

Stability-Aware UAV Trajectory Planning with Sub-Path Decomposition via Diffusion-Reinforced Decision Making

Minjoo Kim, and Soohyun Park

Division of Computer Science, Sookmyung Women's University, Seoul, Republic of Korea

E-mail: minjoo@sookmyung.ac.kr, soohyun.park@sookmyung.ac.kr

Abstract—Autonomous unmanned aerial vehicles (UAVs) deployed in mobile access environments encounter challenges arising from dense and irregular obstacles as well as stringent energy constraints. Reliable trajectory planning in such a context is critical, since UAVs must maintain connectivity and coverage while navigating through cluttered spaces. Existing diffusion-based reinforcement learning approaches have been explored, but they are generally trained on entire trajectories, which limits their generalization and leads to high re-planning latency in safety-critical scenarios. To address these limitations, we propose a sub-path decomposition framework that generates candidate segments with a diffusion model and applies reinforcement learning to select among them. This design enables finer-grained decision-making, improved safety, and low-latency trajectory updates. Simulations in procedurally generated mobile access environments demonstrate higher success rates and lower collision risks compared to heuristic baselines, while maintaining comparable energy efficiency. Furthermore, the diffusion model trained with sub-paths exhibits lower collision rates and shorter trajectories than the model trained on full paths. This combination of generative modeling and reinforcement learning provides a scalable approach to safe UAV navigation in mobile access scenarios.

I. INTRODUCTION

Background and Motivation. Unmanned aerial vehicles (UAVs) have emerged as a platform for applications such as mobile access, surveillance, delivery, and disaster response. In mobile access environments, UAVs are often required to act as aerial base stations or coverage extenders while simultaneously ensuring safe and energy-efficient flight in cluttered spaces with dense, irregular obstacles. These settings present substantial challenges. Vehicles must navigate around buildings, street furniture, and utility poles under on-board energy capacity. Traditional path planning methods such as rapidly-exploring random tree (RRT) and A^* algorithm can provide feasible paths in static maps, but often require frequent global re-planning when local map updates or constraint changes occur, leading to computational overhead and latency. End-to-end trajectory learning and deep reinforcement learning (DRL) offer adaptability by learning navigation policies directly from interaction with the environment, enabling UAVs to react to changes and optimize trajectories in a data-driven manner. Generative models have recently been explored for trajectory generation because they can produce diverse candidate trajectories conditioned on the UAV state, target, and obstacle maps, capturing multi-modality and uncertainty. However, training

on full trajectories and re-planning entire paths in real time can be inefficient. This motivates a sub-path decomposition strategy in which trajectories are segmented into parts. By resampling only the affected sub-paths in response to environmental changes, latency can be reduced. Furthermore, diffusion-based sub-path generation is combined with an on-line selection mechanism inspired by reinforcement learning concepts, where candidate segments are evaluated through a Q-function-style scoring formulation. This scoring balances safety, energy efficiency, and smoothness by rapidly assessing features such as kinematics, obstacle clearance, energy consumption, and model uncertainty, thereby enabling low-latency and safe trajectory updates.

Contribution. The main contributions of this research are summarized as follows.

- Diffusion-based sub-path trajectory generation: A diffusion model generates diverse candidate UAV trajectories segmented into fixed-length time windows, enabling efficient re-planning by resampling only affected sub-segments.
- RL-based trajectory selection with uncertainty awareness: A Q-function guided by a reward formulation rapidly evaluates and selects among candidate sub-paths using features such as kinematics, obstacle clearance, energy consumption, and model uncertainty.
- Integrated framework for UAV navigation in mobile access environments: The combination of diffusion-based trajectory generation and RL-based online selection yields a scalable, low-latency, and energy-efficient navigation strategy tailored to mobile access scenarios.

Organization. The remainder of this paper is organized as follows. Section II provides an overview of prior research, focusing on reinforcement learning (RL) approaches for UAV navigation and diffusion model-based methods for trajectory planning. Section III presents the main concept of the proposed UAV path stabilization approach. Section IV presents a performance evaluation of the proposed algorithms. Lastly, Section V concludes the paper with a summary of the findings and implications for future research directions.

