VOID: VO-Integrated Deep Reinforcement Learning for 3D Dynamic Obstacle Avoidance

Yuting Tao, Mingyang Li, Peng Lu*

Department of Mechanical Engineering, University of Hong Kong, Hong Kong

Email: taoyt@connect.hku.hk; myli0823@connect.hku.hk; lupeng@hku.hk

THE UNIVERSITY OF HONG KONG

ABSTRACT

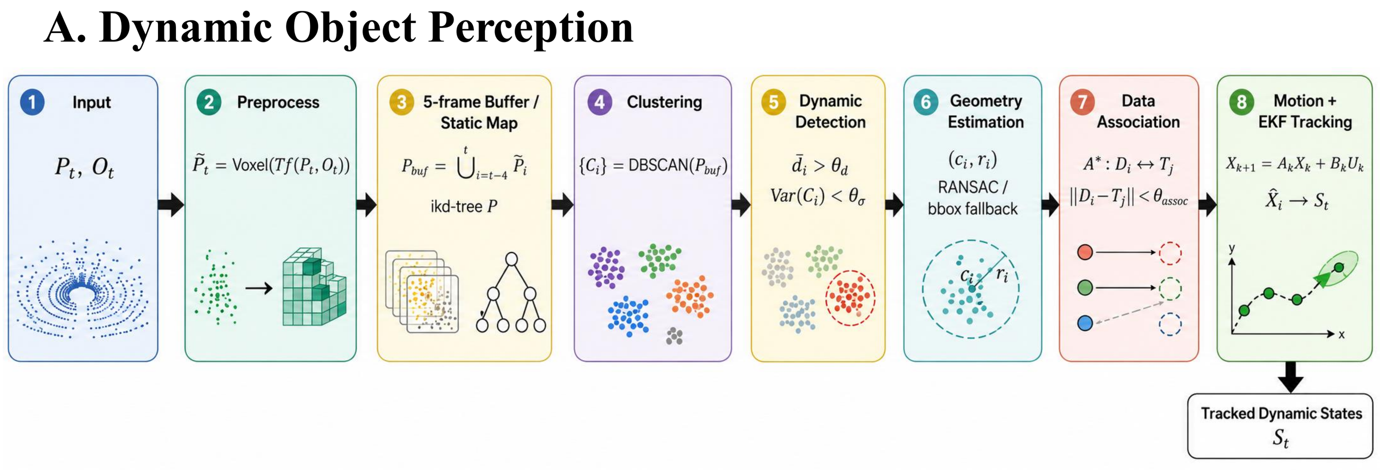
Collision avoidance in dynamic environments with fast-moving obstacles remains a crucial challenge for mobile robots. While end-to-end deep reinforcement learning (DRL) offers streamlined architectures, pure learning-based methods often suffer from limited interpretability and weak generalization. Furthermore, most existing methods assume planar obstacle motion, leading to conservative or inaccurate collision reasoning in 3D space. To address these limitations, we propose a hybrid framework combining **3D LiDAR perception** with a **DRL planner**. The perception module handles **obstacle state extraction** (positions, velocities, and sizes). These extracted states, alongside stacked local grid maps and goal information, feed into a DRL network featuring a **ResNet encoder and residual self-attention** for effective spatiotemporal feature extraction. Crucially, we extend the classical Velocity Obstacle (VO) formulation to **3D** and integrate it into the reward function, ensuring explicit 3D collision reasoning. Simulations and real-world experiments demonstrate that our approach significantly improves navigation performance in complex dynamic environments.

I. Introduction

- **Multi-Feature Tracking**: Extracts dynamic and geometric object features, utilizing independent **Kalman filters** for robust state estimation.
- **Spatio-Temporal DRL:** Integrates **ResNet** and **residual self-attention** to fuse extracted obstacle states and stacked local grid maps.
- **3D Velocity Obstacle:** Extends classical VO to **3D velocity space** within the DRL reward function, expanding the feasible safe velocity set.
- **Validation:** Verifies the framework through extensive **simulations and real-world experiments** across multiple environments.

