Beyond the Map: Learning to Navigate Unseen Urban Dynamics Using Diffusion-Guided Deep Reinforcement Learning

Monu Nagar, Debasis Das

[nagar.3, debasis]@iitj.ac.in

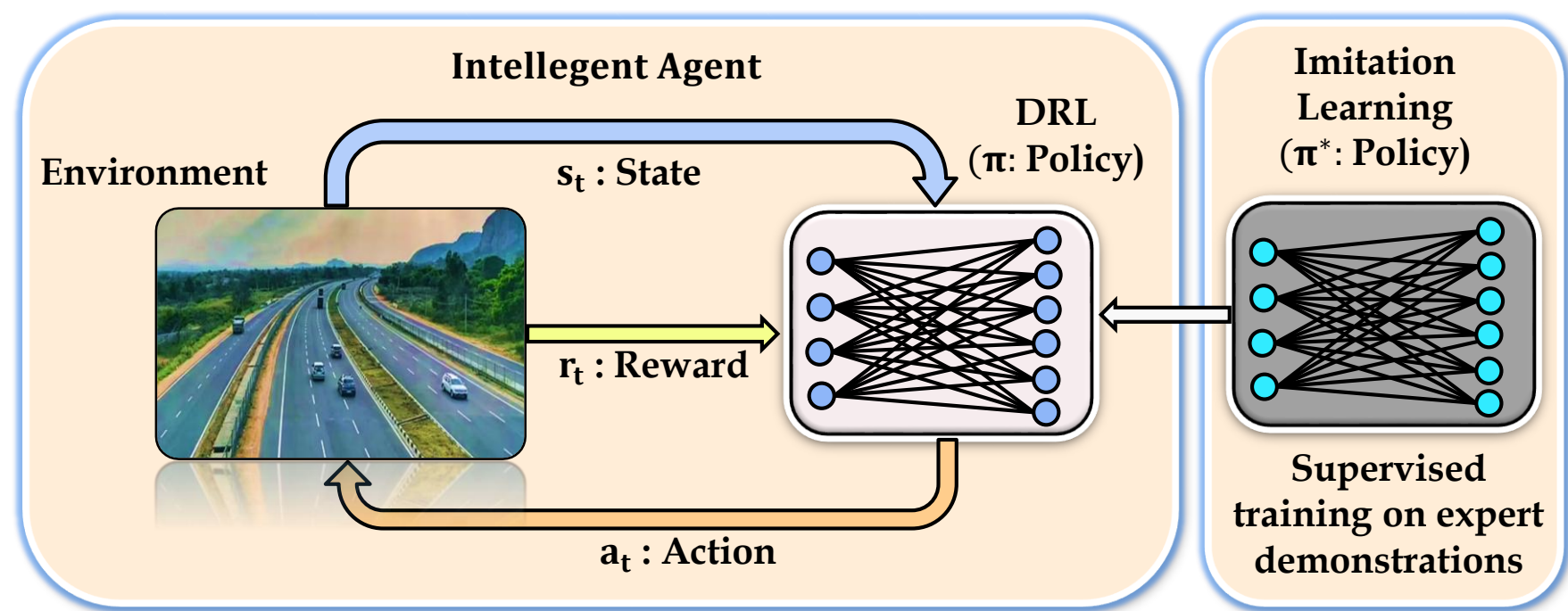Indian Institute of Technology Jodhpur, India

SCAN ME

Problem Statement

➤ Vision-based motion planning is a crucial task for an intelligent agent to effectively learn and adapt to the dynamic characteristics of its environment.

