# Cooperative Decision-Making in Mixed Traffic via Conflict-Aware Heterogeneous Graph Reinforcement Learning

Guohong Zheng[a,b], Yiyang Chen[a,b], Zhigang Wu[a,b], Zhaocheng He[a,b,c] and Haipeng Zeng[a,b,*]

[a]School of Intelligent Systems Engineering, Sun Yat-Sen University, Shenzhen, 518107, China
[b]Guangdong Provincial Key Laboratory of Intelligent Transportation Systems, Shenzhen, 518107, Guangdong, China
[c]Pengcheng Laboratory, Shenzhen, Guangdong, China

ARTICLE INFO

*Keywords*:
Mixed Traffic Flow
Cooperative Decision-Making
Heterogeneous Graph Reinforcement Learning

ABSTRACT

Intersections are critical nodes of urban road networks where heterogeneous traffic flows converge. Yet, they are also major bottlenecks of efficiency and safety, particularly under mixed autonomy where Connected and Automated Vehicles (CAVs) and Human-Driven Vehicles (HDVs) coexist. Traditional control schemes such as traffic lights or optimization-based scheduling often fail to capture the complex and uncertain interactions in such settings. We propose a Conflict-Aware Heterogeneous Graph-based Reinforcement Learning (HGRL) Framework for decentralized intersection management under mixed traffic flows. In our approach, the traffic environment is represented as a heterogeneous interaction graph, where edges encode both cooperative relations and potential conflicts. Building on this representation, a heterogeneous graph-based reinforcement learning controller enables CAVs to make adaptive and coordinated decisions while explicitly accounting for conflict risks. Comprehensive simulations across varying CAV penetration rates demonstrate the effectiveness, robustness, and scalability of the proposed framework. Our method achieves consistent improvements across all key performance indicators compared with strong baselines.

## 1. Introduction

With the acceleration of urbanization and the rapid growth in vehicle ownership, traffic congestion and accidents have emerged as pressing challenges for urban traffic management. According to INRIX (2020) (Pishue, 2021), in New York City, drivers lost 100 hours in 2020, costing 1,486 per driver and 7.7 billion citywide. PEP (Transport, Health, and Environment Pan-European Programme) reported that transport-related congestion, crashes, and environmental impacts cost Europe nearly €820 million in 2019. Intersections, as critical nodes in urban road networks, concentrate large traffic volumes and complex vehicular maneuvers, thereby becoming hotspots for conflicts and accidents. Data from the U.S. Federal Highway Administration indicate that over 2.8 million intersection-related crashes occur annually, accounting for 44% of all traffic accidents (Li et al., 2023b). Consequently, enhancing both the efficiency and safety of traffic flows at intersections has become a central issue in modern traffic management research.

In response to these challenges, national policies increasingly emphasize the deployment of Intelligent Transportation Systems (ITS) and Connected and Automated Vehicles (CAVs). Through vehicle-to-infrastructure (V2I) and vehicle-to-vehicle (V2V) communication, CAV technologies offer the potential to increase throughput, reduce collision risks, improve fuel efficiency, and lower emissions (Di and Shi, 2021). For instance, since 2016 the U.S. National Highway Traffic Safety Administration has required all new vehicles to be equipped with internet connectivity (National Highway Traffic Safety Administration, 2016). Recent policy documents released by the State Council (The State Council, 2021) highlight the strategic importance the Chinese government places on fostering the development of CAVs. Similarly, the U.S. ITS Strategic Plan (2020–2025) underscores the importance of intelligent and connected traffic control technologies, positioning the nation at the forefront of global ITS innovation.

Despite policy advances and technological innovation, effective control strategies for intersections with mixed traffic flows remain underdeveloped. Real-world intersections are characterized by high traffic density, heterogeneous vehicle compositions, and frequent conflict points. In mixed traffic scenarios, where CAVs and human-driven vehicles (HDVs) coexist, variations in human driving behaviors, together with the limited penetration of CAVs undermine the

*Haipeng Zeng
zhenggh8@mail2.sysu.edu.cn (G. Zheng); chenyy553@mail2.sysu.edu.cn (Y. Chen); wuzhg6@mail2.sysu.edu.cn (Z. Wu); hezhch@mail.sysu.edu.cn (Z. He); zenghp5@mail.sysu.edu.cn (H. Zeng)
ORCID(s): 0000-0002-0339-0361 (H. Zeng)

effectiveness of traditional traffic control strategies. Prior approaches—including reservation-based systems (Huang et al., 2023), predictive scheduling (Pan et al., 2022), and machine learning-based adaptive controllers (Zhang et al., 2023) have shown promise in efficiency improvement, yet most are designed for fully connected environments or rely on overly simplified assumptions, thereby failing to capture the complex conflict dynamics inherent in mixed autonomy traffic.

Against this backdrop, explicitly addressing conflict modeling at intersections has become essential. Recent studies have begun to integrate conflict analysis into control frameworks, recognizing that safety and efficiency must be considered jointly (Zhou et al., 2024). However, existing methods still face limitations in accurately capturing heterogeneous interactions.

Motivated by these challenges, this paper proposes a novel intersection management framework that explicitly accounts for conflict relationships in mixed traffic environments. Specifically, we develop a heterogeneous graph-based reinforcement learning approach that models the heterogeneous interactions between CAVs and HDVs at intersections. In the proposed framework, vehicles are represented as nodes in a dynamic interaction graph, where edges encode both cooperative relations and potential conflicts. This graph representation is then used to guide a MARL controller, enabling vehicles to make adaptive, coordinated decisions. By combining graph-based state representation with reinforcement learning, the framework improves intersection efficiency while mitigating safety risks in conflict-prone environments.

The contributions of this paper are threefold:

- **Modeling framework**: We construct a conflict-aware heterogeneous graph representation of mixed traffic intersections, capable of capturing both cooperative and conflict interactions among heterogeneous vehicles.
- **Control strategy**: We develop a heterogeneous graph reinforcement learning (HGRL) framework, in which a heterogeneous graph neural network (HGNN) encodes interaction features via heterogeneous graph attention network (H-GAT) and sparse attention mechanisms, while a deep reinforcement learning (DRL) model learns adaptive policies under mixed traffic conditions.
- **Experimental validation**: We conduct simulations under varying CAV penetration rates and show that our framework outperforms benchmark methods in traffic efficiency and safety. We further perform ablation studies to verify the effectiveness of the proposed design.

## 2. Related Work

### 2.1. Representation Learning for Traffic Intersections

As critical nodes where traffic flows converge and diverge, intersections are highly susceptible to congestion and accidents. Statistics indicate that over one-third of traffic congestion and more than half of traffic accidents occur at intersections (Zainuddin et al., 2018; Namazi et al., 2019; Yu et al., 2013). Therefore, effective intersection management requires a precise quantification of their traffic states, which is the foundation for alleviating congestion and reducing safety risks. Consequently, extracting intersection-level features is essential for applications such as traffic network partitioning (Yu et al., 2021), key node identification (Jayaweera et al., 2017), speed prediction (Li and Lasenby, 2021), and signal control (Wei et al., 2019).

Traditional approaches generally transform road networks into graph structures and analyze them using centrality or connectivity measures (Porta et al., 2006; Zhang et al., 2011, 2017; Yue and Yeh, 2008). While useful, these methods capture only shallow structural information and fail to uncover complex interdependencies between intersections. More recently, network representation learning (NRL) has been introduced to extract latent features while preserving topological and attribute information automatically (Zhang et al., 2018; Hamilton et al., 2017). Classical methods such as DeepWalk and Node2vec (Perozzi et al., 2014; Grover and Leskovec, 2016) learn node embeddings via random walks, whereas Struc2vec (Ribeiro et al., 2017) focuses on structural similarity. However, these approaches are limited in modeling nonlinear or dynamic dependencies. Graph neural networks (GNNs), including GCNs and GATs, have since emerged to jointly leverage node attributes and neighborhood structures (Pareja et al., 2020; Sankar et al., 2020).

In the transportation domain, GNN-based models have shown promise in encoding vehicle interactions under uncertainty. For example, (Klimke et al., 2022) applied GNNs to cooperative autonomous driving, while (Liu et al., 2024) proposed a spatiotemporal dynamic graph framework for ramp merging. Similarly, (Chen et al., 2022) modeled trajectory conflicts via graph-based scheduling methods. These studies highlight the potential of graph-based learning but also reveal the need for more effective representations of conflict relationships in mixed traffic flows.

In this context, our work develops a conflict-aware graph representation of intersections that explicitly encodes both cooperative and conflict vehicle interactions. This representation serves as the foundation of our framework and ensures that key conflict structures—largely overlooked in prior studies—are directly incorporated into the learning process.

