Huiqian LI, Jin HUANG, Zhong CAO, Diange YANG, Zhihua ZHONG, 2023. Stochastic pedestrian avoidance for autonomous vehicles using hybrid reinforcement learning. *Frontiers of Information Technology & Electronic Engineering*, 24(1):131-140. https://doi.org/10.1631/FITEE.2200128

Stochastic pedestrian avoidance for autonomous vehicles using hybrid reinforcement learning

Key words: Pedestrian; Hybrid reinforcement learning; Autonomous vehicles; Decision-making

Corresponding author: Jin HUANG
E-mail: huangjin@tsinghua.edu.cn
ORCID: https://orcid.org/0000-0001-8774-2936

Motivation

1. Interaction with other traffic participants is one of the major considerations in evaluating the autonomous vehicle (AV) safety, which can be intractable even for experienced human drivers. To promote the social acceptance of AVs, researchers must design a robust decision-making system to cope with stochastic and uncertain pedestrian behaviors.

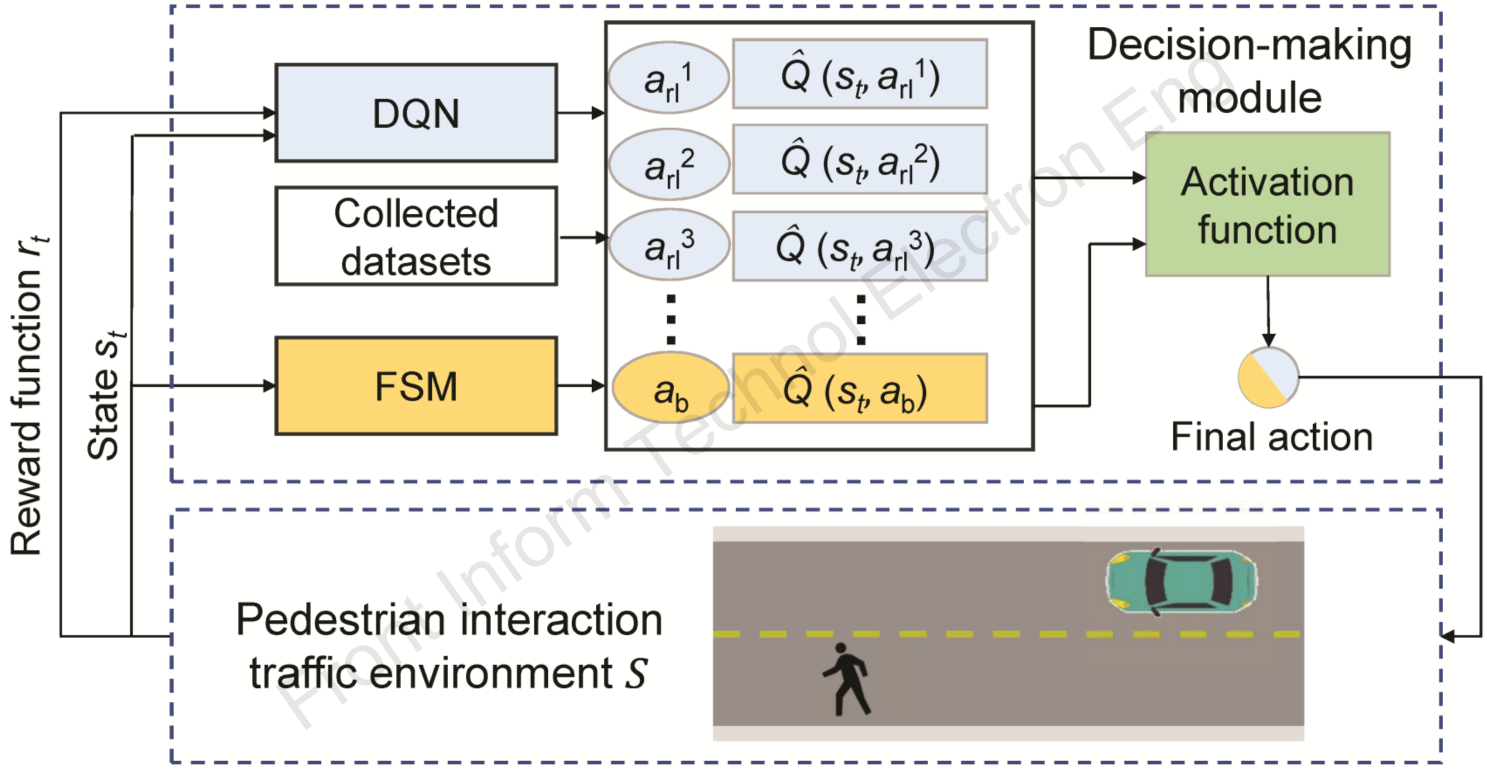
2. Classical methods may fail in some situations where the assumptions in the design process are not consistent with the facts, and they cannot adjust themselves to avoid repeating these mistakes in the future. The performance of learning-based methods cannot be guaranteed without sufficient training.

