Proceedings of the 11th World Congress on Electrical Engineering and Computer Systems and Sciences (EECSS'25)
Paris, France - August, 2025
Paper No. MVML 121
DOI: 10.11159/mvml25.121

Fine-Tuning YOLOv8 for Vehicle Detection: A Deep Learning Approach to Traffic Congestion Monitoring

Ibtihal Mayouche[1], Abdellah Azmani[2]
[1,2]Intelligent Automation and BioMedGenomics Laboratory
FST of Tangier, Abdelmalek Essaâdi University, Tetouan, Morocco
ibtihal.mayouche@etu.uae.ac.ma; a.azmani@uae.ac.ma

Abstract **-** Traffic congestion is a crucial challenge in modern urban areas, causing delays, increased emissions, and inefficiencies. This study explores the application of the YOLOv8 model for vehicle detection in the context of traffic congestion monitoring. By fine-tuning YOLOv8 on a vehicle-specific dataset, the model achieved high precision (90.2%), recall (93.6%), and mean Average Precision (mAP50: 97.3%), showcasing its robustness in diverse traffic scenarios. Evaluation metrics, learning curve analysis, and inference results confirm the effectiveness of the fine-tuned model in accurately detecting vehicles, even in complex conditions. However, challenges such as false negatives and limited dataset diversity highlight areas for improvement. As a perspective, real-time video inference is proposed to monitor traffic streams, detect congestion based on vehicle and pedestrian density, and trigger automated decisions. This research establishes a foundation for intelligent traffic monitoring, with potential applications in improving transportation efficiency and reducing urban congestion.

***Keywords*:** Traffic Congestion, YOLOv8, Vehicle Detection, Fine-Tuning, Objetc Detection, Real-Time Inference, Deep Learning, Intelligent Transportation Systems

1. Introduction

Traffic congestion is a significant challenge faced by modern urban areas, leading to delays, increased pollution, and decreased overall efficiency [1]. Traditional methods for estimating road traffic congestion, such as manual observations and basic sensor-based systems, often lack the real-time analysis capabilities necessary for effective traffic management [2]. To address these challenges, advanced technologies such as deep learning offer the potential for highly accurate and scalable traffic monitoring solutions. Vehicle detection plays an important role in estimating traffic congestion, as it allows for real-time tracking of vehicle presence and movement, enabling more accurate assessments of traffic density. YOLO (You Only Look Once) models, particularly YOLOv8, have emerged as a powerful tool in object detection tasks due to their speed and accuracy. By leveraging these models, it is possible to detect vehicles more reliably and assess road traffic in a timely manner [3]. The primary objective of this research is to explore the application of YOLOv8 for vehicle detection in the context of road traffic congestion monitoring. This study aims to fine-tune the YOLOv8 model on a vehicle-specific dataset and evaluate its performance in accurately detecting vehicles across various traffic conditions. The goal is to assess how well the model generalizes in real-world scenarios and how it can contribute to real-time traffic congestion prediction.

The structure of this research is as follows: Section 2 provides an overview of related work in vehicle detection and traffic monitoring. Section 3 details the methodology, including the dataset, the architecture of YOLOv8, fine-tuning procedures, and performance metrics used in this study. Section 4 presents the results and discussions, underscoring the effectiveness of the fine-tuned model in improving vehicle detection accuracy. Finally, Section 5 concludes the study and suggests directions for future research in using deep learning models for traffic congestion detection.

2. Related work

The application of deep learning techniques to traffic congestion detection has significantly advanced the accuracy and efficiency of traffic monitoring systems. Leveraging surveillance camera images, researchers have developed innovative approaches to address real-world challenges such as varying lighting conditions, complex road configurations, and real-time responsiveness. Hua Cui et al. [4] explored the potential of convolutional neural networks (CNNs) for classifying highway traffic images into "congested" and "non-congested" categories. Using AlexNet and GoogLeNet, their study analyzed a diverse dataset encompassing various road configurations, weather conditions, and times of day. Both models achieved an impressive 98% accuracy on test samples, demonstrating robustness against challenges such as lighting variations, complex backgrounds, and perspective distortions. Notably, their approach eliminated the need for preprocessing techniques like road segmentation. However, limitations related to image scale and perspective occasionally affected recognition performance, especially for borderline congestion cases. Building on this foundation, Xiao Ke et al. [5] proposed a multidimensional approach that fused CNNs with additional visual features to enhance detection accuracy. This method incorporated gray-level co-occurrence matrices, optical flow for speed measurement, and a Gaussian mixture model for background modeling, outperforming traditional methods and demonstrating the advantages of feature fusion in CNN-based traffic analysis.