### 2.2. Intersection Management Methods

Coordinating vehicle movements at intersections has traditionally relied on hierarchical control frameworks, where a central controller schedules arrivals and distributed controllers execute them (Xu et al., 2017). Various methods, such as fuzzy logic, model predictive control, and optimal control (Huang et al., 2023; Du et al., 2018; Malikopoulos et al., 2018), have been explored. Reservation-based strategies, such as FIFO (Dresner and Stone, 2004) and batch-based release (Tachet et al., 2016), attempt to avoid collisions but rarely achieve global efficiency. Optimization-based methods (e.g., integer programming, dynamic programming) can yield optimal schedules, yet the solution space grows exponentially with traffic volume, limiting scalability (Pan et al., 2022; Yao et al., 2023).

Reinforcement learning (RL) has therefore been widely adopted as it circumvents explicit modeling of the scheduling problem. Deep RL further addresses the curse of dimensionality and has shown strong performance in mixed traffic environments (Li et al., 2023a; Zhao et al., 2024). However, most RL approaches neglect detailed vehicle interactions, leading to limited cooperative behavior. Recent advances in graph reinforcement learning (GRL) attempt to overcome this by embedding vehicle interactions into graph structures, enabling more effective policy learning (Shen et al., 2021; Munikoti et al., 2023).

Despite these advances, few studies directly address trajectory planning for CAVs in mixed intersections. Existing works often treat vehicles individually (Zhang and Cassandras, 2019; Chen et al., 2023), or optimize mixed formations (Jiang et al., 2022), but they largely ignore the explicit conflict structures that drive congestion and safety risks.

To address these limitations, we propose a heterogeneous graph-based reinforcement learning strategy that leverages the structural properties of the conflict-aware interaction graph. By integrating cooperative and conflict relations into the policy-learning process, our approach achieves safe and efficient coordination between CAVs and HDVs.

## 3. Problem Statement

### 3.1. Intersection traffic

This study considers a prototypical four-leg unsignalized intersection operating under mixed traffic conditions, as illustrated in Fig. 1 (a). Each approach (North, East, South, West) is divided into three dedicated lanes corresponding to left-turn (L), through (C), and right-turn (R) movements. For example, a northbound left-turn movement is denoted as N-L, a southbound through movement as S-C. Traffic demand is modeled stochastically: vehicle arrivals on each lane follow random processes, and the type of each arriving vehicle is also probabilistic, comprising either CAVs or HDVs. This configuration captures the heterogeneity and randomness inherent in real-world urban intersections.

In this study, conflicts are operationalized as potential interaction risks, identified when the projected trajectories of two vehicle movements overlap within the conflict area of the intersection. For instance, vehicles traveling on N-S and S-L movements will cross paths in the center of the intersection, thereby forming a conflict pair. Table. 1 lists, for each movement direction, the set of other directions with which it conflicts. Conflicts among vehicle movements are a primary source of congestion and accidents at intersections, as they originate from contradictory action decisions between streams. While reward-based penalties discourage unsafe decisions, conflicts may still arise when reinforcement learning vehicles (RVs) attempt to enter the intersection under competing traffic streams. To maintain both safety and flow stability, we employ a conflict-resolution mechanism that post-processes RL outputs before execution. Specifically, an RV may proceed only if the intersection is clear of vehicles from conflicting directions; otherwise, its Go decision is overridden to Stop. When multiple RVs from conflicting streams simultaneously choose Go, priority is given to the stream with the longest waiting time, allowing those vehicles to pass first while others yield.

The following assumptions are made regarding vehicle states and driving behaviors. (1) Full state observability and information sharing. We assume that the positions, velocities, accelerations, and related kinematic attributes of all vehicles are fully observable and shareable across the network. Although HDVs lack onboard sensing and communication capabilities, roadside detection infrastructure (e.g., loop detectors or LiDAR units) is assumed to capture their instantaneous states and broadcast this information to nearby CAVs. (2) HDV driving behavior model. The longitudinal and lateral movements of HDVs are governed by human driving characteristics. Specifically, their

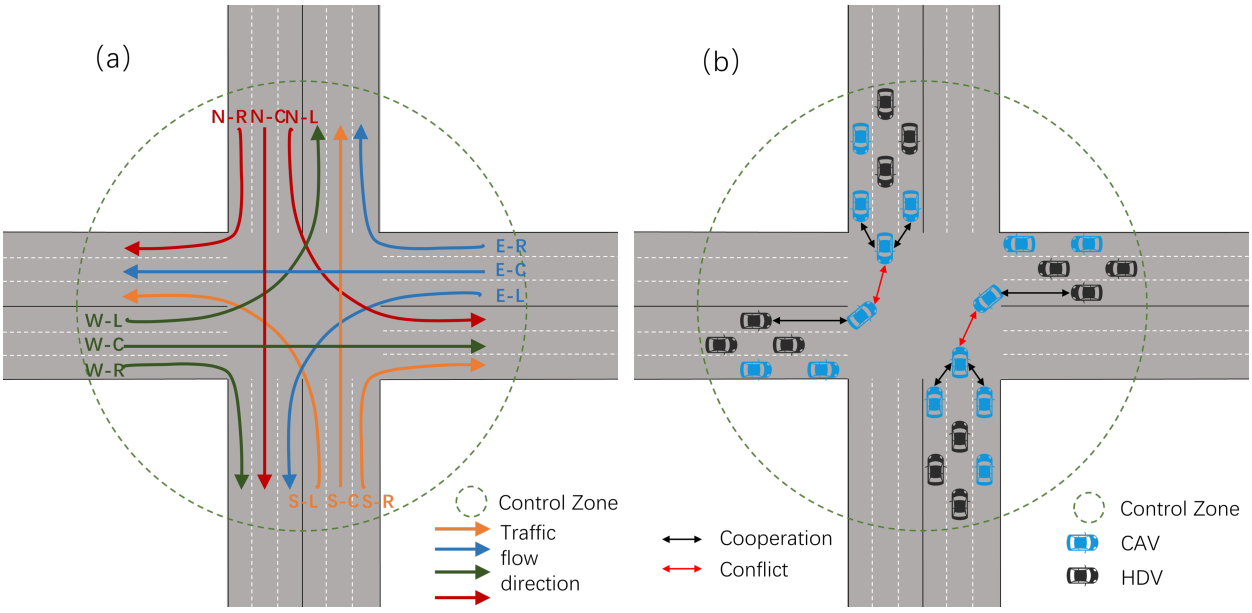


**Figure 1:** Illustration of the intersection environment and conflict determination. (a) Lane-level movement definitions for each approach, including left-turn (L), through (C), and right-turn (R) movements. (b) Mixed traffic scenario with CAVs (blue) and HDVs (black).

**Table 1**
Conflict set $C$: each movement and its conflicting directions (right-turns omitted as non-conflicting).

| Movement | Conflicting directions |
|---|---|
| N-L | {E-S, E-L, S-S, W-L} |
| N-S | {E-S, W-L, S-L, W-S} |
| E-L | {E-S, N-L, S-S, W-S} |
| E-S | {W-L, N-S, N-L, S-S} |
| S-L | {N-S, E-L, W-S, W-L} |
| S-S | {N-L, E-S, E-L, W-S} |
| W-L | {S-L, N-S, E-S, N-L} |
| W-S | {N-S, E-L, S-L, S-S} |

trajectories and car-following behaviors are modeled using the Intelligent Driver Model (IDM) (Treiber et al., 2000), which is widely adopted in traffic simulation.

## 3.2. Decentralized RL for mixed traffic

We formulate the mixed-traffic control problem as a partially observable Markov decision process (POMDP), defined by the following 7-tuple:

$$(S, A, O, T, R, \Omega, \gamma) \tag{1}$$

where $S$ is the set of underlying states, $A$ the set of actions, $O$ the set of observations, $T$ the state-transition function, $R$ the reward function, $\Omega$ the observation function, and $\gamma$ the discount factor.

### *3.2.1. Action space*

Since this study focuses on improving intersection throughput, the control objective is to manage CAVs traversing an unsignalized intersection under mixed traffic conditions. Accordingly, the action space of each CAV is restricted to $A = \{Stop, Go\}$. The selected action determines whether the CAV enters the intersection or remains stopped at the stop line.

In simulation, longitudinal acceleration is computed using the IDM. Outside the control zone, all vehicles follow standard IDM dynamics. Within the control zone, if the CAV chooses "Go", it accelerates at its maximum allowable rate $a_{max}$; conversely, if the action is "Stop", it decelerates according to IDM until it halts at the stop line. In the event of a potential collision, the CAV's emergency braking mechanism overrides the nominal control, applying a deceleration greater than the standard requirement to avoid a crash.

