

A Perception-to-Decision Vehicle Monitoring System for Highway Work-Zone Safety

Seongyong Kim, S.M.ASCE¹; Minjun Chang²;
Sai Machiraju³; and Yong K. Cho, Ph.D., F.ASCE⁴

¹Graduate Student, Dept. of Civil and Environmental Engineering, Georgia Institute of Technology. Email: skim3310@gatech.edu

²Graduate Student, Dept. of Civil and Environmental Engineering, Georgia Institute of Technology. Email: minjunchang@gatech.edu

³Undergraduate Student, Dept. of Electrical and Computer Engineering, Georgia Institute of Technology. Email: saimachi@gatech.edu

⁴Professor, Dept. of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA (corresponding author). Email: yong.cho@gatech.edu

ABSTRACT

Highway work zones are among the most hazardous traffic environments due to temporary layouts, changing access rules, and limited visibility of vehicle behavior near workers and equipment. Although vision-based systems have improved vehicle detection and tracking, most approaches remain perception-focused and lack structured safety interpretation. This study proposes a perception-to-decision vehicle monitoring framework that integrates visual sensing with explicit work-zone semantics and interpretable safety logic. The system adopts a layered architecture consisting of a Perception Layer, Work-Zone Topology Layer, Semantic Event Layer, and Decision Layer. Vehicle trajectories extracted from a single roadside camera are contextualized using representations of zone geometry and access constraints, enabling classification of safety-relevant events such as compliant passage, encroachment, and unauthorized entry. These events are mapped to discrete safety states through rule-based logic designed for practical field deployment. The framework demonstrates a transparent and computationally efficient approach for structured, decision-oriented safety assessment in dynamic highway work zones.

INTRODUCTION

The growing complexity of construction infrastructure has intensified demand for intelligent, IoT-based monitoring systems that support real-time safety management (Ye et al., 2024). Highway work zones represent particularly hazardous environments due to temporary layouts, shifting traffic patterns, and close proximity between workers and live vehicles (Lv et al., 2019; Micko et al., 2023). Despite improvements in traffic control devices and supervision, unsafe vehicle–worker interactions persist, motivating automated monitoring approaches capable of continuous situational awareness.

Recent work-zone safety systems increasingly employ IoT sensing and vision-based monitoring (Zhou et al., 2022). Cameras are widely adopted because of their low cost, flexible deployment, and ability to capture rich contextual information. Advances in real-time vehicle detection and tracking have therefore become central to many monitoring pipelines, enabling continuous observation of traffic behavior near construction activities.

However, most existing systems remain perception-focused, emphasizing detection accuracy while offering limited interpretability from a safety management perspective. Outputs such as bounding boxes and trajectories are not inherently aligned with work-zone rules, access constraints, or operational safety definitions. As a result, translating perception results into actionable safety decisions remains challenging. Deployment constraints further complicate practical implementation. Temporary roadside systems must operate under limitations in power, computation, and communication bandwidth, often requiring edge-based inference on embedded hardware (Chen et al., 2016). These constraints introduce trade-offs among model complexity, latency, and robustness.

To address these limitations, this study proposes a perception-to-decision vehicle monitoring framework that integrates visual sensing with work-zone topology, semantic event interpretation, and rule-based safety logic, with emphasis on transparency and deployability in real highway work zones.

LITERATURE REVIEW

Highway work-zone safety research has produced various Intelligent Transportation Systems (ITS) designed to reduce vehicle intrusions and enhance situational awareness (Elassy et al., 2024). Existing approaches include speed monitoring, intrusion detection, connected warning systems, and threshold-based alert mechanisms. While effective for specific tasks, many rely on isolated sensor triggers and lack the ability to interpret complex vehicle behavior within dynamically changing work-zone configurations.

Vision-based detection and tracking have become enabling technologies for real-time monitoring (Wang et al., 2022). Deep learning models, particularly YOLO-family detectors (Sapkota & Karkee, 2025), provide a balance between accuracy and inference speed suitable for roadside deployment. These systems support trajectory-level analysis; however, evaluation typically remains perception-centric, with limited integration of safety semantics or decision logic.

Edge deployment introduces additional constraints. Embedded platforms face limitations in memory and computational throughput when executing modern detection networks (Adarsh et al., 2020). Prior studies have explored lightweight architectures, quantization, and hardware acceleration to improve feasibility. Hardware platforms range from GPU-equipped systems to single-board computers and neural accelerators (Chen et al., 2016; Nguyen & Hoang, 2025), reflecting trade-offs between performance, cost, and power consumption.