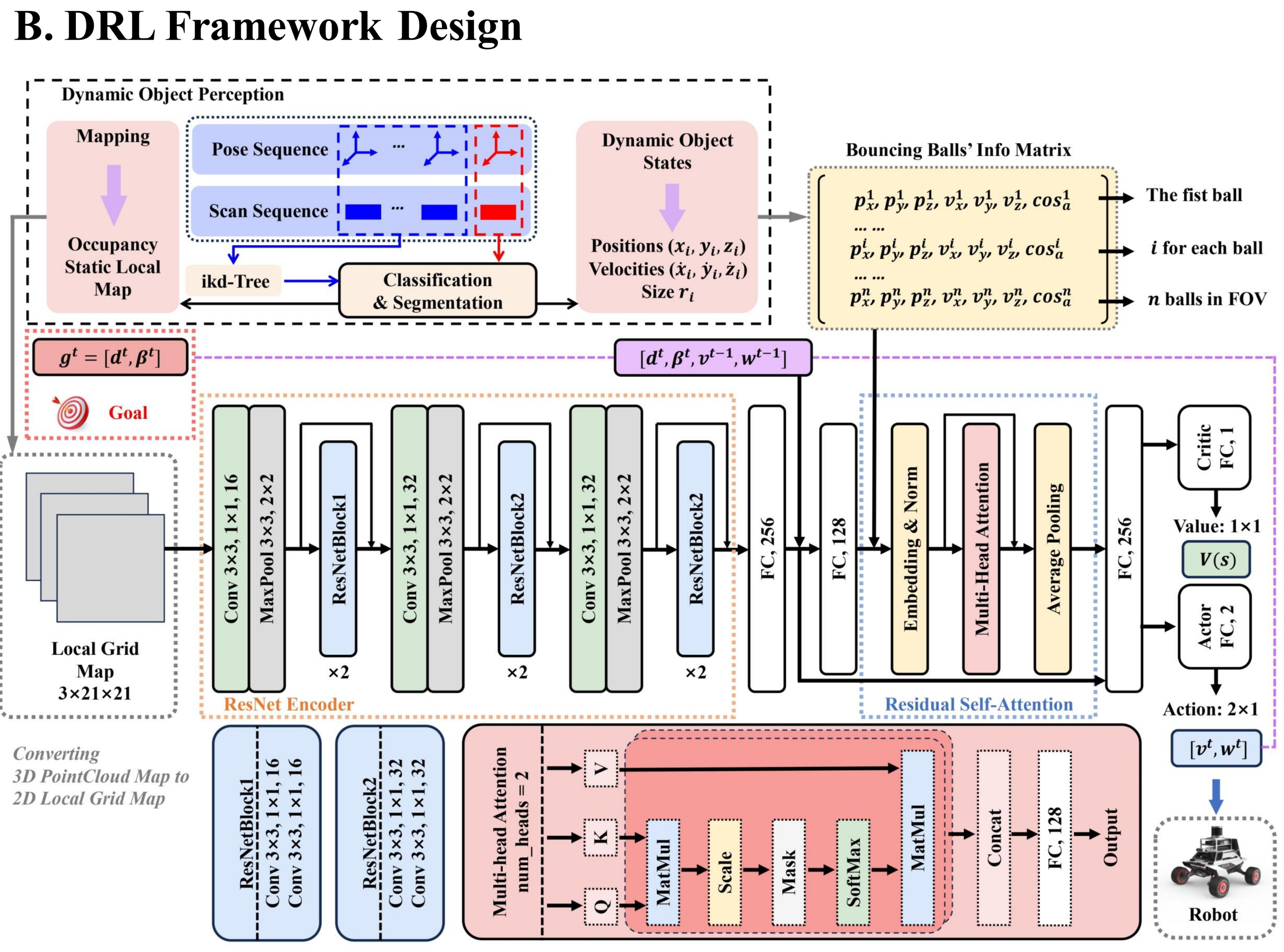
II. Methodology

A. Dynamic Object Perception

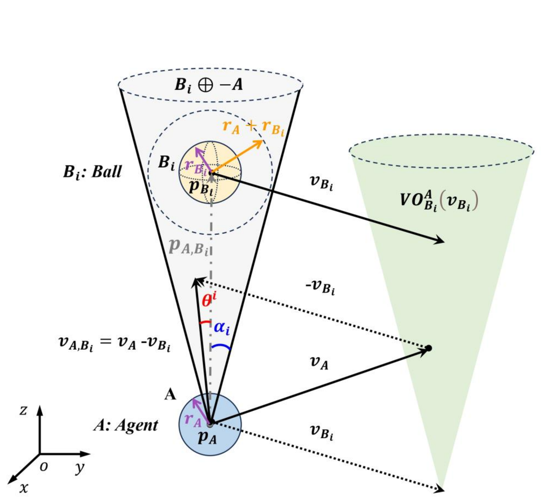


➢ We present a LiDAR-based perception pipeline for dynamic obstacle tracking. Raw point clouds are voxelized and aggregated into a multi-frame static map for robust clustering. Dynamic objects are identified via motion and density consistency, followed by RANSAC-based geometry estimation. Temporal data association links detections to tracks, and an EKF is applied to produce stable motion states S_t.

B. DRL Framework Design



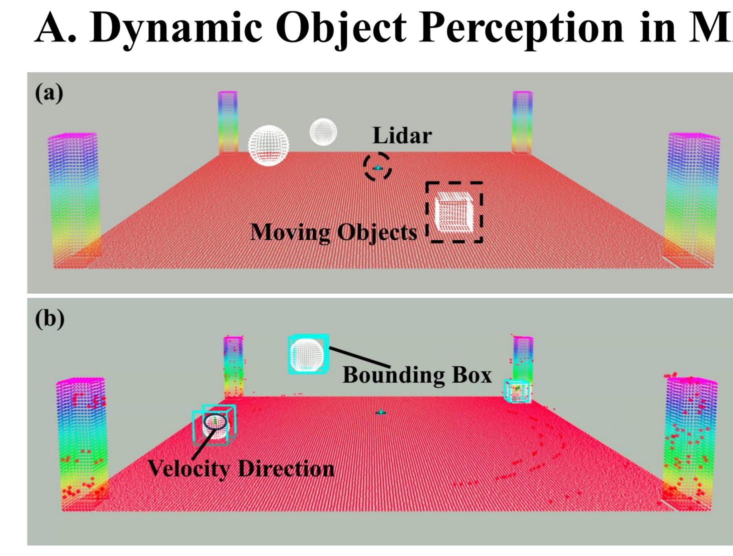
➢ **DRL Navigation Framework Overview.** This framework directly maps multi-modal inputs to continuous robot actions. Static obstacles are captured via a 2D local grid map and encoded by serial **ResNet Blocks** to extract spatial features. Concurrently, real-time dynamic obstacle states from the perception module form a variable-length matrix, which is dynamically aggregated through a **Residual Self-Attention** mechanism to handle an arbitrary number of threats. Finally, the extracted static maps, attention-weighted obstacle features, relative goal vector g^t, and current robot velocities are concatenated to feed **the Actor-Critic heads**, outputting the state value $V(s)$ and optimal control commands $[v^t, w^t]$.



➢ **3D VO Reward.** We extend Velocity Obstacle reasoning to 3D velocity space and use it as a reward term for dynamic collision avoidance. For each moving object, the policy is rewarded when the relative velocity stays outside the 3D collision cone, and penalized when it points toward a potential collision.

