Author/s:
Chan, LS;Zhang, X;Nassir, N;Sarvi, M

Title:
Regulating jaywalking behaviour in adaptive traffic signal control using a novel deep reinforcement learning approach

Date:
2026-06

Citation:
Chan, L. S., Zhang, X., Nassir, N. & Sarvi, M. (2026). Regulating jaywalking behaviour in adaptive traffic signal control using a novel deep reinforcement learning approach. Multimodal Transportation, 5 (2), https://doi.org/10.1016/j.multra.2025.100271.

Persistent Link:
https://hdl.handle.net/11343/361568

Multimodal Transportation 5 (2026) 100271

Contents lists available at ScienceDirect

Multimodal Transportation

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

Full Length Article

Regulating jaywalking behaviour in adaptive traffic signal control using a novel deep reinforcement learning approach

Lok Sang Chan, Xiaocai Zhang*, Neema Nassir, Majid Sarvi

Department of Infrastructure Engineering, Faculty of Engineering & Information Technology, The University of Melbourne, VIC, 3010, Australia

ARTICLE INFO

Keywords:
Multi-objective optimisation
VRU safety evaluation
Pedestrian Behaviour
Deep learning
Adaptive signal control

ABSTRACT

This paper presents a deep reinforcement learning based adaptive traffic signal control framework that explicitly models jaywalking at urban intersections. We integrate a behaviourally grounded Jaywalking Decision model, which endogenises red light violations through waiting time and dynamic gap acceptance, with a Branching Double Deep Q-Network and a comprehensive hybrid action space that controls both phase selection and subphase timing. A multiobjective reward balances delay and jaywalking related safety risk, enabling the controller to respond to non-compliance as it emerges. The framework is evaluated in a multimodal microsimulation of a real intersection in Melbourne across four naturalistic demand scenarios and against both actuated control and established reinforcement learning baselines. Results show that the proposed approach reduces observed pedestrian delay and jaywalking relative to actuated control while achieving a more balanced safety and efficiency profile than single objective or alternative learning architectures. The analysis highlights context-dependent trade-offs that are relevant for policy, since the controller can adapt timing to mitigate non-compliance without assuming full pedestrian obedience. The contributions are threefold: a realistic jaywalking model linked to observable states, a high-granularity action space for multimodal control, and an integrated learning framework that jointly manages delay and safety risk. The proposed framework not only facilitates a more equitable traffic operation system but also offers the first scalable approach to managing risky behaviours in urban traffic environments.

1. Introduction

The pursuit of sustainable transportation development is vital for fostering urban environments that support environmental, social, and economic sustainability (Roy et al., 2025). Central to sustainable mobility is a shift from car dependency to multimodal, transit-oriented development. However, first and last-mile (FLM) connectivity issues persist, hindering the widespread adoption of public transportation systems and active travel modes (Banister, 2008; Boarnet et al., 2017). A study conducted in Brisbane, Australia, emphasised the importance of enhancing public transportation accessibility, intersection density, and pedestrian connectivity to support multimodal travel at the household level (Lu et al., 2022).

Emerging personal mobility alternatives, such as electric scooters and e-bikes, complement walking and cycling to improve FLM connectivity (Hardt and Bogenberger, 2019). Nonetheless, socio-demographic factors and infrastructure availability significantly influence the choice and safety of these active travel modes (Meng et al., 2016). Ensuring safe and efficient pedestrian environments, particularly at intersections, is thus fundamental to reducing private vehicle reliance and improving public health. Unsafe or inefficient

* Corresponding author.
E-mail addresses: loksangc@student.unimelb.edu.au (L.S. Chan), xiaocai.zhang@unimelb.edu.au (X. Zhang), neema.nassir@unimelb.edu.au (N. Nassir), majid.sarvi@unimelb.edu.au (M. Sarvi).

https://doi.org/10.1016/j.multra.2025.100271
Received 10 April 2025; Received in revised form 2 November 2025; Accepted 3 November 2025

intersection designs significantly impact pedestrian safety and discourage walking, undermining sustainability efforts (Kunaratnam et al., 2022; Liao et al., 2020).

The Organization (2023) reported 1.19 million road traffic fatalities in 2021, with VRU accounting for 50 % of these deaths. In Australia, road crash fatalities reached 1,194 in 2022, a 5.8 % increase from 2021 (Infrastructure and, BITRE). VRU represented 37.1 % of these fatalities, with pedestrian deaths rising by a troubling 23.3 %. Intersections are particularly hazardous for vehicle-pedestrian collisions due to their complex traffic dynamics. The growing use of active modes of transportation necessitates safe travel environments, prompting policymakers and researchers to focus on reducing traffic fatalities. Victoria, Australia, for instance, aims to halve fatalities and serious injuries (FSI) by 2030 and eliminate fatal accidents by 2050 (Victoria, 2020).

Numerous studies have addressed the safety concerns of vulnerable road users (VRU) at intersections through artificial intelligent (AI)-based adaptive traffic signal control (ATSC) and operations (Li and Sun, 2019; Ma et al., 2014; Wang et al., 2019a; 2024a). However, these models typically assume rational pedestrian behaviour, presuming that pedestrians obey road rules. This necessitates research to enhance jaywalking behaviour modelling in AI-based ATSC optimisation. Jaywalking presents significant challenges to urban traffic management and safety that should not be overlooked. Research has identified distraction and jaywalking as major contributors to FSI among VRU at intersections (Choi et al., 2013; Samerei et al., 2021; Shiwakoti et al., 2020; Wang et al., 2020). Jaywalking often refers to the violation of pedestrian road rules. In Victoria, Australia, a pedestrian is considered jaywalking if they fail to comply with nine specific rules outlined in the *Road Safety (General) Regulations 2019* (Victoria, 2019). These violations generally include failing to obey signs and traffic lights, as well as crossing at inappropriate locations. Studies have investigated the behaviour underlying jaywalking and found that inefficient allocation of road rights often leads to prolonged waiting times at traffic signals, thereby tempting pedestrians to cross streets unsafely when they perceive a suitable gap (Brosseau et al., 2013; Li, 2013; Raoniar and Maurya, 2022; Zhu et al., 2021).

To address these research gaps, this study introduces an advanced microsimulation-based Jaywalking Decision (JD) model that explicitly captures two behavioural mechanisms often overlooked in adaptive signal control frameworks: pedestrians' tolerance for waiting and their dynamic acceptance of traffic gaps. By endogenising non-compliance through these mechanisms, the JD model generates realistic red-light violations that interact with the control system rather than being treated as exogenous "errors". In parallel, we propose a comprehensive hybrid action space for signal control that spans a rich set of phase configurations and sub-phase timing adjustments. This granularity enables context-aware signal policies that can respond to real-time multimodal conditions and emergent jaywalking behaviour.

The core novelty of the research lies not in the learning backbone itself but in its integration. We couple the behaviourally grounded JD model with a Branching Double Deep Q-Network (BDDQ) and a multi-objective reward that prices both delay and jaywalking-related safety risk. The high-dimensional action space exposes both phase selection and sub-phase duration control within each cycle, allowing the controller to adapt pedestrian protection and clearance timing as non-compliance evolves. To the best of our knowledge, this is the first study to explicitly model jaywalking within a DRL-based ATSC system operating over a high-dimensional, multimodal action space at sub-phase resolution.

The key contributions of this study are summarised as follows:

1. A novel jaywalking decision (JD) model is developed to simulate red-light violations based on empirically grounded representations of pedestrian waiting tolerance and dynamic gap acceptance, thereby advancing the behavioural realism of traffic microsimulation.
2. A distinctly comprehensive hybrid action space is formulated, enabling context-sensitive phase selection and sub-phase duration optimisation. This structure significantly expands the operational flexibility of signal control beyond what existing DRL-based approaches offer.
3. The Branching Double Deep Q-Network for Jaywalking and Delay Optimisation (BDDQ-JDO) is introduced, offering a unique integration of pedestrian behaviour modelling and adaptive traffic control optimisation to jointly reduce jaywalking incidence and improve intersection performance.

By embedding behavioural realism in both the simulation environment and the control logic, the study reconceptualises how adaptive control can manage pedestrian non-compliance. The proposed framework provides a practical pathway for authorities to balance safety and efficiency under varying demand conditions, and demonstrates how explicit modelling of jaywalking, combined with high-granularity control, can yield policy-relevant improvements at urban intersections.

2. Literature review

The increasing promotion of active transportation modes has intensified the complexity of managing urban traffic flow. Advances in Internet of Things (IoT) technologies within transportation provide promising opportunities, particularly through the integration of artificial intelligence (AI) into traffic control systems, aiming to manage these complexities effectively. Existing research has primarily addressed the optimisation of intersection pedestrian crossings from multiple perspectives. This review discusses three critical areas: (1) Multi-objective adaptive traffic signal control (ATSC) optimisation for vulnerable road users (VRU), (2) Deep reinforcement learning (DRL)-based adaptive traffic signal control, and (3) Understanding jaywalking behaviour.

2.1. Multi-objective traffic signal control with VRU

Traffic engineers continuously seek to optimise intersection safety and efficiency, with one direct approach being the segregation of conflicting flows. Grade-separated pedestrian systems (GSPS), including underground pedestrian systems (UPS) and skywalk networks, represent complete separation from street-level traffic, offering both safety and economic benefits, particularly in densely populated urban areas (Cui et al., 2013; Murakami et al., 2021). However, these systems raise concerns about decreased street-level activity and substantial infrastructure costs.

Alternatively, street-level intersection designs such as continuous flow intersections (CFI), incorporating pedestrian crossing facilities, have been explored (Jagannathan and Bared, 2005). Features including marked crosswalks, speed cushions, visibility enhancements, and refuge islands offer protection for VRU while preserving vehicular efficiency (Leden et al., 2006; Tang et al., 2020; Wang et al., 2019b). Despite their safety benefits, such facilities are often limited by site-specific geometry, urban density, and high implementation costs. Consequently, operational modifications, including innovative pedestrian signal phasing, offer a lower-cost alternative.

Research exploring pedestrian signal phasing dates back to the late 20th century, with Abrams and Smith (1977) investigating and recommending phasing strategies tailored to specific pedestrian volumes and vehicle-turning movements. Subsequent studies have examined various operational schemes including the exclusive, concurrent, and early or late release of pedestrians, based on different data sources and demographic settings. These include pedestrian crash records from the United States (Chen et al., 2015; Zegeer et al., 1982) and Israel (Zaidel and Hocherman, 1987), as well as conflict data from Sweden (Gårder, 1989), the United States (Arun et al., 2023; Bechtel et al., 2004; Zhang et al., 2015), and Canada (Kattan et al., 2009). Findings suggest that the proper use of exclusive pedestrian phases is associated with a reduction in pedestrian accidents and conflicts, particularly in areas with moderate pedestrian volumes. However, studies comparing performance before and after the implementation of exclusive signals also reported that their effectiveness on pedestrian safety is influenced by the rate of jaywalkers (Bechtel et al., 2004; Gårder, 1989; Kattan et al., 2009).

In addition to traditional control patterns and signal timing designs, safety-oriented adaptive traffic signal controls (ATSC) are gaining prominence in proactive traffic management. Earlier ATSC optimisation methods focused solely on transportation efficiency (Li et al., 2021; Liu et al., 2017; Wang et al., 2021; Yazdani et al., 2023), but the rise of mode shift has drawn attention to road user safety-oriented optimisation. Modern multi-objective ATSC utilise surrogate safety measures (SSMs) to predict near-future safety risks. A series of studies have emerged aiming to incorporate safety into ATSC, which can be further classified based on the evaluated interaction, such as vehicular (Chan et al., 2025; Du et al., 2022; Gong et al., 2020; Khamis et al., 2012; Li and Sun, 2019; Stevanovic et al., 2015) and multimodal with VRU (Li and Sun, 2019; Ma et al., 2014; 2011; Wang et al., 2019a; 2024a).

Research has demonstrated the effectiveness of multi-objective optimisation in improving pedestrian safety and efficiency at two-stage midblock crosswalks (Ma et al., 2011) and four-approach signalised intersections (Ma et al., 2014), highlighting the importance of balancing pedestrian delays with vehicular congestion through advanced signal timing strategies.

Li and Sun (2019) developed a multi-objective optimisation model for signal settings and lane assignments at intersections, achieving notable improvements in transportation efficiency, energy consumption, and road safety through the application of the cell mapping method and microscopic traffic simulations. Another study introduced a group-based signal timing optimisation model that effectively reduces traffic conflicts and delays at signalised intersections with mixed traffic flows, demonstrating significant improvements in safety and efficiency compared to traditional methods (Wang et al., 2019a). Wang et al. (2024a) proposed an integrated optimisation model combining pedestrian control patterns and signal timing plans, demonstrating significant improvements in traffic efficiency and safety at intersections by reducing pedestrian-vehicle conflicts and delays.

2.2. DRL-based adaptive traffic siganl control

Recent advances in adaptive traffic signal control (ATSC) increasingly employ deep reinforcement learning (DRL) techniques to dynamically optimise intersection operations, responding effectively to fluctuating traffic conditions. Cai and Wei (2024) introduced PN_D3QN, an enhanced DRL model integrating advanced network structures, which significantly reduced vehicle delays and queue lengths compared to traditional control methods.

Addressing challenges posed by limited data availability, Wang et al. (2024c) demonstrated the efficacy of DRL using sparse vehicle trajectory data, significantly improving signal timing with minimal data penetration. Similarly, Wang et al. (2025) presented a vehicle-infrastructure cooperative control framework, integrating trajectory and infrastructure data to optimise traffic flow and reduce delays.

Hybrid action space frameworks within DRL methods have also been explored. Zhang et al. (2023) applied a soft actor-critic algorithm combining discrete and continuous action spaces, achieving improved efficiency and adaptability in complex scenarios. Another recent study by Luo et al. (2024) further demonstrated the potential of hybrid action spaces in traffic signal control, enhancing both efficiency and safety. Zhang et al. (2025) introduced a masked DRL method based on soft actor-critic framework to address the fairness issues among multimodal transportation signal coordination.

Despite these advances, explicit consideration of pedestrian behaviours, particularly jaywalking, remains insufficiently addressed. Moreover, comprehensive hybrid action spaces incorporating detailed phase configurations and duration adjustments remain underexplored. These gaps highlight the need for DRL methods explicitly designed to optimise multimodal interactions and pedestrian safety.

