Multimodal multiscale decision-making and control for urban autonomous vehicles with memory-conditioned dynamic potential field reconstruction

Multimodal Transportation 5 (2026) 100324

Contents lists available at ScienceDirect

Multimodal Transportation

journal homepage: www.elsevier.com/locate/multra

Full Length Article

Multimodal multiscale decision-making and control for urban autonomous vehicles with memory-conditioned dynamic potential field reconstruction

Yanbin Liu [a], Cong Zhang [a,*], Shaohua Cui [b], Guangyu Tian [a], Yugong Luo [a], Lei Zhang [a]

[a] School of Vehicle and Mobility, Tsinghua University, Beijing, 100084, China
[b] Department of Architecture and Civil Engineering, Chalmers University of Technology, Gothenburg, 41296, Sweden

ARTICLE INFO

Keywords:
Autonomous driving
Decision-making and control
Multimodal fusion
Memory-augmented reasoning
Dynamic potential field

ABSTRACT

Urban autonomous driving remains challenging because autonomous vehicles must reason about dense multi-agent interactions, traffic-signal constraints, occlusions, and long-tail events while satisfying real-time onboard computational requirements. To address these challenges, this paper proposes a Multimodal Multiscale Decision-Making and Control framework, termed M3UDMC, for urban autonomous vehicles. The framework integrates multimodal scene representation, memory-augmented risk reasoning, memory-conditioned dynamic potential field reconstruction, and constrained model predictive control within a bi-timescale architecture. Unlike loosely coupled modular pipelines, M3UDMC establishes an explicit information flow from multimodal observation to scene representation, memory-state update, potential-field parameter modulation, MPC risk-cost construction, and control execution. The slow timescale updates semantic memory and scenario-level risk priors, while the fast timescale performs prediction-conditioned potential field reconstruction and risk-aware MPC optimization. The dynamic potential field is not used as a standalone controller; instead, it provides a differentiable risk cost for the constrained MPC formulation, where vehicle dynamics, actuator limits, road boundaries, and minimum-distance constraints are explicitly considered. The proposed framework is evaluated through high-fidelity simulation, hardware-in-the-loop validation, and real-world road tests. Compared with representative baselines, including Apollo 8.0, end-to-end reinforcement learning with DDPG, and fixed-potential MPC, M3UDMC reduces the collision rate from 21.3% and 18.7% to 7.8% in the tested scenarios. Ablation studies further indicate that memory augmentation and dynamic potential field reconstruction contribute to improved decision quality under occlusions, signal transitions, and rare interaction events. The results demonstrate that M3UDMC improves the balance among safety, efficiency, and real-time feasibility in representative urban scenarios, while dense-traffic scalability, parameter adaptation, and cross-city generalization remain important directions for future work.

1. Introduction

Autonomous driving in urban environments remains a challenging task because the ego vehicle must interact with heterogeneous traffic participants under dynamic road, signal, and behavioral constraints. Compared with structured highway scenarios, urban driving involves denser multi-agent interactions, frequent right-of-way changes, occlusions, and rare but safety-critical events such

* Corresponding author.
E-mail address: cong-zhang@mail.tsinghua.edu.cn (C. Zhang).

https://doi.org/10.1016/j.multra.2026.100324
Received 1 April 2026; Received in revised form 9 May 2026; Accepted 18 May 2026

as unprotected crossings, sudden cut-ins, signal transitions, and emergency braking (Kendall et al., 2019). These events require the driving system to perceive the current environment, anticipate the future motion of surrounding agents, and generate safe and smooth control actions within strict computational constraints. Although recent advances in perception, prediction, and control have substantially improved autonomous driving performance, achieving reliable real-time decision-making in complex urban scenarios remains difficult.

Existing autonomous driving methods can be roughly divided into modular pipelines, learning-based decision policies, and optimization-based control approaches. Modular systems decompose the driving task into perception, prediction, planning, and control modules, which improves interpretability and engineering maintainability. Representative industrial platforms, such as Apollo, follow this design philosophy and provide transparent interfaces among functional modules (ApolloAuto 2023). However, the sequential nature of modular pipelines may lead to weak information coupling and error accumulation, especially when upstream perception or prediction uncertainty is not sufficiently reflected in downstream planning and control. End-to-end and reinforcement-learning-based methods provide an alternative by directly learning decision policies from data and interactions (Kendall et al., 2019; Chen et al., 2020). These methods can improve behavioral adaptability in some scenarios, but they usually require extensive training data, careful reward design, and large-scale scenario coverage. Their interpretability and constraint satisfaction also remain challenging in safety-critical urban driving.

Trajectory prediction and interaction-aware planning have recently attracted increasing attention because future motion uncertainty plays a crucial role in autonomous driving decisions. Transformer-based and probabilistic prediction models, such as mmTransformer, AgentFormer, and WayFormer, have improved the modeling of heterogeneous agents, map constraints, and multimodal future trajectories (Liu et al., 2021; Yuan et al., 2021; Nayakanti et al., 2023). Decision-conditioned prediction further incorporates high-level intent or rule-related information into forecasting, which helps align predicted motion with downstream planning objectives. Nevertheless, prediction accuracy alone is insufficient for closed-loop autonomous driving. The predicted future states must be transformed into a control-oriented risk representation that can guide real-time optimization while respecting vehicle dynamics, road boundaries, and safety constraints.

Potential-field-based methods provide an intuitive way to represent driving risk by constructing attraction and repulsion fields around goals, obstacles, lanes, and other traffic participants. They are computationally efficient and interpretable, which makes them attractive for real-time planning. However, conventional potential fields often rely on fixed parameters and manually designed field shapes, making them less adaptive to dynamic urban interactions, signal phase changes, occlusions, and long-tail scenarios. Recent studies have attempted to reconstruct dynamic potential fields to improve context awareness and control adaptability. However, if the potential field is directly used to drive control decisions without an explicit optimization interface, the resulting behavior may still be sensitive to parameter choices and difficult to analyze in terms of feasibility and safety. Therefore, a more rigorous integration between risk-field representation and constrained control is needed.

Model predictive control (MPC) provides a transparent and constraint-aware framework for autonomous vehicle control. By repeatedly solving a finite-horizon optimization problem, MPC can incorporate vehicle dynamics, actuator limits, road boundaries, comfort requirements, and safety-distance constraints (Yu et al., 2021; Mayne et al., 2000). However, standard MPC performance depends heavily on how interaction risk is represented in the objective function and constraints. A fixed or weakly adaptive risk model may cause conservative behavior, delayed response, or insufficient anticipation in complex multi-agent scenarios. Therefore, combining adaptive risk representation with constrained MPC is a promising direction for improving both interpretability and real-time decision quality.

Another relevant direction is memory-augmented and language-enhanced decision-making. Inspired by cognitive decision processes, memory mechanisms can store and retrieve historical interaction patterns, which helps improve decision consistency in recurrent or similar scenarios (Kahneman, 2011; Parisi et al., 2019). Trajectory prediction surveys have also emphasized the importance of historical context and interaction memory for anticipating human and vehicle motion (Rudenko et al., 2020). More recently, language-model-based driving systems have explored the use of semantic reasoning for high-level decision-making in autonomous driving (Besold et al., 2017; Bommasani et al., 2021; Shao et al., 2024; Hu et al., 2025; Kuang et al., 2024; Yuan et al., 2024). Despite these advances, directly mapping memory- or language-based reasoning results to low-level control actions is still risky because it may bypass vehicle dynamics and safety constraints. A suitable interface is needed so that historical context and semantic reasoning can influence control optimization in a quantitative but constrained manner.

Motivated by these observations, this paper proposes a Multimodal Multiscale Decision-Making and Control framework, termed M3UDMC, for autonomous driving in complex urban scenarios. The framework integrates multimodal scene representation, memory-augmented risk reasoning, dynamic potential field reconstruction, and MPC-based control within a bi-timescale architecture. The slow timescale is responsible for updating semantic memory and scenario-level risk priors, while the fast timescale performs dynamic risk-field reconstruction and constrained control optimization. In contrast to a loosely connected modular cascade, the proposed framework uses the memory module to modulate risk-field parameters rather than directly outputting steering, throttle, or braking commands. The reconstructed dynamic potential field is then evaluated within the MPC horizon and incorporated into the optimization objective as a differentiable risk cost. In this way, high-level memory-augmented reasoning affects low-level control through a traceable risk representation, while hard safety requirements remain handled by the constrained MPC formulation.

The proposed dynamic potential field is designed as a prediction-informed and memory-conditioned risk representation. Instead of relying on fixed attraction and repulsion coefficients, it incorporates predicted distributions of surrounding agents, traffic-signal semantics, road-boundary information, and historical interaction context. This design enables the risk field to adapt to changes in interaction intensity, prediction uncertainty, and right-of-way constraints. Meanwhile, the potential field is not treated as an independent controller. Its role is to reshape the MPC cost landscape by assigning higher costs to regions with larger predicted

interaction risk. Vehicle dynamics, actuator limits, road-boundary constraints, and minimum-distance requirements are explicitly considered in the MPC formulation, which improves the physical interpretability and feasibility of the generated control actions.

The main contributions of this paper are summarized as follows.

1. An explicitly coupled multimodal multiscale decision-control framework is proposed for complex urban autonomous driving. The framework integrates multimodal scene representation, memory-augmented risk reasoning, dynamic risk-field reconstruction, and MPC-based control within a bi-timescale architecture, thereby clarifying how high-level reasoning supports real-time constrained control.
2. A memory-conditioned dynamic potential field is developed to represent spatiotemporal interaction risk. Compared with static or manually designed potential fields, the proposed field is reconstructed from predicted agent distributions, traffic-rule semantics, road-boundary information, and historical interaction memory, and is embedded into the MPC objective as a risk-aware cost term.
3. The proposed framework is evaluated through simulation, hardware-in-the-loop validation, and real-world road tests. Comparisons with representative rule-based, learning-based, and optimization-based baselines demonstrate improved safety and decision quality in the tested scenarios, while ablation studies verify the contributions of memory augmentation and dynamic potential field reconstruction.

The remainder of this paper is organized as follows. Section 2 reviews related work on multimodal perception and prediction, autonomous driving decision-making, dynamic potential field modeling, and MPC-based control. Section 3 presents the proposed M3UDMC framework, including the bi-timescale architecture, memory-augmented risk reasoning, dynamic potential field reconstruction, and risk-aware MPC formulation. Section 4 describes the experimental setup, baseline configurations, and evaluation results. Section 5 discusses limitations, parameter sensitivity, latency scalability, and generalization issues. Section 6 concludes the paper.

2. Literature review

2.1. Multimodal perception and trajectory prediction

Multimodal perception and trajectory prediction are fundamental for autonomous driving in complex urban environments. Early autonomous driving perception systems mainly relied on single-modality sensors, such as cameras or LiDAR, to detect surrounding vehicles, pedestrians, and road structures. Datasets such as KITTI, nuScenes, and the Waymo Open Dataset promoted the development of vision- and LiDAR-based perception algorithms and provided standardized benchmarks for object detection and scene understanding in driving scenarios (Cui et al., 2024; Geiger et al., 2012; Caesar et al., 2020; Sun et al., 2020). With the rapid development of deep learning, convolutional neural networks and attention-enhanced detectors substantially improved the recognition of traffic participants in structured road environments (Bochkovskiy et al., 2020). However, single-modality perception remains sensitive to adverse weather, illumination changes, occlusions, and sensor degradation. For example, perception performance can deteriorate significantly under fog, rain, or nighttime conditions, which makes multimodal fusion necessary for robust urban driving (Bijelic et al., 2020; Zhang et al., 2024).

Recent studies have increasingly focused on fusing heterogeneous sensory inputs, including camera images, LiDAR point clouds, radar signals, map information, and traffic signal states (Liao et al., 2024). Multimodal fusion can improve environmental understanding by exploiting complementary characteristics of different sensors. Cameras provide rich semantic information, LiDAR offers accurate geometric structure, and radar improves robustness under adverse weather and long-range detection conditions. Nevertheless, multimodal fusion is not only a feature-level combination problem. In urban scenarios, different sensors usually have different sampling rates, fields of view, noise characteristics, and failure modes. Therefore, temporal synchronization, uncertainty estimation, and reliability-aware weighting are crucial for downstream prediction and decision-making (Feng et al., 2020; Arnold et al., 2022).

Trajectory prediction further extends perception from current-state estimation to future-state inference. Traditional motion prediction methods relied on kinematic models, maneuver recognition, or probabilistic filtering, which are interpretable but limited in capturing complex multi-agent interactions. With the emergence of sequence modeling and attention mechanisms, transformer-based prediction models have shown strong capability in representing long-range temporal dependencies and interactions among heterogeneous traffic participants. For instance, mmTransformer introduced multimodal multi-agent trajectory prediction with transformer-based interaction modeling (Liu et al., 2021). AgentFormer combined agent-aware attention with generative modeling to capture diverse future trajectories in social interaction scenarios (Yuan et al., 2021). WayFormer further improved efficient spatiotemporal representation for motion forecasting by integrating agent history, map information, and interaction features (Nayakanti et al., 2023).

Although these methods have improved prediction accuracy, several challenges remain. First, prediction uncertainty is often not sufficiently propagated to downstream planning and control. A prediction model may output multiple possible trajectories, but the planner must still determine how such uncertainty should influence risk assessment and control optimization. Second, many prediction models emphasize open-loop forecasting metrics, such as average displacement error and final displacement error, while their effect on closed-loop safety and control feasibility is less directly evaluated. Third, decision-conditioned forecasting has shown potential for aligning prediction with planning objectives by incorporating rule-related or intent-level information, but the interface between predicted motion distributions and constrained control remains insufficiently explored. These limitations indicate that trajectory prediction should be coupled with control-oriented risk representation rather than treated as an isolated upstream module.

2.2. Autonomous driving decision-making and control

Autonomous driving decision-making and control methods can be broadly divided into rule-based modular pipelines, learning-based policies, and optimization-based approaches. Rule-based modular systems, represented by industrial platforms such as Apollo and Autoware, decompose autonomous driving into perception, prediction, planning, and control modules. This architecture is interpretable, auditable, and suitable for engineering implementation (ApolloAuto 2023; Kato et al., 2018). In such systems, rule compliance and safety margins can be explicitly encoded through lane-level constraints, traffic regulations, and predefined planning logic. However, modular pipelines may suffer from error accumulation across modules. When perception uncertainty, prediction ambiguity, or abnormal interactions are not sufficiently reflected in downstream planning, the final control behavior may become conservative, delayed, or unsafe in long-tail urban scenarios (Zhang et al., 2026; Khatib, 1986).

