Cooperative Adaptable Lanes for Safer Shared Space and Improved Mixed-Traffic Flow

Rohit K. Dubey[a,*], Javier Argota Sánchez–Vaquerizo[a], Damian Dailisan[a], Dirk Helbing[a,b]

[a] *Computational Social Science, ETH Zürich, Stampfenbachstr. 48, Zürich, 8092, Switzerland*
[b] *Complexity Science Hub Vienna, Josefstädter Str. 39, Vienna, 1080, Austria*

Abstract

With the rapid increase in the percentage of the world's population living in cities, the design of existing transportation infrastructure requires serious consideration. Current road networks, especially in large cities, face acute pressures due to increased demand for vehicles, cyclists, and pedestrians. Although much attention has been given to improve traffic management and accommodate the increased demand via coordinating and optimizing traffic signals, research focused on adapting the static allocation of street spaces and right-of-way dynamically based on mixed traffic flow is still scarce. This paper proposes a multi-agent reinforcement learning (RL) agent approach that cooperatively adapts the individual lane widths and right-of-way access permissions based on real-world mixed traffic flow. In particular, multiple cooperative agents are trained with mixed temporal data that learn to decide suitable lane widths for motorized vehicles, bicycles, and pedestrians, along with whether co-sharing space between pedestrians and cyclists is safe. Using a microscopic traffic simulator model of a four-legged intersection, we trained our RL agent on synthetic data, and tested it on realistic multi-modal traffic data. The proposed approach significantly reduced average waiting time by 73.3%, 5.9%, and 36.6% and reduces the average queue length by 56.1%, 4.4%, and 36.9% with respect to Static, Heuristic, and PPO-based adaptive models, respectively. Moreover, the model learned to adaptively toggle co-sharing of the street space between cyclists and pedestrians as one co-shared lane, keeping the comfort and level of service in accordance with the designer's policy.

1. Introduction

Urban transportation infrastructure is critical in facilitating the mobility of people, goods, and services to achieve connected and sustainable cities. Rapid growth of cities across the world currently poses significant challenges to the static urban infrastructure of the past. Crucial infrastructure in big and populated cities, such as transport and street networks, is under stress due to the exponential increase in mobility needs. Growing traffic congestion is an important factor in environmental pollution, fossil energy depletion, and quality of life in cities. Research has shown excess morbidity and mortality for drivers, commuters, and individuals even at low concentrations of emissions [1]. Thus, reducing traffic congestion with respect to the vehicle queue length at intersections, in traffic jams, and reducing total waiting times in traffic is of utmost importance for comfortable commutes and overall citizen well-being, as well as reduced environmental impacts due to avoidable vehicle emissions. Furthermore, coping and adapting to the diversity of uses in urban space is also crucial to promote inclusion in built environments and spatial planning [2].

Cities must efficiently allocate limited space for various uses, integrating diverse transportation modes, while balancing mobility, safety, and design requirements. This often requires navigating compromises between infrastructure size and existing spatial constraints. Static functional segregation is the common streetscape found in cities. This approach prioritizes safety and coping with the high peak demands of

*Corresponding author
Email address: rohit.dubey@gess.ethz.ch (Rohit K. Dubey)

Submitted preprint *January 2024*

vehicular traffic. Often, this results in oversized roadways, which limit alternative uses [3]. In contrast, sharing spaces and, more recently, adaptive and intelligent systems are promising strategies to cope with changing needs and to leverage the diversity of uses with the limited space in cities. Although pedestrian and bicycle lane separation has primarily succeeded on streets with ample space, many dense cities facing constraints in street space require innovative solutions. New solutions should aim to accommodate vulnerable road users and motorized vehicles by designating specific spaces on streets with fair right-of-way (ROW), thereby addressing potential inequalities in the utilization of urban space [4].

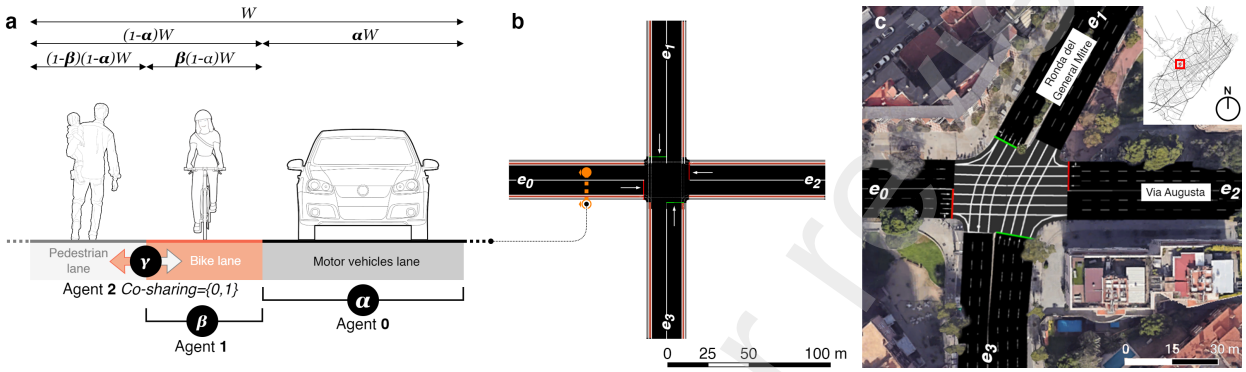

Figure 1: (a) Example of street layout foreseeing seperate rights of way (ROWs). (b) Synthetic 4-arm intersection with four outgoing and four incoming edges. (c) Intersection of Via Augusta with Ronda del General Mitre extracted from the street network of a traffic digital twin of Barcelona. For this research, this layout was modified to allow for adaptable lanes and mixed traffic (vehicles, cyclists and pedestrians). Inset: Location of the intersection in the city.

This paper proposes a reinforcement learning (RL) based **C**ooperative **A**daptable **L**anes **M**odel called **CALM**. Utilizing the Multi-Agent Deep Deterministic Policy Gradient algorithm (MADDPG) [5], our approach aims to enhance safety and optimize mixed-traffic flow with dual objectives: (i) ensuring efficient traffic flow and (ii) prioritizing the safety of vulnerable road users. Three cooperative learning agents (Fig. 1a) jointly manage lane width allocations for cars, cyclists, and pedestrians, including co-sharing arrangements. Unlike static optimization algorithms, cooperative reinforcement learning enables adaptive and collaborative decision-making among agents, offering dynamic responses to changing environments. Our method dynamically allocates street space, minimizing traffic congestion, queue length, and waiting time for all road-user types, while emphasizing the comfort and safety of pedestrians and cyclists. The integration of the widely used open-source traffic simulator SUMO [6] within an OpenAI Gym environment facilitates Deep Reinforcement Learning (DRL) for comprehensive evaluation. A synthetic 4-arm intersection (Fig. 1b) with four bi-directional streets as candidates for adaptability was chosen with synthetic mixed-traffic flow training data. Finally, to showcase the generalizability and strength of the proposed method, we conducted a thorough comparative analysis between centralized and decentralized deployments of adaptive lane models along with stress-testing the network with artificially influenced surge traffic demands.

As we will show, the proposed RL-based cooperative algorithm **CALM** in its decentralized version outperformed static, heuristic and multi-discrete Proximal Policy Optimization (PPO), as well as **CALM** centralized versions of the adaptive lanes. The dynamic allocation of space to the changing demand enabled by **CALM** was finally tested in a realistically calibrated digital twin of an existing busy intersection of the city of Barcelona with real traffic data. Overall, it resulted in shorter waiting times (73.3%, 5.9%, and 36.6%) and queue lengths (56.1%, 4.4%, and 36.9%) with respect to Static, Heuristic, and PPO-based adaptive models, while preserving the Level of Service (LOS) for vulnerable users. We will provide data and code to evaluate the **CALM** model at `https://github.com/ethz-coss/calm_rl`.

2. Related Work

Researchers in the past decade have re-invented the conventional use of shared space for mixed-traffic. Safety, comfort, and reducing traffic congestion have increasingly become vital factors for consideration in

traffic infrastructure design. This section briefly highlights the work done in the area of Intelligent Traffic Control, focusing on using RL. Then, we summarize the limited but crucial effort made toward adaptable street design. Finally, we present the advantages and limitations of shared space with pedestrians and cyclists' LOS.

2.1. Intelligent Control-Based Traffic Optimization Approach

Previous Efforts to enhance traffic management and address increased demand focus on optimizing signals, as seen in Ref. [7], where Q-learning is applied to control vehicle speed, reducing congestion. However, while improving conditions for motorized vehicles, traditional traffic control systems have often neglected pedestrian and cyclist priorities, fairness, and overall service quality. Our approach in contrast, seeks to alleviate congestion while prioritizing the safety and fairness for all road users. Another previous study has employed a single Deep RL agent for controlling the traffic signals at multiple intersections simultaneously [8]. According to it, this produces better traffic management than controlling the traffic signal of single intersections individually. Moreover, Ref. [9] has proposed cellular-automata-based dynamic adaptive reversible lanes that improve static lane use by changing the direction based on real-time fluctuating traffic demand. Ref. [10] proposed a spatial-temporal traffic flow data-based adaptive street layout control that elects ROW of road space assigned to vehicular and pedestrian lanes. Their RL-based method raised the shared pedestrian lane-width by approximately 9%.

