

Multi-Agent and Group Aware Reasoning for Robust Traffic Signal Control Through Jacobian Norm Based on Adversarial Defense

Ali Reza Sattarzadeh

asattarzadeh@deakin.edu.au

School of Engineering, Deakin University, Waurn Ponds, VIC 3216, Australia

Marimuthu Palaniswami

Department of Electrical and Electronic Engineering, The University of Melbourne, Australia

Albert Zomaya

School of Computer Science, The University of Sydney, Australia

Pubudu Nishantha Pathirana

School of Computer Science, The University of Sydney, Australia

Research Article

Keywords: Multi-Agent Reinforcement Learning, Traffic Signal Control, Adversarial Robustness, Jacobian Norm Regularization, Graph-based Coordination, Intelligent Transportation Systems

Posted Date: September 24th, 2025

DOI: <https://doi.org/10.21203/rs.3.rs-7668573/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: The authors declare no competing interests.

Multi-Agent and Group Aware Reasoning for Robust Traffic Signal Control Through Jacobian Norm Based on Adversarial Defense

Ali Reza Sattarzadeh¹ (asattarzadeh@deakin.edu.au)

Marimuthu Palaniswami² (palani@unimelb.edu.au)

Albert Zomaya³ (albert.zomaya@sydney.edu.au)

Pubudu Nishantha Pathirana¹ (pubudu.pathirana@deakin.edu.au)

1. School of Engineering, Deakin University, Waurn Ponds, VIC 3216, Australia
2. Department of Electrical and Electronic Engineering, The University of Melbourne, Australia
3. School of Computer Science, The University of Sydney, Australia

Abstract

Traffic Signal Control (TSC) is a critical component of intelligent transportation systems, directly influencing urban mobility, congestion levels, and environmental sustainability. While Multi-Agent Reinforcement Learning (MARL) has shown significant promise in TSC through decentralized decision-making, it remains vulnerable to adversarial attacks that manipulate learned policies, leading to inefficient traffic management. To overcome this issue, this study proposes a novel Multi-Agent and Group Aware Reasoning Framework with Jacobian Norm-Based Adversarial Defense for robust traffic signal control. Our approach leverages graph-based coordination to enhance multi-agent cooperation while incorporating Jacobian norm regularization to mitigate adversarial perturbations, ensuring both efficiency

and resilience in traffic flow optimization. To evaluate the effectiveness of our method, we conducted extensive simulations using real-world traffic datasets from New York City, Jinan, and Hangzhou under varying congestion levels. The results demonstrate that our approach consistently outperforms state-of-the-art traffic control models in both standard and adversarial conditions. Compared to the strongest baseline state-of-the-art models, our method achieves a 9.2% reduction in Average Travel Time (ATT) in Jinan (from 199.24s to 190.04s), a 9.3% reduction in New York City (from 748.32s to 704.30s), and a 4.5% reduction in New York City (from 1039.15s to 992.26s). Under adversarial perturbations, our Jacobian norm-based model maintains robust performance, achieving an ATT of 711.61s, and 997.57s in complex scenarios, significantly outperforming Deep Perturbation and Adversarial Bayesian Neural Networks. This study sets a new benchmark for intelligent transportation systems with the lowest ATT, shortest queue lengths, and highest resilience to adversarial conditions.

Keywords: Multi-Agent Reinforcement Learning, Traffic Signal Control, Adversarial Robustness, Jacobian Norm Regularization, Graph-based Coordination, Intelligent Transportation Systems

Introduction

Traffic signal control (TSC) is a fundamental aspect of urban traffic management, directly impacting congestion, travel time, and environmental sustainability. Traditional traffic control methods, such as pre-timed and actuated systems, are widely used but lack adaptability and scalability in complex, dynamic traffic networks [1]. These conventional methods rely on fixed or semi-fixed timing plans that fail to adapt to real-time traffic fluctuations, leading to inefficient traffic flow and congestion during peak hours. Furthermore, actuated systems, though more flexible, require costly sensor deployment and

maintenance, making large-scale implementation challenging [2]. Recent advancements in artificial intelligence (AI), particularly multi-agent reinforcement learning (MARL), have emerged as promising solutions for optimizing TSC by enabling decentralized decision-making through cooperative agents [3]. MARL-based approaches allow intersections to function as intelligent agents that learn optimal signal timing strategies based on real-time traffic conditions while interacting with neighboring intersections to enhance overall traffic flow [4]. Unlike traditional centralized methods, which face communication and computational bottlenecks in large-scale networks, MARL distributes control responsibilities, ensuring scalability and adaptability [5]. A crucial factor in multi-agent traffic signal control is effective coordination between agents. Various frameworks, including three-tier multi-agent architectures and holonic multi-agent systems, have been developed to enhance the robustness of TSC against internal and external disturbances [6, 7]. Holonic architectures, inspired by biological and social systems, introduce hierarchical agent structures where local agents optimize their immediate environment while global controllers oversee large-scale coordination, balancing efficiency with adaptability [6]. Additionally, grouped multi-agent reinforcement learning (GMARL) approaches, which cluster intersections into cooperative groups, further improve coordination by reducing action dimensions and capturing broader traffic flow dependencies [8]. Recent advances in deep learning have enabled the integration of graph-based methods, such as graph attention networks (GANs) and graph convolutional networks (GCNs), to improve inter-agent communication and optimize traffic control decisions [9]. These graph-based approaches capture spatial relationships within traffic networks, enabling better coordination among intersections and mitigating localized congestion [10]. By leveraging the underlying road network structure, graph-based controllers provide enhanced adaptability in dynamically changing urban environments, enabling proactive congestion management rather than

reactive signal adjustments. While MARL has shown significant promise in improving TSC, it remains vulnerable to adversarial attacks that can manipulate learned policies and degrade traffic efficiency [11]. In intelligent traffic systems relying on deep reinforcement learning (DRL), adversarial perturbations in sensor data or communication channels can lead to suboptimal decisions, increased congestion, and even safety hazards [12]. These vulnerabilities arise because DRL models rely on reward-driven optimization, making them susceptible to crafted perturbations that mislead agents into incorrect control decisions [13]. For example, small-scale perturbations in vehicle count data or maliciously injected traffic flow anomalies can trigger undesired phase transitions in traffic signals, worsening congestion across intersections.

To address these challenges, Jacobian-based adversarial defense methods have been introduced to enhance the robustness of deep learning models. Jacobian regularization, originally developed for improving generalization in neural networks, has gained renewed attention as an effective mechanism for adversarial defense [14]. By minimizing the Frobenius norm of the Jacobian matrix, this technique reduces model sensitivity to small input perturbations, thereby increasing resilience against adversarial attacks [15, 16]. Empirical studies have demonstrated that Jacobian regularization significantly improves the robustness of neural networks without compromising accuracy, making it a viable defense strategy for MARL-based traffic signal control systems [17, 18]. Moreover, Jacobian regularization fosters improved model interpretability, enabling transportation planners to better understand the reasoning behind control decisions and detect potential adversarial influences [19]. Another notable defense approach involves integrating Jacobian-based constraints with cooperative learning frameworks. For instance, researchers have demonstrated that by incorporating Jacobian norms into reward functions, MARL agents can develop policies that are both robust and interpretable under adversarial conditions [20]. This

approach ensures that multi-agent traffic controllers not only optimize efficiency but also maintain resilience in the presence of adversarial perturbations, such as spoofed traffic sensor inputs or adversarially designed vehicle platoons [21].

This paper proposes a novel integration of multi-agent and group aware reasoning with Jacobian norm-based adversarial defense to achieve robust traffic signal control. Our approach combines the strengths of MARL with advanced coordination mechanisms to improve traffic efficiency while incorporating Jacobian-based regularization to enhance resilience against adversarial attacks. By leveraging graph-based representations and multi-agent communication strategies, our framework ensures that traffic control decisions remain interpretable and secure against perturbations. Additionally, we explore the interplay between centralized and decentralized control strategies, demonstrating how hybrid approaches can balance local adaptability with global traffic optimization [22].

To address the limitations of existing traffic signal control models, we propose a Multi-Agent and Group-Aware Reasoning Framework with Jacobian Norm-Based Adversarial Defense to enhance efficiency and robustness in dynamic urban environments. Traditional MARL-based approaches often struggle with poor inter-agent coordination, leading to suboptimal traffic flow, while existing adversarial defenses fail to generalize effectively to real-world sensor noise and perturbations. Our approach overcomes these challenges by introducing a structured, interpretable, and resilient traffic control system that integrates graph-based multi-agent cooperation, probabilistic traffic forecasting, and Jacobian norm regularization to ensure stability under both normal and adversarial conditions. The key novelties of our approach are:

- We introduce an Adaptive Group Traffic Network (AGT-Net) Encoder that dynamically clusters intersections based on real-time traffic conditions, optimizing inter-agent communication to reduce network-wide congestion.

- Unlike existing methods vulnerable to sensor perturbations and adversarial attacks, our model incorporates Jacobian norm-based regularization to stabilize decision boundaries, enhancing resilience to input manipulations while maintaining optimal traffic flow.
- Our Conditional Variational Autoencoder (CVAE)-based Modified Prediction Module predicts multiple future traffic states, enabling intersections to select efficient signal timing adjustments and prevent sudden disruptions in flow.

By combining multi-agent reasoning, adversarial robustness, and probabilistic traffic forecasting, our approach sets a new benchmark for intelligent traffic signal control, significantly improving travel efficiency, coordination, and resilience against adversarial threats.

Related Works

Multi-Agent Traffic Signal Control

In practice, traffic networks operate in highly dynamic environments. Although several methods have been proposed to improve overall system performance, collaboration between local controllers such as through multi-agent reinforcement learning (MARL) systems has proven to be one of the most effective approaches [23]. Signal control in multi-intersection scenarios can benefit significantly from coordination. With recent advancements in reinforcement learning improving the effectiveness of single-agent traffic signal control, researchers have increasingly explored techniques that enable collaboration among MARL agents.

Multi-agent systems (MAS) play a crucial role in intelligent traffic signal control by enabling decentralized and cooperative decision-making. Traditional centralized traffic

control approaches often face scalability and communication bottlenecks, whereas MAS efficiently adapt to complex, dynamic traffic conditions [5, 24]. MARL has emerged as a dominant approach for traffic control, where each agent manages a single intersection and communicates with neighboring intersections to optimize traffic flow [25]. MARL-based controllers demonstrate improved efficiency over heuristic-based methods, particularly in congested urban environments [4, 26]. A key advancement in MAS for traffic control is the adoption of holonic multi-agent systems, which enhance robustness by effectively handling internal (e.g., communication noise) and external (e.g., varying traffic demand) disturbances [6]. Additionally, communication-aware multi-agent frameworks leverage deep learning techniques, such as graph neural networks (GNNs) and attention mechanisms, to enhance coordination between intersections [27]. Recent approaches explore hierarchical and grouped controllers for improved coordination. A three-tier multi-agent system dynamically optimizes signal phase adjustments using state-space models [7], while independent agent-based control systems model junctions as adjacency matrices for adaptive signal management [28]. Grouped MARL controllers cluster adjacent intersections to reduce action complexity and improve holistic urban-level decision-making [8, 29]. These architectures highlight the growing role of multi-agent coordination in achieving efficient and robust traffic signal control.

Interpretable Adversarial Defense

As multi-agent deep RL systems become more prevalent in traffic control applications, ensuring their robustness against adversarial attacks has emerged as a critical challenge. Traffic control systems increasingly rely on deep learning, making them vulnerable to adversarial attacks that could lead to congestion or safety hazards [30]. Adversarial attacks targeting intelligent traffic control exploit small perturbations in sensor inputs to manipulate decision-making [31, 32].

Several works highlight the vulnerability of DRL-based signal control to malicious perturbations. Qu et al. [33] formulate an attack in which a coalition of vehicles falsifies their sensor reports to “cheat” the controller. They show that colluding vehicles can dramatically reduce their travel time under such attacks. This vulnerability underscores the need for robust defenses. In general, multi-agent DRL, Pinto et al. [34] introduced robust adversarial reinforcement learning (RARL), training a controller against an adversarial agent applying destabilizing forces. RARL formally solves a minimax game and improves robustness and stability in continuous control domains, but it is single-agent and domains (e.g. locomotion) are far from traffic control. Its strength is a theoretically grounded worst-case design; its weakness is applicability (it does not directly extend to cooperative multi-agent traffic settings). More recent work targets cooperative multi-agent DRL directly. Bukharin et al. [35] propose ERNIE, a robust MARL framework that promotes Lipschitz continuity of each agent’s policy via adversarial regularization. ERNIE theoretically controls the policy’s Jacobian (Lipschitz constant) with respect to observations and actions, and empirically yields robustness against noisy observations, changes in dynamics, and malicious agent actions. Its strength is a general regularization principle that can be applied to many MARL algorithms, backed by theory. However, its reliance on adversarial regularizers can introduce optimization instability and overhead. Similarly, Wang et al. [36] introduce a training framework in which a learned attacker perturbs agents’ observations to drive policy robustness. This approach generates diverse adversarial examples during training, helping agents learn robust coordination; its weakness is potential complexity and lack of testing on large-scale.

A parallel concern is that DRL controllers overfit to the training traffic scenario and perform poorly on unseen patterns. Kanis et al. [37] observe that many RL traffic methods test only on in-sample flows, so they propose RLight, a simple yet robust DRL setup that

generalizes to unseen flows. RLight’s contributions include a cluster-aware state representation and an improved MDP formulation, yielding both faster learning and better out-of-sample performance. Notably, Kanis et al. split data from 5 real intersections into training and test sets to explicitly measure generalization. Their model substantially outperforms baselines on held-out traffic, demonstrating that careful representation and validation design can enhance generalization. Strengths of RLight are its empirical robustness and insight into traffic-specific MDP design. Its limitations include scope (tested on a limited real dataset) and lack of explicit mechanisms (it relies on training-test splitting rather than algorithmic regularization) – it may still struggle when faced with traffic patterns very different from any training data.

Other studies address generalization via learning strategies. Müller and Sabatelli [38] study the sim-to-real reality gap for traffic signal RL. They randomize various simulation parameters (domain randomization, DR) and also apply Model-Agnostic Meta-Learning (MAML) to train on one simulator instance and test on another. Both DR and MAML substantially outperformed a standard RL baseline when evaluated on the held-out simulator, indicating these methods can bridge discrepancies between training and deployment environments. The strength of this approach is in systematically exposing policies to diverse conditions (via DR) and enabling quick adaptation (via MAML). However, it requires knowing which parameters to randomize and incurs high training cost. Moreover, it was demonstrated on a single-intersection case; in complex urban networks, the variety of unseen patterns may be far greater.

Graph-based and transfer-learning approaches also aim at generalization [39]. For instance, methods like CoLight (Wei *et al.* [40]) use graph neural networks to capture intersection topology and neighboring interactions, which can in principle transfer policies across different road networks. Their strength is leveraging spatial structure to handle

different layouts; a weakness is they typically assume similar traffic dynamics and have not been explicitly tested for robustness to radically different flows. In summary, existing generalization methods improve resilience to some unseen situations but often rely on extensive engineering or training strategies. They rarely guarantee safety under adversarial shifts or sensor noise. Also, real traffic systems have noisy or failing sensors, bursty demand, and random incidents. A few RL works tackle these disturbances explicitly. Aslani et al. [41] propose A-CATs, an actor-critic approach for a real Tehran network that incorporates RBF-based value approximation and learns from real disturbance data. They show A-CATs “handled sensor noise and stochastic events better than Q -learning” and substantially reduced travel time, fuel use, and emissions compared to fixed-time control. Its strength lies in empirical validation on a real-world setting and explicit modeling of disruptions. But it is a single-intersection, single-agent method, and its reliance on RBF features limits scalability to very large or complex networks. Rodrigues and Azevedo [42] systematically study deep RL under demand surges, incidents, and sensor failures. Using a custom simulation framework, they find that adding Dropout to the policy network substantially improves robustness to sensor failures: “dropout is a key ingredient for developing controllers that are robust to...sensor failures”, as it forces the network to learn redundant representations. They also derive training guidelines (e.g. scenario sampling), but their experiments are limited to one intersection and DQN. The strength is clear evidence that simple regularization (dropout) can mitigate noise. Its weakness is that Dropout is a heuristic measure and may not fully prevent overreaction to structured or adversarial noise. More broadly, neural network smoothing techniques have been applied. For example, Song et al. [43] develop LipsNet, a fully connected actor network normalized to have low Lipschitz constant via a “Multi-dimensional Gradient Normalization” regularizer. In continuous control tests they show LipsNet produces much smoother actions and is empirically robust to Gaussian noise on

observations. The strength of LipsNet is architectural: it enforces bounded Jacobian norm of the policy, thereby directly controlling sensitivity. A limitation is that LipsNet was evaluated on standard control benchmarks, not on traffic network tasks; integrating it into complex traffic simulators and multi-agent interactions remains future work. Finally, the recent Infrastructures review notes that few traffic-RL studies explicitly address robustness: it recommends assessing transferability, noise resilience, or policy degradation under new conditions [44]. Indeed, many DRL-ATSC papers focus on performance under nominal settings, so resilience to sensor noise and out-of-distribution events is often not systematically studied. Existing defensive measures (RBF smoothing, dropout, Lipschitz constraints) provide partial gains but no universal guarantee.