Learning-based methods provide another direction for autonomous driving decision-making. Reinforcement learning and end-to-end driving models can learn policies from interaction data and optimize decision behavior through trial-and-error. Kendall et al. (Kendall et al., 2019) demonstrated that deep reinforcement learning can be used for lane following and vehicle control, while Chen et al. (2020) explored interpretable end-to-end urban autonomous driving with latent reinforcement learning. These methods can improve adaptability in certain scenarios and reduce manual engineering of decision rules. However, they usually require large-scale training data, careful reward design, and extensive scenario coverage. In addition, their learned policies may be difficult to verify, and their behavior under rare safety-critical events remains uncertain. These limitations restrict the direct deployment of purely learning-based policies in complex urban traffic.

Large language models and vision-language models have recently been introduced into autonomous driving to enhance semantic reasoning and decision interpretation. For example, LMDrive explored the use of language models for high-level driving decision-making (Besold et al., 2017; Bommasani et al., 2021; Shao et al., 2024). Such models are promising for understanding traffic rules, natural-language instructions, and high-level driving intentions. However, directly mapping language-based reasoning results to low-level vehicle control remains risky. Large-model reasoning may be computationally expensive, difficult to verify, and sensitive to prompt or context design. More importantly, steering, throttle, and braking commands must satisfy vehicle dynamics, actuator limits, and safety constraints. Therefore, large-model or memory-based reasoning should not replace constrained control; instead, it should provide interpretable priors or risk cues that can be integrated into a verifiable optimization framework.

Optimization-based methods, especially model predictive control, provide a transparent and constraint-aware framework for autonomous vehicle control. MPC repeatedly solves a finite-horizon optimization problem and can explicitly consider vehicle dynamics, road boundaries, actuator constraints, comfort requirements, and collision-avoidance constraints (Yu et al., 2021; Mayne et al., 2000). Compared with black-box policies, MPC has better interpretability and is easier to analyze from the perspective of feasibility and constraint satisfaction. However, standard MPC relies heavily on the quality of the cost function and risk representation. If the surrounding interaction risk is represented by fixed weights or static obstacle penalties, MPC may fail to anticipate rapidly changing risks caused by occlusions, signal transitions, or multi-agent conflicts. Therefore, a key challenge is how to construct an adaptive and control-oriented risk representation that can be embedded into MPC without sacrificing interpretability and real-time feasibility.

2.3. Dynamic potential field and risk-aware MPC

Potential-field methods have long been used for robot motion planning and obstacle avoidance (Khatib, 1986). The basic idea is to construct attractive fields around goals and repulsive fields around obstacles, allowing the planned motion to follow the negative gradient of the resulting field. This formulation is computationally efficient and intuitive, which makes it attractive for real-time autonomous driving. In vehicle planning, potential fields can represent the influence of surrounding vehicles, road boundaries, lane centers, traffic rules, and desired driving targets (Galceran et al., 2015; Schimpe and Diermeyer, 2020). Compared with discrete rule-based planning, potential-field methods provide a continuous representation of spatial risk and can be conveniently integrated with trajectory optimization.

However, conventional potential-field methods have well-known limitations. Fixed attraction and repulsion coefficients may not adapt to rapidly changing traffic contexts. In dense multi-agent urban scenarios, a manually designed field may lead to excessive conservatism, local minima, oscillatory behavior, or abrupt maneuvers. These problems become more pronounced at intersections, where the risk level depends not only on geometric distance but also on signal phase, right-of-way relations, predicted trajectories, and interaction uncertainty. Context-dependent potential-field formulations have been used to improve responsiveness compared with static fields (Galceran et al., 2015; Schimpe and Diermeyer, 2020). Nevertheless, if the potential field directly determines control actions, the resulting behavior may still be sensitive to parameter choices and difficult to evaluate in terms of safety constraints.

Recent studies have attempted to combine risk fields with optimization-based control. In this paradigm, the potential field is not used as an independent controller but as a risk-aware cost or constraint term in trajectory optimization. Such a formulation is more suitable for autonomous driving because vehicle dynamics, actuator limits, lane boundaries, comfort requirements, and safety-distance constraints can be explicitly retained in the optimization problem. For MPC-based autonomous driving, the potential field can reshape the cost landscape by penalizing high-risk regions while leaving the final control command to be determined by constrained optimization. This integration can improve the physical interpretability of the decision-control process and avoid directly converting heuristic potential gradients into steering or braking commands.

Despite this progress, two important issues remain unresolved. First, most existing potential-field designs still rely on manually selected or weakly adaptive parameters. The field intensity and spatial distribution are often not explicitly conditioned on prediction uncertainty, semantic risk, or historical interaction context. Second, the relationship between high-level decision reasoning and low-

level potential-field parameters is often not rigorously defined. As a result, it remains unclear how semantic decisions, memory retrieval, or scenario-level reasoning quantitatively influence the field used by the controller. This gap indicates that a dynamic potential field should be formulated with explicit parameter sources, clear physical meanings, and a well-defined interface to the MPC objective.

2.4. Memory-augmented reasoning and research gap

Memory mechanisms have been widely studied in continual learning and sequential decision-making. Continual learning aims to enable models to acquire new knowledge while retaining previously learned capabilities, thereby alleviating catastrophic forgetting (Parisi et al., 2019). In autonomous driving, memory can be used to store recurrent interaction patterns, rare scenario features, or historical decision outcomes. This is particularly relevant for urban driving, where similar conflict patterns may repeatedly occur at intersections, merging zones, or signal-transition areas. Human motion and vehicle trajectory prediction surveys have also emphasized the importance of historical context and interaction memory for anticipating future behavior (Rudenko et al., 2020).

However, memory-augmented reasoning must be carefully integrated into safety-critical driving systems. If memory retrieval or language-based reasoning directly outputs low-level control commands, the resulting behavior may bypass vehicle dynamics, actuator constraints, and safety verification. Therefore, memory should be used as an intermediate reasoning source rather than a direct controller. A more appropriate design is to let memory provide scenario-level risk priors, interaction cues, or parameter modulation signals, which are then incorporated into a physically constrained planning or control framework. This design preserves the adaptability of memory-based reasoning while maintaining the interpretability and feasibility of optimization-based control.

Based on the above review, three research gaps can be identified. First, multimodal perception and trajectory prediction methods have improved environmental understanding, but prediction uncertainty is not always transformed into a control-oriented risk representation. Second, dynamic potential fields provide an intuitive description of interaction risk, but many existing formulations remain parameter-sensitive and lack an explicit connection to high-level reasoning and constrained control. Third, memory-augmented and language-enhanced decision methods can improve semantic understanding, but their outputs require a traceable and safe interface before influencing low-level vehicle control.

To address these gaps, this paper proposes M3UDMC as an explicitly coupled decision-control framework. The framework does not treat multimodal fusion, memory reasoning, dynamic risk representation, and MPC control as isolated modules. Instead, memory-augmented reasoning provides scenario-level risk priors, the dynamic potential field transforms these priors and predicted agent distributions into a continuous risk representation, and MPC uses the reconstructed field as a risk-aware cost term under vehicle dynamics, road-boundary, actuator, and safety-distance constraints. In this way, the proposed method aims to bridge high-level reasoning and low-level constrained control without relying on direct heuristic conversion from semantic decisions to control commands.

3. Methodology

This section presents the proposed Multimodal Multiscale Decision-Making and Control framework, termed M3UDMC. The framework is designed for autonomous driving in complex urban scenarios, where the ego vehicle must respond to multi-agent interactions, traffic signal constraints, occlusions, and long-tail events under real-time computational constraints. To improve the clarity of the framework presentation, the methodology is organized into four parts. First, the overall architecture and bi-timescale information flow are introduced. Second, the multimodal scene representation and memory-augmented risk reasoning mechanism are formulated. Third, the memory-conditioned dynamic potential field is reconstructed as a control-oriented risk representation. Finally, the dynamic potential field is embedded into a risk-aware MPC formulation, and the computational implementation is discussed.

Different from an end-to-end driving policy or a loosely coupled modular pipeline, M3UDMC establishes an explicit connection between high-level reasoning and low-level control optimization. The memory-augmented reasoning module does not directly output steering, throttle, or braking commands. Instead, it generates structured risk priors and historical interaction cues, which modulate the parameters of the dynamic potential field. The reconstructed potential field is then evaluated within the MPC horizon and introduced into the optimization objective as a differentiable risk cost. In this way, semantic reasoning influences control optimization through a traceable risk representation, while vehicle dynamics, actuator limits, road boundaries, and minimum-distance requirements are handled by the constrained MPC formulation.

3.1. Overall framework and bi-timescale information flow

The overall architecture is shown in Fig. 1. Unlike a simple cascade of independent perception, prediction, decision-making, and control modules, M3UDMC establishes an explicit information flow from multimodal observation to control execution. At each time step, camera images, LiDAR/radar observations, historical trajectories, traffic signal states, and road semantic information are first fused into a unified scene representation. The memory-augmented reasoning module then retrieves relevant historical experiences and generates semantic decision variables, including interaction risk level, driving intent, and dynamic field weights. These semantic variables are not directly converted into steering or braking commands. Instead, they modulate the parameters of the dynamic potential field, which is further incorporated into the MPC objective as a differentiable risk cost. The final control command is obtained through constrained finite-horizon optimization.

The framework adopts a bi-timescale design to balance semantic reasoning and real-time control. The slow timescale updates the memory state, retrieves historical interaction cases, and generates risk priors at a lower frequency. The fast timescale performs

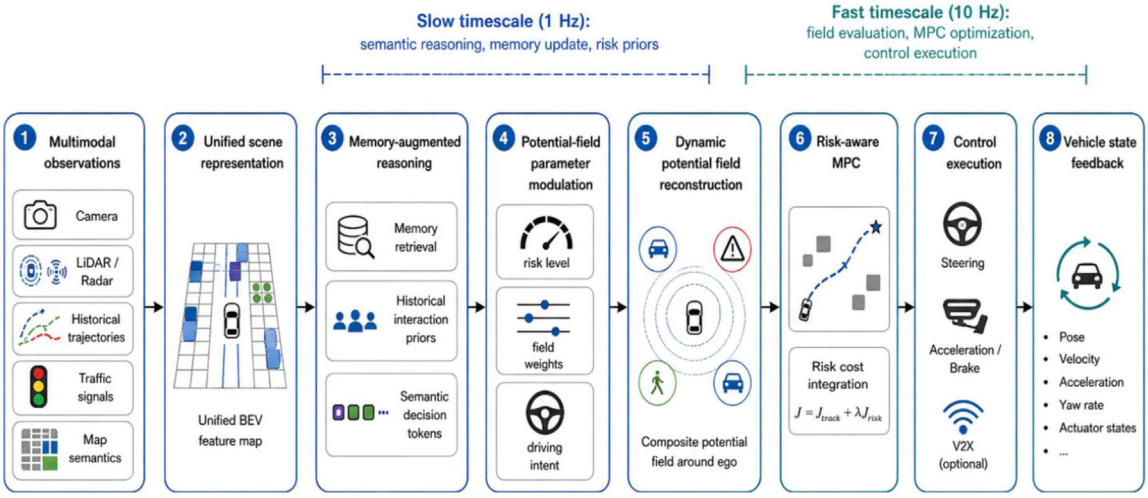

Fig. 1. Overall architecture and information flow of the proposed M3UDMC framework.

surrounding-agent prediction, dynamic potential field reconstruction, MPC optimization, and control execution at a higher frequency. In the implementation used in this paper, the memory-augmented reasoning layer is updated at 1 Hz, while the control layer is executed at 10 Hz. This design reduces the computational burden of memory retrieval and semantic reasoning while preserving the real-time responsiveness of the controller.

Let the multimodal observation at time t be denoted as:

$$\mathbf{O}_t = \left\{ I_t, x_{t-k:t}, S_t \right\} \tag{1}$$

where I_t denotes image observations, $x_{t-k:t}$ historical ego and surrounding-agent trajectories within a time window [t–k, t], and S_t denotes traffic signal states. The overall decision-control mapping is defined as:

$$a_t = \pi\left(\mathbf{O}_t; \theta\right) = \pi_{fast}\left(\pi_{slow}\left(\mathbf{O}_t\right)\right) \tag{2}$$

where π_{slow} produces high-level semantic decisions (risk levels, intent, field weights), and π_{fast} transforms them into continuous control actions.

This explicit coupling clarifies the role of each module. Multimodal fusion provides a structured scene representation; memory-augmented reasoning supplies historical interaction priors; dynamic potential field reconstruction transforms these priors into a continuous risk landscape; and MPC generates dynamically feasible control commands under safety and actuation constraints.

3.2. Multimodal fusion mechanism

The multimodal fusion module integrates heterogeneous sensor information to construct a unified representation of the driving environment. In urban autonomous driving, different modalities show different reliability under adverse conditions. For example, camera-based perception may degrade under fog, glare, or low illumination, whereas LiDAR and radar provide more stable geometric observations but may suffer from sparsity or noise. To address this issue, M3UDMC employs a reliability-aware dynamic fusion mechanism that adapts sensor weights in real time according to environmental context and observation uncertainty.

To dynamically adjust the contribution of each sensor modality based on the context, we introduce a dynamic weighting function $w_i(t)$, which is updated at each time step t. The weight for each sensor modality i is computed using the following formula:

$$w_i(t) = \frac{\alpha_i(t) \cdot \exp\left(-\lambda \cdot \delta_i(t)\right)}{\sum_{j=1}^{N} \alpha_j(t) \cdot \exp\left(-\lambda \cdot \delta_i(t)\right)} \tag{3}$$

where $\alpha_i(t)$ is the reliability score of sensor i at time t, which is determined by factors like signal strength, sensor health, and environmental conditions (e.g., in fog, LiDAR may have higher weight than camera data). $\delta_i(t)$ is the data error or uncertainty associated with the sensor data, calculated as the difference between predicted and observed values. λ is a tuning parameter that controls how sensitive the weights are to the data error. This dynamic weighting mechanism ensures that the system continuously prioritizes the most reliable sensor inputs, improving robustness in dynamic and challenging environments.