2.2. Adaptive Street Design

The challenge of allocating space to diverse uses, in urban planning arises when cities try to combine high capacity for vehicular traffic with diverse uses while aiming at environmental quality, active mobility, and pedestrian-friendly settings in a limited space [11]. However, the resulting trade-off is dynamic, as the demand for space by these diverse and competing, if not incompatible, uses changes over time (e.g., day, week, month, season). Recently, autonomous vehicles have promise to increase the capacity and safety of existing infrastructures [12] while enhancing the sharing of urban space pedestrian-friendly uses in livable streets [13] in more efficient and flexible ways. Traditionally, changing transportation demands has been adopted to by introducing managed lanes and other intelligent transportation systems as much as available technology allowed. However, they have been mainly explored, tested, and implemented in interurban infrastructures. These approaches include diverse solutions such as hard shoulder utilization, reserved lanes, high-occupancy lanes, reversible lanes, ramp metering, dynamic signs, and pricing schemes [14]. In urban environments, some dynamic management methods have been used [15–17], combined with static shared space design principles [18]. More recently, tactical urban planning practices have gained popularity to repurpose urban spaces incrementally and reclaim streets for pedestrians for temporary or more permanent transformations [19]. However, these typically focus on mid and low-frequency changes in transportation demand in the city. Few dynamical approaches try to cope with high-frequency changes (i.e., within the day). Today, public space still prioritizes motor vehicle traffic and size ROW allocation to cope with peak hour requirements [3]. Overall, this results in an under-utilization of space and hinders diversity and alternative uses of space.

2.3. Shared Space: Pedestrian and Cyclist

Shared mobility spaces have become popular as they promise to motivate more citizens to choose active travel modes (e.g., walking, cycling, or using e-bikes) [20]. Urban planners and policymakers also suggest that shared space principles between non-motorized users promote the perception of healthier and more attractive spaces [21]. Although streets incorporating shared-space design elements tend to be tempting and have many potential environmental, health, and social inclusion benefits [22], it has been established that it can also increase pedestrian-cyclist conflicts under certain conditions [23]. Refs. [24, 25] suggest that it is not advisable to co-share the space when pedestrian or cyclist density increases beyond a certain threshold.

Pedestrian–Cyclist shared spaces have garnered researchers' and urban planners' interest due to the limited available public space and intention to promote active mobility. According to multiple studies, the chances of severe accidents in the coexistence of pedestrians and cyclists are rare compared to the coexistence of cyclists and motor vehicles [26, 27]. However, some studies, such as Ref. [28], state that a probability of

conflicts between pedestrians and cyclists can arise in specific situations. Therefore, the strategy of shared lanes between pedestrians and cyclists require more attention and proper planning.

3. Level of Service: Pedestrian–Cyclist Spaces

This section presents a model for the Level of Service (LOS) of dedicated cyclist/pedestrian lanes and co-shared lanes for cyclist and pedestrian traffic. The LOS is a widely recognized indicator for assessing the well-being and satisfaction levels of pedestrians and cyclists while using shared street infrastructure [29]. Existing literature on LOS models predominantly considers either quantitative or qualitative factors, but, to our knowledge, rarely explores the combined influence of both. Since both kinds of factors significantly impact the LOS, our formulation incorporates both quantitative and qualitative measures. The quantitative measures involve assessing the width and occupancy of individual or co-shared lanes, while the qualitative measure is based on the concept of hindrance as proposed in [29]. Hindrance refers to situations where interactions between pedestrians and cyclists result in perceived discomfort. Examples of such events include a fast-moving cyclist passing a pedestrian with minimal lateral and longitudinal speration, a fast-moving pedestrian overtaking a slower pedestrian in a confined space, or a cyclist overtaking another cyclist. To quantify these occurrences, we track the frequency of cyclist–cyclist, pedestrian–pedestrian and cyclist–pedestrian interactions within a proximity of 1 meter.

Equation (1) proposes our formulation ofthe pedestrian-cyclist LOS. The LOS is a function of lane-occupancies (O_p,O_c), hindrance between pedestrian–cyclist (H_{pc}), pedestrian–pedestrian (H_{pp}), and cyclist–cyclist (H_{cc}) weighted by lane-width (w).

$$\mathbf{LOS} = \begin{cases} \frac{z_{lo}*O_p+z_{pc}*H_{pc}+z_{pp}*H_{pp}+z_{pc}*H_{cc}}{z_{lw}*w_{p+c}}, & \gamma = 1 \\ \frac{z_{lo}*O_p+z_{pp}*H_{pp}+O_c+z_{cc}*H_{cc}}{z_{lw}*(w_p+w_c)}, & \gamma = 0 \end{cases} \tag{1}$$

In general the LOS Equation (1) is a weighted sum of multiple influencing factors. For example, lane width and occupancy may have a more substantial impact on perceived safety, as these factors directly correlate with the number of conflicts. On the other hand, Ref. [29] chose to assign more importance (larger z values) to conflicts between pedestrian–cyclists than between pedestrian-pedestrian and between cyclist-cyclist. Computing the weights of the influencing parameters is beyond the scope of this paper, and we defer it to future works. Instead, we employ equal weights for all parameters and set (z_{pc}, z_{pp}, z_{lw}, and $z_{lo} = 1$) to calculate the LOS, to avoid biasing based on intuition.

4. Methodology

Throughout this section, we formulate the traffic problem, detail the relevant background RL algorithms, and introduce the traffic simulation framework used to create a model capable of learning to control adaptive, co-shared spaces. We also provide a heuristic adaptive lane algorithm for comparison.

4.1. Problem Formulation

In recent years, there has been an emerging trend of making city streets more pedestrian and cyclist friendly. Even the notion of shared pedestrian–cyclist lanes has gained popularity because they promote sustainable mobility. However, allocating wide exclusive lanes for each mode of transports is challenging, when street space is scarce. In our limited knowledge, the decision based on static combined pedestrian–cyclist lane is made by transport planners by evaluating the safety and the level of service of the shared lanes, using historical mixed-traffic data. However, such static implementations are unable to adapt to dynamically changing traffic demand.

Thus, the problem at hand involves designing a dynamic shared space allocation system for three types of road users—motorized vehicles, cyclists, and pedestrians—using a multi-agent approach. Our demonstration of such an adaptive street layout focuses on the roads connected to a simple 4-arm intersection (Fig. 1a). To achieve this, an edge $e \in \mathbf{E}$ is modeled as a cooperative reinforcement learning system with three agents

Table 1: Nomenclature.

Notations	Description
W	Width of an one-directional street
e	An edge in a network
$\mathbf{E}$	Set of all edges
α	Proportion of street space for motorized vehicles (Agent 1)
β	Proportion of street space for cyclist (Agent 2)
γ	Binary value that toggles co-shared lanes (Agent 3)
f_v, f_c, f_p	Flow (Motorized vehicle, Cyclist, Pedestrian)
t_{step}	Length of one step (5 minutes)
T	Simulation time period
n_{step}	Number of steps per episode
$t_{episode}$	Episode time period ($n_{step} \times t_{step}$)
w_v, w_c, w_p, w_{p+c}	Lane Width (Motorized vehicle, Cyclist, Pedestrian, Co-shared)
l_v, l_c, l_p	Lane Length (Motorized vehicle, Cyclist, Pedestrian)
D_v, D_c, D_p	Lane Density (Motorized vehicle, Cyclist, Pedestrian)
Δ	Constraint violation penalty
O_v, O_c, O_p	Lane Occupancy (Motorized vehicle, Cyclist, Pedestrian)
$\mathbf{S}$	Joint state space
$\mathbf{A}$	Joint action space
R_t	Cumulative reward of cooperative agents
c_v, c_c, c_p	count (Motorized vehicle, Cyclist, Pedestrian)

working together. Table 1 presents the framework parameters along with their descriptions. **Agent 0** (Motorized Vehicle Lane Agent): This agent is tasked with learning to allocate a portion of the total street space to motorized vehicle lanes. **Agent 1** (Cyclist Lane Agent): The role of this agent is to determine the allocation of widths for cyclist lanes, which affects the proportion of the street dedicated to cyclists. **Agent 2** (Co-sharing Control Agent): This agent is responsible for controlling the co-sharing of cyclist and pedestrian lanes, deciding whether the cyclist and pedestrian lanes should be co-shared or kept apart. The width of an edge (i.e., a one-way street) is denoted by W, with α and $(1-\alpha)$ representing the proportions of motorized and non-motorized lanes, respectively. Similarly, β and $(1-\beta)$ represent the proportions of the cyclist and pedestrian lanes, respectively. Thus, the width of a motorized vehicle lane is αW, that of cyclist lane is $\beta(1-\alpha)W$, and that of pedestrian lane is $(1-\beta)(1-\alpha)W$ (see Fig. 1a). The co-sharing of the pedestrian and cyclist lane is controlled by the parameter γ (a binary value that toggles co-sharing of lanes on and off). In essence, three lane agents collaboratively perform four functions: (1) manage motorized vehicle lane allocation, (2) handle pedestrian-cyclist lane width, (3) handle cyclist lane width, and (4) control the co-sharing arrangement of cyclist and pedestrian lanes. The collaboration among these agents ensures a comprehensive and coordinated approach to optimizing the allocation of street space for motorized vehicles, cyclists, and pedestrians within the modeled 4-arm intersection.