2.3. Understanding jaywalking behaviour

Although research has demonstrated that the application of AI-based ATSC has the potential to enhance the safety and efficiency of urban transportation networks, concerns persist regarding their effectiveness, particularly in relation to pedestrian compliance. Pedestrian behaviour, especially jaywalking, is frequently underrepresented in the existing literature. In a study by Ivan et al. (2017), pedestrian compliance within exclusive phasing crosswalks in Connecticut, United States, was investigated. The findings indicate that pedestrians are not necessarily more likely to jaywalk in exclusive phasing crosswalks. In fact, pedestrians in the study tended to treat all signals as concurrent, disregarding the exclusive phasing. The study also suggested that pedestrian compliance is more strongly correlated with factors such as the length of the crosswalk, vehicle speed limits, and traffic volume. These findings underscore the importance of considering jaywalking behaviour in the optimisation of ATSC, particularly to account for the effects of pedestrian red-light violations.

In examining pedestrian behaviour at crossings, various studies have emphasised the significant influence of demographic and environmental factors on crossing decisions. Demographic factors such as age, gender, education level, trip purpose, and cultural differences are widely reported in the literature. For instance, males are generally more prone to taking risks compared to females (Anik et al., 2021; Brosseau et al., 2013; Dıaz, 2002; Hamed, 2001; Li, 2013; Mamun et al., 2020; O'Hern et al., 2020; Rosenbloom, 2009; Russo et al., 2018; Yang et al., 2015; Zhu et al., 2021). Conversely, females have been found to be more sensitive to traffic volume (Zhu et al., 2021).

Moreover, studies conducted in various countries, including Australia (O'Hern et al., 2020), Bangladesh (Anik et al., 2021), Canada (Brosseau et al., 2013), China (Guo et al., 2011; Li, 2013; Yang et al., 2015; Zhuang et al., 2018), France (Dommes et al., 2015), India (Raoniar and Maurya, 2022), Jordan (Hamed, 2001), and the United States (Mamun et al., 2020; Russo et al., 2018), have consistently reported that the likelihood of pedestrians taking risks decreases with age. This trend may be attributed to the reduced mobility often observed among senior pedestrians (Guo et al., 2011; Russo et al., 2018).

However, a study focusing on pedestrians in Hong Kong, China presents a contrasting observation, where senior pedestrians were found to violate red lights more frequently. This behaviour is potentially due to lower educational attainment and poorer judgement (Zhu et al., 2021). The study further suggested that compliance, across all age ranges, increases in the presence of companions and under conditions of higher traffic volume, particularly when heavy vehicles are present.

An observational study by Jay et al. (2020) comparing pedestrian behaviour at signalised intersections in France and Japan revealed significant cultural differences in hesitation times and subsequent crossing actions. Japanese pedestrians exhibited longer periods of uncertainty and were more likely to follow others when crossing against a red light, highlighting the cultural influence on pedestrian compliance.

Environmental factors, including group size, are critical in understanding red light violations by pedestrians at signalised intersections. The influence of group size is multifaceted and interacts with various other factors. Notably, larger groups of adults, whether moving together or waiting at an intersection, generally exhibit higher compliance with traffic signals (Brosseau et al., 2013; Dıaz, 2002; Dommes et al., 2015; Raoniar and Maurya, 2022; Rosenbloom, 2009; Zhu et al., 2021). In contrast, teenagers in groups are more likely to ignore traffic signal controls (Hamed, 2001; Rosenbloom, 2009).

Additionally, certain environmental characteristics of intersections, such as the presence of crosswalk markings (Anik et al., 2021; Marisamynathan and Vedagiri, 2018), nearby parked vehicles (Dommes et al., 2015), refuge islands (Guo et al., 2011; Marisamynathan and Vedagiri, 2018), countdown timers (Brosseau et al., 2013; Keegan and O'Mahony, 2003), and shorter crosswalk lengths (Anik et al., 2021; Guo et al., 2011; Hamed, 2001; Marisamynathan and Vedagiri, 2018), are associated with a lower likelihood and risk of red light violations.

For instance, research in Ireland demonstrated that installing countdown timers at signalised pedestrian crossings significantly improves compliance with traffic signals. This study found an approximate 11 % reduction in red light violations, underscoring the importance of real-time information in influencing pedestrian behaviour (Keegan and O'Mahony, 2003).

In addition to the aforementioned factors, pedestrian crossing decisions are often evaluated and modelled based on the waiting time at the kerbside or the number of attempts needed before a successful crossing (Anik et al., 2021; Brosseau et al., 2013; Guo et al., 2011; 2019; Hamed, 2001; Li, 2013; Marisamynathan and Vedagiri, 2018; Raoniar and Maurya, 2022; Zhu et al., 2021). Generally, the literature indicates that longer waiting times at intersections significantly increase the likelihood of pedestrians committing violations, particularly among young adults and men. This finding underscores the need for optimised signal timing to improve pedestrian compliance and safety. For instance, the study by Guo et al. (2011) revealed that 50 % of pedestrians would still obey the traffic signal after waiting for 50 s. Similarly, Marisamynathan and Vedagiri (2018) found that 46 % of pedestrians violated the traffic signal to save time during a trip to work or for convenience. Guo et al. (2019) further observed that red-light violations at two-stage crossings are significantly higher in the second stage, particularly at intersections with longer total wait times.

The development of predictive models for pedestrian behaviour has also advanced through the use of mixture models and joint hazard-based duration models. For example, Li (2013) developed a mixture model to predict pedestrians' intended waiting times at signalized intersections, identifying key behavioural patterns and demographic influences on crossing decisions during the red phase. This study revealed a U-shaped distribution of intended waiting times, highlighting two major groups: immediate crossers, who cross upon arrival during the red phase, and gap acceptors, who initially wait but then cross during the red phase when they perceive an acceptable gap in traffic. The overall average intended waiting time for all pedestrian types was found to be 48 s. Similarly, Yang et al. (2015) developed a joint hazard-based duration model combining logit (for immediate crossers) and Weibull accelerated failure time models (for gap acceptors) to accurately estimate pedestrian waiting times at signalised intersections. The results highlighted the significant impact of environmental and demographic factors on crossing behaviour.

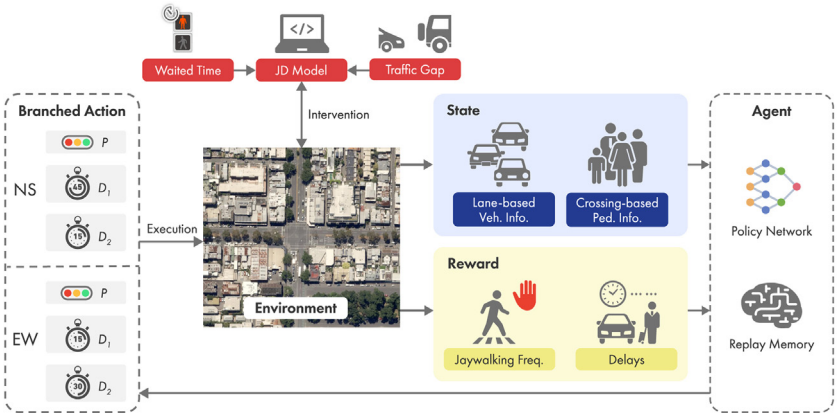

Fig. 1. Overview of the workflow in the proposed BDDQ-JDO framework.

More recently, scholars have begun to study pedestrian crossing decisions through the lens of psychophysical factors, such as gap acceptance, trust in technology, and the decision-making processes associated with approaching vehicles. For instance, Tian et al. (2022) employed a psychophysics-based gap acceptance model to elucidate how higher vehicle speeds lead to increased unsafe pedestrian crossings by leveraging visual looming cues. This study highlights the critical impact of speed on pedestrian decision-making in virtual reality simulations. Similarly, Dong et al. (2024) found that pedestrians are more likely to engage in jaywalking in the presence of autonomous vehicles, particularly in conditions of high visibility and low traffic. These findings suggest the need for regulatory measures such as speed limits and the disabling of autonomous control functions in areas with high pedestrian traffic.

Further research demonstrated that group behavioural cues can increase red-light violation rates in low-risk situations through automatic associative processes. In contrast, vehicle hazard cues reduce red-light violations in high-risk situations by engaging controlled analytical processes, as evidenced by eye-tracking experiments (Du et al., 2024). Additionally, video analysis from China revealed that jaywalkers' crossing decisions are influenced by vehicle behaviours, with aggressive vehicle actions prompting evasive decisions and conservative actions leading to rushing decisions (Zhang et al., 2024).

Building on these insights, there is a clear need to develop advanced modelling approaches that accurately capture the complexity and unpredictability of pedestrian behaviours, particularly jaywalking, under diverse and dynamic traffic scenarios. Existing methodologies often lack comprehensive integration of pedestrian behavioural dynamics within adaptive signal control frameworks, presenting a significant research gap. Addressing this gap would enhance the predictive accuracy of safety interventions, optimally balance multimodal interests, and effectively mitigate pedestrian safety risks through sophisticated and responsive traffic control systems.

3. Methodology

This section begins by outlining the methods for imitating jaywalking behaviours in a micro-simulation environment. The jaywalking decision (JD) model is integrated into a multimodal simulation environment to introduce Safety risks due to jaywalking behaviours. Following that the proposed Branching Double Deep Q-Network for Jaywalking and Delay Optimisation (BDDQ-JDO) method is presented. The workflow of BDDQ-JDO is disputed in Figure 1. It utilised the branching double deep Q-Network to optimise real-time cycle-based traffic signal action based on observed state (lane-based information of vehicles and pedestrians), and previous cycle's action and performances (total user delays and jaywalking frequency). Upon selecting the action, the simulation model executes the cycle. Simultaneously, the JD model accesses information from pedestrians (waiting time and vehicle gap time) to decide whether jaywalking is likely to happen. The JD model then intervenes in the simulation to execute the jaywalking movement. Finally, the performance (rewards) of the executed action is recorded and the exploration of the action is learned by BDDQ-JDO subsequently. Table 1 summarised the definition of key variables which will be further discussed in following the sections.

3.1. Jaywalking behaviour modelling

Jaywalking behaviour is a common phenomenon that is often overlooked in current simulation models. For instance, existing models typically describe jaywalking decisions, particularly red light compliance, as an "error" term. However, jaywalking should not always be considered a "mistake". In reality, jaywalking decisions are influenced by numerous factors, as explored in Section 2.3. In general, the factors behind jaywalking can be classified into two main groups, namely individual and external. Individuals focused on demographic factors such as age, gender, and education, as well as psychological level judgements such as patience to wait and tolerance of "safe to cross" conditions, meanwhile, external factors included prolonged waiting times at traffic signals, nearby travellers' behaviour and intersection geometry, etc.

Table 1
Definition of key variables.

Variable	Definition
Jaywalking decision (JD) model	
t_w	Waited time
t_g	Minimum gap time of the crossing traffic
P_j	Probability of Jaywalking
d	A random dice between 0 and 1. Jaywalking happen if $R_d < P_j$
P_i	Probability of immediate crosser, calibrated: P_i=0.1336
P_n	Probability of never crosser, calibrated: P_n=0.4917
T_i	Minimum threshold for an immediate crosser becoming a gap finder
P_w	Patience term of a pedestrian not to wait further
β_0	Coefficient term within P_w, calibrated: β_0=1.9197
β_1	Coefficient term within P_w, calibrated: β_1=0.0117
β_2	Coefficient term within P_w, calibrated: β_2=7
P_g	Gap acceptance of a pedestrian
Environment and State	
S_c	State observed at the c-th cycle, defined in Equation 5
I_{veh}	Traffic information, defined in Equation 6
$O_{i,j}$	Occupancy matrix of lane i and cell j
$V_{i,j}$	Velocity matrix of lane i and cell j
$V'_{i,j}$	Acceleration matrix of lanes i and cell j
I_{ped}	Pedestrian information, defined in Equation 8
n_k	Number of pedestrians waiting at direction k
Signal plan and action	
E_j	Elementary phase j presented in Figure 3, selectable from 1 to 6
g	Active stage, g=NS (north-south); g=EW (east-west)
P_i^g	Two-phase signal programs i at stage g, presented in Table 2, taking values from 1 to 17
D_k^g	Phase duration of phase k at stage g, k=1 first phase; k=2 second phase, [10, 40]s for the vehicle only phase; [20, 40]s if pedestrian exist
$D'_{i,j,k}$	Minimum phase duration depends on phase type (j), signal program (i) and corresponding position (k)
A_c	Signal action selected for the c-th cycle, defined in Equation 9
D_c	Cycle duration
t_f	Minimum red flashing time for pedestrian crossing, 10s
t_a	Vehicle amber time, if transition exist, 3s
t_r	Minimum red time for clearance, if transition exist, 2s
Optimisation Objectives	
R_d	Overall road user delays
θ_{veh}	Thresholds for vehicle delays = 5km/h
θ_{ped}	Thresholds for pedestrian delays = 0.1km/h
R_j	Jaywalking frequency
R_c	Reward observed after c-th cycle, defined in Equation 14

This study presents a preliminary Jaywalking Decision (JD) model to simulate the probability of pedestrian red-light violations, based on both individual and external behavioural factors. Several assumptions are required for its formulation and calibration within the simulation environment, and these are clarified as follow.

Firstly, pedestrians are categorised into three behavioural types according to their crossing decisions: (i) immediate crossers, who proceed without hesitation; (ii) gap finders, who initially wait but choose to cross when they perceive a safe traffic gap; and (iii) never crossers, who wait until the pedestrian signal turns green. This behavioural classification is grounded in the empirical findings of Li (2013), who identified distinct intention types and their proportions under real-world conditions. Accordingly, the model adopts their observed distribution: 14 % immediate crossers, 36 % gap finders, and 50 % never crossers. These proportions are used as calibration targets during simulation to ensure the behavioural realism of the model.

Secondly, the jaywalking decision process is evaluated independently for each pedestrian and for each crossing lane. It is executed in parallel and updated at one-second intervals. This assumption reflects the decentralised and asynchronous nature of real-world pedestrian behaviour and enables the model to respond dynamically to traffic and signal changes. The environment in which these decisions are evaluated is described in Section 3.2, where state variables such as pedestrian location, traffic arrival patterns, and signal status are continuously updated.

Lastly, the probability of jaywalking (P_j) at any moment is modelled as a function of two key inputs: (i) the accumulated waiting time (t_w) since the pedestrian arrived at the intersection and (ii) the perceived gap time (t_g) in the conflicting traffic stream. The rationale behind this formulation is supported by behavioural studies which show that longer waiting times increase impatience and the likelihood of risk-taking, while acceptable gap sizes depend on pedestrians' perception of safety and urgency. Specifically, the model incorporates a time-dependent function to represent decreasing patience ($P_w(t_w)$), and a discrete mapping of gap acceptance probabilities ($P_g(t_g)$), as elaborated in subsequent equations. Together, these inputs determine the likelihood that a gap finder will decide to cross in the absence of a green signal.