To ensure the accuracy and reliability of the input data, we apply real-time data filtering. The filtering mechanism adapts based on the confidence in the sensor data and the context of the environment. The filtered data (t) is obtained by applying a Kalman filter (or similar filtering method) to each sensor input:

$$x_i(t) = F\left(x_i(t), w_i(t)\right) \tag{4}$$

where $x_i(t)$ is the raw sensor input at time t. $F()$ is the filtering function (e.g., Kalman Filter) that adjusts the data according to its weight $w_i(t)$ and estimates the noise in the sensor readings. The filter ensures that erroneous or unreliable data (e.g., noisy LiDAR scans in a rainstorm) is either corrected or ignored, improving the quality of the fused data.

The fused data $\mathbf{X}_{\text{fuse}}(t)$ is then computed by combining the filtered sensor data from each modality using the weights $w_i(t)$:

$$X_{fuse}(t) = \sum_{i=1}^{N} w_i(t) \cdot x_i(t) \tag{5}$$

where $\mathbf{X}_{\text{fuse}}(t)$ is the fused multimodal data at time t. The weighted sum of the filtered data ensures that the most relevant and reliable sensor information is emphasized.

Once the multimodal data is fused, it is passed into the high-level decision-making module, which uses this comprehensive, real-time understanding of the environment to make decisions. The decision-making function is based on the fused data and includes a decision threshold (t) that determines when an action should be taken (e.g., to stop, turn, or accelerate):

$$\mathbf{u}_{\text{decision}} = f_{\text{decision}}\left(\mathbf{X}_{\text{fuse}}(t), \theta(t)\right) \tag{6}$$

where $\mathbf{u}_{\text{decision}}$ is the output of the decision-making module, representing the high-level decision (e.g., a decision to brake or steer); $f_{\text{decision}}()$ is the decision function that takes the fused multimodal data and computes the appropriate action; $\theta(t)$ is a dynamic decision threshold that changes based on real-time environmental conditions, ensuring adaptive decision-making.

3.3. Memory-conditioned dynamic potential field reconstruction

The dynamic potential field is designed to represent the spatial and temporal risk distribution around the ego vehicle. In the original implementation, the potential field consists of visual, interaction, and rule-related components. To make the formulation clearer and less heuristic, this section reformulates the potential field as a memory-conditioned and prediction-informed risk representation.

M3UDMC integrates a contrastive memory encoder with incremental updating to store and retrieve representative driving experiences. A new sample z_f with outcome y is added to memory M when:

$$\|z_f - \hat{z}_f\|^2 > \in \ or \ \max_i \ cos\left(z_f, m_i\right) < \tau \tag{7}$$

where $\hat{z}_f$ is the autoencoder reconstruction of z_f, are existing memory entries, ϵ is the reconstruction error threshold, and τ is the similarity threshold. Memory retrieval is accelerated by Hierarchical Navigable Small World (HNSW) indexing (Malkov and Yashunin, 2020). A fine-tuned TinyLlama-1.1B model (Zhang et al., 2024), conditioned on retrieved memories and scene descriptors, produces structured decision tokens:

$$y_t = \left\{r_t, \beta_t, u_t^{intent}\right\} \tag{8}$$

where r_t is the estimated risk level, β_t are field-weights vectors, and u_t^{intent} is the semantic intent command. The semantic tokens do not directly output low-level controls. Instead, they modulate the dynamic potential field parameters and influence MPC through a risk-sensitive cost term.

The visual potential field can be written as:

$$U_v = \sum_i \beta_v \cdot \exp\left(-\kappa\theta_i\right) \cdot \frac{1}{\|p - p_i\|} \tag{9}$$

where p is the ego position, p_i is the obstacle position, θ_{i} is the relative angle, k is the angular decay factor, and β_v is the visual weight. The interaction potential field can be written as:

$$U_{\text{int}} = \sum_j \beta_{\text{int}} \exp\left(-\frac{TTC_j}{\tau}\right) \tag{10}$$

where TTC_j is the time-to-collision with agent j, τ is a scaling constant, and β_{int} is the interaction weight. The rule potential filed can be formulated as:

$$U_r = \begin{cases} \beta_r/d & if\ signal = rad, \\ -\beta_r \cdot v & if\ signal = green \end{cases} \tag{11}$$

where d is the distance to the stop line, v is the ego velocity, and β_r is the rule weight. The total potential field is:

$$U = \sigma\left(\beta_v\right)U_v + \sigma\left(\beta_{int}\right)U_{int} + \sigma\left(\beta_r\right)U_r \tag{12}$$

where $\sigma()$ normalizes each component. In the classical potential field method, the repulsion and attraction forces are modeled using fixed coefficients:

$$\begin{aligned} F_{\text{repel}} &= -\beta \cdot \frac{1}{r^2} \\ F_{\text{attr}} &= -\alpha \cdot \frac{1}{r_{\text{target}}^2} \end{aligned} \tag{13}$$

where F_{repel} and F_{attr} are the repulsion and attraction forces, respectively. β and α are the fixed repulsion and attraction coefficients. r is the distance to the obstacle, and r_{target} is the distance to the target position.

However, in M3UDMC, we introduce a dynamic coefficient adaptation to allow for real-time adjustment of these forces, depending on the surrounding conditions. The core idea is to adapt the repulsion and attraction coefficients based on environmental factors such as obstacle proximity, vehicle speed, traffic density, and pedestrian presence. The dynamic repulsion force is adjusted in real time based on the vehicle's proximity to surrounding obstacles. Specifically, the repulsion coefficient (t) is adjusted according to the distance to the nearest obstacle $r_{\text{obs}}(t)$ and the traffic density $\rho(t)$, which captures the crowding of surrounding vehicles or pedestrians:

$$\beta(t) = \beta_0 \cdot \left(1 + \gamma_{\text{repel}} \cdot \frac{r_{\text{obs}}(t)}{r_{\max}}\right) \cdot \left(1 + \kappa_{\text{traffic}} \cdot \rho(t)\right) \tag{14}$$

where β_0 is the base repulsion coefficient; γ_{repel} is a scaling factor for distance-dependent repulsion; $r_{\max}$ is the maximum safe distance; $\rho(t)$ is the traffic density at time t, capturing the number of surrounding vehicles and pedestrians. κ_{traffic} is a factor that adjusts the repulsion based on traffic conditions.

This adjustment ensures that the vehicle increases its repulsion strength in high-density traffic scenarios (to maintain safe distances) and decreases repulsion when the road is clear, improving both efficiency and safety. Similarly, the attraction force toward the target is dynamically adjusted based on the vehicle's current speed and the density of obstacles in the path:

$$F_{\text{attr}}(t) = \alpha_0 \cdot \left(1 + \gamma_{\text{speed}} \cdot v(t)\right) \cdot \frac{1}{r_{\text{target}}^2} \tag{15}$$

where α_0 is the base attraction coefficient. γ_{speed} is a scaling factor for the vehicle's speed $v(t)$, which adjusts the attraction force to facilitate faster movement in clear environments. By incorporating vehicle speed into the attraction force, M3UDMC ensures that the vehicle maintains optimal path-following behavior, accelerating when the path is clear, but adjusting when obstacles appear.

In M3UDMC, the potential field is not a simple, static landscape but a context-sensitive dynamic field that adapts to the environment. For example, the field's shape is altered depending on the longitudinal and lateral positions of the vehicle relative to surrounding traffic. This ensures smoother trajectory planning in narrow spaces or high-density traffic. The dynamic shaping of the potential field is achieved by introducing a context-dependent shaping function S_{field} that adjusts the field based on the vehicle's position and traffic context:

$$S_{\text{field}}(t) = \frac{1}{1 + \alpha_{\text{shape}} \cdot \left(\frac{\left|v(t) - v_{\text{neighbor}}(t)\right|}{max_v}\right)^{\zeta}} \tag{16}$$

where α_{shape} is a coefficient that controls the shaping sensitivity. $v(t)$ and $v_{\text{neighbor}}(t)$ are the speed of the vehicle and the neighboring vehicles' speeds, respectively. $\max_v$ is the maximum vehicle speed in the environment. ζ is a tuning parameter that controls how much the shape is altered based on relative speeds.

This function allows the potential field to shrink or expand based on the relative speed between the vehicle and its neighbors, effectively improving collision avoidance in high-speed scenarios and enhancing maneuverability in tight spaces. Combining the dynamic adjustments for both repulsion and attraction, the final dynamic potential field for M3UDMC is given by:

$$F_{\text{total}}(t) = F_{\text{attr}}(t) + F_{\text{repel}}(t) \tag{17}$$

The resulting dynamic potential field is not minimized as an independent planning objective. Instead, it is evaluated along the predicted ego trajectory and incorporated into the MPC objective as a risk-aware cost term.

It should be emphasized that the dynamic potential field is not used as a standalone controller. Instead, it provides a differentiable risk representation that is embedded into the MPC objective. Fig. 2 illustrates the module-level reconstruction process of the dynamic potential field. The module receives memory retrieval results, semantic decision tokens, and field-weight modulation signals, and outputs a continuous risk cost for MPC optimization.

3.4. Bi-timescale architecture and memory-augmented decision making

The bi-timescale architecture of M3UDMC decouples high-level decision-making from low-level control execution, enabling both components to operate at different frequencies. This architecture is designed to improve real-time performance by allowing high-level reasoning to make deliberative decisions at a lower frequency, while low-level control can operate at a higher frequency to generate timely, responsive commands.

The high-level decision-making module operates at 1 Hz frequency, processing data from various sensors to generate high-level decisions (such as whether to stop, turn, or change lanes). The decision-making process can be modeled by:

$$\mathbf{u}_{\text{decision}}(t) = f_{\text{high}}\left(\mathbf{X}_{\text{fuse}}(t)\right) \tag{18}$$

where $\mathbf{u}_{\text{decision}}(t)$ represents the high-level decisions made by the system at time t, $f_{\text{high}}()$ is the high-level decision function that outputs actions such as lane changes, speed adjustments, or stop/go decisions. This high-level reasoning is done once every second (1 Hz), allowing the system to make strategic decisions based on the overall context of the environment, such as traffic flow, pedestrian behavior, and road conditions. In contrast to high-level decision-making, the low-level control execution operates at a higher frequency, typically 10 Hz, to ensure smooth and responsive actions. The low-level control module takes the high-level decisions and translates them into real-time control commands, such as steering angle, throttle, and brake commands. The low-level control command is computed as:

$$\mathbf{u}_{\text{control}}(t) = f_{\text{low}}\left(\mathbf{u}_{\text{decision}}(t), x_t^{ego}, U_{dyn}\left(p_t\right)\right) \tag{19}$$

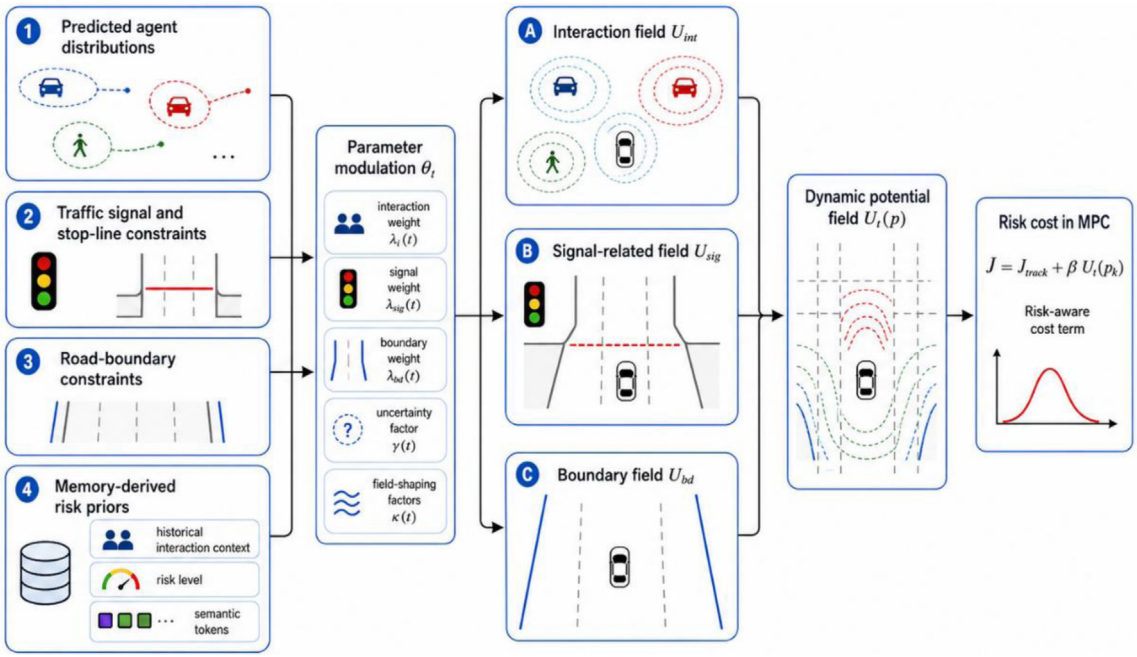

Fig. 2. Memory-conditioned dynamic potential field reconstruction module.

where x_t^{ego} is the current ego-vehicle state, and $U_{dyn}(p_t)$ is the dynamic potential field evaluated near the ego vehicle.

Since the high-level module operates at a lower frequency, the low-level controller does not wait for a new semantic decision at every control step. Instead, it uses the most recent high-level decision and continuously updates the control command based on real-time feedback:

$$u_{control}(t) = f_{low}\left(u_{decision}(t_s), x_t^{ego}, U_{dyn}(p_t)\right), t_s = \max\left\{t' \le t | u_{decision}(t') \; is \; updated\right\} \tag{20}$$

This mechanism allows the system to maintain strategic consistency while responding quickly to immediate traffic changes.