Given the real-time dynamic mixed-traffic demands of vehicles (f_v), pedestrians (f_p), and cyclists (f_c) over the period of a fixed time slot (t_{step}), the goal is to optimize the width of each lane in conjunction with the decision of sharing or not sharing the pedestrians and cyclists lane, to improve the traffic flow efficiency, to reduce congestion, to minimize queue-length, and to increase safety and well-being of non-motorized road users.

Real-world implementations can utilize embedded LEDs on roads (as demonstrated in a case study for dynamically adjusting pedestrian crossing markings [30]) or overhead smart signboards to prescribe road use. In traffic simulations, instantaneous lane changing can cause additional road conflicts (which are analogous to risky maneuvers that can cause accidents). Setting a model value of `lanechange.duration` $\in [3.6, 5.9]$ slows down the lane changing maneuver and allows agents (vehicles, cyclists, and pedestrians) more time to react to other road users [31]. The proposed method is tailored for the future of transportation, envisioning seamless integration of autonomous vehicles and advanced communication (V2V and V2X). Focusing on adaptive lane width, it enables autonomous vehicles to dynamically receive real-time lane change commands, ensuring optimized traffic flow and safety. Acknowledging the importance of notifying users about lane sharing and vehicle intents [32], strategies can incorporate dynamic signage at intervals and smart LED pavement, which adjusts markings in real-time using color-changing LEDs to signify applicable lane sharing [33].

4.2. Reinforcement Learning Background

4.2.1. Markov Decision Process

(MDP) [34] can be used to model traffic control systems. An MDP is defined by the tuple $\langle \mathbf{S}, \mathbf{A}, R, P \rangle$, where $\mathbf{S}$ is the set of all possible states of the environment, $\mathbf{A}$ is the set of actions available to the agent, R is the reward function that determines the immediate reward obtained by the agent after taking an action in a given state, and P is the transition probability function that describes the probability of transitioning to a new state after taking action.

In general, agents may lack access to the complete state representation of the system, instead relying on partial observations $\mathcal{O} := \mathcal{O} \in \mathbf{S}$ of the system. Actions of an agent are selected using a policy $\pi : \mathcal{O} \to \mathbf{A}$, and these actions change the state of the system according to P. For control problems, one common objective is to find the policy π that maximizes the discounted expected reward $R = \sum_{t=0}^{T} \gamma^t r_t$, where γ is a discount factor and T is a time horizon [35].

4.2.2. Multi-Agent Deep Deterministic Policy Gradient

(MADDPG) [5] is a multi-agent extension of the Deep Deterministic Policy Gradient (DDPG) [36] reinforcement learning algorithm. The DDPG algorithm models an agent using an actor–critic algorithm [37] and uses a deep Q-network to approximate the action-value function

$$Q(o, a|\theta) = \mathbb{E}_{o'} \left[r(o, a) + \gamma \mathbb{E}_{a' \sim \pi} \left[Q(o', a'|\theta') \right] \right]. \tag{2}$$

Herein, θ are the parameters of the critic network Q in conjunction with optimizing for an objective $J = \mathbb{E}_{o \sim p^\mu} \left[R(o, a) \right]$ by taking steps in the direction of the gradient $\nabla_\phi J(\phi)$, with the policy network parametrized by ϕ. Choosing deterministic policies $\mu_\phi : \mathcal{O} \to \mathbf{A}$, the gradient can be written as

$$\nabla_\phi J(\phi) = \mathbb{E}_{o \sim p^\mu} \left[\nabla_\phi \mu(o|\phi) \nabla_a Q(o, a|\theta)|_{a=\mu(o)} \right], \tag{3}$$

where p^μ is the state distribution.

The MADDPG algorithm employs a strategy of *combined training* and *decentralized execution*. During training, the critic network $Q^\mu(o, a)$ is replaced by a *combined action–value functions* $Q_i^\mu(o_1, \ldots, o_N, a_1, \ldots, a_N)$, which incorporates observations and actions from all other agents. Then, deployment of the model follows decentralized execution, where agents only use local information to choose an action through their actor network $\pi_1, \ldots, \pi_N$. This approach enables MADDPG to be applied even in partially observable environments, where agent communication is limited.

4.2.3. Proximal Policy Optimization

(PPO) [38] trains a stochastic neural network policy π by directly updating probabilities. To be specific, PPO uses two policies: the first is the current policy $\pi_\phi(a|s)$ that we want to update, and the second is the last policy that we use to collect samples $\pi_{\phi_k}(a|s)$. The weights of the neural network are updated according to

$$\phi_{k+1} = \arg\max_{\phi} \mathop{\mathbb{E}}_{s,a \sim \pi_{\phi_k}} \left[L\left(s, a, \phi_k, \phi\right) \right], \tag{4}$$

where ϕ_k and ϕ represent the weights from the old and current policies, respectively. The objective function is

$$L\left(s, a, \phi_k, \phi\right) = \min \left(\frac{\pi_\phi(a|s)}{\pi_{\phi_k}(a|s)} A^{\pi_{\phi_k}}(s, a), \quad g\left(\epsilon, A^{\pi_{\phi_k}}(s, a)\right) \right), \tag{5}$$

where the clipping hyperparameter ϵ limits the change between the new and old policies. The advantage function $A^\pi(s, a) = Q - \mathbb{E}_\tau[R \mid s_0 = s]$ estimates the difference between the actual discounted reward and the expected reward, based on the value function [39]. The first term in Equation (5) is the objective for policy-gradient methods that push the policy towards actions with greater advantage than the baseline. The second term $g(\epsilon, A)$ accounts for the value of the advantage: $g(\epsilon, A) = (1 + \epsilon)A$ for positive advantage; $g(\epsilon, A) = (1 - \epsilon)A$ otherwise. Clipping the objective function prevents drastic policy updates, thereby regularizing the policy by limiting updates that result in large deviations from the old policy.

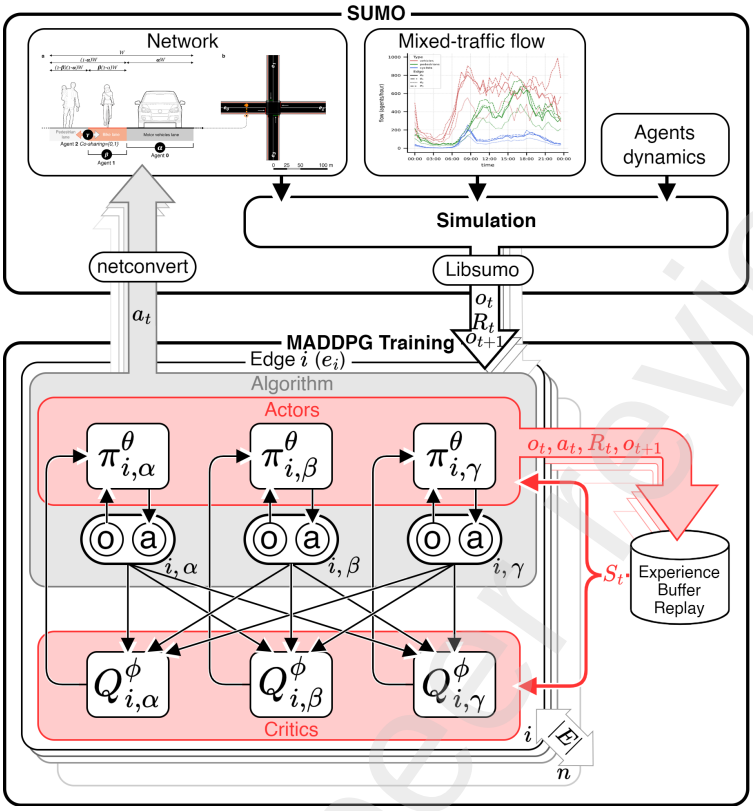

Figure 2: **CALM** Framework: Overview of the proposed SUMO–MADDPG adaptable lanes framework. The framework seamlessly integrates with an OpenAI-Gym environment, featuring two core modules: the SUMO-based traffic simulator and the MADDPG training module. Utilizing `libsumo` and `netconvert`, the framework facilitates the exchange of network updates and transmission of states, actions, and rewards tuples.

*4.3. CALM: **C**ooperative **A**daptable **L**anes **M**odel*

4.3.1. SUMO model of an intersection

SUMO is a free and open-source microscopic traffic simulator, which allows for real-time updates of microscopic traffic parameters (through the `libsumo` Python interface) and manipulating road network configurations (through `netedit`). For our simulations, we model a synthetic 4-arm intersection (Fig. 1b) inspired by a real-world traffic intersection from the city of Barcelona (Fig. 1c). The 4-arm intersection consists of four bidirectional edges ($e_0 - e_3$). Each edge is 100 meters long and 12.6 meters wide, with a 50 km/h speed limit. Edge e has three fixed lanes (l_v, l_p, l_c) that only allow traffic of specific SUMO vehicle classes (passenger, pedestrians, and bicycles), respectively. Each lane follows a sub-lane model with a lateral resolution of 0.1. Default widths of $l_v = 9.6$, $l_p = 1.5$, and $l_c = 1.5$ meters are set at the start of each simulation episode. In the case of pedestrian–cyclists co-sharing of lanes, the width l_c is set to 0 meters, and the width decided by Agent 2 (β) is added to l_p. The access permissions of the pedestrian lane l_p are modified to allow both pedestrian and bicycle vehicle classes. We used fixed-time (FT) traffic signals for the intersection with a fixed cycle duration of 90 seconds, with a green time duration of 40 seconds, each for the North–South and East–West directions. However, we also used actuated traffic signals (ACT) with minimum and maximum green times of 5 and 50 seconds, respectively, as a second baseline for "static" (non-adaptive lanes) scenarios. Finally, we overrode the default values of pedestrian and cyclist widths in SUMO by 0.65 m and 0.75 m, respectively, for more realistic cyclist and pedestrian dynamics.