G. Bindu Madhavi et al. [6] extended the application of CNNs to video-based traffic monitoring systems (VTSS), focusing on detecting both traffic accidents and congestion. Using a continuous prediction technique, their CNN model trained on the Vehicle Accident Image Dataset (VAID) achieved a 93% accuracy in real-time traffic accident detection, showcasing the potential of CNNs for rapid incident response and improved road safety.

Ying Gao et al. [7] addressed the inefficiencies of traditional image-based congestion detection by developing a CNN framework that integrates a traffic parameter layer. By eliminating complex post-processing steps, their model directly estimated congestion from raw images, achieving reliable performance across diverse traffic conditions and weather scenarios. The framework's efficiency in real-time traffic management was demonstrated through reduced processing times. Simplifying CNN-based traffic congestion detection further, Jason Kurniawan et al. [8] trained a model on 1,000 grayscale CCTV images. With minimal preprocessing, their approach achieved an accuracy of 89.5%, emphasizing the viability of CNNs for low-resolution images and their potential for real-time monitoring in resource-constrained settings.

In another notable contribution, Yedi Zhuo et al. [9] leveraged traffic monitoring data from Shanxi Province to develop a semi-supervised CNN model with optimized "Detec-Nets" inspired by DenseNet blocks. Their approach reduced manual labeling efforts while achieving 93% accuracy and a low error rate of 2.46%. The system demonstrated robustness under challenging conditions such as blurred images or the presence of large trailers, making it suitable for deployment in highway monitoring systems.

Adriana-Simona Mihaitaa et al. [10] explored a hybrid approach for predicting and detecting anomalies in traffic congestion. Their study integrated CNNs for spatial feature extraction, RNNs for temporal dynamics, and a CNN-LSTM hybrid for spatio-temporal modeling. Trained on over 36 million data points, this model outperformed traditional methods in both congestion prediction and anomaly detection, offering a robust solution for real-time traffic management.

Asif et al. introduced a Tri-Stage Attention mechanism combining CNNs and RNNs for congestion prediction. Their hybrid model used Multi-Linear Discriminant Analysis (M-LDA) for traffic feature extraction, with CNNs analyzing spatial patterns and RNNs capturing temporal dependencies. Enhanced by an attention mechanism, the model demonstrated superior accuracy and scalability, contributing to intelligent traffic systems.

Ping Wang et al. [11] introduced TrafficNet, a CNN-based architecture tailored for complex freeway environments. By combining AlexNet and VGGNet with Support Vector Machines (SVM) for classification, TrafficNet achieved up to 90% accuracy on a dataset of 30,000 labeled traffic images, outperforming traditional feature extraction techniques. This study underscored the potential of CNNs in handling dynamic and complex traffic scenarios.

NAVIN RANJAN et al. [12] tackled urban traffic challenges using a hybrid CNN-LSTM-Transpose CNN model. Leveraging traffic maps from Seoul's Transportation Operation and Information Service (TOPIS), their model effectively captured spatial and temporal data, achieving superior prediction accuracy while maintaining computational efficiency. This work represents a significant advancement in network-wide congestion prediction and real-time traffic management.

Meng Chen et al. [13] developed PCNN, a deep convolutional neural network that incorporates periodic traffic data for short-term congestion prediction. By transforming time-series data into 2D matrices, PCNN captured multiscale traffic properties, enabling accurate predictions of macro and micro trends. The model outperformed traditional approaches, offering valuable insights for managing recurring congestion patterns.

