J Intel Connect Veh

Research Article

Vehicle-Dynamics-Aware Motion Planning for Pothole-Hazard Mitigation

Xiang Wang[1,†], Scott Piersall[2,†], Zihang Zou[3], Liqiang Wang[2], Rongjie Yu[1,✉]

Cite this article: https://doi.org/10.26599/JICV.2026.9210098

ABSTRACT: Potholes, recognized as the second most frequent pre-crash event, severely compromise traffic safety while posing critical threats to vehicle structural safety and ride comfort. Existing countermeasures, primarily active suspension control and longitudinal speed regulation, struggle to mitigate these structural impacts without increasing the probability of multi-vehicle conflicts, particularly rear-end collisions. Leveraging the extended preview information regarding pothole geometry and locations provided by vehicle-to-everything (V2X) communication, this study proposes a vehicle-dynamics-aware motion planning framework that introduces proactive intra-lane lateral maneuvering to expand conventional mitigation strategies. The framework systematically integrates microscopic tire-pothole impact mechanics, vertical quarter-car dynamics, and car-following behaviors into a unified closed-loop simulation environment. A curriculum learning strategy is incorporated into the Soft Actor-Critic (SAC) algorithm to progressively increase task complexity, thereby mitigating the training instability caused by abrupt and sparse physical impact penalties. Comprehensive simulations covering diverse pothole geometries, preview distances, speed ranges, and car-following interactions validate the framework. For geometrically avoidable hazards, the trained policy executes proactive lateral bypassing, achieving a 95.8% impact-load compliance rate while reducing travel time over the defined pothole-passage interval by approximately 50%. Under unavoidable conditions, the policy performs adaptive speed regulation, achieving a 46.7% impact-load compliance rate while reducing the rear-end collision rate from 13.3% to 1.7% compared with braking-heavy baselines. Overall, this approach expands vehicle capabilities in localized hazard mitigation, demonstrating the potential of integrating low-level physical dynamic boundaries into motion planning frameworks.

KEYWORDS: pothole handling; intelligent connected vehicles; motion planning; vehicle dynamics; traffic safety; reinforcement learning; soft actor-critic (SAC)

1 Introduction

Pothole hazards pose severe operational risks to both human-driven and autonomous vehicles, constituting one of the most frequent pre-crash events (National Highway Traffic Safety Administration, 2002, 2006, 2009). When encountering these road-surface defects, drivers typically suffer from limited preview time and delayed reaction capabilities. Consequently, rather than executing smooth and safe maneuvers, they are frequently forced into instinctive and abrupt evasive actions, such as severe hard braking or sudden steering. These abrupt maneuvers destabilize the ego vehicle, disrupt following traffic, and may contribute to single-vehicle road departures and multi-vehicle rear-end collisions (Ahmad et al., 2023; Khattak et al., 2021). Furthermore, the high-speed traversal of these geometric discontinuities generates transient impact loads capable of damaging chassis and suspension components (Radu et al., 2023; Zhang et al., 2020). Such widespread structural degradation results in substantial economic losses, inflicting an estimated $3 billion in annual vehicle repair costs (AAA East Central, 2024; AAA Western and Central New York, 2023). Therefore, mitigating pothole hazards can be viewed as a coupled challenge of ensuring vehicle structural safety and maintaining interactive traffic stability, which urgently requires proactive interventions through intelligent autonomous driving systems.

To mitigate pothole hazards, vehicles rely primarily on active suspension control and longitudinal speed regulation (Audi UK, 2017; BYD Global, 2023; Huang et al., 2023; Wu et al., 2020; Zhang et al., 2023). These interventions depend heavily on onboard line-of-sight sensors, such as cameras and LiDAR, to detect road-surface defects (Chaaban et al., 2025; Daimler Communications, 2017; Jaguar Land Rover, 2015; Zhao et al., 2024), which inherently restricts the preview horizon due to sensing ranges and visual occlusions. While these systems provide essential localized detection and enable basic impact mitigation (Huang et al., 2023; Liu et al., 2024; Theunissen et al., 2021; Wu et al., 2020), such strategies exhibit inherent physical and interactive limitations. While active suspensions are designed to absorb vertical shocks, they are constrained by limited actuator strokes; under severe pothole excitations, the suspension easily exhausts its travel limit, transmitting damaging rigid impacts directly to the vehicle chassis (Theunissen et al., 2021; Wu et al., 2020). Similarly, although longitudinal speed regulation lessens impact severity, the requisite aggressive braking abruptly disrupts traffic flow and exacerbates rear-end collision risks, particularly in tight car-following scenarios (Du et al., 2022; Sun et al., 2018). Crucially, conventional systems manage vertical impacts and longitudinal kinematics in isolation, forcing a rigid compromise: prioritizing chassis protection endangers rear-end safety, whereas sustaining longitudinal velocity risks severe structural damage (Du et al., 2022; Hussain et al., 2025; Wu et al., 2020; Zhang et al., 2023).

† Xiang Wang and Scott Piersall contributed equally to this work.

[1] College of Transportation, Tongji University, Shanghai 201804, China. [2] Department of Computer Science, University of Central Florida, Orlando FL 32816, USA. [3] Optixway AI, Orlando FL 32792, USA.

✉ Corresponding author. E-mail: yurongjie@tongji.edu.cn

Received: June 18, 2026; Revised: August 6, 2026; Accepted: August 31, 2026

The advent of vehicle-to-everything (V2X) communication and cloud-based hazard services facilitates beyond-visual-range information sharing (DiLorenzo and Yu, 2023; Jaguar Land Rover, 2015). By aggregating and broadcasting defect data across connected networks, these technologies establish an extended preview horizon that overcomes the spatial limitations of conventional line-of-sight sensors (Du et al., 2022; Zhang et al., 2023). This extended spatial awareness empowers intelligent vehicles to acquire pothole information well before reaching the hazard, securing the crucial time window required for proactive interventions. Specifically, the extended preview unlocks the operational space necessary for tactical motion planning, given our empirical finding that 71.4% of potholes can be safely bypassed within the ego-lane. It enables the execution of smooth, coordinated lateral-longitudinal maneuvers—such as proactive intra-lane bypassing or adaptive speed regulation.

To exploit the extended operational space provided by connected hazard information, this study proposes a vehicle-dynamics-aware motion planning framework for proactive pothole-hazard mitigation. Specifically, by constructing a reinforcement learning environment that integrates multi-scale simulations, the proposed framework unifies transient tire-pothole impact mechanics and interactive traffic constraints within a closed-loop evaluation system. The main contributions of this study are summarized as follows:

1) A vehicle-dynamics-aware motion planning framework is proposed to expand vehicle capabilities in mitigating pothole hazards, underpinned by a multi-scale closed-loop simulation environment. This environment systematically integrates microscopic tire-pothole impact mechanics, vertical quarter-car dynamics, and car-following behaviors, thereby providing the multi-domain physical feedback necessary to drive reinforcement learning policy optimization. Through the learned policy, the framework enables flexible maneuvers, including proactive intra-lane lateral bypassing, straddling, and adaptive speed regulation.

2) A curriculum learning strategy is integrated into the Soft Actor-Critic (SAC) framework to mitigate the learning instability caused by abrupt and sparse physical impact penalties. By progressively exposing the SAC agent to increasingly complex pothole geometries and traffic conditions, this tailored training paradigm ensures robust policy convergence in the highly non-convex, multi-objective pothole-handling task.

The remainder of this paper is organized as follows: Section 2 reviews related literature on vehicle control technologies for pothole-hazard driving and reinforcement learning applications. Section 3 presents the proposed methodology, detailing the physics-informed impact mechanics model, the vehicle-dynamics-aware motion planning framework, and the curriculum learning strategy. Section 4 describes the experimental setup. Section 5 presents the experimental results, focusing on learning dynamics and emergent pothole-handling strategies. Section 6 presents an illustrative sequential-hazard case and discusses study limitations. Finally, Section 7 summarizes the main findings and outlines directions for future research.

2 Literature Review

2.1. Vehicle Control Technologies for Pothole-Hazard Driving

Research on pothole hazards has gradually evolved toward integrated vehicle control and planning methodologies that exploit preview information to reduce impact loads and ride discomfort while ensuring executable motion (Chaaban et al., 2025; Jaguar Land Rover, 2015). At the system level, numerous studies have incorporated pothole detection and evasion into autonomous driving architectures, where perception modules provide pothole localization together with geometric information, and subsequent planning or control modules generate feasible avoidance strategies (Chaaban et al., 2025; Daimler Communications, 2017; Jaguar Land Rover, 2015; Zhao et al., 2024). These integrated systems improve the feasibility of pothole-aware autonomous driving; however, many existing formulations still treat pothole handling primarily as a geometric avoidance problem. As a result, planning and control layers often overlook physically informed objectives and constraints that directly characterize transient wheel–pothole interactions and the resulting vertical dynamics, even though these factors are critical for evaluating structural risk and ride comfort during pothole encounters.

Within the control and motion-planning literature, pothole impact mitigation has mainly developed along two complementary directions, namely preview-based suspension control and longitudinal speed regulation. Preview-based suspension control utilizes look-ahead road information to proactively adjust active or semi-active suspension responses, thereby reducing body vibration and wheel-load variation compared with purely reactive strategies (Theunissen et al., 2021). In parallel, longitudinal speed regulation mitigates pothole-induced excitation by adapting the traversal speed before wheel entry, because the resulting vertical disturbance jointly depends on road geometry and vehicle speed (Sun et al., 2018). Building on these insights, recent studies have increasingly combined preview suspension control with velocity planning to jointly balance ride comfort, longitudinal smoothness, travel efficiency, and impact mitigation (Huang et al., 2023; Liu et al., 2024; Wu et al., 2020). Nevertheless, these approaches remain largely centered on longitudinal mitigation of the previewed disturbance, with limited consideration of lane-bounded lateral maneuvering opportunities and interactive traffic constraints.

Beyond comfort-oriented control, pothole-aware planning has increasingly incorporated safety considerations arising from vehicle interactions. Recent work has inferred road-induced vibration from the vertical response of a leading vehicle, thereby enabling anticipatory speed planning while maintaining interaction safety (Zhang et al., 2023). Such developments suggest that pothole-hazard driving is no longer merely a problem of individual vehicle vibration management, but rather an interactive planning and control problem under limited observability.

2.2. Reinforcement Learning for Motion Planning and Vehicle Control

Reinforcement learning (RL) has been widely studied as a unifying framework for motion planning and vehicle control because of its ability to optimize policies in nonlinear and stochastic environments through iterative interaction. By encoding multiple objectives within a structured reward design, RL enables coordinated optimization across safety, comfort, efficiency, and constraint satisfaction. Existing surveys have categorized RL applications throughout the autonomous driving stack, including high-level behavior planning in interactive traffic, motion planning under complex kinematic and dynamic constraints, and low-level tracking and stabilization control (Aradi, 2022; J. Hu et al., 2025; Wu et al., 2024). Collectively, these studies indicate that RL is particularly advantageous in long-tailed or highly interactive scenarios where handcrafted rules are often insufficient.