These assumptions provide a behaviourally informed foundation for modelling jaywalking decisions within a microsimulation framework and are further supported by a calibration procedure that aligns the simulated outcomes with empirical data. Mathemat-

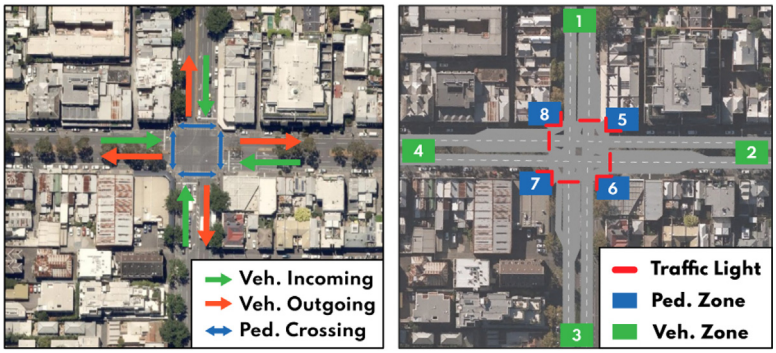

Fig. 2. Illustrated geometric layout of the selected intersection used in the experiments.

ically, the probability of jaywalk (P_j) of a pedestrian during a simulation second is expressed as

$$P_j(t_w, t_g) = \begin{cases} P_i & \text{if } t_w < T_i \\ P_i + (1 - P_i - P_n) \cdot P_w(t_w) \cdot P_g(t_g) & \text{if } t_w \geq T_i \end{cases}, \tag{1}$$

where $T_i = 1$ is the minimum threshold for an immediate crosser becoming a gap finder and P_w and P_g are the probability of the pedestrian deciding not to wait and accept the traffic gap for that moment respectively. Particularly, an exponential growth relationship is used to model the patience to wait with respect to the waited time, i.e.,

$$P_w = \beta_0 \cdot (1 - \exp(-\beta_1 \cdot t_w))^{\beta_2}, \tag{2}$$

where β are the coefficients for calibrations. On the other hand, the gap acceptance of a pedestrian is assumed from psychological studies on pedestrian unsafe behaviour by Tian et al. (2022). For a speed limit of 50 kilometres per hour, the probability of a pedestrian considering the gap between vehicles is "safe to cross" is presented as

$$P_g = \begin{cases} 0 & \text{if } t_g < 2 \\ 6.15\% & \text{if } t_g = 2 \\ 26.12\% & \text{if } t_g = 3 \\ 48.31\% & \text{if } t_g = 4 \\ 75.42\% & \text{if } t_g \geq 5 \end{cases}. \tag{3}$$

Regarding the execution, at each time step the JD model accesses the current waited time and gap time to approximate the probability of jaywalk and a random draw (d), between 0 and 1, is drawn at the same time to control whether jaywalk is executed. For instance

$$\text{Jaywalk} = \begin{cases} True & \text{if } d < P_j \\ False & otherwise \end{cases}. \tag{4}$$

As the probability of jaywalking is positively related to t_w and t_g, the longer the pedestrian waits and the spreader traffic flow increases the upper bound of the execution condition. Thus, allowing jaywalking to happen more easily. For the JD model to be calibrated, the simulation model is utilised to iterate signal cycles at different traffic conditions. The purpose of the calibration is to select a value of P_i, P_n, β_0, β_1 and β_2, such that the JD model will result in a distribution of immediate crosser, gap finder and never crosser matching with the observed distribution after an hour of simulation. The selected value is detailed in Table 1 which is also applied for the subsequent development of BDDQ-JDO.

3.2. Environment and state

The JD model is integrated into a multimodal simulation model to introduce jaywalking which is believed as a risky behaviour that may cause pedestrian safety concerns. The environment consists of a four-lane, four-approach intersection located within the Australian Integrated Multimodal EcoSystems (AIMES) testbed, established by the University of Melbourne in 2016. Vissim, a microsimulation software developed by the Verkehr) (2022), is utilised to simulate the multimodal environment. Figure 2 depicts the geometry of the chosen intersection. Each approach features a dedicated right-turn lane, with left turns permitted only from the leftmost lane. Furthermore, each approach has a bi-directional pedestrian crossing. Given recent advancements in traffic flow detection technologies and the availability of high-resolution data from cameras or connected vehicles, the model assumes the detection of individual vehicle and pedestrian arrivals to gather real-time traffic (I_{veh}) and pedestrian (I_{ped}) information. The state (S_c) representation observed at c-th cycle is denoted as

$$S_c = \left[I_{\text{veh}}, \quad I_{\text{ped}}\right]_c. \tag{5}$$

Combining with the signal action (A_c), which will discussed separately (including definition of (c)) in Section 3.3, forms the input of the BDDQ-JDO. Regarding the traffic information (I_{veh}), the model transformed the real-time traffic data into a lane-based representation.

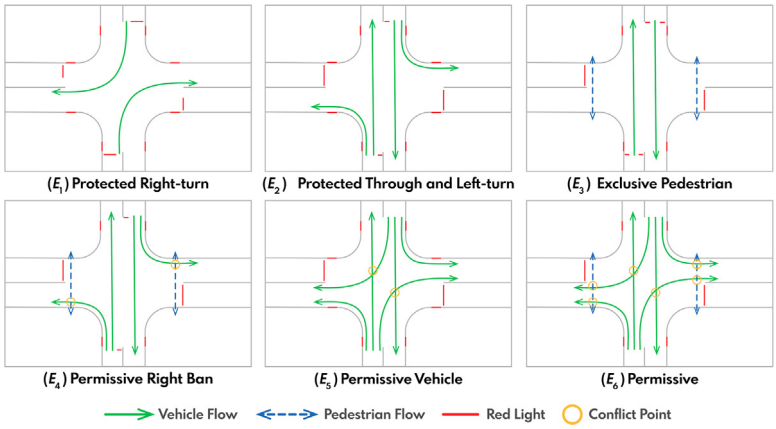

Fig. 3. Elementary phases.

Each incoming lanes (i) are discretised longitudinally into cells of approximately vehicle length and each cell is assigned an index (j). The lane index starts from the right-turn lanes of the southbound approach and increases clockwise. In total, there are 16 lanes for the four-lane, four-approach intersection. Considering information such as occupancy (O), velocity (V), and acceleration (V'), the traffic information (I_{veh}) of cycle c is represented by

$$I_{\text{veh},c} = \left[O, \quad V, \quad V'\right]_c. \tag{6}$$

The structure of occupancy (O), velocity (V), and acceleration (V') matrices are identical but the data types are slightly different. Particularly, the occupancy matrix stored a discrete variable with 1 and 0 standing for occupied and unoccupied cells, respectively, while velocity and acceleration matrices stored continuous variables. The matrix representation of occupancy (O) and velocity (V) is provided as examples, i.e.,

$$O_{i,j} = \begin{bmatrix} o_{1,1} & o_{1,2} & \cdots & o_{1,j} \\ o_{2,1} & o_{2,2} & \cdots & o_{2,j} \\ \vdots & \vdots & \ddots & \vdots \\ o_{i,1} & o_{i,j} & \cdots & o_{i,j} \end{bmatrix} \quad \text{and} \quad V_{i,j} = \begin{bmatrix} v_{1,1} & v_{1,2} & \cdots & v_{1,j} \\ v_{2,1} & v_{2,2} & \cdots & v_{2,j} \\ \vdots & \vdots & \ddots & \vdots \\ v_{i,1} & v_{i,j} & \cdots & v_{i,j} \end{bmatrix}. \tag{7}$$

Similarly, the pedestrian data is converted into a flow-based representation. Considering the four approaches with bi-directional pedestrian crossings, the variable k, with 8 total directions, indexes the different pedestrian flows approaching the crossing and the pedestrian information (I_{ped}) at cycle c is expressed as

$$I_{\text{ped},c} = \left[n_1, \quad n_2, \quad \ldots \quad n_{k-1}, \quad n_k,\right]_c, \tag{8}$$

where n_k represents the number of pedestrians waiting to cross the intersection in the direction k.

For the proposed model with incoming lanes $i = 16$ and lane-cell discretisation $j = 20$, the occupancy matrix and kinematic matrices are

$$O \in \{0,1\}^{16\times 20}, \qquad V, V' \in \mathbb{R}^{16\times 20},$$

and the pedestrian-queue vector is $I_{\text{ped},c} \in \mathbb{R}^8$. The vehicular component therefore comprises three 16×20 matrices, giving $3 \times 16 \times 20 = 960$ elements in total. Including the pedestrian component adds a further 8 elements, so the overall state has 968 elements. All inputs are normalised before being passed to the network.

3.3. Signal plan and action space

This study proposes a deep reinforcement learning (DRL) based traffic signal control system designed to learn from past experiences and dynamically adjust signal operations in real-time. In conventional actuated traffic signals, decisions are typically based on the presence of specific road users, and the extension of the vehicle green phase depends on traffic flow density. Additionally, pedestrian signals in such systems are pre-timed, limiting flexibility. In contrast, the proposed DRL-based system offers a higher degree of freedom in selecting signal operations, allowing for more adaptive traffic management. The signal plan in this study incorporates a conceptual stage-based two-phase traffic signal that alternates between the north-south (NS) and east-west (EW) stages. Each stage (g) is composed of two out of the six elementary phases detailed in Figure 3, which operate sequentially. These elementary phases are designed with varying priority levels. Elementary phases E_1 and E_2 prioritise turning movements, while E_3 provides exclusive green time for pedestrians. Phases E_4, E_5, and E_6 accommodate different pairs of conflicting traffic movements.

Seventeen signal programs (P_i^g) were chosen from a pool of 30 possible permutations (refer to Table 2). These programs were selected to ensure that every traffic flow had at least one designated exit opportunity within a stage. Back to a general level, a signal

Table 2
Definition of preset signal program.

Program P_i^g	Sequence	Program P_i^g	Sequence	Program P_i^g	Sequence
P_1^g	$E_1 \rightarrow E_4$	P_7^g	$E_3 \rightarrow E_5$	P_{13}^g	$E_5 \rightarrow E_4$
P_2^g	$E_1 \rightarrow E_6$	P_8^g	$E_3 \rightarrow E_6$	P_{14}^g	$E_5 \rightarrow E_6$
P_3^g	$E_4 \rightarrow E_1$	P_9^g	$E_5 \rightarrow E_3$	P_{15}^g	$E_6 \rightarrow E_4$
P_4^g	$E_6 \rightarrow E_1$	P_{10}^g	$E_6 \rightarrow E_3$	P_{16}^g	$E_6 \rightarrow E_5$
P_5^g	$E_2 \rightarrow E_6$	P_{11}^g	$E_4 \rightarrow E_5$	P_{17}^g	$E_6 \rightarrow E_6$
P_6^g	$E_6 \rightarrow E_2$	P_{12}^g	$E_4 \rightarrow E_6$		

P_i = Signal program i, E_j = Elementary phase j

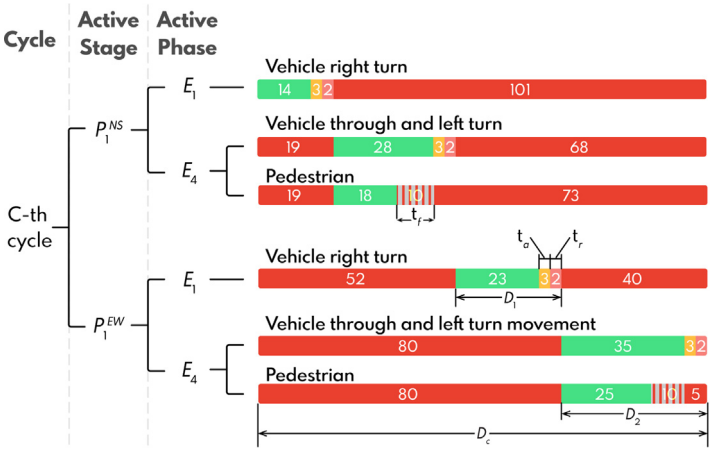

Fig. 4. Example signal plan.

cycle (c) is defined to encompass two stages, in a sequence of an NS stage and EW stage by default. Therefore, as signal action (A) of the c-th cycle is defined as

$$A_c = \left[P_i^{NS}, \quad D_1^{NS}, \quad D_2^{NS}, \quad P_i^{EW}, \quad D_1^{EW}, \quad D_2^{EW}\right]_c. \tag{9}$$

In addition to the selection of signal programs, variations in phase durations (D_k^g), where k corresponds to the relative position within a stage (g), were also considered. Mathematically, the total duration of a signal cycle (D_c) is formulated as

$$D_c = D_1^{NS} + D_2^{NS} + + D_1^{EW} + D_2^{EW}. \tag{10}$$

The phase durations are designed to include transition times, encompassing both vehicle amber (t_a) and clearance red times (t_r), or flashing red for pedestrians (t_f). This implies the implemented elementary phase logic will autonomously adjust to satisfy minimum time requirements (D') based on its position within the stage and the associated road user types. For phases accommodating both vehicles and pedestrians, the flashing pedestrian signal is guaranteed to terminate before the vehicle signal transitions to amber. Therefore, minimum duration requirements are established for each phase within their corresponding programs. For example, assume during an arbitrary cycle, P_1 is assigned to both stages with the following settings:

$$A_c = \left[P_1^{NS}, \quad 19, \quad 33, \quad P_1^{EW}, \quad 28, \quad 40\right]_c. \tag{11}$$

The corresponding signal plan is illustrated in Figure 4. From the example, the cycle duration (D_c) is 120 s, however, when focusing on sub-phases the effective green time is adjusted internally to fulfil the minimum time requirements. For example in the 2nd phase of the EW stage, since E_4 consists of pedestrians and vehicle movement, the effective green time of pedestrian light (25 s) can be calculated by

$$t_g = D_2^{EW} - t_a - t_r - t_f. \tag{12}$$

Detailed information regarding all signal settings is presented in Table 1.

3.4. Optimisation objective (reward mechanism)

In the context of DRL, the model is trained to maximise the cumulative reward. Considering the application in adaptive traffic signals, this study experiments with the impact of optimising one or both of the following objectives: (1) overall road user delays (R_d), including both vehicles and pedestrians and (2) jaywalking frequency (R_j). To the best of the authors' knowledge, BDDQ-JDO is the first attempt to address safety concerns among intersections through an indirect objective, specifically, regulating risky behaviour.