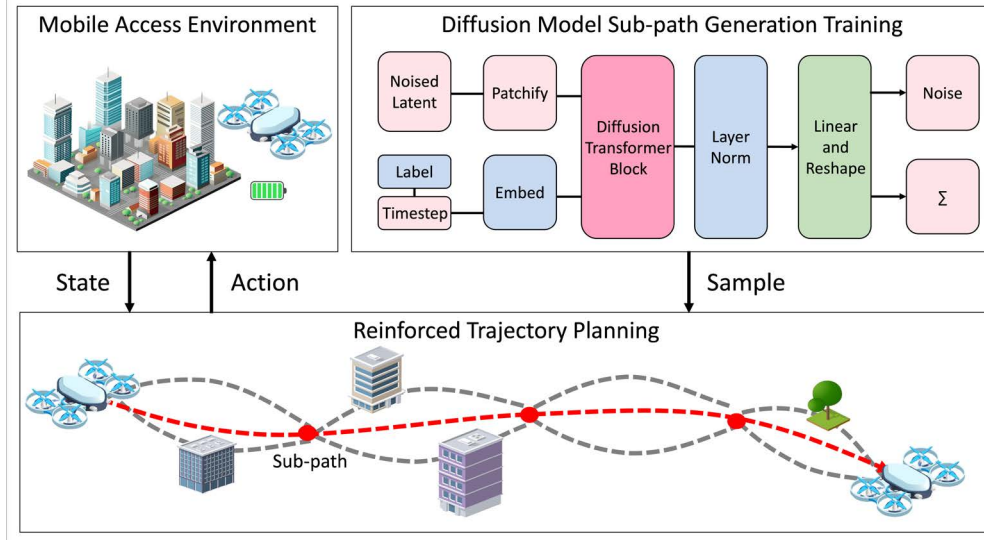


Fig. 1: System overview

II. RELATED WORK

A. Reinforcement Learning for UAV Navigation

UAV navigation is regarded as a robust solution for autonomous flight, obstacle avoidance and trajectory optimization in complex and uncertain environments. While traditional path planning algorithms such as A^* and Dijkstra remain effective in static or low-dimensional scenarios, their reliance on pre-computed maps and high re-planning costs limits their applicability in dynamic settings. DRL emerged as a promising alternative by formulating navigation as a Markov Decision Process (MDP) and allowing agents to learn policies through interaction with the environment [1]–[3]. By leveraging neural function approximators, DRL methods are capable of handling high-dimensional state spaces and can adapt to previously unseen situations [4]. Representative algorithms include value-based methods such as the deep q-network (DQN), policy gradient methods such as proximal policy optimization (PPO), and hybrid approaches such as Actor-Critic algorithms [5], [6]. These approaches have been widely applied to UAV navigation tasks, including collision avoidance in cluttered 3D environments, autonomous exploration, and energy-aware trajectory planning [7].

Nevertheless DRL-based navigation is not without limitations. A well-known drawback is its high sample complexity, requiring large values of training data and extensive simulation before deployment [8]. Furthermore DRL policies often suffer from limited generalization when transferred from simulation to the real world due to domain shift, and real-time responsiveness can be hindered by slow re-planning speeds in highly dynamic environments [9]. To address these challenges recent advances have incorporated transformer-based architectures for better temporal reasoning, model-based RL to improve sample efficiency, safe RL to enforce safety guarantees during exploration, and self-supervised pretraining to enhance representation learning [10], [11]. Such extensions highlight the

growing potential for adaptive UAV navigation in real-world urban scenarios.