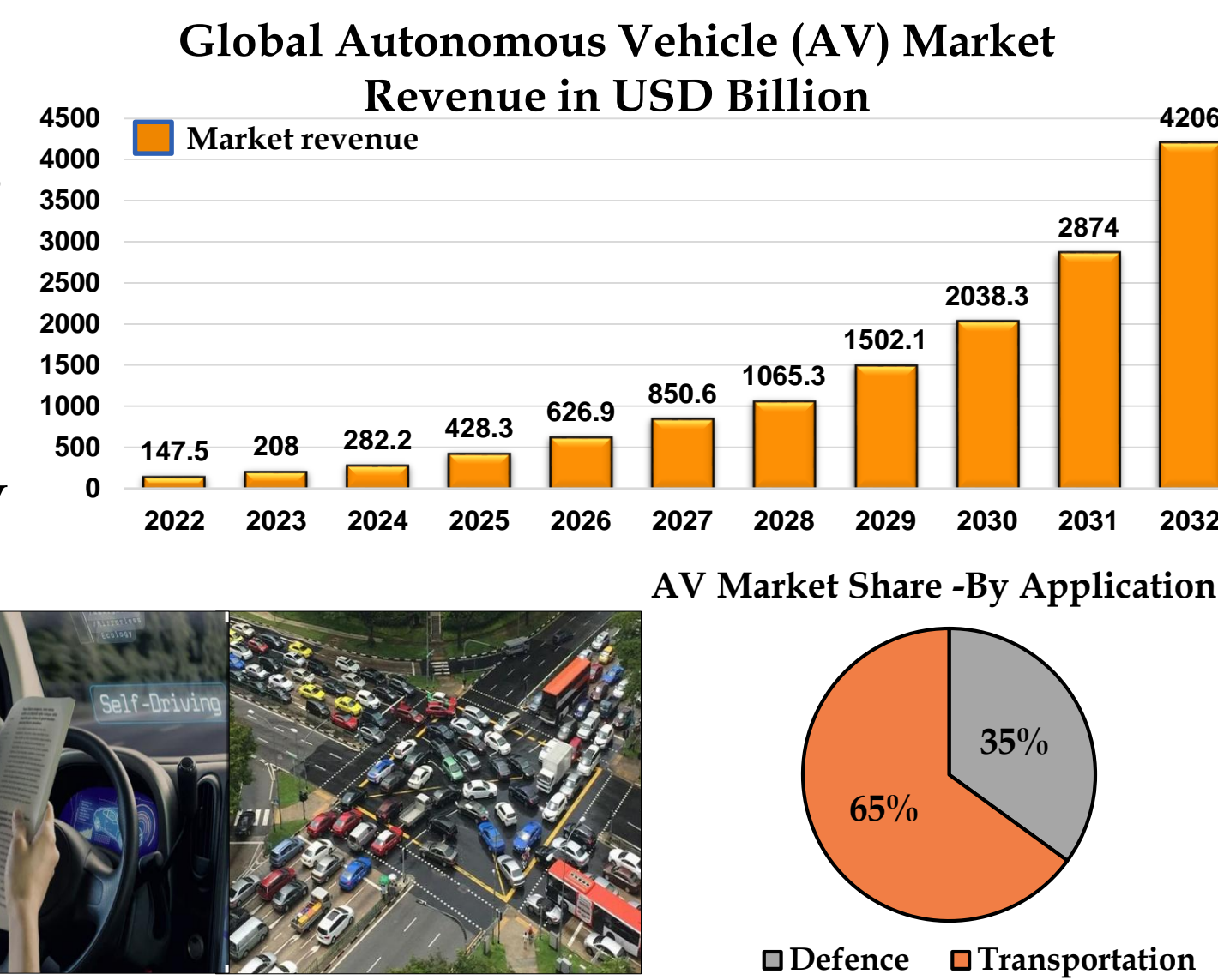➤ **Challenges in IL-Based DRL Approaches:**
- Performance limited by the quality of the expert data.
- Impossible to collect expert data for all driving scenarios.
- Distribution mismatch between training and real-world data causes poor generalization to new states.

Applications

Motion Planning — Intelligent Agent

Motivation

➤ Greater Road Safety
➤ Greater Independence
➤ Reduced Congestion
➤ Increase Productivity
➤ Enhanced accessibility

Global Autonomous Vehicle (AV) Market Revenue in USD Billion

AV Market Share -By Application
65% Transportation, 35% Defence

➤ According to US market reports [5], transportation applications dominate the AV market with 65% share across self-driving cars, trucks, and public transit. This dominance creates significant opportunities for AI-driven solutions in autonomous driving.

Key Contributions

➤ We propose a Diffusion-Guided Deep Reinforcement Learning (**DGDRL**) framework, which integrates diffusion models with Soft Actor-Critic to enhance task generalization and reduce dependency on IL.

➤ We introduce a novel modified Partially Observable Markov Decision Process (**mPOMDP**) to optimize the behavior of the SAC policy for handling uncertainty in dynamic environments.