The YOLO framework has also gained prominence in traffic congestion detection. Pranamesh Chakraborty et al. [14] compared YOLO with Deep Convolutional Neural Networks (DCNNs) for classifying traffic conditions, achieving 91.5% and 90.2% accuracy, respectively. Both models performed robustly across varied scenarios, maintaining high AUC values even under challenging nighttime conditions. Sundas Iftikhar et al. [15] extended YOLO's application to UAV-based traffic monitoring. By addressing challenges such as object scale and recognition accuracy, their study highlighted the advantages of UAVs for mobility and wide coverage in traffic monitoring, particularly in smart city implementations. Lastly, Saif Bashar and Abdulamir Abdullah Karim [16] explored advanced computer vision techniques for vehicle detection and tracking, showcasing YOLO's effectiveness in real-time traffic analysis. Their findings contribute to urban traffic optimization and enforcement, aligning closely with ongoing efforts to leverage YOLO for congestion detection.

3. Materials and Methods

This section outlines the materials and methodologies employed to develop and evaluate the YOLOv8 model for vehicle detection, which is essential for assessing traffic congestion.

3.1. Dataset

The Top-View Vehicle Detection Image Dataset was specifically curated to fine-tune YOLOv8 for detecting vehicles such as cars, trucks, and buses from aerial perspectives. Comprising 626 images, it is split into 536 training images and 90 validation images, all resized to 640x640 pixels to standardize input dimensions and optimize model performance. To enhance generalization, the training set undergoes data augmentation, including horizontal flipping, while the validation set remains unaltered to provide unbiased evaluations. Annotations are provided in the YOLO format, with bounding boxes defined using normalized coordinates (class, x_center, y_center, width, height). Label files are included only for images containing detectable objects, improving data processing efficiency. This diverse dataset captures various traffic scenarios, simulating real-world environments for robust model training. Its configuration is managed via a data.yaml file, specifying dataset paths, the number of classes (1), and the class name ('Vehicle'). With its meticulous design, this dataset establishes a reliable foundation for training YOLOv8 models, enabling precise and efficient vehicle detection in traffic monitoring applications.

3.2. Yolov8 Architecture

The YOLO framework, first introduced by Redmon et al. in 2016, revolutionized object detection by offering an end-to-end network capable of simultaneously detecting object locations and classifying their labels. Over the years, the model has undergone continuous advancements, culminating in its eighth iteration, YOLOv8, released in January 2023 [17]. This latest version incorporates several key architectural improvements:

- **Backbone**: YOLOv8's backbone leverages a variation of the Cross Partial Stage (CSP) network [18], which divides feature maps into segments for separate convolution operations. This design reduces computational complexity while preserving the model's learning capacity. The backbone is built upon the C2f module, an enhanced version of CSP influenced by the ELAN structure from YOLOv7 [19] . Additionally, the inclusion of the SPPF (Spatial Pyramid Pooling – Fast) module enhances detection performance across multiple scales.
- **Neck**: The model's neck incorporates the PAN-FPN (Path Aggregation Network and Feature Pyramid Network) architecture to achieve efficient multi-scale feature fusion. By combining the strengths of FPN and PAN, upper layers process higher-level contextual information, while lower layers preserve detailed spatial localization.
- **Head**: YOLOv8 adopts a decoupled head architecture, separating the classification task from the regression of bounding boxes. Unlike earlier anchor-based approaches, this version employs an anchor-free mechanism, where objects are identified based on their centers, and distances from the center to the bounding box edges are predicted directly. This simplifies the process by eliminating the dependency on predefined anchor boxes.

This refined architecture makes YOLOv8 highly efficient for object detection tasks, offering faster processing and improved accuracy across diverse applications [20].