B. Diffusion Model Based Trajectory Planning

Diffusion models have recently emerged as generative frameworks in computer vision, particularly for image synthesis, by leveraging a forward process that gradually adds noise to data and a reverse process that removes noise to reconstruct realistic samples [12]. This iterative refinement mechanism has been extended to trajectory generation, enabling the synthesis of feasible and realistic paths such as drone flight routes or robotic trajectories [13], [14]. Unlike conventional approaches that predict a trajectory in a single step, diffusion-based methods benefit from iterative refinement through multiple denoising stages [15]. This property makes them particularly effective characterized by uncertainty and complex constraints. Recent research on diffusion models for trajectory generation highlights several advantages. First, they provide flexible and practical path generation under diverse constraints, such as varying goal locations. Second, the iterative denoising process naturally yields not only a single solution but also a diverse set of candidate trajectories, thereby enhancing solution diversity [16], [17]. Third, diffusion models can be trained using real-world trajectory data alone, improving generalization ability and practical applicability. Moreover, compared to earlier generative frameworks such as GANs and VAEs, diffusion models mitigate issues such as training instability and mode collapse while offering strong representational power for spatiotemporal data. Their conditional generation capability further supports the incorporation of task-specific constraints, including obstacle avoidance, goal conditions, and energy limitations. These properties make diffusion models promising for real-world applications such as autonomous driving, robotics, and drone mission planning, and position them as a central direction for future trajectory generation research.

III. PROPOSED METHOD

A. System Model

UAV Trajectory Metrics. Fig. 1 presents the architecture of the proposed system. In the mobile access environment, the UAV prioritizes energy efficiency, with diffusion models employed for trajectory generation. Furthermore, sub-path-based trajectory planning is carried out using an online selection mechanism inspired by reinforcement learning concepts through a Q-function evaluation. The UAV trajectory planning problem is modeled as a discrete-time system with simplified kinematics and energy accounting. The state at time t is

$$y_t = [p_t, v_t, g - p_t, E_t], \quad (1)$$

where $p_t \in \mathbb{R}^3$ and $v_t \in \mathbb{R}^3$ denote the UAV position and velocity, g is the goal position, and $E_t \in [0, 1]$ is the residual energy ratio. Given a candidate trajectory $\tau = \{p_1, \dots, p_T\}$, we define the following geometric quantities,

$$L(\tau) = \sum_{i=1}^{T-1} \|p_{i+1} - p_i\|, \quad (2)$$

$$\kappa(\tau) = \sum_{i=2}^{T-1} \arccos\left(\frac{(p_{i+1} - p_i) \cdot (p_i - p_{i-1})}{\|p_{i+1} - p_i\| \|p_i - p_{i-1}\|}\right), \quad (3)$$

$$d_{\min}(\tau) = \min_i d(p_i; O), \quad (4)$$

where $L(\tau)$ is the total Euclidean path length, $\kappa(\tau)$ is the aggregate turning angle that serves as a smoothness measure, and $d_{\min}(\tau)$ is the minimum Euclidean clearance from the trajectory to the obstacle set O with $d(\cdot; O)$ denoting the point-to-obstacle distance.

Energy consumption. Energy consumption follows a multirotor UAV propulsion model with optional compute power. For segment speed V and local turn radius r (estimated from three consecutive points), the propulsion power is

$$P_{\text{prop}}(V, r) = P_0 \left(1 + \frac{3V^2}{U_{\text{tip}}^2}\right) + P_i \phi(V, r) + \frac{1}{2} d_0 \rho s A V^3, \quad (5)$$

where $\phi(V, r) = \sqrt{1 + \frac{V^4}{r^2 g^2}}$ for turning flight, and $\phi(V, \infty) = \sqrt{1 + \frac{V^4}{4v_0^4}} - \frac{V^2}{2v_0^2}$ for straight or hover. Here P_0 and P_i are profile and induced power coefficients, U_{tip} is the rotor tip speed, v_0 is the induced velocity in hover, d_0 is the fuselage drag ratio, ρ is air density, s is rotor solidity, and A is rotor disc area. The onboard compute power is modeled as

$$P_{\text{comp}}(u) = P_{\text{idle}} + u (P_{\text{max}} - P_{\text{idle}}), \quad u \in [0, 1], \quad (6)$$

where P_{idle} and P_{max} are idle and maximum compute power, and u is the utilization factor. Hence the total energy along a trajectory τ is

$$E(\tau) = \sum_{i=1}^{T-1} \left(P_{\text{prop}}(V, r_i) + P_{\text{comp}}(u_i) \right) \Delta t_i, \quad \Delta t_i = \frac{\|p_{i+1} - p_i\|}{V}, \quad (7)$$

where r_i is the local turn radius and Δt_i is the duration of each segment under cruise speed V .