A parallel line of research has investigated RL for vehicle control problems characterized by strong dynamic coupling and parametric uncertainty. In such settings, conventional controllers may be sensitive to modeling inaccuracies or varying operating conditions. Representative applications include suspension vibration mitigation, preview-based chassis control, and speed planning over rough or degraded road surfaces, all of which involve nonlinear dynamics and competing performance objectives (Du et al., 2022; Wang et al., 2023; Zhao et al., 2024). Recent developments have increasingly incorporated physics-guided and constraint-aware learning formulations, in which action spaces are restricted to physically meaningful control variables and safety-critical bounds are explicitly enforced to improve deployability and reduce unsafe exploration (Zhao et al., 2024).

Safety concerns have further driven the development of safe RL for autonomous driving, since unconstrained exploration is incompatible with safety-critical applications. Recent reviews have emphasized constrained optimization, risk-sensitive objectives, and formal safety filters designed to enhance constraint satisfaction during both training and deployment (Fang et al., 2025; Inamdar et al., 2024). Hybrid architectures that combine RL with model predictive control (MPC) have also emerged, leveraging the structural guarantees and constraint-handling capability of MPC while allowing RL to adapt controller parameters across varying conditions (Wang et al., 2023).

These studies suggest that practical RL-based planning and control benefits from structured formulations that explicitly unify physical constraints, interaction safety, and real-time feasibility within a closed-loop evaluation framework.

In the context of pothole-hazard driving, the key challenge lies not merely in selecting RL as an optimization tool, but in constructing an RL-compatible problem formulation and training environment that can accurately capture wheel–pothole impact mechanics, vertical vehicle dynamics, and interaction-safety constraints within surrounding traffic. This perspective motivates the development of multi-physics simulation environments and maximum-entropy RL formulations capable of learning motion planning strategies under sparse reward feedback and heterogeneous pothole geometries while maintaining physically interpretable safety compliance (Aradi, 2022; Fang et al., 2025; Inamdar et al., 2024).

3 Methodology

This section presents the proposed hierarchical framework for pothole-aware motion planning. As illustrated in Fig. 1, the framework integrates tactical motion planning with detailed modeling of vehicle-road interaction mechanics within a closed-loop, multi-physics simulation environment. The architecture simulates interactive traffic dynamics concurrently with mechanism-informed structural impacts to assess risks and define failure boundaries. Within this hierarchy, a high-level RL agent processes the scenario state and outputs a tactical action consisting of a target speed and a target lateral offset.

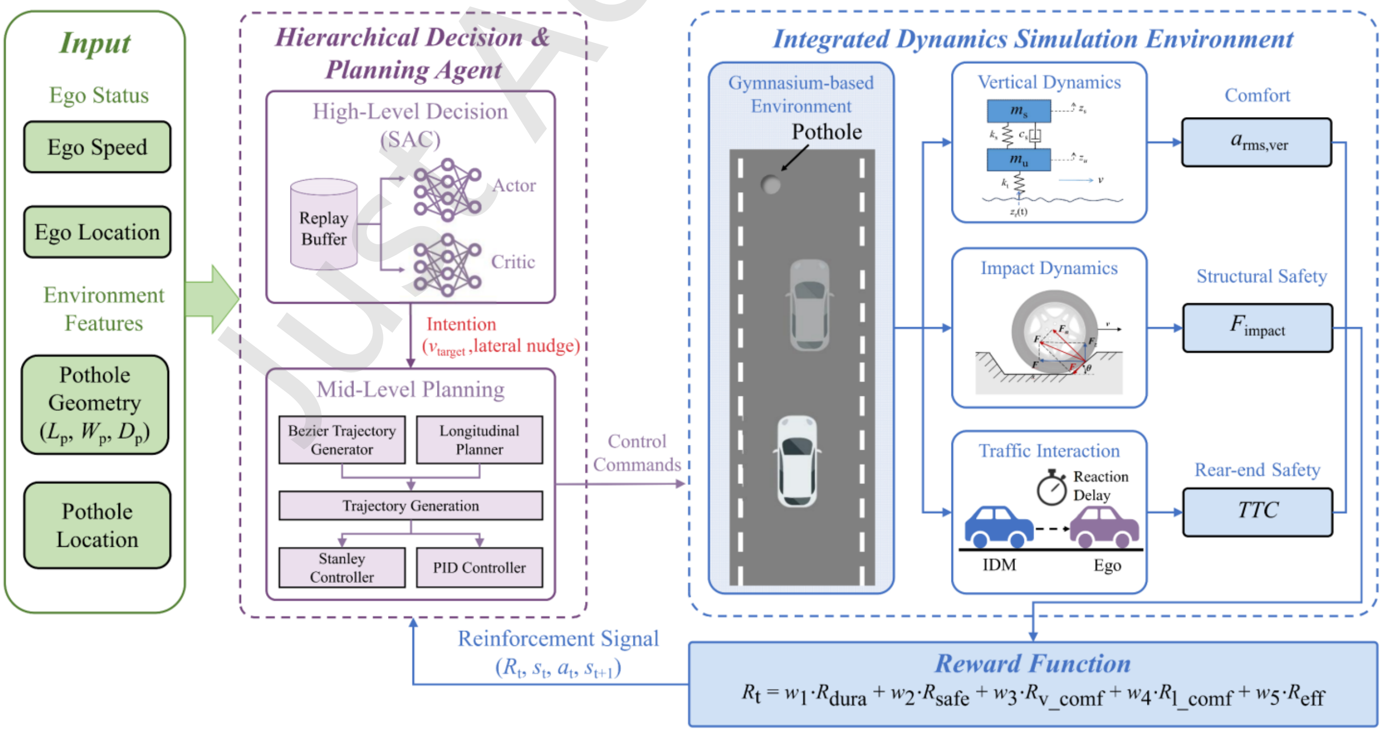


Fig. 1. Overall framework of the proposed vehicle-dynamics-aware hierarchical autonomy for pothole hazards.

3.1. Impact Mechanics Modeling

To quantitatively evaluate the risk of structural damage during pothole traversal, we model the problem from a multibody contact dynamics perspective rather than relying solely on classical vertical vibration theory (Radu et al., 2023; Zhang et al., 2020). This formulation is necessary to capture the impulsive loads associated with tire and rim failures.

The adopted quarter-car and tire–pothole impact models quantify vertical response and transient wheel-impact loads but do not represent coupled lateral–vertical dynamics, vehicle roll, lateral load transfer, or tire-force saturation.

3.1.1. *Suspension Dynamics via Quarter-Car Model*

The vehicle vertical response is modeled using a 2-DOF quarter-car system consisting of a sprung mass $\boldsymbol{m_s}$ and an unsprung mass $\boldsymbol{m_u}$. The upward vertical direction is defined as positive. The variables $\boldsymbol{z_s}$ and $\boldsymbol{z_u}$ denote the dynamic displacements of the sprung and unsprung masses from their static equilibrium positions, respectively, while $\boldsymbol{z_r}$ denotes the road-profile displacement relative to the nominal road plane. Because the equations are expressed relative to static equilibrium, gravitational and static tire-load terms do not explicitly appear.

Defining the suspension deflection and relative velocity as

$$\Delta z = z_s - z_u, \quad \Delta \dot{z} = \dot{z}_s - \dot{z}_u,$$

the equations of motion are

$$\begin{cases} m_s \ddot{z}_s = -F_s, \\ m_u \ddot{z}_u = F_s - F_{\mathrm{tire,z}}, \end{cases} \tag{1}$$

where $\boldsymbol{F_s}$ is the force transmitted through the suspension and $\boldsymbol{F_{\mathrm{tire,z}}}$ is the signed vertical tire-restoring force associated with the road-profile input.

The suspension force comprises the linear spring force, asymmetric damping force, and bump-stop force:

$$F_s = k_s \Delta z + F_{\mathrm{damp}} + F_{\mathrm{bs}}, \tag{2}$$

$$F_{\mathrm{damp}} = \begin{cases} c_{s,\mathrm{reb}} \Delta \dot{z}, & \Delta \dot{z} \geq 0, \\ c_{s,\mathrm{comp}} \Delta \dot{z}, & \Delta \dot{z} < 0, \end{cases} \tag{3}$$

and

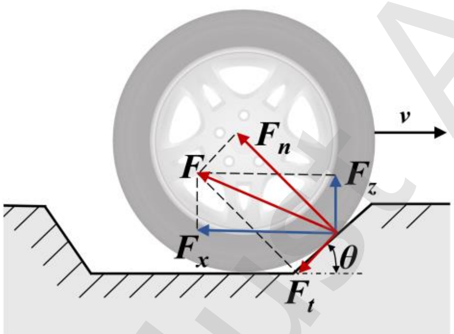


Fig. 2. Proposed mechanism-informed model for tire–pothole edge impact analysis.

$$F_{\mathrm{bs}} = \begin{cases} k_b(\Delta z - z_{\max}), & \Delta z > z_{\max}, \\ k_b(\Delta z + z_{\max}), & \Delta z < -z_{\max}, \\ 0, & |\,\Delta z\,| \leq z_{\max}, \end{cases} \tag{4}$$

Under the continuous-contact approximation, the signed vertical tire-restoring force is expressed as

$$F_{\mathrm{tire,z}} = k_t(z_u - z_r) + c_t(\dot{z}_u - \dot{z}_r), \tag{5}$$

where $\boldsymbol{k_t}$ and $\boldsymbol{c_t}$ are the vertical stiffness and damping coefficient of the tire, respectively.

Eqs. (1)–(5) describe the road-profile-induced quarter-car response. The wheel–pothole edge-impact model introduced in Section 3.1.2 is evaluated separately, and its contact force is not superimposed on Eq. (1). The quarter-car model provides the sprung-mass acceleration and dynamic tire-load response, whereas the edge-impact model provides the peak contact load used in the structural-durability evaluation. This separation avoids double counting of the pothole excitation.

3.1.2. *Wheel–Pothole Impact Dynamics*

When the wheel encounters the leading edge of a pothole at high speed, a short-duration but high-magnitude impact load is generated (Radu et al., 2023; Zhang et al., 2020). To capture this process, we adopt a reduced-order contact-impact formulation that accounts for the wheel's forward motion and the local wheel–edge contact geometry, as illustrated in Fig. 2. At the onset of contact, the critical contact angle θ_c is determined by the pothole depth D_p and the effective tire radius R_t:

$$\theta_c = \arccos\left(1 - \frac{D_p}{R_t}\right) \tag{6}$$

where D_p denotes the pothole depth and R_t denotes the effective tire radius. The normal contact force is modeled using a penetration-based nonlinear contact law with a Hunt–Crossley-type velocity correction:

$$F_n = \max(K_c \delta^n (1 + \alpha_{hc} v_c), 0) \tag{7}$$

where K_c is the effective contact stiffness, δ is the normal penetration depth, n is the contact force exponent, α_{hc} is the velocity correction coefficient, and v_c is the compressive normal relative velocity at the contact point.