#### *3.2.2. Observation space*

To enhance generalizability across intersections with different topologies, we define the observation space by separating ego-level and global-level information.

**Ego information.** The ego state of each CAV is encoded as a 4-tuple:

$$S_{ego} = \{x, y, speed, acc\}, \tag{2}$$

where $x$ and $y$ denote the vehicle's center-point coordinates in a Cartesian reference frame; $speed$ represents the scalar magnitude of its velocity, and $acc$ represents the scalar magnitude of its acceleration.

**Global information.** The global information is characterized by two lane-level indicators for each inbound lane of the intersection: the queue length $l_i$ and the average waiting time $w_i$. The value of $l_i$ is computed as the number of vehicles lined up in lane $i$ before reaching the stop line, while $w_i$ is calculated as the average waiting time of these vehicles. Together, $\{l_i, w_i\}$ quantify the spatial and temporal levels of congestion, enabling CAVs to anticipate traffic pressure beyond their own trajectory.

The complete observation for each controlled CAV is thus:

$$O = \{S_{ego}, S_{global}\} \tag{3}$$

This formulation ensures that each agent bases its decision not only on its own kinematic state but also on the broader congestion context of the intersection. In simulation, both ego and global features are directly extracted from the microscopic states of all vehicles. In real-world applications, ego information can be obtained via onboard sensors, while global information may be estimated through V2V communication among CAVs.

#### *3.2.3. Reward function*

To balance traversal efficiency, emission, and safety, we define a composite reward function for each RL vehicle as follows, inspired by prior designs in Wang et al. (2025). But we merge jerk and fuel consumption into a unified eco-driving component, which provides a clearer trade-off between smoothness and energy efficiency.

$$r(s_t, a_t, s_{t+1}) = w_w\, r_w + w_v\, r_v + w_j\, p_j + p_c, \tag{4}$$

where $w_w, w_v, w_j$ are weighting coefficients, and $p_c$ are penalties for conflict behaviors. The terms are defined as follows:

- **Waiting-time reward** ($r_w$): Encourages vehicles to minimize excessive idling. If the vehicle chooses to proceed, a positive reward proportional to its reduction in waiting time is given; otherwise, a penalty is applied.

$$\begin{cases} -w^{t+1,j}, & \text{if } a^t = \text{Stop}; \\ w^{t+1,j}, & \text{otherwise.} \end{cases} \tag{5}$$

- **Speed reward** ($r_v$): Normalized instantaneous speed of the ego vehicle,

$$r_v = \frac{v}{v_{\max}}, \tag{6}$$

where $v$ is the current velocity and $v_{\max}$ is the maximum allowable speed. This term promotes efficient progression through the intersection.

- **Jerk penalty** ($p_j$): Penalizes abrupt acceleration changes to encourage smooth and fuel-efficient driving,

$$p_j = \frac{|a_t - a_{t-1}|}{J_{\max}}, \tag{7}$$

where $a_t$ and $a_{t-1}$ denote current and previous accelerations, respectively, and $J_{\max}$ is a saturation constant.

- **Conflict penalty** ($p_c$): A fixed penalty of $-1$ is imposed if the vehicle's action results in a potential conflict with another vehicle in the intersection control zone.

In our implementation, the weights are set to $w_w = 1.0$, $w_v = 0.3$, and $w_j = -1.0$, ensuring a balanced trade-off between efficiency, comfort, and safety. The normalized structure of each component ensures comparability in scale and enhances the stability of policy learning.

## 4. Methodology

This section provides a detailed exposition of the decision-making problem formulated using a heterogeneous graph-based reinforcement learning within mixed autonomous traffic flows, including Basic Framework, Heterogeneous Graph Representation, Heterogeneous Graph Reinforcement Learning (HGRL).

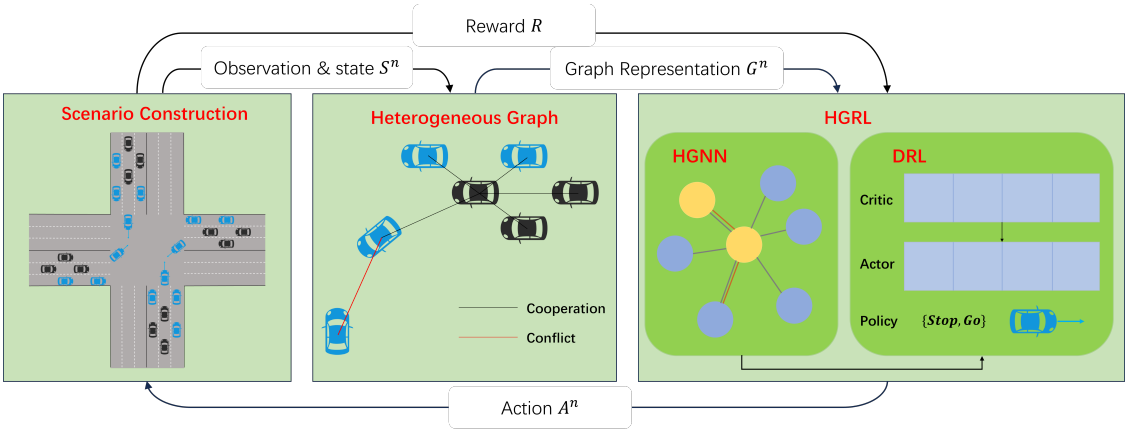


**Figure 2:** Framework of our method for mixed traffic control. It consists of three main modules: (1) Scenario Construction module, where traffic situations are generated; (2) Heterogeneous Graph module, which models cooperation and conflict relations between vehicles; (3) HGRL module, where the HGNN model encodes graph embeddings and the DRL model learns policies (Stop/Go) based on actor–critic structure.

### 4.1. Basic Framework

The basic framework of the GRL-based decision-making system for RVs is illustrated in Fig 2. The basic framework consists of three parts: the scenario construction module, the heterogeneous graph representation module, the HGRL module, including the HGNN model and the DRL model. Specifically, the scenario construction module provides the dynamic environment of mixed traffic flows, including both CAVs and HDVs, which is defined in 3.1; the heterogeneous graph representation module encodes vehicles as graph nodes and their cooperative or conflict relations as edges, thereby transforming raw traffic states into structured graph features; In the HGRL module, the HGNN model serves as the core feature extractor, learning high-dimensional embeddings that capture both local and global interaction patterns; the DRL model leverages the extracted graph embeddings to train a policy network, which outputs discrete driving decisions (Stop/Go) for each controlled CAV through an actor–critic architecture.

### 4.2. Heterogeneous Graph Representation

Building upon the intersection environment defined in Section 3.1, we represent the mixed traffic scenario as a heterogeneous undirected graph $G = \{V, E\}$, where $V = \{v_i, i \in \{1, 2, \cdots, n\}\}$ is the set of nodes, each $v_i$ encoding

the state of the $i_{th}$ vehicle; $E = \{e_{ij}, i, j \in \{1, 2, \cdots, n\}\}$ is the set of edges, where $e_{ij}$ represents the pairwise interaction between vehicle $i$ and vehicle $j$. Specifically, $n$ denotes the number of RVs in the driving loop. The construction of this graph involves three components: node feature matrix, edge attributes, and the adjacency matrix.

**Node feature matrix**. The node feature matrix encapsulates the kinematic observations of all RL vehicles in the control zone. Formally, it is defined as:

$$N = [x_1, x_2, \cdots, x_n]^T, \tag{8}$$

where each $x_i$ corresponds to the observation vector of the RL vehicle, ensuring consistency with the reinforcement learning formulation described in formulation 3.

**Edge attributes**. Edges encode pairwise interactions between vehicles and consist of cooperation edges and conflict edges. The definition involves both connectivity judgment (whether an edge exists) and edge values (what information the edge carries).

- Cooperative edges. A cooperative edge is created if the Euclidean distance between vehicles $i$ and $j$ does not exceed the communication threshold $\rho_c$:

$$e_{ij}^{coop} = \begin{cases} d_{ij}, & d_{ij} \leq \rho_c, \\ 0, & d_{ij} > \rho_c, \end{cases} \tag{9}$$

  where $\rho_c$ denotes the communication range of CAVs and $d_{ij}$ the Euclidean distance between vehicle $i$ and vehicle $j$. This criterion implies that any two CAVs within the specified communication range can establish a link and exchange information, thereby enabling cooperative interaction. In our experiments, we set $\rho_c = 50m$.