Main idea

1. To incorporate the strengths of the classical methods and learning-based methods, we establish a hybrid reinforcement learning (HRL) framework in which the ego vehicle can safely interact with pedestrians. The HRL framework starts with a rule-based policy, and the reinforcement learning (RL) policy is integrated to enhance performance.

2. To achieve trustworthy performance improvement, we propose a reliable activation function design to determine whether to activate the rule-based policy or RL policy.

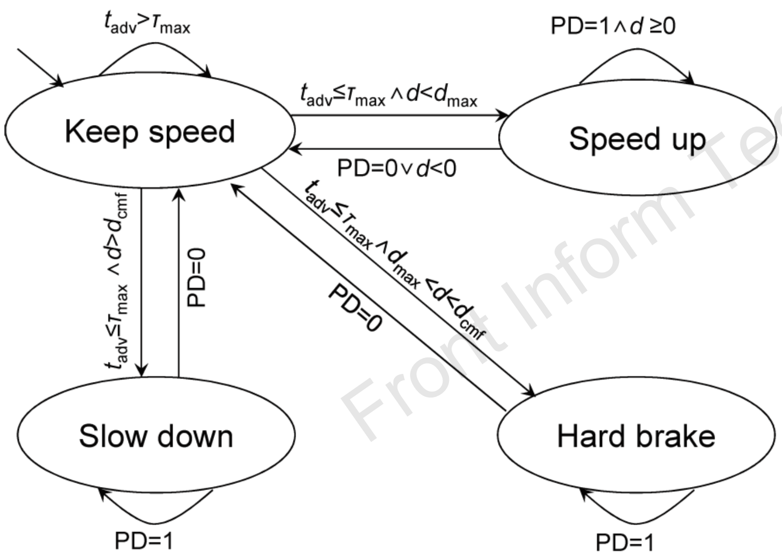
Framework



Framework of the proposed hybrid reinforcement learning (HRL) pedestrian avoidance strategy

Method

1. Baseline rule-based policy. The basic policy designs a finite state machine (FSM) with four modes: keep speed, slow down, hard brake, and speed up.



Basic finite state machine policy

Symbol	Definition
t_{adv}	Time advantage, $\frac{d_y}{v_{\mathrm{p}}} - \frac{d}{v_{\mathrm{v}}}$
τ_{max}	Time advantage threshold
a_{max}	Value of maximum deceleration
a_{cmf}	Value of comfortable deceleration
d_{max}	Maximum braking distance, $\frac{v_v^2}{2a_{\mathrm{max}}}$
d_{cmf}	Comfortable braking distance, $\frac{v_v^2}{2a_{\mathrm{cmf}}}$
PD	1 for pedestrian detected, 0 for not

Variable definitions of the basic policy

Method

2. Hybrid reinforcement learning. We evaluate the rule- and learning-based policies with a uniform criterion. Then we design an activation function to select the final policy. In this way, we can guarantee that the performance of the final policy is not worse than that of the rule-based policy.

$$s = (d, d_y, \psi_{\mathrm{p}}, v_{\mathrm{v}}, v_{\mathrm{p}}) \in \mathcal{S},$$

$$\mathcal{A} = \{a_{\mathrm{ks}}, a_{\mathrm{sd}}, a_{\mathrm{hb}}, a_{\mathrm{su}}\},$$

$$r_1 = \begin{cases} 0, \text{ no collision,} \\ -1, \text{ there is a collision.} \end{cases}$$

$$r_2 = \frac{v}{v_{\mathrm{des}}} - 1,$$

$$a_{\mathrm{rl}} = \underset{a \in \mathcal{A}}{\operatorname{argmax}} Q(s_t, a).$$

$$\pi_{\mathrm{hrl}} = \pi_{\mathrm{rule}} + \frac{\pi_{\mathrm{rl}} - \pi_{\mathrm{rule}}}{1 + \exp(-w\mathcal{C}(\pi_{\mathrm{rl}}, \pi_{\mathrm{rule}}, s))},$$

$$\mathcal{C}(\pi_{\mathrm{rl}}, \pi_{\mathrm{rule}}, s) = Q(s, \pi_{\mathrm{rl}}(s)) - Q(s, \pi_{\mathrm{rule}}(s)) - c_{\mathrm{thre}},$$