4.3.2. CALM: A SUMO–MADDPG-based Traffic Simulation

We introduce a SUMO–MADDPG adaptable lanes framework, integrating with an OpenAI-Gym environment to interface with SUMO (see Fig. 2). The logical flow consists of two main modules: the SUMO-based traffic simulator and the MADDPG training module. Leveraging `libsumo` and `netconvert`, we facilitate the exchange of network updates and transfer states, actions, and rewards tuples to the MADDPG training module. The simulation involves loading a traffic network with mixed traffic flow, running for a one-day period (Fig. 3). SUMO simulates steps, concluding each training epoch by commanding netconvert to adjust lane widths and decide on the sharing of cycle and pedestrian lanes. The new tuples are then sent to the training module, producing a new set of actions for subsequent training phases. This continues for a pre-decided number of training episodes. This framework was designed to use MADDPG, which enables us to deal with smaller state and action spaces compared to single-agent RL algorithms. The simulation environment models the multi-agent reinforcement learning system with 3 cooperative RL agents for one edge and 12 agents for all four edges in a 4-arm intersection. Below, we outline the critical elements of this framework:

- $\mathbf{E}$ is the set of all edges $e \in \mathbf{E}$.
- α, β, and γ parameters controlled by Agent0 (vehicle), Agent1 (cyclist), and Agent2 (co-sharing lanes) per e.
- $\mathbf{S}$ represents the collective partial states observed by each agent. $\mathbf{S} = \sum_1^n s_{e_i}$ and i denotes individual agents (1 to 3 for one edge and 1 to 12 in the 4-arm intersection).
- $\mathbf{A}$ represents the collective actions a_1, a_2 and a_3 taken by α, β, and γ, respectively.

The workflow of **CALM** consists of a stochastic traffic simulation phase and an off-policy actor–critic-based learning phase. At step $t = 0$, **CALM** loads the default 4-arm network configuration and initiates the mixed-training flow route files. A warm-up period of 300 steps is simulated before starting the observation phase. The observation phase runs from 300 to 600 simulation steps. At the end of 600 simulation steps, the model predicts actions, and the corresponding reward from the environment is recorded. **CALM** learns policies that select α, β, and γ iteratively.

Observations: **CALM** uses continuous observations from the environment. Instead of assigning each agent the complete observation space, we assigned to each agent observations of lane widths w, occupancies O, and densities D of the edge (see Appendix) for other possible observation sets). The widths are indirect representations of the actions α and β, and we also observe the co-sharing action γ. The observation vector for each agent is listed below:

- $S_0 = [w_v, w_p + w_c, O_v, D_v]$
- $S_1 = [w_c, w_p, O_c.O_p]$
- $S_2 = [w_v, w_p, w_c, O_v, D_v, O_c.O_p, \gamma, 1-\gamma, D_c, D_p]$

Actions: Three discrete sets for determining the action space for multiple cooperative agents by deterministic policy function with uniform random noise for agents are:

- **Agent 0** $(\alpha) = \left\{ \frac{x}{12.6} \mid x \in \{3.2, 5.6, 6.4, 7.8, 9.6\} \right\}$
- **Agent 1** $(\beta) = \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9\}$
- **Agent 2** $(\gamma) = \{0, 1\}$

Agent 0 allocates a portion of the total road width W to motorized vehicles, while **Agent 1** distributes the remaining space between cyclist and pedestrian lanes. **Agent 2** manages separating or co-sharing the cyclist and pedestrian lanes. Because small incremental changes in the width of the motorized vehicle lane are impractical, we designed instead the actions of α to range from the minimum to maximum lane width,

in increments of the widths that allow one additional vehicle with the sublane model. The actions for β are more granular with increments of 0.1, and finally, γ has a binary choice between 0 and 1. Additionally, we apply this action to both the upstream and downstream edges for simplicity (e.g., eastbound directions of e_0 and e_2).

Rewards: In the context of a cooperative multi-agent system, the reward is a cumulative sum of all individual multi-agent rewards (i.e., each agent receives the same reward). $R_t = \sum_{i=0}^{n} R_i$, where $n = 3$, is the number of agents considering a single edge of a 4-arm intersection. Equation (6) proposes how the observation from the environment was used to compute the agents' reward:

$$
\begin{aligned}
R_0 &= -\frac{\sum_t^{n_{step}} c_v}{w_v * l_v * n_{step}} \\
R_1 &= \begin{cases} \Delta, & \text{if } w_p < 1 \text{ or } w_c < 1 \\ -(\frac{c_p}{w_p * l_p * n_{step}} + \frac{c_c}{w_c * l_c * n_{step}}), & \text{otherwise} \end{cases} \\
R_2 &= \begin{cases} +0.75, & \text{if } \gamma = 1, \ \& \ D_s/n_{step} < \lambda \\ +0.75, & \text{if } \gamma = 0, \ \& \ D_p/n_{step} + D_c/n_{step} > 2\lambda \\ -0.75 & \text{otherwise} \end{cases}
\end{aligned} \tag{6}
$$

R_0 and R_1 incentivize the agents to minimize the average density of vehicular, pedestrian, and cyclist lanes per episode. Additionally, R_1 strongly penalizes the agent with Δ for generating cyclist and pedestrian lanes that are too narrow. R_2 incentivizes co-sharing when the densities of the shared lane are below λ and discourages co-sharing when densities are above 2λ (high). The extra factor of 2 accounts for the summing of the density values of two types of lanes. Safety and LOS (Section 3) depend on multiple factors, such as the traffic flow rate, width of shared spaces, occupancy, and shared lane density. Therefore, we introduce a lane density threshold λ as a heuristic parameter for deciding when to share pedestrian and cyclist lanes. R_2 would depend on the policy selected by a traffic planner but can be tuned to make the model learn to respect different policies based on LOS. Lane density is calculated using Equation (8), where it is a function of mixed-traffic flow, lane length, and lane width. See Appendix for how we select a value for λ.

4.4. Heuristic-Based Algorithm

We developed a robust adaptive model to serve as a heuristic baseline method. In this model, lane widths for cars, bicycles, and pedestrians are determined based on the measured flows/demand, with proportions set as follows:

$$
\begin{aligned}
\alpha &= \text{clip}\left(\frac{f_{veh}}{\sum f_i}, \frac{3.2}{12.6}, \frac{9.6}{12.6}\right) \\
\beta &= \text{clip}\left(\frac{f_{cyclist}}{\sum f_i}, \frac{1.5}{(1-\alpha)12.6}, 0.9\right) \\
\gamma &= \begin{cases} 1 & D_t \leq 2\lambda \\ 0 & D_t > 2\lambda \end{cases}
\end{aligned} \tag{7}
$$

$$
\begin{aligned}
D_i &= \frac{c_i}{w_i * l_i}, \text{for } i \in \{p, c, p+c\} \\
D_t &= \begin{cases} D_p + D_c, & \text{if } \gamma = 0 \\ D_s, & \text{otherwise} \end{cases}
\end{aligned} \tag{8}
$$

The heuristic algorithm Equation (7) allocates the actions based on the proportional flows of the vehicle types. The total road density D_t is then determined using Equation (8), where D_p, D_c, and D_s represent pedestrian, cyclist, and shared lanes density respectively. c_p and c_c represent the pedestrian and cyclist count on a lane at each simulation step. We allow co-sharing if the combined density of shared lanes or the total density of each lane (represented as D_t for both) is lower than the required threshold 2λ. Otherwise, segregation of users in these lanes is preferred for safety and comfort reasons.

5. Simulation Environment

In this section, we elaborate upon the training and validation simulation setup. We validate **CALM** against two street networks (1. a synthetic 4-arms intersection, 2. a real-world intersection) and against two mixed-traffic flow scenarios (1. a random traffic condition to stress-test the model, 2. a real-world mixed-traffic flow distribution over 24 hours). Moreover, we propose a heuristic algorithm for adaptable street design as a robust baseline to evaluate whether the RL-based method can outperform the expert-based mathematical model.

We ran our simulations on a laptop equipped with an 8-core Apple M2 processor and 8GB of RAM.

5.1. Training and Testing Simulation Setup

Below we elucidate the street network configurations and the mixed-traffic flow demands employed for training **CALM**.

5.1.1. Street Network Configurations

Using the adaptive street layout applied to the synthetic 4-arm intersection (see Fig. 1a), we trained **CALM** for two scenarios: in Scenario 1, only a single edge (e.g., e_0 in Fig. 1b) was made adaptive, while in Scenario 2, all four edges of the intersection (i.e., e_0, e_1, e_2, and e_3) were made adaptive. We used the traffic flows defined in Table 2 for training the model, while we used the mixed-traffic flow demands in Fig. 3 for testing.