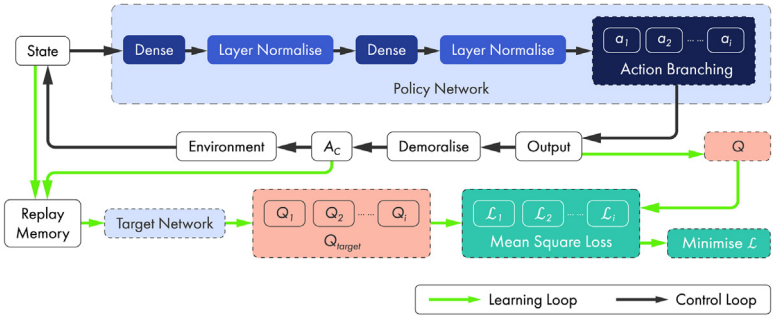

Fig. 5. Model architecture.

As discussed in Section 2.3, jaywalking is associated with prolonged waiting time and dispersed traffic. Therefore, optimising R_j is considered proactive strategy for mitigating excessive pedestrian waiting times and minimising underutilised right-of-way. In addition, it holds potential for addressing safety concerns associated with sudden disruptions to traffic flow caused by jaywalkers.

Since the objective is to minimise both delays and frequency of risky behaviour, the reward observed after the c-th cycle (R_c) is defined as the negative sum of the total delays and jaywalking frequency. For the delays calculation, vehicles and pedestrian is considered delayed when the speed of vehicle (V_{veh}) and pedestrian (V_{ped}) is below certain thresholds, denoted by θ_{veh} and θ_{ped} respectively. Thus, R_d is increased by 1 user-second under the following conditions: $V_{veh} < \theta_{veh}$ and $V_{ped} < \theta_{ped}$, where θ_{veh} and θ_{ped} is set to 5 and 0.1 kilometres per hour respectively. Furthermore, the online min-max normalisation approach is adopted such that the magnitude of R_d and R_j is comparable during the training process. Online normalisation is a process where reward normalisation is performed in real-time when new data arrives. Mathematically, the normalisation process is expressed as

$$\hat{x}_c = \frac{x_c - \min(\min_{c-1}, x_c)}{\max(\max_{c-1}, x_c) - \min(\min_{c-1}, x_c)}, \tag{13}$$

where $\hat{x}_c$ is the normalised value x_c, and $\min_{c-1}$ and $\max_{c-1}$ are the minimum and maximum from previous cycles, respectively. This ensures the reward is scaled in real-time and both R_d and R_j are comparable in the range of $[0, 1]$. Therefore, the total reward R_c after the c-th cycle is expressed as

$$R_c = -\left[\hat{R}_d + \hat{R}_j\right]_c, \tag{14}$$

which has a range of $[-2, 0]$.

3.5. Model architecture and training

The BDDQ-JDO model is fundamentally based on a Branching Double Deep Q-Network (BDDQN), designed to address the challenges of controlling six discrete traffic signal parameters simultaneously in a high-dimensional action space. It integrates two critical techniques in deep reinforcement learning (DRL): (1) Double Deep Q-Network (DDQN), and (2) action branching. This combination is specifically adapted to enhance learning stability and control performance in the context of multimodal urban traffic.

DDQN was introduced by Van Hasselt et al. (2016) and is widely recognised as a refinement of the standard Deep Q-Network (DQN). In conventional DQN, a single network is responsible for both action selection and evaluation, often leading to overestimation of Q-values. DDQN addresses this limitation by using two networks. The policy network selects the best action, while the target network evaluates its value. This decoupling of selection and evaluation processes stabilises the learning dynamics. In our study, both the policy and target networks adopt the same architecture, as shown in Figure 5.

Action branching is particularly important in our application due to the six-dimensional nature of the action space A_c. If each sub-action has n discrete choices (e.g., 17 signal programs or 30 timing intervals), the joint action space without branching would involve at least n^6 possible combinations, making learning computationally intractable. To overcome this, we apply action branching to decompose the joint action into six independent sub-actions. Each sub-action is optimised separately, allowing the agent to explore and learn more efficiently in high-dimensional environments. This design enables practical control of complex signal timing parameters such as green splits and phase selection across different approaches.

The policy network and the target network share the same architecture. Each network consists of two fully connected (dense) layers with 1024 and 512 units, respectively. Both layers use ReLU activation and are followed by layer normalisation to stabilise training. The shared latent representation then branches into six independent output heads. These heads have output dimensions $[17, 30, 30, 17, 30, 30]$, which correspond to the six sub-actions in the control decision, i.e. $A_c = [P_i^{NS}, D_1^{NS}, D_2^{NS}, P_i^{EW}, D_1^{EW}, D_2^{EW}]$. The heads with 17 outputs select the signal program for each stage (from 17 candidate phase sequences), and the heads with 30 outputs select discretised duration settings for each sub-phase within a stage. This branching structure allows each component of the signal plan to be optimised separately while still being conditioned on the full shared state representation. During training, the target network parameters are updated towards the policy network parameters using a soft update with rate $\tau = 0.005$, providing stable target estimates.

Table 3
Training hyperparameters.

Parameter	Value
Learning rate (α)	0.0001
Batch size	128
Discount factor (γ)	0.99
Target update rate (τ)	0.005
Replay buffer size	2,000,000
Episodes	500
Exploration noise (Gaussian σ)	0.2

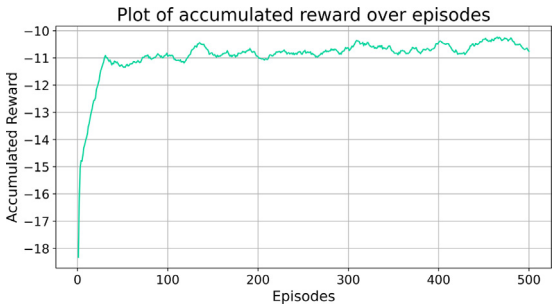

Fig. 6. The accumulated reward for each episode during training (example).

The Q-value for each sub-output in BDDQ-JDO is computed using the DDQN formulation:

$$Q_{target,i} = r_c + \gamma\left(\sum_{i=1}^{6} Q'_i(S_{c+1}, \arg\max_{a_{c,i}} Q_i(S_c, a_{c,i}; \theta); \theta^-)\right), \tag{15}$$

where γ is the discount factor, S_{c+1} is the next state, and $a_{c,i}$ is the sub-action of A_c selected by the current policy network. θ and θ^- are the weights of the policy and target networks, respectively. The training process of the BDDQN aims to minimise the mean squared loss ($\mathcal{L}$) between the predicted Q-values and the target Q-values, which is written as

$$\mathcal{L} = \frac{1}{6}\sum_{i=1}^{6}\left(Q_{target,i} - Q_i(S_c, a_{c,i}; \theta)\right)^2. \tag{16}$$

The training hyperparameters are summarised in Table 3. These settings are aligned with best practices in DRL and were tuned empirically. Exploration is supported by Gaussian noise injection with a standard deviation of 0.2, and a large replay buffer (2 million transitions) is maintained for stable learning.

During training, exploration is achieved by adding zero-mean Gaussian noise with standard deviation $\sigma = 0.2$ to each action head before selecting the discrete $\arg\max$ action. During evaluation, the greedy (noise-free) $\arg\max$ action is used. On average, the model takes approximately 0.00014 s to decide an action, with a 95 % confidence interval of ±0.00026 s, demonstrating its computational efficiency and suitability for real-time decision-making. Convergence is assessed through the stabilisation of cumulative reward and loss curves over training episodes.

Training stability and convergence were monitored throughout learning. Figure 6 shows the cumulative reward per episode for a representative training run of BDDQ-JDO. The reward increases sharply during the initial exploration phase (roughly the first 50 episodes), then remains relatively stable for the remainder of training with only minor fluctuations. After this point, both the cumulative reward and the training loss exhibit low variance across episodes, indicating that the learning process has effectively converged. Based on this observed stabilisation, all reported models were trained for 500 episodes.

To evaluate the effectiveness of various model components, an extensive set of benchmark models was implemented. These models were designed to isolate the effects of action branching, reward objectives, and the inclusion of jaywalking behaviour. The full configuration and purpose of each benchmarked model are detailed in Section 4. This design enables a structured analysis of the contribution of individual components and underscores the advantages of the proposed BDDQ-JDO framework.

4. Result and discussion

BDDQ-JDO has been trained and evaluated across four distinct demand scenarios, ranging from A to D. Each demand scenario represents a typical hour of traffic input capturing the average traffic and pedestrian volume every fifteen minutes with fluctuation over time. Details of each scenario are presented along with the original destination (OD) matrices in the following sections.

Multiple methods have also been developed for benchmarking and comparisons. In this study, three aspects of the setting are considered, including, the optimisation objectives, algorithm or method in use and consideration of jaywalking behaviour. As a

Table 4
Model settings.

Model	Jaywalking	Algorithm	Objective
AS-WOJ	Disabled	Actuated	Delay (Vehicle only)
AS-WJ	Enabled	Actuated	Delay (Vehicle only)
BDDQ-JDO	Enabled	BDDQN	Delay and Jaywalk
BDDQ-DO	Enabled	BDDQN	Delay
BDDQ-JO	Enabled	BDDQN	Jaywalk
DDPG-JDO	Enabled	DDPG	Delay and Jaywalk
BDQ-JDO	Enabled	BDQN	Delay and Jaywalk

result, seven setting groups are chosen, together with four demand scenarios, twenty-eight models are benchmarked. The settings are highlighted in Table 4 and the description of each setting group is as follows:

1. Actuated Signal Without Jaywalking (AS-WOJ): AS-WOJ employs an actuated signal without intervention from the JD model. Examples of actuated signals include the Sydney Coordinated Adaptive Traffic System (SCATS), a well-known system that utilises sensors such as inductive loop detectors and push buttons to detect vehicles and pedestrians (Sims, 1979; Sims and Dobinson, 1980). AS-WOJ determines the extension of green time based on detected traffic gaps. It is the only model in which jaywalking behaviours are absent, which serves as a reference to assess whether the JD model introduces safety concerns into the simulated environment.
2. Actuated Signal With Jaywalking (AS-WJ): This version of the actuated signal includes intervention from the JD model. It is considered the baseline model in this study. Subsequent discussions on the percentage of improvement or deterioration will be based on this model.
3. Branching Double Deep Q-Network for Jaywalking and Delay Optimisation (BDDQ-JDO): This is the proposed framework which incorporated the BDDQN and action branching techniques discussed in Section 3. In this model, both delay and jaywalking are considered as the optimisation objective.
4. Branching Double Deep Q-Network for Delay Optimisation (BDDQ-DO): A variant of BDDQ-JDO which only considers delay in its optimisation objective. It provides insights between multiple objectives and single objective optimisation.
5. Branching Double Deep Q-Network for Jaywalking Optimisation (BDDQ-JO): Another variant of BDDQ-JDO which focused on jaywalking mitigation.
6. Deep Deterministic Policy Gradient for Jaywalking and Delay Optimisation (DDPG-JDO): This model adopted the Deep Deterministic Policy Gradient (DDPG) framework. It is a well-known DRL model designed for continuous control (Lillicrap, 2015; Wang et al., 2024b), which is often found in DRL traffic signal control research (Casas, 2017). DDPG-JDO is also trained to optimise both the delay (R_d) and jaywalking (R_j) rewards.
7. Branching Deep Q-Network for Jaywalking and Delay Optimisation (BDQ-JDO): BDQ is another benchmarked DRL method, which is fundamentally built on DQN (Mnih et al., 2015) with action branching. Similar to BDDQ-JDO, it considers both objectives but with a single Q-Network to select and evaluate policy.

In the following sections, we evaluate the models using five performance metrics (PM), each selected to capture key dimensions of intersection safety and efficiency:

1. Jaywalking count (PM1): This metrics reflects the behavioural safety compliance of pedestrians and the effectiveness of control strategies in deterring risky crossing behaviour.
2. Pedestrian Delays (PM2): Measures the total time pedestrians spend waiting before crossing, accumulated over all pedestrian agents. This reflects the level of service for foot traffic and the responsiveness of the control system to pedestrian demand.
3. Pedestrian safety costs (PM3): Estimates the monetary cost of pedestrian-vehicle near misses by incorporating the likelihood of a crash and potential severity if one were to occur. This cost-based approach enables objective evaluation of safety risks from a pedestrian perspective.
4. Vehicle delays (PM4): Captures the cumulative delay experienced by vehicles, expressed in user-seconds. It reflects traffic efficiency and overall network throughput.
5. Vehicle safety costs (PM5): Represents the economic impact of vehicular near misses, based on conflict profiles, time-to-collision metrics, and estimated impact energy. It provides a complementary safety perspective focused on vehicular traffic.

The delay metrics (PM2 and PM4) are computed by aggregating the waiting time experienced by individual agents across the simulation. The safety costs (PM3 and PM5) are derived using post-simulation trajectory analysis and conflict detection, then monetised using the cost-based frameworks proposed by Chan et al. (2025), respectively. These frameworks incorporate surrogate safety indicators and crash cost modelling to estimate safety risk of each near misses between pedestrian-vehicles and vehicle-vehicle interactions. All results reported represent the cumulative values of each PM over one hour of simulation. To ensure statistical robustness, each scenario was simulated 500 times, and the mean and standard deviation were computed across runs.

Table 5
Vehicle OD matrices for Scenario A.

Period	O\D	1	2	3	4
0:00-15:00	**1**	0	50	100	30
	2	35	0	25	80
	3	100	50	0	20
	4	25	80	35	0
15:00-30:00	**1**	0	55	110	33
	2	39	0	28	88
	3	110	55	0	22
	4	28	88	39	0
30:00-45:00	**1**	0	65	130	39
	2	46	0	33	104
	3	130	65	0	26
	4	33	104	46	0
45:00-60:00	**1**	0	45	90	27
	2	32	0	23	72
	3	90	45	0	18
	4	23	72	32	0

Table 6
Pedestrian OD matrices for Scenario A.

Period	O\D	5	6	7	8
00:00-15:00	**5**	0	40	0	40
	6	40	0	40	0
	7	0	40	0	40
	8	40	0	40	0
15:00-30:00	**5**	0	44	0	44
	6	44	0	44	0
	7	0	44	0	44
	8	44	0	44	0
30:00-45:00	**5**	0	52	0	52
	6	52	0	52	0
	7	0	52	0	52
	8	52	0	52	0
45:00-60:00	**5**	0	36	0	36
	6	36	0	36	0
	7	0	36	0	36
	8	36	0	36	0

Table 7
Statistical table for Scenario A.

