DOI: 10.35598/mcfpga.2026.008

Closed-Loop YOLO System for Adaptive Traffic Control

Volodymyr Hryha
ORCID 0000-0001-5458-525X
Dept. of Computer Engineering and Electronics
Vasyl Stefanyk Carpathian National University
Ivano-Frankivsk, Ukraine
volodymyr.hryha@cnu.edu.ua

Vitaliy Martyniuk
ORCID 0009-0001-4310-6302
Dept. of Computer Engineering and Electronics
Vasyl Stefanyk Carpathian National University
Ivano-Frankivsk, Ukraine
vitalii.martyniuk@cnu.edu.ua

Ihor Kohut
ORCID 0000-0001-5524-1790
Dept. of Computer Engineering and Electronics
Vasyl Stefanyk Carpathian National University
Ivano-Frankivsk, Ukraine
ihor.kohut@cnu.edu.ua

Liudmyla Hryha
ORCID 0000-0002-6260-7559
Dept. of Computer Engineering and Electronics
Vasyl Stefanyk Precarpathian National University
Ivano-Frankivsk, Ukraine
liudmyla.hryha@cnu.edu.ua

***Abstract* - A Safe Road Manager (SRM) system is developed in which a video-surveillance module (YOLO neural network, 106 layers, ONNX) and an adaptive traffic-light control module are for the first time united in a closed loop via a Feedback Coordinator. An extended sector load coefficient with a dynamic predictive term, an adaptive camera distribution algorithm for 1–4 cameras, and a cross-sector balancing method are proposed. The system is implemented in C# on .NET and validated by unit (NUnit) and functional tests.**

Keywords -SRM; YOLO; closed-loop control; adaptive traffic regulation; computer vision; feedback coordination; C#.

I. Introduction

Growing urban traffic intensity and the inability of traditional fixed-cycle signal control systems to adapt to real intersection load are the primary causes of road accidents in pedestrian crossing zones [1]. Existing solutions fall into two independent classes: video-surveillance systems (Axis Camera Station, Milestone XProtect, Qognify VisionHub) and regulation systems (SiTraffic Concert/Siemens, ATC/Hitachi, ATMS/Conduent) [2,3]. The fundamental problem is the absence of feedback between them: detection failures are directly propagated into control decisions without any correction, and cross-sector synchronization at the intersection is not performed at all.Recent advances in deep learning have made single-stage detectors the method of choice for real-time traffic video analysis [4, 5]. Compared to two-stage methods such as Faster R-CNN [10], single-pass architectures like YOLO achieve a favourable trade-off between detection accuracy and inference latency, making them suitable for embedded deployments where GPU resources are unavailable. State-of-the-art variants including YOLOv4 [9] further demonstrate that architectural improvements to the backbone and neck can push mean Average Precision above 43 % on MS COCO while maintaining real-time throughput.

The goal of this work is to eliminate this gap by constructing a unified closed loop in which video surveillance and regulation interact in real time. The SRM system is the first to integrate a custom-trained YOLO neural network, a Feedback Coordinator, and a cross-sector balancing algorithm into a single C#/.NET software platform.Unlike solutions that treat detection and signal control as independent pipelines, the proposed SRM architecture enforces bidirectional data exchange: every scheduling decision made by the regulation module is conditioned on counter reliability scores produced by the Feedback Coordinator, and every anomaly suppressed by the Coordinator is logged with a timestamp and sector identifier for post-hoc audit. This design follows the closed-loop control paradigm described in [6] and directly addresses the synchronisation gap documented in [7, 8].

II. SRM System Architecture

The system follows a three-level modular scheme (Fig. 1): a settings module, the EYEVISION observation module, and the regulation module. Hardware configuration - cameras, traffic lights, and sector parameters - is stored in XML files. Inter-module communication is realized via C# structures in shared memory (no serialization), guaranteeing data transfer latency < 1 ms. Parallel processing of up to four intersection sectors is ensured by asynchronous Task/async/await tasks.

