benchmarking of modern deep learning object-detection models in traffic surveillance and intelligent transportation research. In Figure 1 dataset typical traffic situation is presented, with several traffic participans groups annotaded wth bounding boxes.

In this paper, a recent YOLOv11 model nas been used, trained on selected dataset [13] which comprises relevant traffic images. Static images of traffic scenarios have been used, but the next steps will include computer vision from IMX219 camera connected to MIPI CSI-2 Jetson Camera module monitoring real time traffic.

Inference tasks have been conducted using the Edge AI device NVIDIA Jetson Orin Nano (Figure 2), which features multicore Ampere GPU, a 6-core ARM CPU and achieves up to 40 TOPS performance depending on specific configuration. A general overview of the specifications of the Jetson Orin Nano used in this research is summarized in Table 1. Fundamental details about the operating system, JetPack version, as well as general information on the CPU, GPU, memory utilization, temperature and energy consumption can be tracked using jtop, a real-time monitoring tool for NVIDIA Jetson devices. Device has been used in its recent Super Developer Mode, signidicantly increasing performance.

*Fig. 1 Typical dataset traffic situation image with 21 traffic participans in groups annotaded wth bounding boxes*

**Table 1** Specifications of NVIDIA Jetson Orin Nano device

| Category | Specification |
|---|---|
| **CPU** | 6-core Arm Cortex-A78AE v11.2 64-bit CPU |
| **GPU** | NVIDIA Ampere, 1,024 CUDA cores, 32 Tensor Cores |
| **RAM** | 8 GB 128-bit LPDDR5 (unified) |
| **Software** | Ubuntu 22.04, JetPack 6.2 (CUDA 12.6.68, cuDNN 9.3, TensorRT 10.3.0.30) |
| **Power** | 7W – 25W |

*Fig. 2 Experimental setup with NVIDIA Jetson Orin Nano*

The model was trained for 40 epochs and achieved a high mean Average Precision at 50% Intersection over Union (mAP@50 ≈ 0.929) and a solid mean Average Precision averaged over Intersection over Union thresholds from 0.50 to 0.95 (mAP@50–95 ≈ 0.76), confirming its strong capability for detection tasks and its suitability for deployment in real-world transport engineering environments concerning road intersection management. The overall accuracy achieved by the model is 88%.

Figure 3 shows the Precision–Recall curve for some evaluated classes - bicycle, bus, car, and motorbike - together with the aggregated performance curve summarizing results across all classes. The curves exhibit stable precision across a wide range of recall values, indicating that the model maintains reliable detection performance, even as recall increases. Although small differences between individual class curves are present, variations remain minimal, suggesting that the model generalizes well across all evaluated object categories.

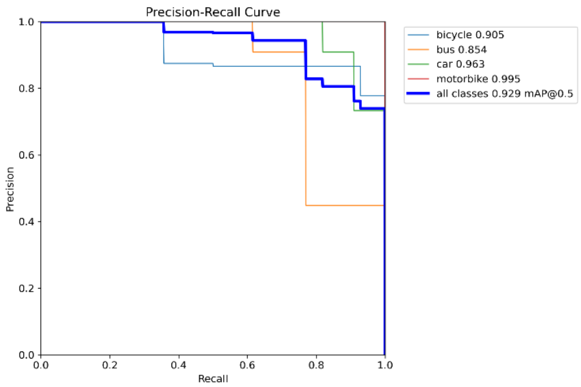


***Fig. 3** The model's Precision-Recall curve*

Custom wrappers were developed to enable seamless use of XAI techniques—SHAP, LIME, and Grad-CAM—in combination with the YOLO model, allowing flexible experimentation and evaluation. Implementing these methods on the Jetson platform required targeted adjustments and performance optimizations, including installing a CUDA-enabled PyTorch build to fully utilize GPU acceleration.

Grad-CAM, originally intended for CNN-based classification, had to be reworked for object detection by incorporating YOLO's multi-scale feature maps and bounding box logic. The key modification involved capturing high-level activation maps from the model's second to last convolutional layer prior to inference and then applying the standard Grad-CAM computation to those features.

The SHAP wrapper relies on a revised prediction function that extracts YOLOv11's average detection confidence values. Using these outputs, SHAP's KernelExplainer is executed with GPU batching to compute the corresponding SHAP values, with GPU execution being mandatory due to the heavy computational load.

Likewise, the LIME wrapper includes a tailored classification utility that retrieves YOLOv11 class probability estimates and applies LIME ImageExplainer with GPU-batched evaluation to derive explanations over segmented regions. To improve the fidelity of the generated explanations, the number of perturbation samples per image was set high, taking advantage of the GPU resources.

## 4 RESULTS AND DISCUSSION

Grad-CAM benefits strongly from GPU acceleration because it relies on a single forward pass and uses activations from a chosen layer, avoiding the repeated forward–backward computations required by LIME and SHAP, which consequently gain less from GPU execution.