Model	PM1 (Jaywalking)		PM2 (Ped. Delay)		PM3 (Ped. Safety)		PM4 (Veh.Delay)		PM5 (Veh. Safety)	
	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std
AS-WOJ	0.00	0.00	73,844.73	1,333.20	128.46	1,512.53	77,333.76	1,524.53	10,051.28	32,384.87
AS-WJ	193.85	8.08	57,263.14	1,276.38	481.83	2,851.07	79,063.10	1,823.68	9,993.04	31,788.82
BDDQ-JDO	77.75	9.85	36,872.38	1,590.67	420.86	2,762.81	86,405.79	4,319.87	29,807.28	347,728.33
BDDQ-DO	85.19	9.29	38,340.63	1,581.57	629.08	3,404.50	83,079.77	4,800.73	12,986.99	41,204.90
BDDQ-JO	72.74	8.62	35,541.10	1,594.87	689.87	3,455.68	81,814.17	4,461.14	31,505.73	349,407.01
DDPG-JDO	52.79	7.36	23,131.14	760.19	920.76	3,805.94	35,513.73	1,929.61	54,276.61	494,803.62
BDQ-JDO	59.24	8.98	32,434.53	1,447.11	695.30	3,403.83	63,185.92	4,550.09	14,176.84	62,881.45

4.1. Scenario A: peak hour demand

The first scenario represents a typical peak-hour demand situation. The north-south approach serves as the major flow, while the east-west approach functions as the minor flow. Tables 5 and 6 outline the vehicle and pedestrian demands, respectively. The demand fluctuates every quarter hour. The accumulated reward throughout the entire set of episodes during the training of BDDQ-JDO is plotted in Figure 6. During testing, five hundred episodes were found to be sufficient for the models presented in this study to converge.

Table 7 reports the statistical performance and Figure 7 visualises the comparison with the details of settings of each model. The introduction of the JD model primarily affected pedestrian-related metrics. Notably, the jaywalking count increased to 194 in AS-WJ, demonstrating the extent of pedestrian non-compliance when given the opportunity. Interestingly, pedestrian delays decreased substantially, from approximately 73,800 to 57,300 user-seconds. This aligns with the intention behind jaywalking, as pedestrians are

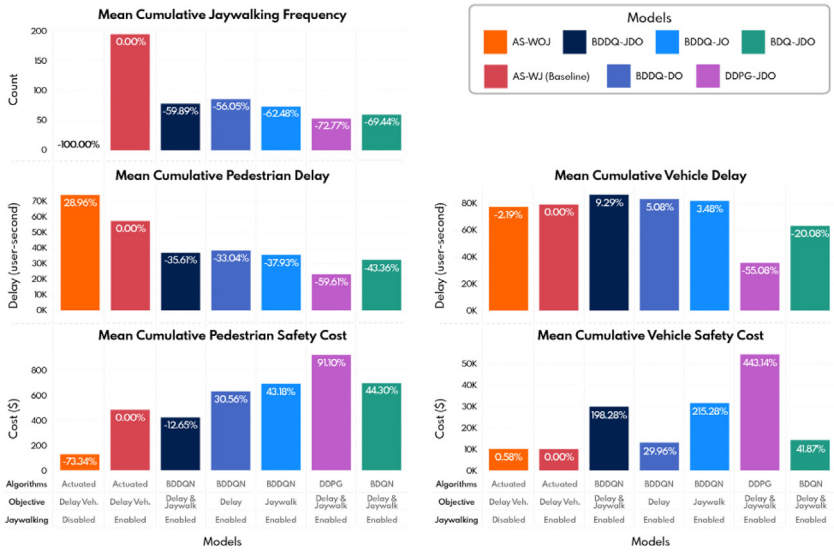

Fig. 7. Comparison of PMs across various models after the testing period for Scenario A, with percentages shown relative to the baseline model.

attempting to cross faster, and the reduction in pedestrian delay reflects their ability to bypass signal timings to achieve shorter waiting times. However, this came at the cost of significantly higher pedestrian safety risks, with the safety cost increasing by approximately 275 %, from 128.46 to 481.83. Vehicle-related impacts exhibited minimal variation, with only a slight increase in vehicle delay and relatively stable vehicle safety costs.

Focusing on the comparison of BDDQ-JDO across different optimisation objectives, including AS-WJ, BDDQ-DO, and BDDQ-JO. BDDQ-JDO outperformed AS-WJ in pedestrian-related metrics, reducing jaywalking, pedestrian safety costs, and delays by 60 %, 13 %, and 36 %, respectively. However, this improvement came at the expense of vehicle-related performance. Pedestrian delays were the lowest with BDDQ-JO, at approximately 35,500 user seconds, indicating its focus on jaywalking mitigation effectively reduced delays. Jaywalking is associated with waiting time and traffic gaps, and BDDQ-JO's training optimises these factors to improve pedestrian flow. Meanwhile, vehicle safety costs were the lowest for BDDQ-DO without significantly compromising other performance metrics. This indicates that optimising solely for delay may lead to a relatively stable environment across all performance metrics in high-demand situations. However, the balance between safety and efficiency still depends on the complexity of traffic conditions.

When evaluating the impact of different algorithm choices, BDDQ-JDO shows notable differences in traffic performance compared to models such as AS-WJ, DDPG-JDO, and BDQ-JDO. DDPG-JDO, which employs a continuous control approach, had the lowest pedestrian delays, reducing it by 60 % compared to the baseline AS-WJ model. Vehicle delays were also reduced by around 55 %, from approximately 79,100 to 35,500 user seconds. However, despite these improvements, DDPG-JDO was not stable in terms of safety performance, as it significantly compromised pedestrian and vehicle safety costs. The pedestrian safety cost increased by approximately 91 %, while vehicle safety costs escalated drastically by 443 %, indicating a substantial trade-off in safety to achieve lower delays. In contrast, BDDQ-JDO offered a more balanced solution in terms of both pedestrian safety and delays but sacrificed vehicle-related performances. For BDQ-JDO, similar patterns were observed as in BDDQ-DO, which did not compromise excessively on any single performance metric, achieving better average performance in high-demand scenarios.

4.2. Scenario B: moderate demand

The second scenario depicts an off-peak hour with moderate demand. Similar to the first scenario, the north-south approach serves as the major flow, while the east-west approach functions as the minor flow. Tables 8 and 9 outline the vehicle and pedestrian demands, respectively.

For this scenario, as depicted in Table 10 and Figure 8, the introduction of jaywalking in AS-WJ resulted in a jaywalking count of 120, consistent with Scenario A. Pedestrian delays decreased by approximately 17 %, from 32,886 to 27,259 user seconds. Pedestrian delays decreased by 17 %, from 32,886 to 27,259 user-seconds, showing that jaywalking reduces wait times but increases safety risks, by 79 %, consistent with Scenario A. These findings are consistent across both peak and off-peak scenarios, indicating that the JD model generally decreases observed pedestrian delays but raises safety concerns.

In objective comparison, BDDQ-JO proved more effective than AS-WJ in balancing delays and safety for both pedestrians and vehicles. Focusing only on reducing delay, BDDQ-DO led to significant increases in safety costs, with pedestrian and vehicle safety costs rising by 91 % and 443 %, respectively, compared to AS-WJ. Overall, BDDQ-JDO and BDDQ-JO performed similarly, though BDDQ-JO tended to provide slightly better outcomes for pedestrians and vehicle safety in this scenario.

Additionally, the algorithm comparison shows that BDQ-JDO was effective at minimising pedestrian delays during off-peak hours, achieving the lowest delay of approximately 17,500 user seconds. However, this improvement came at a cost, with pedestrian safety

Table 8
Vehicle OD matrices for Scenario B.

Period	O\D	1	2	3	4
00:00-15:00	**1**	0	25	50	15
	2	18	0	13	40
	3	50	25	0	10
	4	13	40	18	0
15:00-30:00	**1**	0	28	55	17
	2	20	0	14	44
	3	55	28	0	11
	4	14	44	20	0
30:00-45:00	**1**	0	33	65	20
	2	23	0	17	52
	3	65	33	0	13
	4	17	52	23	0
45:00-60:00	**1**	0	23	45	14
	2	16	0	12	36
	3	45	23	0	9
	4	12	36	16	0

Table 9
Pedestrian OD matrices for Scenario B.

Period	O\D	5	6	7	8
00:00-15:00	**5**	0	20	0	20
	6	20	0	20	0
	7	0	20	0	20
	8	20	0	20	0
15:00-30:00	**5**	0	22	0	22
	6	22	0	22	0
	7	0	22	0	22
	8	22	0	22	0
30:00-45:00	**5**	0	26	0	26
	6	26	0	26	0
	7	0	26	0	26
	8	26	0	26	0
45:00-60:00	**5**	0	18	0	18
	6	18	0	18	0
	7	0	18	0	18
	8	18	0	18	0

Table 10
Statistical table for Scenario B.

Model	PM1 (Jaywalking)		PM2 (Ped. Delay)		PM3 (Ped. Safety)		PM4 (Veh.Delay)		PM5 (Veh. Safety)	
	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std
AS-WOJ	0.00	0.00	32,886.36	792.29	89.37	1,296.57	34,122.93	960.21	3,874.83	8,080.62
AS-WJ	119.58	8.39	27,259.00	812.73	159.59	1,806.15	34,197.59	973.07	4,018.38	7,941.42
BDDQ-JDO	60.43	7.87	18,123.24	915.18	182.49	1,672.35	29,779.80	1,876.39	4,929.32	32,938.13
BDDQ-DO	63.66	7.52	18,884.89	857.01	535.66	3,245.14	28,554.57	1,550.47	20,989.63	347,479.05
BDDQ-JO	57.95	7.56	17,944.71	928.08	81.09	1,015.43	29,907.98	1,791.61	3,212.59	22,393.30
DDPG-JDO	61.95	8.54	18,160.73	919.00	167.79	1,368.06	32,244.97	1,899.61	4,300.31	29,138.07
BDQ-JDO	56.44	7.58	17,487.16	919.41	357.80	2,975.81	32,081.31	1,979.83	19,165.21	348,780.20

costs increasing by approximately 124 % and vehicle safety costs rising by 377 % compared to AS-WJ. These results highlight the significant trade-off between reducing delays and maintaining safety. DDPG-JDO showed similar behaviour to BDDQ-JDO without any significant differences in overall performance.

4.3. Scenario C: high inflow demand

Scenario C simulates a special case of high in-flow demand, often observed during periods like school drop-offs or before major events, where a large number of vehicles and pedestrians are moving towards a specific destination. For example with the geometry illustrated in Figure 2, assuming the destination is located in the north-east area, a high traffic and pedestrian volume is simulated moving towards Zone 5 and Zone 2 for pedestrians and vehicles, respectively. This introduced asymmetric truing movements and crossing activities. The demands for vehicles and pedestrians are presented in Tables 11 and 12 respectively.

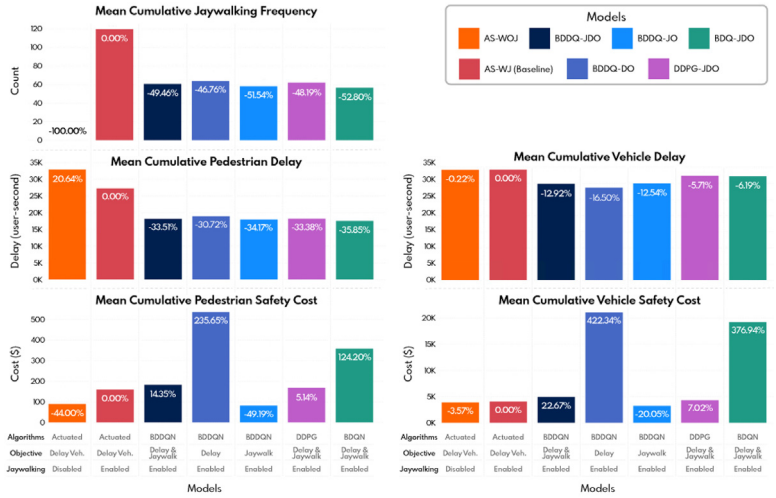

Fig. 8. Comparison of PMs across various models after the testing period for Scenario B, with percentages shown relative to the baseline model.

Table 11
Vehicle OD matrices for Scenario C.

Period	O\D	1	2	3	4
00:00-15:00	**1**	0	80	20	10
	2	30	0	25	35
	3	25	80	0	5
	4	15	80	15	0
15:00-30:00	**1**	0	88	22	11
	2	33	0	28	39
	3	28	88	0	6
	4	17	88	17	0
30:00-45:00	**1**	0	104	26	13
	2	39	0	33	46
	3	33	104	0	7
	4	22	104	22	0
45:00-60:00	**1**	0	72	18	9
	2	27	0	23	32
	3	23	72	0	4
	4	15	72	15	0

Table 12
Pedestrian OD matrices for Scenario C.

Period	O\D	5	6	7	8
00:00-15:00	**5**	0	18	0	16
	6	44	0	12	0
	7	0	36	0	39
	8	46	0	15	0
15:00-30:00	**5**	0	20	0	18
	6	49	0	13	0
	7	0	40	0	43
	8	51	0	17	0
30:00-45:00	**5**	0	23	0	21
	6	57	0	16	0
	7	0	47	0	51
	8	60	0	20	0
45:00-60:00	**5**	0	16	0	14
	6	40	0	11	0
	7	0	32	0	35
	8	42	0	14	0

Table 13
Statistical table for Scenario C.