Our proposed approach addresses these gaps by directly penalizing the policy’s Jacobian norm while training with adversarial perturbations. Intuitively, controlling the Jacobian (policy gradient) enforces local Lipschitz continuity: small changes in the observed traffic state led to small changes in action. Prior work suggests this is effective for robustness. For example, Bukharin et al. [35] argue that limiting a policy’s Lipschitz constant yields robustness to noise and dynamics shifts, and explicit Lipschitz regularization has been shown to reduce sensitivity to adversarial examples. By combining this with adversarial attacks during training (e.g. fast-gradient sign methods), our method ensures the policy learns to operate in worst-case scenarios similar in spirit to Pinto et al.’s [34] RARL but generalized to multi-agent ATSC. This joint strategy has demonstrated benefits: contemporary work finds that combining Lipschitz regularization with adversarial training significantly improves RL robustness. Specifically, the Jacobian-norm term smooths the learned value/policy function, addressing overreaction to random noise and modest distributional shifts. The adversarial attack component actively probes the policy with targeted perturbations (e.g. falsified vehicle counts or delays), so that the controller learns not to be fooled. Thus, our

method directly tackles the limitations of earlier approaches: it does not merely split data for cross-validation (as in RLight) or rely on domain randomization that may miss critical worst-cases (as in Müller & Sabatelli [38]) but explicitly optimizes for worst-case stability. It extends ERNIE’s insight by applying Lipschitz regularization via the Jacobian norm, but in a simpler mini-max training loop rather than a complex Stackelberg game. And it generalizes Rodrigues’s finding by ensuring robustness not just to random sensor dropouts but to adversarially chosen perturbations across all agents. Overall, by bridging adversarial training with differential regularization of the policy, the proposed method enhances policy resilience in all the stressed dimensions (attacks, new traffic patterns, noisy sensors) where prior work had partial solutions.

Jacobian Norm for Interpretable Adversarial Defense

Traffic control systems increasingly rely on deep learning, making them vulnerable to adversarial attacks that could lead to congestion or safety hazards [30]. Adversarial attacks targeting intelligent traffic control exploit small perturbations in sensor inputs to manipulate decision-making [31, 32]. Jacobian norm regularization has been introduced as a robust defense mechanism, minimizing the Frobenius norm of the Jacobian matrix of neural network outputs with respect to inputs to enhance stability [45]. This technique, recently revisited in adversarial defense research, reduces sensitivity to input perturbations, thereby improving robustness [20]. The theoretical foundation of Jacobian-based regularization lies in its ability to flatten decision boundaries, increasing resilience against adversarial manipulations [15]. Empirical studies confirm that Jacobian-regularized models exhibit superior robustness against attack methods such as FGSM and DeepFool [17, 20]. In traffic control, Jacobian norm-based defenses effectively secure reinforcement learning models against adversarial attacks [12, 46]. Advanced implementations integrate Jacobian constraints into multi-agent architectures, ensuring real-time robustness in dynamic

environments [47]. The intersection of Jacobian norm regularization with interpretable adversarial defense mechanisms provides both security and transparency, enabling traffic engineers to efficiently diagnose potential threats and system vulnerabilities [16, 18].

Group Aware Reasoning in Traffic Signal Control

In multi-agent traffic systems, grouping strategies improve coordination by enabling agents to share information and make collaborative decisions. Treating closely located intersections as a single entity enhances global optimization and reduces complexity compared to independent controllers [8, 29]. Grouped reinforcement learning approaches such as GPLight, dynamically cluster intersections based on real-time traffic flow and road topology, reducing training complexity while maintaining coordination efficiency [48]. Coordination graphs facilitate structured cooperation by periodically pairing adjacent intersections based on observed demand patterns. These graph-based approaches leverage multi-hop communication and knowledge-sharing to mitigate congestion in connected urban road networks. Inter-agent communication plays a vital role in group-aware reasoning. Techniques such as graph convolutional networks (GCNs) and hypergraph-based learning enable traffic controllers to model complex interactions beyond direct neighboring intersections [49, 50]. Additionally, frameworks like Friend-DQN and moral reward-based approaches encourage cooperation by optimizing joint objectives rather than isolated performance metrics [51, 52]. Robust group-aware reasoning is essential for ensuring scalability and adaptability in multi-agent traffic control. Dynamic clustering, coordination graphs, and enhanced communication protocols are shaping the next generation of intelligent transportation systems, enabling real-time adaptation to urban traffic conditions while maintaining system robustness and efficiency [53, 54].

Methodology (Multi-Agent Problem Formulation)

The goal of this research is to forecast future traffic signals for multiple agents and infer interaction relationships based on past traffic history by studying traffic prediction in different scenarios with adaptive reasoning. The whole traffic condition, represented mathematically as $\mathbb{X} \in \mathbb{R}^{N \times (T_p + T_f) \times 2}$, is made up of the previous traffic condition of N intersections in the simulation scenarios $\mathbb{X}^- \in \mathbb{R}^{N \times T_p \times 2}$, and the traffic condition, $\mathbb{X}^+ \in \mathbb{R}^{N \times T_f \times 2}$, where T_p and T_f denote their respective history horizon. The recent traffic history and traffic condition of the i^{th} agent are given by $X_i^- = \mathbb{X}_{i,:}^- = [x_i^0, x_i^1, \dots, x_i^{T_p-1}] \in \mathbb{R}^{T_p \times 2}$ as well as $X_i^+ = \mathbb{X}_{i,:}^+ = [x_i^{T_p}, x_i^{T_p+1}, \dots, x_i^{T_p+T_f-1}] \in \mathbb{R}^{T_f \times 2}$, respectively. We aim to develop a diversified prediction model $g(\cdot)$ to forecast future traffic signals of multiple agents, capturing the multimodality of various possible futures. Also, the projected future traffic signals, $\hat{\mathbb{X}}^+ = g(\mathbb{X}^-)$, will be as near as feasible to the future traffic signals $\mathbb{X}^+$. Agent interaction modeling is required in the prediction model $g(\cdot)$ to enhance accuracy of future prediction. Furthermore, the model $g(\cdot)$ explicitly captures interaction impacts among agents, such as interaction between different intersections with additional semantic modeling to account for these impacts. Since interaction factors can change over time, semantic modeling can be further extended to dynamic semantic modeling. Adaptive semantic modeling is an unsupervised process, relying solely on the traffic signals of individual agents without prior knowledge of underlying relationships.

The prediction model $g(\cdot)$ builds on earlier techniques that model pair-wise interactions, such as [55-58]. Nonetheless, it is typical for numerous agents to communicate in group settings. We examine interactions in $g(\cdot)$ on both a pair and group level in this study using an adaptive cross-scale supergraph. Furthermore, we consider interactions between different intersections and their interaction strengths in the prediction model $g(\cdot)$, whereas other

adaptive semantic modelling approaches, including those found in [58, 59], focus only on interactions between intersections.

Architecture overview: Multi-Agent

The traffic condition in complex scenarios is simulated as an adaptive cross-scale super-graph, where agents are nodes and interactions are hyperedges, forming the foundation of our proposed Adaptive Group Traffic Net (AGT-Net). This adaptive cross-scale super-graph is then used to train agent and time-varying interaction representations. To depict adaptive agent interaction, we consider (K) developing steps. At each step, a sub-traffic condition over a time period serves as input, and the corresponding agent representations are the output.

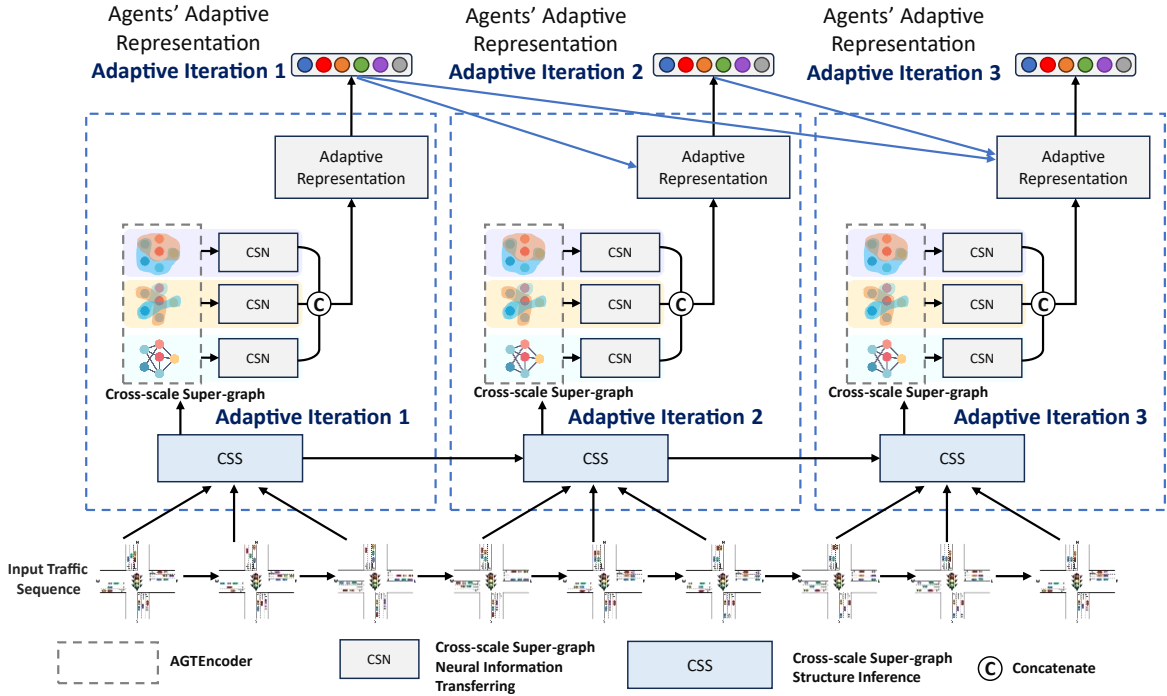


Figure 1: Overview AGT-Net architecture. AGT-Net encodes traffic signals through multiple adaptive iterations, utilizing the AGTEncoder for each iteration to model both cross-scale temporal and group interactions. The AGTEncoder employs an adaptive cross-scale super-graph, constructed via cross-scale super-graph structure inference and neural information transfer, to derive representations and interaction relationships.

We employ an AGT-Encoder in each adaptive iteration, consisting of three main functions: cross-scale super-graph neural information transfer, cross-scale supergraph structure inference, and adaptive representation evolution. The cross-scale super-graph topology is deduced from the input sub-traffic condition. Neural information transfer captures the connections and interactions among agents, while adaptive representation evolution models the temporal dependencies of learned agent patterns, integrating them to produce adaptive agent representations. Let $\mathbb{X} \in \mathbb{R}^{N \times (T_p + T_f) \times 2}$ represent the whole input traffic condition of AGT-Net mathematically. The input sub-traffic condition at the k -th adaptive iteration is $\mathbb{X}^{[k]} = \mathbb{X}_{:,k\tau:k\tau+T_E-1,:} \in \mathbb{R}^{N \times T_E \times 2}$, indicating that we use an adaptive length T_E as the input sub-traffic condition and conduct one adaptive step by every developing gap τ . The overall AGT-Net architecture is shown in Figure 1, where $T_E = \tau = 3$.

Network Components and Implementation Details

The traffic prediction system employs several specialized subnetworks, each with a distinct architecture, to process information and generate predictions. All Multi-Layer Perceptrons (MLPs) within these subnetworks utilize the ReLU activation function unless otherwise specified. Within the AGTEncoder's cross-scale super-graph topology inference, an initial traffic embedding for each agent is generated by an MLP, $f_{\text{init}}(\cdot)$ that outputs a feature vector of dimension $d = 32$. During the cross-scale super-graph neural information transfer stage, several MLPs are utilized. For calculating the weights in the collective representation $e_i^{(s,k)}$, an MLP denoted as $\mathcal{F}_l^{(s)}(\cdot)$ is used. The agent connection strength $r_i^{(s,k)}$ is inferred using an MLP $\mathcal{F}_r^{(s)}(\cdot)$, followed by a sigmoid activation function to constrain the output. Neural interaction categories $c_{i,l}^{(s,k)}$ are determined by an MLP $\mathcal{F}_c^{(s)}(\cdot)$, followed by the addition of Gumbel (0,1) noise, division by a temperature $\tau = 0.5$, and finally the application of a SoftMax function. There are $L = 6$ categories for the zeroth scale

(pairwise interactions) and $L = 10$ for other scales (group-wise interactions). Each per-category function, $\mathcal{F}_l^{(s)}(\cdot)$, is a three-layer MLP with a hidden layer size of 128, taking the sum of agent representations as input and outputting a vector of dimension $d = 32$. In the hyperedge-to-node phase, agent embeddings are updated by a trainable MLP, $f_v^{(a)}(\cdot)$, which processes the concatenation of the current agent embedding and associated interaction embeddings, outputting a vector of dimension $d = 32$. The adaptive latent representations mechanism uses an MLP, $\mathcal{F}_{pc}(\cdot)$, for positional encoding. The core of this module is a two-layer transformer with 8 attention heads, which takes the current agent representation (dimension $d(S + 1)$) as query and previous dynamic representations with positional encoding as keys and values. In the Prediction Module, the future traffic encoding module, $\mathcal{F}_f(\cdot)$, is implemented as a one-layer bidirectional LSTM with a hidden dimension of 32. The parameters for the prior $(\mu_p^{[k]})$ and posterior $(\mu_q^{[k]})$ distributions of the latent codes (dimension $d_z = 20$) are generated by separate linear layers, $\mathcal{F}_p(\cdot)$ and $\mathcal{F}_q(\cdot)$ respectively. The Decoder architecture features a GRU cell for recurrently calculating hidden states; this GRU has one layer with a hidden dimension of 128. The parameters for the bivariate GMM are then generated from the GRU's hidden state by a single linear layer, $\mathcal{F}_{GMM}(\cdot)$. Finally, the prediction modified module, $\mathcal{M}_{\text{ref}}(\cdot)$, is a two-layer MLP. This MLP has a first hidden layer with 256 units and a second hidden layer with 512 units, designed to smooth the final predicted traffic flows.