➤ We conduct extensive testing across multiple CARLA towns under varying weather conditions and traffic densities (Empty, Regular, and Dense scenarios). where DGDRL achieves 95% task completion success rate with enhanced safety

Actor Network Function (π)

$$L_\phi^\pi = -\mathbb{E}_{o_t \sim D}\left[\mathbb{E}_{\tilde{a}_t \sim \pi_\phi(\cdot|o_t)}\left(\min_{k=1,2} Q_{\theta_k}(o_t, \tilde{a}_t) - \alpha \log \pi_\phi(\tilde{a}_t|o_t)\right)\right]$$
$$+ \lambda \, \mathbb{E}_{o_t' \sim \text{Diffusion}(o_t)}\left[D_{KL}(\pi_\phi(\cdot|o_t) \,\|\, \pi_\phi(\cdot|o_t'))\right]$$

Proposed Method : Diffusion-Guided Deep Reinforcement Learning

$o_1(t = 1s)$ $o_2(t = 2s)$ $o_2(t = 3s)$ $o_2(t = 4s)$

a) POMDP   b) Proposed POMDP

Replay Memory : State, Action, Reward, Next-State
V_k : Speed and Steering Angle
I_k : 256x256x3 Image, ▯: Conv2d Layer
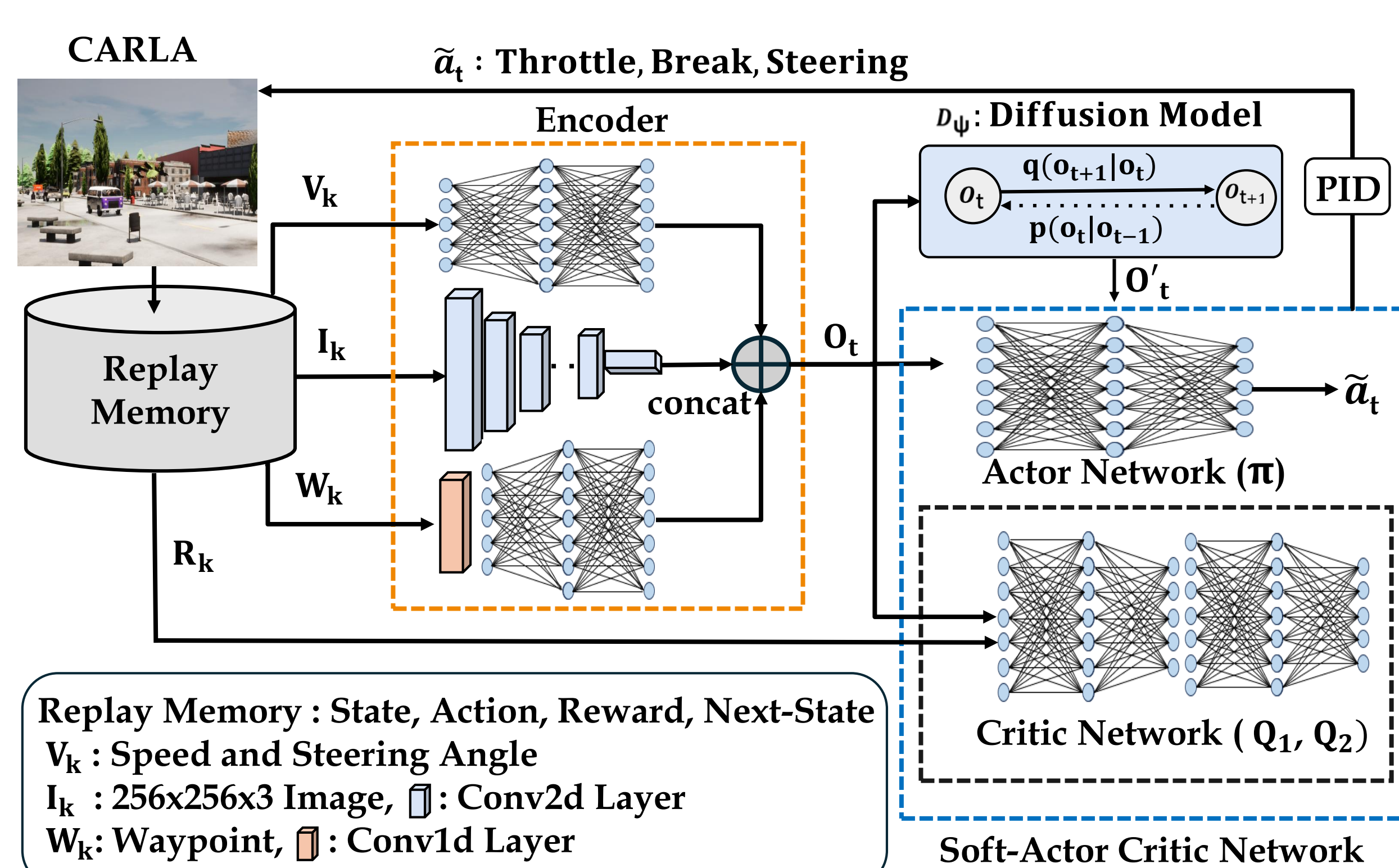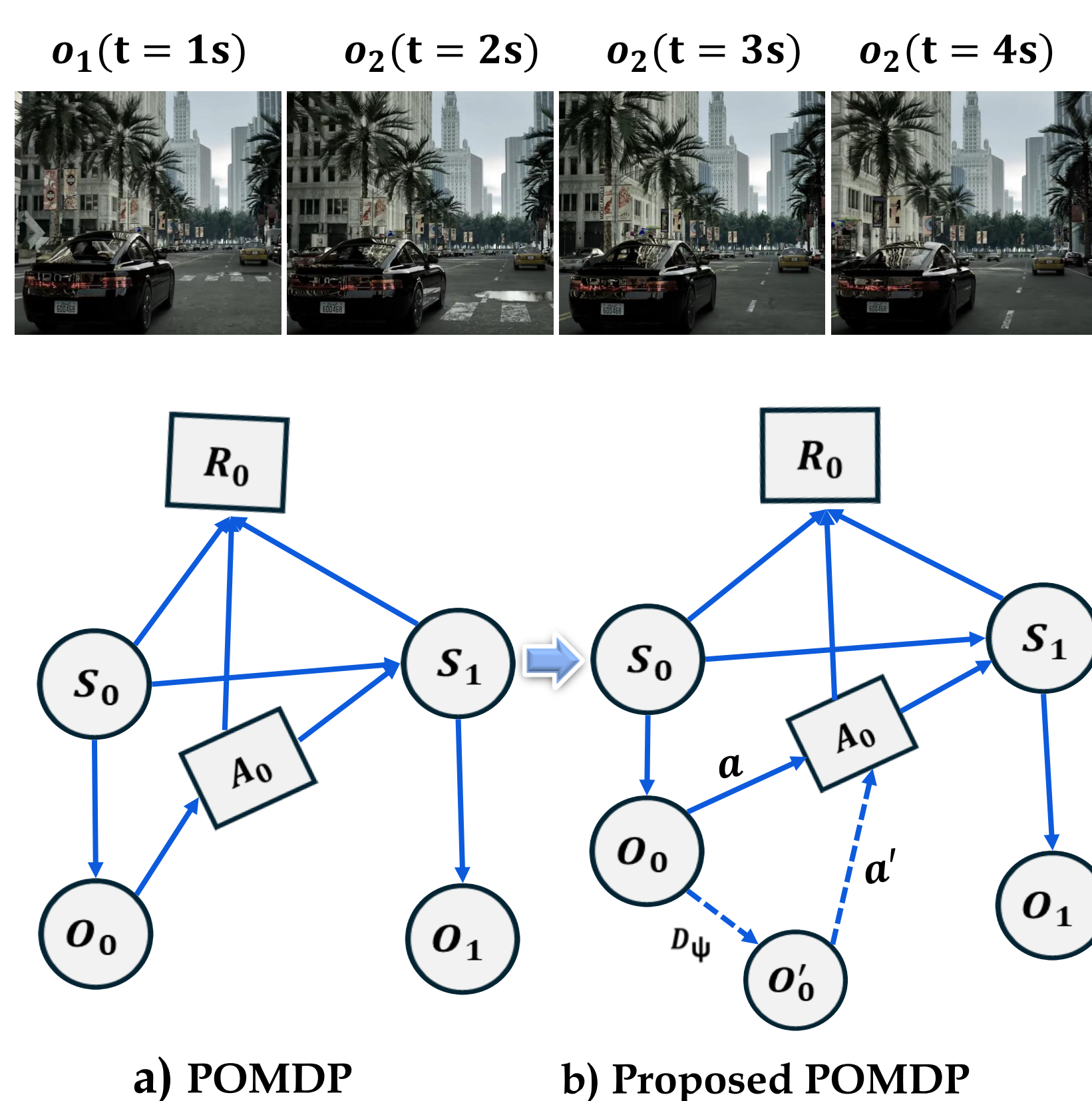W_k : Waypoint, ▯: Conv1d Layer