For benchmarking on the NVIDIA Jetson, CPU-only and GPU-accelerated runtimes were compared at a batch size of 16. With GPU support, Grad-CAM achieved significant reduction in execution time, while LIME reached smaller but notable improvement. SHAP, however, could not be evaluated on the CPU due to excessive latency and memory consumption, therefore overall results without GPU acceleration were not considered.

To meaningfully demonstrate GPU capabilities, LIME was evaluated with many perturbation samples, whereas SHAP used fewer samples to maintain feasible resource usage. Timing results across different batch sizes confirmed that Grad-CAM consistently delivers by far the highest throughput among the three XAI techniques (Figure 4).

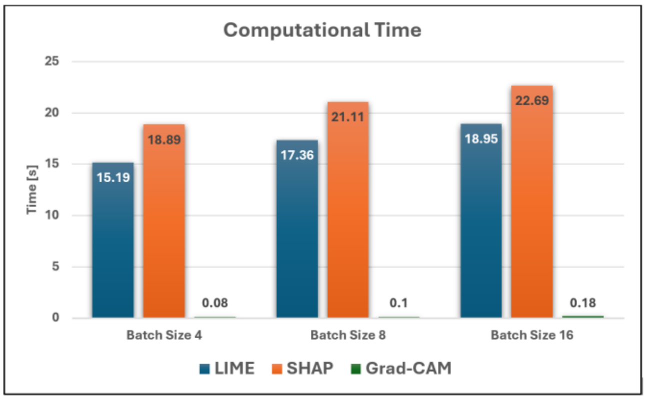


***Fig. 4** Computational time vs batch size for SHAP, LIME and Grad-CAM*

Figure 5 presents visual explanations generated by Grad-CAM and LIME. Grad-CAM produces heatmaps in which warmer colors (red/yellow) indicate the most influential areas for the predicted class. LIME highlights image regions through superpixel segmentation, where red segments correspond to features with stronger influence on the prediction compared to green segments, as determined by the local surrogate model. XAI output provides clear, verifiable justification for the system's decision in real-world road intersection traffic scenario.

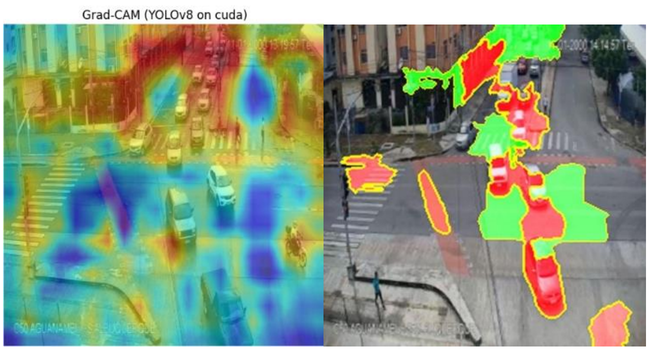


***Fig. 5** Grad-CAM heatmaps and LIME superpixel explanations illustrating key decision-making regions*

In our experiments, XAI integration causes only minimal, acceptable performance degradation while significantly boosting transparency. The next step in future research would assume inclusion of decision layer, which would enforce adaptive traffic policy. It would include the logic that transforms the YOLO output (vehicle and other traffic participants counts, queues lengths, speeds) into a definitive traffic signal control action (e.g., adjusting green light phase duration) deciding on resolving conqestion in any group of traffic participants (vehicles, motorbikes, pedestrians, cyclists, etc.) where XAI explanations would be used to provide explanation for the action performed.

Research results confirm a clear trade-off between the interpretability and efficiency of XAI methods on resource-constrained devices. While SHAP and LIME offered useful explanations, they proved too demanding for real-time deployment. Grad-CAM emerged as the most suitable choice due to significant GPU acceleration, underscoring that XAI method selection must align with available hardware capabilities.

From the wider perspective, results confirm that XAI models fundamentally contribute to trust, safety, and operational usability in AI systems. In road intersection management, visual explanations of traffic decisions allow human operators to rapidly understand AI behavior, react faster to unexpected traffic situations, and efficiently correct misclassified vehicle or pedestrian movements. This reduces risk and increases acceptance of emerging AI-driven ITS technologies.

## 5 CONCLUSION AND FUTURE WORK

Our results confirm that integrating XAI into deep learning systems for smart road intersection management on Edge AI devices is feasible but presents substantial difficulties. Successful implementation necessitates a relatively capable edge device, like the NVIDIA Jetson Orin Nano, and an essential selection of XAI methods for practical deployment. Combining XAI techniques (SHAP, LIME, Grad-CAM) with the computationally intensive YOLO model adds significant complexity, especially given the resource constraints of edge devices for real-time inference.