5.1.2. Mixed-Traffic Flow Demands

Unlike previous research, [10], the mixed-traffic flow data for training was generated based on a heuristic rather than a bimodal travel time distribution [40]. The mixed-traffic demands for each vehicle class were divided into three categories, low, medium, and high, based on the real-world traffic data from multiple real-world street networks observed and adjusted to capacity estimations for equivalent road sections [41]. Table 2 presents the different flow rates for each vehicle class. Unfortunately, disaggregated flow data for all three vehicle classes for the same street are not readily available. Hence, to generate mixed-traffic flow data, we used a combination of real measurements of vehicles, cyclists, and pedestrians from different locations with similar characteristics, adjusting known aggregated daily counts with daily curve demands [42–44]. To generate mixed-traffic training scenarios, we bootstrapped combinations of low, medium, and high demands for all modalities. During each RL episode, we employed a heuristic to modify the vehicle flow by introducing a random perturbation between 1–5% at every simulation step. This perturbation aligns with the observed range of variations in off-peak 30-minute traffic flow observations [44–46]. At the start of each training episode, a random flow file is loaded from the training data, and the demand is perturbed.

Table 2: Mixed-Traffic Flow Demands based on real-world traffic flow (units/hour). A range for low, medium, and high traffic demand is employed to generate synthetic mixed-traffic training flows.

	Low	**Medium**	**High**
vehicles	140–350	740–950	1290–1500
pedestrian	1–165	108–622	236–750
cyclist	1–172	42–213	75–246

5.2. Validation Simulation Setup

Next, we describe the street network configurations and the mixed-traffic flow demands employed during validation.

5.2.1. Street Network Configurations

To validate our trained model, we chose the busy crossing of Via Augusta with Ronda del General Mitre in Barcelona as a real-world traffic intersection, which is presented in Fig. 1c and was extracted from a traffic digital twin of the city of Barcelona [46]. This multilane 4-way intersection features many of the design, planning, and engineering trade-offs due to mixed-used traffic needs addressed by the model. From the motor vehicle perspective, it is a signal-controlled intersection, part of the arterial network of the city. Cycle-wise, Via Augusta is currently being transformed to allocate new dedicated space for bike lanes [47]. Finally, both streets are planned to be part of the green corridors for the city without full pedestrian prioritization [48, 49]. Overall, it combines a heavy mixed use of streets within the actual urban transformation process, which wants to give more room in public space to non-motorized uses.

5.2.2. Mixed-Traffic Flow Demand

To realistically validate the model, we use a semi-realistic mixed approach that additionally tries to estimate a hypothesized more diverse mixed-use of the intersection envisioned by the current city plans. The estimation process is divided into two phases: (i) First, we estimate the demand curve in 30-minute timeslots for the entire day to account for the time-dependent demand. (ii) Then, we scale the curve to the estimated daily total count of that particular type of traffic on a given street section.

To obtain the motor vehicle demand, the traffic counts of the street sections converging into the intersection are extracted from the calibrated traffic digital twin of the entire city of Barcelona [46]. Fig. 3 shows this realistic reference time series for the motor vehicle demand at this crossing in 30-minute periods spanning an entire day, calibrated with real traffic counts [44]. It reflects the changing pace of both main streets. As in the case of the training demand dataset, real counts from other streets of Barcelona and similar cities are used to overcome the lack of precise counting of pedestrians and bikes in the studied intersection. To estimate the 24-hour curve demand for pedestrians, we rely on the real measured pedestrian flow of spaces in other Spanish city centers [43]. For the cyclist demand curves, we average out the daily curves observed in the most frequented cycle lanes as measured by the city (over 2000 cyclists/day as an average during the year) and the two closest cycle lanes counting points [44]. Finally, demand curves are scaled up in the second stage according to expected daily counts.

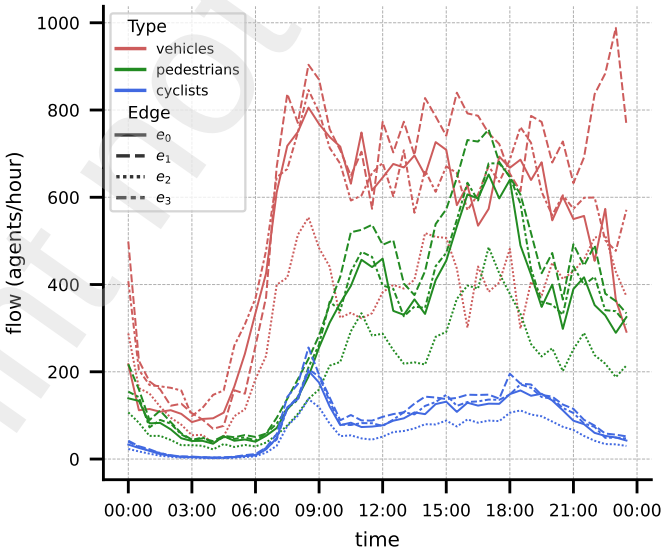

Figure 3: Mixed-traffic flow demands from a realistically calibrated simulation, used for testing the algorithm at the intersection Via Augusta with Ronda del General Mitre in Barcelona.

5.2.3. Validation Measures

CALM aims to assign road spaces to motorized, pedestrian, and cyclist users by minimizing the traffic delay and well-being of non-motorized vulnerable users. Therefore, we chose the average queue length (i.e., the count of standstill vehicles with speeds below 0.01 m/s) and waiting time for all three vehicle classes as validation measures. Moreover, we employed the proposed LOS (Section 3) measure to validate the overall safety and comfort of pedestrians and cyclists.

6. Results and Discussion

6.1. Training Simulation Results

Two effective and most commonly used RL algorithms, MADDPG and PPO, were employed to train the CALM model on a single edge e_0 of a 4-arms intersection. Edge e_0 in Fig. 1a illustrates an example edge with three cooperative adaptive agents. Only the best-performing algorithm between PPO, MADDPG, and Heuristic on a single edge (see Sec. 4.4) was elected and employed to train the 4-arm synthetic and real-world intersection network (i.e., traffic on all four arms). One key difference between the algorithms is that, in MADDPG, each agent is responsible for each discrete action, while in PPO, a single agent determines three discrete actions (multi-discrete). We used the PPO implementation in `StableBaselines3` [50]. .Both algorithms were trained for 1500 episodes with 20 actions in each episode (total of 30000 actions) Training MADDPG took ∼8 hours compared to ∼10 hours for PPO. MADDPG training time was significantly faster than the multi-discrete version of PPO, which can be attributed to the lower dimension observation space of the MADDPG model.

Fig. 4 showcases the learning curve for MADDPG and multi-discrete PPO, respectively. The best-performing hyperparameters are shown in Table 3. While training MADDPG, we observed that the algorithm is sensitive to the individual agent's observation space. Therefore, we decided to investigate the matter further. Appendix details the ablation study with changing observation space on the validation measures.

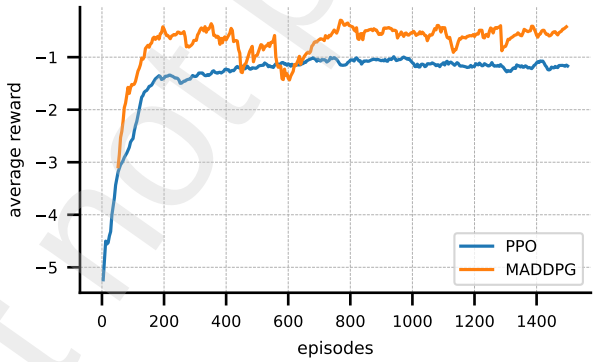

Figure 4: Learning curves of MADDPG and PPO. MADDPG achieves higher average rewards than PPO for the same training duration.

6.2. Validation Simulation Results

To validate the proposed algorithm and highlight the generalizability, we test it against synthetically induced surge (from 3 PM to 8 PM) in the mixed-traffic scenarios of a simulated real-world intersection at Via Augusta with Ronda del General Mitre in Barcelona as discussed in Section 5.2. Validation was performed against Static (non-adaptive lanes) and Heuristic (see Sec. 4.4) models. In Fig. 5, we present the results against three elected measures.

Table 3: Hyperparameters for MADDPG and multi-discrete PPO.

Notations	MADDPG	Multi-discrete PPO
Number of Episodes	1500	1500
Batch size	1024	64
Episode Length	20	20
Hidden Dimension	[64,64]	[64,64]
learning Rate	0.01	0.0003
Gamma	0.95	0.99
Soft update	0.01	–
Gae Lambda	–	0.95
vf coef	–	0.5
max grad norm	–	0.5

6.2.1. Real-World Network Configuration

We evaluated **CALM**'s performance in controlling a single intersection edge, using mixed-traffic flow demand based on Barcelona traffic, with an additional surge period from 3 PM to 8 PM (see Fig. 5 and Table 4). Our results show that adaptive lane models achieve better pedestrian and cyclist metrics when compared to a static lane configuration. Vehicle lane metrics—where static is expected to perform best—only suffer in the case of heuristic and PPO, with MADDPG performing similarly, if not slightly worse than a static configuration. Aside from the benefits to pedestrian and cyclist metrics, we also find that adaptive algorithms can provide lower (better) LOS scores to non-motorized lanes. Inspecting the resulting co-sharing policies, we see that PPO learns to always co-share lanes, while MADDPG stops sharing lanes at 07:30, which is around the time vehicle demand starts to sharply rise (see Fig. 3). The novel approach of dynamically determining optimal instances for shared usage of lanes by cyclists and pedestrians, guided by learned logic contingent upon mixed-traffic flow density, presents a significant advancement in addressing the challenges posed by static lane configurations. This methodology effectively mitigates issues associated with the misallocation of space during periods of low peak density for pedestrians and cyclists, concurrently with high vehicular traffic. Furthermore, the potential for accidents arising from the co-sharing of a lane during instances of elevated pedestrian and cyclist density is addressed through the incorporation of cooperative agents (i.e., by deciding to not co-share lanes), offering a comprehensive solution to enhance safety and optimize traffic flow in such dynamic scenarios.