Model	PM1 (Jaywalking)		PM2 (Ped. Delay)		PM3 (Ped. Safety)		PM4 (Veh.Delay)		PM5 (Veh. Safety)	
	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std
AS-WOJ	0.00	0.00	48,999.66	1,048.40	397.19	2,505.93	50,660.15	1,172.34	4,167.47	19,061.68
AS-WJ	148.59	8.92	38,961.51	983.27	599.38	3,411.25	50,089.16	1,210.10	4,575.08	9,996.20
BDDQ-JDO	64.08	8.30	25,396.80	1,208.86	552.48	3,489.36	53,297.15	1,781.92	2,202.69	9,874.47
BDDQ-DO	66.39	8.24	25,504.06	1,180.44	342.69	2,828.46	54,224.59	1,650.64	1,538.44	7,111.74
BDDQ-JO	60.40	7.75	24,391.81	1,139.44	635.30	3,528.12	52,169.36	2,631.12	4,113.58	27,882.28
DDPG-JDO	69.78	9.49	27,753.30	986.36	68.94	1,073.34	59,005.57	1,740.26	521.41	3,007.64
BDQ-JDO	45.18	6.49	19,686.47	743.98	815.90	3,734.92	38,165.50	1,999.30	20,117.30	346,659.49

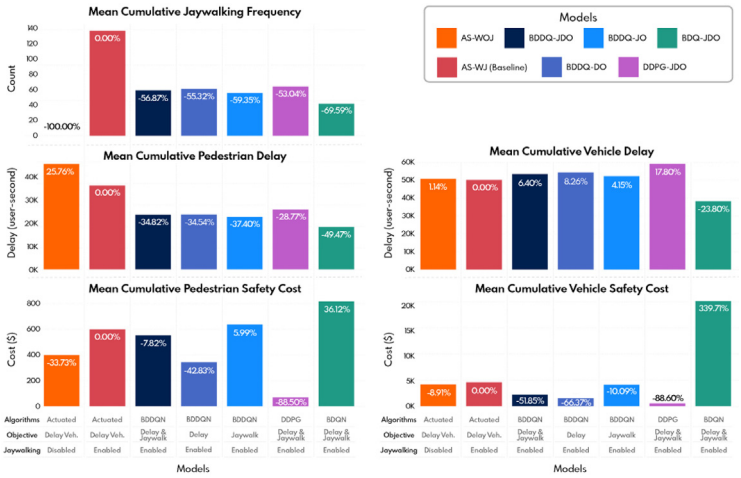

Fig. 9. Comparison of PMs across various models after the testing period for Scenario C, with percentages shown relative to the baseline model.

In this scenario, which involves heavy turning movements, jaywalking had a more pronounced impact compared to previous scenarios. As shown in Table 13 and Figure 9, AS-WJ recorded a mean jaywalking frequency of 150, which the pedestrian reached the destination 20 % faster. However, this increased movement of non-compliance came with increased pedestrian safety costs, rising by 50 % with respect to AW-WOJ, illustrating the complexity of managing jaywalking behaviour at intersections with substantial turning movements and asymmetric flow.

The BDDQ-JDO model showcased its adaptability in handling complex demand patterns by effectively managing the PM1 (−57 %), PM2 (−35 %), PM3 (−8 %) and PM5 (−52 %) better than AS-WJ. However, there is a slight increase in PM4 (+6 %), indicating some trade-offs in managing complex demand. In terms of optimisation objectives, BDDQ-DO achieved the lowest pedestrian safety cost and the lowest vehicle safety cost, reducing by 43 % and 66 %, respectively, compared to AS-WJ. By focusing on delay minimisation, BDDQ-DO provided stable outcomes without significantly compromising safety metrics, making it a strong candidate for overall performance in this scenario. In contrast, BDDQ-JO prioritised reducing jaywalking, which led to improved pedestrian-related metrics but less balanced vehicle performance.

The algorithm comparison further demonstrated significant differences among the models. DDPG-JDO was particularly effective in reducing both pedestrian and vehicle safety costs, achieving the lowest pedestrian safety cost and vehicle safety costs among all models. However, DDPG-JDO has a relatively high delay especially for vehicles, meaning it trade-offs the delays for safety in this scenario. BDQ-JDO, on the other hand, achieved the lowest pedestrian and vehicle delay, reaching a 49 % and 24 % reduction, respectively, compared to AS-WJ. However, this came at a substantial cost, with pedestrian safety costs increasing by 36 % and vehicle safety costs rising by 340 %, indicating it may not be as effective in balancing multiple objectives. This analysis underlines that while some algorithms may excel in specific metrics, they often do so by sacrificing others, making BDDQ-JDO a more consistently balanced choice for complex traffic scenarios.

4.4. Scenario D: High outflow demand

The last scenario is the opposite of Scenario C, representing a high outflow demand case. This could occur when traffic leaves a car park after an event, with a large number of vehicles and pedestrians moving away from a specific destination. In the OD matrices presented in Tables 14 and 15, a high proportion of vehicles and pedestrians are moving away from zone 2 and zone 5, respectively.

The effects of jaywalking in Scenario D were similar to those in earlier scenarios, with an increase in the jaywalking count for AS-WJ to 140, which caused pedestrians to cross the intersection 20 % faster, reducing delays from 45,000 to 36,000 user seconds.

Table 14
Vehicle OD matrices for Scenario D.

Period	O\D	1	2	3	4
00:00-15:00	**1**	0	18	25	15
	2	42	0	38	55
	3	25	15	0	19
	4	17	10	16	0
15:00-30:00	**1**	0	20	28	17
	2	46	0	42	61
	3	28	17	0	21
	4	19	11	18	0
30:00-45:00	**1**	0	23	33	20
	2	55	0	49	72
	3	33	20	0	25
	4	22	13	21	0
45:00-60:00	**1**	0	16	23	14
	2	38	0	34	50
	3	23	14	0	17
	4	15	9	14	0

Table 15
Pedestrian OD matrices for Scenario D.

Period	O\D	5	6	7	8
00:00-15:00	**5**	0	42	0	46
	6	12	0	36	0
	7	0	22	0	16
	8	7	0	35	0
15:00-30:00	**5**	0	46	0	51
	6	13	0	40	0
	7	0	24	0	18
	8	8	0	39	0
30:00-45:00	**5**	0	55	0	60
	6	16	0	47	0
	7	0	29	0	21
	8	9	0	46	0
45:00-60:00	**5**	0	38	0	41
	6	11	0	32	0
	7	0	20	0	14
	8	6	0	32	0

Table 16
Statistical table for Scenario D.

Model	PM1 (Jaywalking)		PM2 (Ped. Delay)		PM3 (Ped. Safety)		PM4 (Veh.Delay)		PM5 (Veh. Safety)	
	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std
AS-WOJ	0.00	0.00	45,204.34	967.27	160.47	1,864.25	32,897.97	973.74	2,716.19	10,615.88
AS-WJ	137.29	8.80	36,363.05	960.05	330.71	4,542.48	32,827.08	984.21	3,481.02	17,819.43
BDDQ-JDO	68.18	8.15	24,973.12	1,188.75	371.94	2,434.99	31,978.31	2,402.04	3,792.84	28,796.03
BDDQ-DO	68.35	8.43	24,741.21	1,120.77	299.08	1,824.77	30,956.22	1,921.89	1,693.81	12,997.11
BDDQ-JO	58.34	7.97	23,167.94	1,110.56	298.95	2,177.81	29,715.97	2,284.73	1,247.04	10,886.19
DDPG-JDO	58.41	7.46	22,266.63	806.68	498.90	2,801.47	45,137.01	1,984.62	1,296.98	5,507.58
BDQ-JDO	79.23	8.90	27,328.36	1,143.22	476.30	2,773.94	33,096.10	2,048.05	3,392.78	21,534.10

Pedestrian safety costs also rose by 106 %, underscoring the consistent trend of integrating the JD model and the associated safety risks of jaywalking behaviour (Table 16 and Figure 10).

When evaluating different optimisation objectives, BDDQ-JO demonstrated superior average performance across all metrics compared to BDDQ-JDO and BDDQ-DO. Specifically, BDDQ-JO achieved the lowest pedestrian and vehicle safety costs, reducing by 10 % and 64 %, respectively, compared to AS-WJ. In contrast, BDDQ-JDO, which aimed to optimise both jaywalking mitigation and delay, showed slightly higher pedestrian and vehicle safety costs when compared to AS-WJ. Perhaps, BDDQ-JO, focusing solely on jaywalking, is a strong candidate for targeting the asymmetric outflow traffic demands.

Lastly, from an algorithmic perspective, DDPG-JDO proved highly effective in reducing pedestrian delay, achieving the lowest value of 22,300 user seconds, which represents a 39 % reduction compared to AS-WJ. Additionally, DDPG-JDO exhibited excellent vehicle safety performance, with safety costs of 1,300, significantly lower than those observed in BDDQ-JDO. Meanwhile, BDQ-JDO displayed mixed results, achieving a pedestrian delay reduction of 25 % compared to AS-WJ, but at the cost of increased safety risks

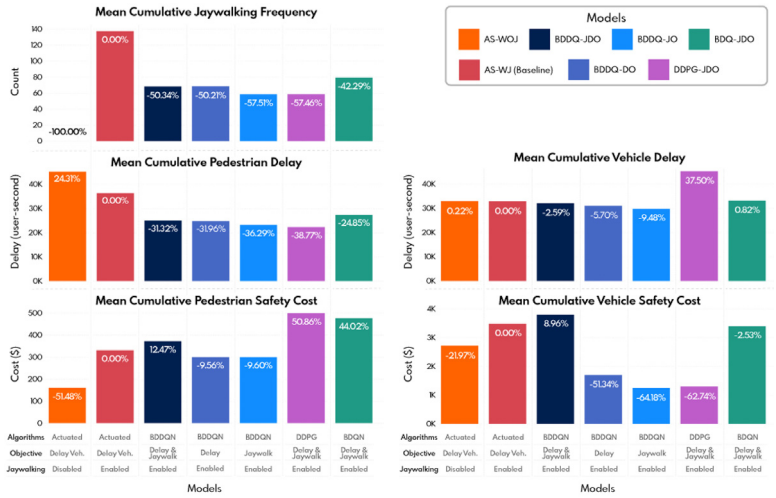

Fig. 10. Comparison of PMs across various models after the testing period for Scenario D, with percentages shown relative to the baseline model.

by 44 %. This suggests that the algorithm struggled to maintain a balance between safety and delay minimisation under high outflow demand conditions.

4.5. Summary

The analysis of performance metrics across the four scenarios provides valuable insights into the implications of jaywalking behaviour modelling and adaptive traffic signal control (ATSC). The key findings and their broader implications are summarised as follows:

1. Incorporating jaywalking behaviour in the simulation consistently increases pedestrian safety risks while reducing pedestrian delays. This reflects real-world pedestrian tendencies where longer waits and light traffic encourage riskier crossings. Signal timing strategies could be adjusted to help reduce jaywalking by minimising pedestrian wait times without compromising safety. For instance, reintroducing pedestrian phases later within a green phase or overlapping non-conflicting phases may help mitigate unsafe crossings. Traditional actuated signals often tie pedestrian phases rigidly to the start of vehicle phases, which can be restrictive.
2. DRL-based ATSC methods generally outperform conventional actuated signals that rely on vehicle gap detection. This supports the viability of deep reinforcement learning for managing complex multimodal traffic environments. Future signal control could further improve by allowing more independent pedestrian phase decisions within safe constraints.
3. Multi-objective approaches, such as BDDQ-JDO, tend to achieve balanced performance across competing goals like delay reduction and safety improvement. In contrast, single-objective approaches like BDDQ-DO or BDDQ-JO excel in specific metrics. Authorities may choose among these based on contextual needs. For example, a setting with high pedestrian volume, such as a major event, may justify prioritising safety using BDDQ-JO. In contrast, on arterial roads with minimal pedestrian activity, the use of BDDQ-DO might be acceptable if the associated safety trade-off is minor.
4. DDPG and BDQ algorithms show less consistent results across different scenarios. These methods may compromise safety to achieve lower delays or favour vehicle flow at the expense of pedestrians. This highlights the importance of aligning algorithm choice with site-specific operational goals, or using adaptive methods that can adjust priorities depending on the context.
5. While BDDQ-JDO offers balanced performance, its outcomes still vary across demand scenarios. This suggests opportunities to explore more dynamic frameworks. Future work could investigate adaptive weighting of objectives or Pareto frontier-based methods to provide flexible control policies that respond to changing conditions and better support operator decision-making.

The comparative performance of BDDQ-JDO, BDDQ-DO, and BDDQ-JO also has practical implications for deployment. In pedestrian-dominant settings with elevated conflict risk, such as school zones, major transit hubs, or post-event dispersal, it may be appropriate to prioritise safety and non-compliance mitigation, consistent with the behaviour of the jaywalking-focused controller (BDDQ-JO), even if this leads to higher vehicle delay. In contrast, on vehicle-priority corridors with low pedestrian activity and strong pressure to maintain throughput, a delay-oriented controller such as BDDQ-DO may be acceptable, provided that the associated increase in safety cost remains within tolerable bounds. In mixed-use urban environments with substantial volumes of both vehicles and pedestrians, the balanced multi-objective controller (BDDQ-JDO) offers a practical compromise by avoiding extreme degradation in either safety or delay. These distinctions indicate that controller selection can be policy-driven and site-specific rather than one-size-fits-all. Taken together, these insights offer practical direction for traffic signal optimisation. They demonstrate the importance of tailoring optimisation strategies to local priorities and suggest pathways to improve both safety and efficiency in urban intersections.

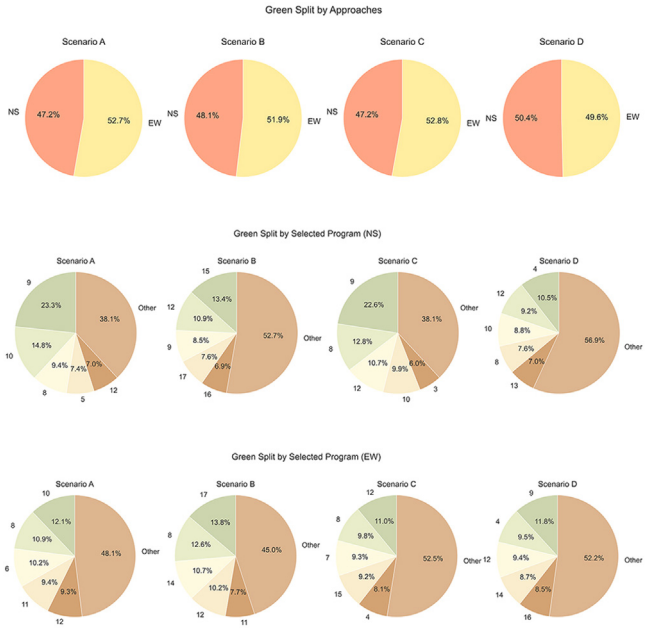

Fig. 11. BDDQ-JDO green splits by approaches and top 5 selected program.

Furthermore, Figure 11 illustrates the green split of BDDQ-JDO across the four scenarios after 500 h evaluation period, reported in terms of allocation by approach and by the top five selected signal programs.

The scenario-specific policy analysis shows that the proposed model does not simply allocate green time according to a fixed ”major road vs minor road” rule. Instead, it reshapes stage order and sub-phase durations in response to the build-up of pedestrians and turning movements. Although the north–south (NS) direction is nominally the major road in Scenarios A and B, the controller frequently assigns slightly more total green time (approximately 52 % in Scenarios A, B, and C) to the east–west (EW) approach. This departs from conventional actuated control, which would typically continue favouring the dominant vehicle approach. The learned policy can be interpreted as preventive: by periodically serving the nominally ”minor” approach, including its pedestrian movements, the controller lowers pedestrian waiting time and suppresses the pressure to jaywalk. In effect, it accepts a small efficiency loss on the NS approach to avoid the much higher safety cost that emerges when pedestrians begin accepting risky gaps.