- Conflict edges. A conflict edge is established if the ego observation of vehicle $i$ identifies vehicle $j$ as a potential conflict partner. Formally,

$$e_{ij}^{conf} = \begin{cases} obs_{ego}^{i}, & \text{if vehicle } i \text{ conflict with vehicle } j, \\ 0, & \text{otherwise}, \end{cases} \tag{10}$$

  In this case, the connectivity is determined by the conflict-detection mechanism, while the edge weight is directly taken from the ego observation vector of vehicle $i$, described in formulation 2.

**Adjacency matrix**. The final adjacency matrix integrates both cooperative and conflict relations:

$$D = D_{coop} + D_{conf}, \tag{11}$$

where $D_{coop} = \{e_{ij}^{coop}\}_{i,j=1}^{n}$ encodes distance-based cooperative interactions, and $D_{conf} = \{e_{ij}^{conf}\}_{i,j=1}^{n}$ encodes conflict interactions derived from ego observations. By jointly incorporating proximity-driven cooperation and safety-critical conflict detection, the adjacency matrix provides a comprehensive representation of inter-vehicle interactions, which is then processed by the HGNN model.

## 4.3. Heterogeneous Graph Reinforcement Learning (HGRL)

### *4.3.1. HGNN Model*

The proposed HGNN model is designed to extract high-level representations from the heterogeneous graph constructed in Section.4.2, where node observations and edge attributes encode both the kinematic states of vehicles and their cooperative or conflict interactions. As illustrated in Fig. 3, the HGNN integrates graph attention mechanisms, fully connected layers, and nonlinear activation functions, thereby enabling adaptive feature extraction tailored to mixed-traffic intersection control.

The heterogeneous graph can be represented as $\bar{G} = (\bar{N}, \bar{D})$, where the node feature matrix $\bar{N} \in \mathbb{R}^{N \times f}$ and the cooperative and conflict adjacency matrices $\bar{D}_{coop}, \bar{D}_{conf} \in \mathbb{R}^{N \times N}$. To capture heterogeneous relations, we adopt a heterogeneous graph attention network (H-GAT), where parallel GAT are applied to cooperative and conflict edges separately. The attention mechanism dynamically adjusts the contribution of neighboring nodes based on relational importance. Formally, the attention weight function is given as:

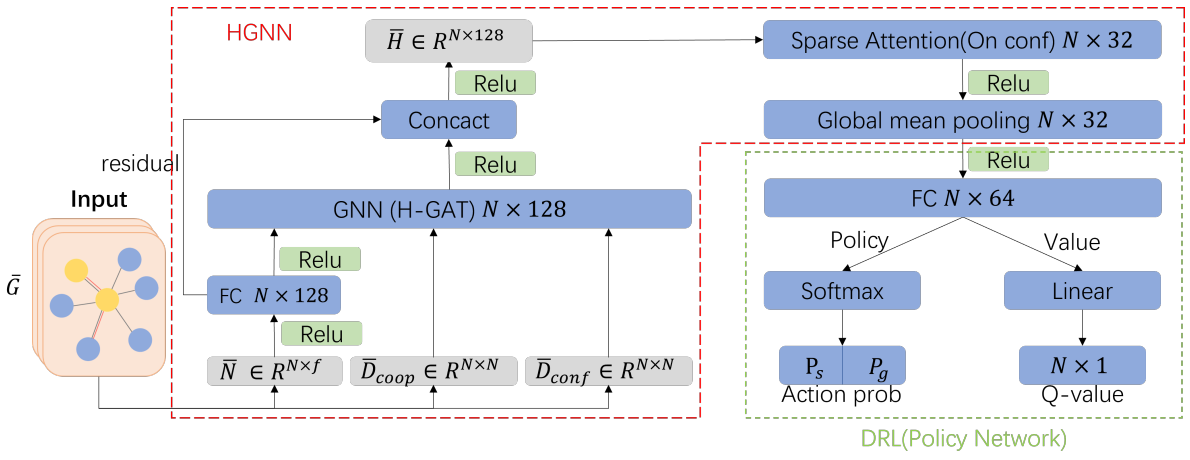


**Figure 3:** Detailed architecture of the proposed HGRL module. The HGNN model (left) captures cooperative and conflict features from the heterogeneous traffic graph, producing graph embeddings. The DRL policy network (right) takes these embeddings as input and outputs action probabilities (Stop/Go) and value estimates under an actor–critic structure.

$$Attention(e_{ij}) = a^{\top} LeakyRelu(W \cdot [h_i \| h_j]) \tag{12}$$

where $\|$ is the concatenate operator, $a^{\top}$ is the learnable attention vector, $W$ is a learnable weight matrix, $h_i$ and $h_j$ are the feature of the $i_{th}$ and $j_{th}$ node respectively. The outputs of cooperative and conflict GAT heads are then concatenated and passed through nonlinear transformations to form node-level embeddings $H \in \mathbb{R}^{N\times128}$.

To further emphasize safety-critical relations, a sparse attention mechanism is applied to conflict edges. Since cooperative edges are typically dense but low-risk, while conflict edges represent fewer yet high-impact interactions, sparsity encourages the model to filter out redundant conflicts and retain the most critical ones. This design enhances both interpretability and computational efficiency. Then, the learned embeddings are aggregated through fully connected layers and global mean pooling to produce compact graph-level representations ($N \times 32$). which serve as the input to the policy network.

### *4.3.2. DRL Model*

The graph embeddings generated by the HGNN model are integrated into a reinforcement learning framework based on Proximal Policy Optimization (PPO) (Schulman et al., 2017). PPO is an actor–critic algorithm that provides stable and sample-efficient policy updates by employing a clipping mechanism to constrain gradient steps. This makes it particularly suitable for large-scale, multi-agent traffic environments.

In this architecture, the actor network corresponds to the policy network, which maps HGNN embeddings into action probabilities ($P_s$, $P_g$) for each CAV, representing the decision to Stop or Go. The critic network estimates the value of each state, thereby guiding the policy updates with a learned baseline. To reduce variance in policy gradient estimation, we adopt Generalized Advantage Estimation (GAE).

The clipped surrogate objective for the actor is defined as:

$$L^{CLIP}(\theta) = \mathbb{E}_t[\min(r_t(\theta)\hat{A}_t, clip(r_t(\theta), 1-\epsilon, 1+\epsilon)\hat{A}_t)], \tag{13}$$

where $r_t(\theta) = \frac{\pi_\theta(a_t|s_t)}{\pi_{\theta_{old}}(a_t|s_t)}$ measures the ratio of new to old policy probabilities, $\hat{A}_t$ is the estimated advantage at timestep $t$. The critic network is trained by minimizing the mean squared error:

$$L_{critic}(\phi) = \mathbb{E}_t[(r_t + \gamma V_\phi(s_{t+1}) - V_\phi(s_t))^2], \tag{14}$$

where $r_t$ is the immediate reward and $\gamma$ is the discount factor with a range of $[0, 1]$.

As illustrated in Fig.3, the policy network consists of fully connected layers and activation functions to output both action probabilities and value estimates for each CAV. Let $\Phi_{\text{policy}}$ denote the policy operator. The network first projects the aggregated embedding $H'$ into a latent representation through a fully connected layer ($N \times 64$). This representation is then divided into two branches: a policy head, which applies a softmax to produce the action probabilities $p_s, p_g$, and a value head, which outputs the state-value $Q(s^{n,T}, a^n)$ via a linear layer:

$$Q(s^{n,T}, a^n) = \Phi_{\text{policy}}(H'). \tag{15}$$

Here $Q(s^{n,T}, a^n)$ denotes the expected return of taking action $a^n$ under observation $s^{n,T}$. This value serves as a guiding principle for determining the driving behavior of CAVs.

## 5. Experiment and results

### 5.1. Mixed Traffic

Training and evaluation of RL algorithms require a suitable simulation environment. We chose the open-source simulator Simulation of Urban Mobility (SUMO) (Behrisch et al., 2011). It includes human driving models, configurable traffic networks and flows, and mechanisms for enforcing traffic rules, safety rules, and physical constraints. The built-in collision avoidance mechanism and human driving model will act as the downstream modules of self-driving software and the human driver, respectively. These mechanisms and models will ensure collision-free driving of a vehicle. This assumption has been widely adopted by previous studies.