To further enhance system stability, we implement a dual-loop control mechanism that combines open-loop and closed-loop control strategies. The high-level module sets goals or targets, while the low-level control continuously adjusts the vehicle's trajectory and speed to meet these targets. The system ensures that even in unpredictable traffic conditions, such as sudden cut-ins or pedestrian crossings, the vehicle can react immediately to ensure safety. To enhance stability and responsiveness, the framework adopts a dual-loop control structure:

$$\tau_t = \tau_t^{ref} + \Delta\tau_t^{fb} \tag{21}$$

where τ_t^{ref} is the target trajectory provided by the high-level decision module, and $\Delta\tau_t^{fb}$ is the real-time feedback correction generated by the low-level controller. This dual-loop structure helps M3UDMC maintain a balance between strategic consistency and real-time responsiveness under the evaluated conditions.

Memory-augmented decision making is a key feature of M3UDMC that significantly enhances its ability to handle rare and long-tail events that may not be well-represented in the training data. Traditional decision-making frameworks in autonomous driving systems often rely on real-time sensor data and trained models, but they may struggle when faced with novel or infrequent situations. Our memory-augmented approach allows the system to recall past experiences and use them to inform current decision-making, improving both safety and robustness.

The memory module is structured to store past experiences that include both high-level decisions and sensor observations. These experiences are stored in a memory bank, which is continuously updated as the vehicle encounters new situations. Each memory entry consists of the sensor input at a specific time *t*, such as camera, LiDAR, and radar data; the high-level decision made by the system at that time (e.g., accelerate, turn, stop); the outcome or result of the decision (e.g., successful lane change, collision avoidance). The memory bank *M* is updated as:

$$\mathcal{M}(t) = \mathcal{M}(t-1) \cup \left\{\mathbf{x}_{\text{sensor}}(t), \mathbf{u}_{\text{decision}}(t), \mathbf{y}_{\text{outcome}}(t)\right\} \tag{22}$$

where $\mathbf{x}_{\text{sensor}}(t)$ is the sensor data at time *t*. $\mathbf{u}_{\text{decision}}(t)$ is the decision made by the system at time *t*. $\mathbf{y}_{\text{outcome}}(t)$ is the outcome or result of the decision made at time *t*. This memory bank allows the system to maintain a history of experiences that can be referenced during decision-making in similar future situations. When the vehicle encounters a new scenario, it uses the memory bank to retrieve relevant past experiences that resemble the current situation. The memory retrieval function retrieves the most similar experiences

based on the current input data:

$$\mathbf{u}_{\text{mem}}(t) = \text{MemoryRetrieve}\left(\mathbf{x}_{\text{input}}(t), \mathcal{M}\right) \tag{23}$$

where $\mathbf{x}_{\text{input}}(t)$ is the current sensor input. $\mathcal{M}$ is the memory bank; $\mathbf{u}_{\text{mem}}(t)$ is the memory-augmented decision based on similar past experiences. The MemoryRetrieve function uses a similarity measure (such as cosine similarity, Euclidean distance, or other distance metrics) to find past experiences that match the current input. Once the relevant memories are retrieved, they are used to guide the decision-making process, allowing the system to make more informed decisions when faced with unfamiliar situations. Incorporating past experiences into decision-making can significantly improve the system's response to rare events. The memory-augmented decision is adjusted by integrating the retrieved memory information into the decision-making process. This adjustment can be mathematically expressed as:

$$\mathbf{u}_{\text{adjusted}}(t) = f_{\text{adjust}}\left(\mathbf{u}_{\text{current}}(t), \mathbf{u}_{\text{mem}}(t)\right) \tag{24}$$

where $\mathbf{u}_{\text{current}}(t)$ is the current decision made by the system based on real-time data. $\mathbf{u}_{\text{mem}}(t)$ is the memory-augmented decision retrieved from the memory module. $f_{\text{adjust}}()$ is a function that adjusts the current decision based on the information from past experiences, ensuring that the final decision incorporates lessons learned from previous situations.

This function allows M3UDMC to fine-tune its decisions based on past successful outcomes, improving its ability to handle rare or unexpected situations that were not well-represented in training data. To prevent the memory bank from becoming overloaded and to ensure the system focuses on relevant experiences, we implement a forgetting mechanism that updates the memory bank over time. This ensures that outdated or irrelevant experiences are discarded, and only the most recent and useful memories are retained. The memory update rule is:

$$\mathcal{M}(t) = \text{Forget}(\mathcal{M}(t-1), \theta(t)) \tag{25}$$

where $\theta(t)$ is a forgetting threshold based on factors such as the age of memories and their relevance to current situations. $\text{Forget}()$ is a function that removes less relevant experiences from the memory bank.

3.5. Prediction-conditioned trajectory representation and MPC control

Historical ego motion is processed by multi-kernel depthwise separable convolutions:

$$h_k = \text{DSConv}_k\left(x_{t-k:t}\right), k \in \{3, 5, 7\} \tag{26}$$

where each kernel size k captures short-, medium-, and long-term motion trends. These are concatenated and fed into an interaction-aware GRU:

$$r_t = \sigma\left(W_r\left[x_t, h_{t-1}, z_d\right]\right) \tag{27}$$

$$u_t = \sigma\left(W_u\left[x_t, h_{t-1}, z_d\right]\right) \tag{28}$$

where z_d is the decision embedding, and are gating matrices. The future trajectory distribution is modeled as a Gaussian mixture (Bishop, 1994):

$$p(y|x) = \sum_{m=1}^{M} \pi_m N\left(y\middle|\mu_m, \sum\nolimits_m\right) \tag{29}$$

where π_m are mixing coefficients, μ_m are mode means, and Σ_m are covariance matrices. Additionally, the controller solves:

$$\min_{u0:T} \sum_{t=0}^{T} \left(\left\| x_t - x_t^{ref} \right\|_Q^2 + \left\| u_t \right\|_R^2 + U\left(x_t\right) \right) \tag{30}$$

subject to

$$x_{k+1} = f\left(x_k, u_k\right)$$

$$x_k \in X_{road}$$

$$d_{i,k} \geq d_{safe}, \forall i \in C_t$$

where x_k is the predicted ego state, u_t is the control input, x_t^{ref} is the reference state, Q, R, and P are weighting matrices, X_{road} is the drivable road area, d_{safe} is the minimum safety distance, and C_t is the set of conflict-relevant agents.

The reconstructed potential field is therefore used as a risk-aware cost-shaping term, rather than as a direct control law. Safety-critical requirements are enforced by explicit MPC constraints, including vehicle dynamics, road boundaries, actuator limits, and minimum-distance constraints. Under bounded prediction errors and feasible MPC constraints, the planned trajectory satisfies the prescribed drivable-area and safety-distance conditions. Iterative LQR is used to accelerate the MPC solution, with a typical solve time of approximately 5 ms in the tested configuration. Neural modules are optimized with TensorRT and distributed across the

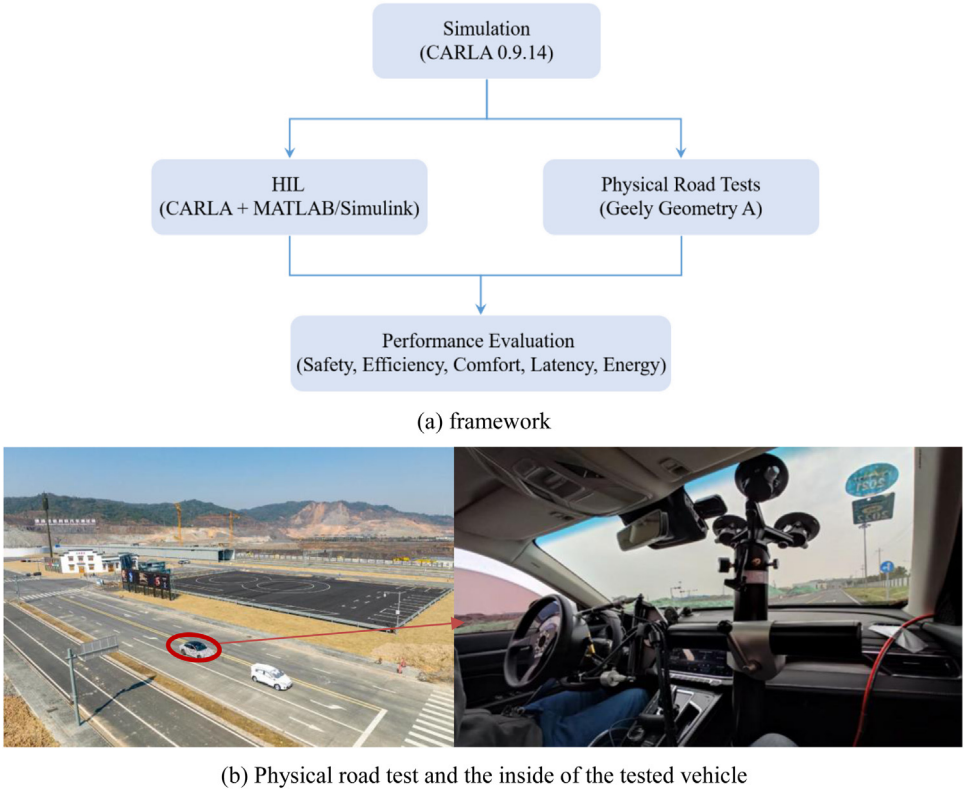

Fig. 3. Evaluation framework of our experiment.

heterogeneous computing units of an NVIDIA Orin AGX. Vision processing is executed on GPU tensor cores, trajectory planning is executed on CUDA cores, and decision reasoning is executed on deep-learning accelerators. Multi-rate inputs are synchronized within 50 ms. A safety monitor triggers emergency braking when the safety-risk threshold exceeds 0.7. The complete system operates at up to 15 Hz, with an average end-to-end latency of approximately 400 ms and a power consumption of 18 W.

4. Experimental results and performance analysis

The evaluation of the proposed M3UDMC framework was conducted through a three-layer validation pipeline, combining high-fidelity simulation, hardware-in-the-loop (HIL) testing, and real-world deployment. Simulation experiments were performed on CARLA 0.9.14, which provides photorealistic rendering and dynamic traffic scenarios. The scenarios included unprotected left turns, pedestrian crossings, signal transition zones, and sudden braking events. Environmental conditions were varied across daytime, nighttime, rain, and fog in order to test robustness under sensor degradation. Hardware-in-the-loop (HIL) validation was carried out by coupling CARLA with MATLAB/Simulink. This configuration enabled the dynamic potential field controller to be executed in real time with closed-loop feedback, ensuring that safety overrides and latency constraints could be evaluated without physical risk. As shown in Fig. 3(a), the evaluation framework integrates simulation, HIL, and real-world deployment.

Physical road tests, shown as Fig. 3(b), were conducted on a Geely Geometry A prototype vehicle, equipped with an NVIDIA Orin AGX (275 TOPS) for onboard computation and an OXTS RT3000 GNSS-INS for high-precision ground-truth localization. Tests were performed on campus roads, covering intersections, roundabouts, and mixed traffic with pedestrians and bicycles.

To benchmark the proposed M3UDMC framework, we implemented and compared it against several state-of-the-art baselines. These include Apollo 8.0, representing a rule-based modular pipeline (ApolloAuto 2023); End-to-End DDPG, representing reinforcement learning policies trained in CARLA (Chen et al., 2020); and Fixed-Potential MPC, representing conventional model predictive control with static β coefficients (Yu et al., 2021). Additionally, we also included three recent, end-to-end learning-based models: UniAD, which leverages deep learning for integrated decision-making; VAD, which focuses on vehicle action decision-making using reinforcement learning; and DriveVLM (Tian et al., 2024), which integrates vision-language models for enhanced decision-making in diverse traffic environments. These models were selected to provide a comprehensive comparison across traditional model-based methods and cutting-edge learning-based systems, enabling an evaluation of the strengths and weaknesses of M3UDMC. The following evaluation metrics were defined and rigorously quantified:

Table 1
Quantitative comparison of M3UDMC and baseline methods.

Method	CR(%)	TE(%)	RC	DL(ms)	EC (W)	Control Execution (Hz)	End-to-End Latency (ms)
Apollo 8.0 (Pipeline)	21.3	Baseline	0.09	150	0.60	1	300
End-to-End DDPG	18.7	12	0.15	310	0.55	2	350
Fixed-Potential MPC	16.4	–8	0.10	98	0.65	10	180
UniAD	7.8	18	0.10	110	0.62	5	380
VAD	9.2	15	0.12	130	0.60	6	410
DriveVLM	10.5	20	0.09	125	0.63	5	390
M3UDMC (ours)	**7.8**	**23**	**0.11**	**120**	**0.62**	**10**	**300**

Collision rate (CR) — the proportion of trials ending in a collision, defined as:

$$CR = \frac{N_{col}}{N_{total}} \times 100\% \tag{31}$$

where N_{col} is the number of trials that resulted in any collision (with vehicles, pedestrians, or static obstacles), and N_{total} is the total number of trials conducted.

Travel efficiency (TE) — the relative improvement in travel time compared with the rule-based baseline, defined as:

$$TE = \frac{T_{baseline} - T_{method}}{T_{baseline}} \times 100\% \tag{32}$$

where T_{method} is the average time to complete the route using the evaluated method, and $T_{baseline}$ is the corresponding time under Apollo 8.0.

Ride comfort (RC) — quantified by the variance of jerk, i.e., the derivative of acceleration:

$$RC = \mathrm{Var}\left(\frac{da_t}{dt}\right) \tag{33}$$

where a_t denotes longitudinal and lateral acceleration at time t. Lower values of RC indicate smoother maneuvers.

Decision latency (DL) — the average computational delay between perception input and control output, measured as:

$$DL = \frac{1}{N}\sum_{i=1}^{N}\left(t_i^{out} - t_i^{in}\right) \tag{34}$$

where t_i^{in} and t_i^{out} denote the input and output timestamps of the decision module for trial i.

Energy consumption (EC) — the average system power usage during operation, computed as:

$$EC = \frac{1}{T}\int_0^T P(t)dt \tag{35}$$

where $P(t)$ is the instantaneous power draw and T is the total duration of the trial.