B. Sub-Path Diffusion for Candidate Generation

To improve adaptability to changing environments, the full UAV trajectory is decomposed into fixed-length time-step sub-paths. Each sub-path is generated independently, which enables efficient local re-planning when environmental changes occur or new static obstacles are detected. Candidate sub-paths are generated using a conditional diffusion model, where the diffusion process consists of a forward noise-adding stage and a reverse denoising stage. Given a sub-path trajectory x_0 , the forward process gradually adds Gaussian noise.

$$q(x_t | x_{t-1}) = \mathcal{N}\left(x_t; \sqrt{1 - \beta_t} x_{t-1}, \beta_t \mathbf{I}\right), \quad (8)$$

where β_t is the noise schedule. The reverse denoising process reconstructs x_0 from $x_T \sim \mathcal{N}(0, I)$ using a Transformer-based denoiser ϵ_θ .

$$p_\theta(x_{t-1} | x_t, c) = \mathcal{N}\left(x_{t-1}; \frac{1}{\sqrt{\alpha_t}} \left(x_t - \frac{\beta_t}{\sqrt{1 - \alpha_t}} \epsilon_\theta(x_t, t, c)\right), \sigma_t^2 \mathbf{I}\right), \quad (9)$$

where c encodes conditional information including UAV state, goal, and static obstacle maps. The denoising network ϵ_θ is trained to predict the injected noise diffusion step. The training loss is defined as

$$\mathcal{L}(\theta) = \mathbb{E}_{x_0, t, \epsilon \sim \mathcal{N}(0, I)} [\|\epsilon_\theta(x_t, t, c) - \epsilon\|_2^2] \quad (10)$$

where $x_t = \sqrt{\alpha_t} x_0 + \sqrt{1 - \alpha_t} \epsilon$ is the conditional context. This objective enforces accurate denoising and enables conditional generation of paths. During online navigation, multiple candidate sub-paths are sampled via denoising diffusion implicit models (DDIM), providing a diverse set of feasible trajectories. The set of K candidate trajectories at time t is represented as

$$\mathcal{C}_t = \{(\tau^{(k)}, \Sigma^{(k)})\}_{k=1}^K, \quad (11)$$

where $\tau^{(k)}$ denotes the k -th trajectory and $\Sigma^{(k)}$ its variance estimates. The variance provides an uncertainty score,

$$U(\tau^{(k)}) = \frac{1}{T} \sum_{i=1}^T \|\Sigma_i^{(k)}\|_2, \quad (12)$$

which is used as part of the reward shaping in the subsequent decision process. Collision risk for a candidate τ is approximated via Monte Carlo sampling.

$$p_{\text{col}}(\tau) = \frac{1}{S} \sum_{s=1}^S \mathbb{1}_{\{\tau^{(s)} \cap O \neq \emptyset\}}. \quad (13)$$

At each re-planning step, the diffusion process yields a diverse set of candidate sub-paths, each annotated with an uncertainty score and an estimated collision probability. While such diversity enhances robustness to environmental variability, it necessitates a principled selection mechanism.

C. Reinforcement Learning-Based Path Selection

While the diffusion model provides diverse sub-path candidates, effective online decision making requires selecting the most suitable option under uncertain conditions. To model this process, the UAV trajectory planning problem is formulated as a Markov Decision Process (MDP),

$$\mathcal{M} = \{\mathcal{S}, \mathcal{A}, \mathcal{P}, r, \gamma\}, \quad (14)$$

where $\mathcal{S}$ is the state space, $\mathcal{A}$ the action space, $\mathcal{P}$ the transition dynamics, r the reward function, and γ the discount factor. The state $s_t \in \mathcal{S}$ at time t is defined as

$$s_t = [p_t, v_t, g - p_t, E_t, L, C, D_{\min}, h_{\text{start}}, h_{\text{end}}, U], \quad (15)$$

where p_t and v_t are position and velocity, $g - p_t$ is the relative goal position, E_t is the residual energy, L is path length, C is trajectory curvature, $D_{\min}$ is the minimum clearance to obstacles, h_{start} and h_{end} are the altitudes of the sub-path endpoints, and U is the uncertainty score estimated from diffusion variance. The action space is discrete, corresponding to candidate sub-path selection:

$$a_t \in \mathcal{A} = \{1, 2, \dots, K\}, \quad (16)$$

where K is the number of diffusion-generated candidate sub-paths.