Despite advances in detection, tracking, and edge optimization, existing literature rarely provides a structured mechanism for transforming perception outputs into interpretable safety decisions within work-zone contexts. The separation between sensing, inference, and decision logic remains a critical gap. This motivates the development of a unified perception-to-decision framework that bridges real-time vision sensing with topology-aware interpretation and actionable safety assessment in highway work zones.

METHODOLOGY

Perception Layer. The perception layer provides vehicle-level observations required for downstream semantic interpretation. Its role is to detect and track vehicles from a fixed roadside camera, without embedding any construction-specific reasoning.

Vehicles are detected in each video frame using a YOLO-based object detector (Sapkota & Karkee, 2025). The detector outputs bounding boxes and confidence scores corresponding to vehicles. The specific YOLO variant is not fixed, allowing flexibility across different deployment environments.

Detected vehicles are associated across frames using a multi-object tracking algorithm to generate continuous trajectories. Each tracked vehicle is assigned a persistent identifier. For each frame, a representative vehicle position is estimated and used as the spatial reference for subsequent processing. The perception layer outputs time-stamped vehicle trajectories, which are passed directly to the work-zone topology and semantic event layers.

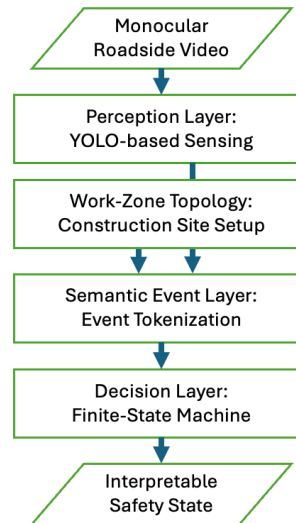


Figure 1. System overview of the perception-to-decision vehicle monitoring framework.

Work-Zone Topology Representation. The monitored scene is partitioned into a small set of polygonal zones that capture safety-relevant spatial semantics:

- OPEN: Normal travel lanes where public traffic is permitted.
- TRANSITION: Merge region where vehicles are expected to adjust trajectory due to shoulder closure.
- RESTRICTED: Taper, closed shoulder or buffer region where public traffic is not permitted.
- WORK: Active construction workspace where construction operations occur.

Highway shoulder construction sites typically exhibit a linear spatial progression along the traffic direction. Accordingly, the zones can be interpreted as an ordered sequence: OPEN → TRANSITION → RESTRICTED → WORK (see Figure 2). This simplified, zone-only representation integrates spatial boundaries and legality semantics directly into zone membership, reducing configuration complexity.

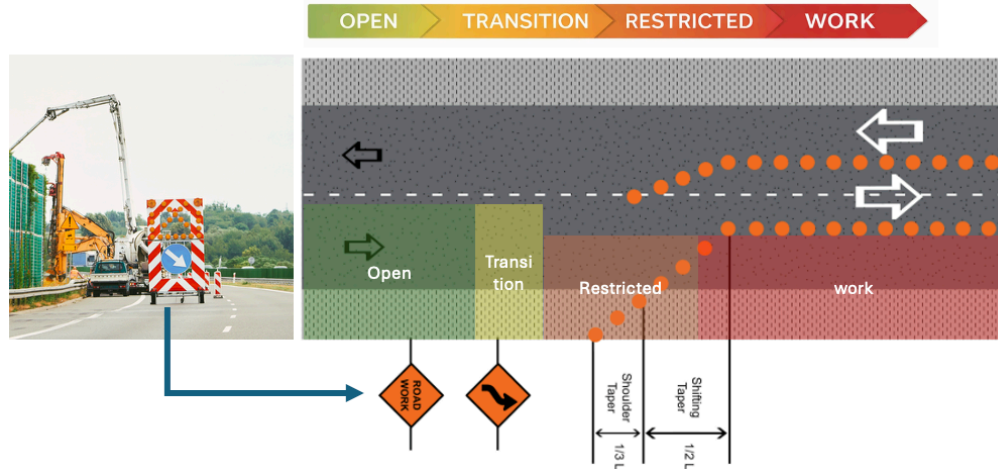


Figure 2. Conceptual illustration of work-zone topology.

Semantic Event Layer. The semantic event layer converts vehicle trajectories into a controlled vocabulary of construction-aware event tokens. The system produces discrete semantic descriptors grounded in work-zone topology and vehicle motion. For each tracked vehicle, a time-ordered sequence of event tokens is generated. The core token categories used in the current implementation are as follows.

