

GUIDE: A Diffusion-Based Autonomous Robot Exploration Framework Using Global Graph Inference

Zijun Che^{1*}, Yinghong Zhang^{1*}, Shengyi Liang¹, Boyu Zhou², Jun Ma¹, Jinni Zhou¹

Abstract—Autonomous exploration in structured and complex indoor environments remains a challenging task, as existing methods often struggle to appropriately model unobserved space and plan globally efficient paths. To address these limitations, we propose GUIDE, a novel exploration framework that synergistically combines global graph inference with diffusion-based decision-making. We introduce a region-evaluation global graph representation that integrates both observed environmental data and predictions of unexplored areas, enhanced by a region-level evaluation mechanism to prioritize reliable structural inferences while discounting uncertain predictions. Building upon this enriched representation, a diffusion policy network generates stable, foresighted action sequences with significantly reduced denoising steps. Extensive simulations and real-world deployments demonstrate that GUIDE consistently outperforms state-of-the-art methods, achieving up to 18.3% faster coverage completion and a 34.9% reduction in redundant movements.

I. INTRODUCTION

Autonomous exploration remains a cornerstone of modern robotics research, with pivotal applications in scenarios where unknown environment coverage is critical: environmental monitoring, warehouse logistics, and search-and-rescue operations [1], [2]. A defining challenge in these tasks is efficiently covering all reachable areas under stringent constraints—limited time, finite energy, and constrained computational resources. Despite decades of progress, existing exploration strategies still struggle to appropriately model unobserved space and plan globally efficient paths.

Current exploration methodologies can be broadly categorized into model-based and learning-based approaches, each exhibiting fundamental limitations in addressing the global coverage challenge. Early model-based techniques, including frontier-based methods [3], [4] and sampling-based exploration strategies [5], rely exclusively on observed map information to determine exploration directions. While these approaches demonstrate reasonable performance in structured environments, their inherent myopia toward unobserved areas frequently results in redundant revisits, inefficient path planning, and suboptimal coverage—particularly in environments

This work was supported by the Guangdong Provincial Educational Science Planning Project (Grant No. 2025GXJK0629), Guangzhou Higher Education Teaching Research and Reform Project (Grant No. 2024YBJG094), and Guangzhou Municipal Educational Science Planning Project (Grant No. 202316577). (Zijun Che and Yinghong Zhang contributed equally to this work.) (Corresponding Author: Jinni Zhou.)

¹The Hong Kong University of Science and Technology (Guangzhou), Guangzhou, China. {zche624, yzhang780, sliang318}@connect.hkust-gz.edu.cn, jun.ma@ust.hk, eejinni@hkust-gz.edu.cn

²Southern University of Science and Technology, Shenzhen, China. zhouby@sustech.edu.cn

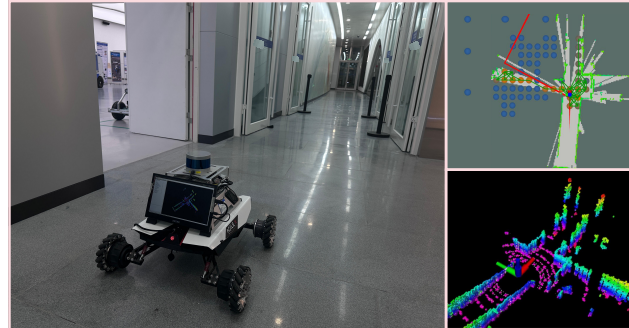


Fig. 1. Exploration in a structured indoor environment with a mobile robot using GUIDE. Left: the robot's actual position in the real world. Top right: the robot leverages the global graph representation to generate exploration trajectories, where blue nodes indicate predictions of unexplored areas and the red line denotes the generated action sequence. Bottom right: the 3D occupancy map constructed by the robot during exploration.

with complex topologies. Coverage path-based methods [6], [7] attempt to address this limitation by incorporating explicit coverage objectives; however, their reliance on uniform grid decomposition implicitly assumes environmental regularity, leading to performance degradation in spaces with irregular layouts or varying structural complexity.

More recently, learning-based approaches have emerged as promising alternatives, yet they too face significant challenges in achieving comprehensive spatial understanding. The first category employs neural networks to directly map observed environments to exploration actions [8]–[10]. Although these methods improve adaptability to specific environments, they fundamentally operate with limited information—encoding only observed areas while remaining agnostic to the structure of unknown spaces. This inherent constraint severely limits their capacity to achieve globally efficient exploration, often requiring extensive training across diverse environments to achieve moderate performance. The second category explicitly predicts unobserved areas and associated information gain [11], [12]. While conceptually promising, these approaches typically utilize predicted maps only for local planning rather than incorporating them into a comprehensive global planning framework, thereby failing to fully leverage the predictive information for long-horizon path optimization.

These limitations collectively highlight a critical research gap: the absence of a unified framework that effectively integrates predictions of unknown areas with globally optimized exploration planning. Specifically, existing methods

lack mechanisms to (1) construct a comprehensive environmental representation that coherently combines observed information with predictions of unexplored areas, (2) leverage credible predictions to guide exploration decisions, and (3) generate stable, long-horizon trajectories that maximize coverage efficiency while minimizing redundant movements.