The reward function is formulated as a weighted combination of interpretable components that balance safety, efficiency, and stability:

$$r_t = w_g r_g(t) + w_c r_c(t) + w_p r_p(t) + w_d r_d(t) + w_e r_e(t) + w_u r_u(t) + w_v r_v(t) + w_l r_l(t), \quad (17)$$

where $r_g(t)$ and $r_c(t)$ are terminal rewards for goal reaching and collision, $r_p(t)$ and $r_d(t)$ represent goal-directed progress and residual distance, $r_e(t)$ captures energy consumption, $r_u(t)$ penalizes uncertainty, $r_v(t)$ enforces trajectory smoothness, and $r_l(t)$ encourages maintaining clearance from obstacles. The weights $\{w_g, w_c, w_p, w_d, w_e, w_u, w_v, w_l\}$ control the trade-offs among competing objectives.

The Q-function is then defined as

$$Q(s, a) = \mathbb{E} \left[\sum_{t=0}^{\infty} \gamma^t r(s_t, a_t) \mid s_0 = s, a_0 = a \right], \quad (18)$$

and the sub-path to execute is selected by

$$a^* = \arg \max_{a \in \mathcal{A}} Q(s, a). \quad (19)$$

This Q-function search is not a full reinforcement learning algorithm but rather an online selection mechanism inspired by RL concepts, enabling principled evaluation and selection among diffusion-generated candidates. By incorporating safety, energy efficiency, smoothness, and uncertainty into the reward formulation, the UAV achieves low-latency and reliable re-planning in cluttered mobile access environments.

IV. PERFORMANCE EVALUATION

A. Experimental Setup

Training data are generated using sampling-based motion planning in synthetic 3D environments populated with randomly placed static obstacles. Each scenario specifies a start position, a goal position, and obstacles modeled as axis-aligned bounding boxes and spheres. A large-scale dataset with diverse configurations is prepared to improve generalization. The trajectory generator is trained with a conditional denoising diffusion implicit model. Training uses 50 epochs, batch size 64, horizon $T = 64$, and $N = 1000$ diffusion steps with v -prediction. Conditions include the start-goal pair and obstacle maps. At test time, a receding-horizon planning scheme is adopted. Full trajectories are divided into fixed-length sub-segments (length 32, stride 16). During re-planning, only the affected sub-segment is resampled, reducing latency while preserving global consistency. For each sub-path, $K = 12$ candidate trajectories are sampled and evaluated by a learned Q-function using features such as path length, curvature, minimum clearance, energy consumption, and diffusion variance as an uncertainty signal. Reward weights $\{w_g, w_c, w_p, w_d, w_e, w_u, w_v, w_l\}$ are tuned to balance safety, efficiency, and smoothness. Performance is measured over multiple test episodes with random start and goal positions.

Compared methods. The following planners are compared under an identical evaluation protocol with the same obstacle maps, start-goal pairs, re-planning horizon, and energy model, ensuring that performance differences stem only from the training regime and selection policy.

- **Proposed:** Segmented diffusion trained on sub-paths with receding-horizon sampling and RL-based Q-function selection.
- **Non-Segmented:** The same diffusion architecture trained end-to-end on full trajectories without sub-path segmentation, using the same planning and RL algorithm interface at test time.
- **Greedy:** A lightweight baseline that, at each step, moves the UAV along the feasible direction with the largest projected progress toward the goal while avoiding immediate collisions. This heuristic is fast but lacks long-horizon reasoning, which can yield suboptimal routes in clutter.