Cross-scale super-graph topology inference

We propose inferring a cross-scale super-graph to represent interactions among agents at different intersection group sizes, based on the k^{th} adaptive iteration input $\mathbb{X}^{[k]}$. The super-graph, denoted as $\mathcal{G}^{[k]} = \{\mathcal{G}^{(0,k)}, \mathcal{G}^{(1,k)}, \dots, \mathcal{G}^{(S,k)}\}$, captures connections between agents at the

k -th adaptive iteration, where $\mathcal{V}$ is the set of N agents, and $\mathcal{G}^{(s,k)} = (\mathcal{E}^{(s,k)}, \mathcal{V})$ represents the super-graph of the s -th scale. Group-wise relations are represented by the hyperedge set $\mathcal{E}^{(s,k)} = \{e_1^{(s,k)}, e_2^{(s,k)}, \dots, e_{M_s}^{(s,k)}\}$, where each connections show the relations. In addition, the edges of a particular super-graph, $\mathcal{G}^{(0,k)} = (\mathcal{V}, \mathcal{E}^{(0,k)})$, represent the best pair-wise agent connections. Also, $H^{(s,k)} \in \mathbb{R}^{|\mathcal{V}| \times |\mathcal{E}^{(s,k)}|}$ is used to describe the structure of each $\mathcal{G}^{(s,k)}$. Since the graph topology is not predefined, interactions between traffic signals at intersections are in a stealth state and difficult to capture. We create hyperedges by clustering agents with strongly linked traffic signals, quantified by projecting the signals into a high-dimensional feature space to infer the cross-scale super-graph at the k^{th} iteration. For the input traffic signal $X_i^{[k]}$ at the k^{th} adaptive iteration, we define its initial traffic signal latent feature as $q_i^{[k]} = f_{\text{init}}(X_i^{[k]}) \in \mathbb{R}^d$, where $f_{\text{init}}(\cdot)$ is a MLP and i denotes the agent number. Also, the relation between agents at the k^{th} adaptive iteration is reflected in an affinity matrix $A^{[k]} \in \mathbb{R}^{N \times N}$. The element that is $(i, j)^{th}$ in $A^{[k]}$ is:

$$A_{i,j}^{[k]} = q_i^{[k]\top} q_j^{[k]} / (\|q_i^{[k]}\|_2 \|q_j^{[k]}\|_2), \quad (1)$$

At the k^{th} adaptive iteration, indicates the correlation between two agents. To guarantee that the traffic representations contain unit measure for stable relation estimation, we normalize the representations. To preserve and smooth historical agent behavior, the $A^{[k]}$ at position k ($k > 0$) is modified by adding the weight total of the $A^{[k-1]}$ as follows:

$$A^{[k]} = (A^{[k]} + \alpha A^{[k-1]}) / (1 + \alpha), \quad (2)$$

where $\alpha \in [0,1]$ modifies the weight of the prior affinity matrix. Hyperedges are then constructed at different scales based on $A^{[k]}$ at the k^{th} adaptive iteration. Pairwise connections are taken into account at the 0^{th} scale, with $H^{(0,k)}$ containing M_0 edges, each connecting a node to its top-scoring nodes. For each scale, group-wise relationships are modelled by identifying high-density submatrices in $A^{[k]}$. We designate a series of rising group sizes $\{M^{(s)}\}_{s=1}^S$. In addition, strongly correlated group is recognized at any scale s for each node. By resolving the optimization issue of seeking a $M^{(s)} \times M^{(s)}$ submatrix, this hyperedge is obtained as follow:

$$\begin{aligned} e_i^{(s,k)} = \arg \max_{\Omega^{(s,k)} \subseteq \mathcal{V}} & \|A_{\Omega^{(s,k)}, \Omega^{(s,k)}}^{[k]}\|_{1,1} ; \\ \text{such that } |\Omega^{(s,k)}| &= M^{(s)}, v_i \in \Omega^{(s,k)}, i = 1, \dots, N, \end{aligned} \quad (3)$$

where $\|\cdot\|_{1,1}$ denotes the sum of absolute values of all elements. The goal is to identify and connect the most strongly connected agents, with restrictions ensuring a fixed group size and inclusion of the i^{th} node. One hyperedge is created per node at each scale. To solve the optimization problem in Eq. 3, we employ different approaches based on the number of agents: for systems with a small number of agents, an enumeration algorithm can identify the optimal solution, while for larger systems, a greedy algorithm provides an efficient approximation. Thus, $\{H^{(0,k)} \in \mathbb{R}^{N \times NM^{(0)}}, \{H^{(s,k)} \in \mathbb{R}^{N \times N}\}_{s=1}^S\}$ is utilized for a cross-scale super-graph of the k^{th} adaptive iteration. Here, $H^{(0,k)}$ and $H^{(s,k)}$ ($s \geq 1$) denote the $NM^{(0)}$ edges and N hyperedges to illustrate pairwise and group-wise interactions within $M^{(s)}$ agents, respectively. Our cross-scale supergraph differs from the typical cross-scale graph by maintaining a stable node count while varying hyperedge sizes across scales [60, 61]. Using the same affinity matrix to select edges and hyperedges offers two advantages: (i) low computational cost for identifying high-order relations, and (ii) stable and informative

training via back-propagation. Our topology inference approach is novel compared to collaborative prediction and neural relational inference methods [62, 63] in two ways: (i) it actively infers network structure through learning, unlike fully connected architectures, and (ii) it captures both group- and pair-wise links at several scales, unlike single-scale pair-wise approaches.

Cross-scale super-graph neural information transfer

We adapt a cross-scale super-graph neural information transferring algorithm to iteratively capture the interaction between agents as well as features through forward and backward path between nodes and hyperedges, as shown in Figure 2. The control pattern is learned based on agents signal control at the k^{th} adaptive iteration, given the inferred cross-scale super-graph. First, agent representations are derived from each agent's traffic signal. Agent features are then combined to capture the interaction between agents and features iteratively through forward path between node as well as hyperedge. Every agent representation is updated based on relation interaction representations during backward path between hyperedge and node. Finally, each agent's representation is obtained by integrating information across scales. We demonstrate an interaction representation with three components to support a relational reasoning: (i) neural interaction strength, capturing interaction intensity, (ii) neural interaction category, reflecting interactions between different intersections of the agents, and (iii) per- intersection function, modeling the interaction phase. For the $e_i^{(s,k)}$, at the s -th scale of the super-graph of adaptive iteration k , its interaction representation is produced as follows:

$$e_i^{(s,k)} = r_i^{(s,k)} \sum_{\ell=1}^L c_{i,\ell}^{(s,k)} \mathcal{F}_\ell^{(s)} \left(\sum_{v_j \in e_i^{(a,k)}} v_j \right) \in \mathbb{R}^d, \quad (4)$$

where $r_i^{(s,k)}$ denotes the connection strength. The likelihood of the ℓ -th agent interaction between different intersections among L potential intersections is indicated by $c_{i,\ell}^{(s,k)}$, the ℓ -th component of an intersection $c_i^{(s,k)} \in [0,1]^L$. We designate a learnable intersection function $\mathcal{F}_\ell^{(s)}(\cdot)$ implemented by MLPs for every intersection at scale s . These components are trained within an end-to-end framework.

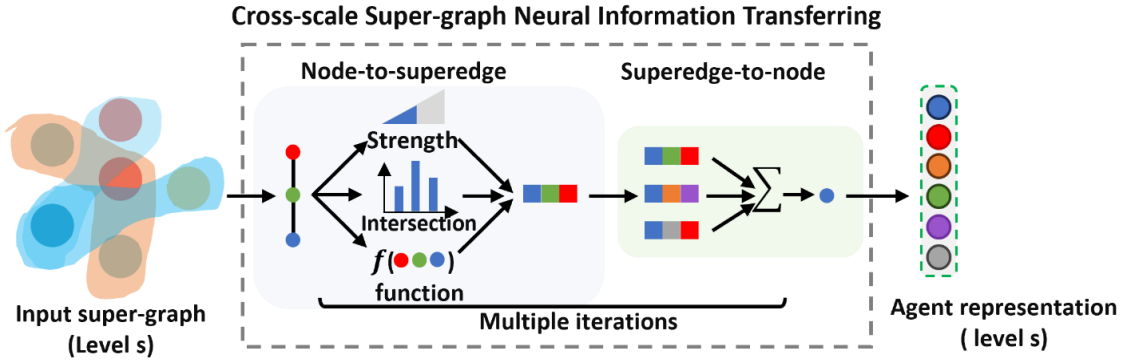


Figure 2: Cross-scale super-graph neural information transfer. In the node-to- super-graph stage, the proposed model used related agents' representation to extract their interaction representation with three components.

Latent features at the s^{th} level are utilized to derive the agent interaction between different intersections $c_i^{(s,k)}$ and agent interaction strength $r_i^{(s,k)}$ at the k^{th} adaptive iteration. These latent features, representing a hidden state, reflects the features of agents participating in a group. The latent feature $z_i^{(s,k)}$ of $e_i^{(s,k)}$ at the s^{th} scale is obtained by adding up all the weighted agent feature relations with the hyperedge $e_i^{(s,k)}$. At the k^{th} adaptive iteration, the weight $w_j^{(s,k)}$ represents the j^{th} node's contribution at level s . The agent interaction between different intersections, $c_i^{(s,k)}$, and agent interaction strength, $r_i^{(s,k)}$, are derived from the collective representation as follows:

$$\begin{aligned} r_i^{(s,k)} &= \sigma \left(\mathcal{F}_r^{(s)}(z_i^{(s,k)}) \right), \\ c_i^{(s,k)} &= \text{softmax} \left(\left(\mathcal{F}_c^{(s)}(z_i^{(s,k)}) + g \right) / \gamma \right), \end{aligned} \quad (5)$$

where g denotes the sigmoid function, which is an independent and identically distributed random vector sampled from the Gumbel distribution $(0,1)$, and $\sigma(\cdot)$. We approximate the discrete distribution continuously using the Gumbel softmax [64]. MLPs are utilized to simulate the neural interaction between different intersections and strength at the s^{th} level, respectively, using functions $\mathcal{F}_r^{(s)}(\cdot)$ and $\mathcal{F}_c^{(s)}(\cdot)$. By using a shared latent feature, the inference of agent interaction intensity, agent interaction between different intersections, and per-intersection function takes group behavior into account. In addition, agent latent feature is updated by considering all the interactions learned from agent groups. Also, the hyperedges related with v_i in the s^{th} level of super-graph at the k^{th} adaptive iteration are denoted as $\mathcal{E}_i^{(s,k)} = \{e_j \mid v_i \in e_j, e_j \in \mathcal{E}^{(s,k)}\}$. At the k^{th} adaptive iteration, the i^{th} agent's latent feature in the s^{th} level is updated as:

$$v_i^{(s,k)} \leftarrow f_v^{(s)} \left(\left[v_i^{(s,k)}, \sum_{e_j \in \mathcal{E}_i^{(s,k)}} e_j \right] \right) \in \mathbb{R}^d, \quad (6)$$

where $[\cdot, \cdot]$ indicates the concatenation operator and $f_v^{(s)}(\cdot)$ is a trainable MLP. The latent feature $v_i^{(s,k)}$ is iteratively refined through forward and backward paths between nodes and hyperedges at level s during the k^{th} adaptive iteration. Latent features $v_i^{(0,k)}, v_i^{(1,k)}, \dots, v_i^{(S,k)}$ are computed across S levels in parallel phase. As a result, the agent latent feature is concatenated over all scales to get the latent feature at the k^{th} adaptive iteration of the i^{th} agent. Our super-graph agent message transferr is unique in three ways from other

approaches [65], [60], and [66]. First, unlike Kipf et al. and Graber et al. [63, 67], which use message transfer on regular graphs, our method employs neural message passing on a sequence of hypergraphs, enabling richer information propagation and aggregation. Second, our technique infers interactions between intersections and their strength using a three-component representation, whereas these methods only infer the interactions between different intersections. Third, the interaction between different intersections and strength estimated by the proposed technique are the intermediate properties involved in agent message transferring, whereas these approaches predict interactions as the final output following message transfer. Our strategy not only enhances prediction performance but also improves relational thinking in agent message transfer.

Adaptive latent representations

The latent feature of an agent is localized within a specific time frame at each adaptive iteration. We construct an adaptive latent representation technique to upgrade each agent's latent features by integrating historical features from prior adaptive iterations, thereby improving the temporal interdependence of latent features across adaptive iterations. The current latent feature serves as the query, while prior adaptive latent features act as the key and value in a transformer structure, which records global temporal relationships to facilitate this evolution. The adaptive latent feature is denoted as $\tilde{V}_i^{[-k]} = [\tilde{v}_i^{[0]} + p^{[0]}, \dots, \tilde{v}_i^{[k-1]} + p^{[k-1]}] \in \mathbb{R}^{d(S+1) \times k}$ prior to the k^{th} adaptive iteration, where $p^{[k]} \in \mathbb{R}^{d(S+1) \times k}$ is a positional vector, and $[\cdot, \cdot]$ denotes concatenation. In addition, a positional encoding is applied to each developing step, computed as $p^{[k']} = \mathcal{F}_{pe} \left(-1 + 2 \frac{k'}{k} \right)$, where $\mathcal{F}_{pe}(\cdot)$ is a MLP function. The adaptive latent feature of the i^{th} agent at the k^{th} adaptive iteration is generated as follows:

$$\tilde{v}_i^{[k]} = v_i^{[k]} + MHA \left(v_i^{[k]}, \tilde{V}_i^{[-k]}, \tilde{V}_i^{[-k]} \right) \in \mathbb{R}^{d(s+1)}, \quad (7)$$

Multi-head attention $MHA(\cdot)$ is used in the proposed model. Figure 3 illustrates the adaptive representation evolution process. Using adaptive representation evolution, the model leverages temporal relations and adaptive representations to integrate interaction information of past timestamps into an adaptive cross-scale super-graph. Our adaptive representation offers two benefits over earlier approaches [63, 67] that also address the adaptation of the agents' state. First, the proposed model not only captures longer-range temporal dependencies, but also benefits from a global receptive field, allowing it to consider all adaptive iterations.

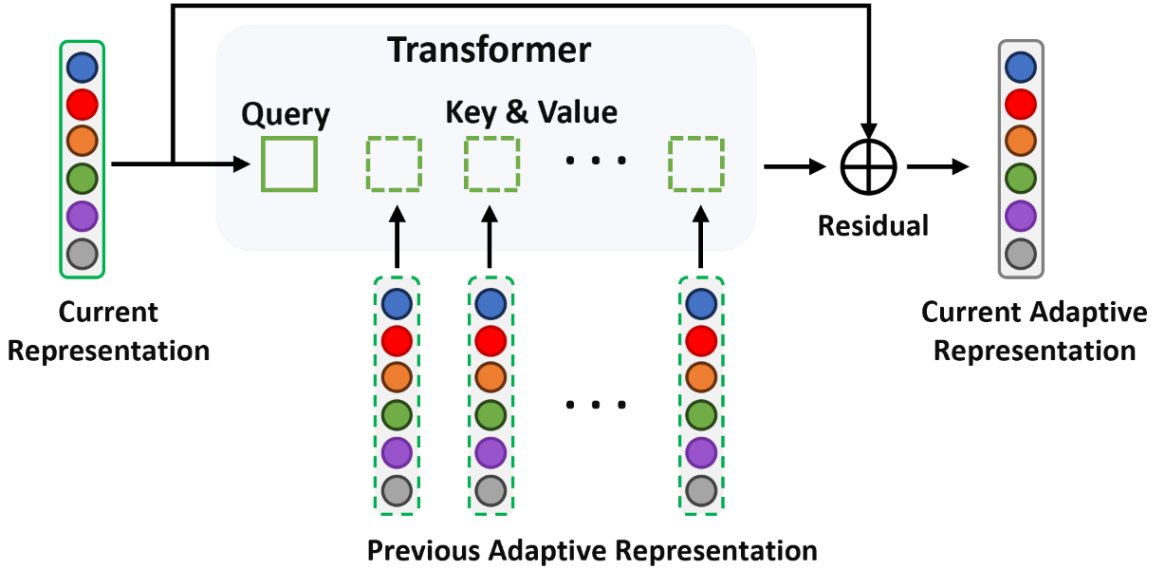


Figure 3: Overview of adaptive representation. The history of representation is fed as input in a transformer.

Additionally, building upon the superior capability of transformer-based architectures in capturing long-term dependencies, our adaptive representation evolution method further enhances this ability compared to LSTM and GRU [68]. Second, while earlier methods

focused solely on the evolution of interaction categories, our approach focuses on the evolution of entire agents, including multilevel group interactions between different intersections, resulting in a more comprehensive evolution of agent states.

Prediction Module

This study proposes a control approach centered on the AGT-Net as the main component. The approach anticipates future traffic signals in a recursive manner. The conditional variational autoencoder (CVAE) architecture serves as the foundation for the prediction system. As shown in Figure 4, the prediction system consists of an encoding phase based on AGT-Net, a decoding module, and a control module. The encoding module uses AGT-Net to compute adaptive agent representations, generating a latent feature presentation based on previous or projected traffic signals. The decoding module then utilizes these latent features and adaptive agent representations to predict future traffic signals for a specified period. By including the anticipated traffic signals as a component into the next input, the encoding-decoding process could initiate again. The ultimate output of our prediction system is the modified anticipated traffic signals, obtained by combining the projected traffic signals from each encoding-decoding step and modifying them collectively. The adaptive length $T_E = T_p$ is selected to fully utilize the known previous traffic signals $\mathbb{X}^- \in \mathbb{R}^{N \times T_P \times 2}$. In addition, the length of projected traffic signals is set by the decoder to the developing gap $T_D = \tau$ to ensure non-overlapping predictions. To simplify the following equations, let $\hat{\mathbb{X}}_{in}^{[k]} = [\hat{X}^{kT_D}, \dots, \hat{X}^{kT_D+T_E-1}] \in \mathbb{R}^{N \times T_E \times 2}$ and $\mathbb{X}_{out}^{[k]} = [X^{kT_D+T_E}, \dots, X^{(k+1)T_D+T_E-1}] \in \mathbb{R}^{N \times T_D \times 2}$ represent the ground-truth and the input, where $[k]$ illustrates the k^{th} phase and $\hat{\cdot}$ denotes the predicted sequence. For historical timestamps $t < T_p$, we set $\hat{X}^t = X^t$ since the historical sequence is known.