This mechanism is particularly visible in Scenario A (peak demand), where the NS direction most frequently selects Program 9 (23.3 %). Program 9 consists of a permissive vehicle-only phase followed by an exclusive pedestrian phase. Operationally, this has two effects. First, the permissive phase clears high-volume turning and through movements without forcing a full vehicle stop. Second, before pedestrians wait long enough to begin crossing illegally, the controller inserts a protected pedestrian phase. This protected window ”releases” the accumulated pedestrian queue under crowded conditions. The result is that jaywalking is reduced and pedestrian delay drops relative to actuated control, but vehicle green is sometimes cut earlier than a purely delay-minimising controller (e.g., BDDQ-DO) would allow. This explains why BDDQ-JDO, although more balanced overall, can produce higher vehicle delay than BDDQ-DO in high-demand periods. It deliberately interrupts vehicle flow earlier to maintain pedestrian compliance and reduce safety exposure.

In Scenario B (moderate demand), the policy is more flexible. On the NS approach, the controller distributes its choices across Programs 12 and 15 rather than relying on a single dominant program. Both of these programs include permissive phases with targeted turn restrictions (for example, limiting filter right turns via E4) and permissive phases (E6) in different sequence. This indicates that, under less saturated conditions, the controller manages conflicts by selectively constraining high-risk turns rather than always inserting a full exclusive pedestrian phase. On the EW approach in Scenario B, Program 17 (a fully permissive sequence) is frequently selected. This is consistent with relatively light EW traffic and fewer pedestrians: in that context, the model can maintain throughput with minimal interruption because the safety risk is lower and pedestrians are less likely to accumulate long waits. In other words, when pedestrian pressure is low, the controller behaves more like BDDQ-DO (delay-focused); when pressure is high, it behaves more like BDDQ-JO (compliance-focused).

Scenarios C and D highlight how the learned policy adapts to directional asymmetry. Scenario C represents a high inflow condition (e.g., pre-event arrival), where demand is concentrated into one quadrant, generating heavy turning flows and frequent pedestrian crossings in a specific direction. In this case, the most frequently selected program for the NS direction is again Program 9. Here, Program 9 acts as a conflict management tool. It allows high inbound vehicle demand to clear permissively, then quickly follows with an exclusive pedestrian phase to clear the pedestrian queues that form in the same area. This sequencing reduces the incentive for pedestrians to force a crossing during the permissive vehicle movement. At the same time, it prevents long multi-cycle deferrals of

pedestrians that would otherwise drive up jaywalking counts and pedestrian safety cost. The trade-off is that vehicles approaching from lower-priority directions may see increased delay compared with a pure throughput-oriented strategy, which matches the observed increase in PM4 in Scenario C for BDDQ-JDO relative to AS-WJ.

Scenario D, by contrast, represents a high outflow condition (e.g., post-event dispersal), where pedestrians and vehicles are predominantly moving away from a common origin. Here, the controller often applies Program 9 to the EW direction rather than the NS direction. This ”mirroring” of strategy reflects that the dominant outbound demand has shifted laterally. The controller learns to prioritise whichever approach is carrying that outbound flow and then inserts a protected pedestrian phase to discharge the corresponding pedestrian surge. This behaviour aligns with the strong performance of BDDQ-JO in Scenario D, where safety and compliance-oriented control is particularly valuable during dispersal. In this setting, aggressively serving pedestrians early reduces jaywalking pressure even under very uneven directional loading. At the same time, delay-oriented control (BDDQ-DO) performs well for vehicle safety cost, because it tends to maintain long greens for the major outbound movement and minimise complex turning conflicts.

Across all four scenarios, a consistent pattern emerges:

1. BDDQ-JDO behaves like a compromise policy. It alternates permissive vehicle flow with short, well-timed exclusive pedestrian phases to keep pedestrian queues from becoming ”desperate,” which in turn suppresses jaywalking and associated safety cost.
2. BDDQ-DO tends to hold long green phases for dominant vehicle movements. This reduces vehicle delay but can prolong pedestrian waits, which increases pedestrian safety exposure and can allow more risky crossings.
3. BDDQ-JO acts as a compliance-preserving controller. It serves pedestrians proactively and frequently, reducing jaywalking and pedestrian safety cost, especially in conditions with high pedestrian pressure or directional surges.

These behavioural differences explain the quantitative trade-offs reported earlier. For example, when BDDQ-JDO shows slightly higher vehicle delay than BDDQ-DO, it is because it is ”spending” part of the vehicle green time to provide pedestrians with a protected phase before they resort to unsafe crossings. Likewise, when BDDQ-DO exhibits lower delay but higher safety cost, it is because pedestrians are being held longer and are more likely to cross during permissive vehicle movement rather than wait for a protected interval. Finally, algorithms such as BDQ-JDO and DDPG-JDO tend to exploit permissive phases aggressively to drive down delay, but in doing so they often allow high-conflict states to persist. This explains why these algorithms can achieve impressive delay reductions while at the same time producing large safety costs in Scenarios C and D.

Overall, these patterns illustrate that the learned policies are not only optimising numerical rewards but are also learning interpretable timing strategies, for example, when to preserve a dominant vehicle green, when to insert an exclusive pedestrian phase to release pressure, and when to constrain specific turning movements to reduce conflicts. This supports the deployability of such controllers, since operators can relate observed timing plans (e.g., Program 9 followed by short pedestrian clearance) to explicit policy goals such as ”protect pedestrians during event egress” or ”hold throughput on the major approach during off-peak”. Future work could extend this analysis to multi-intersection coordination, asymmetric geometries, and real-world data to support scalable, policy-driven ATSC deployment.

5. Conclusion

This paper presented a deep reinforcement learning approach, the Branching Double Deep Q-Network for Jaywalking and Delay Optimisation (BDDQ-JDO), for adaptive traffic signal control in a multimodal setting with explicit treatment of pedestrian non-compliance. The framework was designed with two objectives. First, to promote fairness by accounting for individual pedestrians in delay estimation. Second, to reduce risky behaviour that arises from long waits and inefficient right-of-way allocation, for example, when green time is served to approaches with low or uneven traffic, which creates opportunities for jaywalking.

To achieve these objectives, a hybrid action space is constructed, incorporating different phase configurations and durations. Unlike traditional fixed-time or actuated traffic signals, the proposed action space allows the selection of phases such as permissive or protected pedestrian crossings, as well as phases that prioritise turning vehicles from crossing conflicts. Furthermore, it permits the selection of various detailed duration settings, down to the sub-phase level. The proposed action is based on a cycle-level, two-stage, two-phase traffic signal, meaning that each cycle consists of two stages (North-South and East-West), with two phases inside each stage. Thus, each action includes six variables corresponding to the signal program for each stage and the durations of phases within each stage. Overall, the action space contains over two hundred thousand discrete choices.

Subsequently, the BDDQ-JDO model employs a branching architecture to efficiently manage this high-dimensional action space, allowing for effective exploration of different actions and enhancing the model’s learning capabilities. Extensive experiments were conducted using four different simulated traffic scenarios based on a real intersection in Melbourne. The model’s performance was compared against other models that had different optimisation objectives and algorithms. The evaluation used five performance metrics: jaywalking frequency, pedestrian safety and delays, and vehicle safety and delays. The results showed that the DRL-based methods outperformed the baseline models, such as actuated signals. Moreover, the BDDQ-JDO model demonstrated stability and generalised well across all demand scenarios. Perhaps the results suggest that a model capable of dynamically switching optimisation objectives could be more robust in addressing various traffic conditions.

5.1. Limitations and future work

This study has several limitations. First, all experiments were conducted in a controlled microsimulation environment for a single four-leg intersection in the AIMES testbed. Although four demand scenarios were designed to reflect peak, off-peak, concentrated inflow, and concentrated outflow conditions, these scenarios were generated from synthetic OD matrices rather than fully observed time-synchronised field data. As a result, the reported performance should be interpreted as indicative rather than directly generalisable to broader network settings. Future work will extend the evaluation to coordinated multi-intersection networks, different geometries, and empirically observed demand patterns to assess scalability.

Second, while the controller adapts to different traffic and pedestrian conditions, the present study does not include a deployment-level policy analysis. The results suggest that different optimisation modes correspond to different operational priorities. A safety-oriented configuration such as BDDQ-JO may be most appropriate in environments with high pedestrian exposure (e.g., event egress, school zones), where reducing risky crossing behaviour is critical. A delay-oriented configuration such as BDDQ-DO may be suitable on major vehicle corridors with low pedestrian activity, provided safety impacts remain acceptable. The balanced formulation, BDDQ-JDO, is intended for mixed-use urban intersections where neither safety nor efficiency can be neglected. These use cases will be further examined in future work with site-specific data.

There is also scope to enhance behavioural realism. The current Jaywalking Decision (JD) model represents red-light violations at designated crossings, driven by pedestrian waiting time and gap acceptance. In practice, non-compliant behaviour can include mid-block crossings, off-crossing manoeuvres, or reactions to exogenous triggers (e.g., arrival of public transport). Incorporating these behaviours would strengthen the behavioural credibility of the simulation.

Related to this, pedestrians are represented using three behavioural types with fixed proportions. This provides a tractable basis for analysis but does not capture context-dependent variation. Future work will investigate adaptive behavioural classes and perform sensitivity analysis to quantify how alternative behavioural assumptions influence safety and delay outcomes.

From a control perspective, the current action space is limited to a two-stage cycle structure with two sequential phases per stage. Extending the action space to include asymmetric, approach-specific phase configurations would allow finer control in highly unbalanced or time-varying demand conditions.

Finally, although the proposed controller generates actions in real time (on the order of 10^{-4} seconds per decision), a full computational performance assessment and hyperparameter sensitivity study were not conducted. Systematic analysis of computational load, learning stability under different training settings, and robustness to parameter variation will form an important part of future work, particularly for real-time deployment.

Declaration of competing interest

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

CRediT authorship contribution statement

Lok Sang Chan: Writing – original draft, Software, Methodology, Data curation. **Xiaocai Zhang:** Writing – review & editing, Methodology. **Neema Nassir:** Writing – review & editing, Project administration, Methodology, Conceptualization. **Majid Sarvi:** Writing – review & editing, Project administration, Conceptualization.

Acknowledgements

This research is funded by ARC LP200301389, Kapsch TrafficCom Australia, RACQ, and iMOVE CRC, the Cooperative Research Centres program, an Australian Government initiative.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at 10.1016/j.multra.2025.100271

References

Abrams, C.M., Smith, SA., 1977. Selection of pedestrian signal phasing. Transportation Research Record 629, 1–6.

Anik, M.A.H., Hossain, M., Habib, M.A., 2021. Investigation of pedestrian jaywalking behaviour at mid-block locations using artificial neural networks. Safety Science 144, 105448.

Arun, A., Lyon, C., Sayed, T., Washington, S., Loewenherz, F., Akers, D., Ananthanarayanan, G., Shu, Y., Bandy, M., Haque, M.M., 2023. Leading pedestrian intervals – Yay or Nay? A before-after evaluation of multiple conflict types using an enhanced non-stationary framework integrating quantile regression into Bayesian hierarchical extreme value analysis. Accident Analysis & Prevention 181, 106929. doi:10.1016/j.aap.2022.106929.

Banister, D., 2008. The sustainable mobility paradigm. Transport Policy 15 (2), 73–80.

Bechtel, A.K., MacLeod, K.E., Ragland, D.R., 2004. Pedestrian scramble signal in Chinatown Neighborhood of Oakland, California: An evaluation. Transportation Research Record 1878 (1), 19–26. doi:10.3141/1878-03.

Boarnet, M.G., Giuliano, G., Hou, Y., Shin, E.J., 2017. First/last mile transit access as an equity planning issue. Transportation Research Part A: Policy and Practice 103, 296–310.