B. Results and Analysis

Fig. 4 shows representative trajectories sampled by the diffusion model given obstacle maps, revealing diverse navigation modes such as direct narrow corridors and longer safer detours. This diversity supplies a rich candidate set for RL-based selection. Table I reports a quantitative comparison over 10 test scenarios using path length, total energy, curvature, accumulated acceleration energy, minimum clearance, and collision rate. The Q-based selection favors safer and smoother trajectories overall.

Fig. 2 compares mean per-timestep energy over elapsed time. The Proposed segmented diffusion with RL is lowest, the Non-Segmented baseline is highest, and Greedy

TABLE I: Comparison of Trajectory Performance Metrics between Proposed, Non-Segmented and Greedy Baseline

Metric	Proposed	Non-Segmented	Greedy
Path length	100.14	120.51	98.71
Energy consumption	2837.45	3522.35	3046.03
Curvature	1.90	3.04	2.12
Acceleration energy cost	4.03	20.99	4.69
Clearance	1.24	2.03	0.74

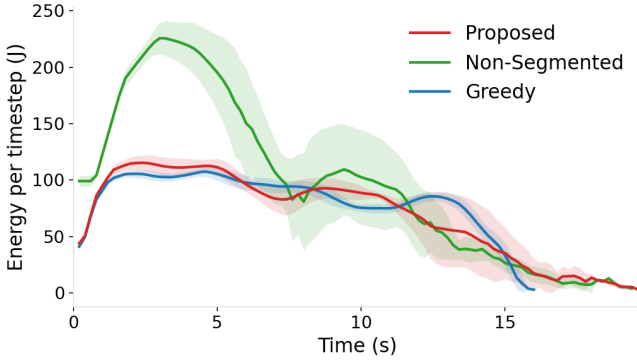


Fig. 2: Mean per timestep energy over elapsed time for Proposed, Non-Segmented, and Greedy. Each step's energy is the sum of propulsion power, affected by forward speed and turning/curvature, and onboard compute power set by the workload level.

lies in between. Table I shows total energy of 2837.45J for Proposed, 3046.03J for Greedy, and 3522.35J for Non-Segmented. Normalized by path length, energy per meter further favors Proposed 28.33J/m over Greedy 30.86J/m and Non-Segmented 29.23J/m. Fig. 3 shows mean curvature per timestep. Proposed consistently lowers instantaneous curvature relative to both baselines. In Table I, Proposed achieves the lowest total curvature 1.90 and low acceleration energy 4.03 while keeping path length close to Greedy, 100.14m and 98.71m respectively, and much shorter than Non-Segmented 120.51m. For clearance, Proposed provides higher margins than Greedy, 1.24m and 0.74m respectively, and lower than Non-Segmented 2.03m.

Qualitative examples in Fig. 4 reflect these trends. Greedy tends to be shorter with sharper turns and smaller obstacle margins. Non-Segmented often over-extends path length. Proposed balances length with smoother turns and improved energy efficiency. Overall, the segmented diffusion and Q-selection framework improves energy efficiency and smoothness while keeping path length competitive.

V. CONCLUSIONS AND FUTURE WORK

This paper presented a framework for UAV trajectory planning in mobile access settings that combines diffusion-based sub-path generation with an RL-based selection module. The diffusion model produces diverse candidate trajectories conditioned on start-goal states and obstacle maps, capturing multimodality and uncertainty. The RL-based selector evaluates

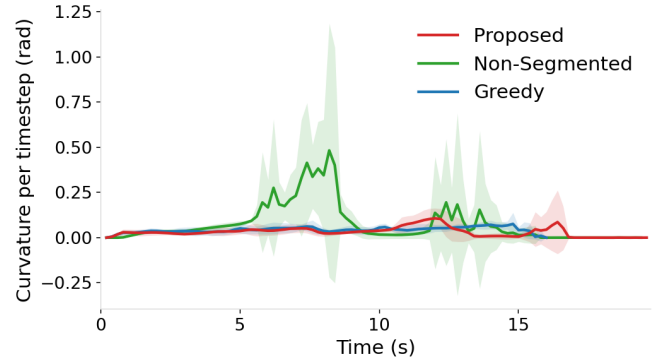


Fig. 3: Mean curvature per timestep over elapsed time for Proposed, Non-Segmented, and Greedy. Curvature at each step is the magnitude of the direction change between consecutive path segments along the 3D trajectory. Higher spikes indicate sharper turns.