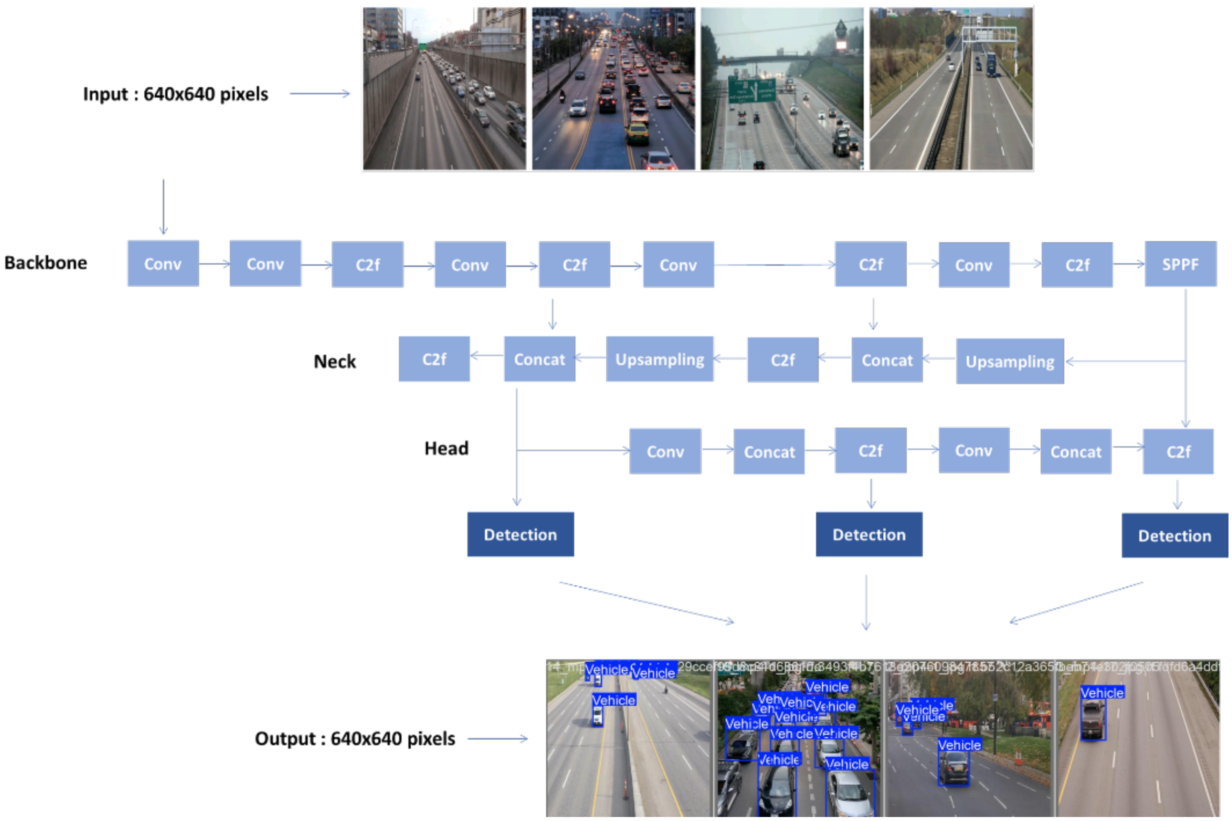

Fig. 1: The architecture of YOLOv8.

3.3. Fine tuning the YOLOv8 Model

In this phase of our study, we fine-tune the pre-trained YOLOv8 object detection model using transfer learning, adapting it specifically to the Top-View Vehicle Detection Image Dataset. Leveraging the model's pre-trained weights—originally optimized on the COCO dataset, which encompasses a diverse range of object classes—we bypass the need for training from scratch, significantly saving time and computational resources. This fine-tuning process allows the model to adapt to the unique characteristics of our dataset, including aerial perspectives capturing vehicles from top-down views, the spatial patterns and proportions typical of vehicles in traffic scenarios, and the variability of real-world highway environments. Through this tailored adaptation, the model becomes proficient in accurately detecting and localizing cars, trucks, and buses in complex traffic scenes. This approach leverages the reliability of pre-trained weights alongside task-specific adaptations, striking an optimal balance between computational efficiency and detection accuracy. Consequently, the model delivers outstanding performance in practical applications such as highway monitoring and traffic management.

3.4. Metrics

To assess the performance of the YOLOv8 model during and after training, we rely on a combination of loss functions, precision-recall, f1score metrics, and confusion matrix analysis.

The training losses include the box loss, which measures the error in bounding box predictions, cls loss for classification errors, and dfl loss (Distribution Focal Loss), which refines the accuracy of box predictions [21].

During evaluation, the confusion matrix provides a visual representation of true positives (TP), false positives (FP), and false negatives (FN), helping to identify specific areas where the model may be underperforming as shown in Fig.2. These metrics, used together, offer a comprehensive and precise assessment of the model's performance, ensuring its suitability for real-world applications such as traffic monitoring.