In this research, we conducted training in a typical unsignalized intersection measuring 30m × 30m. The entrance and exit lanes had a length of 100m. The entrance and exit lanes of the four branches within the intersection were included in the study area. To examine the intersection's performance in handling high-density traffic scenarios with frequent interactions, the traffic volume for the entrance was randomly set between 900 and 1200 vehicles per hour. To evaluate the effectiveness of CAVs in this context, we tested different CAV market penetration rates (MPRs): 50% to 100%. The type of vehicle (CAV or HDV) for each vehicle was randomly assigned. For HDV, the acceleration is computed using IDM. For CAV, when it is outside the control zone, IDM is used to determine the acceleration; when it is inside the control zone, the decisions are determined by the policy, which is computed by RL model.

### 5.2. Experiment set-up

Our evaluation metrics contain average waiting time, average travel time, and average fuel consumption of all vehicles, which comprehensively reflect both the operational efficiency of the intersection and the environmental impact of traffic control strategies. These indicators are widely adopted in the literature as they enable a balanced assessment of mobility performance and sustainability outcomes.

We evaluate our method by comparing it to four baselines: (1) TL: the default traffic signal program deployed in the intersection; (2) NoTL: no traffic lights; (3) Wang (Wang et al., 2025): the state-of-the-art RL traffic controller with 50% ~100% CAV penetration rate; and (4) Yan (Yan and Wu, 2021): the state-of-the-art RL traffic controller with 100% CAV penetration rate for unsignalized intersections.

### 5.3. Overall performance

To establish a benchmark for subsequent evaluations under varying penetration rates, we first examine the overall performance of all models in a fully cooperative setting where the penetration rate of CAVs is 100%. This scenario represents the maximum achievable performance of each method without interference from human-driven vehicles and thus provides a reliable reference for later comparisons in mixed-traffic conditions.

Table. 2 reports the average performance of five representative approaches across three widely adopted indicators: travel time, fuel consumption, and waiting time. The baselines include conventional traffic light control (TL), a no-traffic-light scenario (NoTL), and two representative learning-based methods (Wang and Yan), against which our proposed approach (Ours) is evaluated.

The comparative analysis underscores the clear superiority of our proposed approach. In terms of travel time, our model reduces delays by more than 58% compared with conventional traffic light control and by nearly 64% relative to the uncontrolled NoTL setting. Even against advanced learning-based baselines, our method still achieves reductions of approximately 13% (relative to Wang) and 29% (relative to Yan), highlighting its efficiency advantage. A similar trend is observed for fuel consumption, where our approach yields a 33% reduction compared with TL and a 30% reduction

**Table 2**
Comparison of models under averaged metrics

| Metrics (Average) | Models (100% penetration rate) | | | | |
|---|---|---|---|---|---|
| | TL | NoTL | Wang | Yan | **Ours** |
| Travel time | 148.00 | 171.63 | 70.81 | 86.44 | **61.82** |
| Fuel consumption | 1111.22 | 1041.19 | 788.07 | 1068.02 | **745.44** |
| Waiting time | 15.61 | 27.87 | 6.21 | 9.28 | **4.41** |

compared with Yan, while also slightly outperforming Wang. With respect to waiting time, our model delivers the most pronounced improvement, lowering delays by over 70% relative to TL and more than 50% relative to Yan. These results indicate that the proposed framework not only accelerates traffic flows but also reduces unnecessary idling and fuel expenditure, thus delivering simultaneous gains in efficiency, sustainability, and traffic smoothness—a combination rarely achieved by existing methods.

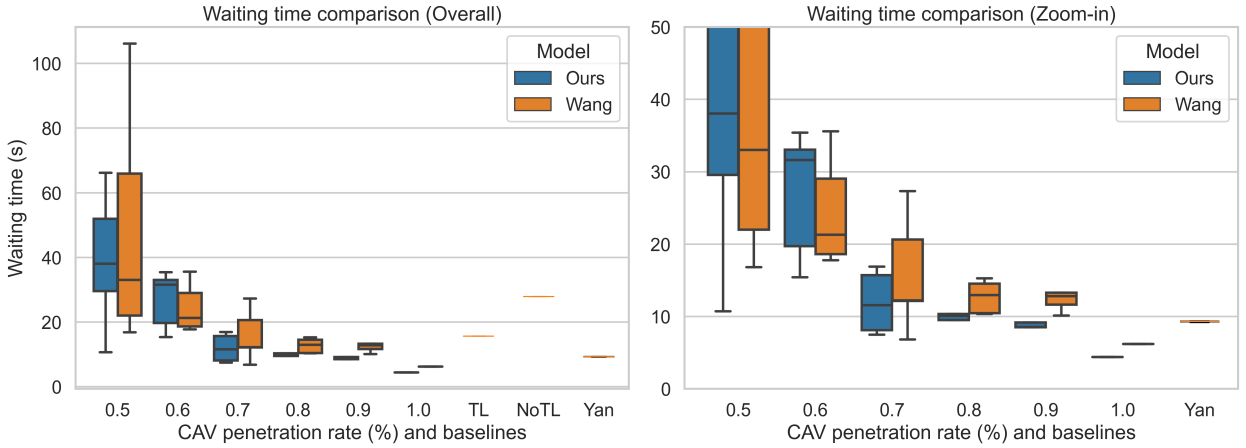


**Figure 4:** The overall results in average fuel consumption at the intersection. The RIGHT sub-figure displays a zoomed-in version of the LEFT sub-figure, excluding the Yan methods. Our method consistently outperforms the other four baselines.

### *5.3.1. Waiting time*

Waiting time measures the average duration that vehicles remain idle before passing through the intersection and serves as a direct reflection of queuing and congestion. Figure 4 reports the waiting time distributions of different methods under varying penetration rates.

The results show that waiting time decreases as the penetration rate increases and gradually converges to a stable low level. The proposed method achieves reductions ranging from 19.5% at low penetration to over 31.4% at high penetration (0.9–1.0). This trend can be explained by the decreasing proportion of HDVs, which enables cooperative CAVs to coordinate their decisions more effectively, thereby reducing hesitation and mitigating queue spillback. Our method further amplifies this effect through its graph-based modeling framework, which explicitly captures inter-vehicle interactions and conflict relations. This design allows vehicles to negotiate shared space more efficiently, minimizing idle periods. Consequently, our approach delivers not only shorter but also more stable waiting times, highlighting its robustness under mixed-traffic conditions.

### *5.3.2. Travel time*

Travel time is a fundamental indicator of traffic efficiency. Compared with delay, another widely used metric, travel time also reflects the overall mobility of vehicles at intersections. Therefore, we adopt travel time as the primary measure of efficiency. The distributions of travel time for the different models are shown in Fig. 5.

As illustrated, our method achieves consistently lower travel times across most penetration rates compared with Wang and other baselines. At penetration levels of 0.5 and 0.6, the relative advantage over Wang diminishes, with

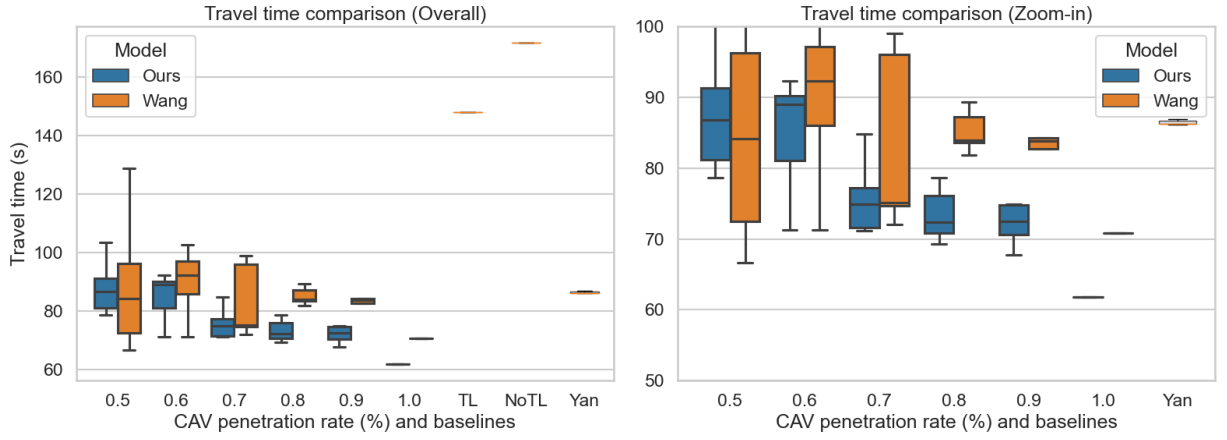


**Figure 5:** The overall results in average travel time at the intersection. The RIGHT sub-figure displays a zoomed-in version of the LEFT sub-figure, excluding the Yan methods. When the RV penetration rate reaches or exceeds 70%, our method consistently outperforms the other four baselines.