candidates online to balance safety, efficiency, and smoothness. Experiments in test scenarios show that the Proposed method achieves the lowest mean per-timestep energy and the lowest total energy, while also reducing curvature and acceleration effort. Path length is comparable to Greedy and shorter than Non-Segmented, and clearance is higher than that of Greedy but lower than that of Non-Segmented. Normalizing by distance, energy per meter further favors the Proposed method. These findings indicate that integrating generative modeling with RL-based selection yields trajectories that are more energy-efficient, smoother, and safer without sacrificing path length.

Several directions remain open for future work. First, extending the framework to account for dynamic obstacles and moving agents would enhance applicability in highly interactive environments. Second, incorporating richer energy models that consider aerodynamics, payload variations, and communication workloads can further improve realism. Third, online adaptation mechanisms such as continual learning or model predictive updates may reduce sensitivity to distributional shifts in unseen environments. Finally, integrating the framework with multi-UAV coordination strategies could enable scalable aerial networks for mobile access and disaster-response scenarios.

ACKNOWLEDGMENT

This research was supported by the MSIT (Ministry of Science and ICT), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2024-RS-2024-00436887) supervised by the IITP (Institute for Information & Communications Technology Planning & Evaluation).

REFERENCES

- [1] C. Yan, X. Xiang, and C. Wang, "Towards real-time path planning through deep reinforcement learning for a uav in dynamic environments," *Journal of Intelligent & Robotic Systems*, vol. 98, no. 2, pp. 297–309, 2020.

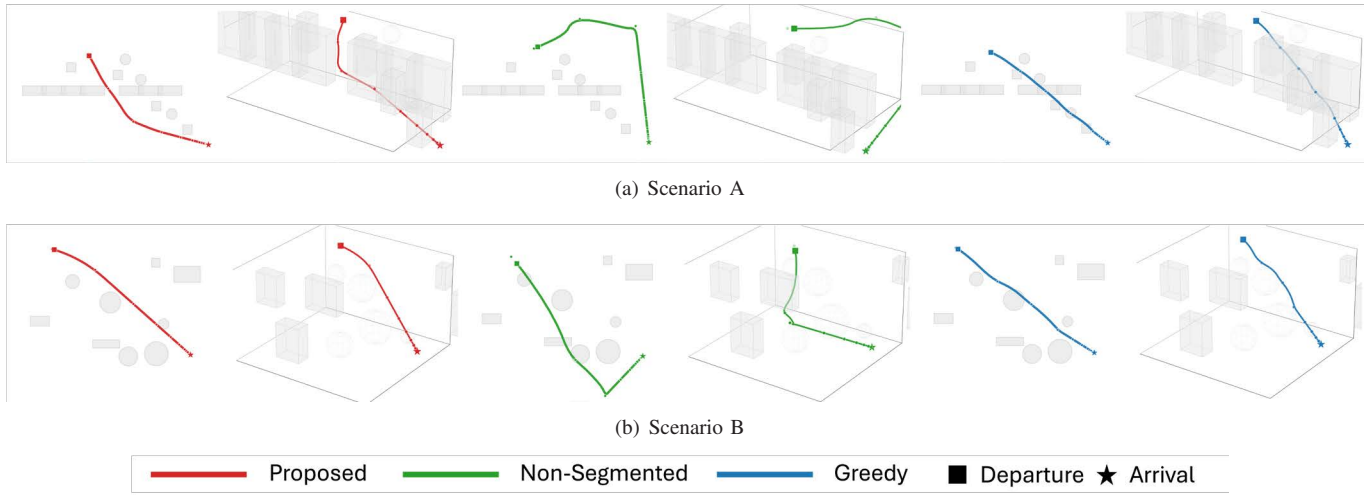


Fig. 4: Top-view and 3D-view of representative trajectories generated by each algorithm across different obstacle scenarios.