Soft-Actor Critic Network

Figure 1: The DGDRL framework integrates encoder blocks, diffusion mode, and Soft Actor-Critic networks to address uncertainty and improve generalization in dynamic environments.
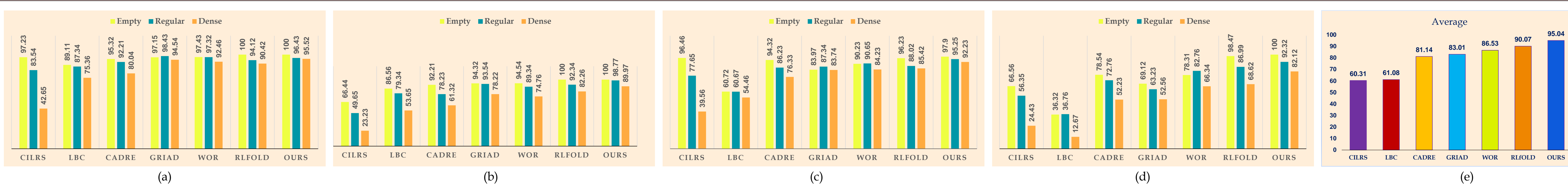
Results

Figure 2: DGDRL performance comparison showing success rates (%) on CARLA NoCrash benchmark across two towns (T01, T02) and weather conditions (W01, W02): (a) T01, W01; (b) T01, W02; (c) T02, W01; (d) T02, W02
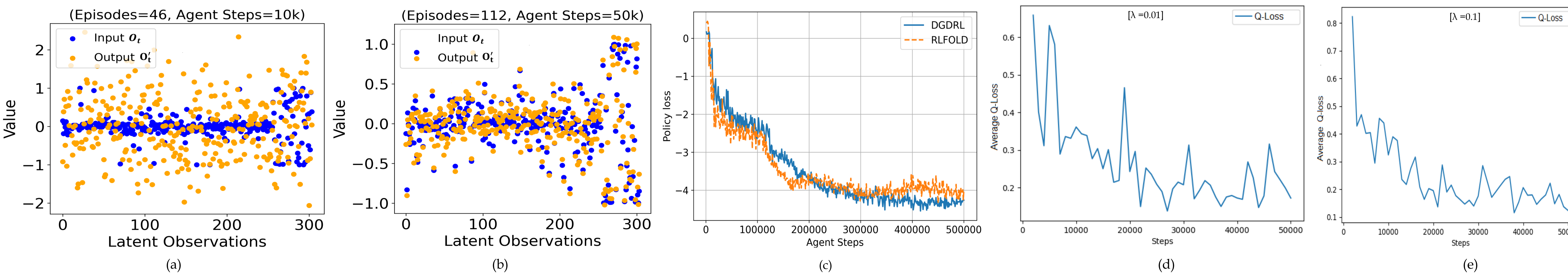