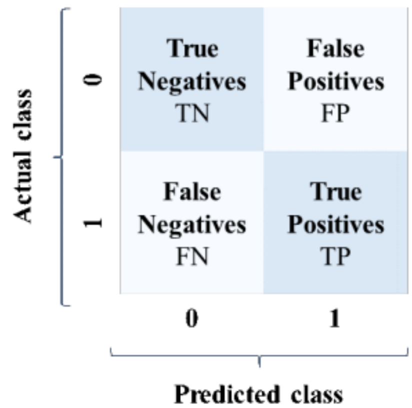

Fig. 2: Confusion Matrix. [22]

Precision measures the proportion of correct predictions among all positive detections Eq.1, while recall evaluates the model’s ability to detect all instances of a given class Eq.2.

$$Recall = \frac{TP}{TP + FN} \tag{1}$$

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

The F1 score, defined as the harmonic mean between precision and recall, provides a balanced metric, particularly useful in cases of class imbalance Eq.3.

$$F1 = 2 \times \left(\frac{Precision \times Recall}{Precision + Recall} \right) \tag{3}$$

Additionally, mAP (Mean Average Precision) is used to evaluate the model’s detection performance at various thresholds of Intersection over Union (IoU). The mAP@50 represents the average precision at a fixed IoU threshold of 0.50, while mAP@50-95 offers a stricter evaluation by averaging precision over IoU thresholds ranging from 0.50 to 0.95 [23].

Finally, Fitness combines key evaluation metrics such as precision, recall, and mAP into a single score to summarize the model's overall effectiveness. This metric simplifies the process of comparing different training runs or models.

4. Results & Discussion

This part presents the results of the YOLOv8 model's performance on the Top-View Vehicle Detection Image Dataset and discusses the insights derived from the analysis of key evaluation metrics, including learning curves, confusion matrix, precision-recall measures and inference performance.

4.1. Model Learning Curve Analysis

The learning curves for box loss, classification loss, and distribution focal loss demonstrate a significant reduction in loss values during the initial epochs, followed by a gradual stabilization as training progresses. This pattern, combined with the close alignment of the training and validation loss curves, suggests that the model is effectively learning without overfitting. It indicates that the model is well-tuned to the dataset and generalizes well, avoiding the pitfalls of bias or excessive variance. The smoothness observed in the learning curves, particularly during the latter epochs, suggests that the model is reaching a state of equilibrium, where additional training does not significantly improve performance. This finding implies that 100 epochs are sufficient for training the YOLOv8 model, and extending the training duration would likely offer diminishing returns in terms of further performance gains.

4.2. Confusion Matrix Analysis

The confusion matrix for the YOLOv8 vehicle detection model reveals strong accuracy, confirming the model's effectiveness in detecting vehicles as illustrated in Fig.3. The model successfully identifies the presence of a vehicle in 95% of instances, which reflects its robust detection capability. However, in the remaining 5% of cases, the model fails to detect a vehicle that is actually present, which suggests a need for improvement in reducing false negatives and enhancing the model's sensitivity. This confusion matrix result highlights that while the model performs well, there is still room for improvement in minimizing false negatives, which could be critical in real-time applications like traffic monitoring.

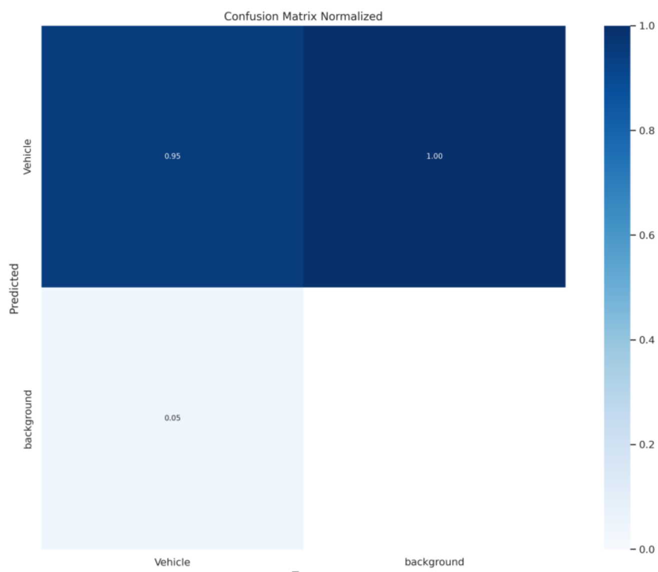

Fig. 3: Confusion matrix results.