Tangential interaction is modeled using Coulomb sliding friction:

$$F_f = \mu F_n \tag{8}$$

where μ is the sliding friction coefficient. The resultant contact force is decomposed into longitudinal and vertical components according to the instantaneous contact angle θ:

$$\begin{cases} F_x = F_n \sin\theta + F_f \cos\theta \\ F_z = F_n \cos\theta - F_f \sin\theta \end{cases} \tag{9}$$

To quantify the severity of the edge strike, we define the peak resultant wheel-hub load during the impact interval as

$$F_{peak} = \max_{t \in [0, T_{imp}]} \sqrt{F_x^2(t) + F_z^2(t)} \tag{10}$$

where T_{imp} denotes the effective contact duration. This physics-based metric is used as the impact-load indicator in the subsequent tactical decision-making process. The key parameters of the impact model are summarized in Table 3.

3.2. *Hierarchical Planning Framework*

The proposed framework adopts a vehicle-dynamics-aware hierarchical architecture that separates tactical action generation from trajectory realization. The internal architecture of the planning agent is illustrated in Fig. 3. As shown in Fig. 3, the system operates through a closed-loop interaction between three core modules: a high-level strategic layer, a low-level execution layer, and a physics-informed simulation environment. Unlike conventional continuous control where the policy operates at a high frequency (e.g., 10 Hz), our tactical decision-making is formulated as an episodic, single-step optimization problem. Upon previewing a hazard, the agent selects a one-step tactical decision, and the episode terminates after the low-level controllers execute the maneuver and physical outcomes are evaluated. Consequently, the process is modeled as a continuous-action contextual bandit problem, or equivalently, a single-step Markov Decision Process (MDP) defined by the tuple $(\mathcal{S}, \mathcal{A}, \mathcal{R})$ without a temporal discount factor.

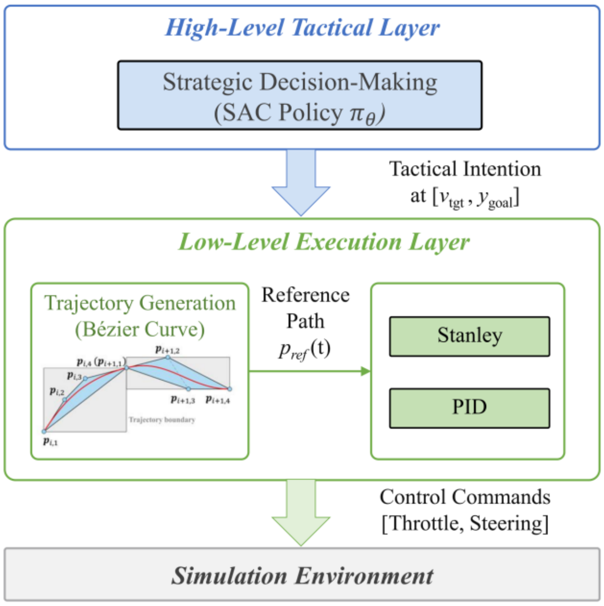


Fig. 3. Architecture of the hierarchical decision-making and low-level execution layers.

- **State Space$\mathcal{S}$:** Encodes the initial mission context $s0$ at the exact moment the pothole is previewed, including the ego vehicle's kinematics (e.g., initial velocity $\boldsymbol{v_{ego}}$), the pothole's geometric attributes (length L_p, width W_p, depth D_p), and the relative state of surrounding traffic.
- **Action Space$\mathcal{A}$: A continuous** space representing the one-step tactical action $\mathbf{u} = [\boldsymbol{v_{tgt}}, \boldsymbol{y_{tgt}}]$. Here, $\boldsymbol{v_{tgt}}$ denotes the target traversal velocity, and $\boldsymbol{y_{tgt}}$ specifies the target lateral offset within the lane boundaries.
- Reward $\mathcal{R}$: Maps the selected tactical action and its resulting episode outcome to a scalar return that encodes trade-offs among passage efficiency, rear-end safety, ride comfort, and structural durability.

The high-level policy $\pi_\theta(\boldsymbol{u} \mid s_0)$ is optimized using the Soft Actor-Critic (SAC) algorithm (Haarnoja et al., 2018). Although typically applied to sequential tasks, SAC is particularly advantageous in this single-step context due to its maximum-entropy framework. In our highly constrained environment, naive optimization frequently causes the policy to collapse into suboptimal local minima (e.g., passive hard braking without lateral avoidance). SAC's entropy regularization promotes robust exploration in the continuous tactical action space. Under this single-step formulation, the critic is trained from the terminal reward of the selected tactical action, while the entropy term helps preserve exploration diversity. The selected tactical action is subsequently executed by the low-level layer, where a quintic Bézier generator constructs a smooth reference path and the Stanley and PID controllers generate the steering and longitudinal commands. The lateral offset, steering command, and curvature-dependent execution speed are bounded during trajectory generation and tracking, as detailed in Section 4.1.

3.3. Reward Function

The reward function is designed to balance passage efficiency, rear-end interaction safety, ride comfort, and structural durability. Because the framework operates as a single-step decision process and the low-level controller is designed to complete the maneuver, premature termination cannot be exploited to avoid penalties. Accordingly, the episode return is formulated as the negative weighted sum of physically grounded penalty terms. The total reward $\boldsymbol{R_{total}}$ is defined as

$$\boldsymbol{R_{total} = W_{eff}\, r_{eff} + W_{rear}\, r_{rear} + W_{comf}\, r_{comf} + W_{dura}\, r_{dura}} \quad (11)$$

where $\boldsymbol{W_{eff}, W_{rear}, W_{comf}, W_{dura} > 0}$ are scalar weights that prioritize competing objectives, and each $\boldsymbol{r \le 0}$ is a physically grounded penalty term computed over the episode rollout.

3.3.1. Passage Efficiency

To discourage overly conservative behaviors such as unnecessary stopping or crawling, the agent is penalized for the squared relative delay. To support magnitude normalization across heterogeneous scenarios, the efficiency penalty is normalized by the reference travel time:

$$\boldsymbol{r_{eff} = -\left(\frac{\max\left(0,\, T_{act} - T_{ref}\right)}{T_{ref}}\right)^2} \quad (12)$$

where T_{act} denotes the elapsed time from episode initialization, with the ego vehicle located at $x = 0$, until it reaches $x = x_p + 5$ m, where x_p is the longitudinal position of the pothole center. The reference travel time is defined as $T_{ref} = (x_p + 5)/v_{init}$, representing the time required to cover the same distance at the initial vehicle speed v_{init}.

3.3.2. Traffic Safety

Rear-end collision risk induced by aggressive braking is quantified using the inverse time-to-collision (iTTC) relative to the following vehicle. The traffic-safety penalty is accumulated over the episode rollout as

$$\boldsymbol{r_{rear} = -\sum_{k=1}^{N}\left(\max(0,\, iTTC_k - iTTC_{thr})\right)^2} \quad (13)$$

Following empirical car-following safety analyses (Hussain et al., 2025), we set $iTTC_{thr} = 0.33$ s^{-1} (corresponding to TTC = 3.0 s) to maintain a conservative car-following safety margin.

3.3.3. Ride Comfort

Ride comfort is evaluated using ISO-weighted sprung-mass accelerations in the vertical and longitudinal directions. To accurately reflect human perception, the raw accelerations are filtered using the ISO 2631-1 frequency weighting, yielding ISO-weighted sprung-mass accelerations az,w and ax,w. The unified comfort penalty is defined as

$$\boldsymbol{r_{comf} = -\left(RMS\left(a_{z,w}\right)^2 + w_{lon}\, RMS\left(a_{x,w}\right)^2\right)} \quad (14)$$

where $\boldsymbol{w_{lon}}$ balances longitudinal braking intensity against the primary vertical suspension response. Here $\boldsymbol{RMS(x) = \sqrt{\frac{1}{N}\sum_{k=1}^{N} x_k^2}}$ is computed over episode samples, and the ISO 2631-1 weighting is realized in discrete time at the simulator sampling interval $\boldsymbol{\Delta t}$.

3.3.4. Structural Durability

This component explicitly enforces the vehicle-dynamics-aware safety constraint, targeting the risk of wheel and rim damage during edge strikes. The penalty activates when the peak resultant force F_{peak} exceeds a warning threshold F_{warn}:

$$\boldsymbol{r_{dura} = -\left(\frac{\max\left(0,\, F_{peak} - F_{warn}\right)}{F_{lim} - F_{warn}}\right)^2} \quad (15)$$

where $\boldsymbol{F_{lim}}$ denotes the catastrophic rim failure limit. Based on full-vehicle impact experiments (Zhang et al., 2020), we set $\boldsymbol{F_{warn}} = \mathbf{10.6}\ \boldsymbol{kN}$ to represent the onset of potential mechanical damage and $\boldsymbol{F_{lim}} = \mathbf{15.15}\ \boldsymbol{kN}$ as the hard physical boundary.

3.3.5. Reward Calibration

The objective weights are empirically tuned following the principle of magnitude normalization, such that the unweighted penalty components (r_{eff}, r_{rear}, r_{comf}, r_{dura}) remain within the same order of magnitude under typical pothole encounter scenarios. This normalization prevents any single metric with inherently large numerical scale from dominating gradient updates. Building upon this baseline, scalar weights are further adjusted to reflect safety-first priorities in mixed traffic, where durability and comfort are prioritized over efficiency. The detailed parameter configuration is summarized in Table 1.

Table 1. Configuration of reward function weights.

Objective component	Symbol	Value
Passage Efficiency	W_{eff}	1.0
Traffic Safety	W_{rear}	10.0
Unified Ride Comfort	W_{comf}	10.0
Longitudinal sub-weight	w_{lon}	1.0
Structural Durability	W_{dura}	20.0

3.4. Curriculum Learning Strategy

In the high-speed pothole-handling task, the agent faces a highly non-convex reward landscape caused by the coupling among impact safety, ride comfort, traffic interaction, and lane-bounded maneuvering. Standard reinforcement learning often struggles in such settings because useful policies are difficult to discover through unconstrained exploration alone, especially when severe impact events induce sparse and high-variance penalties (Lin et al., 2025; Narvekar et al., 2020). To improve training stability, we employ a two-stage curriculum learning strategy that gradually increases task complexity, following the general principle of progressive task sequencing in curriculum-based reinforcement learning (Lin et al., 2025; Narvekar et al., 2020). The core idea is to first establish a reliable longitudinal protection policy, and then introduce more diverse pothole geometries that require coordinated lateral–longitudinal decision-making.

Stage I: Longitudinal anchoring. In the first stage, the agent is primarily exposed to large and difficult-to-bypass potholes. This stage simplifies the decision structure and encourages the policy to learn the dominant relationship between speed reduction and impact-load mitigation, thereby establishing a stable safety-oriented baseline.

Stage II: Lateral–longitudinal coordination. After the initial anchoring stage, the training distribution is expanded to include a mixture of large unavoidable potholes, wheel-path potholes, and randomly located moderate potholes. This stage requires the agent to jointly optimize target speed and lateral offset, enabling the emergence of coordinated avoidance, straddling, and compliant traversal behaviors.