- [2] R. Xie, Z. Meng, L. Wang, H. Li, K. Wang, and Z. Wu, "Unmanned aerial vehicle path planning algorithm based on deep reinforcement learning in large-scale and dynamic environments," *IEEE Access*, vol. 9, pp. 24 884–24 900, 2021.
- [3] G. Tong, N. Jiang, L. Biye, Z. Xi, W. Ya, and D. Wenbo, "UAV navigation in high dynamic environments: A deep reinforcement learning approach," *Chinese Journal of Aeronautics*, vol. 34, no. 2, pp. 479–489, 2021.
- [4] C. Wang, J. Wang, J. Wang, and X. Zhang, "Deep-reinforcement-learning-based autonomous uav navigation with sparse rewards," *IEEE Internet of Things Journal*, vol. 7, no. 7, pp. 6180–6190, 2020.
- [5] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski et al., "Human-level control through deep reinforcement learning," *nature*, vol. 518, no. 7540, pp. 529–533, 2015.
- [6] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal policy optimization algorithms," *arXiv preprint arXiv:1707.06347*, 2017.
- [7] R. B. Grando, J. C. de Jesus, and P. L. Drews-Jr, "Deep reinforcement learning for mapless navigation of unmanned aerial vehicles," in *2020 Latin American Robotics Symposium (LARS), 2020 Brazilian Symposium on Robotics (SBR) and 2020 Workshop on Robotics in Education (WRE)*. IEEE, 2020, pp. 1–6.
- [8] C. Zhang, L. Ma, and A. Schmitz, "A sample efficient model-based deep reinforcement learning algorithm with experience replay for robot manipulation," *International Journal of Intelligent Robotics and Applications*, vol. 4, no. 2, pp. 217–228, 2020.
- [9] K. Wang, J. Ma, K. L. Man, K. Huang, and X. Huang, "Sim-to-real transfer with domain randomization for maximum power point estimation of photovoltaic systems," in *2021 IEEE International Conference on Environment and Electrical Engineering and 2021 IEEE Industrial and Commercial Power Systems Europe (EEEIC/I&CPS Europe)*. IEEE, 2021, pp. 1–4.
- [10] S. Huang, Z. Wang, P. Li, B. Jia, T. Liu, Y. Zhu, W. Liang, and S.-C. Zhu, "Diffusion-based generation, optimization, and planning in 3d scenes," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023, pp. 16 750–16 761.
- [11] F. Zhang and G.-H. Yang, "Adaptive safety-certified reinforcement learning for constrained optimal control of autonomous robots with uncertainties," *IEEE Internet of Things Journal*, 2025.
- [12] J. Ho, A. Jain, and P. Abbeel, "Denoising diffusion probabilistic models," *Advances in neural information processing systems*, vol. 33, pp. 6840–6851, 2020.
- [13] Y. Zhu, Y. Ye, S. Zhang, X. Zhao, and J. Yu, "Difftraj: Generating gps trajectory with diffusion probabilistic model," *Advances in Neural Information Processing Systems*, vol. 36, pp. 65 168–65 188, 2023.
- [14] J. Carvalho, A. T. Le, M. Baierl, D. Koert, and J. Peters, "Motion planning diffusion: Learning and planning of robot motions with diffusion models," in *2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2023, pp. 1916–1923.
- [15] W. Mao, C. Xu, Q. Zhu, S. Chen, and Y. Wang, "Leapfrog diffusion model for stochastic trajectory prediction," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2023, pp. 5517–5526.
- [16] I. Bae, Y.-J. Park, and H.-G. Jeon, "Singulartrajectory: Universal trajectory predictor using diffusion model," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024, pp. 17 890–17 901.
- [17] T. Yun, S. Yun, J. Lee, and J. Park, "Guided trajectory generation with diffusion models for offline model-based optimization," *Advances in Neural Information Processing Systems*, vol. 37, pp. 83 847–83 876, 2024.