4.3. Model Evaluation Insights

The YOLOv8 model's performance on the validation set is noteworthy, the results are summarized in Table 1. With a precision of 90.2%, the model consistently makes correct predictions, with few false positives. The recall score of 93.6% highlights the model’s ability to detect most of the relevant instances, emphasizing its effectiveness in identifying vehicles in various traffic scenarios. The mean Average Precision (mAP) at 50% Intersection over Union (IoU) is 97.3%, indicating the model’s accuracy in detecting objects with significant overlap with the ground truth. Even when the IoU threshold range is expanded from 50% to 97%, the model maintains a solid mAP of 74.1%, showcasing its robustness across a range of localization requirements. The fitness score of 76.4% demonstrates a good balance between precision, recall, and the IoU of the predictions, reaffirming the model’s capability to effectively perform object detection tasks.

Table 1: Evaluation metrics of YOLOv8 model.

Metrics	Value
Precision	0.902
Recall	0.936
mAP50	0.973
Map50-95	0.741
Fitness	0.764

4.4. Inference Performance

To gauge the model's effectiveness in generalizing, we conducted inference on both the validation set and an unseen test image, followed by testing on real-world data. The results highlighted the importance of fine-tuning YOLOv8 for vehicle detection. Before fine-tuning, the pre-trained YOLOv8 model correctly detected a person as a "person" and failed to detect a car that was behind another vehicle. This was due to the model’s broader training, which was not optimized specifically for vehicle detection. After fine-tuning on a vehicle-specific dataset, the model no longer detected the person, as it was now restricted to detecting only vehicles. As a result, it successfully identified and classified all vehicles, including the previously missed car, showing a clear improvement in detecting only the relevant objects (vehicles) with higher accuracy as presented in Fig.4.

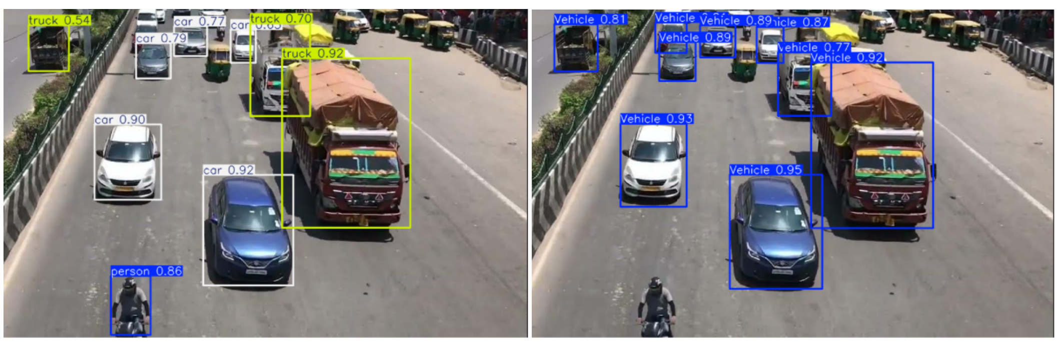

Fig. 4: YOLOv8 performance before (left) and after (right) fine-tuning on a vehicle dataset

5. Conclusion

This study successfully fine-tuned the YOLOv8 model for vehicle detection, achieving impressive precision, recall, and mean Average Precision (mAP). The model demonstrated robustness in addressing traffic monitoring challenges, though issues like false negatives and limited data diversity need to be addressed for improved generalization. As part of future work, our study provides several implications. First, the integration of **real-time video inference systems** represents a critical step toward more dynamic and responsive traffic management. These systems would analyze ongoing traffic streams, assess congestion levels based on vehicle and pedestrian density, and trigger automated responses in high-density scenarios. By incorporating live video analysis, congestion detection could become more precise and timely, enabling immediate actions to manage traffic flow effectively. Second, building on this, **Mobility as a Service (MaaS)** can play a pivotal role in enhancing real-time traffic management. By combining the data from real-time congestion detection systems with multimodal transport platforms, MaaS could optimize the use of urban transport infrastructure. This would provide users with real-time insights into congestion levels and alternative transport options, helping reduce congestion and improve the overall efficiency of urban mobility by guiding users toward the best available transport solutions based on current traffic conditions. Third, to ensure the integrity and security of these systems is essential. **Intelligent and Resilient Urban Network Defender** (IRUNd) would secure communication between MaaS services, ensuring data integrity is maintained and the system remains resilient to cyber threats. This would guarantee a smooth and reliable experience for users, even in complex and evolving urban environments, where secure data flows and resilient infrastructures are critical. The integration of real-time video systems, MaaS, and IRUNd holds significant potential for advancing intelligent traffic management, creating a more secure, efficient, and sustainable urban mobility ecosystem.