III. Experiments

A. Dynamic Object Perception in MARSIM



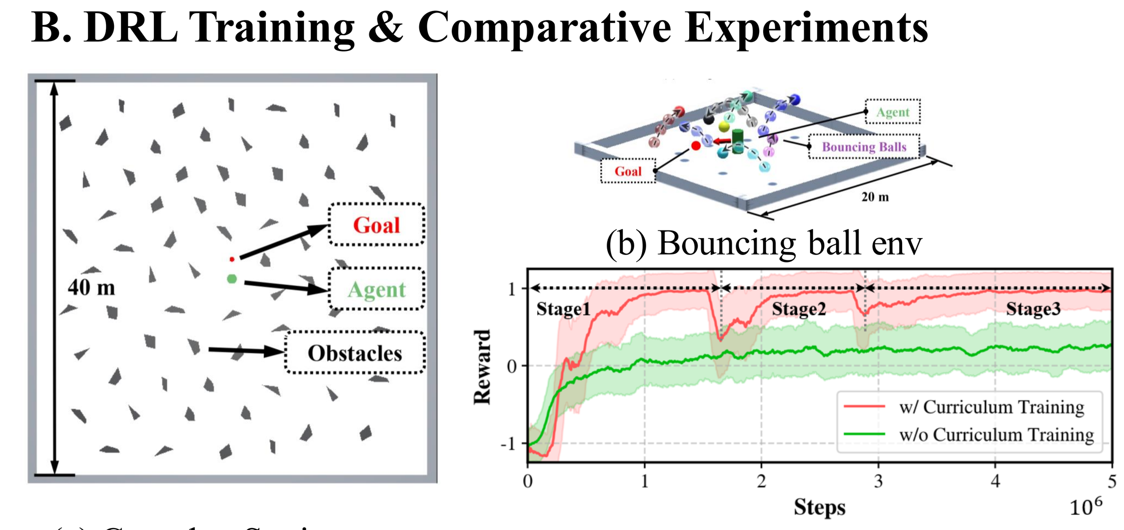
DYNAMIC PERCEPTION COMPARISON

Size (m)	0.2	0.4	0.6	0.8	1.6
	e_{pos} / e_{vel}	e_{pos} / e_{vel}	e_{pos} / e_{vel}	e_{pos} / e_{vel}	e_{pos} / e_{vel}
FAPP [1]	0.08 / 3.19	0.16 / 2.77	0.28 / 2.13	0.42 / 1.96	0.80 / 2.35
Ours	0.12 / 2.83	0.18 / 2.58	0.22 / 1.91	0.31 / 1.74	0.43 / 2.12

➢ Our LiDAR-based perception module detects and tracks dynamic objects in MARSIM, estimating object centers, sizes, and velocities for downstream planning. Compared with FAPP, it is more robust to large objects by estimating geometric centers rather than relying on surface points, reducing position error by 21.4% at 0.6 m. The gravity - aware model further improves velocity estimation for free-falling obstacles.

[1] M. Lu, X. Fan, H. Chen, and P. Lu, "FAPP: Fast and adaptive perception and planning for uavs in dynamic cluttered environments," IEEE Trans. Robot., vol. 41, pp. 871–886, 2025.