Design of the reinforcement learning policy and activation function

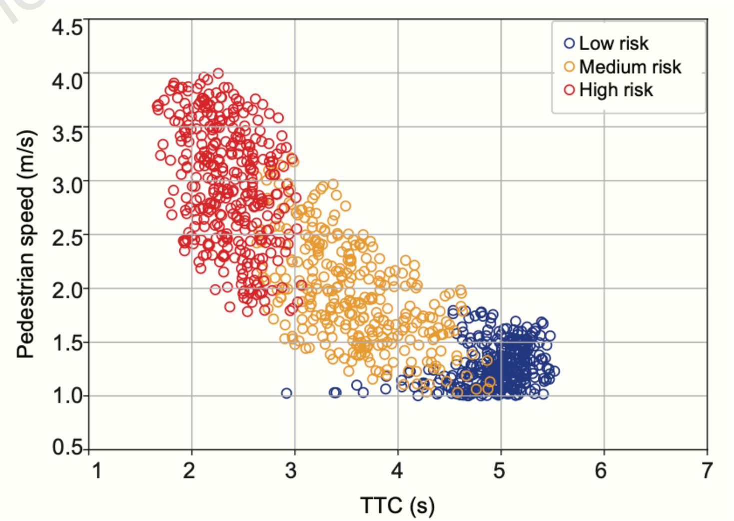
Method

3. Accelerated simulation. To simulate the randomness of pedestrians, we set several types of pedestrians with different behavior patterns and risk levels. Owing to the rareness of safety-critical events, we design a pedestrian generation method for accelerated evaluation to demonstrate the performance of the algorithm.

Type	Speed v_p (m/s)	Orientation ψ_p (°)
Normal	$[1, 2]$	0
Random	$[1.5, 4]$	$[-30, 30]$

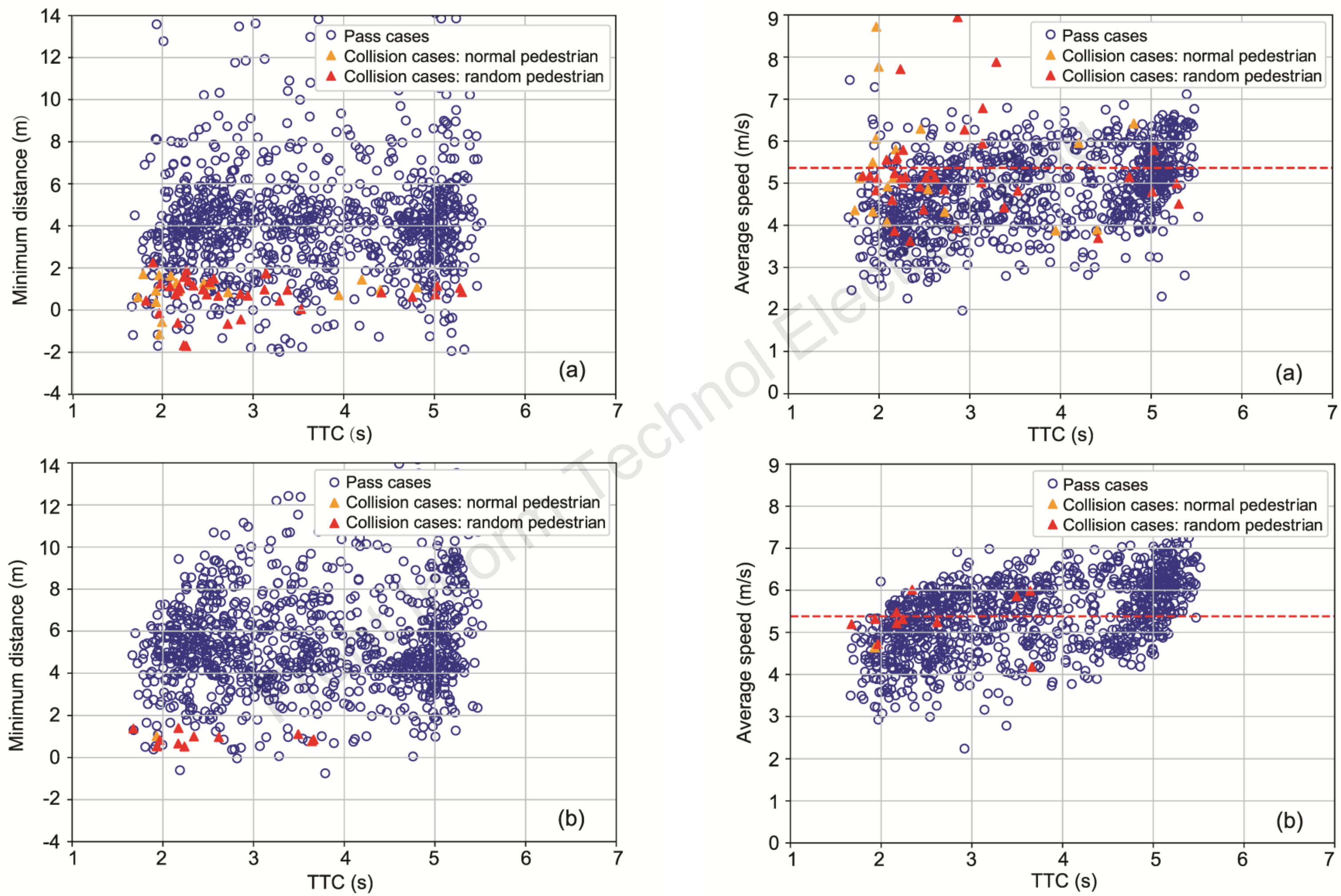
Risk level	Required deceleration (m/s^2)
High	$[-a_{max}, -4.1)$
Medium	$[-4.1, -2.3)$
Low	$[-2.3, 0)$
Trivial	$[0, +\infty)$