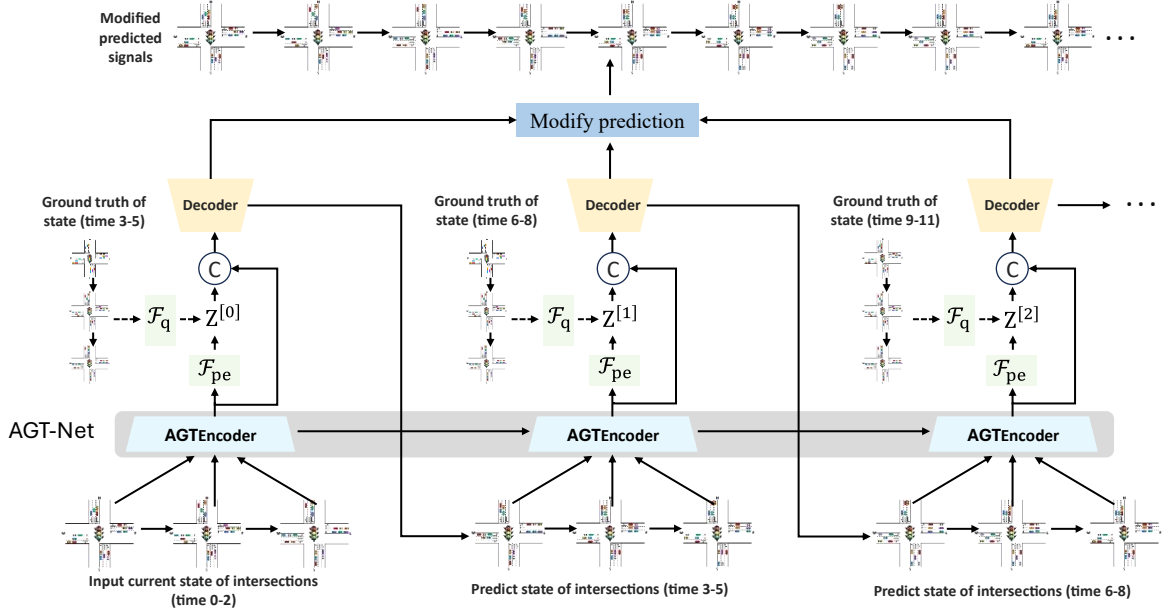


Figure 4: Overview of the prediction model architecture for encoding and decoding with iteration length 3. We utilized the AGT-Net based on the AGTEncoder, decoder, and modified prediction module to create the prediction model. To create a final traffic signal, the model utilized all future timestamp signals and modified them to capture final signal.

Encoder

The distribution parameters of the latent feature are generated in the encoding phase and then samples the latent representation $Z^{[k]} \in \mathbb{R}^{N \times d_z}$ from the distribution, where d_z represents the latent features dimension. Moreover, $\hat{\mathbb{X}}_{\text{in}}^{[k]} \in \mathbb{R}^{N \times T_E \times 2}$ and, $\mathbb{X}_{\text{out}}^{[k]} \in \mathbb{R}^{N \times T_D \times 2}$ denote the inputs and related ground-truth, respectively. In addition, the $\text{Cat}(d_z, \mu)$ denotes the categorical distribution that is adapted for all N agents to simulate the multi-intersection of the future traffic signals. Moreover, μ_i denotes the distribution for categorical distribution to satisfy $\|\mu_i\|_1 = 1, \mu_{i,j} \in [0,1], j \in \{0,1, \dots, d_z - 1\}$ for the i^{th} agent. Each agent's representation is then samples from the distribution of one-hot discrete categorical latent features.

The distribution parameter is calculated as follows:

$$\begin{aligned}
\tilde{V}^{[k]} &= \text{AGT}(\hat{\mathbb{X}}_{\text{in}}^{[k]}), V_{\text{gt}}^{[k]} = \mathcal{F}_f(\mathbb{X}_{\text{out}}^{[k]}), \\
\mu_p^{[k]} &= \mathcal{F}_p(\tilde{V}^{[k]}), \mu_q^{[k]} = \mathcal{F}_q([\tilde{V}^{[k]}, V_{\text{gt}}^{[k]}]), \\
p(Z^{[k]} | \hat{\mathbb{X}}_{\text{in}}^{[k]}) &= \text{Cat}(d_z, \mu_p^{[k]}), \\
q(Z^{[k]} | \hat{\mathbb{X}}_{\text{in}}^{[k]}, \mathbb{X}_{\text{out}}^{[k]}) &= \text{Cat}(d_z, \mu_q^{[k]}).
\end{aligned} \tag{8}$$

Adaptive agent representations are captured by the proposed model based on the AGT-Net (Eq. 8). Two linear layers are utilized to calculate distribution parameters for the prior $\mu_p^{[k]}$ and posterior $\mu_q^{[k]}$ states in k^{th} encoding phase, as shown in Eq. 8. During the training phase, potential future traffic signals are sampled from latent features $Z^{[k]}$. Since ground-truth traffic signals are unavailable during the evaluation phase, this study tries to sample $Z^{[k]}$ from the previous distribution $Z^{[k]} \sim p(Z^{[k]} | \hat{\mathbb{X}}_{\text{in}}^{[k]})$. The encoding phase results in latent features $Z^{[k]}$ concatenated with Adaptive agent representations $\tilde{V}^{[k]}$: $V_E^{[k]} = [Z^{[k]}, \tilde{V}^{[k]}]$. The design of the encoding phase aims to address the stochasticity of each actor's behaviour. To capture multimodal behavior, latent features are sampled from their distribution. Also, both possible multimodal behavior and present movement information are incorporated by concatenating the latent features with the adaptive representations of the current agents. Figure 5 shows the prediction module design, which comprises a decoding phase, a control phase, and an encoding phase based on AGT-Net. The latent features are created during the encoding phase, and the Traffic signals are produced during the decoding phase utilizing the latent features. The encoding-decoding process iterates by assigning the output traffic signals as inputs for the subsequent encoding phases. The prediction module integrates these predictions and applies a modified prediction module to obtain the final prediction results after projecting the traffic signals for each future timestamp.

Decoding phase

The encoding phase output $V_E^{[k]}$ is used by the decoder to anticipate the k^{th} encoding-decoding phase $\hat{X}_{out}^{[k]} = [\hat{X}^{kT_D+T_E}, \dots, \hat{X}^{(k+1)T_D+T_E-1}]$, as shown in Figure 5. Rather than deterministically regressing the traffic signals coordinates, the model computes the distribution of the projected traffic signals to accurately depict the uncertainty of the expected traffic signals. A bivariate Gaussian distribution is used to generate the projected traffic signals' distribution. The next signal control of the agents is obtained by sampling from this bivariate Gaussian distribution, with parameters are generated step-by-step by each GRU cell. The l^{th} generating step ($l \in \{0, \dots, T_D - 1\}$) in the k^{th} decoding phase is defined as follows:

$$\begin{aligned} M^{[k],l} &= \text{GRU} \left(M^{[k],l-1}, [V_E^{[k]}, \hat{X}^{kT_D+T_E+l-1}] \right), \\ \Delta \hat{X}^{kT_D+T_E+l} &\sim \text{GMM} \left(\mathcal{F}_{\text{GMM}}(M^{[k],l}) \right), \\ \hat{X}^{kT_D+T_E+l} &= \hat{X}^{kT_D+T_E+l-1} + \Delta \hat{X}^{kT_D+T_E+l}. \end{aligned} \quad (9)$$

The hidden states of l^{th} step are recurrently calculated in Eq. 9 as $M^{[k],l}$. A linear layer $\mathcal{F}_{\text{GMM}}(\cdot)$ is used to produce the GMM distribution parameters. By carrying out T_D decoding phases, the model generates the predicted signals for the k th encoding-decoding process: $\hat{X}_{out}^{[k]} = [\hat{X}^{kT_D+T_E}, \dots, \hat{X}^{(k+1)T_D+T_E-1}] \in \mathbb{R}^{N \times T_D \times 2}$. While using a bivariate Gaussian distribution to characterize the uncertainty of traffic signals enhances prediction diversity. It may also provide erroneous next inputs into the learning phase of the encoding-decoding phase, making the training phase less effective. To address this, a multiple sampling technique is utilized, where traffic signals are samples numerous times and the most correct one is selected for subsequent predictions during the learning stage, stabilizing the learning process. The decoding process is iterated K_{ms} times to obtain K_{ms} predicted traffic signals.

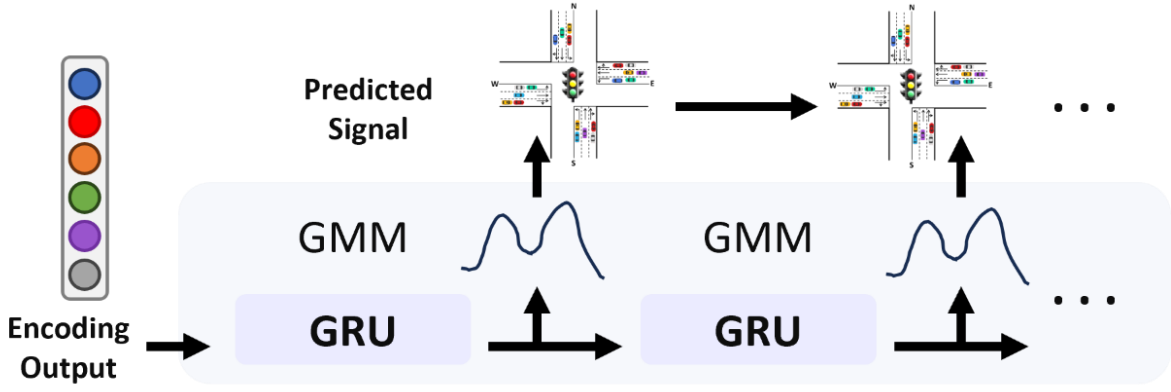


Figure 5: Overview of the prediction module architecture that handled uncertainty to predict future signal by GMM when the parameters of GMM is captured by GRU.

Modified Prediction

We collect all decoder outputs to generate a comprehensive traffic signal forecast, $\hat{\mathbf{X}}^+ = [\hat{\mathbf{X}}^{T_p}, \hat{\mathbf{X}}^{T_p+1}, \dots, \hat{\mathbf{X}}^{T_p+T_r-1}]$. The encoding-decoding phase is repeated $\left\lceil \frac{T_r}{T_D} \right\rceil$ times. Nevertheless, since the GMM model samples feature for each timestamp individually throughout the decoding process, the projected coordinates may lack correlation, potentially leading to unsmoothing issues. To address this, we propose a modified prediction module to improve the correlation of anticipated coordinates across all timestamps by modifying the entire set of future signals. In particular, $\hat{\mathbf{X}}_{\text{ref}}^+ = \mathcal{M}_{\text{ref}}(\hat{\mathbf{X}}^+)$ may be used to express the prediction modified, where $\mathcal{M}_{\text{ref}}(\cdot)$ is the modified method similar to MLPs.

Training loss

To train the prediction system, we optimized the model by utilizing the loss function $\mathcal{L} = \mathcal{L}_{\text{elbo}} + \mathcal{L}_{\text{modified}}$.

$$\begin{aligned}
\mathcal{L}_{\text{elbo}} = & \beta_1 \sum_{k=0}^{\lceil \frac{T_i}{T_D} \rceil - 1} - \mathbb{E}_q \left[\log p \left(\hat{\mathbb{X}}_{\text{out}}^{[k]} \mid Z^{[k]}, \hat{\mathbb{X}}_{\text{in}}^{[k]} \right) \right] \\
& + \beta_2 \sum_{k=0}^{\lceil \frac{T_i}{T_D} \rceil - 1} \text{KL} \left(q \left(Z^{[k]} \mid \mathbb{X}_{\text{out}}^{[k]}, \hat{\mathbb{X}}_{\text{in}}^{[k]} \right) \parallel p \left(Z^{[k]} \mid \hat{\mathbb{X}}_{\text{in}}^{[k]} \right) \right)
\end{aligned} \tag{10}$$

Here, $\mathcal{L}_{\text{elbo}}$ denotes the negative evidence lower-bound, where β_1, β_2 represent the hyperparameters. $\text{KL}(\cdot \parallel \cdot)$ represents the KL divergence, and $\mathcal{L}_{\text{modified}} = \beta_3 \|\hat{\mathbb{X}}_{\text{ref}}^+ - \mathbb{X}^+\|_2^2$, with β_3 as the hyperparameter. The prediction modified loss is $\mathcal{L}_{\text{modified}} = \beta_3 \|\hat{\mathbb{X}}_{\text{ref}}^+ - \mathbb{X}^+\|_2^2$. We include the KL-divergence and all negative log-likelihoods of each encoding-decoding operation in the negative evidence lower-bound $\mathcal{L}_{\text{elbo}}$.

Preliminary Jacobian norm for TSC

Assume that the proposed model $F_\theta(X): X \mapsto \hat{Y} \in \mathbb{R}^{N \times K}$ maps the inputs $X \in \mathbb{R}^{N \times D}$ to the output as control signals for intersections. The proposed model parameters are represented by θ , and $\hat{Y} \in \mathbb{R}^{N \times K}$ yields the projections of F_θ . The goal is to identify a set of parameters θ^* on the training set that minimizes error between predicted control signal $\hat{Y}$ and true signal control $Y \in \mathbb{R}^{N \times K}$ in order to train the model F_θ . It is also possible to express $\theta^* = \arg \min_\theta H(Y, \hat{Y})$, where H denotes the loss function to minimize the error. Alternatively, it can be expressed as $\sum_{n=1}^N \sum_{k=1}^K -Y_{n,k} \log F_\theta(X_{n,k})$. This study utilized the Jacobian matrix J for the input of the proposed model as below:

$$J(\mathbf{x}) := \nabla_{\mathbf{x}} F_\theta(\mathbf{x}) = \left[\frac{\partial F_\theta(\mathbf{x})}{\partial x_1}, \dots, \frac{\partial F_\theta(\mathbf{x})}{\partial x_D} \right], \tag{11}$$

where x dimensionality is represented by $D = h \times w \times c$. Although the proposed model improves the performance on the training set, its performance rapidly decreases when noisy samples are included. The proposed model is considered robust against this attack based on the Eq. 12 even when a tiny noise z is added to the input with $p=\infty$ and ϵ serving as the scaling factor.

$$\arg \max_{k \in K} F_{\theta}^k(x) = \arg \max_{k \in K} F_{\theta}^k(x + \epsilon z), \forall \epsilon \in B_p(\epsilon) = \epsilon: \|\epsilon\|_p \leq \epsilon, \quad (12)$$

Also, adversarial training [69] enhances the robustness of the model by matching distribution between the adversarial test data as well as the training data through the reducing loss function as follows:

$$\min_{\theta} \rho(\theta), \text{ when } \rho(\theta) = \mathbb{E}_{(x,y)} \left[\max_{\epsilon \in B(\epsilon)} H(F_{\theta}(x + \epsilon z), y) \right], \quad (13)$$

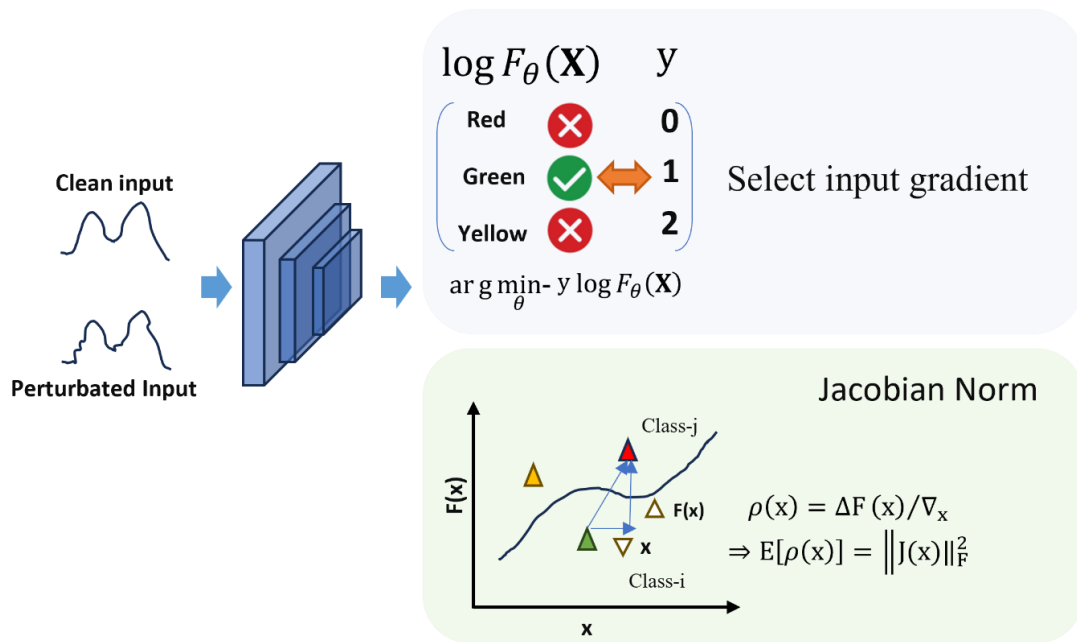


Figure 6: Overview of adversarial robustness architecture that utilized Jacobian normalization.