For fair comparison, all methods were evaluated under the same scenario set, perception inputs, vehicle dynamics constraints, and computational platform unless otherwise specified. The simulation scenarios were generated using identical traffic configurations and random seeds when applicable. The rule-based, learning-based, vision-language-model-based, and optimization-based baselines were tested using the same initial conditions, road geometries, traffic signal settings, and surrounding-agent trajectories. In addition, the key hyperparameters and system configurations of each method, including planning frequency, safety-distance thresholds, speed limits, MPC horizon, weighting matrices, reinforcement-learning parameters, and large-model inference settings, are reported in the Appendix to improve reproducibility and transparency.

4.1. Quantitative results

A comprehensive evaluation was conducted across more than 280 urban driving scenarios, including unprotected left turns (32%), pedestrian crossings (28%), signal transition zones (22%), and emergency braking (18%). Environmental conditions spanned daytime (45%), nighttime (30%), and adverse weather such as rain and fog (25%).

Table 1 presents a detailed comparison of M3UDMC against several state-of-the-art autonomous driving methods, including Apollo 8.0, End-to-End DDPG, Fixed-Potential MPC, UniAD, VAD, and DriveVLM. Notably, the proposed M3UDMC framework achieved a 7.8% collision rate, significantly outperforming traditional methods such as Apollo 8.0 (21.3%) and End-to-End DDPG (18.7%). Additionally, it demonstrated superior travel efficiency, with a 23% improvement over reinforcement learning baselines, driven by its ability to dynamically adjust the decision-making process and optimize trajectory planning in real-time. This efficiency boost is especially notable during complex urban navigation, where M3UDMC adapted rapidly to changes in the environment. In terms of ride comfort, M3UDMC exhibited a jerk variance of 0.11, which is comparable to other state-of-the-art method such as UniAD (0.10) and DriveVLM (0.09), while outperforming DDPG (0.15). This indicates that M3UDMC delivers a smooth and comfortable ride, ensuring minimal disruptions for passengers despite the challenges of urban driving. The decision latency of M3UDMC was maintained at 120 ms, which is significantly lower than that of End-to-End DDPG (310 ms) and competitive with methods such

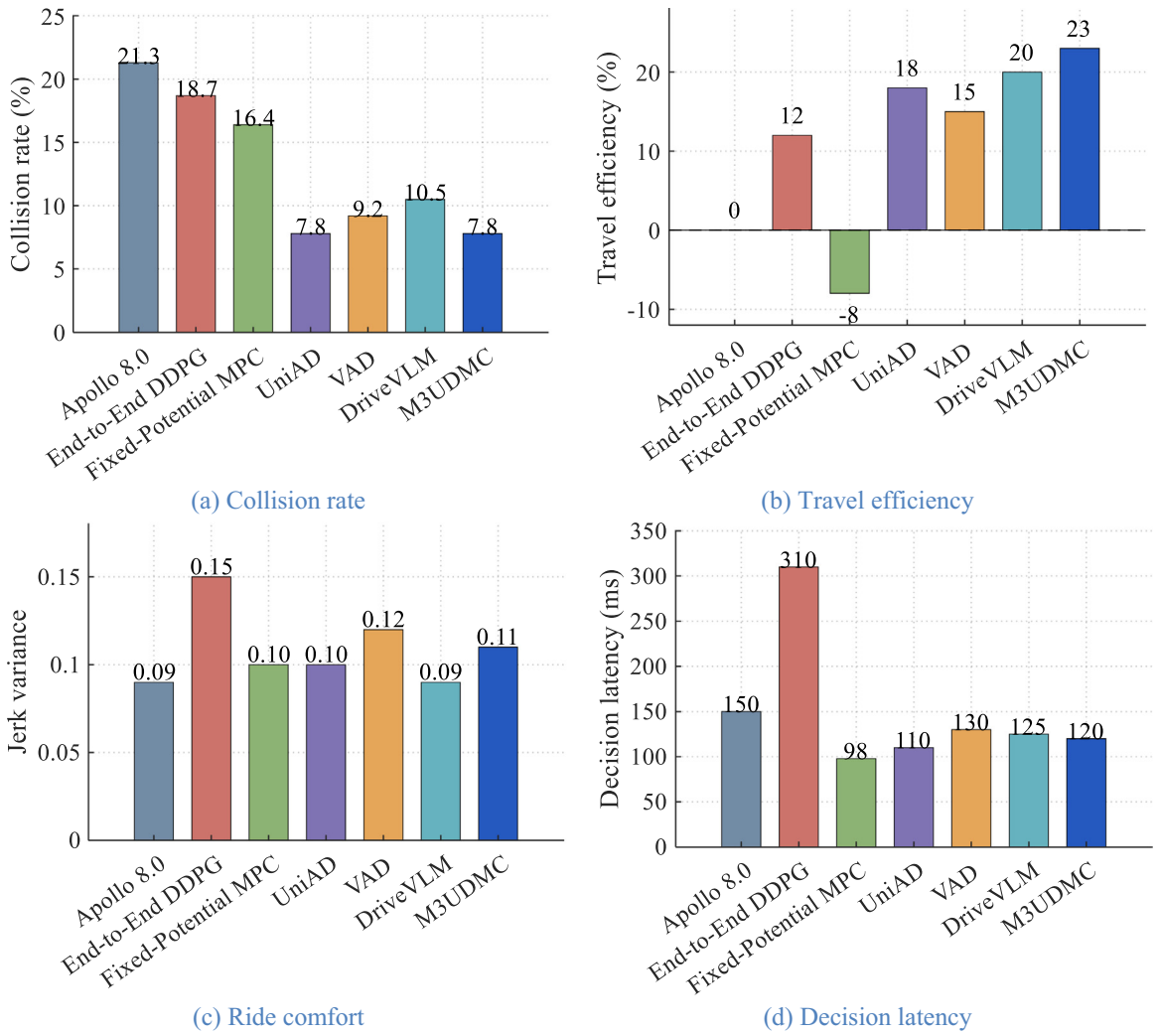

Fig. 4. Quantitative comparison of M3UDMC with baseline methods across collision rate, efficiency, comfort, and latency.

as UniAD (110 ms). This ensures that M3UDMC remains highly responsive to real-time changes in the environment. The system's end-to-end latency of 300 ms is slightly higher than MPC (180 ms), but still allows for real-time performance, making M3UDMC a strong choice for dynamic traffic environments. Regarding energy consumption, M3UDMC demonstrated a consumption rate of 18 W, which is competitive within the tested methods and reflects a balance between high performance and energy efficiency. The system's energy efficiency is calculated at 0.62 J/W, which is favorable compared with other approaches like End-to-End DDPG (0.55 J/W), showing that M3UDMC can operate effectively without excessive power demands.

The best result for each metric is highlighted in bold. For travel efficiency, a positive value indicates reduced travel time compared with the reference baseline, whereas a negative value indicates increased travel time. For collision rate, jerk variance, decision latency, end-to-end latency, and power consumption, lower values indicate better performance. For control frequency, a higher value indicates better real-time execution capability.

To provide a clearer visual comparison among different methods, the quantitative results are further plotted in Fig. 4. Unlike the previous single-axis visualization, the revised figure presents each metric in an independent subplot because collision rate, travel efficiency, ride comfort, and latency have different physical meanings and numerical scales. In particular, the travel efficiency subplot includes both positive and negative ranges, and a zero-reference line is added to avoid misleading interpretation when a method produces a negative efficiency gain.

As shown in Fig. 4, M3UDMC achieves a low collision rate while maintaining favorable travel efficiency and real-time responsiveness. Compared with Apollo 8.0 and End-to-End DDPG, the proposed method substantially reduces the collision rate, indicating that the memory-conditioned risk reasoning and dynamic potential field are effective in safety-critical scenarios. Compared with UniAD, which achieves a comparable collision rate, M3UDMC provides higher control frequency and lower end-to-end latency, suggesting a more balanced trade-off between safety and real-time feasibility. The travel efficiency results also show that some baseline

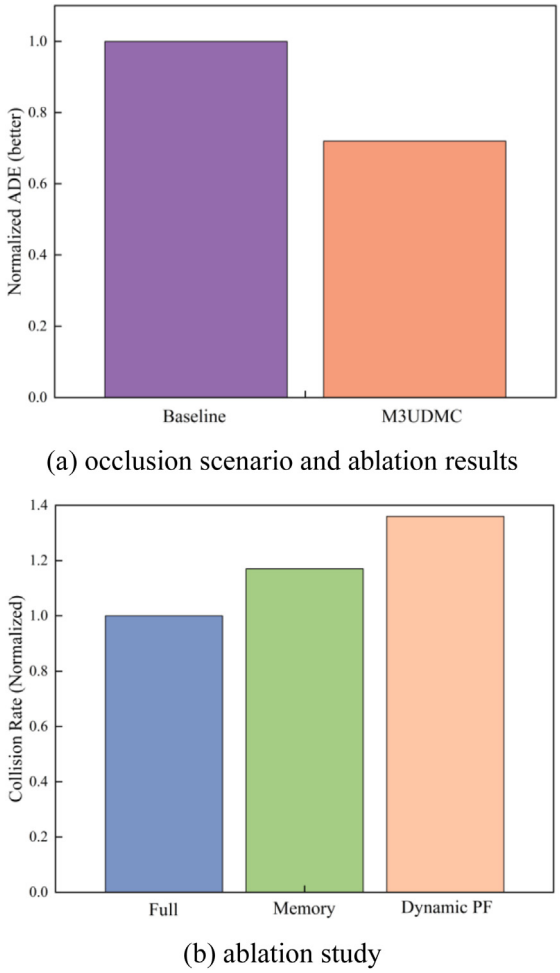

Fig. 5. Representative case studies and ablation analysis.

methods may become overly conservative or inefficient under complex interactions, whereas M3UDMC maintains positive efficiency improvement by jointly considering risk, interaction context, and control feasibility.

Overall, M3UDMC achieves a balanced performance across safety, efficiency, comfort, and real-time responsiveness in the evaluated scenarios. The results suggest that the proposed memory-conditioned risk representation and MPC-based control improve decision quality compared with representative baselines, while further validation is still needed in denser and more diverse real-world environments.

4.2. Case studies and ablation analysis

To gain deeper insights, three representative cases were analyzed as shown in Fig. 5 shown. In signal transition scenarios, M3UDMC updated the potential field coefficients within 600 ± 50 ms following a signal change. This rapid adaptation yielded a smooth deceleration trajectory, whereas Fixed-Potential MPC exhibited abrupt braking, resulting in higher jerk peaks. The smoother profile highlights the advantage of dynamic β reconfiguration in safety-critical situations.

In occlusion scenarios (Fig. 5a), where pedestrians emerged from blind spots, M3UDMC reduced the average displacement error (ADE) by 29% compared with baselines. This improvement significantly lowered the number of emergency braking events, thereby decreasing collision risks by 37%. The ablation study (Fig. 5b) underscores the importance of memory augmentation and dynamic potential fields. Removing the memory module increased collision frequency by 18%, while disabling dynamic potential adaptation raised emergency braking events by 34%. These results confirm that both components are critical for safe and reliable performance in rare or long-tail scenarios.

The practicality of M3UDMC for embedded deployment was validated on an NVIDIA Orin AGX. Module-wise profiling revealed average latencies of 15.6 ms for perception, 4.3 ms for trajectory modeling, and 95 ms for decision generation. The total end-to-end latency was kept within 300 ms, fulfilling real-time control requirements in dynamic environments.

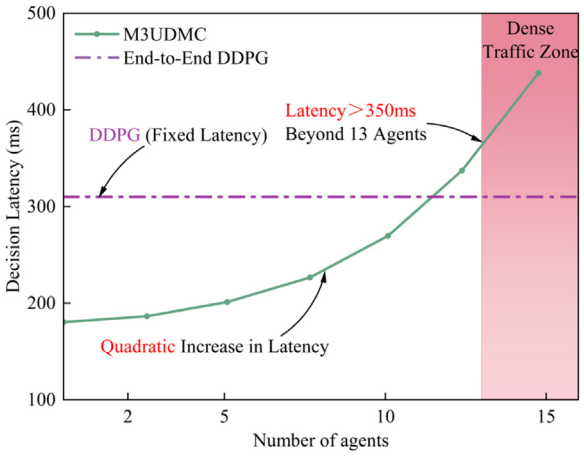

Fig. 6. Scalability evaluation of decision latency with increasing number of interacting agents.

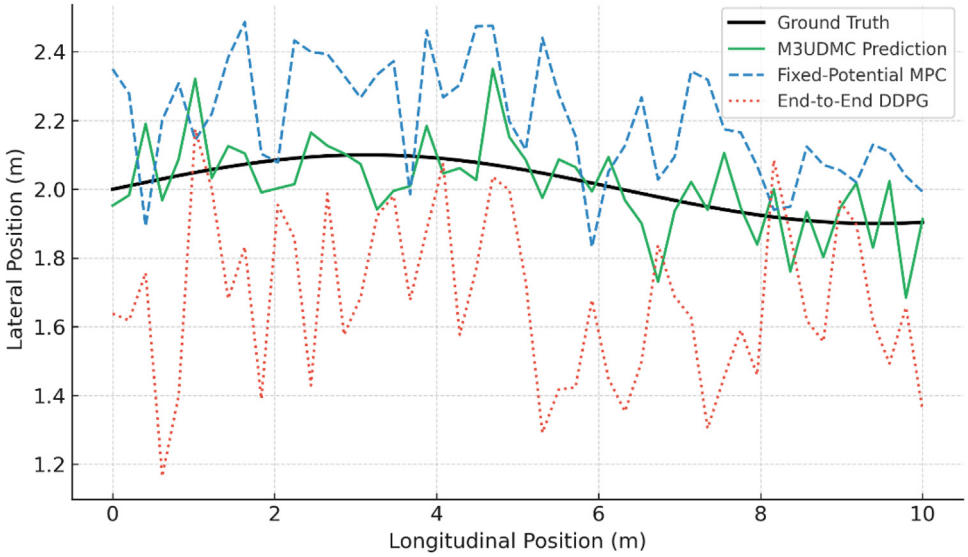

Fig. 7. Trajectory comparison under occlusion scenario.

Scalability was further assessed under scenarios involving up to 16 interacting agents. As shown in Fig. 6, decision latency increased approximately quadratically with the number of agents. When the number of agents exceeded 13, latency occasionally surpassed 350 ms, highlighting current challenges in dense traffic scenarios. While RL methods such as DDPG maintained constant latency, this was at the cost of significantly higher collision rates. These results suggest that future improvements in agent importance sampling and interaction encoder optimization will be crucial for enhancing scalability in dense urban environments.