Table 4: Quantitative results on a Single Edge of a 4-arm intersection using the proposed, static, heuristic, and multi-discrete PPO methods. Average waiting time is measured in seconds and average queue length is the number of stopped vehicles on a lane. Results highlighted in blue showcase the best performing model for a specific measure.

	CALM (proposed model)		**Static-FT**		**Heuristic**		**PPO**	
Average	Waiting Time	Queue Length	Waiting Time	Queue Length	Waiting Time	Queue Length	Waiting Time	Queue Length
Vehicle Lane	34.47	2.65	36.63	2.61	52.94	3.24	43.27	2.66
Pedestrian Lane	44.58	2.64	226.54	8.12	30.50	2.24	61.69	5.47
Cyclist Lane	6.98	0.52	60.05	2.48	8.01	0.58	30.14	1.07
Overall Avg.	28.67	1.93	107.74	4.40	30.48	2.02	45.03	3.06

Overall, the adaptive lane model produces the best results in most evaluation metrics, while distributing fairly the ROWs street space conditional on the dynamic nature of mixed-traffic flow rate. Moreover, the density threshold parameters provide additional flexibility to the transportation engineer for deciding what LOS measure is safe and acceptable to initiate the co-sharing of pedestrian and cyclist lanes. These results indicate that **CALM** is the best-performing model, and we then extend it beyond the single-edge case.

6.2.2. Centralized versus Decentralized Model

Finally, to investigate how generalizable and applicable the **CALM** model is to real-world scenarios, we compared a centralized vs. decentralized deployment of the adaptive lane models. The decentralized approach involved training the **CALM** model on a single edge and deploying it across all four arm edges

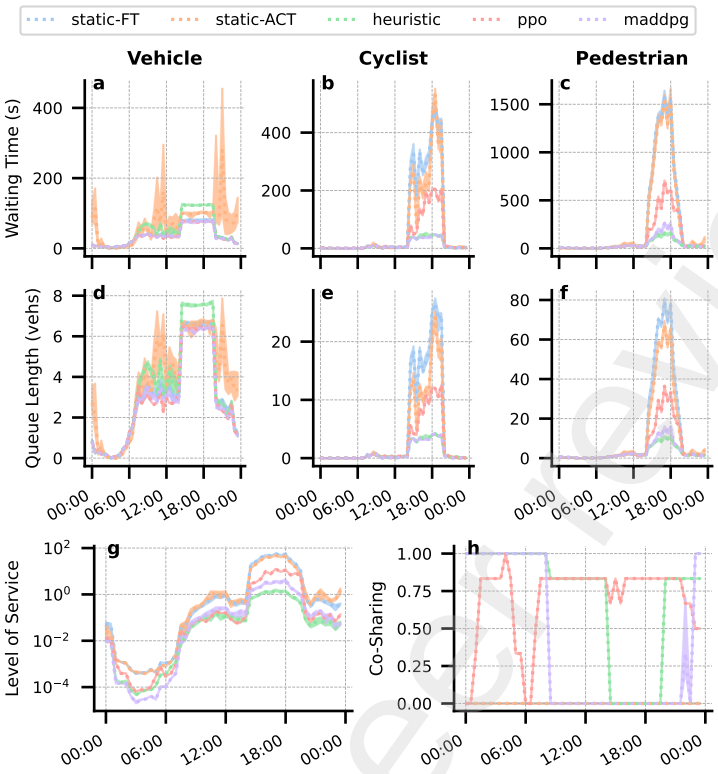

Figure 5: Traffic measures evaluated on a single edge for the different algorithms used in this work. The adaptive algorithms (heuristic, PPO, and MADDPG) all achieve better measures than static (fixed-time and actuated), except for heuristic and PPO having worse waiting times and queue lengths for vehicles (a and d). PPO learns to co-share more than MADDPG and heuristic.

during testing. On the other hand, the centralized approach entailed training the model simultaneously on all four edges. In both cases, the models were trained for 1500 episodes.

Comparison of mean waiting times and queue lengths show lower values for both metrics when using the decentralized **CALM** (Fig. 6). The decentralized model also has better LOS values and a more stable co-sharing policy through time for all four arms compared to the centralized version (Fig. 7). Despite having access to the overall network configuration during training, the centralized framework fails to learn an effective adaptive lane model. In contrast, the decentralized training approach proves to be more robust and capable of efficient generalization based on mixed-traffic flow scenarios observed during training on a single edge. Although these preliminary results are derived from a single intersection and should undergo rigorous testing, they indicate the potential for adapting the streets of a district or an entire city using the decentralized version of the adaptive lane model.

7. Conclusion and Future Work

We propose an RL-based cooperative adaptable lanes model (**CALM**) for safer shared space and improved mixed-traffic flow. The popular open-source microscopic traffic simulation SUMO was employed in conjunction with the OpenAI Gym environment for training and validation. After conducting a myriad of simulation and ablation studies, we observed that cooperative multi-agent methods using the MADDPG architecture outperformed static, heuristic, and multi-discrete PPO versions of adaptive lane models. Critical traffic metrics such as average waiting times and average queue lengths were minimized using **CALM**, and the ROWs were distributed fairly based on each road user's dynamic and current needs. Moreover, the

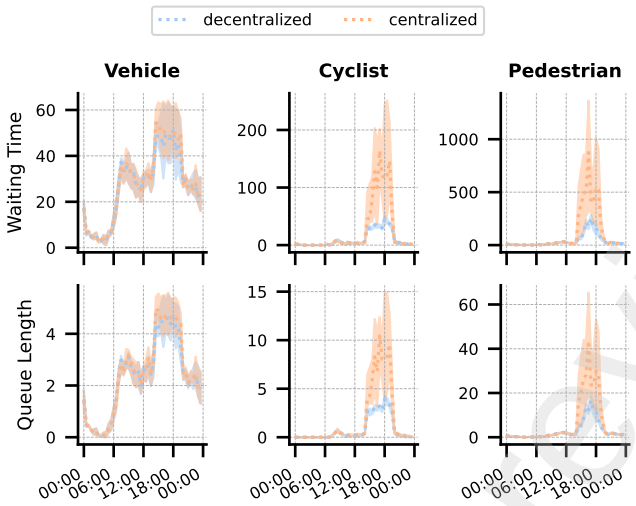

Figure 6: Comparing the impact of centralized and decentralized variation of **CALM** on the validation metrics. The results shown are averages over the 4 edges.

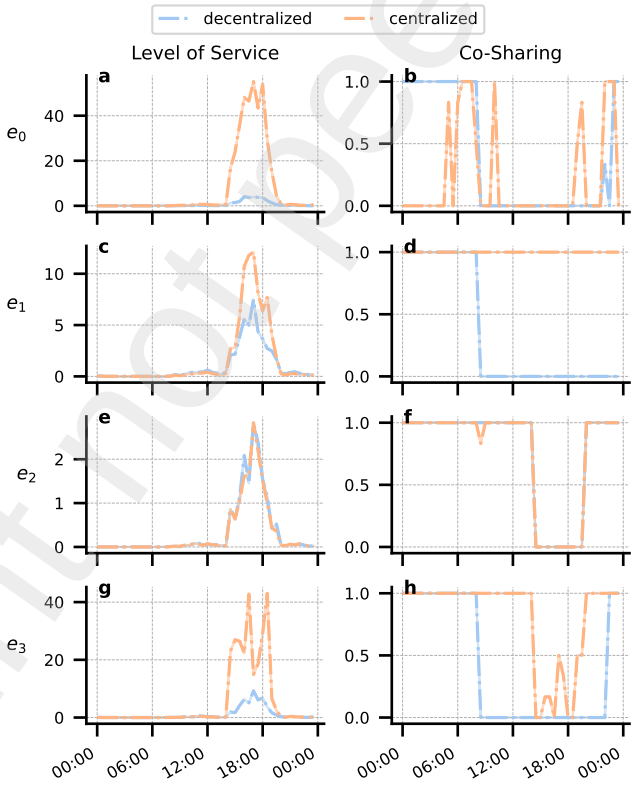

Figure 7: Comparing the impact of a centralized and decentralized variation of CALM on co-sharing and Level of Service (LOS) on individual edges. Lower LOS is better. The decentralized model performs well on all edges, with stable co-sharing policies.

LOS for vulnerable users was computed and preserved within a designer-selected, allowed threshold during lane adaption. Finally, we showcased that the proposed **CALM** model in its decentralized version outper-

formed the centralized model and provided optimistic expectancy for testing it over more extensive street networks. Our ablation studies revealed that co-sharing agents required more than unique observation (i.e., simulation-based traffic and street information) for robust learning of co-sharing.