Attacks

The proposed method by Szegedy, C., et al [69] applied noise to sequences to increase the approximation of the loss ($x_{FGSM} = x + \epsilon \cdot \text{sign } \nabla_x H(y, \hat{y})$), generating adversarial instances. It employs a non-linear loss function with small linear steps, iteratively using a small ρ (e.g., $\epsilon = 0.01$) to induce miss signal control to carry out attack. Additionally, the method presented by [70], introduces uniform random noise at the start, creating adversarial samples through iterative execution of the following form: $x_{PGD}^{l+1} = \Pi_{x+S}[x_{PGD}^t + \epsilon, \text{sign } \nabla_x H(y, \hat{y})]$, where Π , ϵ , and $x + S$ denote the projection operator, the gradient step, and the perturbation set, respectively. To enhance the likelihood of the ground-truth and reduce the probability of predicted labels, the proposed method presented by [71] repeatedly searches for inputs that can be subtly modified, generating instances that are difficult for humans to notice due to minimal changes in feature dimensions. To stabilize the update directions, the proposed method presented by [72] incorporates iterative attack based on the momentum, improving the transferability of adversarial instances. Similarly, the approach by [73] uses the Adabelief optimizer to create adversarial samples that resemble training examples, further enhancing the ability to transfer adversarial samples. The proposed method by [74] successfully looks forward and enhances the ability to transfer adversarial instances by using Nesterov's accelerated gradient into repeated attacks.

Defenses

To improve robustness, adversarial cases are incorporated into the training process [75]. We extend the proposed model to perform FGSM [69] on the learning phase, as implemented in [76], by computing the loss function of the model as the average of normal and adversarial losses. Gradients are prohibited from spreading in order to thwart the FGSM attack [69]. FGSM perturbations are calculated using the anticipated labels rather than the genuine labels to stop label leakage. By training an initial model with the ground truth signal control and then using the model's outputs, distillation is employed as a defensive strategy. By leveraging predicted output from the proposed model to annotate the inputs in the second training phase, class knowledge extracted by distillation is used to a different DNN architecture. The purpose of this concept is to increase DNN resistance to disturbances [77]. The predicted output of the proposed model (which we refer to as $\hat{z}_k$) within a softmax layer is divided by a temperature T as follows: $F_{T,\theta}(X_{n,k}) = \frac{e^{\hat{z}_k(X_{n,k})/T}}{\sum_{i=1}^K e^{\hat{z}_i(X_{n,k})/T}}$. When $T \rightarrow \infty$, the predicted output will converge to $1/K$. The defense according to distillation is expressed as:

$$\begin{aligned} \theta^0 &= \arg \min_{\theta} \sum_{n=1}^N \sum_{k=1}^K -Y_{n,k} \log F_{T,\theta}(X_{n,k}), \theta^* \\ &= \arg \min_{\theta} \sum_{n=1}^N \sum_{k=1}^K -F_{T,\theta^0}(X_{n,k}) \log F_{T,\theta}(X_{n,k}). \end{aligned} \quad (14)$$

Jacobian norm

This study examines the connection between linearized robustness and the regularization term based on the Jacobian norm. It also defines the robustness against adversarial perturbations using the difference between the noise-free data and the closest perturbed data that could lead to an incorrect signal control. The models's prediction is accurate when the

distance is small, as the disturbed and noise-free inputs are harder to differentiate (see Figure 6). Adversarial robustness states that $i^* = \arg \max_i F_\theta^i(x)$ corresponds to the top prediction for x as bound, while $j^* = \arg \max_{j \neq i^*} F_\theta^j(x + \epsilon z)$ refers to the highest prediction for the perturbed sample $\hat{x} = x + \epsilon z$. The $\|J(x)\|_F^2$ denotes Jacobian norm to sets an upper bound on robustness. Given a classifier F^2 for x with $\epsilon \ll 1$, we can represent the linearized robustness of $F_\theta^i(x)$ as the prediction of class i .

$$\rho(x) := \min_{j^* \neq i^*} \frac{F^{i^*}(x) - F^{j^*}(x)}{\|\nabla_x F^{i^*}(x) - \nabla_x F^{j^*}(x)\|}. \quad (15)$$

The difference in logits, $g := \nabla_x (F^{i^*} - F^{j^*})(x)$, represents the Jacobian and the input matching is $\alpha(x, g) = \langle x, g \rangle$. The decision boundary satisfies $\rho(x) \leq \frac{\alpha(x, g) + C}{\|g\|}$, where C is a positive constant. In addition, $\hat{\rho} = z^T J(x)^T J(x) z$ based on the Gaussian z calculates $\text{Tr}(J(x)^T J(x))$:

$$\mathbb{E}_z[\rho] = \frac{\text{Tr}(J(x)^T J(x) \mathbb{E}[zz^T])}{\|g\|} = \frac{\|J(x)\|_F^2}{\|g\|}. \quad (16)$$

An expectation over a mini batch X of m samples is given by $\mathbb{E}[\rho] = \mathbb{E}_x[\|J_X\|_F^2]$. Robustness against adversarial noise is defined as the minimum perturbation required to alter the model's predictions. Specifically, it measures the distance between clean input data and the smallest adversarial modification that causes the model to change its output. Also, the distance between a classifier's nearest decision boundary and its robustness at a certain location can be used to gauge the latter. However, if the locally adaptable area surrounding the x is big enough, the decision boundary is derived for model based on locally adaptable scoring functions.

Enhanced robustness with Jacobian norm: The DNN's predicted response should remain close to the accurate forecast under perturbations, expressed mathematically as $\Omega = F_\theta(x) - F_\theta(\hat{x})$. For noisy input $\hat{x}_i = x_i + \epsilon z_i$, where x_i represents the i^{th} element of the noise-free input x . It should be noted that the variable Ω quantifies the disparity between the forecasts pertaining to unaltered data and its altered counterpart. For every inference, the noise term follows a Gaussian distribution z with zero mean and variance σ^2 . Observe that the variance of the noise term is the same as that of x . In addition, the proposed robustness and the upper-bound is estimated through the prediction Jacobian with respect to input x . As a result, we can calculate the gradient with respect to the difference between the predictions, Ω , using the gradient back-propagation method as follows:

$$\begin{aligned}\nabla_x \Omega &= \nabla_x (F_\theta(x) - F_\theta(\hat{x})) \\ &= \nabla_x F_\theta(x) - \nabla_x F_\theta(\hat{x}) \Rightarrow \|\nabla_x F_\theta(x) - \nabla_x F_\theta(\hat{x})\|_F^2 \leq \|J(x)\|_F^2.\end{aligned}\tag{17}$$

The gradient based on the inputs is considered to compute the Jacobian norm, which makes a local linear estimation. However, the gradients are calculated based on the input, which is not interpretable due to noisy nature. Thus, the model is trained to minimize Jacobian norm and a gradient regularization to overcome this difficulty.

Selective input gradient regularization

Drucker et al. [45] proposed the concept of input gradient regularization, which trains the model by reducing both the loss and the rate of change of the loss in relation to the input characteristics. The cross-entropy could be used to construct the energy in the following way:

$$\begin{aligned} \theta^* = \arg \min_{\theta} & \sum_{n=1}^N \sum_{k=1}^K -Y_{n,k} \log F_{\theta}(X_{n,k}) \\ & + \lambda_m \sum_{n=1}^N \sum_{d=1}^D \left(\frac{\partial}{\partial x_d} \sum_{k=1}^K -Y_{n,k} \log F_{\theta}(X_{n,k}) \right)^2 \end{aligned} \quad (18)$$

where λ_m denotes the hyperparameter that modulates the penalty severity. This type of input gradient regularization ensures the KL divergence between true labels and predictions remains stable under small input changes, reducing sensitivity to perturbations brought on by disturbances by this twofold backpropagation. Since gradient regularization makes the model's gradients smoother, the model naturally becomes less sensitive to small changes in the input. The unification of gradient regularization with the Jacobian norm is defined as follows:

$$\begin{aligned} \theta^* = \arg \min_{\theta} & \sum_{n=1}^N \sum_{k=1}^K -Y_{n,k} \log F_{\theta}(X_{n,k}) \\ & + \lambda_m \sum_{n=1}^N \sum_{d=1}^D \left(\frac{\partial}{\partial x_d} \sum_{k=1}^K -Y_{n,k} \log F_{\theta}(X_{n,k}) \right)^2 + \lambda_j \|J(x)\|_F^2, \end{aligned} \quad (19)$$

where λ_j and λ_m denote the weights for the normalization and regularization, respectively. The proposed model's ability to make local predictions is largely dependent on the combined optimization over regularization as well as normalization, as expressed in Eq. (19). Upon encountering transferable adversarial cases designed to avoid local maxima through the accumulation of stable gradient directions, the proposed approach makes smoother gradients, ultimately stabilizing the output of the model. However, only extremely close training instances are constrained by the normalization and gradient regularization in Eq. (19). Consequently, the adversarial perturbation issue remains unsolved. Therefore, to address

these limitations and choose the most discriminative features that are both interpretable and resilient to perturbations, we propose a saliency map based on perturbations in the next section. In deep learning, saliency mapping is used for interpreting input properties that are thought to be crucial for the output of the neural network [78]. Saliency maps have been presented as a way to explain DNN decision making, as domain specialists are more interested in how interpretable a DNN could be. Gradients can be used directly to assess how much an input characteristic influences an output. However, gradient-based saliency map techniques frequently over smooth the gradients with low quality [79]. We suggest using the perturbation-based saliency map, which is created from the noisy input gradient and is represented as $\mathcal{M}_d = f(\nabla_x F_\theta(\hat{x}))$. It follows the saliency map's mission of emphasizing the significance of input characteristics. The behavior of a DNN may be highly interpreted by computing a noisy-based saliency map, which is obtained by altering the input and tracking the resulting changes. To be more precise, we mimic the input images and highlight the most noticeable areas to calculate this saliency map. In accordance with Zhang, Z., et al. [80], we create a saliency map using a Generative Adversarial Network (GAN) that resembles the original image visually. We employ a matching model (f) to translate the JN to the input domain given $J(\hat{x}) = \nabla_x F_\theta(\hat{x})$: $J' = f(J(\hat{x}))$. Thus, the classifier and the alignment network may be used to frame the generator $G(x, y)$ as follows: $G_{\theta, \phi}(x, y) = f(\nabla_x F_\theta(\hat{x}))$. Consequently, the generator may be trained to approximate the $p_{J'}$ distribution so that it resembles p_x . After obtaining $\mathcal{M}_d$, we use the salient map to compute the input gradients, allowing us to choose the most resilient gradients for back-propagation. The goal with selective input gradient regularization and Jacobian norm is described mathematically as:

$$\begin{aligned}
\theta^* = \arg \min_{\theta} & \sum_{n=1}^N \sum_{k=1}^K -Y_{n,k} \log F_{\theta}(X_{n,k}) \\
& + \lambda_m \sum_{n=1}^N \sum_{d=1}^D \left(\frac{\partial}{\partial x_d} 1(\beta - \mathcal{M}_d) \sum_{k=1}^K \right. \\
& \left. - Y_{n,k} \log F_{\theta}(X_{n,k}) \right)^2 + \lambda_j \|J(x)\|_F^2
\end{aligned} \tag{20}$$

We propose that two terms are equally important to the total optimization, therefore, unless otherwise stated, we use $\lambda_m = \lambda_j = 0.5$ as the default setup. The term $\|J(x)\|_F^2$ represents the JN. This ensures that the linearized robustness of the output is maintained in the presence of perturbed data, thereby reducing errors made by the trained model. The gradients can then be smoothed by the input gradient regularization, which is regulated by λ_m , to be noise-free. As a result, to enhance the interpretability of the proposed model, we incorporate the perturbation-based saliency map $\mathcal{M}_d$ in Eq. (20). The indicator function 1 returns 1 if $\beta - \mathcal{M}_d \geq 0$ and 0 otherwise. Algorithm 1 depicts the whole suggested method's training process.

Experiments

In this section, we present a comprehensive evaluation of our proposed integration of Multi-Agent and Group Aware Reasoning with Jacobian Norm-Based Adversarial Defense for robust traffic signal control. Our experimental design encompasses diverse real-world traffic scenarios, which evaluates the system's adaptability to varying levels of congestion and examines its robustness against adversarial attacks. Through extensive comparative analyses against state-of-the-art traffic signal control methods, we validate the effectiveness of our approach in optimizing urban traffic flows while ensuring resilience against perturbations. To evaluate the effectiveness of our approach, we conduct simulations using

the CityFlow simulator, a high-performance traffic simulation environment widely used in intelligent transportation research. CityFlow enables large-scale and realistic simulations of urban road networks, providing accurate insights into traffic dynamics under different control policies. The road networks for our experiments are derived from real-world urban traffic maps, ensuring their applicability to practical deployment scenarios. The traffic demand levels and intersection configurations vary across different datasets to comprehensively assess the adaptability of our proposed method. Our experiments incorporate multiple configurations to examine the impact of traffic density variations, control responsiveness, and scalability across networks of different sizes. The primary simulation parameters include a total simulation time of 3600 seconds per episode to capture long-term traffic dynamics, with decision intervals set at 5 seconds per step to allow for real-time adaptive signal control. A yellow light duration of 2 seconds is used before transitions to ensure safety, while a homogeneous policy is applied across all intersections to maintain consistency in evaluation. Each experiment is run for multiple trials, and the results are averaged to minimize stochastic variability. For a thorough comparative study, we benchmark our approach against several existing methods, including Fixed-Time [81] Control, which follows pre-defined phase timing based on historical traffic data, and Max-Pressure (MP) Control, a reactive mechanism that dynamically adjusts signals based on local queue pressure. We also compare against reinforcement learning-based methods such as PressLight, which optimizes phase selection through reward-driven learning, CoLight, a multi-agent reinforcement learning method leveraging graph attention networks, and MPLight, a policy optimization framework designed for adaptive traffic control. Finally, our proposed method integrates multi-agent reasoning with Jacobian norm-based adversarial defense to enhance both efficiency and robustness.

Evaluation Metrics

To quantitatively assess the performance of our method, we utilize key performance indicators widely adopted in traffic signal control research. The average travel time measures the mean duration taken by vehicles to traverse the network, with lower values indicating improved traffic efficiency. The Queue Length (QL) represents the average number of stopped vehicles at each intersection, where a reduction signifies decreased congestion. Additionally, we evaluate the Robustness to Adversarial Attacks, which measures the model’s ability to maintain stable performance under perturbations, including sensor noise and manipulated vehicle flow data.

Traffic Datasets

To ensure a rigorous evaluation, we employ three well-established real-world traffic datasets from different urban environments: Jinan (China), Hangzhou (China), and New York City (USA) (Table 1). The Jinan dataset consists of a 3×4 grid of 12 intersections with three different levels of traffic demand. The Hangzhou dataset features a 4×4 grid of 16 intersections with two distinct traffic demand levels. The New York City dataset represents a large-scale urban traffic scenario with 196 intersections arranged in a 28×7 grid, reflecting varying congestion patterns throughout the day. These datasets provide real-time vehicle trajectories and historical flow data collected from urban traffic monitoring systems. They capture a wide range of traffic conditions, from moderate flow scenarios to severe congestion during peak hours. By utilizing these datasets, we ensure that our model is evaluated under conditions that reflect real-world complexities, allowing us to measure its effectiveness in both localized and network-wide signal control optimization.

Table 1: An Overview of Traffic Datasets with Different Traffic Demands.

Network	Type	Volume (Veh)	Arrival Rate (Veh/min)			
			Mean	Std.	Max	Min
Jinan (3×4)	S1	6295	104.92	19.79	136.00	50.00
	S2	4365	72.75	15.15	101.00	43.00
	S3	5494	91.57	9.51	111.00	69.00
Hangzhou (4×4)	S4	2983	49.72	8.24	67.00	40.00
	S5	6984	116.40	63.72	230.00	39.00
New York (7×28)	S6	10676	177.92	27.06	222.00	79.00
	S7	15862	264.35	35.10	320.00	79.00

Baseline Methods

We compare the policies learned by AGT-Net to a variety of traditional and learning-based baseline techniques. For all learning-based approaches, we meticulously train a particular strategy for each traffic demand within distinct road network, in accordance with latest studies in the area. Following each training, we assess the policies using ten assessment events with various seeds and present the average vehicle journey duration. Although we performed the majority of the baselines in our setup, certain data were not obtained.

- ***Fixed-Time Control***: A traditional rule-based approach that follows pre-defined phase schedules based on historical traffic data. While simple and computationally efficient, it lacks adaptability to real-time traffic fluctuations, making it ineffective in dynamic urban environments, especially under congestion.
- ***Max-Pressure (MP) Control***: A reactive control method that adjusts signal timings based on queue pressure at intersections. It dynamically redistributes green time to lanes with the highest pressure, improving efficiency over fixed-time methods.

However, its localized decision-making fails to capture long-range traffic dependencies, making it suboptimal in large, interconnected networks.