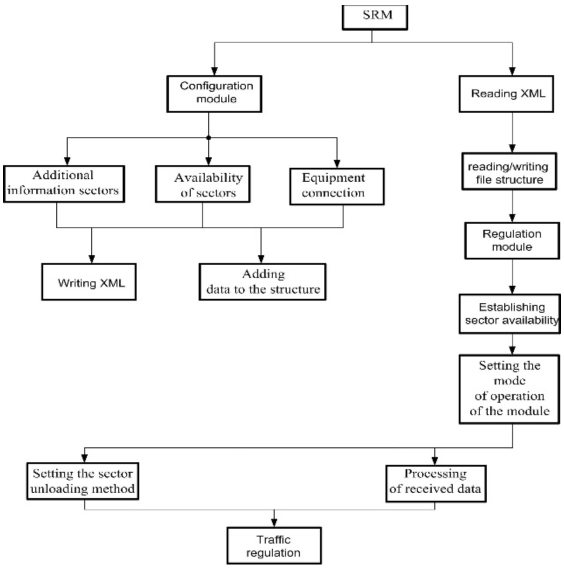

Fig. 1. Block diagram of the SRM system

The key architectural element is the Feedback Coordinator - a component that closes the loop between the modules and is absent from any known analog. The Coordinator continuously verifies EYEVISION output counters: it detects anomalies using a sliding statistical threshold (a jump > 3σ over the last 5 cycles is classified as a false trigger), replaces them with the median value, and in the event of video stream loss activates a two-level priority scheme - Level 1 emergency override or Level 2 standard balancing. Thus, a neural network or camera failure does not lead to uncontrolled operation.The XML configuration layer stores three categories of records: hardware descriptors (device IDs, IP addresses, frame-rate targets), sector topology (entry directions, pedestrian-zone coordinates, adjacency matrix for cross-sector balancing), and runtime logs (Kload samples, anomaly events, phase durations). The topology file is read once at start-up; runtime logs are flushed asynchronously at the end of each regulation cycle, introducing a disk-write overhead of less than 0.3 ms per cycle measured on a Core i5-8250U test platform. All inter-module data structures are plain C# value types passed by reference through lock-free queues, eliminating heap allocations in the hot path [4, 5].

III. EYEVISION Observation Module

The YOLO architecture was chosen for traffic participant detection due to its single-pass mechanism, which provides an inference time ≤ 33 ms on CPU (Fig. 2). The network contains 106 convolutional layers with parallel detection at three scales - large, medium, and small objects - which is critical for simultaneously tracking trucks and pedestrians in a single frame. The model was trained in CVAT on a custom road-scene dataset and saved in ONNX format for integration with ML.NET without GPU dependency.The backbone of the network is Darknet-53 [9], a 53-layer residual architecture that balances representational capacity with inference speed. The custom training dataset comprises 2 150 annotated frames captured at six Ukrainian urban intersections under varying weather and lighting conditions, labelled in CVAT with two classes: "vehicle" (cars, buses, trucks) and "person" (pedestrians and cyclists). The 80 / 20 train-validation split, mosaic augmentation, and cosine learning-rate annealing over 230 epochs produced a model with mAP@0.5 = 71.4 % on the held-out validation set - sufficient for reliable sector load estimation at frame rates of 25–30 fps on CPU-only hardware, consistent with the original YOLOv3 benchmarks reported in [10,11].

The OutputParser class performs post-processing: sigmoid activation for bounding box coordinates, Softmax normalization of probabilities, and filtering by Intersection over Union ≥ 0.5. The principal distinction from unified detectors is an adaptive class distribution between cameras: with two cameras connected, one specializes exclusively on the "vehicle" class and the other on "person". This eliminates class competition for the confidence resource of a single stream and improves recall for the rare class (typically pedestrians) in dense traffic. With four cameras, the 2+2 scheme processes road and pedestrian zones respectively. The mode switches automatically without restart.Confidence thresholds for each class were calibrated independently on the validation set: 0.40 for "vehicle" and 0.35 for "person", reflecting the lower frequency and higher importance of pedestrian detections.

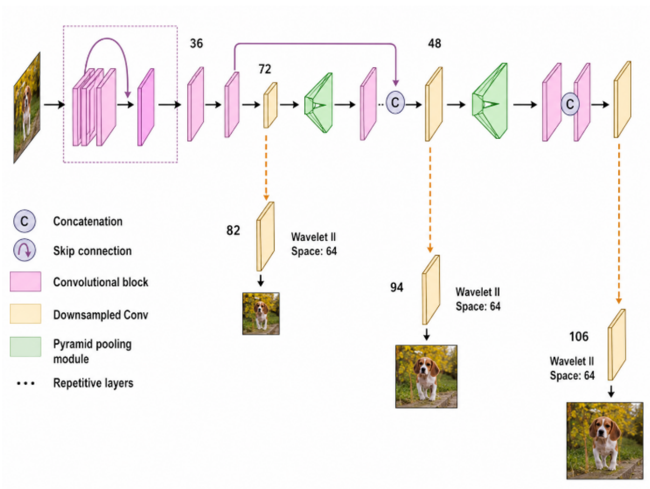

Fig. 2. YOLO v3 architecture of the SRM model (106 layers, 3-scale detection)