Pedestrian behavior patterns and risk levels



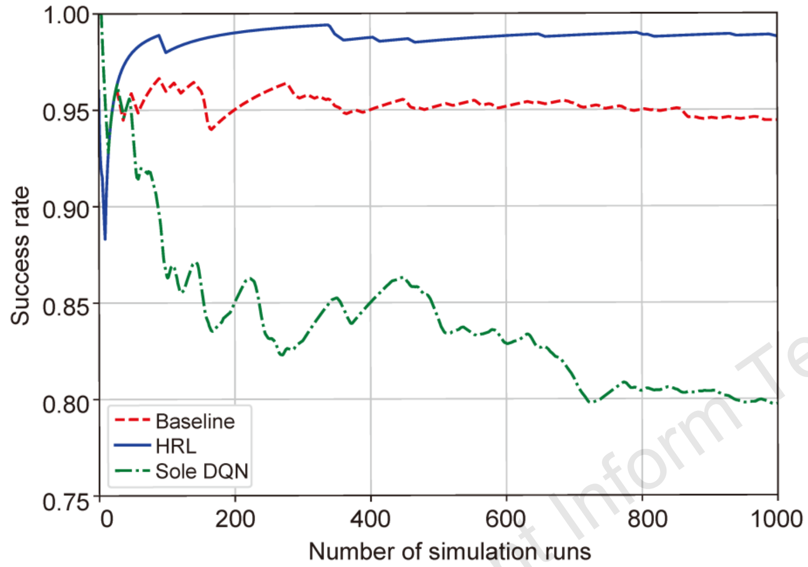
Initial conditions generated according to different risk levels

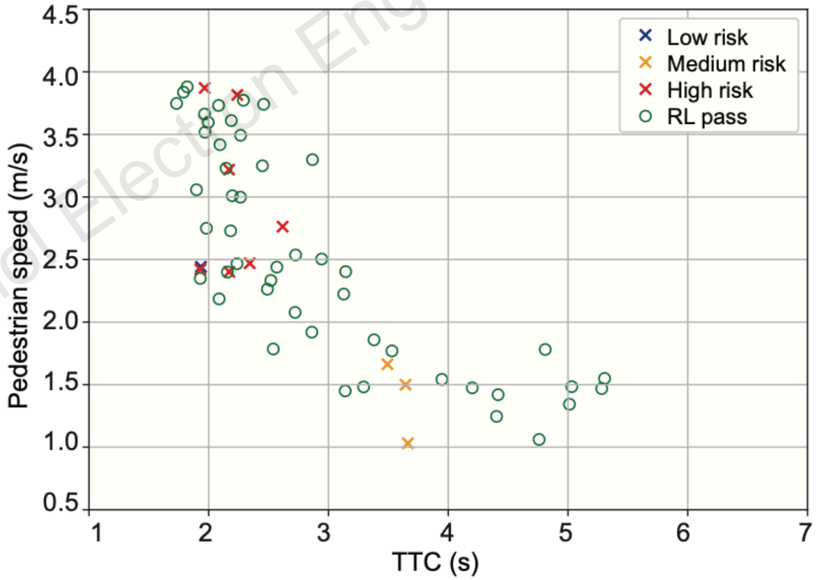
Major results



Relationships between TTC and minimum distance (left) and average speed (right)

Major results





Success rate of the ego vehicle passing the crosswalk

Performance of the proposed method in the cases where the baseline method fails

Conclusions

1. We propose a hybrid reinforcement learning strategy for autonomous vehicles that must handle interaction with stochastic pedestrians. The proposed strategy can adjust the ego vehicle's longitudinal acceleration to improve the success rate when interacting with pedestrians in different risk levels and behavior models.

2. The method is tested on an urban road on the CARLA simulator. The results show that the hybrid reinforcement learning method can increase the success rate by 4.4% compared with the basic rule-based method, and benefits from the introduction of reinforcement learning.