B. DRL Training & Comparative Experiments

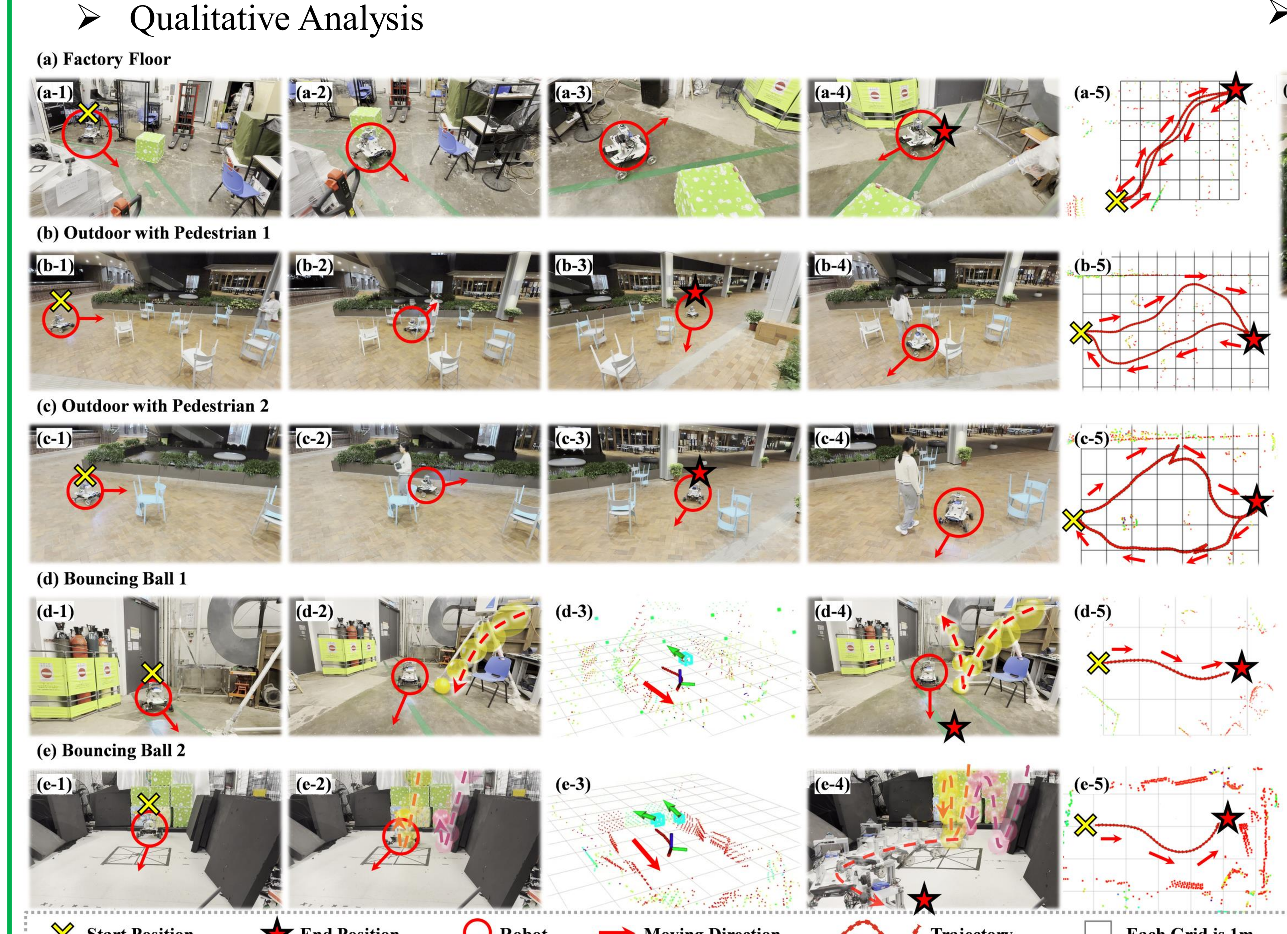


(a) Complex Static env (b) Bouncing ball env (c) Curriculum Learning

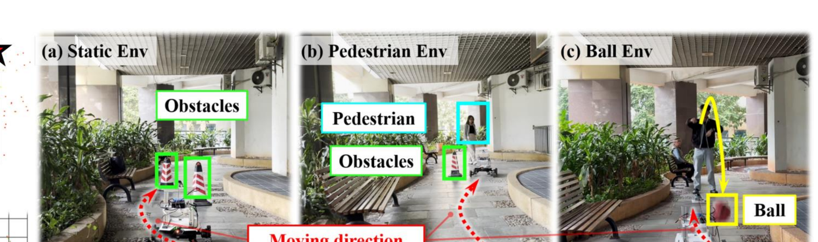
Methods	SR (%)	Time (s) / Std	Length (m) / Std	Speed (m/s)
	Max_speed = 1 m/s, Ball_num = 2			
APF [2]	50	10.22 / 1.63	10.16 / 1.64	**0.99 / 0.01**
VO [3]	49	9.90 / 0.61	9.11 / 0.26	0.92 / 0.03
DRL_VO [4]	63	10.36 / 0.91	9.12 / 0.77	0.88 / 0.27
Carl_MC [5]	66	10.02 / 1.03	**8.62 / 0.58**	0.86 / 0.07
Ours	**76**	**8.99 / 0.66**	8.86 / 0.68	**0.99 / 0.10**
	Max_speed = 1 m/s, Ball_num = 4			
APF	36	10.53 / 1.30	10.48 / 1.30	**0.99 / 0.01**
VO	36	**9.88 / 0.67**	9.12 / 0.30	0.93 / 0.03
DRL_VO	37	10.73 / 1.02	8.91 / 0.23	0.83 / 0.52
Carl_MC	39	10.24 / 1.00	**8.48 / 0.24**	0.83 / 0.06
Ours	**53**	9.92 / 0.77	9.82 / 0.67	**0.99 / 0.13**
	Max_speed = 1 m/s, Ball_num = 8			
APF	13	11.03 / 1.58	10.95 / 1.60	0.99 / 0.01
VO	15	10.20 / 0.76	9.27 / 0.35	0.91 / 0.04
DRL_VO	15	11.57 / 0.87	9.61 / 0.22	0.83 / 0.46
Carl_MC	16	10.15 / 1.12	**8.63 / 0.34**	0.85 / 0.07
Ours	**34**	**9.57 / 0.72**	9.57 / 0.41	**1.00 / 0.07**