Acknowledgements

This research was supported by the National Center for Scientific and Technical Research (CNRST) under the PhD-ASsociate Scholarship (PASS) program. Additional support was provided by the Ministry of Higher Education, Scientific Research, and Innovation, as well as the Digital Development Agency (DDA) and the CNRST of Morocco, under the Smart DLSP Project - AL KHAWARIZMI Program.

References

[1] S. Xu, C. Sun, and N. Liu, « Road congestion and air pollution -Analysis of spatial and temporal congestion effects », Sci. Total Environ., vol. 945, p. 173896, oct. 2024, doi: 10.1016/j.scitotenv.2024.173896.

[2] F. Zhao, H. Li, K. Niu, J. Shi, and R. Song, « Application of Deep Learning-Based Intrusion Detection System (IDS) in Network Anomaly Traffic Detection », 8 juillet 2024. doi: 10.20944/preprints202407.0595.v1.

[3] M. Mahayash, N. Punia, N. Tripathi, and Sarvamangala, « Smart approach for vehicle detection and counting », présenté à 1ST INTERNATIONAL CONFERENCE ON EMMA-2021, Chennai, India, 2024, p. 020057. doi: 10.1063/5.0185927.

[4] H. Cui, G. Yuan, N. Liu, M. Xu, and H. Song, « Convolutional neural network for recognizing highway traffic congestion », J. Intell. Transp. Syst., vol. 24, no 3, p. 279-289, mai 2020, doi: 10.1080/15472450.2020.1742121.

[5] X. Ke, L. Shi, W. Guo, and D. Chen, « Multi-Dimensional Traffic Congestion Detection Based on Fusion of Visual Features and Convolutional Neural Network », IEEE Trans. Intell. Transp. Syst., vol. 20, no 6, p. 2157-2170, juin 2019, doi: 10.1109/TITS.2018.2864612.

[6] G. B. Madhavi, A. D. Bhavani, Y. S. Reddy, A. Kiran, N. T. Chitra, and P. C. S. Reddy, « Traffic Congestion Detection from Surveillance Videos using Deep Learning », in 2023 International Conference on Computer, Electronics & Electrical Engineering & their Applications (IC2E3), Srinagar Garhwal, India: IEEE, juin 2023, p. 1-5. doi: 10.1109/IC2E357697.2023.10262545.

[7] Y. Gao, J. Li, Z. Xu, Z. Liu, X. Zhao, and J. Chen, « A novel image-based convolutional neural network approach for traffic congestion estimation », Expert Syst. Appl., vol. 180, p. 115037, oct. 2021, doi: 10.1016/j.eswa.2021.115037.

[8] J. Kurniawan, S. G. S. Syahra, C. K. Dewa, and Afiahayati, « Traffic Congestion Detection: Learning from CCTV Monitoring Images using Convolutional Neural Network », Procedia Comput. Sci., vol. 144, p. 291-297, 2018, doi: 10.1016/j.procs.2018.10.530.

[9] Y. Zhuo, P. Wang, J. Sun, and Y. Jin, « Traffic Congestion Detection Based on the Image Classification with CNN », in Advances in Guidance, Navigation and Control, vol. 644, L. Yan, H. Duan, and X. Yu, Éd., in Lecture Notes in Electrical Engineering, vol. 644. , Singapore: Springer Singapore, 2022, p. 4569-4580. doi: 10.1007/978-981-15-8155-7_378.

[10] A.-S. Mihaita, H. Li, and M.-A. Rizoiu, « Traffic congestion anomaly detection and prediction using deep learning », 2020, arXiv. doi: 10.48550/ARXIV.2006.13215.