Wang occasionally achieving marginally shorter travel times. Nevertheless, the variance of our results remains significantly smaller, indicating that the proposed framework provides more stable system performance in mixed-traffic environments dominated by HDV-induced disturbances.

At medium to high penetration rates (0.7–1.0), our approach demonstrates clear superiority, yielding both lower mean travel times and reduced variability. This suggests that the explicit modeling of cooperative and conflict relations enables more efficient coordination among vehicles and mitigates congestion at conflict points. When compared with traditional baselines such as TL, NoTL, and Yan, our method consistently achieves the lowest travel times, underscoring its robustness and scalability across a wide spectrum of traffic compositions.

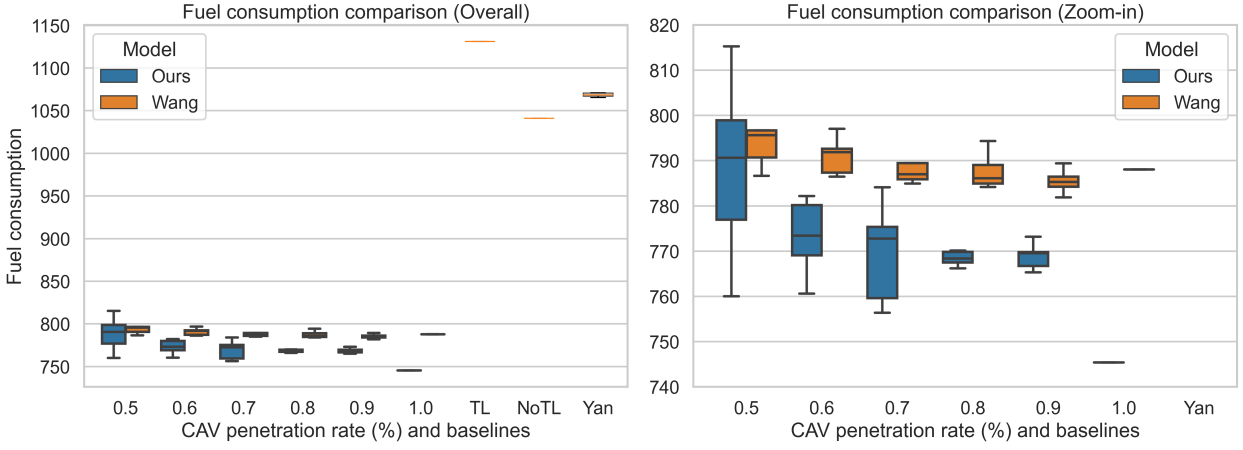


**Figure 6:** The overall results in average fuel consumption at the intersection. The RIGHT sub-figure displays a zoomed-in version of the LEFT sub-figure, excluding the Yan methods. Our method consistently outperforms the other four baselines.

### *5.3.3. Fuel consumption*

Fuel consumption is an important indicator of environmental sustainability, as it reflects both energy efficiency and emissions reduction. Fig. 6 illustrates the distribution of fuel consumption across different penetration rates for our method and the Wang baseline, with additional comparisons to TL, NoTL, and Yan.

Overall, our method achieves consistently lower fuel consumption, with reductions ranging from 0.9% at low penetration to 5.4% under full penetration compared with Wang. These improvements are less pronounced than those

**Table 3**
Ablation settings of different model variants

| Model | Graph Structure | Graph Encoder | Conflict Sparse Attention (SAT) |
|---|---|---|---|
| DRL | None | – | – |
| GNN | Homogeneous | GCN | – |
| GRL | Heterogeneous | GCN | – |
| GRL_A | Heterogeneous | H-GAT | – |
| GRL_SAT | Heterogeneous | GCN | ✓ |
| Ours | Heterogeneous | H-GAT | ✓ |

observed for travel time, primarily because fuel consumption is influenced not only by cooperation but also by idling and acceleration patterns, which remain affected by HDVs in mixed traffic.

The relatively modest gains highlight that while higher penetration improves coordination and reduces stop-and-go behavior, inefficiencies caused by HDV variability cannot be completely eliminated. Nonetheless, our model delivers noticeable advantages due to the jerk penalty in the reward function, which discourages abrupt acceleration and deceleration. By promoting smoother driving trajectories and reducing idling in queues, our method successfully lowers unnecessary fuel use, achieving superior performance across all penetration rates compared with Wang and other baselines.

## 5.4. Ablation study

To evaluate the contribution of each component in our framework, we design a comprehensive set of ablation experiments, as summarized in Table. 3. The ablated models differ primarily in three aspects: the use of graph structure, the choice of graph encoder, and the inclusion of the conflict-aware sparse attention mechanism. Each variant can thus be regarded as a different version of the proposed HGNN model illustrated in Fig. 3.

Several points about the adopted models should be emphasized as follows:

- The DRL model and GNN model are selected as the baseline models. The DRL model does not employ any graph structure; instead, it aggregates raw observations from the environment as input for decision-making. The GNN model, in contrast, replaces the heterogeneous graph with a homogeneous one, without explicitly modeling conflict relations.
- The GRL model replaces the graph neural network in our model with a standard GCN, to examine the effectiveness of the H-GAT encoder under the heterogeneous graph setting.
- If the model does not incorporate the sparse attention mechanism, which highlights conflict edges, it is replaced with a fully connected layer.

### *5.4.1. Analysis of results*

The training reward curve is plotted in Fig. 7. Moreover, the evaluation metrics are averaged as the final training results when the model converges, as shown in Table. 4.

The DRL and GNN models serve as preliminary baselines to evaluate the effect of incorporating graph structure. Both models lag behind the heterogeneous variants in terms of convergence speed and final reward plateau. The DRL model, which directly aggregates raw observations without any graph representation, converges slowly and remains less stable throughout training. The GNN model, which employs a homogeneous graph, performs slightly better than DRL in terms of training dynamics; however, it still fails to capture heterogeneous relations, such as cooperation and conflict. As reflected in the evaluation result, although its travel time and fuel consumption are not markedly worse than those of GRL, it suffers from the highest waiting time among all models. These results indicate that the explicit construction of heterogeneous graphs is essential for effectively modeling interaction dynamics in mixed-autonomy traffic.

The GRL models exhibits notable deficiencies in stability and efficiency. Although the mean episode reward increases progressively with training, the curve remains characterized by substantial oscillations and fails to converge

**Table 4**
The Evaluation Result of Ablation Study

| Metrics (Average) | Models | | | | | |
|---|---|---|---|---|---|---|
| | DRL | GNN | GRL | GRL_A | GRL_SAT | **Ours** |
| Travel time | 69.26 | 65.19 | 67.57 | 67.92 | 62.38 | **61.82** |
| Fuel consumption | 787.97 | 788.86 | 791.94 | 790.57 | 779.47 | **745.44** |
| Waiting time | 7.94 | 8.79 | 5.64 | 6.98 | 5.79 | **4.41** |

within a narrow interval. This pattern suggests that a conventional graph convolutional encoder lacks the representational capacity required to adequately capture the intricate interaction structures present in mixed-autonomy traffic, thereby constraining both robustness and decision reliability.

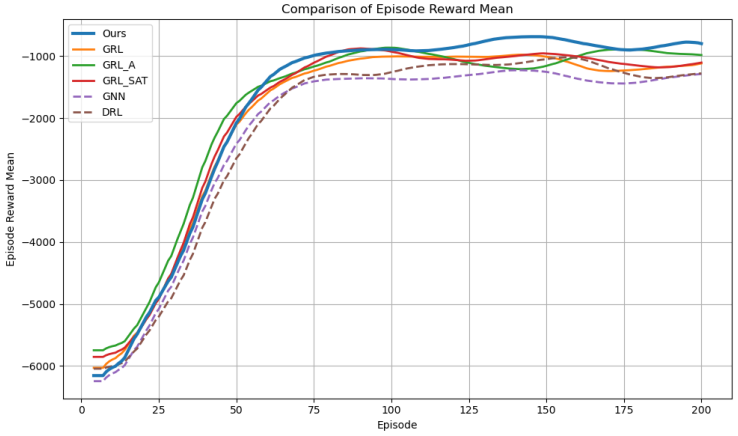


**Figure 7:** The training reward curve of the implementation models in the ablation study.

Substituting the GNN with a graph attention network (GRL_A) yields a steeper rise of the reward curve during the initial training phase and accelerates convergence. This indicates that attention mechanisms over nodes and their incident edges enable the policy to prioritize salient local interactions and exploit structural cues more effectively. However, because the attention weights are distributed densely across both critical and non-critical neighbors, the influence of genuinely decisive relations is diluted. As a consequence, fluctuations re-emerge in the mid-to-late stages of training, and the efficiency improvements do not consistently extend to safety or stability.