The experimental evaluation provides compelling evidence of the effectiveness of the proposed M3UDMC framework. By dynamically reconstructing potential fields, the system was able to substantially reduce collision rates, achieving more than a 60% improvement compared with state-of-the-art baselines. At the same time, the bi-timescale architecture facilitated smoother and more efficient driving, reducing average travel times by over 20% while preserving ride comfort. The deployment results further demonstrated that the framework can sustain real-time operation on embedded automotive hardware with only moderate energy consumption, confirming its feasibility for practical applications.

As illustrated in Fig. 7, trajectory prediction under occlusion demonstrates the advantage of the proposed framework. When pedestrians emerged from partially obstructed regions, predictions by M3UDMC aligned closely with the ground-truth trajectories, exhibiting an average displacement error reduction of approximately 29% compared with MPC and more than 40% compared with end-to-end DDPG. In contrast, MPC predictions tended to lag and deviate laterally, while RL-based trajectories showed unstable oscillations. These findings emphasize that potential-conditioned prediction, informed by decision semantics, significantly enhances anticipation of hidden agents and reduces collision risks in long-tail urban scenarios.

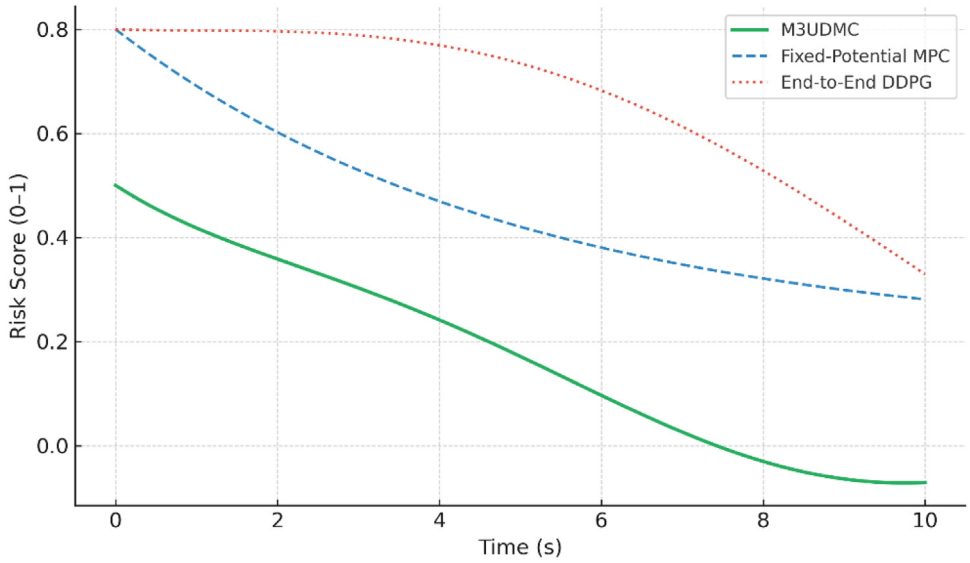

Fig. 8. Risk score evolution during a signal transition.

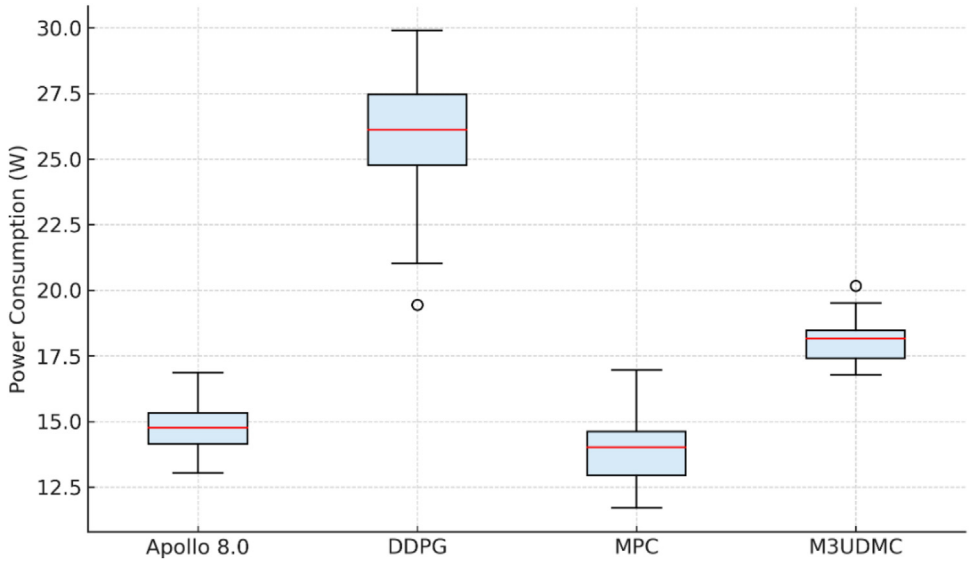

Fig. 9. Energy consumption distribution across 50 runs.

Fig. 8 presents the temporal evolution of risk scores during a signal transition event. The risk profile generated by M3UDMC decays smoothly following the onset of a green signal, maintaining bounded values and avoiding abrupt fluctuations. By comparison, MPC maintained elevated risk levels for extended durations, and RL policies exhibited oscillatory risk responses, occasionally overshooting safety thresholds. The smooth risk attenuation of M3UDMC is consistent with its dynamic β-field adaptation, which allows rapid yet stable reconfiguration in response to traffic-rule changes. This property contributes directly to safer and more comfortable driving behaviors during critical decision phases.

Energy consumption analysis across 50 independent runs is visualized in Fig. 9. M3UDMC exhibited a median power draw of approximately 18 W with narrow interquartile range, indicating both efficiency and stability in embedded deployment. By contrast, RL-based controllers consumed substantially more energy, with medians around 26 W and wide variability, reflecting unstable utilization of computational resources. MPC consumed less power on average ($\approx$14 W) but failed to achieve comparable safety and efficiency. These results confirm that M3UDMC achieves a balanced compromise, providing improved performance without excessive energy overhead, which is critical for long-duration on-vehicle operation.

Taken together, these findings indicate that M3UDMC shows advantages over several representative modular and end-to-end learning baselines in the evaluated scenarios. Nevertheless, limitations remain in handling sensor degradation under adverse weather,

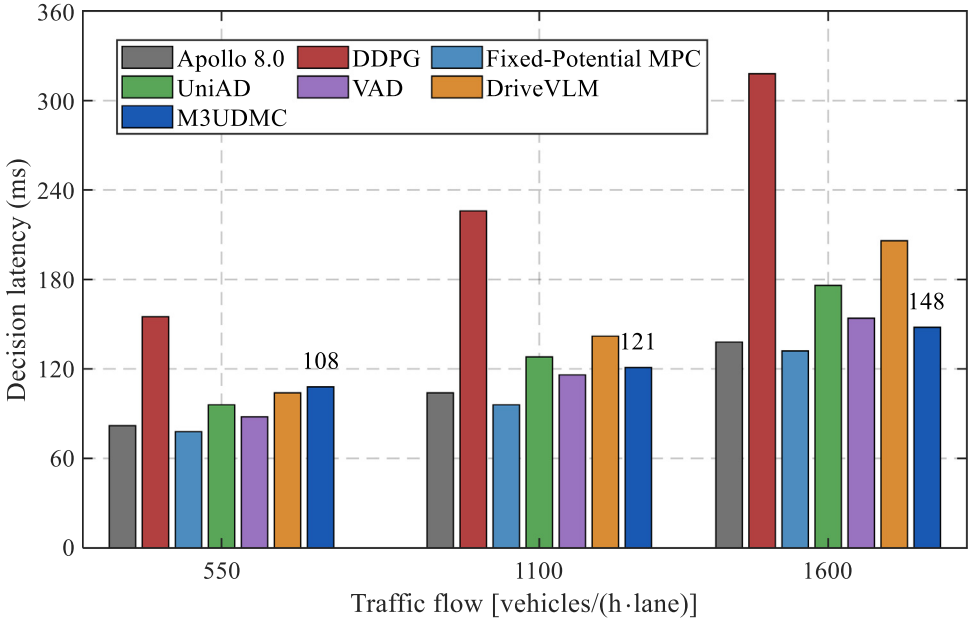

Fig. 10. Decision latency comparison under different traffic-flow conditions.

in adapting to traffic systems with different geographical or regulatory characteristics, and in scaling efficiently when the number of interacting agents becomes very large. These insights provide a clear foundation for the future research directions outlined in Section 5 and Section 6.

While we tested M3UDMC in 280 simulation scenarios, we acknowledge that this number may seem limited for validating claims such as a 62% collision reduction. The scenarios included a variety of urban conditions (e.g., intersections, pedestrian crossings, and lane changes) with environmental factors such as rain and fog. To ensure the robustness of the system, we also conducted HIL testing and physical road tests on campus roads, which provided a controlled but sufficiently diverse set of conditions. However, we recognize that testing in more diverse real-world environments will be a focus for future work to improve the robustness of the system further.

4.3. Latency and scalability analysis

Decision latency is a key factor affecting the deployability of autonomous-driving decision-control systems. In dense urban traffic, the number of surrounding agents increases significantly, which directly affects interaction modeling, potential-field reconstruction, and trajectory optimization. To evaluate the computational scalability of different methods, Fig. 10 compares the decision latency under three traffic-flow levels, namely 550, 1100, and 1600 vehicles/(h·lane). These traffic-flow levels correspond to low-, medium-, and high-density urban traffic conditions, respectively. The compared methods include Apollo 8.0, End-to-End DDPG, Fixed-Potential MPC, UniAD, VAD, DriveVLM, and the proposed M3UDMC.

As shown in Fig. 10, the decision latency of all methods increases as the traffic flow becomes denser. This trend is expected because more surrounding agents need to be perceived, evaluated, and incorporated into the decision-making process. End-to-End DDPG shows the most significant latency increase, rising from 155 ms under low traffic flow to 318 ms under high traffic flow. DriveVLM also exhibits a clear increase in latency, reaching 206 ms under the high-density condition, mainly due to the additional computational cost of vision-language reasoning. In comparison, M3UDMC maintains a relatively moderate latency increase, from 108 ms at 550 vehicles/(h·lane) to 148 ms at 1600 vehicles/(h·lane). Although its latency is slightly higher than Fixed-Potential MPC and VAD under low-density traffic, M3UDMC shows better scalability under denser conditions. This is because the proposed framework does not reconstruct interaction risks for all perceived agents indiscriminately. Instead, conflict-relevant agents are selected through risk-aware filtering, and the dynamic potential field is reconstructed mainly around the spatiotemporal conflict region of the ego vehicle.

Compared with Fixed-Potential MPC, M3UDMC requires additional computation for memory-conditioned risk modulation and dynamic potential-field reconstruction. However, the latency increase remains acceptable for real-time decision-control execution. These results indicate that M3UDMC achieves a practical balance between interaction-awareness and computational efficiency. Nevertheless, the increasing latency under high traffic flow also suggests that extremely dense urban scenarios remain challenging, and further optimization through sparse risk-field representation, parallel computation, and hardware-aware deployment is still necessary.

The latency increase with traffic density mainly originates from interaction-field reconstruction and MPC evaluation. If N_a denotes the number of surrounding agents and N_p denotes the prediction horizon, the computational complexity of interaction-field evaluation

is approximately $\mathbf{O}(N_a\ N_p)$. This scaling raises deployment concerns in extremely dense urban scenarios. In the current implementation, conflict-relevant agents are selected from the perceived surrounding agents, and the dynamic potential field is reconstructed based on these agents. However, when the number of vehicles, pedestrians, and cyclists increases rapidly, explicit modeling of all agents becomes computationally expensive.

Three mitigation strategies can be adopted. First, risk-aware agent filtering can retain only agents located within the potential spatiotemporal conflict region of the ego vehicle. Second, low-risk agents can be represented by an aggregated background potential field rather than individual interaction fields. Third, potential-field reconstruction and candidate trajectory evaluation can be parallelized on embedded GPUs or dedicated accelerators. After risk-aware filtering, the effective complexity can be reduced to $\mathbf{O}(N_c\ N_p)$, where N_c is the number of conflict-relevant agents and is usually much smaller than N_a. Although the current system satisfies real-time requirements in the evaluated scenarios, extremely dense traffic remains a challenging case and should be further addressed through sparse risk-field representation and hardware-aware optimization.

5. Discussion

The experimental results indicate that M3UDMC achieves a favorable balance among safety, efficiency, comfort, latency, and energy consumption in the evaluated urban driving scenarios. Compared with Apollo 8.0, End-to-End DDPG, Fixed-Potential MPC, UniAD, VAD, and DriveVLM, the proposed framework achieves a lower collision rate while maintaining real-time feasibility on embedded hardware. The improvement mainly results from three coupled mechanisms: memory-augmented reasoning for long-tail scenarios, dynamic potential field reconstruction for adaptive risk representation, and bi-timescale control for coordinating semantic decisions and fast actuation.

However, the results should be interpreted with appropriate caution. Since the compared methods represent different technical paradigms, their performance can be affected by implementation details, hyperparameter tuning, reward-function design, and computational optimization. To reduce unfair comparison, all methods were evaluated under the same scenario distribution, perception inputs, vehicle platform, and computational hardware. For reinforcement-learning baselines, the learning rate, batch size, exploration noise, discount factor, reward weights, and training steps should be reported. For MPC-based baselines, the same vehicle model, actuator constraints, and reference trajectory settings should be used. For large-model-based baselines, model size, context length, temperature, and fine-tuning strategy should be reported. These details are summarized in the Appendix to improve transparency.

5.1. Interpretation of performance improvements

The superior safety performance of M3UDMC is mainly attributed to the dynamic coupling between memory retrieval and potential-field reconstruction. In signal transition, occlusion, and sudden cut-in scenarios, the memory module retrieves similar historical interaction patterns and provides risk priors for the current scene. These priors modulate the dynamic potential-field parameters, allowing the risk field to expand, contract, or reweight its components according to the current traffic context. Compared with Fixed-Potential MPC, which uses static coefficients, M3UDMC can adapt the field intensity and spatial distribution when the surrounding traffic state changes.