- ***Advanced Max-Pressure (Advanced-MP)***: An enhanced version of MP that incorporates multi-step lookahead and improved flow estimation. It performs better in mitigating congestion but still struggles with scalability and lacks multi-agent coordination, limiting its performance in complex urban scenarios.
- ***CoLight***: A deep reinforcement learning (RL)-based approach that utilizes graph attention networks (GATs) to facilitate communication between neighboring intersections. It significantly improves efficiency in moderate congestion levels but fails to generalize well in high-density networks due to its limited interaction modeling.
- ***MPLight***: An RL-based traffic signal control method that enhances policy learning using multi-agent proximal policy optimization (MAPPO). While it offers better adaptability than CoLight, its performance deteriorates in extreme traffic conditions and remains vulnerable to adversarial perturbations.
- ***Advanced-CoLight***: A refinement of CoLight that incorporates better cooperative learning mechanisms between intersections, improving performance in large-scale networks. However, it still lacks an explicit adversarial defense mechanism, making it susceptible to sensor noise and input perturbations.
- ***Advanced-MPLight***: An improved version of MPLight that enhances cooperative learning and long-term policy optimization. Despite offering better traffic flow efficiency, it faces stability issues in adversarial settings, leading to erratic phase decisions when sensor inputs are manipulated.
- ***DenseLight***: A spatial-temporal deep learning model that optimizes signal control using dense multi-agent coordination. It significantly reduces congestion in mid-

sized urban networks but struggles with computational scalability in large-scale networks.

- ***SocialLight***: A distributed cooperation learning model that enables network-wide traffic signal optimization. It shows strong performance in coordinated traffic control but lacks robustness in high-adversarial scenarios, leading to higher variance in performance under noisy conditions.
- ***CoordLight***: The strongest graph-based cooperative RL model before our approach, achieving low travel times by integrating multi-hop inter-agent communication. However, it lacks explicit adversarial defenses, making it vulnerable to sensor manipulations and perturbations.

Results and Discussion

Table 2 presents a detailed performance comparison of different traffic signal control methods across various urban traffic scenarios. The comparison is based on key performance indicators, including average travel time and queue length, which are critical in evaluating traffic efficiency. The results highlight the superior performance of our proposed Multi-Agent and Group Aware Reasoning with Jacobian Norm-Based Adversarial Defense (Ours) method in optimizing urban traffic flow while ensuring robustness against adversarial perturbations. Traditional methods such as Fixed-Time Control, show the weakest performance across all scenarios due to their inability to dynamically adjust to real-time traffic variations. In particular, S6 and S7, representing high-density urban networks, experience extreme travel times of 1507.12s and 1660.29s, respectively. Similarly, queue length remains excessively high across all scenarios, confirming the inefficiency of this rigid, non-adaptive approach. Max-Pressure (MP) Control performs notably better than Fixed-

Time Control by dynamically adjusting traffic signals based on queue pressure. However, despite lowering ATT in S1–S5, it struggles in large-scale networks, with ATT values of 1179.55s (S6) and 1535.77s (S7). The Advanced Max-Pressure (Advanced-MP) variant further modifies traffic signal adjustments, achieving the best results among rule-based methods, with ATT reductions of up to 10% over standard MP. However, its performance deteriorates in large urban settings due to its localized optimization strategy. Deep reinforcement learning-based approaches, including CoLight and MPLight, introduce multi-agent learning to optimize traffic flow. CoLight achieves significant reductions in ATT compared to traditional methods, particularly in S1–S3, with ATT values of 276.33s, 237.14s, and 278.16s, respectively. However, it still exhibits high variability in performance under congestion, as seen in S6 (1221.77s) and S7 (1476.18s). MPLight, which incorporates policy optimization, performs slightly better, particularly in S3 (261.45s) and S5 (282.14s). However, its performance remains inconsistent in highly congested environments. The Advanced-CoLight and Advanced-MPLight variants improve upon their base models by introducing better cooperative learning strategies, resulting in lower ATT and QL values across most scenarios. For example, Advanced-CoLight achieves ATT values of 1025.47s (S6) and 1341.13s (S7), outperforming standard CoLight by 8–10%, while Advanced-MPLight achieves a similar trend but lags slightly in performance under heavy congestion. Recent state-of-the-art cooperative and graph-based models, such as DenseLight, SocialLight, and CoordLight [82], further enhance traffic efficiency by incorporating advanced inter-agent communication strategies. DenseLight, which applies spatial-temporal learning for traffic control, significantly outperforms all previous methods, particularly in S1–S5, where it achieves ATT reductions of over 15% compared to Advanced-CoLight. However, its performance in S6 and S7 (803.42s and 1106.69s, respectively), while improved, is still affected by scalability limitations. SocialLight, which leverages distributed

cooperation learning, performs slightly better than DenseLight in some cases, particularly in S1 (217.92s) and S2 (211.75s), but struggles in high-density networks (S6: 771.92s, S7: 1108.44s). Among these approaches, CoordLight [82] achieves the best overall results, reducing ATT to 748.32s (S6) and 1039.15s (S7), making it the strongest baseline prior to our proposed method. Our proposed approach outperforms all existing methods by achieving the lowest ATT across all scenarios, demonstrating both scalability and robustness. Compared to CoordLight [82], our model reduces ATT by 9.2% (S1), 6.2% (S2), 5.4% (S3), 9.3% (S6), and 4.5% (S7), with a corresponding reduction in QL, indicating more efficient traffic management. In particular, our method achieves ATT values of 704.30s (S6) and 992.26s (S7), outperforming CoordLight [82] by a significant margin in high-congestion settings. These improvements stem from the integration of Jacobian norm-based adversarial defense, which enhances the robustness of learned policies against sensor noise and adversarial manipulations. Unlike conventional reinforcement learning models that degrade under adversarial conditions, our method maintains stable performance, ensuring both efficiency and resilience in real-world deployment. In summary, Fixed-Time and MP-based methods are ineffective in dynamic environments, deep reinforcement learning-based approaches improve adaptability but struggle with adversarial robustness, and graph-based cooperative models enhance coordination but exhibit limitations in extreme congestion scenarios. Our Multi-Agent and Group Aware Reasoning with Jacobian Norm-Based Adversarial Defense achieves the best overall performance, combining the lowest ATT, shortest queue lengths, and highest resilience against adversarial attacks, setting a new benchmark for intelligent traffic signal control systems.

Table 2: Average Travel Times (Sec) Comparing performance In Each Scenarios based on the Table 1.

Methods	Scenarios	Average Travel Time ↓	Queue Length ↓
Fixed-Time [81]	S1	346.36	35.6
	S2	289.74	31.8
	S3	316.7	33.5
	S4	432.32	39.2
	S5	359.44	35.2
	S6	1507.12	53.7
	S7	1660.29	68.8
MaxPressure [83]	S1	273.96	25.1
	S2	245.38	23.2
	S3	245.81	23.8
	S4	288.54	28.0
	S5	348.98	32.2
	S6	1179.55	42.6
	S7	1535.77	61.7
Advanced-MP [84]	S1	253.61	25.3
	S2	238.62	21.1
	S3	235.21	23.5
	S4	279.47	28.3
	S5	318.67	31.2
	S6	1060.41	48.7
	S7	1371.52	58.2
CoLight [40]	S1	276.33+-3.98	30.4
	S2	237.14+-0.73	28.7
	S3	278.16+-1.13	30.5
	S4	271.07+-2.11	28.5
	S5	297.26+-9.47	31.7
	S6	1221.77+-19.83	49.9
	S7	1476.18+-29.22	58.6
MPLight [85]	S1	300.93+-3.57	31.5
	S2	259.10+-8.90	29.1
	S3	261.45+-1.74	30.3
	S4	343.47+-2.89	33.5
	S5	282.14+-7.20	30.3
	S6	1168.49+-23.99	46.4
	S7	1597.24+-18.57	60.1
Advanced-CoLight [84]	S1	253.95+-0.77	28.4
	S2	234.03+-0.61	26.9
	S3	244.51+-0.66	27.5
	S4	269.62+-0.51	28.5
	S5	308.62+-4.58	32.7
	S6	1025.47+-34.33	44.2
	S7	1341.13+-14.33	56.9
	S1	250.04+-0.65	28.2

Advanced-MPLight [84]	S2	230.92+-0.36	26.5
	S3	227.72+-0.58	25.8
	S4	285.28+-0.41	30.5
	S5	296.76+-4.59	30.7
	S6	1304.60+-18.12	52.2
	S7	1604.48+-22.34	66.9
DenseLight [86]	S1	226.97	24.5
	S2	215.82	22.1
	S3	239.58	25.3
	S4	248.43	26.5
	S5	272.27	27.3
	S6	803.42	38.4
	S7	1106.69	49.1
SocialLight [87]	S1	217.92+-0.40	23.4
	S2	211.75+-0.40	22.9
	S3	210.46+-0.32	22.5
	S4	254.86+-0.23	27.5
	S5	288.55+-0.65	30.7
	S6	771.92+-5.04	37.2
	S7	1108.44+-15.39	49.7
CoordLight [82]	S1	199.24+-0.20	21.2
	S2	198.21+-0.11	20.5
	S3	191.05+-0.01	19.7
	S4	248.45+-0.07	26.5
	S5	250.87+-0.21	26.7
	S6	748.32+-4.60	39.2
	S7	1039.15+-10.05	47.9
Ours	S1	190.04+-0.65	19.5
	S2	185.92+-0.36	19.1
	S3	180.72+-0.58	17.3
	S4	225.28+-0.41	23.5
	S5	216.16+-4.59	22.3
	S6	704.30+-8.12	35.4
	S7	992.26+-5.03	44.1

Ablation Studies

To better understand the contributions of different components in our proposed method, we conducted extensive ablation studies focusing on three key aspects: (i) the number of agents in each scenario (Table 3), (ii) the feature extraction and integration techniques (Table 4), and (iii) the impact of different modules, including multi-sampling (MS) and modified

prediction (MP) (Table 5). The results demonstrate the significance of each component in achieving robust and interpretable traffic signal control.

Impact of the Number of Agents in Each Scenario

Table 3 examines the effect of varying the number of agents involved in the control process across different levels of communication. Each level consists of a structured hierarchy of agent groups, such as (1,2,4,8), meaning there are 4 levels where each level contains 1, 2, 4, and 8 agents, respectively. The results indicate a clear trend: increasing the number of agents in a structured manner significantly enhances performance in terms of average travel time. For example, in Scenario S6, the ATT decreases from 1207.19s at level 1 to 1023.01s at level (1,2,4), demonstrating the effectiveness of multi-level communication. The best performance is achieved at the (1,2,4,8,12) level, where ATT is minimized across all scenarios (e.g., 874.65s in S6 and 1151.36s in S7), highlighting the optimal grouping strategy for multi-agent control. However, further increasing the agent levels to (1,2,4,8,16) and (1,2,4,8,16,196) slightly degrades performance in low-density scenarios like S1 and S2 (e.g., ATT increases from 190.04s to 202.69s in S1). This suggests that excessive agent clustering introduces communication overhead, leading to diminishing returns. However, in high-density networks like S6 and S7, the additional agents provide substantial benefits, further reducing ATT to 704.30s and 992.26s, respectively. This emphasizes the need for a balanced grouping strategy that maximizes coordination without incurring excessive computational overhead.

Table 3: Effect of number of agents in each scenario for each group in term of ATT metric.

Scenarios Level	S1	S2	S3	S4	S5	S6	S7
1	239.28	237.91	234.32	289.01	281.18	1207.19	1602.57
1,2	229.00	220.11	213.80	280.37	269.72	1138.31	1460.92
1,2,4	216.48	204.54	199.01	267.45	255.69	1023.01	1354.41
1,2,4,8	199.81	194.69	187.77	253.55	247.43	981.80	1209.81
1,2,4,8,12	190.04	185.92	180.72	241.01	229.11	874.65	1151.36
1,2,4,8,16	202.69	192.33	188.43	225.28	216.16	812.42	1024.40
1,2,4,8,16,196	209.56	201.02	195.56	238.55	232.47	704.30	992.26

Effect of Feature Extraction and Integration

Feature extraction plays a critical role in the effectiveness of our model. Table 4 compares different methods, including Multi-Layer Perceptron (MLP), Non-Local Attention (NLT), and element-wise summation (EWS), against our proposed approach. The results demonstrate that simple MLP-based feature extraction leads to suboptimal performance, as seen in the relatively high ATT values. For instance, in Scenario S5, the ATT with MLP is 258.43s, significantly higher than our approach at 216.16s. The NLT mechanism improves upon MLP by introducing global attention, allowing the model to capture long-range dependencies. However, its effectiveness varies across different scenarios, particularly in highly dynamic urban settings where real-time coordination is essential. The EWS approach fails to effectively integrate multi-agent interactions, leading to increased ATT in some scenarios (e.g., 1156.35s in S6, compared to 704.30s with our method). Our proposed approach significantly outperforms all alternatives, consistently achieving the lowest ATT

values across all scenarios. This demonstrates the effectiveness of our advanced feature extraction and integration mechanism, which captures both spatial and temporal dependencies while preserving interpretability. The results reinforce the necessity of integrating structured feature learning techniques to enhance traffic signal control.

Table 4: Effect of feature extraction and integration in each scenario in term of ATT metric.

Scenarios Method	S1	S2	S3	S4	S5	S6	S7
MLP	228.14	219.80	202.69	292.33	258.43	1170.34	1522.11
(None-local attention) NLT	218.02	204.40	199.66	257.45	285.63	1044.12	1124.76
element-wise summation (EWS)	229.71	234.32	244.73	280.37	269.72	1156.35	1230.79
Ours	190.04	185.92	180.72	225.28	216.16	704.30	992.26

Impact of Different Modules

To further investigate the contributions of different modules, Table 5 presents an ablation study on the impact of Multi-Sampling (MS) and Modified Prediction (MP). The baseline Gaussian Mixture Model (GMM) without MS or MP exhibits the highest ATT values, confirming that a simple probabilistic approach fails to capture the underlying complexity of traffic dynamics. For instance, in Scenario S6, GMM alone results in an ATT of 1142.14s. Introducing MS improves performance by allowing the model to generate multiple future predictions, enhancing its ability to handle uncertainties in traffic flow. The reduction in ATT across all scenarios highlights the benefit of sampling diverse possible traffic signal

transitions rather than relying on a single deterministic prediction (e.g., S6 ATT decreases to 974.42s with MS enabled). Adding MP further modifies these predictions by smoothing signal inconsistencies and reducing stochastic errors. The best overall performance is achieved when both MS and MP are incorporated, resulting in the lowest ATT values across all scenarios (e.g., ATT in S6 further drops to 704.30s). This confirms that an ensemble-based approach, coupled with modifies traffic signal is essential for maintaining stability and efficiency in real-world traffic signal control.

Table 5: Effect of different modules in each scenario in term of ATT metric (Multi-sampling (MS), Modified Prediction (MP)).

Scenarios			S1	S2	S3	S4	S5	S6	S7
Module									
GMM	MS	MP							
x	x	x	244.40	241.66	259.00	290.81	278.43	1142.14	1417.15
✓	x	x	234.32	222.19	226.48	264.19	245.63	974.42	1163.36
✓	✓	x	209.71	200.12	204.73	243.23	229.72	864.35	1030.12
✓	✓	✓	190.04	185.92	180.72	225.28	216.16	704.30	992.26

The results from Table 3, Table 4, and Table 5 highlight the interplay between agent coordination, feature extraction, and prediction enhancement techniques in achieving robust traffic control. Increasing the number of agents systematically improves performance, but excessive clustering introduces overhead, necessitating an optimal balance. The best results occur when structured multi-agent cooperation is complemented by advanced feature extraction mechanisms. For instance, in Scenario S6, ATT decreases from 1207.19s with a single agent level to 704.30s with our full method, showcasing the combined effect of

structured agent coordination and feature learning. Similarly, feature extraction alone is insufficient; even with optimal agent grouping, subpar prediction techniques lead to higher ATT values. The integration of Multi-Sampling and Prediction modified addresses uncertainties in traffic flow, ensuring robust decision-making and yielding significant ATT reductions across all scenarios. The highest gains are observed when agent coordination, feature extraction, and predictive enhancements are all combined, reinforcing the need for a holistic approach to intelligent traffic signal control. Overall, the ablation study results demonstrate the critical importance of structured multi-agent grouping, advanced feature extraction mechanisms, and robust prediction strategies. Our findings confirm that optimal agent selection, coupled with effective learning techniques and modified modules, significantly enhances the resilience and efficiency of traffic signal control systems. By integrating these components, our proposed method consistently achieves superior performance, setting a new benchmark for intelligent traffic management in dynamic urban environments.