In contrast, the GRL_SAT variant introduces a sparse attention mechanism restricted exclusively to conflict edges. By concentrating model capacity on those interactions that determine right-of-way and collision risk, this approach attenuates the influence of irrelevant neighbors and enhances discriminative power. Empirically, the corresponding reward curve exhibits smoother dynamics and convergence to a narrower plateau. The model tends to adopt more conservative actions at potential conflict points, thereby sacrificing a degree of instantaneous aggressiveness in exchange for greater reliability and stability of traffic flow.

Finally, the complete model (Ours) integrates the attentive graph encoder with conflict-aware sparse attention in a unified framework. The reward curve rises rapidly and stabilizes at a higher plateau than all ablated configurations, reflecting both faster learning dynamics and improved convergence properties. Consistent with these observations, the model achieves the most favorable trade-off across travel time, fuel consumption, and waiting time. Collectively, these findings demonstrate that while attention is essential for surpassing the limitations of basic GNNs, it is the conflict-aware sparsification of attention that translates representational gains into robust, safe, and efficient decision-making. The synergy of GAT and conflict-edge SAT thus constitutes the critical factor enabling the proposed framework to deliver state-of-the-art performance in mixed-traffic environments.

## 6. Discussion

This paper proposed a conflict-aware HGRL framework, tailored for mixed traffic flows of CAVs and HDVs. By explicitly modeling both cooperative and conflict relationships among vehicles, the framework captures heterogeneous traffic interactions that are often overlooked by conventional graph-based or rule-based approaches. The integration of sparse conflict-aware attention enables the model to extract safety-critical patterns. The HGNN representation is further incorporated into a reinforcement learning–based policy network, which allows adaptive decision-making under varying penetration rates of CAVs.

As the evaluation results show, our method can significantly improve the systemic utility. By designing from the perspective of individual CAVs, the algorithm not only ensures the operational efficiency of each agent but also enhances the performance of the entire intersection. This suggests that when agents can access global information, they can better coordinate with other CAVs and HDVs, allowing for more informed decisions before entering conflict zones and avoiding indecisive or inefficient behaviors.

Our method can also be applied to diverse intersection scenarios, including non-orthogonal layouts and signalized intersections:

Firstly, this adaptability stems from the fact that the model does not rely on the explicit topological structure of the intersection or the vehicles. Instead, decisions are guided by real-time dynamic states, including vehicle positions, queue lengths, and conflict relations derived from these variables. As a result, the irregular geometry of an intersection does not pose a limitation; the conflict identification remains consistent so long as positional and movement data are available. This abstraction away from rigid geometric assumptions allows the approach to handle diverse and less structured road environments robustly.

Secondly, the proposed method can further be extended to signalized intersections, where traffic signals impose additional constraints on vehicle behavior. In such cases, the signal phase and timing information can be incorporated into the model as supplementary environmental variables. This enables vehicles to account not only for conflict relations but also for the dynamic state of the traffic lights when making decisions. The inclusion of signal states strengthens the coordination process, ensuring that decisions respect regulatory control while still optimizing conflict avoidance. As a result, the method remains effective under signalized conditions, demonstrating its versatility across both unregulated and regulated intersection scenarios.

Beyond individual intersections, the proposed method can be effectively extended to multi-intersection traffic networks. This scalability is first supported by the nature of the model's input, which relies on vehicle states such as position, speed, and queue length, as well as conflict identification based on flow directions. Since these definitions do not depend on the physical boundaries of a single intersection, they can be generalized to networks by incorporating upstream and downstream flows into the conflict classification process. Second, the framework is inherently aligned with the principle of global information sharing: while in a single intersection, agents coordinate based on local conflicts, in a network environment, the same mechanism can be applied by broadening the scope of global information to include adjacent intersections, such as their signal phases or queue conditions. Third, the agent-based design ensures scalability because each vehicle continues to make decisions at the local level—focused on its immediate conflict zone—while the consistency of decision rules across agents enables coordinated performance at the network level. Finally, the method is naturally compatible with upstream–downstream coupling effects. In multi-intersection networks, vehicles often need to anticipate spillback or congestion from neighboring intersections, and since the model already incorporates dynamic traffic states as decision inputs, it can seamlessly integrate these extended interactions into the decision-making process. Collectively, these characteristics explain why the method can transition from single-intersection scenarios to multi-intersection networks while maintaining both local efficiency and system-level coordination.

Nevertheless, several limitations must be acknowledged. First, HDV behaviors in this study are modeled using the IDM model, which primarily captures longitudinal car-following dynamics. While IDM is a widely adopted baseline in mixed traffic simulation, it does not fully represent the uncertainty and diversity of human decision-making at intersections, particularly in real-world situations. Second, although sparse conflict-aware attention improves stability, the marginal gains in fuel consumption suggest that further refinement of eco-driving objectives may be necessary. Third, the current conflict resolution mechanism relies on predefined priority rules, which may not always align with real-world fairness or regulatory standards. Finally, while our framework shows scalability to more complex or networked intersections, additional validation is required to assess computational efficiency and robustness under large-scale deployment scenarios.

Addressing these limitations will be the focus of future research. In particular, integrating real-world traffic datasets, incorporating communication uncertainty, and designing more flexible conflict-resolution policies will be crucial to bridging the gap between simulation and deployment.

## 7. Conclusion

Mixed traffic intersections present considerable challenges for safe and efficient operation, as CAVs must coordinate not only with each other but also with HDVs under complex and dynamic conditions. Addressing this problem, we proposed a conflict-aware HGRL Framework that explicitly models vehicle interactions to support cooperative and efficient decision-making.

Our main contributions are threefold. First, we developed a conflict-aware heterogeneous graph representation of mixed traffic intersections that explicitly captures both cooperative and conflict interactions among heterogeneous vehicles. This provides a unified modeling framework for analyzing the complex dynamics between CAVs and HDVs. Second, we design an HGRL framework that integrates a heterogeneous graph neural network with a DRL-based policy network to achieve safe and efficient multi-agent coordination. This design ensures that individual CAVs improve their own performance while contributing to system-level efficiency and safety. Third, we conducted comprehensive simulation experiments under varying penetration rates of CAVs, validating the proposed approach against established baselines. Moreover, ablation studies confirmed the necessity of each component in the framework.

While the results are promising, several challenges remain. First, HDV behaviors are modeled using the IDM model, which captures only basic car-following dynamics and fails to reflect the diversity of real-world decision-making at intersections. Second, computational overhead may increase when scaling to large and dense networks, which can affect real-time applicability. Future work will therefore focus on incorporating more realistic behavioral models and optimizing conflict detection and decision-making processes of the framework to ensure scalability in real-world deployments.

## Acknowledgments

This work is supported by grants from the National Key R&D Program of China (No. 2023YFB4301900) and Science and Technology Planning Project of Guangdong Province (No. 2023B1212060029). The authors declare no competing financial interests in this paper.

## References

Behrisch, M., Bieker, L., Erdmann, J., Krajzewicz, D., 2011. Sumo–simulation of urban mobility: an overview, in: Proceedings of SIMUL 2011, The Third International Conference on Advances in System Simulation, ThinkMind.

Chen, C., Xu, Q., Cai, M., Wang, J., Wang, J., Li, K., 2022. Conflict-free cooperation method for connected and automated vehicles at unsignalized intersections: Graph-based modeling and optimality analysis. IEEE Transactions on Intelligent Transportation Systems 23, 21897–21914.

Chen, P.C., Liu, X., Lin, C.W., Huang, C., Zhu, Q., 2023. Mixed-traffic intersection management utilizing connected and autonomous vehicles as traffic regulators, in: Proceedings of the 28th Asia and South Pacific Design Automation Conference, pp. 52–57.

Di, X., Shi, R., 2021. A survey on autonomous vehicle control in the era of mixed-autonomy: From physics-based to ai-guided driving policy learning. Transportation research part C: emerging technologies 125, 103008.

Dresner, K., Stone, P., 2004. Multiagent traffic management: A reservation-based intersection control mechanism, in: Autonomous Agents and Multiagent Systems, International Joint Conference on, IEEE Computer Society. pp. 530–537.

Du, Z., HomChaudhuri, B., Pisu, P., 2018. Hierarchical distributed coordination strategy of connected and automated vehicles at multiple intersections. Journal of Intelligent Transportation Systems 22, 144–158.

Grover, A., Leskovec, J., 2016. node2vec: Scalable feature learning for networks, in: Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining, pp. 855–864.

Hamilton, W.L., Ying, R., Leskovec, J., 2017. Representation learning on graphs: Methods and applications. arXiv preprint arXiv:1709.05584 .