However, the proposed model also has limitations. Firstly, the mixed-traffic flow is one-directional and does not account for left, right, and u-turns. We did not include turns to avoid the delay caused by traffic congestion at an intersection, which would introduce anomalies that result in *"vehicle teleportation"* in SUMO. Secondly, we do not account for queue spillbacks and congestion effects from the more extensive surrounding network and treat the network part of interest in isolation. Finally, we mirrored the upstream edge adaptation to the downstream edges (i.e., streets immediately after the intersection), thereby excluding bottleneck effects that can originate due to the mismatch of lane width.

In the future, we aim to investigate the efficacy of the distributed system by stress-testing the distributed RL model trained on one intersection and applying it to larger networks (e.g., a district or city). Finally, the formulation of LOS that currently gives equal weights to its variable may be improved by computing the precise significance of each influencing parameter. Additionally, implementing this dynamic allocation of ROWs in real settings remains a challenge, which should be further explored from the engineering, practical, legal, ethical, and cognitive points of view.

Acknowledgments

RKD, JASV, and DH acknowledge support by the Semantic Urban Elements module financed by the Future Cities Lab Global of the Singapore–ETH Centre, which was established collaboratively between ETH Zurich and the National Research Foundation Singapore. DD thanks the European Union's Horizon 2020 Research and Innovation Programme, the Distributed Intelligence and Technology for Traffic and Mobility Management (DIT4TraM) Project under Grant 953783.

References

[1] B. Brunekreef, Air pollution and life expectancy: Is there a relation?, 1997. doi:10.1136/oem.54.11.781.
[2] D. Helbing, S. Mahajan, R. H. Fricker, A. Musso, C. I. Hausladen, C. Carissimo, D. Carpentras, E. Stockinger, J. Argota Sanchez-Vaquerizo, J. C. Yang, M. C. Ballandies, M. Korecki, R. K. Dubey, E. Pournaras, Democracy by Design: Perspectives for digitally assisted, participatory upgrades of society, J. Comput. Sci. (2023) 102061. doi:10.1016/j.jocs.2023.102061.
[3] P. D. Norton, Fighting Traffic: The Dawn of the Motor Age in the American City, MIT Press, Cambridge, MA, 2008.
[4] D. Prytherch, Legal geographies-codifying the right-of-way: Statutory geographies of Urban mobility and the street, Urban Geogr. 33 (2012) 295–314. doi:10.2747/0272-3638.33.2.295.
[5] R. Lowe, Y. I. Wu, A. Tamar, J. Harb, O. Pieter Abbeel, I. Mordatch, Multi-agent actor-critic for mixed cooperative-competitive environments, Adv. Neural Inf. Process. Syst. 30 (2017).
[6] M. Behrisch, L. Bieker, J. Erdmann, D. Krajzewicz, SUMO–simulation of urban mobility: An overview, in: Proc. SIMUL 2011 Third Int. Conf. Adv. Syst. Simul., ThinkMind, 2011.
[7] E. Walraven, M. T. Spaan, B. Bakker, Traffic flow optimization: A reinforcement learning approach, Eng. Appl. Artif. Intell. 52 (2016) 203–212.
[8] A. Paul, S. Mitra, Deep reinforcement learning based traffic signal optimization for multiple intersections in ITS, in: 2020 IEEE Int. Conf. Adv. Netw. Telecommun. Syst. ANTS, IEEE, 2020, pp. 1–6.
[9] D. Pérez-Méndez, C. Gershenson, M. E. Lárraga, J. L. Mateos, Modeling adaptive reversible lanes: A cellular automata approach, PloS one 16 (2021) e0244326.
[10] Q. Ye, Y. Feng, E. Candela, J. Escribano Macias, M. Stettler, P. Angeloudis, Spatial-temporal flows-adaptive street layout control using reinforcement learning, Sustainability 14 (2021) 107.
[11] D. Banister, Unsustainable transport: City transport in the new century, in: Unsustainable Transport: City Transport in the New Century, Routledge Taylor & Francis Group, 2005, pp. 1–292. doi:10.4324/9780203003886.
[12] O. Tengilimoglu, O. Carsten, Z. Wadud, Implications of automated vehicles for physical road environment: A comprehensive review, Transp. Res. Part E Logist. Transp. Rev. 169 (2023) 102989. doi:10.1016/j.tre.2022.102989.
[13] W. Riggs, B. Appleyard, M. Johnson, A design framework for livable streets in the era of autonomous vehicles, Urban Plan. Transp. Res. 8 (2020) 125–137. doi:10.1080/21650020.2020.1749123.
[14] T. Schönhofer, K. Bogenberger, A comprehensive review on managed lanes in europe, in: 101th Annu. Meet. Transp. Res. Board, Washington D.C., USA, 2022.
[15] S. Lämmer, D. Helbing, Self-control of traffic lights and vehicle flows in urban road networks, J. Stat. Mech. Theory Exp. 2008 (2008). doi:10.1088/1742-5468/2008/04/P04019. arXiv:0802.0403.

[16] R. Mangiaracina, A. Perego, G. Salvadori, A. Tumino, A comprehensive view of intelligent transport systems for urban smart mobility, Int. J. Logist. Res. Appl. 20 (2017) 39–52. doi:10.1080/13675567.2016.1241220.
[17] D. Metz, Tackling urban traffic congestion: The experience of London, Stockholm and Singapore, Case Stud. Transp. Policy 6 (2018) 494–498. doi:10.1016/j.cstp.2018.06.002.
[18] B. Hamilton-Baillie, Shared space: Reconciling people, places and traffic, Built Environ. 34 (2008) 161–181. doi:10.2148/benv.34.2.161.
[19] M. Lydon, A. Garcia, Tactical Urbanism, volume 1, Island Press/Center for Resource Economics, Washington, DC, 2015. doi:10.5822/978-1-61091-567-0.
[20] S. Shaheen, Shared Mobility: The Potential of Ridehailing and Pooling, Springer, 2018.
[21] E. Ravazzoli, G. P. Torricelli, Urban mobility and public space. A challenge for the sustainable liveable city of the future, J. Public Space 2 (2017) 37–50.
[22] S. Barr, S. Lampkin, L. Dawkins, D. Williamson, Shared space: Negotiating sites of (un)sustainable mobility, Geoforum 127 (2021) 283–292. doi:10.1016/j.geoforum.2021.11.012.
[23] R. Imrie, Auto-disabilities: The case of shared space environments, Environ. Plan. A 44 (2012) 2260–2277.
[24] D. Beitel, J. Stipancic, K. Manaugh, L. Miranda-Moreno, Assessing safety of shared space using cyclist-pedestrian interactions and automated video conflict analysis, Transp. Res. Part Transp. Environ. 65 (2018) 710–724.
[25] A. Nikiforiadis, S. Basbas, F. Mikiki, A. Oikonomou, E. Polymeroudi, Pedestrians-cyclists shared spaces level of service: Comparison of methodologies and critical discussion, Sustainability 13 (2021) 361.
[26] RH. Grzebieta, AM. McIntosh, S. Chong, Pedestrian-cyclist collisions: Issues and risk, in: Australas. Coll. Road Saf. Conf., 2011, pp. 1–12.
[27] S. O'Hern, J. Oxley, Pedestrian injuries due to collisions with cyclists Melbourne, Australia, Accident Anal. Prev. 122 (2019) 295–300. doi:10.1016/j.aap.2018.10.018.
[28] S. Chong, R. Poulos, J. Olivier, W. L. Watson, R. Grzebieta, Relative injury severity among vulnerable non-motorised road users: Comparative analysis of injury arising from bicycle–motor vehicle and bicycle–pedestrian collisions, Accident Anal. Prev. 42 (2010) 290–296. doi:10.1016/j.aap.2009.08.006.
[29] A. Nikiforiadis, S. Basbas, M. I. Garyfalou, A methodology for the assessment of pedestrians-cyclists shared space level of service, J. Clean. Prod. 254 (2020) 120172.
[30] Umbrellium, Case Study: Make Roads Safer, More Responsive & Dynamic, 2017. https://umbrellium.co.uk/case-studies/south-london-starling-cv/, last checked on 2023-08-29.
[31] T. Toledo, D. Zohar, Modeling duration of lane changes, Transportation Research Record 1999 (2007) 71–78. URL: https://journals.sagepub.com/doi/10.3141/1999-08http://journals.sagepub.com/doi/10.3141/1999-08. doi:10.3141/1999-08.
[32] T. Tran, C. Parker, M. Tomitsch, A Review of Virtual Reality Studies on Autonomous Vehicle–Pedestrian Interaction, IEEE Trans. Hum. Mach. Syst. 51 (2021) 641–652. doi:10.1109/THMS.2021.3107517.
[33] T. Andersen, Shared space is a new type of design based on mutual consideration, 2019. URL: https://cyclingsolutions.info/shared-space/.
[34] R. Bellman, A markovian decision process, J. Math. Mech. (1957) 679–684.
[35] R. S. Sutton, A. G. Barto, Reinforcement Learning: An Introduction, MIT press, 2018.
[36] D. Silver, G. Lever, N. Heess, T. Degris, D. Wierstra, M. Riedmiller, Deterministic policy gradient algorithms, in: Proc. 31st Int. Conf. Mach. Learn., 2014, pp. 387–395.
[37] V. Konda, J. Tsitsiklis, Actor-critic algorithms, Adv. Neural Inf. Process. Syst. 12 (1999).
[38] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, O. Klimov, Proximal policy optimization algorithms, 2017. arXiv:1707.06347.
[39] J. Schulman, P. Moritz, S. Levine, M. Jordan, P. Abbeel, High-dimensional continuous control using generalized advantage estimation, 2015. arXiv:1506.02438.
[40] M. A. Taylor, S. Somenahalli, Travel time reliability and the bimodal travel time distribution for an arterial road, Road Transp. Res. J. Aust. N. Z. Res. Pract. 19 (2010) 37–50.
[41] National Academies of Sciences and Medicine (U.S.). Engineering Board. Transportation Research, Highway Capacity Manual: A Guide for Multimodal Mobility Analysis., Transportation Research Board, Washington, D.C, 2016.
[42] TC Group Solutions, Traffic index, 2023.
[43] J. Argota Sánchez-Vaquerizo, D. Cardoso Llach, The social life of small urban spaces 2.0, in: Communications in Computer and Information Science, volume 1028, Springer, Singapore, 2019, pp. 295–310. doi:10.1007/978-981-13-8410-3_21.
[44] Ajuntament de Barcelona, Mobility gauging detail of the city of Barcelona - Datasets, 2019.
[45] M. E. Hallenbeck, M. Rice, B. Smith, C. Cornell-Martinez, J. Wilkinson, Vehicle Volume Distributions by Classification, Technical Report, Washington State Transportation Center, Chaparral Systems Corporation, 1997.
[46] J. Argota Sánchez-Vaquerizo, Getting real: The challenge of building and validating a large-scale digital twin of barcelona's traffic with empirical data, ISPRS Int. J. Geo-Inf. 11 (2021) 24. doi:10.3390/ijgi11010024.
[47] Ajuntament de Barcelona, Bicycle | barcelona city council, 2023.
[48] S. Rueda, Les superilles per al disseny de noves ciutats i la renovació de les existents. El cas de Barcelona, Pap. Regió Metrop. Barc. Territ. Estratègies Planejament (2017) 78–93.
[49] Ajuntament de Barcelona, Medida de gobierno. Superilla Barcelona para regenerar Barcelona y sus barrios, 2022.
[50] A. Raffin, A. Hill, A. Gleave, A. Kanervisto, M. Ernestus, N. Dormann, Stable-baselines3: Reliable reinforcement learning implementations, J. Mach. Learn. Res. 22 (2021) 12348–12355.
[51] J. T. Kim, S. Ha, Observation space matters: Benchmark and optimization algorithm, in: 2021 IEEE Int. Conf. Robot. Autom. ICRA, IEEE, 2021, pp. 1527–1534. doi:10.1109/icra48506.2021.9561019.