The improvement in travel efficiency is not achieved by simply increasing target speed. Instead, M3UDMC reduces unnecessary conservatism by distinguishing high-risk and low-risk regions in the local driving area. When the predicted interaction risk is low, the attraction component toward the reference trajectory becomes more dominant. When the predicted interaction risk increases, the interaction and rule-related components impose stronger penalties on unsafe regions. This adaptive trade-off allows the ego vehicle to maintain efficiency without sacrificing safety.

The ride-comfort results show that M3UDMC does not introduce excessive control oscillation despite the use of dynamic potential fields. This is because the potential field is embedded into the MPC objective as a smooth risk-aware cost term, while acceleration, steering, and control variation are explicitly constrained or penalized. Therefore, the controller avoids directly following the negative gradient of the potential field, which is a common source of oscillatory behavior in classical artificial potential-field methods.

The dynamic potential field is a central component of M3UDMC, but it should not be interpreted as an independent safety controller. Its role is to provide a continuous and differentiable risk representation for MPC. The final control command is generated by solving a constrained finite-horizon optimization problem, rather than by directly applying a potential-field force.

This distinction is important. Classical potential-field methods may suffer from local minima, discontinuous repulsion, and oscillatory trajectories when the field gradient is directly mapped to control commands. In M3UDMC, the potential field reshapes the MPC cost landscape, while hard constraints enforce road-boundary limits, actuator feasibility, and minimum safety distance. Thus, safety-critical requirements are handled by constrained optimization, whereas the potential field improves risk awareness during trajectory selection.

The framework does not claim unconditional global safety or global optimality. The safety of the generated trajectory is conditional on bounded prediction error, reliable perception input, and the feasibility of the MPC problem. When these conditions hold, the planned trajectory satisfies the prescribed drivable-area and safety-distance constraints. When these conditions are violated, for example, under severe sensor degradation, long-term occlusion, incorrect signal recognition, or abnormal agent behavior outside the evaluated distribution, additional fallback strategies such as emergency braking and rule-based safety monitoring remain necessary.

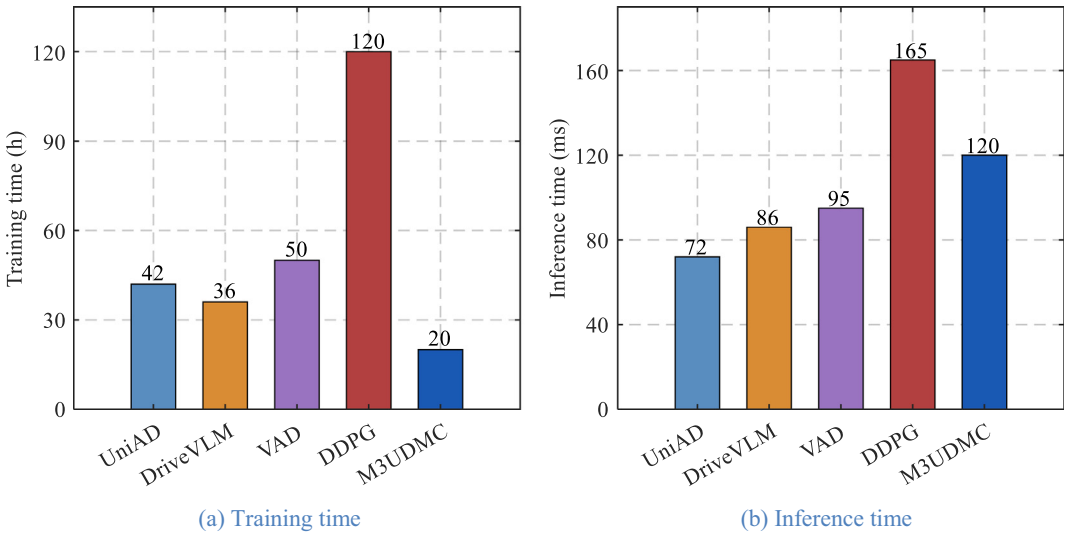

Fig. 11. Training time and inference time comparison for learning-based and hybrid decision-control methods.

5.2. Training cost, inference cost, and embedded feasibility

In addition to online decision latency, training cost and inference time are also important for evaluating the practicality of learning-based and hybrid autonomous-driving methods. Pure end-to-end reinforcement-learning methods usually require extensive simulation interactions, while large-model-based methods introduce additional inference overhead. To further compare the computational characteristics of different approaches, Fig. 11 reports the training time and inference time of UniAD, DriveVLM, VAD, End-to-End DDPG, and the proposed M3UDMC.

Fig. 11(a) compares the training time of different methods. End-to-End DDPG requires the longest training time, approximately 120 h, because it relies on trial-and-error interactions with the simulation environment to learn driving policies. VAD, UniAD, and DriveVLM require 50 h, 42 h, and 36 h, respectively, reflecting the training or adaptation cost of data-driven perception-planning models. In contrast, M3UDMC requires approximately 20 h of training, which is significantly lower than the other learning-based methods. This reduction is mainly because M3UDMC does not rely entirely on end-to-end policy learning. Instead, it combines memory-augmented reasoning, dynamic potential-field reconstruction, and model-based MPC optimization, thereby reducing the amount of policy exploration required during training. Fig. 11(b) compares the inference time of different methods. UniAD, DriveVLM, and VAD show relatively lower inference times of 72 ms, 86 ms, and 95 ms, respectively. M3UDMC requires approximately 120 ms, which is higher than these data-driven baselines but lower than End-to-End DDPG, whose inference time reaches 165 ms. The additional inference cost of M3UDMC mainly comes from memory retrieval, risk-prior modulation, dynamic potential-field reconstruction, and MPC optimization. However, this cost brings improved interpretability and constraint awareness, since the final control command is generated through explicit risk-aware optimization rather than an implicit policy output.

Overall, Fig. 11 shows that M3UDMC achieves a favorable trade-off between training efficiency, inference cost, and decision-control reliability. Compared with End-to-End DDPG, it substantially reduces training time and inference latency. Compared with purely data-driven methods, it introduces a moderate inference overhead but provides a more interpretable decision-control process with explicit safety constraints. These results support the feasibility of deploying M3UDMC on embedded autonomous-driving platforms, while also indicating that further acceleration of memory retrieval and MPC evaluation would be beneficial for large-scale real-time applications.

The hybrid nature of M3UDMC improves flexibility but also introduces parameter sensitivity. Several parameters directly influence the final driving behavior. The risk-cost weight β affects the balance between safety and efficiency. The repulsion and attraction coefficients influence the conservativeness of obstacle avoidance and reference tracking. The spatial attenuation parameters determine how far surrounding agents affect the ego vehicle. The memory similarity threshold affects whether retrieved experiences are relevant to the current scenario. The MPC weighting matrices Q, R, and P determine the trade-off among tracking accuracy, control effort, and terminal convergence.

Therefore, reproducibility requires detailed reporting of system configurations and hyperparameters. For reproducibility, the manuscript provides the planning frequency, safety-distance thresholds, speed limits, MPC prediction horizon, control interval, weighting matrices, dynamic potential-field coefficients, reinforcement-learning hyperparameters, reward-function weights, training steps, and large-model inference settings in the Appendix.

The main configuration of M3UDMC, including decision frequency, control frequency, synchronization window, onboard computing platform, safety override threshold, memory retrieval strategy, and MPC solver, is summarized in the experimental setup. For parameters that are scenario-dependent, such as MPC weights, potential-field coefficients, safety-distance thresholds, and baseline hyperparameters, we report the selected values or the tuning protocol in Appendix A. For public baseline methods, the official config-

urations are followed unless otherwise specified, and only parameters modified for the present evaluation are explicitly reported. It should be noted that the reported parameters are calibrated for the tested vehicle platform and experimental scenarios. When transferring M3UDMC to a different vehicle platform, road geometry, traffic density, or driving culture, recalibration may be required. Future work should investigate adaptive parameter tuning and uncertainty-aware calibration to reduce dependence on manual adjustment.

The simulation, hardware-in-the-loop, and real-world experiments demonstrate the effectiveness of M3UDMC in representative urban scenarios, including unprotected interactions, pedestrian crossings, signal transitions, occlusions, and emergency braking. Nevertheless, the current validation is still limited in several aspects.

First, the real-world tests were conducted in limited urban environments. Although they include intersections, roundabouts, pedestrians, bicycles, and mixed traffic, they do not fully cover unseen cities with different road geometries, traffic rules, signal-control policies, or driving cultures. The geographic generalization of M3UDMC therefore requires further large-scale validation. Second, the framework depends on the reliability of multimodal perception and trajectory prediction. Severe sensor degradation, adverse weather, long-term occlusion, incorrect object tracking, or wrong traffic-signal recognition may lead to inaccurate risk-field reconstruction. In these cases, the memory module may retrieve irrelevant historical cases, and the potential field may overestimate or underestimate the actual interaction risk. Third, long-tail scenarios remain difficult to exhaustively validate. Although memory-augmented reasoning improves the handling of rare cases by retrieving similar past experiences, it cannot guarantee correct behavior in all unseen combinations of aggressive agents, ambiguous right-of-way, dense traffic, and degraded perception. Fourth, M3UDMC still requires careful calibration of potential-field and MPC parameters. The memory-conditioned modulation improves adaptability compared with fixed-potential MPC, but it does not eliminate the need for parameter tuning across different platforms and traffic environments.

These limitations suggest that M3UDMC should be regarded as an interpretable and adaptive step toward integrated decision-control systems, rather than a complete solution to all urban autonomous-driving challenges.

6. Conclusion

This paper proposed M3UDMC, a Multimodal Multiscale Decision-Making and Control framework for urban autonomous driving. The framework integrates multimodal scene representation, memory-augmented risk reasoning, memory-conditioned dynamic potential field reconstruction, and risk-aware MPC within a bi-timescale architecture. In contrast to loosely coupled modular pipelines, M3UDMC establishes a traceable information path from multimodal observation to scene representation, memory-state update, potential-field parameter modulation, MPC risk-cost construction, and final control execution. The memory module does not directly generate steering, throttle, or braking commands. Instead, it provides structured risk priors and historical interaction cues that modulate the parameters of the dynamic potential field.

The proposed dynamic potential field was formulated as a prediction-informed and memory-conditioned risk representation rather than a standalone controller. By combining visual, interaction, and rule-related risk components, the field describes spatiotemporal driving risk under multi-agent interactions, signal constraints, occlusions, and road-boundary limitations, and is embedded into the MPC objective as a differentiable risk cost. Vehicle dynamics, actuator limits, drivable-area constraints, and minimum-distance requirements are explicitly handled by the constrained optimization formulation. The proposed framework was evaluated through high-fidelity simulation, hardware-in-the-loop experiments, and real-world road tests. The results show that M3UDMC achieves improved safety and decision quality compared with representative rule-based, learning-based, vision-language-model-based, and optimization-based baselines in the tested scenarios. In particular, compared with Apollo 8.0 and end-to-end DDPG, M3UDMC reduces the collision rate from 21.3% and 18.7% to 7.8%, respectively. The case studies and ablation results further indicate that memory-augmented reasoning and dynamic potential field reconstruction are beneficial for handling occlusions, signal transitions, and long-tail interaction events, while the latency and embedded-computing analyses show that the bi-timescale design helps balance semantic reasoning capability and real-time control responsiveness.

Despite these improvements, several limitations remain. The performance of M3UDMC still depends on the reliability of multimodal perception and trajectory prediction. Severe sensor degradation, long-term occlusion, incorrect traffic-signal recognition, or abnormal behaviors outside the evaluated distribution may affect risk-field reconstruction and decision quality. In addition, the potential-field coefficients, MPC weights, and safety thresholds still require calibration when transferring the framework to different vehicle platforms, traffic densities, or road structures. Computational latency also increases with the number of surrounding agents, which may challenge deployment in extremely dense urban traffic. Future work will focus on adaptive parameter tuning, uncertainty-aware safety verification, risk-aware agent pruning, parallel implementation, and larger-scale validation in unseen cities and more diverse real-world traffic environments.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Yanbin Liu: Visualization, Validation, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Cong Zhang:** Validation, Supervision. **Shaohua Cui:** Writing – original

draft, Formal analysis, Data curation. **Guangyu Tian:** Writing – review & editing, Validation, Supervision, Methodology. **Yugong Luo:** Supervision, Resources, Project administration. **Lei Zhang:** Writing – review & editing, Resources.

Appendix

Key implementation settings of M3UDMC

The main system-level settings of M3UDMC are summarized in Table A1. These parameters are fixed across all simulation, hardware-in-the-loop, and real-world experiments unless otherwise specified.

The high-level reasoning module is executed at a lower frequency to reduce the computational burden of memory retrieval and semantic reasoning. The low-level controller runs at a higher frequency to maintain closed-loop responsiveness. Multi-rate sensor inputs are synchronized within a 50 ms temporal window. A rule-based safety monitor is used as a fallback layer, and emergency braking is triggered when the safety-risk score exceeds the predefined threshold.

Several parameters in M3UDMC are scenario-dependent, including the MPC weighting matrices, potential-field coefficients, safety-distance threshold, memory similarity threshold, and risk-cost weight. Instead of independently tuning these parameters on the test scenarios, all scenario-dependent parameters are selected using validation scenarios and then kept fixed during testing. The parameter-selection protocol follows three principles. First, safety-related parameters are calibrated before efficiency-related parameters. The minimum safety-distance threshold, safety override threshold, and interaction-risk weight are selected to avoid collision or near-collision events in validation scenarios. Second, comfort-related parameters are adjusted after the safety constraints are satisfied. These parameters include the control-variation penalty and acceleration/steering bounds. Third, efficiency-related parameters, such as the attraction coefficient and reference-tracking weight, are tuned only after safety and comfort requirements are met.The dynamic potential field parameters are selected according to their physical roles. The interaction-risk coefficient controls the penalty imposed on predicted conflict regions. The spatial attenuation coefficient determines the influence range of surrounding agents. The signal-related coefficient strengthens the penalty near stop lines during signal transition phases. The boundary-related coefficient penalizes trajectories close to lane or road boundaries. These coefficients are kept fixed during testing after validation-based calibration.