Robust Traffic Signal Control under Adversarial Perturbations

Table 6 presents a detailed comparison of the proposed Jacobian norm-based adversarial defense against state-of-the-art methods in traffic signal control when subjected to perturbed data generated via Projected Gradient Descent (PGD). The results demonstrate the effectiveness of our approach in maintaining robust and interpretable traffic control policies under adversarial conditions. Unlike traditional deep learning-based methods that exhibit high sensitivity to input perturbations, our approach significantly reduces performance degradation by incorporating Jacobian-based regularization, ensuring that small variations in sensor data do not lead to erratic or suboptimal traffic signal decisions. The results highlight that our Jacobian norm-based method consistently achieves the lowest ATT across all scenarios, significantly outperforming alternative approaches. For instance, in Scenario S6,

where congestion is most severe, our method achieves an ATT of 711.61s, whereas Deep Perturbation (DP) [81] records 1381.49s, Adversarial Bayesian Neural Networks (Adv-BNN) [82] reports 1278.71s, and Jacobian Adversarially Regularized Networks (JARN) [83] reaches 1306.20s. These results indicate that standard deep learning and Bayesian neural network-based methods struggle to maintain stable performance under adversarial conditions, leading to increased congestion and inefficient traffic flow. The vulnerability of these methods arises from their high sensitivity to adversarial noise, which causes disruptions in traffic signal timing and results in a cascading effect of delays across intersections. Among competing methods, Deep Perturbation (DP) shows the worst performance due to its lack of explicit adversarial defense, making it highly susceptible to manipulated sensor inputs. While Adv-BNN slightly improves robustness by incorporating Bayesian uncertainty estimation, it still underperforms in high-density networks, as it lacks mechanisms to mitigate structured adversarial attacks effectively. JARN, which integrates Jacobian-based regularization, provides some robustness benefits but lacks group-aware reasoning and adaptive multi-agent coordination, which are crucial for maintaining optimal control in large-scale urban environments. In contrast, our method integrates Jacobian norm regularization with MARL, ensuring both adversarial robustness and traffic efficiency.

The key factor behind our method’s superior performance is its ability to minimize the norm of the input-output Jacobian, thereby constraining the sensitivity of traffic control decisions to small perturbations. This Jacobian-based gradient regularization flattens decision boundaries, making the model more resistant to adversarial perturbations. Additionally, selective input gradient regularization further enhances stability, ensuring that the model maintains robust and consistent decisions even in highly dynamic environments. Unlike standard MARL approaches that degrade under adversarial conditions, our method exhibits a consistent ATT improvement across all scenarios, with reductions of 49.6% in S6

and 42.3% in S7 compared to DP, highlighting its effectiveness in high-traffic-density settings. Overall, these findings confirm that Jacobian norm-based adversarial defense plays a crucial role in ensuring the robustness of multi-agent traffic signal control systems. By integrating multi-agent cooperation, group-aware reasoning, and adversarial robustness, our approach sets a new benchmark for resilient and efficient traffic management, outperforming existing solutions in both adversarial and standard traffic conditions.

Table 6: Effect of the proposed method on perturbed data by projected gradient descent in each scenario in term of ATT metric to compare with state of the art method.

Scenarios Method	S1	S2	S3	S4	S5	S6	S7
DP [88]	246.44	237.65	241.54	322.62	289.03	1381.49	1422.21
Adv-BNN[89]	238.51	226.48	232.60	289.37	315.75	1278.71	1224.46
JARN [90]	239.84	255.72	276.03	301.40	292.20	1306.20	1181.92
Jacobian norm (Ours)	195.45	191.17	184.92	230.44	222.72	711.61	997.57

Discussion

The results of our proposed Multi-Agent and Group Aware Reasoning framework with Jacobian Norm-Based Adversarial Defense demonstrate significant improvements in traffic efficiency, coordination, and robustness against adversarial perturbations. However, beyond performance metrics, it is essential to examine the practical implications, implementation challenges, and the reasoning behind key design choices. Each module in our approach—AGT-Encoder, Decoder, Modified Prediction Module, Jacobian Norm Regularization, Multi-Agent Learning, and Group Aware Reasoning, was carefully designed to address fundamental challenges in real-world traffic signal control. Additionally, the choice of loss functions was motivated by the need to balance prediction accuracy, robustness, and adaptability in dynamic urban environments. Traditional multi-agent reinforcement learning

(MARL) methods often treat each intersection as an independent agent, leading to suboptimal decision-making due to a lack of coordination. This results in localized congestion, inefficient green waves, and delays spreading across networks. To overcome this limitation, we introduced the Adaptive Group Traffic Network (AGT-Net) Encoder, which dynamically infers a cross-scale super-graph structure, allowing agents to communicate effectively while adapting to real-time congestion patterns. Unlike conventional MARL models that rely on fixed neighborhood interactions, our approach enables dynamic grouping of intersections based on current traffic states, ensuring that highly interdependent intersections can share decision-making responsibilities. This group-aware reasoning is inspired by real-world traffic management practices, where synchronized signals on major roads improve flow efficiency while localized control adapts to real-time conditions on secondary roads. The result is a more scalable and adaptable coordination strategy that balances local responsiveness with global optimization.

Traffic flow prediction remains a critical challenge due to the inherent stochasticity of vehicle movements. Reactive models that adjust signals based solely on current traffic states often fail to anticipate congestion buildup before it occurs, leading to inefficient phase transitions. To address this, our framework integrates a Decoder with a Modified Prediction Module, designed to improve forecasting accuracy through a Conditional Variational Autoencoder (CVAE)-based approach. Unlike deterministic models that produce single-output predictions, our method generates multiple future traffic scenarios, allowing the system to adaptively choose the most efficient traffic signal transitions based on probabilistic forecasting. Additionally, our Modified Prediction Module modified these forecasts by incorporating signal smoothing techniques, preventing sudden signal shifts that could disrupt vehicle flow. This ensures that predicted traffic states remain consistent over multiple time steps, enhancing both stability and efficiency in real-world deployment.

A key aspect of intelligent traffic signal control is robustness against adversarial attacks and sensor noise, which can degrade performance by manipulating vehicle count inputs or injecting perturbations into sensor data. Many existing MARL-based methods are highly sensitive to adversarial perturbations, causing unpredictable phase transitions, congestion spikes, and increased travel times. Our approach addresses this vulnerability through Jacobian Norm-Based Adversarial Defense, which stabilizes decision-making by reducing the sensitivity of model outputs to small input perturbations. By incorporating Jacobian-based gradient regularization, our method effectively flattens decision boundaries, ensuring that adversarial manipulations do not lead to disruptive signal phase adjustments. The experimental results confirm that our model maintains stable traffic control even under adversarial attacks, with ATT improvements of 49.6% in S6 compared to Deep Perturbation (DP). This resilience is critical for real-world applications, where sensor anomalies or malicious interference could otherwise compromise urban traffic networks. The selection of loss functions in our model was driven by the need to balance robustness and prediction accuracy. The Negative Log-Likelihood (NLL) loss was used to ensure that our model could effectively learn the multimodal nature of traffic flow, preventing overfitting to a single predicted outcome. The Kullback-Leibler (KL) Divergence Loss was integrated to enforce distribution alignment, ensuring that future predictions remain consistent with real-world traffic patterns. Additionally, the Jacobian Norm Regularization Loss plays a crucial role in enhancing model interpretability and adversarial robustness by reducing output sensitivity to perturbations. To further refine predictions, we introduced a modified Prediction Loss, which corrects inconsistencies in multi-step forecasting, ensuring that traffic signal phase transitions remain smooth and continuous over time. Together, these loss functions allow the system to learn stable, generalizable, and interpretable traffic control policies that maintain efficiency in both standard and adversarial conditions. Despite the strong performance of our

approach, real-world deployment presents several challenges. Computational complexity remains a concern, as our graph-based multi-agent coordination mechanism requires efficient optimization for real-time decision-making. While our model is significantly more scalable than centralized control systems, further optimization is needed to ensure low-latency inference in large-scale urban environments. Another challenge lies in multi-agent communication delays, where real-world intersections may not always have synchronized data exchange due to network latency or sensor failures. Future work should explore asynchronous coordination strategies to ensure that agents can make resilient decisions even in the presence of delayed or incomplete information. Additionally, while our Jacobian-based adversarial defense significantly enhances robustness, real-world adversarial threats may be more sophisticated than those simulated in our experiments. Attackers could employ adaptive strategies that evolve over time, requiring continuous adversarial training and reinforcement learning updates to maintain long-term resilience. Integrating real-time feedback loops from smart infrastructure sensors could further enhance our model's ability to detect and adapt to evolving threats, ensuring sustained robustness in deployment.

In conclusion, our proposed framework represents a major advancement in intelligent traffic signal control, successfully integrating multi-agent coordination, group-aware reasoning, probabilistic forecasting, and adversarial robustness into a unified system. The results demonstrate that our approach not only improves travel efficiency but also ensures stability under adversarial conditions, setting a new benchmark for resilient and interpretable traffic management. Future research should focus on further reducing computational overhead, improving communication robustness, and expanding adversarial defense strategies to enhance the real-world applicability of intelligent transportation systems.

Limitations

Despite the strong performance and robustness of our proposed Multi-Agent and Group Aware Reasoning framework with Jacobian Norm-Based Adversarial Defense, several limitations must be considered for real-world deployment. These limitations highlight potential challenges in scalability, computational efficiency, real-world adaptability, and adversarial defense generalization, which future research can address. One major limitation is the computational complexity introduced by multi-agent coordination and graph-based reasoning. While our method is significantly more scalable than centralized traffic control systems, the dynamic clustering of agents and real-time super-graph inference require substantial computational resources. In large-scale urban networks with hundreds of intersections, the overhead of real-time message passing and adaptive decision-making could lead to delays in signal updates, potentially reducing responsiveness in fast-changing traffic conditions. Optimizing the model for real-time inference on edge devices or distributed processing architectures is essential for practical deployment.

Another challenge is the communication overhead in multi-agent environments. Our approach assumes synchronized and reliable data exchange between traffic intersections, but in real-world implementations, network delays or sensor failures could disrupt information sharing, leading to inconsistencies in signal coordination. While our model is designed to handle localized disturbances, extreme cases of data loss or delayed communication could degrade system performance. Future work should explore asynchronous decision-making mechanisms and decentralized learning strategies to improve robustness in unreliable communication environments. While our Jacobian norm-based adversarial defense significantly enhances model robustness, it may not generalize perfectly to highly adaptive and intelligent adversarial attacks. Our experiments demonstrate strong resilience against Projected Gradient Descent (PGD)-based perturbations, but real-world attackers could

design more sophisticated attack strategies, such as adaptive perturbations that evolve over time or coordinated adversarial attacks across multiple intersections. Implementing real-time adversarial detection mechanisms and adaptive model retraining could help counteract evolving threats in future intelligent transportation systems. Finally, our model relies on accurate traffic sensor data, but real-world urban environments often experience data inconsistencies, sensor malfunctions, and missing observations due to environmental factors or infrastructure limitations. Although our approach includes multi-sampling and prediction modified techniques to handle uncertainty, severe data quality issues could still impact model accuracy. Integrating multi-modal traffic data sources, such as camera-based vehicle detection, GPS trajectories, and connected vehicle data, could improve the reliability and robustness of our system in real-world settings.

In summary, while our proposed framework introduces state-of-the-art efficiency, robustness, and coordination for traffic signal control, addressing these limitations in computational scalability, real-world communication constraints, adversarial adaptation, and data reliability will be critical for future advancements in intelligent transportation systems.

Conclusion

In this paper, we introduced a novel Multi-Agent and Group Aware Reasoning framework with Jacobian Norm-Based Adversarial Defense to enhance the robustness and efficiency of traffic signal control in adversarial environments. By combining structured multi-agent coordination with graph-based reasoning and Jacobian norm regularization, our approach addresses key limitations of existing MARL methods, particularly in dynamic and uncertain urban settings. Experimental evaluations across multiple urban traffic scenarios demonstrate that our method consistently reduces average travel time and queue length, outperforming both traditional baselines and adversarially robust models such as Deep

Perturbation and Adv-BNN. Notably, under PGD-based attacks, our system maintains performance levels where others degrade significantly, validating its resilience. Ablation studies further highlight the contributions of hierarchical agent grouping, feature extraction modules, and the combined use of multi-sampling with prediction modification, achieving up to 38.3% improvement in high-congestion scenarios. These components collectively ensure scalable, interpretable, and stable control policies under varying traffic conditions.

Despite its effectiveness, the framework still faces challenges related to computational scalability, communication constraints, and the generalization of adversarial defenses. Addressing these limitations through distributed learning, asynchronous control, and integration of multi-modal real-world sensor data remains a promising direction for future research. Overall, this study establishes a new standard for resilient and intelligent multi-agent traffic control, demonstrating that adversarially robust MARL systems are essential for reliable deployment in real-world intelligent transportation systems.

References

- [1] A. K. Azad, T. Atkison, and A. Shah, "A Review on Machine Learning in Intelligent Transportation Systems Applications," *The Open Transportation Journal*, vol. 18, no. 1, 2024.
- [2] Z. Wang, S. Cong, and S. Cheng, "Fixed Point Learning Based Intelligent Traffic Control System," *IOP Conference Series: Materials Science and Engineering*, vol. 261, 2017.
- [3] X. Tan, Y. Zhou, and X. Jiao, "Traffic signal control based on deep reinforcement learning using state fusion and trend reward," *Engineering Applications of Artificial Intelligence*, vol. 159, p. 111701, 2025/11/15/ 2025, doi: <https://doi.org/10.1016/j.engappai.2025.111701>.
- [4] T. T. Nguyen, N. D. Nguyen, and S. Nahavandi, "Deep Reinforcement Learning for Multiagent Systems: A Review of Challenges, Solutions, and Applications," *IEEE Transactions on Cybernetics*, vol. 50, pp. 3826-3839, 2018.
- [5] Y. Xu, Y. Zhang, and M. Liu, "Multiagent Based Decentralized Traffic Light Control for Large Urban Transportation System," *Math. Probl. Eng.*, vol. 2014, pp. 1-13, 2014.
- [6] M. Abdoos, "Robustness Analysis of Holonic Multi-Agent Systems: Application to Traffic Signals Control," *Tehnicki vjesnik - Technical Gazette*, 2018.
- [7] P. Rani and D. Kumar, "Multi-Agent Based Traffic Signal Optimization under Fully Actuated Control," *Int. J. Innov. Technol. Explor. Eng.*, 2019.
- [8] H. Wang, H. Chen, Q. Wu, C. Ma, and Y. Li, "Multi-Intersection Traffic Optimisation: A Benchmark Dataset and a Strong Baseline," *IEEE Open Journal of Intelligent Transportation Systems*, vol. 3, pp. 126-136, 2021.
- [9] H. Wei, G. Zheng, V. Gayah, and Z. Li, "A survey on traffic signal control methods," *arXiv preprint arXiv:1904.08117*, 2019.
- [10] C. Wang, Y. Li, J. Chen, J. Zhang, and Y. Xue, "Cooperative traffic signal control for a partially observed vehicular network using multi-agent reinforcement learning," *Engineering Applications of Artificial Intelligence*, vol. 160, p. 111813, 2025/11/15/ 2025, doi: <https://doi.org/10.1016/j.engappai.2025.111813>.