Zone state tokens describe the spatial context of the vehicle:

- $\text{zone.state} \in \{\text{OPEN, TRANSITION, RESTRICTED, WORK}\}$
- $\text{zone.transition} = (\text{zone}_{\{t-1\}} \rightarrow \text{zone}_{\{t\}})$
- $\text{zone.legality} \in \{\text{allowed, restricted}\}$

Motion tokens summarize vehicle dynamics using simple kinematic bins computed over short temporal windows:

- $\text{motion.speed} \in \{\text{stopped, slow, normal, fast}\}$
- $\text{motion.proximity} \in \{\text{far, near}\}$

Time-to-zone (TTC) tokens provide predictive timing cues based on the approximately linear structure of highway shoulder work zones. TTC denotes the estimated time-to-crossing into a downstream zone along the vehicle's direction of travel:

- $\text{ttc.restricted} \in \{<1 \text{ s}, 1-3 \text{ s}, >3 \text{ s}, \text{inf}\}$
- $\text{ttc.work} \in \{<1 \text{ s}, 1-3 \text{ s}, >3 \text{ s}, \text{inf}\}$

Authorization tokens encode access constraints associated with construction operations. Note that public traffic observed within the WORK zone is treated as unauthorized:

- $\text{access.authorization} \in \{\text{authorized, unauthorized, unknown}\}$

Decision Layer. The decision layer interprets event token sequences using an interpretable finite-state machine (FSM). A separate FSM instance is maintained for each tracked vehicle, enabling behavior recognition over temporal episodes rather than isolated frames (see Figure 3). The FSM consists of the following states:

-
- The diagram illustrates the Zone State Machine (ZSM) with the following states and transitions:
- Normal** (top center) transitions to **Approaching** (top right) with the condition `zone.state = TRANSITION`.
 - Approaching** transitions to **Suspected Encroachment** (bottom right) with the condition `if ttc.restricted in 1-3s & motion.proximity = "near"`.
 - Suspected Encroachment** transitions to **Encroachment Confirmed** (bottom left) with the condition `zone.transition = "allowed→restricted"`.
 - Encroachment Confirmed** transitions to **Unauthorized Entry** (top left) with the condition `zone.state = WORK & access.authorization = unauthorized`.
 - Unauthorized Entry** transitions back to **Normal** with the condition `WARN (critical)`.
 - Suspected Encroachment** also transitions to **Encroachment Confirmed** with the condition `WARN (primary)`.
 - Encroachment Confirmed** transitions back to **Unauthorized Entry** with the condition `WARN (high priority)`.
- Global exits (apply to ANY state):**
- Resolved** —(compliance restored / end condition)
 - Lost** —(tracking lost / timeout)

State transitions are triggered by combinations of semantic event tokens with persistence requirements to suppress noise. The FSM supports both escalation and early resolution, allowing vehicles to return to a resolved state if compliance is restored, while ensuring that safety-critical violations are fully captured. In addition, global exits to Resolved and Lost Track are permitted from any state.

The decision layer also produces warning outputs as explicit decision-level actions. Warning is treated as an output field rather than a semantic token, reflecting system decisions rather than observed scene properties. For each completed episode, the FSM outputs a behavior label, an associated warning record, and the semantic event tokens responsible for the state transitions.

Data Collection for Feasibility Study. Preliminary video data were collected along Interstate 85 in Georgia to evaluate the feasibility of the proposed framework in a highway shoulder environment (Figure 4). Video was recorded from a camera mounted at a fixed roadside location, capturing realistic traffic flow and vehicle motion patterns commonly observed near shoulder work

zones. This fixed-camera setup is consistent with the assumptions supports stable spatial reasoning, zone assignment, and time-to-crossing estimation without requiring online calibration.



Figure 4. Roadside video data collection setup along I-85 in GA.

In parallel with data collection, a proof-of-concept sensing system was developed to assess the feasibility of edge-based deployment. The prototype integrates low-cost hardware with a YOLO-based computer vision pipeline capable of detecting and tracking vehicles in real time (Sapkota & Karkee, 2025). Vehicle bounding boxes enable proximity and speed estimation through a calibrated pixel-to-distance mapping, while the use of an embedded AI accelerator supports higher frame rates under practical power constraints. An interactive interface was developed for constructing work-zone topology through polygonal zone definition, enabling explicit encoding of spatial and access semantics for downstream event interpretation (Fig. 5).



Figure 5. Zone editing interface for constructing work-zone topology for downstream event interpretation.

Interpretable Decision Outcomes and Warning Generation. Using the token streams produced by the semantic event layer, vehicle behavior is evaluated through the interpretable decision structure, resulting in behavior states and corresponding warning outputs. For each tracked vehicle,

the decision layer maintains a temporal state progression and emits warning outputs when predefined conditions are met. Warning is treated as a decision-level output, characterized by its severity, triggering reason, and associated decision state.

Table 1 shows an example decision trace for a single vehicle approaching and entering a restricted shoulder region. The example illustrates how early semantic cues lead to progressive state transitions and warning escalation.