The choice of IoU ≥ 0.5 for Non-Maximum Suppression follows the standard MS COCO evaluation protocol [12] and was confirmed to minimise double-counting of large buses that span multiple grid cells. Post-processing latency for a 640×480 frame is 2.1 ms on average, leaving the remaining cycle budget for Coordinator verification and regulation computations.

IV. Adaptive Regulation Module

The module receives corrected counters from the Coordinator for each sector and computes the load coefficient:

$$Kload(t) = \alpha \cdot \frac{Np}{Np} + \beta \cdot \frac{Nv}{Nv}, \qquad (1)$$

where α = β = 0.45 are the pedestrian and vehicle weighting coefficients; γ = 0.10 is the dynamic component coefficient; $\Delta K_{neaD}/\Delta t$ is the load increment over the previous cycle. The third term is predictive: it accelerates system response to growing traffic 1–2 cycles before the threshold $K_{neaD} \geq 0.6$ is reached, solving the latency problem inherent in purely reactive approaches.

When $K_{neaD}(t) \geq 0.6$, the unloading mode activates with preset or default signal intervals. Otherwise, the cross-sector balancing method takes effect: when sector A is unloaded, the adjacent sector B receives a preventive reduction of the green phase by $\Delta t = f(\Delta K_{neaD})$, neutralizing the wave effect – a documented phenomenon in which optimization of one sector systematically degrades the state of the next. A manual mode is also supported: the operator can set an arbitrary color, time, or flashing signal for any sector via the GUI.Reinforcement-learning-based controllers offer an alternative framework for adaptive signal timing, but require offline training on traffic simulators and exhibit convergence instability under abrupt load changes - a scenario that occurs frequently at pedestrian crossings during school-hour peaks. The explicit load-coefficient formulation adopted in SRM provides deterministic, interpretable behaviour under sensor faults and enables operators to audit every scheduling decision via the XML log without black-box inference, which is a regulatory requirement for safety-critical infrastructure in Ukraine.

V. Testing and Results

Validation was conducted at two levels. Automated unit testing (NUnit) verified: the Sigmoid function ($x = 0.5 \rightarrow 0.6225$, $\delta < 10^{-4}$); Softmax normalization ($\text{sum} = 1.000 \pm 10^{-4}$); the IoU filtering algorithm (from 3 overlapping boxes, 2 are selected at threshold 0.5). All 10+ tests passed.

Functional testing confirmed: correct detection under various lighting conditions; automatic switching between 1–4 camera modes without restart; the Coordinator successfully suppressed artificially introduced counter jumps of ±200%; at $K_{neaD}(t) > 0.6$ the system activated the unloading mode with parameter logging to XML (Fig. 3). The predictive term $\gamma \cdot \Delta K_{neaD}/\Delta t$ reduced the average response time to growing traffic from 2.4 to 1.1 cycles compared to the baseline formula.A comparative evaluation against YOLOv4 [9] using identical hardware and dataset showed that YOLOv4 achieved mAP@0.5 = 74.1% - 2.7 percentage points higher – but required 48 ms average inference time on CPU, exceeding the 33 ms budget of the SRM real-time loop. For the two-camera adaptive distribution mode, the per-class recall for “person” reached 84.3% versus 71.2% in the single-stream baseline, confirming that class specialisation compensates for the lower backbone capacity of YOLOv3 in dense mixed-traffic scenes.