Appendix A. Ablation Studies

Appendix A.1. Lane Density Threshold for Co-Sharing

We performed ablation studies on the lane density threshold value λ as a policy design parameter. We generated a synthetic dataset to identify the candidate values of λ. The dataset was comprised of simulated 24000 datapoints for random variations of α, β, and γ, along with varying f_v, f_c, and f_p. We removed all data points that violated constraints (i.e., $l_v < 3.6$ and $l_p/l_c < 1$) and those that did not co-share (i.e., $\gamma = 0$). From the distribution of the density values of the co-shared lane, we determined candidate density thresholds $\lambda = \{1.75, 4.87, 9.27, 16.69, 35.64\}$, corresponding to five quantiles (5%, 25%, 50%, 75%, and 95%) of the distribution.

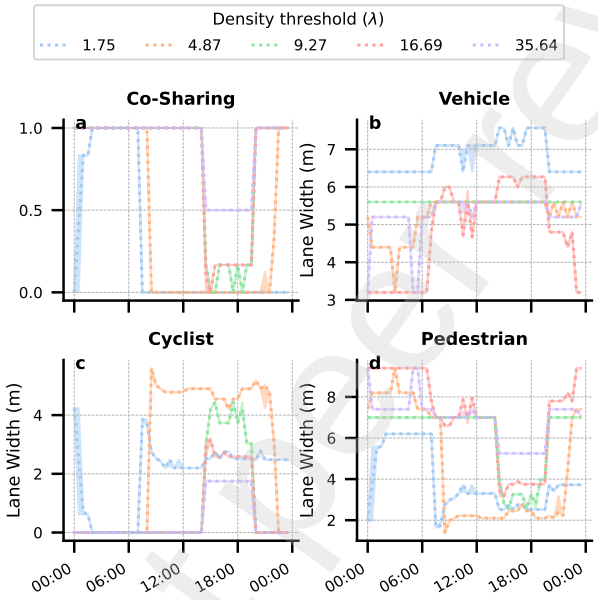

Figure A.1: Co-sharing and lane widths of CALM for different density thresholds λ. Co-sharing stops before the rise in traffic demand at 8 AM for $\lambda \in \{1.75, 4.87\}$. The model learns to allocate most space for cyclists for $\lambda = 4.87$ while keeping the vehicle lanes wide enough for two cars.

We then simulated CALM using these different values of λ. In particular, we focused on the co-sharing of lanes based on the density threshold λ (Fig. A.1). We found that at lower values of λ, the model decided to disallow co-sharing at earlier timeslots (near the AM spike in demand for $\lambda \in \{1.75, 4.87\}$) and reverted to co-sharing once the traffic flows dissipate beyond 9 PM. At higher density thresholds ($\lambda \in \{9.27, 16.69, 35.64\}$), the model was more tolerant to high densities and only disallowed co-sharing from 3 PM to 8 PM, which was the period of the surge in demand. However, allowing for higher densities while co-sharing should result in worse LOS values, which, depending on the overall policy objectives of a city planner, can be undesirable. Since the density threshold is a parameter that can be chosen to suit city planning goals, a planner can select a density threshold that avoids co-sharing during the peak congestion hours of the street. Thus, we chose a density threshold of $\lambda = 4.87$ as the basis of the rest of this work. This choice also promoted cyclist lane use while keeping the vehicle lanes wide enough for two cars.

Appendix A.2. Observation Spaces

Reinforcement learning algorithms are known to be sensitive to problem formulation, specifically concerning observation spaces [51]. Therefore, to understand the impact of varying local observation of an agent on the policy learned and to identify the most effective subset of state spaces for each agent in CALM, we performed ablation studies with three varying sets of observation spaces (Table A.5). These are the Minimal

Table A.5: Three varying observation spaces (proposed, minimal, and maximal sets) for the ablation study.

Agent	Proposed Set			Minimal Set			Maximum Set		
	0	1	2	0	1	2	0	1	2
Vehicle Lane Width	✓		✓	✓			✓	✓	✓
Pedestrian Lane Width		✓	✓		✓	✓	✓	✓	✓
Cyclist Lane Width		✓	✓		✓	✓	✓	✓	✓
Pedestrian+Cyclist Lane Width	✓						✓	✓	✓
Vehicle Lane Occupancy	✓		✓	✓			✓	✓	✓
Pedestrian Lane Occupancy		✓	✓		✓	✓	✓	✓	✓
Cyclist Lane Occupancy		✓	✓		✓	✓	✓	✓	✓
Vehicle Lane Density	✓		✓	✓			✓	✓	✓
Pedestrian Lane Density			✓			✓	✓	✓	✓
Cyclist Lane Density			✓			✓	✓	✓	✓
Co-sharing (On/Off)			✓			✓	✓	✓	✓

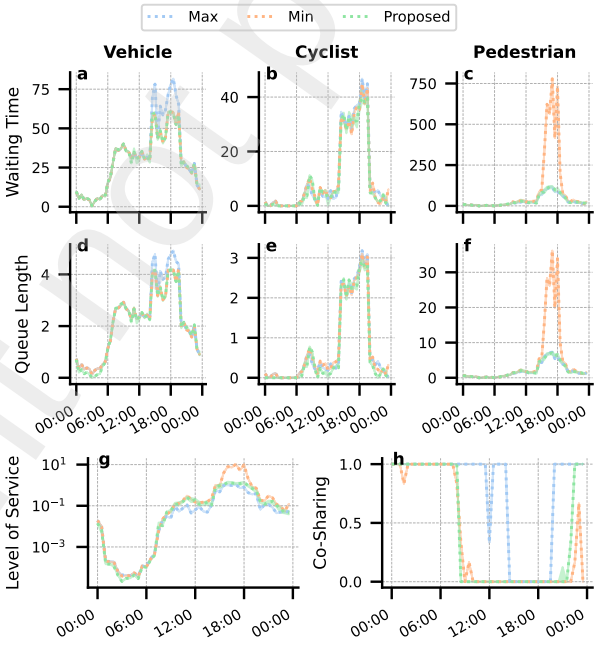

Figure A.2: Ablation of the observation space sets on the traffic measures. All sets perform similarly for vehicle and cyclist measures, except the Minimal set (Min), which underperforms on the pedestrian measures.

(fewest observations that are local and unique to each lane agent), Maximum (each agent receives complete information regarding all the lanes), and Proposed sets (combines unique and redundant information for each agent, intuitively selected based on ablation). Fig. A.2 presents the impact of different observation spaces on the validation measures. We noticed the Minimum set under-performed (for pedestrians) while the intuitively selected Proposed and Maximum sets performed similarly.