To address these challenges, we propose **GUIDE**, a novel exploration framework that synergistically combines global graph inference with diffusion-based decision-making. At its core, GUIDE constructs a region-evaluation global graph representation that integrates both observed environmental data and predictions of unexplored areas. This representation is enhanced through a region-level evaluation mechanism that prioritizes significant regional structural inferences, effectively creating an informative yet compact environmental model that prioritizes credible structural inferences while appropriately discounting uncertain predictions. Building upon this enriched representation, GUIDE employs a diffusion policy network that generates stable, foresighted action sequences with significantly reduced denoising steps compared to conventional approaches—enabling efficient long-horizon planning that effectively balances immediate information gain with comprehensive coverage objectives. The reduced computational overhead ensures real-time responsiveness—a critical advantage for resource-constrained robotic platforms.

We rigorously evaluate GUIDE across diverse simulation environments with varying structural complexities and through real-world deployments on physical robotic platforms. Quantitative results demonstrate consistent improvements over state-of-the-art methods, with our approach achieving up to 18.3% faster coverage completion and a 34.9% reduction in redundant movements across benchmark environments. Qualitative analysis further showcases GUIDE’s superior capability in structural inference and adaptive exploration behavior. The main contributions of this work are threefold:

- 1) We introduce a region-evaluation global graph inference module that constructs a unified environmental representation by integrating observed information with predictions of unexplored areas. It incorporates a novel region-evaluation mechanism that assesses the reliability and decision relevance of predicted areas, enabling robust and reliable exploration planning under uncertainty.

- 2) We develop a diffusion-based decision-making framework that explicitly leverages the global graph representation to generate stable, long-horizon exploration trajectories, significantly reducing the computational burden of conventional diffusion policies while producing foresighted and efficient exploration paths.

- 3) We conduct comprehensive evaluations across multiple simulation environments and real-world scenarios, demonstrating GUIDE’s superior performance in both structural inference precision and exploration efficiency, establishing a new benchmark for autonomous exploration systems.

II. RELATED WORK

Research on autonomous exploration is broadly divided into two paradigms. Model-based methods rely on explicitly designed planning schemes defined over environment representations to guide exploration, while learning-based methods use data-driven algorithms to learn exploration strategies from interaction or training. The following subsections review the developments in these two directions.

A. Model-Based Exploration Methods

Classical exploration starts with frontier-based methods, first proposed by Yamauchi [13], which target the boundary between known and unknown areas. Sampling-based planners such as NBVP [5] adopt a receding-horizon RRT scheme, but their local nature often results in myopic decisions and frequent revisits. Later extensions with history-aware strategies, motion primitives [14], and graph-based formulations [15] improve efficiency and consistency, yet still lack explicit long-horizon planning.

Hierarchical planning methods address this limitation by introducing a global visiting order. FUEL [3] maintains an incremental frontier structure and solves an ATSP tour over frontier clusters to guide global exploration, reducing detours and improving efficiency. TARE [4] coordinates sparse global routing with dense local refinement in a coarse-to-fine tree, preserving scalability and connectivity. Under sensing constraints, recent work improves viewpoint determination to curb revisits and wasted motion [16], [17]. However, these methods still rely exclusively on observed map information, making them inherently myopic toward unobserved areas and prone to inefficient planning and suboptimal coverage.

Coverage-path-based methods address this limitation by incorporating explicit coverage objectives to improve global path consistency and multi-robot coordination [18], [19]. Notably, FALCON [7] exploits cues from unexplored areas to compute a global guidance path over unknown space, controlling local frontier visitation and suppressing backtracking. In the same spirit, dual-layer planners combine global region routing with local viewpoint and trajectory optimization to accelerate exploration while maintaining short and consistent tours [20]. However, their reliance on uniform grid decomposition implicitly assumes environmental regularity, leading to performance degradation in spaces with irregular layouts or varying structural complexity.

B. Learning-Based Exploration Methods

Recent studies have explored learning-based methods to improve exploration under partially observed environments. Early work showed that reinforcement learning can acquire strategies directly from onboard sensing without explicit priors [21], [22], though policies are often myopic. To mitigate this, graph-based methods introduced structural reasoning to manage uncertainty and support coordination [23], [24], while attention mechanisms capture multi-scale dependencies and predict information gain, enabling more consistent long-horizon planning in large-scale environments [8], [9]. Subsequent advances include hierarchical