C. Real-world Experiments

➢ Qualitative Analysis



➢ Quantitative Analysis



THE REAL-WORLD EXPERIMENT COMPARISON

Metrics	Success Rate (%)	Path Length / Std (m)
Static Env		
Carl_MC [5]	90.0	**7.59 / 0.32**
Ours	**100.0**	8.03 / 0.79
Pedestrian Env		
Carl_MC	66.7	**8.52 / 0.23**
Ours	**90.9**	8.75 / 0.67
Bouncing-ball Env		
Carl_MC	18.2	**7.67 / 0.43**
Ours	**63.6**	7.82 / 0.31

Metrics	Execution Time / Std (s)	Average Speed / Std (m/s)
Static Env		
Carl_MC [5]	17.65 / 3.23	0.43 / 0.13
Ours	**16.06 / 3.08**	**0.50 / 0.07**
Pedestrian Env		
Carl_MC	19.36 / 2.78	0.44 / 0.17
Ours	**17.86 / 2.57**	**0.49 / 0.07**
Bouncing-ball Env		
Carl_MC	17.43 / 3.11	0.44 / 0.11
Ours	**15.64 / 2.24**	**0.50 / 0.06**

[2] C. W. Warren, "Global path planning using artificial potential fields," in Proc. IEEE Intl. Conf. on Robot. and Autom. IEEE Computer Society, 1989, pp. 316–317

[3] J. Van den Berg, M. Lin, and D. Manocha, "Reciprocal velocity obstacles for real-time multi-agent navigation," in Proc. IEEE Intl. Conf. on Robot. and Autom. Ieee, 2008, pp. 1928–1935.

[4] Z. Xie and P. Dames, "DRL-VO: Learning to navigate through crowded dynamic scenes using velocity obstacles," IEEE Trans. Robot., vol. 39, no. 4, pp. 2700–2719, 2023.

[5] Y. Tao, M. Li, X. Cao, and P. Lu, "Mobile robot collision avoidance based on deep reinforcement learning with motion constraints," IEEE T. Intell. Veh., pp. 1–11, 2024.

IV. Conclusion

This work presents a hybrid navigation framework integrating dynamic obstacle perception with a DRL planner. Extending our previous work, we actively detect and track dynamic obstacles via point clouds to improve state estimation accuracy. The estimated obstacle states, stacked local grid maps, and goal info are processed by a DRL network using a ResNet encoder and residual self-attention for spatiotemporal feature extraction. Furthermore, we extend the classical VO to 3D and embed it into the reward function, explicitly modeling obstacle motion and expanding the safe velocity space. Experiments demonstrate effective navigation in complex dynamic scenarios. Future work will address current limitations in long-horizon trajectory prediction and terrain adaptation.

V. Useful Links

ArcLab Website — Our Lab

ArcLab Github — Github

Paper Video — YouTube

Tao Yuting — My Twitter

Publication Note

This poster is based on prior published work: Y. Tao, M. Li, J. Kang, X. Cao, and P. Lu, "VOID: VO-Integrated Deep Reinforcement Learning for 3-D Dynamic Obstacle Avoidance," IEEE Transactions on Industrial Electronics, 2026, doi: 10.1109/TIE.2026.3684260.

THE UNIVERSITY OF HONG KONG

IROS 2026 PITTSBURGH