Algorithm 1 Training and interaction procedure of the proposed hierarchical framework

Require: Replay buffer $\mathcal{B}$; total episodes $\mathbf{M}$.

Ensure: Trained tactical policy $\boldsymbol{\pi_\theta}$.

Initialize stochastic policy $\boldsymbol{\pi_\theta}$ and value networks $\boldsymbol{Q_{\phi_1}}, \mathbf{Q_{\phi_2}}$.
for episode $\mathbf{e} = \mathbf{1}$ **to** $\boldsymbol{M}$ do
  if $\mathbf{e \le 300}$ **then**
    Sample a training scenario from the Stage-I distribution.
  else
    Sample a training scenario from the Stage-II mixed distribution.
  end if
  Observe initial preview state $\mathbf{s_0}$ (vehicle kinematics, pothole geometry, and traffic context).
  Tactical decision (single-step high-level):
  Generate tactical action $\mathbf{u_e} = [\mathbf{v_{tgt}}, \mathbf{y_{tgt}}] \sim \boldsymbol{\pi_\theta}(\cdot \mid \mathbf{s_0})$.
  Execution and physical interaction (continuous low-level):
  Generate a reference path using a quintic Bézier curve based on $\mathbf{u_e}$.
  Execute closed-loop tracking via Stanley and PID controllers until the hazard is passed.
  Compute the episodic physical outcomes, including peak impact force $\mathbf{F_{peak}}$ and traffic-safety indicators.
  Policy evaluation and update:
  Calculate the episodic return $\mathbf{R_{total}}$ using Eq. (11).
  Store terminal transition $(\mathbf{s_0}, \mathbf{u_e}, \mathbf{R_{total}}, \mathbf{s_{term}}, \mathbf{d = 1})$ in replay buffer $\mathcal{B}$.
  Update $\boldsymbol{\pi_\theta}$ and $\mathbf{Q_{\phi_{1,2}}}$ using the SAC objective for terminal one-step returns.
end for

In the implemented training schedule, Stage I occupies the first 300 episodes. From episode 301 onward, the training distribution switches to a mixed regime consisting of 30% large unavoidable potholes, 40% wheel-path potholes, and 30% randomly located potholes. The complete training and interaction procedure, including the stage-dependent scenario sampling strategy, is summarized in Algorithm 1. This schedule-based curriculum avoids abrupt exploration over the full task space at the start of training and improves convergence robustness in the coupled multi-objective environment.

4 Experimental Setup

This section describes the experimental setup used to train and evaluate the proposed vehicle-dynamics-aware motion planning method for pothole handling. We first introduce a Gymnasium-based multi-physics environment that integrates quarter-car vertical dynamics, tire–pothole impact interaction, and reaction-time-delayed car-following behavior. We then present the empirical parameterization adopted for pothole geometry sampling. Baseline strategies are summarized in Section 4.3, and SAC implementation details are provided in Section 4.4.

4.1. Integrated Multi-Physics Simulation Environment

The environment follows the Gymnasium interface to provide a standardized agent–environment interaction loop for reinforcement learning (Towers et al., 2024). Each episode is initialized such that the ego vehicle approaches a pothole with an initial speed uniformly sampled from [60, 120] km/h, and a stochastic preview distance sampled from [10, 100] m to emulate variability in detection range.

The ego vehicle is initialized at $x = 0$, and the pothole center is positioned at $x = x_p$ according to the sampled preview distance. Each episode terminates when the ego vehicle reaches $x = x_p + 5$ m.

At each decision step, the agent observes a state vector that encodes (i) the current longitudinal speed, (ii) the preview distance to the pothole, and (iii) the pothole geometry and lateral alignment relative to the ego lane. The agent outputs a two-dimensional continuous tactical action $\boldsymbol{u} = [v_{\mathrm{tgt}}, y_{\mathrm{tgt}}]$, where v_{tgt} denotes the target traversal speed and y_{tgt} denotes the target vehicle-center lateral offset. The admissible lateral offset is determined by the lane width and vehicle footprint, thereby keeping the commanded vehicle position within the lane. The high-level policy is therefore responsible only for tactical target generation, while the low-level trajectory generator and controllers realize the maneuver subject to the prescribed execution bounds.

During lateral maneuver execution, the target vehicle-center offset is bounded within [-0.85, 0.85] m, and the front-wheel steering command generated by the Stanley controller is limited to $\pm 30°$. For each generated Bézier path, the maximum curvature $\kappa_{\max}$ is calculated, and the execution speed is constrained by

$$v_{\mathrm{exec}} \leq \sqrt{\frac{a_{\mathrm{lat,max}}}{\kappa_{\max}}}, \tag{16}$$

where $a_{\mathrm{lat,max}} = 4.0\ \mathrm{m/s^2}$. Accordingly, the nominal path-based lateral acceleration $v_{\mathrm{exec}}^2 \kappa_{\max}$ is maintained within the prescribed limit. These bounds constrain the commanded maneuver at the planning and control levels; they do not constitute comprehensive full-vehicle handling-stability validation.

The ego vehicle's vertical dynamics are modeled by a 2-DOF quarter-car system, capturing sprung-mass and unsprung-mass dynamics for comfort evaluation and impact-transient coupling. Key parameters used in the simulator are summarized in Table 2.

Table 2. Quarter-car model parameters.

Parameter	Symbol	Value
Sprung mass	m_s	350 kg
Unsprung mass	m_u	45 kg
Suspension stiffness	k_s	3.0×10^4 N/m
Tire stiffness	k_t	2.0×10^5 N/m
Tire damping	c_t	200 N·s/m
Compression damping	$c_{s,comp}$	1000 N·s/m
Rebound damping	$c_{s,reb}$	3000 N·s/m
Maximum suspension travel	z_{max}	0.10 m
Bump-stop stiffness	k_b	5.0×10^5 N/m

To explicitly capture edge-strike transients beyond acceleration-only surrogates, the simulator invokes the mechanism-informed tire–pothole impact model introduced in Section 3.1. For each episode, the impact solver returns the peak resultant impact load, which is used by the impact-load objective and enables constraint-aware tactical learning under pothole hazards. The main parameters used in the impact model are summarized in Table 3.

Table 3. Parameters used in the wheel–pothole impact model.

Parameter	Symbol	Value
Effective tire radius	R_t	0.31 m
Normal contact stiffness	K_c	2.0×10^6 N/m^n
Contact force exponent	n	1.0
Velocity correction coefficient	α_{hc}	0.25
Sliding friction coefficient	μ	0.80
Equivalent wheel mass	m_w	50 kg

Interactive traffic is modeled by a following vehicle governed by the Intelligent Driver Model (IDM) (Treiber et al., 2000). To reflect realistic rear-end pressure during pothole handling, the follower is initialized with short time headways consistent with naturalistic observations (Zhu et al., 2016). Moreover, a perception–reaction delay is incorporated to avoid idealized instantaneous responses; we follow empirical reaction-time ranges reported in car-following studies (Chang and Chon, 2005). Table 4 reports the IDM and reaction-time settings used for simulation.

Table 4. Follower IDM and reaction-time parameters.

Parameter	Symbol	Setting
Initial time headway	T_{init}	$\mathcal{U}(0.5,1.2)$s
Desired time headway (IDM)	T	$\mathcal{U}(0.5, T_{\mathrm{init}})$s
Desired speed	v_0	$\mathcal{U}(33,36)$m/s
Max acceleration	a	1.4 m/s^2
Comfortable deceleration	b	2.0 m/s^2
Hard braking limit	b_{max}	$\mathcal{U}(4,6)$m/s^2
Reaction time	τ_r	$\mathcal{U}(0.8,1.2)$s
Integration step (traffic)	Δt	0.02 s

4.2. Empirical Parameterization of Pothole Geometries

Scenario generation is based on empirical pothole measurements to ensure realistic defect severities in both training and evaluation. Each pothole is parameterized by its geometric dimensions, namely length L_p, width W_p, and depth D_p. These parameters are estimated from two independent data sources (Fig. 4), including the Road Surface Reconstruction Dataset (RSRD) (Zhao et al., 2024) and an in-house dataset comprising 130,325 pothole records (annotations) from 109,501 road images collected in Shanghai and Zhejiang Province during a July–August 2025 survey campaign. Each record corresponds to one pothole and includes its image reference, defect type, measured width, length, height/depth, and area. In this study, the two sources play an equal role: both are used to statistically characterize pothole size distributions and enlarge the coverage of observed defect geometries, rather than introducing appearance- or texture-level variation.

Aggregating the measurements yields a long-tailed severity profile: most potholes are of moderate size, while a non-negligible subset exhibits large width and depth that may induce critical edge-strike loads. Accordingly, we sample $L_p \in [0.2,1.0]$m, $W_p \in [0.3,1.0]$ m, and $D_p \in [0.03,0.15]$ m to construct realistic yet diverse pothole hazards, together with the preview-distance variability described in Section 4.1. This data-driven parameterization enables controlled scenario diversity for policy learning and evaluation. The empirical distributions of length, width, height/depth, and area in the in-house data, together with RSRD, were used to determine these simulation ranges and retain severe upper-tail cases. Evaluation potholes are generated by the stratified sampling protocol in Section 5.2.1 rather than drawn directly from the image records.

(a) RSRD samples. (b) In-house dataset samples.

Fig. 4. Representative pothole samples from the two data sources used for geometry parameterization.

Geometric avoidability is determined from the available lateral clearance. With a lane width W_{lane} = 3.5 m, a vehicle width W_{veh} = 1.8 m, and a total clearance allowance c = 0.10 m, the required passage width is $W_{\text{req}} = W_{\text{veh}} + c$ = 1.90 m. The admissible vehicle-center offset is

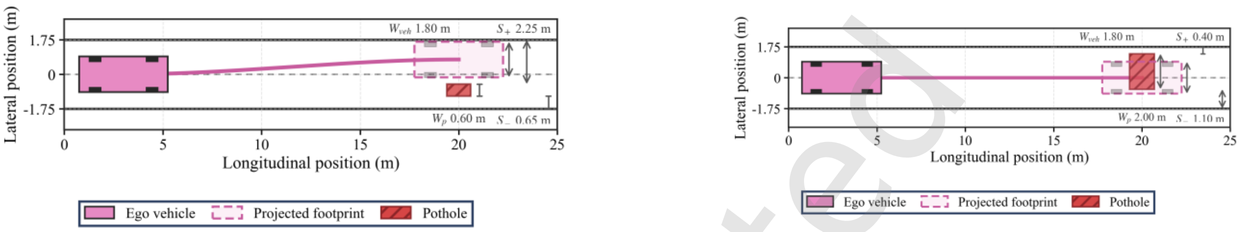
$$|y_e| \le \frac{W_{\text{lane}} - W_{\text{veh}}}{2} = 0.85\text{m} \tag{17}$$

For a pothole of width W_p centered at y_p, the free passages above and below the pothole are

$$S_+ = \frac{W_{\text{lane}}}{2} - y_p - \frac{W_p}{2}, S_- = \frac{W_{\text{lane}}}{2} + y_p - \frac{W_p}{2} \tag{18}$$

A pothole is classified as avoidable when either S_+ or S_- is no smaller than W_{req}; otherwise, it is classified as unavoidable.