Of the three methods tested, Grad-CAM proved most suitable due to the challenging nature of YOLO's single-stage detection process. Grad-CAM exhibited significant performance improvement with Jetson's GPU acceleration. Conversely, LIME and SHAP were less suitable because they require multiple forward and backward passes, leading to less notable GPU speedup and making them impractical for real-time intersection applications.

The primary limitation that surfaced in our research is that resource constraints on current edge devices enforce a critical trade-off between the interpretability provided by complex XAI methods (like SHAP and LIME) and the real-time efficiency required for practical deployment, making simpler techniques like Grad-CAM the feasible option.

Future research should focus on full-scale experiments integrating XAI with real-world camera systems at intersections. This includes possibly exploring other XAI variants, especially CAM methods suitable for YOLO. Ultimately, future work must investigate edge-cloud hybrid explanation strategies to effectively balance interpretability demands with the strict real-time computational constraints of smart intersection management.

Explainability is essential in the next generation of safe and widely adopted Intelligent Transportation Systems (ITS) because it is the crucial bridge between complex AI decision-making and human trust, enabling the necessary transparency for rapid operator intervention, regulatory compliance, and continuous improvement of system performance in safety-critical scenarios [14].


## ACKNOWLEDGMENT

This work has been supported by the Ministry of Science, Technological Development and Innovation of the Republic of Serbia [Grant Numbers: 451-03-137/2025-03/200109 and 451-03-137/2025-03/200102].


## REFERENCES


1. Saadi, A., Abghour, N., Chiba, Z., Moussaid, K. and Ali, S., 2025. A survey of reinforcement and deep reinforcement learning for coordination in intelligent traffic light control. *Journal of Big Data*, 12(1), pp.1-20.
2. Khan, S., Ghazal, T.M., Alyas, T., Waqas, M., Raza, M.A., Ali, O., Khan, M.A. and Abbas, S., 2025. *Towards Transparent Traffic Solutions: Reinforcement Learning and Explainable AI for Traffic Congestion*. International Journal of Advanced Computer Science and Applications (IJACSA), 16(1), pp.503-511.
3. Liu, X. et al, (2025). From Black Box to Glass Box: A Practical Review of Explainable Artificial Intelligence (XAI). *AI*, 6(11), 285.
4. Liu, G., Shi, H., Kiani, A., Khreishah, A., Lee, J., Ansari, N., Liu, C. and Yousef, M.M., 2022. Smart Traffic Monitoring System Using Computer Vision and Edge Computing, *IEEE Transactions on Intelligent Transportation Systems*, 23(8), pp.12027–12038.
5. Agrahari, A., Dhabu, M.M., Deshpande, P.S., Tiwari, A., Baig, M.A. and Sawarkar, A.D., 2024. Artificial intelligence-based adaptive traffic signal control system: A comprehensive review, *Electronics*, 13(19), p.3875.
6. Prajapati, M., Upadhyay, A. K., Patil, H. & Dongradive, J. , 2024, A Review of Deep Reinforcement Learning for Traffic Signal Control, *International Journal for Multidisciplinary Research (IJFMR)*, 6(1), pp. 1-13.
7. Singh, R. and Gill, S.S., 2023, Edge AI: A survey, *Internet of Things and Cyber-Physical Systems*, 3, pp. 71–92.
8. Gong, T., Zhu, L. and Tang, T., 2023, Edge Intelligence in Intelligent Transportation Systems: A Survey, *IEEE Transactions on Intelligent Transportation Systems*, 24(9), pp. 8919-8944.
9. Wei, J. et al., 2025, A Review of YOLO Algorithm and Its Applications in Autonomous Driving Object Detection, *IEEE Access*, early access, DOI: 10.1109/ACCESS.2025.3573376.
10. Wang, J. and Shang, P., 2024, Edge Computing Application of Expressway Intelligent Transportation System Based on IoT Technology, *Computing and Informatics*, 43(4), pp. 974–992.
11. Milić, P., Marinković, D., Klinge, S. and Ćojbašić, Ž., 2023. Reissner-Mindlin based isogeometric finite element formulation for piezoelectric active laminated shells, *Tehnički vjesnik*, 30(2), pp.416-425.
12. Abusitta, A., Li, M.Q. and Fung, B.C., 2024, Survey on Explainable AI: Techniques, challenges and open issues. *Expert Systems with Applications*, 255, p.124710.
13. Sarıdoğan, Y.B. (2023) Traffic Detection Project [dataset]. Kaggle. Available at: https://www.kaggle.com/datasets/yusufberksardoan/traffic-detection-project, (Accessed 15th November 2025).
14. A. Kuznietsov et. Al, 2024, Explainable AI for Safe and Trustworthy Autonomous Driving: A Systematic Review," *IEEE Transactions on Intelligent Transportation Systems*, vol. 25, no. 12, pp. 19342-19364.



Contact address:
Žarko Ćojbašić
Faculty of Mechanical Engineering, University of Niš
18104 Niš
A. Medvedeva 14
E-mail: zcojba@ni.ac.rs