[11] P. Wang, L. Li, Y. Jin, and G. Wang, « Detection of unwanted traffic congestion based on existing surveillance system using in freeway via a CNN-architecture trafficnet », in 2018 13th IEEE Conference on Industrial Electronics and Applications (ICIEA), Wuhan: IEEE, mai 2018, p. 1134-1139. doi: 10.1109/ICIEA.2018.8397881.

[12] N. Ranjan, S. Bhandari, H. P. Zhao, H. Kim, and P. Khan, « City-Wide Traffic Congestion Prediction Based on CNN, LSTM and Transpose CNN », IEEE Access, vol. 8, p. 81606-81620, 2020, doi: 10.1109/ACCESS.2020.2991462.

[13] M. Chen, X. Yu, and Y. Liu, « PCNN: Deep Convolutional Networks for Short-Term Traffic Congestion Prediction », IEEE Trans. Intell. Transp. Syst., vol. 19, no 11, p. 3550-3559, nov. 2018, doi: 10.1109/TITS.2018.2835523.

[14] P. Chakraborty, Y. O. Adu-Gyamfi, S. Poddar, V. Ahsani, A. Sharma, and S. Sarkar, « Traffic Congestion Detection from Camera Images using Deep Convolution Neural Networks », Transp. Res. Rec. J. Transp. Res. Board, vol. 2672, no 45, p. 222-231, déc. 2018, doi: 10.1177/0361198118777631.

[15] S. Iftikhar, M. Asim, Z. Zhang, A. Muthanna and J. Chen, « Target Detection and Recognition for Traffic Congestion in Smart Cities Using Deep Learning-Enabled UAVs: A Review and Analysis », Appl. Sci., vol. 13, no 6, p. 3995, mars 2023, doi: 10.3390/app13063995.

[16] S. B. Neamah and A. A. Karim, « Real-time Traffic Monitoring System Based on Deep Learning and YOLOv8 », ARO- Sci. J. KOYA Univ., vol. 11, no 2, p. 137-150, nov. 2023, doi: 10.14500/aro.11327.

[17] C. H. Kang and S. Y. Kim, « Real-time object detection and segmentation technology: an analysis of the YOLO algorithm », JMST Adv., vol. 5, no 2-3, p. 69-76, sept. 2023, doi: 10.1007/s42791-023-00049-7.
[18] X. Wang and J. Liu, « Tomato Anomalies Detection in Greenhouse Scenarios Based on YOLO-Dense », Front. Plant Sci., vol. 12, p. 634103, avr. 2021, doi: 10.3389/fpls.2021.634103.
[19] W. Liu, Z. Shen, and S. Xu, « CF-YOLO: a capable forest fire identification algorithm founded on YOLOv7 improvement », Signal Image Video Process., vol. 18, no 8-9, p. 6007-6017, sept. 2024, doi: 10.1007/s11760-024-03288-w.
[20] F. Solimani, A. Cardellicchio, G. Dimauro, A. Petrozza, S. Summerer, F. Cellini and V. Renò, « Optimizing tomato plant phenotyping detection: Boosting YOLOv8 architecture to tackle data complexity », Comput. Electron. Agric., vol. 218, p. 108728, mars 2024, doi: 10.1016/j.compag.2024.108728.
[21] H. Rezatofighi, N. Tsoi, J. Gwak, A. Sadeghian, I. Reid, and S. Savarese, « Generalized Intersection Over Union: A Metric and a Loss for Bounding Box Regression », in 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA: IEEE, juin 2019, p. 658-666. doi: 10.1109/CVPR.2019.00075.
[22] I. Boukrouh, F. Tayalati, and A. Azmani, « Feature Selection and Classification Performance: A Multi-Dataset Comparative Analysis Using Boruta Algorithm and Random Forest », présenté à The 10th World Congress on Electrical Engineering and Computer Systems and Science, août 2024. doi: 10.11159/cist24.174.
[23] C. C. Mullins, T. J. Esau, Q. U. Zaman, and P. J. Hennessy, « Optimizing data collection requirements for machine learning models in wild blueberry automation through the application of DALL-E 2 », Smart Agric. Technol., vol. 10, p. 100764, mars 2025, doi: 10.1016/j.atech.2024.100764.