Avoidable cases must additionally intersect at least one nominal tire-track band for $y_e = 0$, ensuring wheel–pothole contact under straight traversal; rejection sampling enforces these conditions. Representative cases are shown in Fig. 5.



(a) An avoidable pothole. (b) An unavoidable pothole.

Fig. 5. Geometric classification of potholes.

4.3. Baseline Strategies

To evaluate the effectiveness of the proposed method, we benchmark against four representative strategies spanning passive execution, longitudinal-only learning and optimization, and geometry-driven lateral avoidance.

1. **Constant Speed (CS):** A passive baseline that maintains the initial speed and lane-center trajectory throughout the scenario, providing a reference for unmitigated pothole traversal.
2. **RL-based Speed Control (RL-SC):** Drawing on prior work (Du et al., 2022), this baseline learns longitudinal speed regulation to reduce vertical excitation and energy consumption, but does not use lateral avoidance.
3. **Optimization-based Speed Planning (Opt-SP):** Following prior work (Zhang et al., 2023), this baseline optimizes a longitudinal speed profile under a lane-keeping assumption to mitigate road-induced excitation, but does not use lateral maneuvering space.
4. **Heuristic Avoidance System (HAS):** Following prior work (Chaaban et al., 2025), this baseline applies geometry-based clearance rules to generate lateral maneuvers. It bypasses the pothole when feasible and otherwise decelerates, without explicitly considering impact-load constraints.
5. **Curriculum ablation (Proposed w/o Curriculum):** This variant retains the observation and two-dimensional action spaces, SAC architecture, reward formulation, training budget, hyperparameters, and following-vehicle initialization of the complete method, but samples the target scenario distribution from the beginning instead of using the staged curriculum. The comparison between the two variants therefore isolates the effect of curriculum scheduling within the proposed framework.

4.4. Reinforcement Learning Implementation Details

The proposed method is trained using Soft Actor-Critic (SAC), which serves as the optimizer for the single-step tactical policy. Table 5 summarizes the main hyperparameters used in training. Unless otherwise stated, Section 5 reports the mean and sample standard deviation across five independent repetitions.

Table 5. Main hyperparameters used for SAC training.

Parameter	Symbol	Value
Actor learning rate	η_π	3×10^{-4}
Critic learning rate	η_Q	3×10^{-4}
Target network update rate	τ	0.005
Batch size	B	64
Replay buffer size	$\lvert\mathcal{B}\rvert$	1×10^5
Updates per episode	--	8
Warm-up episodes	--	200
Entropy temperature	α	auto$\mathcal{H}_{\text{tgt}} = -2$
Hidden width	--	256

5 Experimental Results

This section reports the experimental results and analyzes the performance of the proposed method in pothole-handling scenarios. We first examine the learning dynamics to assess the proposed two-stage training strategy (Section 5.1). We then compare the proposed method with representative baselines under factorized scenario settings (Section 5.2), followed by representative case studies that illustrate the learned decision logic (Section 5.3). The proposed method and its curriculum ablation use the training hyperparameters summarized in Table 5, and the reported results are aggregated over multiple random seeds.

5.1. Training Performance and Convergence Analysis

We examine the training performance to assess the role of the proposed two-stage training strategy in stabilizing policy optimization under coupled safety, comfort, and interaction objectives. Specifically, we compare the proposed two-stage training scheme with the proposed method without curriculum learning (Proposed w/o Curriculum), trained under the same hyperparameter setting but without curriculum staging. Fig. 6 plots the episodic return during training, where solid lines denote the mean over five random seeds and the shaded regions indicate the corresponding variability. The vanilla SAC baseline shows persistently higher return variance and a lower final return plateau. This pattern is consistent with sparse, high-magnitude penalties induced by rare but severe events, such as excessive impact loads or rear-end conflicts, which amplify gradient noise and destabilize value estimation during training.

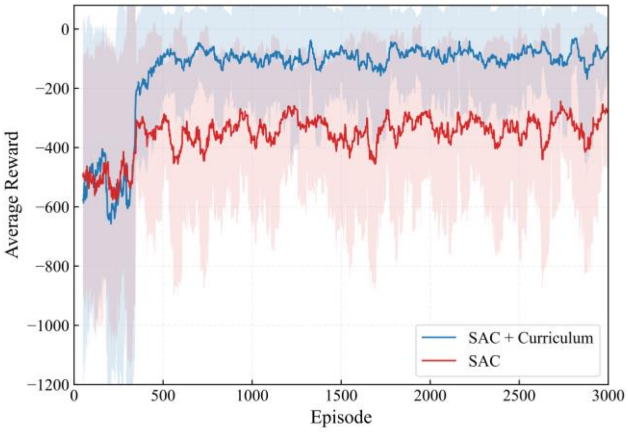


Fig. 6. Training curves of episodic return.

In contrast, the proposed staged training yields a faster rise in return and markedly reduced variability across seeds. In Stage 1, the agent is exposed to simplified conditions that emphasize the dominant relationship between longitudinal speed regulation and impact-load mitigation, allowing it to learn a reliable baseline for satisfying the impact-load constraint. In Stage 2, more challenging lateral maneuvering and interactive constraints are introduced, which initially causes a transient fluctuation in return as the policy adapts to the expanded feasible set. After adaptation, the learning curve stabilizes with substantially lower variance, indicating that the staged strategy improves optimization stability and leads to more consistent convergence in the coupled multi-objective environment.

5.2. Comparative Performance Analysis

To evaluate the robustness of the proposed framework, we use a stratified sampling protocol to generate evaluation episodes. This section outlines the experimental design, defines the performance metrics, and presents a quantitative comparison against the baseline strategies.

5.2.1 Experimental Design and Metrics

Experimental Protocol: Evaluation scenarios are constructed using a stratified factorial protocol over three key dimensions:

(i) Preview distance (Short: 10-30 m; Medium: 30-60 m; Long: 60-100 m);

(ii) Pothole depth (Shallow: 3-8 cm; Deep: 8-15 cm); and

(iii) Avoidability (Avoidable vs. Unavoidable, determined by whether a feasible geometric solution exists within the lane boundaries).

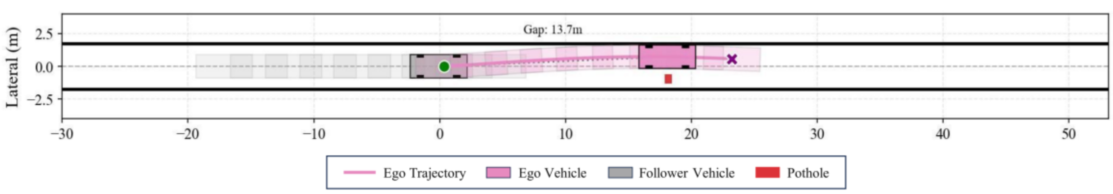
This yields 3 × 2 × 2 = 12 core scenario groups. Within each repetition, 20 scenarios are generated for each group with randomized traffic behavior (IDM parameters) and ego vehicle initial speed (v_{init} ∈ [60, 120] km/h), yielding 240 evaluation episodes per method. The complete evaluation is repeated five times using independent training and evaluation seeds. All methods share the same scenario manifest within each repetition; Tables 6-8 report the mean ± sample standard deviation across the five repetition-level summaries.

Performance Metrics. To provide a multidimensional assessment of the tactical decision-making capability, we report six episode-level metrics:

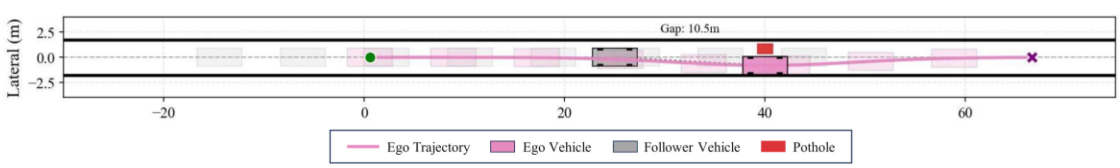
- **Avoidance rate (%):** The percentage of episodes in which the tire footprint does not intersect the pothole polygon, resulting in zero impact force.
- **Impact-load compliance rate (%):** The percentage of episodes in which the peak resultant wheel-impact force satisfies $F_{peak} \leq F_{lim}$. This metric specifically evaluates compliance with the prescribed impact-force limit within the adopted tire–pothole model.
- **Collision rate (%):** The percentage of episodes involving a rear-end collision with the following vehicle (gap≤ **0**).
- **Travel time (s):** The elapsed time from episode initialization, when the ego vehicle is located at $x = 0$, until it reaches a point 5 m beyond the pothole center, i.e., $x = x_p + 5$ m.
- **Vertical / Longitudinal Acceleration RMS (m/s²):** The root mean square (RMS) of the sprung mass vertical acceleration ($\boldsymbol{a_z}$) and longitudinal acceleration ($\mathbf{a_x}$), characterizing ride comfort and braking intensity, respectively.

5.2.2 Emergent Pothole-Handling Strategies

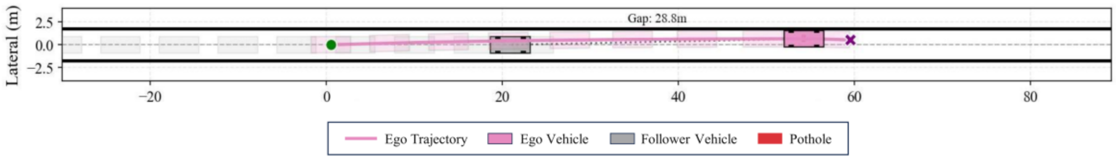
Before analyzing aggregated statistics, we examine the physical validity of the learned policy. As illustrated in Fig. 7, the proposed method demonstrates three distinct in-lane handling strategies: left nudge, right nudge, and straddling (center pass). These behaviors are not manually specified, but instead emerge from the reward design, and provide the kinematic basis for the quantitative performance improvements reported below.



(a)



(b)



(c)

Fig. 7. Visualization of emergent in-lane pothole avoidance behaviors learned by the proposed method. (a) Left nudge. (b) Right nudge. (c) Straddling (center pass).

5.2.3 Quantitative Results

We compare Constant Speed (CS), RL-based Speed Control (RL-SC), Optimization-based Speed Planning (Opt-SP), Heuristic Avoidance System (HAS), Proposed w/o Curriculum, and the complete Proposed method. RL-SC and Opt-SP use longitudinal actions only, whereas HAS and both proposed variants can use intra-lane lateral motion. Comparisons with the longitudinal-only planners therefore reflect the benefit of the joint longitudinal-lateral planning formulation together with method-specific differences; the paired comparison between Proposed and Proposed w/o Curriculum is used specifically to assess curriculum scheduling. The results are reported for all scenarios (Table 6), avoidable scenarios (Table 7), and unavoidable scenarios (Table 8).

Table 6. Performance statistics across all 240 evaluation episodes.