Table 1. FSM Decision outcomes and warning generation.

time (s)	Decision State	warn.state	warn.level	warn.reason
8.43	Approaching	zone.state = TRANSITION	primary	ttc_low
9.12	Suspected Encroach.	ttc.restricted = 1–3 s	high priority	ttc_low
10.02	Encroach. Confirmed	zone.transition allowed→restricted	critical	entered_restricted
11.45	Resolved	return to allowed zone	none	—

This example demonstrates that the proposed framework produces transparent, stage-wise decision outcomes, where warning severity increases as safety conditions deteriorate and resolves once compliance is restored. Because each decision and warning is explicitly tied to semantic event tokens, the resulting behavior interpretation remains traceable and human understandable.

Discussion. Given limited volume of roadside scenarios, the system successfully converts vehicle trajectories into structured semantic event tokens and interpretable decision outcomes, including progressive warning generation. These results confirm that construction-aware semantics and decision logic can be derived with deployability and scalability.

A key observation is that the zone-based topology and time-to-crossing cues provide sufficient structure to capture meaningful behavioral progression, from compliant approach to restricted-zone entry. The resulting decision outputs are transparent and traceable, allowing each warning to be explained through explicit semantic tokens and state transitions. This interpretability is important for construction safety applications, where decision rationale must be understandable to site managers and operators.

CONCLUSION

This study presented a perception-to-decision vehicle monitoring framework for highway work zones that integrates vision-based detection with topology-aware interpretation and rule-based safety logic. Rather than treating detection and tracking as end goals, the proposed system structures perception outputs within explicit work-zone geometry and access constraints to derive safety-relevant semantic events and decision states. The layered architecture supports transparent and explainable safety assessment while remaining compatible with edge-based deployment under temporary roadside constraints.

Future work will focus on real-world field validation and end-to-end system integration. In particular, the framework will be connected to physical alerting mechanisms, such as wearable or roadside warning devices, to evaluate latency, reliability, and operational usability. Quantitative evaluation using safety-oriented metrics, including alert accuracy, missed hazardous events, and false positives, will be conducted through expanded data collection across diverse work-zone configurations and traffic conditions.

ACKNOWLEDGEMENT

This material is based upon work supported by the Georgia Department of Transportation (GDOT) under Project No. RP 25-08. The views expressed herein are those of the authors and do not necessarily reflect the views of GDOT.

REFERENCES

- Adarsh, P., Rathi, P., & Kumar, M. (2020). YOLO v3-Tiny: Object Detection and Recognition using one stage improved model. *2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS)*, 687–694.
- Chen, Y.-H., Krishna, T., Emer, J. S., & Sze, V. (2016). Eyeriss: An energy-efficient reconfigurable accelerator for deep convolutional neural networks. *IEEE Journal of Solid-State Circuits*, 52(1), 127–138.
- Elassy, M., Al-Hattab, M., Takruri, M., & Badawi, S. (2024). Intelligent transportation systems for sustainable smart cities. *Transportation Engineering*, 16, 100252.
- Lv, Z., Hu, B., & Lv, H. (2019). Infrastructure monitoring and operation for smart cities based on IoT system. *IEEE Transactions on Industrial Informatics*, 16(3), 1957–1962.
- Micko, K., Papcun, P., & Zolotova, I. (2023). Review of IoT sensor systems used for monitoring the road infrastructure. *Sensors*, 23(9), 4469.
- Nguyen, H. H., & Hoang, M. S. (2025). LEAF-YOLO: Lightweight Edge-Real-Time Small Object Detection on Aerial Imagery. *Intelligent Systems with Applications*, 25, 200484.
- Sapkota, R., & Karkee, M. (2025). Ultralytics YOLO evolution: An overview of YOLO26, YOLO11, YOLOv8 and YOLOv5 object detectors for computer vision and pattern recognition. *ArXiv Preprint ArXiv:2510.09653*.
- Wang, Z., Zhan, J., Duan, C., Guan, X., Lu, P., & Yang, K. (2022). A review of vehicle detection techniques for intelligent vehicles. *IEEE Transactions on Neural Networks and Learning Systems*, 34(8), 3811–3831.

- Ye, Z., Wei, Y., Yang, S., Li, P., Yang, F., Yang, B., & Wang, L. (2024). IoT-enhanced smart road infrastructure systems for comprehensive real-time monitoring. *Internet of Things and Cyber-Physical Systems*, 4, 235–249.
- Zhou, H., Xu, C., Tang, X., Wang, S., & Zhang, Z. (2022). A review of vision-laser-based civil infrastructure inspection and monitoring. *Sensors*, 22(15), 5882.