Brosseau, M., Zangenehpour, S., Saunier, N., Miranda-Moreno, L., 2013. The impact of waiting time and other factors on dangerous pedestrian crossings and violations at signalized intersections: A case study in Montreal. Transportation Research Part F: Traffic Psychology and Behaviour 21, 159–172.
Cai, C., Wei, M., 2024. Adaptive urban traffic signal control based on enhanced deep reinforcement learning. Scientific Reports 14 (1), 14116.
Casas, N., 2017. Deep deterministic policy gradient for urban traffic light control. arXiv preprint arXiv:1703.09035.
Chan, L.S., Nassir, N., Zhang, X., Yazdani, M., Sarvi, M., 2025. Preemptive crash risk reduction through a real-time cost-based safety prediction model (RECOSAM) for traffic signal control. Computers and Electrical Engineering 128, 110639.
Chen, L., Chen, C., Ewing, R., 2015. Left-turn phase: Permissive, protected, or both? A quasi-experimental design in New York City. Accident Analysis & Prevention 76, 102–109. doi:10.1016/j.aap.2014.12.019.
Choi, J., Kim, S., Kim, S., Hong, H., Baik, S., 2013. Pedestrian crashes during jaywalking: Can we afford to overlook? In: 16th International Conference Road Safety on Four Continents. Beijing, China (RS4C 2013). 15-17 May 2013. Statens väg-och transportforskningsinstitut.
Cui, J., Allan, A., Lin, D., 2013. The development of grade separation pedestrian system: A review. Tunnelling and Underground Space Technology 38, 151–160. doi:10.1016/j.tust.2013.06.004.
Dıaz, E.M., 2002. Theory of planned behavior and pedestrians' intentions to violate traffic regulations. Transportation Research Part F: Traffic Psychology and Behaviour 5 (3), 169–175.
Dommes, A., Granié, M.-A., Cloutier, M.-S., Coquelet, C., Huguenin-Richard, F., 2015. Red light violations by adult pedestrians and other safety-related behaviors at signalized crosswalks. Accident Analysis & Prevention 80, 67–75. doi:10.1016/j.aap.2015.04.002.
Dong, X., Guerra, E., Daziano, R.A., 2024. Will automated vehicles encourage more jaywalking? Results from a stated preference survey. Transportation Research Part F: Traffic Psychology and Behaviour 103, 217–229.
Du, X., Ma, J., Zhang, M., Wang, J., Liang, C., 2024. The attentional guidance and facilitating effects of group behavioral cues on individual college pedestrians' jaywalking decisions. Traffic Injury Prevention 25 (5), 733–740.
Du, W., Ye, J., Gu, J., Li, J., Wei, H., Wang, G., 2022. Safelight: A reinforcement learning method toward collision-free traffic signal control. In: Proceedings of the AAAI Conference on Artificial Intelligence, Vol. 37, pp. 14801–14810.
Gårder, P., 1989. Pedestrian safety at traffic signals: A study carried out with the help of a traffic conflicts technique. Accident Analysis & Prevention 21 (5), 435–444. doi:10.1016/0001-4575(89)90004-3.
Gong, Y., Abdel-Aty, M., Yuan, J., Cai, Q., 2020. Multi-objective reinforcement learning approach for improving safety at intersections with adaptive traffic signal control. Accident Analysis & Prevention 144, 105655. doi:10.1016/j.aap.2020.105655.
Guo, H., Gao, Z., Yang, X., Jiang, X., 2011. Modeling pedestrian violation behavior at signalized crosswalks in China: A hazards-based duration approach. Traffic Injury Prevention 12 (1), 96–103.
Guo, Y., Wang, X., Meng, X., Wang, J., Liu, Y., 2019. Analysis of red-light violation behavior of pedestrian two-stage crossing at a signalized intersection. Civil Engineering Journal 5 (2), 429–436.
Hamed, M.M., 2001. Analysis of pedestrians' behavior at pedestrian crossings. Safety Science 38 (1), 63–82.
Hardt, C., Bogenberger, K., 2019. Usage of e-scooters in urban environments. Transportation Research Procedia 37, 155–162.
Infrastructure, B.o., (BITRE), T.R.E., 2023. Road Trauma Australia 2022 statistical summary. Technical Report. BITRE, Canberra ACT.
Ivan, J.N., McKernan, K., Zhang, Y., Ravishanker, N., Mamun, S.A., 2017. A study of pedestrian compliance with traffic signals for exclusive and concurrent phasing. Accident Analysis & Prevention 98, 157–166. doi:10.1016/j.aap.2016.10.003.
Jagannathan, R., Bared, J.G., 2005. Design and performance analysis of pedestrian crossing facilities for continuous flow intersections. Transportation Research Record 1939 (1), 133–144.
Jay, M., Régnier, A., Dasnon, A., Brunet, K., Pelé, M., 2020. The light is red: Uncertainty behaviours displayed by pedestrians during illegal road crossing. Accident Analysis & Prevention 135, 105369.
Kattan, L., Acharjee, S., Tay, R., 2009. Pedestrian scramble operations: Pilot study in calgary, alberta, canada. Transportation Research Record 2140 (1), 79–84. doi:10.3141/2140-08.
Keegan, O., O'Mahony, M., 2003. Modifying pedestrian behaviour. Transportation Research Part A: Policy and Practice 37 (10), 889–901.
Khamis, M.A., Gomaa, W., El-Shishiny, H., 2012. Multi-objective traffic light control system based on Bayesian probability interpretation. 2012 15th International IEEE conference on intelligent transportation systems (ITSC) 995–1000. doi:10.1109/ITSC.2012.6338853.
Kunaratnam, V., Schwartz, N., Howard, A., Mitra, R., Saunders, N., Cloutier, M.-S., Macpherson, A., Fuselli, P., Rothman, L., 2022. Equity, walkability, and active school transportation in Toronto, Canada: A cross-sectional study. Transportation Research Part D: Transport and Environment 108. doi:10.1016/j.trd.2022.103336.
Leden, L., Gårder, P., Johansson, C., 2006. Safe pedestrian crossings for children and elderly. Accident Analysis & Prevention 38 (2), 289–294. doi:10.1016/j.aap.2005.09.012.
Li, B., 2013. A model of pedestrians' intended waiting times for street crossings at signalized intersections. Transportation Research Part B: Methodological 51, 17–28.
Li, X., Sun, J.-Q., 2019. Intersection multi-objective optimization on signal setting and lane assignment. Physica A: Statistical Mechanics and its Applications 525, 1233–1246. doi:10.1016/j.physa.2019.04.223.
Li, Z., Yu, H., Zhang, G., Dong, S., Xu, C.-Z., 2021. Network-wide traffic signal control optimization using a multi-agent deep reinforcement learning. Transportation Research Part C: Emerging Technologies 125, 103059.
Liao, B., van den Berg, P.E.W., van Wesemael, P.J.V., Arentze, T.A., 2020. Empirical analysis of walkability using data from the Netherlands. Transportation Research: Part D: Transport and Environment 85. doi:10.1016/j.trd.2020.102390.
Lillicrap, TP., 2015. Continuous control with deep reinforcement learning. arXiv preprint arXiv:1509.02971.
Liu, Y., Liu, L., Chen, W.-P., 2017. Intelligent traffic light control using distributed multi-agent Q learning. In: 2017 IEEE 20th International conference on intelligent transportation systems (ITSC). IEEE, pp. 1–8.
Lu, Y., Prato, C.G., Sipe, N., Kimpton, A., Corcoran, J., 2022. The role of household modality style in first and last mile travel mode choice. Transportation Research Part A: Policy and Practice 158, 95–109.
Luo, H., Bie, Y., Jin, S., 2024. Reinforcement learning for traffic signal control in hybrid action space. IEEE Transactions on Intelligent Transportation Systems 25 (6), 5225–5241.
Ma, W., Liu, Y., Head, K.L., 2014. Optimization of pedestrian phase patterns at signalized intersections: A multi-objective approach. Journal of Advanced Transportation 48 (8), 1138–1152. doi:10.1002/atr.1256.
Ma, W., Liu, Y., Xie, H., Yang, X., 2011. Multiobjective optimization of signal timings for two-stage, midblock pedestrian crosswalk. Transportation Research Record 2264 (1), 34–43.
Mamun, S., Caraballo, F.J., Ivan, J.N., Ravishanker, N., Townsend, R.M., Zhang, Y., 2020. Identifying association between pedestrian safety interventions and street-crossing behavior considering demographics and traffic context. Journal of Transportation Safety and Security 12 (3), 441–462. doi:10.1080/19439962.2018.1490369.
Marisamynathan, S., Vedagiri, P., 2018. Modeling pedestrian crossing behavior and safety at signalized intersections. Transportation Research Record 2672 (31), 76–86. doi:10.1177/0361198118759075.
Meng, M., Koh, P.P., Wong, Y.D., 2016. Influence of socio-demography and operating streetscape on last-mile mode choice. Journal of Public Transportation 19 (2), 38–54.
Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A.A., Veness, J., Bellemare, M.G., Graves, A., Riedmiller, M., Fidjeland, A.K., Ostrovski, G., 2015. Human-level control through deep reinforcement learning. Nature 518 (7540), 529–533.
Murakami, J., Villani, C., Talamini, G., 2021. The capital value of pedestrianization in Asia's commercial cityscape: Evidence from office towers and retail streets. Transport Policy 107, 72–86. doi:10.1016/j.tranpol.2021.04.017.
O'Hern, S., Stephens, A.N., Estgfaeller, N., Moore, V., Koppel, S., 2020. Self-reported pedestrian behaviour in Australia. Transportation Research Part F: Traffic Psychology and Behaviour 75, 134–144. doi:10.1016/j.trf.2020.10.002.

Organization, W.H., 2023. Global status report on road safety 2023. World Health Organization, Geneva.
Raoniar, R., Maurya, A.K., 2022. Pedestrian red-light violation at signalised intersection crosswalks: Influence of social and non-social factors. Safety Science 147, 105583.
Rosenbloom, T., 2009. Crossing at a red light: Behaviour of individuals and groups. Transportation Research Part F: Traffic Psychology and Behaviour 12 (5), 389–394.
Roy, K., Chan, L.S., Zhang, X., Nassir, N., 2025. Multi-task deep learning for joint prediction of traffic emissions and travel delay. Transportation Research Part D: Transport and Environment 146, 104846.
Russo, B.J., James, E., Aguilar, C.Y., Smaglik, E.J., 2018. Pedestrian behavior at signalized intersection crosswalks: Observational study of factors associated with distracted walking, pedestrian violations, and walking speed. Transportation Research Record 2672 (35), 1–12. doi:10.1177/0361198118759949.
Samerei, S.A., Aghabayk, K., Shiwakoti, N., Karimi, S., 2021. Modelling bus-pedestrian crash severity in the state of Victoria, Australia. International Journal of Injury Control and Safety Promotion 28 (2), 233–242.
Shiwakoti, N., Tay, R., Stasinopoulos, P., 2020. Development, testing, and evaluation of road safety poster to reduce jaywalking behavior at intersections. Cognition, Technology & Work 22, 389–397.
Sims, A.G., 1979. The Sydney coordinated adaptive traffic system. In: Engineering foundation conference on research directions in computer control of urban traffic systems, 1979, Pacific Grove, California, USA.
Sims, A.G., Dobinson, K.W., 1980. The Sydney coordinated adaptive traffic (SCAT) system philosophy and benefits. IEEE Transactions on Vehicular Technology 29 (2), 130–137.
Stevanovic, A., Stevanovic, J., So, J., Ostojic, M., 2015. Multi-criteria optimization of traffic signals: Mobility, safety, and environment. Engineering and Applied Sciences Optimization (OPT-i) - Professor Matthew G. Karlaftis Memorial Issue 55, 46–68. doi:10.1016/j.trc.2015.03.013.
Tang, L., Liu, Y., Li, J., Qi, R., Zheng, S., Chen, B., Yang, H., 2020. Pedestrian crossing design and analysis for symmetric intersections: Efficiency and safety. Transportation Research Part A: Policy and Practice 142, 187–206. doi:10.1016/j.tra.2020.10.012.
Tian, K., Markkula, G., Wei, C., Lee, Y.M., Madigan, R., Merat, N., Romano, R., 2022. Explaining unsafe pedestrian road crossing behaviours using a psychophysics-based gap acceptance model. Safety Science 154, 105837.
Van Hasselt, H., Guez, A., Silver, D., 2016. Deep reinforcement learning with double q-learning. In: Proceedings of the AAAI conference on artificial intelligence, Vol. 30.
Verkehr), P.G.P.P.T., 2022. PTV Vissim 2022 User Manual. 2022 PTV AG, Karlsruhe, Germany 1352.
Victoria, T. G. o., 2019. Road safety (general) regulations 2019 (Vic).
Victoria, T. G. o., 2020. Road Safety action plan 2021-2023 (Vic).
Wang, F., Tang, K., Li, K., Liu, Z., Zhu, L., 2019. A Group-based signal timing optimization model considering safety for signalized intersections with mixed traffic flows. Journal of Advanced Transportation 2019, e2747569. doi:10.1155/2019/2747569.
Wang, J., Huang, H., Xu, P., Xie, S., Wong, SC., 2020. Random parameter probit models to analyze pedestrian red-light violations and injury severity in pedestrian–motor vehicle crashes at signalized crossings. Journal of Transportation Safety & Security 12 (6), 818–837.
Wang, J., Zhang, Y., Zhao, J., Shang, C., Wang, X., 2024. Unified strategy for cooperative optimization of pedestrian control patterns and signal timing plans at intersections. Journal of Intelligent Transportation Systems 29 (2), 170–196.
Wang, S., Zhang, X., Li, J., Wei, X., Lau, H.C., Dai, B.T., Huang, B., Xiao, Z., Fu, X., Qin, Z., 2024. Fuel-saving route planning with data-driven and learning-based approaches-a systematic solution for harbor tugs. In: Proceedings of the thirty-third international joint conference on artificial intelligence, pp. 7483–7490.
Wang, T., Cao, J., Hussain, A., 2021. Adaptive traffic signal control for large-scale scenario with cooperative group-based multi-agent reinforcement learning. Transportation Research Part C: Emerging Technologies 125, 103046.
Wang, X., Jerome, Z., Wang, Z., Zhang, C., Shen, S., Kumar, V.V., Bai, F., Krajewski, P., Deneau, D., Jawad, A., Jones, R., Piotrowicz, G., Liu, H.X., 2024. Traffic light optimization with low penetration rate vehicle trajectory data. Nature Communications 15 (1), 1306.
Wang, X., Lu, L., Zhang, Z., Wang, Y., Li, H., 2025. Introducing the vehicle-infrastructure cooperative control system by quantifying the benefits for the scenario of signalized intersections. Transportation Research Part A: Policy and Practice 192, 104378.
Wang, T., Zhao, J., Li, C., 2019. Pedestrian delay model for continuous flow intersections under three design patterns. Mathematical Problems in Engineering 2019 (1), 1016261.
Yang, X., Abdel-Aty, M., Huan, M., Peng, Y., Gao, Z., 2015. An accelerated failure time model for investigating pedestrian crossing behavior and waiting times at signalized intersections. Accident Analysis & Prevention 82, 154–162. doi:10.1016/j.aap.2015.04.036.
Yazdani, M., Sarvi, M., Asadi Bagloee, S., Nassir, N., Price, J., Parineh, H., 2023. Intelligent vehicle pedestrian light (IVPL): A deep reinforcement learning approach for traffic signal control. Transportation Research Part C: Emerging Technologies 149, 103991. doi:10.1016/j.trc.2022.103991.
Zaidel, D.M., Hocherman, I., 1987. Safety of pedestrian crossings at signalized intersections. Transportation Research Record 1141, 1–6.
Zegeer, C.V., Opiela, K.S., Cynecki, M.J., 1982. Effect of pedestrian Signals and Signal Timing on Pedestrian Accidents. Technical Report.
Zhang, M., Chen, K., Zhu, J., 2023. An efficient planning method based on deep reinforcement learning with hybrid actions for autonomous driving on highway. International Journal of Machine Learning and Cybernetics 14 (10), 3483–3499.
Zhang, X., Chan, L.S., Nassir, N., Sarvi, M., 2025. Towards fair lights: A multi-agent masked deep reinforcement learning for efficient corridor-level traffic signal control. Communications in Transportation Research 5, 100203.
Zhang, Y., Mamun, S.A., Ivan, J.N., Ravishanker, N., Haque, K., 2015. Safety effects of exclusive and concurrent signal phasing for pedestrian crossing. Accident Analysis & Prevention 83, 26–36. doi:10.1016/j.aap.2015.06.010.
Zhang, Z., Li, H., Ren, G., 2024. Evade or rush? Investigating jaywalkers' sequential crossing decisions at mid-blocks without crossing facilities. Travel Behaviour and Society 36, 100799.
Zhu, D., Sze, NN., Bai, L., 2021. Roles of personal and environmental factors in the red light running propensity of pedestrian: Case study at the urban crosswalks. Transportation Research Part F: Traffic Psychology and Behaviour 76, 47–58.
Zhuang, X., Wu, C., Ma, S., 2018. Cross or wait? Pedestrian decision making during clearance phase at signalized intersections. Accident Analysis & Prevention 111, 115–124.