Method	Avoidance Rate (%)	Travel Time (s)	Impact-load compliance rate (%)	Collision Rate (%)	Vertical RMS (m/s^2)	Longitudinal RMS (m/s^2)
CS	0.0 ± 0.0	**2.45 ± 0.04**	22.1 ± 2.2	**0.0 ± 0.0**	1.72 ± 0.03	**0.00 ± 0.00**
RL-SC	0.0 ± 0.0	4.63 ± 0.29	55.4 ± 3.6	12.5 ± 2.6	1.47 ± 0.02	1.79 ± 0.04
Opt-SP	0.0 ± 0.0	4.74 ± 0.28	55.4 ± 3.6	12.5 ± 2.4	1.46 ± 0.02	1.79 ± 0.04
HAS	15.4 ± 1.7	4.74 ± 0.28	62.1 ± 2.3	12.5 ± 2.4	1.26 ± 0.02	1.79 ± 0.04
Proposed w/o Curriculum	41.1 ± 1.9	2.98 ± 0.19	69.4 ± 3.6	1.4 ± 0.6	1.19 ± 0.07	1.52 ± 0.11
Proposed	**46.2 ± 1.7**	2.71 ± 0.17	**71.2 ± 1.3**	0.8 ± 0.5	**1.08 ± 0.06**	1.41 ± 0.08

Overall Performance (Table 6): The proposed method achieves the highest avoidance rate (46.2%) and impact-load compliance rate (71.2%), while keeping the collision rate at 0.8%. Rather than optimizing a single outcome, this combination indicates a favorable balance among pothole avoidance, wheel-impact mitigation, and rear-end interaction risk. Relative to Proposed w/o Curriculum, the complete method improves the mean value of all six reported metrics, supporting the contribution of staged training to the overall balance. These findings are interpreted as simulation-based evidence within the adopted vehicle model and prescribed motion limits.

Table 7. Performance statistics in 120 avoidable pothole scenarios.

Method	Avoidance Rate (%)	Travel Time (s)	Impact-load compliance rate (%)	Collision Rate (%)	Vertical RMS (m/s^2)	Longitudinal RMS (m/s^2)
CS	0.0 ± 0.0	2.46 ± 0.05	25.8 ± 4.0	**0.0 ± 0.0**	1.09 ± 0.05	**0.00 ± 0.00**
RL-SC	0.0 ± 0.0	4.75 ± 0.24	56.7 ± 5.9	11.7 ± 3.2	1.13 ± 0.04	1.81 ± 0.04
Opt-SP	0.0 ± 0.0	4.75 ± 0.24	56.7 ± 5.9	11.7 ± 3.2	1.13 ± 0.04	1.81 ± 0.04
HAS	30.8 ± 3.4	4.75 ± 0.24	70.0 ± 3.2	11.7 ± 3.2	0.72 ± 0.04	1.79 ± 0.04
Proposed w/o Curriculum	87.2 ± 3.9	2.84 ± 0.21	92.5 ± 2.6	1.5 ± 1.1	0.18 ± 0.18	1.71 ± 0.15
Proposed	**92.5 ± 4.2**	**2.36 ± 0.21**	**95.8 ± 1.4**	**0.0 ± 0.0**	**0.07 ± 0.04**	1.21 ± 0.14

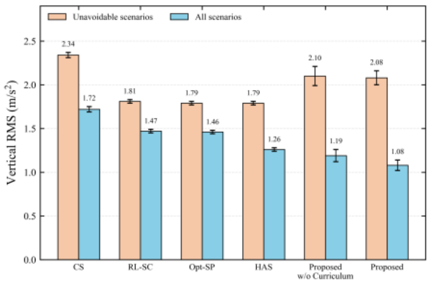
Performance in Avoidable Scenarios (Table 7): When lateral bypass is feasible, the proposed method achieves a 92.5% avoidance rate, compared with 30.8% for HAS, and reduces vertical RMS to 0.07 m/s². Together with the approximately 50% reduction in travel time over the defined pothole-passage interval relative to longitudinal-only braking strategies, these results show that exploiting intra-lane lateral space is the main source of the impact-mitigation and efficiency advantage in avoidable scenarios. The curriculum ablation further indicates that staged training improves the coordination of lateral maneuvering and speed control.

Table 8. Performance statistics in 120 unavoidable pothole scenarios.

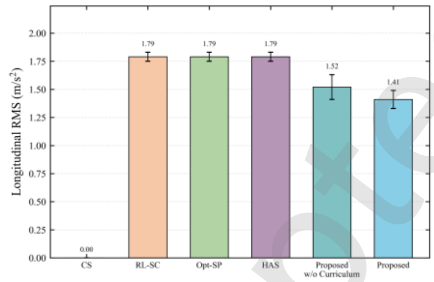
Method	Avoidance Rate (%)	Travel Time (s)	Impact-load compliance rate (%)	Collision Rate (%)	Vertical RMS (m/s^2)	Longitudinal RMS (m/s^2)

CS	0.0 ± 0.0	**2.44 ± 0.04**	18.3 ± 3.2	**0.0 ± 0.0**	2.34 ± 0.03	**0.00 ± 0.00**
RL-SC	0.0 ± 0.0	4.51 ± 0.37	**54.2 ± 2.0**	13.3 ± 1.8	1.81 ± 0.02	1.77 ± 0.05
Opt-SP	0.0 ± 0.0	4.73 ± 0.35	**54.2 ± 2.0**	13.3 ± 1.8	**1.79 ± 0.02**	1.78 ± 0.05
HAS	0.0 ± 0.0	4.73 ± 0.35	**54.2 ± 2.0**	13.3 ± 1.8	**1.79 ± 0.02**	1.78 ± 0.05
Proposed w/o Curriculum	0.0 ± 0.0	3.25 ± 0.38	41.8 ± 6.7	2.8 ± 1.3	2.10 ± 0.11	2.10 ± 0.11
Proposed	0.0 ± 0.0	3.06 ± 0.16	46.7 ± 5.4	1.7 ± 0.7	2.08 ± 0.08	1.61 ± 0.06

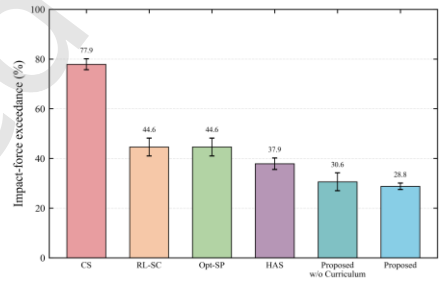
Trade-offs in Unavoidable Scenarios (Table 8): When lateral avoidance is infeasible, the braking-heavy baselines achieve higher impact-load compliance (54.2%) but incur a 13.3% collision rate. The proposed method reduces the collision rate to 1.7% while retaining 46.7% impact-load compliance, indicating a more balanced compromise between wheel-impact mitigation and rear-end interaction risk. Its improvement over Proposed w/o Curriculum in both measures further suggests that staged training also benefits longitudinal speed regulation when avoidance is unavailable.



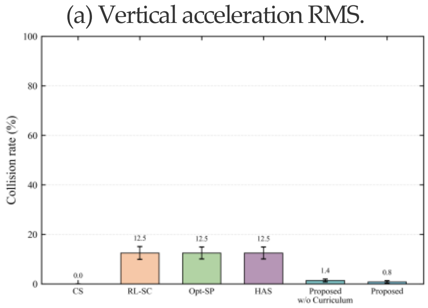
(a) Vertical acceleration RMS.



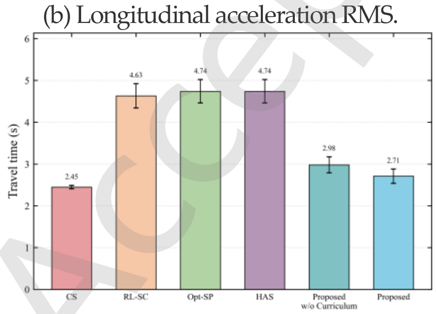
(b) Longitudinal acceleration RMS.



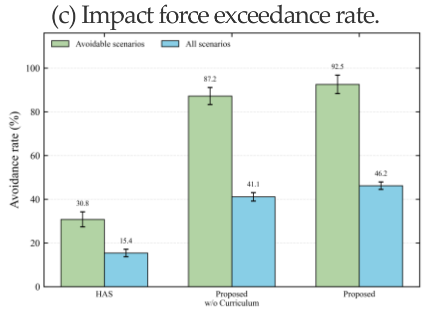
(c) Impact force exceedance rate.



(d) Collision rate.



(e) Travel time.



(f) Avoidance rate.

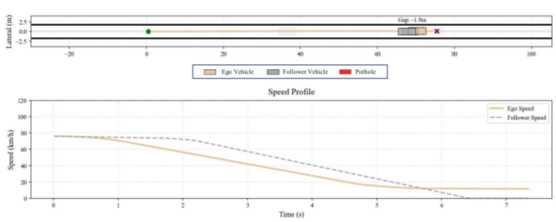
Fig. 8. Distributional comparison of key metrics across evaluation episodes. Error bars denote the standard deviation across five independent repetitions.

5.3. Case Studies of Representative Scenarios

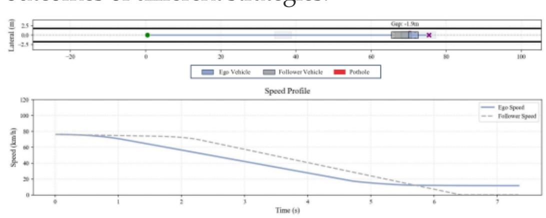
To complement the aggregate statistics in Section 5.2, we present two representative episodes to interpret the behavioral mechanisms of the proposed method. The selected cases are drawn from the avoidable and unavoidable groups, respectively, and are chosen to highlight two recurrent decision patterns observed in evaluation: (i) lateral avoidance under elevated rear-end collision risk and (ii) speed regulation to satisfy the impact-load limit when avoidance is infeasible. In both cases, we compare the proposed method against the baselines defined in Section 4.3 using the same scenario realization and traffic seed.

5.3.1 Case I: Lateral avoidance under elevated rear-end collision risk

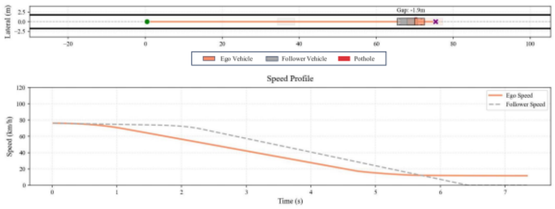
This case belongs to the avoidable group, where a deep pothole can be bypassed within lane boundaries. The ego vehicle approaches the hazard at $\boldsymbol{v}_{init} = \mathbf{76.1}$ km/h with a closely following vehicle, creating a constrained car-following condition with limited braking margin. Fig. 9 compares the time histories and outcomes of different strategies.