Figure 3: DGDRL training analysis across multiple metrics: (a-b) Latent observation comparison (o_t vs o_t') at 10K and 50K steps showing diffusion model's progressive enhancement in guiding policy network to handle environmental uncertainties, (c) Policy loss evolution over 500K training steps, and (d-e) Q-network loss curves comparing λ values (0.01 vs 0.1) during initial 50K steps to illustrate hyperparameter impact on learning dynamics.
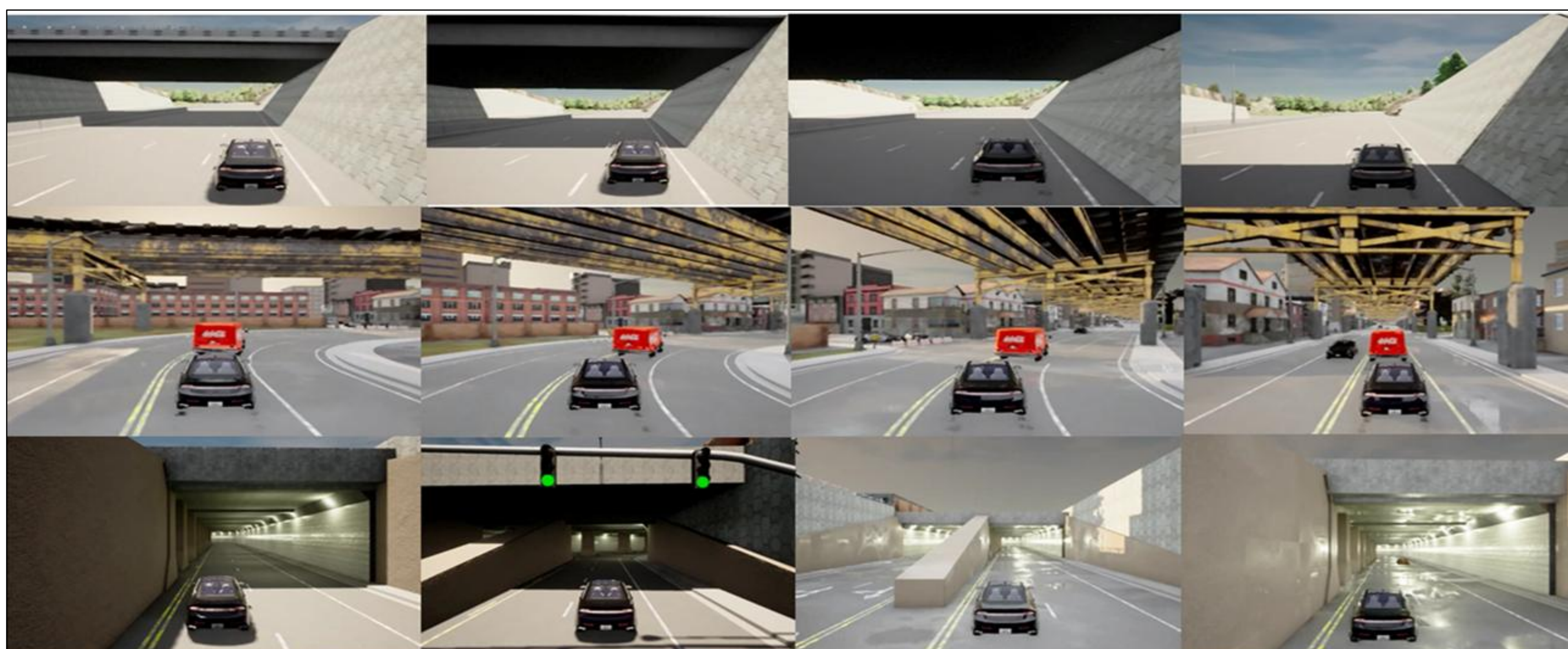

Figure 4: DGDRL generalization analysis in Town02 dense setup. Rows 1-2: successful navigation through complex scenarios including underpasses. Bottom row: failure case where the model encounters a tunnel and halts, unable to make decisions. These examples highlight both model adaptability and limitations in specific environments.

Conclusion and Future Work

➤ **Model:** The proposed Diffusion-Guided Deep Reinforcement Learning (DGDRL) model integrates diffusion models with Soft Actor-Critic under a modified POMDP framework.

➤ **Performance Benchmark:** 95.04% success rate across diverse environments (e.g., new town and weather)

➤ **Key Improvements:** Enhanced generalization to new task and improve pedestrian safety.

➤ **Future Work:** Targets complex scenarios such as tunnels, lane change , and broader domain applications.

Acknowledgement

References

[1] [Coelho et al., 2024] Daniel Coelho, Miguel Oliveira, and Vitor Santos. Rlfold: Reinforcement learning from online demonstrations in urban autonomous driving. In Proceedings of the AAAI Conference on Artificial Intelligence, volume 38, pages 11660–11668, 2024

[2] [Chekroun et al., 2023] Raphael Chekroun, Marin Toromanoff, Sascha Hornauer, and Fabien Moutarde. Gri: General reinforced imitation and its application to visionbased autonomous driving. Robotics, 12(5):127, 2023.

[3] [Zhao et al., 2022] Yinuo Zhao, Kun Wu, Zhiyuan Xu, Zhengping Che, Qi Lu, Jian Tang, and Chi Harold Liu. Cadre: A cascade deep reinforcement learning framework for vision-based urban driving. In Proceedings of the AAAI, volume 36, pages 3481–3489, 2022.

[4] [Jain and Unhelkar, 2024] Abhinav Jain and Vaibhav Unhelkar. Go-dice: Goal-conditioned option-aware offline imitation learning via stationary distribution correction estimation. In Proceedings of the AAAI conference on artificial intelligence, volume 38, pages 12763–12772, 2024.

[5] [Pangarkar, 2025] Tajammul Pangarkar, Autonomous Vehicles Statistics 2025, https://www.news.market.us/autonomous-vehicles-statistics