- [11] Y. Xiang *et al.*, "Congestion attack detection in intelligent traffic signal system: Combining empirical and analytical methods," *Security and Communication Networks*, vol. 2021, no. 1, p. 1632825, 2021.
- [12] K. L. Tan, S. Poddar, A. Sharma, and S. Sarkar, "Deep Reinforcement Learning for Adaptive Traffic Signal Control," *ArXiv*, vol. abs/1911.06294, 2019.
- [13] S. Bouktif, A. Cheniki, A. Ouni, and H. El-Sayed, "Parameterized-action based deep reinforcement learning for intelligent traffic signal control," *Engineering Applications of Artificial Intelligence*, vol. 159, p. 111422, 2025/11/01/ 2025, doi: <https://doi.org/10.1016/j.engappai.2025.111422>.
- [14] D. Jakubovitz and R. Giryes, "Improving dnn robustness to adversarial attacks using jacobian regularization," in *Proceedings of the European conference on computer vision (ECCV)*, 2018, pp. 514-529.
- [15] A. Noack, I. Ahern, D. Dou, and B. A. Li, "Does Interpretability of Neural Networks Imply Adversarial Robustness?," *ArXiv*, vol. abs/1912.03430, 2019.
- [16] B. M. Le, S. Tariq, and S. S. Woo, "Bridging Optimal Transport and Jacobian Regularization by Optimal Trajectory for Enhanced Adversarial Defense," in *Asian Conference on Computer Vision*, 2023.
- [17] C. Finlay, A. M. Oberman, and B. Abbasi, "Improved robustness to adversarial examples using Lipschitz regularization of the loss," *ArXiv*, vol. abs/1810.00953, 2018.
- [18] D. Liu, L. Wu, L. Liu, H. Zhao, F. Boussaid, and Bennamoun, "Jacobian Norm with Selective Input Gradient Regularization for Improved and Interpretable Adversarial Defense," 2022.
- [19] A. Chan, Y. Tay, Y. S. Ong, and J. Fu, "Jacobian Adversarially Regularized Networks for Robustness," *ArXiv*, vol. abs/1912.10185, 2019.
- [20] D. Jakubovitz and R. Giryes, "Improving DNN Robustness to Adversarial Attacks using Jacobian Regularization," in *European Conference on Computer Vision*, 2018.
- [21] X. Peng, H. Gao, S. Chen, H. Wang, and H. M. Zhang, "Combat Urban Congestion via Collaboration: Heterogeneous GNN-based MARL for Coordinated Platooning and Traffic Signal Control," *ArXiv*, vol. abs/2310.10948, 2023.
- [22] Y. Zhang *et al.*, "Toward Dependency Dynamics in Multi-Agent Reinforcement Learning for Traffic Signal Control," *ArXiv*, vol. abs/2502.16608, 2025.
- [23] A. R. Sattarzadeh and P. N. Pathirana, "Unification of probabilistic graph model and deep reinforcement learning (UPGMDRL) for multi-intersection traffic signal control," *Knowledge-Based Systems*, vol. 305, p. 112663, 2024/12/03/ 2024, doi: <https://doi.org/10.1016/j.knosys.2024.112663>.
- [24] Q. Wu, J. Wu, J. Shen, B. Yong, and Q. Zhou, "An Edge Based Multi-Agent Auto Communication Method for Traffic Light Control," *Sensors (Basel, Switzerland)*, vol. 20, 2020.
- [25] H. Goel, Y. Zhang, M. Damani, and G. Sartoretti, "SocialLight: Distributed Cooperation Learning towards Network-Wide Traffic Signal Control," in *Adaptive Agents and Multi-Agent Systems*, 2023.
- [26] L. Canese *et al.*, "Multi-Agent Reinforcement Learning: A Review of Challenges and Applications," *Applied Sciences*, vol. 11, p. 4948, 2021.
- [27] T. Chu, J. Wang, L. Codecà, and Z. Li, "Multi-agent deep reinforcement learning for large-scale traffic signal control," *IEEE transactions on intelligent transportation systems*, vol. 21, no. 3, pp. 1086-1095, 2019.
- [28] M. A. Basmassi, S. Boudaakat, J. A. Chentoufi, L. Benameur, A. Rebbani, and O. Bouattane, "Evolutionary reinforcement learning multi-agents system for intelligent traffic light control: new approach and case of study," *International Journal of Electrical and Computer Engineering (IJECE)*, 2022.
- [29] Z. Zhang, J. Qian, C. Fang, G. Liu, and Q. Su, "Coordinated Control of Distributed Traffic Signal Based on Multiagent Cooperative Game," *Wirel. Commun. Mob. Comput.*, vol. 2021, pp. 6693636:1-6693636:13, 2021.
- [30] Y. Xiang *et al.*, "Congestion Attack Detection in Intelligent Traffic Signal System: Combining Empirical and Analytical Methods," *Security and Communication Networks*, 2021.
- [31] M. Wu, M. Wicker, W. Ruan, X. Huang, and M. Kwiatkowska, "A Game-Based Approximate Verification of Deep Neural Networks with Provable Guarantees," *ArXiv*, vol. abs/1807.03571, 2018.
- [32] N. Papernot, P. McDaniel, S. Jha, M. Fredrikson, Z. B. Celik, and A. Swami, "The Limitations of Deep Learning in Adversarial Settings," *2016 IEEE European Symposium on Security and Privacy (EuroS&P)*, pp. 372-387, 2015.
- [33] A. Qu, Y. Tang, and W. Ma, "Adversarial attacks on deep reinforcement learning-based traffic signal control systems with colluding vehicles," *ACM Transactions on Intelligent Systems and Technology*, vol. 14, no. 6, pp. 1-22, 2023.
- [34] L. Pinto, J. Davidson, R. Sukthankar, and A. Gupta, "Robust adversarial reinforcement learning," in *International conference on machine learning*, 2017: PMLR, pp. 2817-2826.

- [35] A. Bukharin *et al.*, "Robust multi-agent reinforcement learning via adversarial regularization: Theoretical foundation and stable algorithms," *Advances in Neural Information Processing Systems*, vol. 36, pp. 68121-68133, 2023.
- [36] C. Wang, Y. Ran, L. Yuan, Y. Yu, and Z. Zhang, "Robust Multi-Agent Reinforcement Learning against Adversaries on Observation," 2023.
- [37] S. Kanis, L. Samson, D. Bloembergen, and T. Bakker, "Back to basics: Deep reinforcement learning in traffic signal control," *arXiv preprint arXiv:2109.07180*, 2021.
- [38] A. Müller and M. Sabatelli, "Bridging the reality gap of reinforcement learning based traffic signal control using domain randomization and meta learning," in *2023 IEEE 26th International Conference on Intelligent Transportation Systems (ITSC)*, 2023: IEEE, pp. 5271-5278.
- [39] A. R. Sattarzadeh, P. N. Pathirana, and M. Palaniswami, "Implicit sensing self-supervised learning based on graph multi-pretext tasks for traffic flow prediction," *Neural Computing and Applications*, vol. 37, no. 2, pp. 1041-1066, 2025.
- [40] H. Wei *et al.*, "Colight: Learning network-level cooperation for traffic signal control," in *Proceedings of the 28th ACM international conference on information and knowledge management*, 2019, pp. 1913-1922.
- [41] M. Aslani, M. S. Mesgari, and M. Wiering, "Adaptive traffic signal control with actor-critic methods in a real-world traffic network with different traffic disruption events," *Transportation Research Part C: Emerging Technologies*, vol. 85, pp. 732-752, 2017.
- [42] F. Rodrigues and C. L. Azevedo, "Towards robust deep reinforcement learning for traffic signal control: Demand surges, incidents and sensor failures," in *2019 IEEE intelligent transportation systems conference (ITSC)*, 2019: IEEE, pp. 3559-3566.
- [43] X. Song *et al.*, "LipsNet: A smooth and robust neural network with adaptive Lipschitz constant for high accuracy optimal control," in *International Conference on Machine Learning*, 2023: PMLR, pp. 32253-32272.
- [44] P. Michailidis, I. Michailidis, C. R. Lazaridis, and E. Kosmatopoulos, "Traffic Signal Control via Reinforcement Learning: A Review on Applications and Innovations," *Infrastructures*, vol. 10, no. 5, p. 114, 2025.
- [45] H. Drucker and Y. LeCun, "Improving generalization performance using double backpropagation," *IEEE transactions on neural networks*, vol. 3, pp. 991-7, 1992.
- [46] A. J. Havens, Z. Jiang, and S. Sarkar, "Online Robust Policy Learning in the Presence of Unknown Adversaries," in *Neural Information Processing Systems*, 2018.
- [47] A. Bukharin *et al.*, "Robust Multi-Agent Reinforcement Learning via Adversarial Regularization: Theoretical Foundation and Stable Algorithms," *ArXiv*, vol. abs/2310.10810, 2023.
- [48] Y. Liu *et al.*, "GPLight: Grouped Multi-agent Reinforcement Learning for Large-scale Traffic Signal Control," in *International Joint Conference on Artificial Intelligence*, 2023.
- [49] K. Wang, Z. Shen, Z. Lei, and T. Zhang, "Towards Multi-agent Reinforcement Learning based Traffic Signal Control through Spatio-temporal Hypergraphs," *ArXiv*, vol. abs/2404.11014, 2024.
- [50] M. Azad-Manjiri, M. Afsharchi, and M. Abdoos, "DDPGAT: Integrating MADDPG and GAT for optimized urban traffic light control," *IET Intel. Transport Syst.*, 2025.
- [51] T. Chu, J. Wang, L. Codecà, and Z. Li, "Multi-Agent Deep Reinforcement Learning for Large-Scale Traffic Signal Control," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, pp. 1086-1095, 2019.
- [52] S. Wang and S. Wang, "A Novel Multi-Agent Deep RL Approach for Traffic Signal Control," *2023 IEEE International Conference on Pervasive Computing and Communications Workshops and other Affiliated Events (PerCom Workshops)*, pp. 15-20, 2023.
- [53] J. Schmidt, F. Dreyer, S. A. Hashimi, and S. Stober, "TransferLight: Zero-Shot Traffic Signal Control on any Road-Network," *ArXiv*, vol. abs/2412.09719, 2024.
- [54] W. Guo, W. Li, Z. Zhang, L. Zhang, L. Li, and D. Li, "Scalable Multi-Objective Optimization for Robust Traffic Signal Control in Uncertain Environments," *ArXiv*, vol. abs/2409.13388, 2024.
- [55] Z. Yang, R. Zhang, G. Pandey, N. Masoud, and H. X. Liu, "A hierarchical vehicle behavior prediction framework with traffic signals and interactive agents," *IEEE Trans. Intell. Transp. Syst.*, vol. 24, no. 10, pp. 11066-11079, 2023.
- [56] S. Cao, L. Wu, R. Zhang, Y. Chen, J. Li, and Q. Liu, "A spatial-temporal gated hypergraph convolution network for traffic prediction," *IEEE Trans. Veh. Technol.*, vol. 73, no. 7, pp. 9546-9559, 2024.
- [57] M. Wegener, F. Herrmann, L. Koch, R. Savelsberg, and J. Andert, "Longitudinal vehicle motion prediction in urban settings with traffic light interaction," *IEEE Trans. Intell. Veh.*, vol. 8, no. 1, pp. 204-215, 2021.
- [58] Z. He *et al.*, "Pairwise and hyper-correlations based spatiotemporal neural networks for traffic speed predictions," in *2023 24th IEEE international conference on Mobile Data Management (MDM)*, 2023: IEEE, pp. 235-244.

- [59] H. Twaddle and F. Busch, "Binomial and multinomial regression models for predicting the tactical choices of bicyclists at signalised intersections," *Transportation research part F: traffic psychology and behaviour*, vol. 60, pp. 47-57, 2019.
- [60] M. Li, S. Chen, Y. Zhao, Y. Zhang, Y. Wang, and Q. Tian, "Dynamic multiscale graph neural networks for 3d skeleton based human motion prediction," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 214-223.
- [61] M. Li, S. Chen, Y. Zhao, Y. Zhang, Y. Wang, and Q. Tian, "Multiscale spatio-temporal graph neural networks for 3d skeleton-based motion prediction," *IEEE Trans Image Process*, vol. 30, pp. 7760-7775, 2021.
- [62] Y. Hu, S. Chen, Y. Zhang, and X. Gu, "Collaborative motion prediction via neural motion message passing," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 6319-6328.
- [63] T. Kipf, E. Fetaya, K.-C. Wang, M. Welling, and R. Zemel, "Neural relational inference for interacting systems," in *International conference on machine learning*, 2018: Pmlr, pp. 2688-2697.
- [64] C. J. Maddison, A. Mnih, and Y. W. Teh, "The concrete distribution: A continuous relaxation of discrete random variables," *arXiv preprint arXiv:1611.00712*, 2016.
- [65] Y. Liu *et al.*, "GPLight: Grouped Multi-agent Reinforcement Learning for Large-scale Traffic Signal Control," in *IJCAI*, 2023, pp. 199-207.
- [66] M. Azad-Manjiri, M. Afsharchi, and M. Abdoos, "DDPGAT: Integrating MADDPG and GAT for optimized urban traffic light control," *IET Intelligent Transport Systems*, vol. 19, no. 1, p. e70000, 2025.
- [67] C. Graber and A. G. Schwing, "Dynamic neural relational inference," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 8513-8522.
- [68] A. R. Sattarzadeh, P. N. Pathirana, and V. T. Huynh, "Traffic State Estimation with Spatio-Temporal Autoencoding Transformer (STAT Model)," *IEEE Access*, 2025.
- [69] C. Szegedy *et al.*, "Intriguing properties of neural networks," *arXiv preprint arXiv:1312.6199*, 2013.
- [70] A. Madry, A. Makelov, L. Schmidt, D. Tsipras, and A. Vladu, "Towards deep learning models resistant to adversarial attacks," *arXiv preprint arXiv:1706.06083*, 2017.
- [71] N. Papernot, P. McDaniel, S. Jha, M. Fredrikson, Z. B. Celik, and A. Swami, "The limitations of deep learning in adversarial settings," in *2016 IEEE European symposium on security and privacy (EuroS&P)*, 2016: IEEE, pp. 372-387.
- [72] Y. Dong *et al.*, "Boosting adversarial attacks with momentum," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 9185-9193.
- [73] B. Yang, H. Zhang, Z. Li, Y. Zhang, K. Xu, and J. Wang, "Adversarial example generation with adabelief optimizer and crop invariance," *Appl Intell*, vol. 53, no. 2, pp. 2332-2347, 2023.
- [74] J. Lin, C. Song, K. He, L. Wang, and J. E. Hopcroft, "Nesterov accelerated gradient and scale invariance for adversarial attacks," *arXiv preprint arXiv:1908.06281*, 2019.
- [75] A. Ilyas, S. Santurkar, D. Tsipras, L. Engstrom, B. Tran, and A. Madry, "Adversarial examples are not bugs, they are features," *Advances in neural information processing systems*, vol. 32, 2019.
- [76] N. Papernot, I. Goodfellow, R. Sheatsley, R. Feinman, and P. McDaniel, "cleverhans v2. 0.0: an adversarial machine learning library," *arXiv preprint arXiv:1610.00768*, vol. 10, 2016.
- [77] N. Papernot, P. McDaniel, X. Wu, S. Jha, and A. Swami, "Distillation as a defense to adversarial perturbations against deep neural networks," in *2016 IEEE symposium on security and privacy (SP)*, 2016: IEEE, pp. 582-597.
- [78] J. Xing, T. Nagata, X. Zou, E. Neftci, and J. L. Krichmar, "Achieving efficient interpretability of reinforcement learning via policy distillation and selective input gradient regularization," *Neural Networks*, vol. 161, pp. 228-241, 2023.
- [79] A. Ross and F. Doshi-Velez, "Improving the adversarial robustness and interpretability of deep neural networks by regularizing their input gradients," in *Proceedings of the AAAI conference on artificial intelligence*, 2018, vol. 32, no. 1.
- [80] T. Nishi, K. Otaki, K. Hayakawa, and T. Yoshimura, "Traffic signal control based on reinforcement learning with graph convolutional neural nets," in *2018 21st International conference on intelligent transportation systems (ITSC)*, 2018: IEEE, pp. 877-883.
- [81] P. Koonce, "Traffic signal timing manual," United States. Federal Highway Administration, 2008.
- [82] Y. Zhang, H. Goel, P. Li, M. Damani, S. Chinchali, and G. Sartoretti, "Coordlight: Learning decentralized coordination for network-wide traffic signal control," *IEEE Transactions on Intelligent Transportation Systems*, 2025.
- [83] P. Varaiya, "The max-pressure controller for arbitrary networks of signalized intersections," in *Advances in dynamic network modeling in complex transportation systems*: Springer, 2013, pp. 27-66.

- [84] L. Zhang, Q. Wu, J. Shen, L. Lü, B. Du, and J. Wu, "Expression might be enough: Representing pressure and demand for reinforcement learning based traffic signal control," in *International Conference on Machine Learning*, 2022: PMLR, pp. 26645-26654.
- [85] C. Chen *et al.*, "Toward a thousand lights: Decentralized deep reinforcement learning for large-scale traffic signal control," in *Proceedings of the AAAI conference on artificial intelligence*, 2020, vol. 34, no. 04, pp. 3414-3421.
- [86] J. Lin, Y. Zhu, L. Liu, Y. Liu, G. Li, and L. Lin, "Denselight: Efficient control for large-scale traffic signals with dense feedback," *arXiv preprint arXiv:2306.07553*, 2023.
- [87] H. Goel, Y. Zhang, M. Damani, and G. Sartoretti, "Socallight: Distributed cooperation learning towards network-wide traffic signal control," *arXiv preprint arXiv:2305.16145*, 2023.
- [88] M. Lecuyer, V. Atlidakis, R. Geambasu, D. Hsu, and S. Jana, "Certified robustness to adversarial examples with differential privacy," in *2019 IEEE symposium on security and privacy (SP)*, 2019: IEEE, pp. 656-672.
- [89] X. Liu, Y. Li, C. Wu, and C.-J. Hsieh, "Adv-bnn: Improved adversarial defense through robust bayesian neural network," *arXiv preprint arXiv:1810.01279*, 2018.
- [90] A. Chan, Y. Tay, Y. S. Ong, and J. Fu, "Jacobian adversarially regularized networks for robustness," *arXiv preprint arXiv:1912.10185*, 2019.