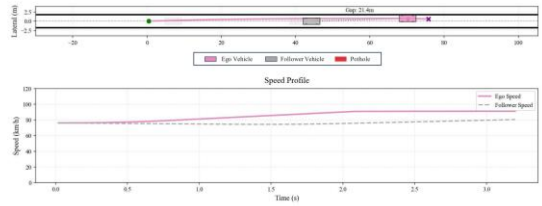
(a) HAS.



(b) Opt-SP.



(c) RL-SC.



(d) Proposed method.

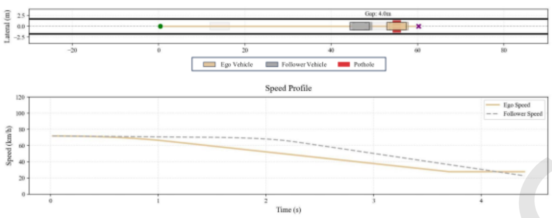
Fig. 9. Case I (avoidable, $v_{\mathrm{init}} = 76.1\,\mathrm{km/h}$): Comparison of baseline strategies and the proposed method under tight car-following constraints. (a) HAS. (b) Opt-SP. (c) RL-SC. (d) Proposed method.

Across the baselines, the dominant response is aggressive deceleration to reduce the pothole-induced vertical excitation. Although this behavior mitigates impact severity, it does not explicitly account for rear-end collision risk under tight car-following conditions. Consequently, the rapid speed reduction leads to a violation of the rear-end safety margin and results in a rear-end collision in this episode.

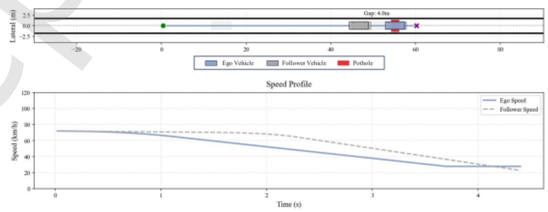
In contrast, the proposed method adopts a joint decision that couples lateral feasibility with interaction safety. Rather than relying on strong braking, it executes a controlled lateral bypass maneuver while maintaining a flow-consistent speed profile. This behavior preserves the rear-end safety margin and simultaneously avoids the pothole, yielding zero structural impact in this episode. This case illustrates that, when avoidance is feasible, joint lateral–longitudinal planning can preserve rear-end safety while eliminating the impact load.

5.3.2 Case II: Impact-load-constrained speed regulation in unavoidable hazards

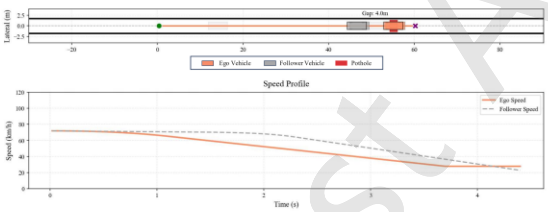
This case belongs to the unavoidable group, where the hazard spans the lane and geometric avoidance is infeasible. The ego vehicle approaches an unavoidable pothole at $v_{init} = 68.1$ km/h. Under this setting, the decision primarily concerns how to reduce the impact load without inducing excessive rear-end risk. Fig. 10 compares the resulting speed profiles and safety outcomes.



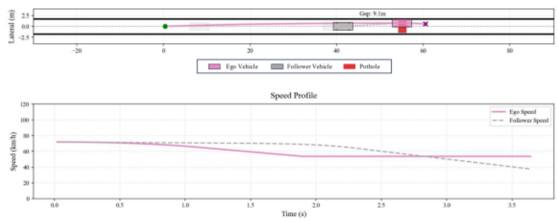
(a) HAS.



(b) Opt-SP.



(c) RL-SC.



(d) Proposed method.

Fig. 10. Case II (*unavoidable*, $v_{\mathrm{init}} = 68.1\mathrm{km/h}$): Comparison under an unavoidable hazard, highlighting impact-load-constrained speed regulation. (a) HAS. (b) Opt-SP. (c) RL-SC. (d) Proposed method.

Baseline methods again reduce speed to a near-minimum level to minimize impact. Although such over-braking can reduce the instantaneous impact load, it substantially compresses the rear gap and increases rear-end risk, which is undesirable in mixed traffic.

The proposed method employs an impact-load-constrained regulation strategy, guided by the calibrated impact-load limit ($F_{\mathrm{lim}} = 15.15\mathrm{kN}$). It decelerates to a compliant yet non-extreme speed of 29.1 km/h, resulting in an impact force of 12.53 kN (below F_{lim}). By avoiding excessive braking, the proposed method preserves a larger rear-end safety margin while ensuring compliance with impact-load constraints. This case further illustrates the trade-off discussed in Section 5.2: in unavoidable hazards, strict impact minimization may conflict with interaction safety. The proposed approach addresses this trade-off through risk-prioritized, threshold-aware speed regulation.

6 Discussion

The experimental results presented in Section 5 show that the proposed framework consistently outperforms the baseline strategies in single-hazard scenarios. We additionally provide preliminary qualitative evidence from one sequential-hazard case.

6.1. Illustrative Evaluation in a Sequential-Hazard Scenario

Although the training curriculum primarily considers single-pothole scenarios, road defects may occur consecutively in practice. We therefore use a representative sequential-hazard case to qualitatively examine how the trained policy responds to successive potholes.

As shown in Fig. 11, the agent performs a lateral nudge to bypass the first pothole, stabilizes the vehicle trajectory, and subsequently prepares for the next maneuver. The velocity profile also shows that the agent recovers speed during the safe interval between the two hazards rather than adopting a passive stop-and-go strategy.

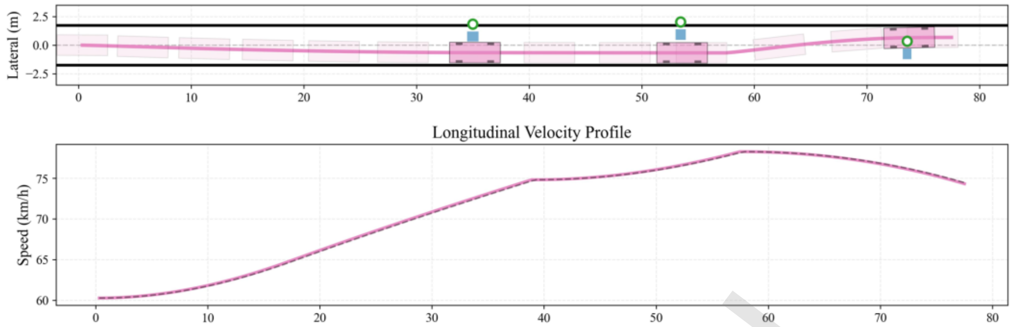


Fig. 11. Illustrative response in a sequential-hazard scenario. The ego vehicle coordinates successive maneuvers and recovers speed during the safe interval between potholes.

This case illustrates the policy's ability to coordinate successive lateral and longitudinal adjustments within one episode. A systematic quantitative evaluation across diverse multi-hazard scenarios remains a subject for future work.

6.2. Limitations and Implementation Considerations

Several limitations should be considered when interpreting the present results. First, the simulations assume accurate knowledge of pothole location and geometry. Localization errors, geometry-estimation uncertainty, and V2X communication delays are not explicitly modeled; therefore, robustness to imperfect preview information remains to be evaluated. Second, the quarter-car model captures vertical suspension responses and transient impact loads but does not represent their coupling with lateral and roll dynamics. Although the target lateral offset, steering command, and curvature-based nominal lateral acceleration are bounded, these constraints only restrict the commanded maneuver within the prescribed operating range. They do not reproduce lateral load transfer, tire-force saturation, or vehicle-roll responses during high-speed avoidance. Further evaluation using a coupled full-vehicle or high-fidelity multibody dynamics model is therefore required.

From an implementation perspective, recent autonomous-driving research demonstrates mature camera, LiDAR, multimodal, and vision-language perception for road-scene understanding, while pothole-specific sensor fusion further supports localized surface-hazard acquisition (Y. Hu et al., 2025; Pan et al., 2024; Cai et al., 2025). Given a validated pothole state, policy decision-making satisfies the real-time requirement, and the trajectory-generation/controller stack provides planning-constraint checks before execution; end-to-end validation under perception uncertainty remains future work.

7 Conclusion

This study presented a vehicle-dynamics-aware motion planning framework for pothole-hazard autonomous driving. In contrast to existing methods that mainly emphasize geometric avoidance or isolated suspension control, the proposed approach combines a mechanism-informed tire–pothole impact model with curriculum-based reinforcement learning, enabling the agent to balance impact-load compliance, ride comfort, and traffic safety under tight physical constraints.

The primary findings of this study are summarized as follows.

（1）Improved impact-load compliance. In unavoidable hazard scenarios, the proposed framework achieves an impact-load compliance rate of 46.7%, compared with 18.3% for the passive constant-speed baseline. This result shows that the learned policy can effectively reduce excessive impact loads even when full geometric avoidance is not possible.

（2）Improved interaction safety. In unavoidable scenarios, the proposed method reduces the rear-end collision rate to 1.7%, compared with 13.3% for the braking-heavy baselines. This result indicates that the learned policy avoids overly aggressive braking and instead balances impact mitigation with interaction safety in surrounding traffic.

（3）Adaptive pothole-handling behavior. The learned policy exhibits lateral nudging, straddling, and threshold-aware speed regulation. One illustrative sequential-hazard case provides preliminary qualitative evidence of coordinated successive maneuvers; systematic multi-hazard evaluation remains future work.

In conclusion, this study bridges the gap between conventional geometric path planning and vehicle–road interaction mechanics, presenting a resilient solution for autonomous mobility in unpredictable infrastructure settings. This work shifts the safety paradigm from static obstacle avoidance to dynamic impact management. Future research will focus on integrating perceptual uncertainty, such as noisy depth estimation, and validating the proposed algorithms on full-scale vehicle platforms to address real-world sensing and actuation latency.

Author contributions

Xiang Wang: Writing – original draft, Validation, Software, Methodology, Formal analysis, Data curation, Conceptualization. **Scott Piersall:** Writing – review & editing, Supervision, Methodology, Funding acquisition. **Zihang Zou:** Software, Data curation. **Liqiang Wang:** Formal Analysis, Methodology. **Rongjie Yu:** Writing – review & editing, Formal Analysis, Methodology.

Replication and data sharing

The source data within this research can be made accessible upon request via email to the corresponding author.

Acknowledgements

This study was jointly sponsored by the National Key R&D Program of China(2023YFC3081700) and Shanghai Science & Technology Innovation Action Plan Project under Grant 24ZR1452800.

Declaration of competing interest

The authors have no competing interests to declare that are relevant to the content of this article.

References

AAA East Central, 2024. AAA: Potholes can be a black hole for your wallet.

AAA Western and Central New York, 2023. AAA: Massive bump, 57% increase in drivers who sustain vehicle damage from potholes.

Ahmad, N., Arvin, R., Khattak, A. J., 2023. Exploring pathways from driving errors and violations to crashes: The role of instability in driving. *Accident Analysis & Prevention*, **179**, 106876.

Aradi, S., 2022. Survey of deep reinforcement learning for motion planning of autonomous vehicles. *IEEE Transactions on Intelligent Transportation Systems*, **23**, 740–759.