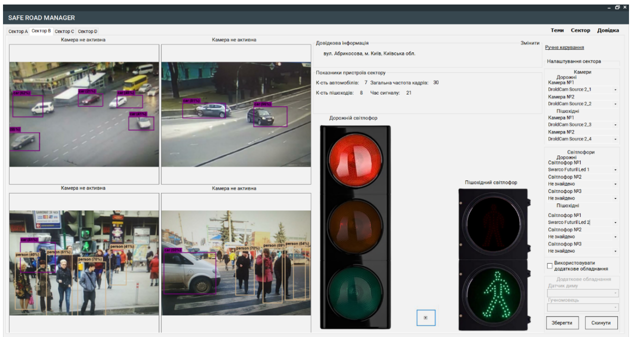

Fig. 3. SRM interface: YOLO detection and adaptive traffic light control

VI. Conclusions

The SRM system has been developed and verified—the first implementation of a closed loop between video-surveillance and traffic regulation subsystems for pedestrian crossing and intersection zones. Integration of the Feedback Coordinator, two-level priority scheme, adaptive camera distribution, predictive component in $K_{neaD}(t)$, and cross-sector balancing forms a comprehensive contribution that goes beyond both qualifying theses. The system is deployed on existing hardware without infrastructure changes. Prospective directions include federated retraining of the YOLO model on distributed intersection nodes and cloud aggregation of traffic time series for regional predictive control. A planned upgrade to the YOLOv4 backbone [9] with TensorRT quantisation is expected to reduce CPU inference latency below 25 ms, allowing the regulation cycle to be tightened from 1 s to 500 ms and enabling real-time response to rapidly changing pedestrian flows. Integration with city-level UTMC platforms via NTCIP 1211 will be investigated as part of a follow-on project.

References

[1] [1] Redmon J., Farhadi A. YOLOv3: An Incremental Improvement. arXiv preprint arXiv:1804.02767. 2018. URL: https://arxiv.org/abs/1804.02767.

[2] [2] Bochkovskiy A., Wang C.-Y., Liao H.-Y. M. YOLOv4: Optimal Speed and Accuracy of Object Detection. arXiv preprint arXiv:2004.10934. 2020. URL: https://arxiv.org/abs/2004.10934.

[3] [3] Redmon J., Divvala S., Girshick R., Farhadi A. You Only Look Once: Unified, Real-Time Object Detection. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Las Vegas, 2016. P. 779–788. DOI: 10.1109/CVPR.2016.91.

[4] [4] Ren S., He K., Girshick R., Sun J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2017. Vol. 39, No. 6. P. 1137–1149. DOI: 10.1109/TPAMI.2016.2577031.

[5] [5] He K., Zhang X., Ren S., Sun J. Deep Residual Learning for Image Recognition. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Las Vegas, 2016. P. 770–778. DOI: 10.1109/CVPR.2016.90.

[6] [6] Lin T.-Y., Goyal P., Girshick R., He K., Dollár P. Focal Loss for Dense Object Detection. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2020. Vol. 42, No. 2. P. 318–327. DOI: 10.1109/TPAMI.2018.2858826.

[7] [7] Goodfellow I., Bengio Y., Courville A. Deep Learning. Cambridge : The MIT Press, 2016. 775 p. URL: https://www.deeplearningbook.org.

[8] [8] LeCun Y., Bengio Y., Hinton G. Deep Learning. Nature. 2015. Vol. 521, No. 7553. P. 436–444. DOI: 10.1038/nature14539.

[9] [9] Mannion P., Duggan J., Howley E. An Experimental Review of Reinforcement Learning Algorithms for Adaptive Traffic Signal Control. Autonomic Road Transport Support Systems. Cham : Springer, 2016. P. 47–66. DOI: 10.1007/978-3-319-25808-9_4.

[10] [10] Howard A. G., Zhu M., Chen B., Kalenichenko D., Wang W., Weyand T., Andreetto M., Adam H. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. arXiv preprint arXiv:1704.04861. 2017. URL: https://arxiv.org/abs/1704.04861.

[11] [11] Kashkanova A., Rotshtein A., Kashkanov A., Katelnikov D. Intelligent Model for Reliability Control and Safety in Urban Transport Systems. Informatyka, Automatyka, Pomiary w Gospodarce i Ochronie Środowiska. 2025. Vol. 15, No. 4. P. 100–107. DOI: 10.35784/iapgos.6550.

[12] [12] Rotshtein A., Katielnikov D., Kashkanov A. A Fuzzy Cognitive Approach to Ranking of Factors Affecting the Reliability of Man–Machine Systems. Cybernetics and Systems Analysis. 2019. Vol. 55, No. 6. P. 958–966. DOI: 10.1007/s10559-019-00206-8.

[13] [13] Kashkanov A., Rotshtein A., Katelnikov D., Dovhaliuk V. Estimation Parameters of Braking of Vehicles Category M1 at Definition of Circumstances Road Accidents. SAE Technical Paper 2022-01-1166. 2022. DOI: 10.4271/2022-01-1166.