Huang, X., Lin, P., Pei, M., Ran, B., Tan, M., 2023. Reservation-based cooperative ecodriving model for mixed autonomous and manual vehicles at intersections. IEEE Transactions on Intelligent Transportation Systems 24, 9501–9517.

Jayaweera, I., Perera, K., Munasinghe, J., 2017. Centrality measures to identify traffic congestion on road networks: A case study of sri lanka. IOSR Journal of Mathematics 13, 13–19.

Jiang, S., Pan, T., Zhong, R., Chen, C., Li, X.a., Wang, S., 2022. Coordination of mixed platoons and eco-driving strategy for a signal-free intersection. IEEE Transactions on Intelligent Transportation Systems 24, 6597–6613.

Klimke, M., Völz, B., Buchholz, M., 2022. Cooperative behavior planning for automated driving using graph neural networks, in: 2022 IEEE Intelligent Vehicles Symposium (IV), IEEE. pp. 167–174.

Li, D., Lasenby, J., 2021. Spatiotemporal attention-based graph convolution network for segment-level traffic prediction. IEEE Transactions on Intelligent Transportation Systems 23, 8337–8345.

Li, D., Zhu, F., Chen, T., Wong, Y.D., Zhu, C., Wu, J., 2023a. Coor-plt: A hierarchical control model for coordinating adaptive platoons of connected and autonomous vehicles at signal-free intersections based on deep reinforcement learning. Transportation Research Part C: Emerging Technologies 146, 103933.

Li, J., Yu, C., Shen, Z., Su, Z., Ma, W., 2023b. A survey on urban traffic control under mixed traffic environment with connected automated vehicles. Transportation research part C: emerging technologies 154, 104258.

Liu, Q., Tang, Y., Li, X., Yang, F., Gao, X., Li, Z., 2024. Sif-stgdan: A social interaction force spatial-temporal graph dynamic attention network for decision-making of connected and autonomous vehicles, in: 2024 IEEE Intelligent Vehicles Symposium (IV), IEEE. pp. 376–383.

Malikopoulos, A.A., Cassandras, C.G., Zhang, Y.J., 2018. A decentralized energy-optimal control framework for connected automated vehicles at signal-free intersections. Automatica 93, 244–256.

Munikoti, S., Agarwal, D., Das, L., Halappanavar, M., Natarajan, B., 2023. Challenges and opportunities in deep reinforcement learning with graph neural networks: A comprehensive review of algorithms and applications. IEEE transactions on neural networks and learning systems 35, 15051–15071.

Namazi, E., Li, J., Lu, C., 2019. Intelligent intersection management systems considering autonomous vehicles: A systematic literature review. Ieee Access 7, 91946–91965.

National Highway Traffic Safety Administration, 2016. Vehicle-to-Vehicle Communication Technology for Light Vehicles. Technical Report FMVSS 150. US Dep. Transp.

Pan, X., Chen, B., Timotheou, S., Evangelou, S.A., 2022. A convex optimal control framework for autonomous vehicle intersection crossing. IEEE transactions on intelligent transportation systems 24, 163–177.

Pareja, A., Domeniconi, G., Chen, J., Ma, T., Suzumura, T., Kanezashi, H., Kaler, T., Schardl, T., Leiserson, C., 2020. Evolvegcn: Evolving graph convolutional networks for dynamic graphs, in: Proceedings of the AAAI conference on artificial intelligence, pp. 5363–5370.

Perozzi, B., Al-Rfou, R., Skiena, S., 2014. Deepwalk: Online learning of social representations, in: Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining, pp. 701–710.

Pishue, B., 2021. 2021 inrix global traffic scorecard. INRIX, Washington, USAM Scorecard Rep. 1.

Porta, S., Crucitti, P., Latora, V., 2006. The network analysis of urban streets: a primal approach. Environment and Planning B: planning and design 33, 705–725.

Ribeiro, L.F., Saverese, P.H., Figueiredo, D.R., 2017. struc2vec: Learning node representations from structural identity, in: Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining, pp. 385–394.

Sankar, A., Wu, Y., Gou, L., Zhang, W., Yang, H., 2020. Dysat: Deep neural representation learning on dynamic graphs via self-attention networks, in: Proceedings of the 13th international conference on web search and data mining, pp. 519–527.

Schulman, J., Wolski, F., Dhariwal, P., Radford, A., Klimov, O., 2017. Proximal policy optimization algorithms. arXiv preprint arXiv:1707.06347 .

Shen, S., Fu, Y., Su, H., Pan, H., Qiao, P., Dou, Y., Wang, C., 2021. Graphcomm: A graph neural network based method for multi-agent reinforcement learning, in: ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE. pp. 3510–3514.

Tachet, R., Santi, P., Sobolevsky, S., Reyes-Castro, L.I., Frazzoli, E., Helbing, D., Ratti, C., 2016. Revisiting street intersections using slot-based systems. PloS one 11, e0149607.

The State Council, 2021. National comprehensive three-dimensional transportation network planning outline.

Treiber, M., Hennecke, A., Helbing, D., 2000. Congested traffic states in empirical observations and microscopic simulations. Physical review E 62, 1805.

Wang, D., Li, W., Zhu, L., Pan, J., 2025. Learning to control and coordinate mixed traffic through robot vehicles at complex and unsignalized intersections. The International Journal of Robotics Research 44, 805–825.

Wei, H., Xu, N., Zhang, H., Zheng, G., Zang, X., Chen, C., Zhang, W., Zhu, Y., Xu, K., Li, Z., 2019. Colight: Learning network-level cooperation for traffic signal control, in: Proceedings of the 28th ACM international conference on information and knowledge management, pp. 1913–1922.

Xu, B., Ban, X.J., Bian, Y., Wang, J., Li, K., 2017. V2i based cooperation between traffic signal and approaching automated vehicles, in: 2017 IEEE Intelligent Vehicles Symposium (IV), IEEE. pp. 1658–1664.

Yan, Z., Wu, C., 2021. Reinforcement learning for mixed autonomy intersections, in: 2021 IEEE International Intelligent Transportation Systems Conference (ITSC), IEEE. pp. 2089–2094.

Yao, Z., Jiang, H., Jiang, Y., Ran, B., 2023. A two-stage optimization method for schedule and trajectory of cavs at an isolated autonomous intersection. IEEE Transactions on Intelligent Transportation Systems 24, 3263–3281.

Yu, J., Wang, L., Gong, X., 2013. Study on the status evaluation of urban road intersections traffic congestion base on ahp-topsis modal. Procedia-Social and Behavioral Sciences 96, 609–616.

Yu, Q., Li, W., Yang, D., Zhang, H., 2021. Partitioning urban road network based on travel speed correlation. International journal of transportation science and technology 10, 97–109.

Yue, Y., Yeh, A.G.O., 2008. Spatiotemporal traffic-flow dependency and short-term traffic forecasting. Environment and Planning B: Planning and Design 35, 762–771.

Zainuddin, N., Shah, S., Hashim, M., Roslam, M., Tey, L., 2018. Comparison of operational performance before and after improvement: Case study at pengkalan weld, pulau pinang, in: AIP Conference Proceedings, AIP Publishing LLC. p. 020027.

Zhang, D., Yin, J., Zhu, X., Zhang, C., 2018. Network representation learning: A survey. IEEE transactions on Big Data 6, 3–28.

Zhang, J., Li, S., Li, L., 2023. Coordinating cav swarms at intersections with a deep learning model. IEEE Transactions on Intelligent Transportation Systems 24, 6280–6291.

Zhang, W., Wang, S., Tian, X., Yu, D., Yang, Z., 2017. The backbone of urban street networks: Degree distribution and connectivity characteristics. Advances in Mechanical Engineering 9, 1687814017742570.

Zhang, Y., Cassandras, C.G., 2019. An impact study of integrating connected automated vehicles with conventional traffic. Annual Reviews in Control 48, 347–356.

Zhang, Y., Wang, X., Zeng, P., Chen, X., 2011. Centrality characteristics of road network patterns of traffic analysis zones. Transportation research record 2256, 16–24.

Zhao, R., Li, Y., Wang, K., Fan, Y., Gao, F., Gao, Z., 2024. Centralized cooperation for connected autonomous vehicles at intersections by safe deep reinforcement learning. IEEE Transactions on Mobile Computing 23, 12830–12847.

Zhou, D., Hang, P., Sun, J., 2024. Reasoning graph-based reinforcement learning to cooperate mixed connected and autonomous traffic at unsignalized intersections. Transportation Research Part C: Emerging Technologies 167, 104807.