Audi UK, 2017. As seen by the new Audi A8: Predictive active suspension.

BYD Global, 2023. BYD launches DiSus intelligent body control system.

Cai, Y., Deng, M., Xu, X., Wang, W., Xu, X., 2025. Road pothole recognition and size measurement based on the fusion of camera and LiDAR. IEEE Access, 13, 46210–46227.

Chaaban, O., Daher, L., Ghaddar, Y., Zantout, N., Asmar, D., Daher, N., 2025. An End-to-End Pothole Detection and Avoidance System for Autonomous Ground Vehicles. In: *2025 International Conference on Control, Automation, and Instrumentation (IC2AI)*, 1–6.

Chang, K., Chon, K., 2005. A car-following model applied reaction times distribution and perceptual threshold. *Journal of the Eastern Asia Society for Transportation Studies*, **6**, 1888–1903.

Daimler Communications, 2017. Press information: Road surface scan and MAGIC BODY CONTROL (S-Class).

DiLorenzo, T., Yu, X., 2023. Use of ice detection sensors for improving winter road safety. *Accident Analysis & Prevention*, **191**, 107197.

Du, Y., Chen, J., Zhao, C., Liu, C., Liao, F., Chan, C.-Y., 2022. Comfortable and energy-efficient speed control of autonomous vehicles on rough pavements using deep reinforcement learning. *Transportation Research Part C: Emerging Technologies*, **134**, 103489.

Fang, Z., He, J., Qin, P., Chen, H., Zhang, C., Sun, S., 2025. Enhancing autonomous driving safety in real lane-changing scenarios under friction variability: A friction-adaptive shield reinforcement learning framework. *Accident Analysis & Prevention*, **223**, 108265.

Haarnoja, T., Zhou, A., Abbeel, P., Levine, S., 2018. Soft Actor-Critic: Off-Policy Maximum Entropy Deep Reinforcement Learning with a Stochastic Actor. In: *Proceedings of the 35th International Conference on Machine Learning*, 1861–1870.

Hu, J., Wang, Y., Cheng, S., Xu, J., Wang, N., Fu, B., et al., 2025. A survey of decision-making and planning methods for self-driving vehicles. *Frontiers in Neurorobotics*, **19**, 1451923.

Hu, Y., Hu, L., Kong, Q., Fan, B., 2025. A survey on end-to-end perception and prediction for autonomous driving. Machine Intelligence Research, 22, 999–1030.

Huang, T., Wang, J., Pan, H., 2023. Adaptive bioinspired preview suspension control with constrained velocity planning for autonomous vehicles. *IEEE Transactions on Intelligent Vehicles*, **8**, 3925–3937.

Hussain, Z., Mohammed, S. S., Dias, C., Hussain, Q., Alhajyaseen, W. K. M., 2025. Empirical analysis of car-following behavior: Impacts of driver demographics, leading vehicle types, and speed limits on driver behavior and safety. *Transportation Research Part F: Traffic Psychology and Behaviour*, **108**, 188–205.

Inamdar, R., Sundarr, S. K., Khandelwal, D., Sahu, V. D., Katal, N., 2024. A comprehensive review on safe reinforcement learning for autonomous vehicle control in dynamic environments. *e-Prime - Advances in Electrical Engineering, Electronics and Energy*, **10**, 100810.

Jaguar Land Rover, 2015. Jaguar Land Rover develops technology to detect potholes.

Khattak, Z. H., Fontaine, M. D., Li, W., Khattak, A. J., Karnowski, T., 2021. Investigating the relation between instantaneous driving decisions and safety critical events in naturalistic driving environment. *Accident Analysis & Prevention*, **156**, 106086.

Lin, S., Mi, Q., Gao, T., 2025. A Survey of Curriculum Learning in Deep Reinforcement Learning. In: *2025 IEEE 15th Annual Computing and Communication Workshop and Conference (CCWC)*, 1141–1147.

Liu, Z., Si, Y., Sun, W., 2024. Ride comfort oriented integrated design of preview active suspension control and longitudinal velocity planning. *Mechanical Systems and Signal Processing*, **208**, 110992.

Narvekar, S., Peng, B., Leonetti, M., Sinapov, J., Taylor, M. E., Stone, P., 2020. Curriculum learning for reinforcement learning domains: A framework and survey. *Journal of Machine Learning Research*, **21**, 1–50.

National Highway Traffic Safety Administration, 2002. The 100-Car Naturalistic Driving Study. U.S. Department of Transportation.

National Highway Traffic Safety Administration, 2006. The impact of driver inattention on near-crash/crash risk. U.S. Department of Transportation.

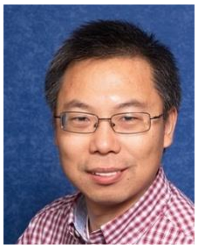
National Highway Traffic Safety Administration, 2009. Comparing real-world behaviors of drivers with high versus low rates of crashes and near-crashes. U.S. Department of Transportation.
Pan, C., Yaman, B., Nesti, T., Mallik, A., Allievi, A. G., Velipasalar, S., Ren, L., 2024. VLP: Vision Language Planning for Autonomous Driving. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 14760–14769.
Radu, A. I., Tolea, A. B., Trusca, D. D., Ispas, N., Nastasoiu, M., 2023. Development of a multibody model used to study the impact between a vehicle wheel with a pothole. *International Journal of Automotive Technology*, **24**, 1441–1446.
Sun, J., Liu, Z., Wu, J., 2018. An Optimal Driving Strategy for the Ride Comfort Performance of Intelligent Vehicles on Uneven Roads. IFAC-PapersOnLine, 51, 984–989.
Theunissen, J., Tota, A., Gruber, P., Dhaens, M., Sorniotti, A., 2021. Preview-based techniques for vehicle suspension control: A state-of-the-art review. *Annual Reviews in Control*, **51**, 206–235.
Towers, M., Kwiatkowski, A., Terry, J., Balis, J. U., De Cola, G., Deleu, T., et al., 2024. Gymnasium: A Standard Interface for Reinforcement Learning Environments. arXiv preprint arXiv:2407.17032.
Treiber, M., Hennecke, A., Helbing, D., 2000. Congested traffic states in empirical observations and microscopic simulations. *Physical Review E*, **62**, 1805–1824.
Wang, L., Yang, S., Yuan, K., Huang, Y., Chen, H., 2023. A combined reinforcement learning and model predictive control for car-following maneuver of autonomous vehicles. *Chinese Journal of Mechanical Engineering*, **36**, 80.
Wu, J., Huang, C., Huang, H., Lv, C., Wang, Y., Wang, F.-Y., 2024. Recent advances in reinforcement learning-based autonomous driving behavior planning: A survey. *Transportation Research Part C: Emerging Technologies*, **164**, 104654.
Wu, J., Zhou, H., Liu, Z., Gu, M., 2020. Ride comfort optimization via speed planning and preview semi-active suspension control for autonomous vehicles on uneven roads. *IEEE Transactions on Vehicular Technology*, **69**, 8343–8355.
Zhang, J., Xiao, B., Tang, B., Jiang, Y., Yu, T., Li, D., 2023. Safe and comfortable motion planning for bumpy road driving. In: *2023 IEEE 26th International Conference on Intelligent Transportation Systems (ITSC)*, 5706–5711.
Zhang, Y., Zhou, T., Yang, X., Wang, L., Zang, M., 2020. Transient dynamic behavior of tire traversing obstacles in full vehicle scenario. *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, **234**, 2001–2013.
Zhao, S., Zhang, J., He, C., Ji, Y., Huang, H., Hou, X., 2024. Autonomous vehicle extreme control for emergency collision avoidance via reachability-guided reinforcement learning. *Advanced Engineering Informatics*, **62**, 102801.
Zhao, T., Xie, Y., Ding, M., Yang, L., Tomizuka, M., Wei, Y., 2024. A road surface reconstruction dataset for autonomous driving. *Scientific Data*, **11**, 459.
Zhu, M., Wang, X., Wang, X., 2016. Car-Following Headways in Different Driving Situations: A Naturalistic Driving Study. In: *CICTP 2016: Green and Multimodal Transportation and Logistics*, 1419–1428.

Author biography

Xiang Wang received the B.Sc. degree in transportation engineering from the Harbin Institute of Technology, Weihai, China, in 2024. He is currently working toward the M.Sc. degree with the College of Transportation Engineering, Tongji University, Shanghai, China. His research interests include motion planning, vehicle dynamics control, autonomous vehicle safety and intelligent vehicles.

Scott Piersall received the M.S. degree in computer science from the University of North Florida, Jacksonville, FL, USA, in 2000. He is currently working toward the Ph.D. degree in computer science at the University of Central Florida, Orlando, FL, USA. He is a Visiting Instructor with the School of Computing, University of North Florida, where he teaches undergraduate courses in systems programming, enterprise COBOL, and graduate courses in parallel computing. His research interests include machine learning, large language models, program analysis, parallel computing, and distributed systems.

Zihang Zou received the B.Sc. degree in electrical and electronics engineering from Sun Yat-sen University, Guangzhou, China, in 2010, the M.Sc. degree in electrical and electronics engineering from the University of Central Florida (UCF), Orlando, FL, USA, in 2011, and the Ph.D. degree from the College of Engineering and Computer Science, UCF, in 2025. He is currently a Lead Research Scientist at Optixway AI, USA. His research focuses on safe, reliable, and legally compliant artificial intelligence, with interests spanning trustworthy generative AI, intellectual property protection for generative models, lawful autonomous driving, transportation safety, and AI systems designed for real-world deployment. His work covers neural watermarking and ownership verification for generative models, crash-risk prediction, and traffic-law-aware AI, with a broader emphasis on aligning AI systems with real-world rules, operational constraints, and societal needs. More broadly, he is interested in advancing robust, compliant, and socially responsible AI for safety-critical and infrastructure-related applications.



Liqiang Wang received the Ph.D. degree in computer science from Stony Brook University, Stony Brook, NY, USA, in 2006. He was a Faculty Member with the Department of Computer Science, University of Wyoming, Laramie, WY, USA, from 2006 to 2015, and a Visiting Research Scientist with the IBM T.J. Watson Research Center from 2012 to 2013. He is currently a Professor with the Department of Computer Science, University of Central Florida, Orlando, FL, USA, where he is also the Director of the Big Data Lab. His research focuses on big data computing and analytics, including improving model accuracy and security, and optimizing performance of parallel computing systems. Dr. Wang was a recipient of the NSF CAREER Award in 2011, the Castagne Faculty Fellowship (2013--2015), and the UCF Mid-Career Refresh Award in 2020.

Rongjie Yu received the B.Sc. degree in civil engineering from Tongji University, Shanghai, China, in 2010, and the M.Sc. degree in transportation engineering and the Ph.D. degree from the University of Central Florida, Orlando, FL, USA, in 2012 and 2013, respectively. He is currently a Professor with the College of Transportation Engineering, Tongji University, Shanghai, China. His research interests include traffic safety analysis, driving behavior modeling, proactive safety management, and autonomous vehicle safety evaluation.