For the MPC module, the weighting matrices Q, R, and P are selected to balance trajectory tracking, control effort, and terminal consistency. In all experiments, the same vehicle model, control interval, actuator limits, and road-boundary constraints are used for the proposed method and MPC-based baselines. This setting ensures that the comparison focuses on the effect of memory-conditioned dynamic potential field reconstruction rather than differences in vehicle dynamics or actuation constraints.

Key implementation settings of the proposed M3UDMC framework.

Item	Setting
High-level reasoning frequency	1 Hz
Low-level control frequency	10 Hz
Maximum system update frequency	15 Hz
Input synchronization window	50 ms
Safety override threshold	0.7
Onboard computing platform	NVIDIA Orin AGX
State weight Q	diag(2.0,2.0,1.0,0.5)
Control weight R	diag(0.2,0.5)
Terminal weight P	diag(5.0,5.0,2.0,1.0)
Control variation weight ρ	0.5
Risk-cost weight β	8.0
Base repulsion coefficient β_0	2.0
Base attraction coefficient α_0	0.8
Critical TTC threshold	2.5 s
Control interval Δt	0.1s
Language model backbone	TinyLlama-1.1B

References

ApolloAuto, 2023. Apollo 8.0 [Computer Software]. GitHub [Accessed: April 9, 2026]. https://github.com/ApolloAuto/apollo.

Arnold, E., Dianati, M., de Temple, R., Fallah, S., 2022. Cooperative perception for 3D object detection in driving scenarios using infrastructure sensors. IEEE Trans. Intell. Transp. Syst. 23 (3), 1852–1864. doi:10.1109/TITS.2020.3028424.

Besold, T.R., d'Avila Garcez, A.S., Bader, S., Bowman, H., Domingos, P., Hitzler, P., Kühnberger, K.-U., Lamb, L.C., Lowd, D., Lima, P.M.V., de Penning, L., Pinkas, G., Poon, H., Zaverucha, G. (2017). Neural-symbolic learning and reasoning: a survey and interpretation. arXiv Preprint arXiv:1711.03902. https://arxiv.org/abs/1711.03902.

Bijelic, M., Gruber, T., Mannan, F., Kraus, F., Ritter, W., Dietmayer, K., Heide, F., 2020. Seeing through fog without seeing fog: deep multimodal sensor fusion in unseen adverse weather. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 11679–11689. doi:10.1109/CVPR42600.2020.01170.

Bishop, C.M., 1994. Mixture Density Networks (Technical Report NCRG/94/004). Neural Computing Research Group, Aston University.

Bochkovskiy, A., Wang, C.-Y., Liao, H.-Y.M. (2020). YOLOv4: optimal speed and accuracy of object detection. arXiv Preprint arXiv:2004.10934. https://arxiv.org/abs/2004.10934.

Bommasani, R., Hudson, D.A., Adeli, E., Altman, R., Arora, S., von Arx, S., Bernstein, M.S., Bohg, J., Bosselut, A., Brunskill, E., Brynjolfsson, E., Buch, S., Card, D., Castellon, R., Chatterji, N., Chen, A., Creel, K., Davis, J.Q., Demszky, D., Donahue, C., Doumbouya, M., Durmus, E., Ermon, S., Etchemendy, J., Ethayarajh, K., Fei-Fei, L., Finn, C., Gale, T., Gillespie, L.E., Goel, K., Goodman, N., Grossman, S., Guha, N., Hashimoto, T., Henderson, P., Hewitt, J., Ho, D.E., Hong, J., Hsu, K., Huang, J., Icard, T., Jain, S., Kalluri, P., Karamcheti, S., Keeling, G., Khani, F., Koh, P.W., Krass, M., Krishna, R., Kuditipudi, R., Kumar, A., Ladhak, F., Lee, M., Lee, T., Leskovec, J., Levent, I., Li, X.L., Ma, T., Malik, A., Manning, C.D., Mirchandani, S., Mitchell, E., Munyikwa, Z., Nair, S., Narayan, A., Newman, B., Nie, A.,

Niebles, J.C., Nilforoshan, H., Nyarko, J., Ogut, G., Orr, L., Papadimitriou, I., Park, J.S., Piech, C., Portelance, E., Potts, C., Raghunathan, A., Reich, R., Ren, H., Rong, F., Roohani, Y.H., Ruiz, C., Ryan, J., Ré, C., Sadigh, D., Sagawa, S., Santhanam, K., Shih, A., Srinivasan, K.P., Tamkin, A., Taori, R., Thomas, A.W., Tramèr, F., Wang, R.E., Wang, W., Wu, B., Wu, J., Wu, Y., Xie, S.M., Yasunaga, M., You, J., Zaharia, M., Zhang, M., Zhang, T., Zhang, X., Zhang, Y., Zheng, L., Zhou, K., Liang, P., 2021. On the opportunities and risks of foundation models. arXiv Preprint. https://arxiv.org/abs/2108.07258.

Caesar, H., Bankiti, V., Lang, A.H., Vora, S., Liong, V.E., Xu, Q., Krishnan, A., Pan, Y., Baldan, G., Beijbom, O., 2020. nuScenes: a multimodal dataset for autonomous driving. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 11621–11631.

Chen, J., Li, S.E., Tomizuka, M. (2020). Interpretable end-to-end urban driving with latent deep reinforcement learning. arXiv Preprint arXiv:2001.08726. https://arxiv.org/abs/2001.08726.

Cui, C., et al., "On-board vision-language models for personalized autonomous vehicle motion control: system design and real-world validation," Nov. 17, 2024, ArXiv: arXiv:2411.11913. doi: 10.48550/arXiv.2411.11913.

Feng, D., Haase-Schütz, C., Rosenbaum, L., Hertlein, H., Gläser, C., Timm, F., Wiesbeck, W., Dietmayer, K., 2020. Deep multi-modal object detection and semantic segmentation for autonomous driving: datasets, methods, and challenges. IEEE Trans. Intell. Transp. Syst. 22 (3), 1341–1360. doi:10.1109/TITS.2020.2972974.

Galceran, E., Eustice, R.M., Olson, E., 2015. Toward integrated motion planning and control using potential fields and torque-based steering actuation for autonomous driving. In: Proceedings of the 2015 IEEE Intelligent Vehicles Symposium. IEEE, pp. 304–309.

Geiger, A., Lenz, P., Urtasun, R., 2012. Are we ready for autonomous driving? The KITTI vision benchmark suite. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. IEEE, pp. 3354–3361. doi:10.1109/CVPR.2012.6248074.

Hu, Z., Xu, M., Cheng, Q., 2025. Multimodal Large-language model empowering next-generation autonomous driving systems. J. Intell. Connected Vehicles 8 (2), 9210059-1–9210059-3. doi:10.26599/JICV.2025.9210059.

Kahneman, D., 2011. Thinking, fast and slow. Farrar, Straus and Giroux.

Kato, S., Tokunaga, S., Maruyama, Y., Maeda, S., Hirabayashi, M., Kitsukawa, Y., Monrroy, A., Ando, T., Fujii, Y., Azumi, T., 2018. Autoware on board: enabling autonomous vehicles with embedded systems. In: Proceedings of the 9th ACM/IEEE International Conference on Cyber-Physical Systems. IEEE, pp. 287–296. doi:10.1109/ICCPS.2018.00035.

Kendall, A., Hawke, J., Janz, D., Mazur, P., Reda, D., Allen, J.-M., Lam, V.-D., Bewley, A., Shah, A., 2019. Learning to drive in a day. In: Proceedings of the IEEE International Conference on Robotics and Automation. IEEE, pp. 8248–8254. doi:10.1109/ICRA.2019.8793742.

Khatib, O., 1986. Real-time obstacle avoidance for manipulators and mobile robots. Int. J. Robot. Res. 5 (1), 90–98. doi:10.1177/027836498600500106.

Kuang, S., Liu, Y., Wang, X., Wu, X., Wei, Y., 2024. Harnessing multimodal large language models for traffic knowledge graph generation and decision-making. Commun. Transport. Res. 4, 100146. doi:10.1016/j.commtr.2024.100146.

Liao, H., et al., 2024. GPT-4 enhanced multimodal grounding for autonomous driving: leveraging cross-modal attention with large language models. Commun. Transport. Res. 4, 100116. doi:10.1016/j.commtr.2023.100116.

Liu, Y., Zhang, J., Fang, L., Jiang, Q., Zhou, B., 2021. Multimodal motion prediction with stacked transformers. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 7577–7586.

Malkov, Y.A., Yashunin, D.A., 2020. Efficient and robust approximate nearest neighbor search using hierarchical navigable small world graphs. IEEE Trans. Pattern Anal. Machine Intell. 42 (4), 824–836. doi:10.1109/TPAMI.2018.2889473.

Mayne, D.Q., Rawlings, J.B., Rao, C.V., Scokaert, P.O.M., 2000. Constrained model predictive control: stability and optimality. Automatica 36 (6), 789–814. doi:10.1016/S0005-1098(99)00214-9.

Nayakanti, N., Al-Rfou, R., Zhou, A., Goel, K., Refaat, K.S., Sapp, B., 2023. Wayformer: motion forecasting via simple and efficient attention networks. In: Proceedings of the IEEE International Conference on Robotics and Automation. IEEE, pp. 2980–2987.

Parisi, G.I., Kemker, R., Part, J.L., Kanan, C., Wermter, S., 2019. Continual lifelong learning with neural networks: a review. Neural Networks 113, 54–71. doi:10.1016/j.neunet.2019.01.012.

Rudenko, A., Palmieri, L., Herman, M., Kitani, K.M., Gavrila, D.M., Arras, K.O., 2020. Human motion trajectory prediction: a survey. Int. J. Robotics Res. 39 (8), 895–935. doi:10.1177/0278364920917446.

Schimpe, A., Diermeyer, F. (2020). Steer with me: a predictive, potential field-based control approach for semi-autonomous, teleoperated road vehicles. arXiv preprint arXiv:2006.15718. https://arxiv.org/abs/2006.15718.

Shao, H., Hu, Y., Wang, L., Song, G., Waslander, S.L., Liu, Y., Li, H., 2024. LMDrive: closed-loop end-to-end driving with large language models. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 15120–15130.

Sun, P., Kretzschmar, H., Dotiwalla, X., Chouard, A., Patnaik, V., Tsui, P., Guo, J., Zhou, Y., Chai, Y., Caine, B., Vasudevan, V., Han, W., Ngiam, J., Zhao, H., Timofeev, A., Krivokon, M., Gao, A., Joshi, A., Zhang, Y., Shlens, J., Chen, Z., Anguelov, D., 2020. Scalability in perception for autonomous driving: waymo Open Dataset. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 2446–2454.

Tian, X., Gu, J., Li, B., Liu, Y., Wang, Y., Zhao, Z., Zhan, K., Jia, P., Lang, X., Zhao, H. (2024). DriveVLM: the convergence of autonomous driving and large vision-language models. arXiv Preprint arXiv:2402.12289.

Yu, S., Hirche, S., Huang, J., 2021. Model predictive control for autonomous vehicles: a survey. IEEE Trans. Intell. Vehicles 6 (4), 677–692.

Yuan, Y., Weng, X., Ou, Y., Kitani, K.M., 2021. AgentFormer: agent-aware transformers for socio-temporal multi-agent forecasting. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 9813–9823.

Yuan, K., et al., 2024. Evolutionary decision-making and planning for autonomous driving based on safe and rational exploration and exploitation. Engineering 33, 108–120. doi:10.1016/j.eng.2023.03.018.

Zhang, H., Xiao, L., Cao, X., Foroosh, H., 2024a. Multiple adverse weather conditions adaptation for object detection via causal intervention. IEEE Trans. Pattern Anal. Machine Intell. 46 (3), 1742–1756. doi:10.1109/TPAMI.2022.3166765.

Zhang, P., Zeng, G., Wang, T., Lu, W. (2024). TinyLlama: an open-source small language model. arXiv preprint arXiv:2401.02385. https://arxiv.org/abs/2401.02385.

Zhang, C., Wei, B., Liu, Y., Labi, S., Feb. 2026. World model-based long-tail and scenario-specific generation for autonomous driving. J. Intell. Connected Vehicles doi:10.26599/JICV.2026.9210080.

Yanbin Liu received the B.S. degree from Fuzhou University, Fuzhou, China, in 2008 and M.S degree in Software engineering from the Zhejiang University, Hangzhou,China, in 2010.He is currently pursuing the Ph.D. degree in Tsinghua University, Beijing, China. His research focuses on intelligent transportation systems, intelligent connected vehicles, and multimodal large language models, as well as their engineering applications.

Cong Zhang received the Ph.D. degree in civil and construction department from Purdue University, West Lafayette, USA, and the M.S. degree in industrial engineering from Pennsylvania State University, State College, USA in 2021. She is currently a post-doctoral research fellow of the Emerging Transportation Solution Lab at the School of Vehicle and Mobility in Tsinghua University, China. Her research interests include large language model in autonomous driving, human-vehicle interactions.

Shaohua Cui is currently a Post-Doctoral Researcher with the Department of Architecture and Civil Engineering, Chalmers University of Technology (CTH). He received the M.S. degree from the Systems Science Institute, School of Traffic and Transportation, Beijing Jiaotong University in 2019, and the Ph.D. degree in transportation engineering from Beihang University in 2023. His research interests include traffic flow analysis, vehicle control, adaptive control, and robust control.

Guangyu Tian received the B.S. degree and Ph.D. degree from Tsinghua University, Beijing, China, in 1986 and 1995. He is currently a Professor with the School of Vehicle and Mobility, Tsinghua University. His main interests include Electric vehicle integration and control.

Yugong Luo (Member, IEEE) received the B.S. and M.S. degrees from Chongqing University, Chongqing, China, in 1996 and 1999, respectively, and the Ph.D. degree from Tsinghua University, Beijing, China, in 2003. He is currently a Professor with the School of Vehicle and Mobility, Tsinghua University. His research interests include intelligent connected electric vehicle dynamics and control, and vehicle noise control.

Lei Zhang serves as Vice President of Alibaba Cloud and Chief Scientist of Industrial AI; he also holds a joint professorship at Tsinghua University for national-level high-level talents. His main interests include artificial intelligence and autonomous vehicles.