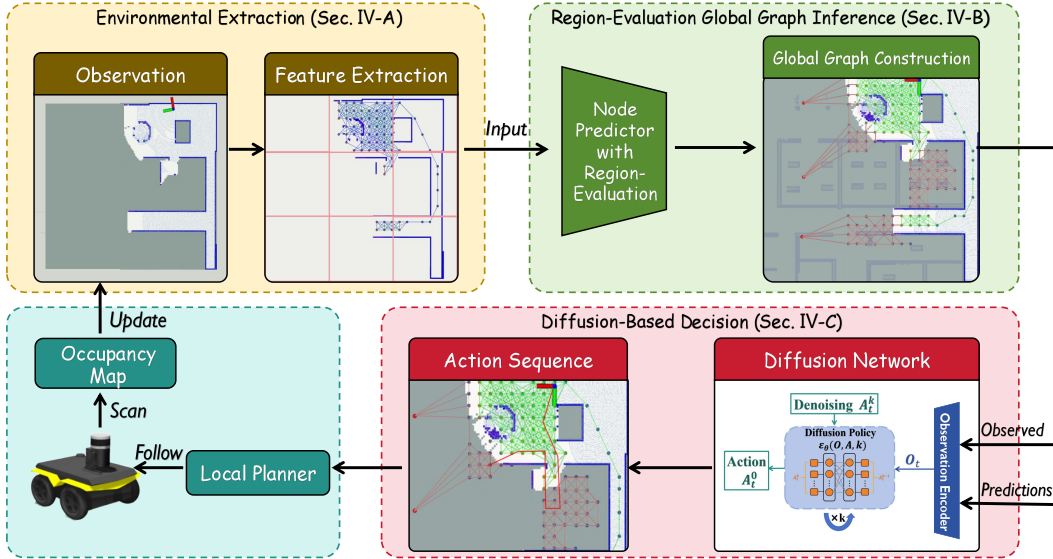


Fig. 2. Overview of the proposed **GUIDE** framework with three modules: (**Sec. IV-A**) Environmental Extraction updates the observation, samples free nodes, and decomposes the whole space; (**Sec. IV-B**) Region-Evaluation Graph Inference leverages observations to predict nodes and filters them for reliability; (**Sec. IV-C**) Diffusion-Based Decision conditions a diffusion policy on the global graph to generate stable long-horizon trajectories.

decision networks [25], dynamic graph methods [26], and hybrid schemes combining frontier heuristics [27]. More recently, diffusion-based policies generate complete trajectories for stable and coherent exploration [10], and multi-agent reinforcement learning with graph attention enables decentralized coordination under constrained sensing [28]. While these methods improve adaptability, they remain constrained to encoding observed regions without modeling unexplored structures, fundamentally limiting long-horizon, globally efficient exploration.

Therefore, some works explicitly predict unobserved areas to provide guidance and mitigate myopic planning. Learned occupancy completion has been shown to improve information-gain estimation [11], [29], while over-reliance can lead to greedy and inefficient coverage [30]. To mitigate this, recent works fine-tune LaMa [31] on the KTH floor plan dataset [32] and use ensembles to model variance, yielding probabilistic information gain [12], [33]. Although these methods enhance inference of unobserved structures, they are mostly applied to local planning and cannot fully exploit predictions for long-horizon optimization, while dense map representations further incur high computational cost.

III. SYSTEM OVERVIEW

We consider the task of robot exploration in a previously unknown environment. The robot incrementally constructs a 2D occupancy map O_t from onboard sensor data, labeling each grid cell as free, occupied, or unknown. The planner is required to compute safe and feasible trajectories that expand O_t over time until all reachable unexplored areas are revealed, thereby yielding a complete occupancy map of the accessible space while minimizing exploration cost, measured in distance traveled or time.

As illustrated in Fig. 2, the proposed **GUIDE** consists of three modules. The **Environmental Extraction module** (Sec. IV-A) updates the occupancy map O_t , samples free nodes, and decomposes the entire space into fixed-size regions. The **Region-Evaluation Global Graph Inference module** (Sec. IV-B) leverages observed information to predict nodes in unexplored areas, employing a region-evaluation mechanism to assess region reliability and filter predictions, thereby constructing a credible global graph representation. The **Diffusion-Based Decision module** (Sec. IV-C) conditions a diffusion policy directly on this global graph to generate foresighted action sequences with fewer denoising steps, producing stable long-horizon trajectories. The robot iteratively executes trajectories and updates O_t with new sensor data, enabling exploration guided jointly by observations and region-evaluated inferences for improved efficiency and consistency.

IV. METHODOLOGY

A. Environmental Extraction

Graph representations abstract the environment by modeling free positions as nodes and collision-free connections as edges. Unlike methods restricted to observed space, we construct a unified global graph that also includes placeholders for unexplored areas. Formally, the environment is represented as $G = (V, E)$, where each node $v \in V$ encodes its lattice coordinate at resolution d_n , a type label (known/unknown), a visitation flag, neighborhood indices, and a utility score, which quantifies the node's potential for exploration. From the occupancy map O_t , we sample free cells to form the known free node set:

$$\mathcal{V}_f = \{ (x, y) \mid O_t(x, y) = \text{free at resolution } d_n \}, \quad (1)$$

while the occupied cells form the obstacle node set \mathcal{V}_o , which are used only as structural references.

To cover the entire environment, we decompose the space into fixed-size regions:

$$\mathcal{G} = \{g_1, g_2, \dots, g_N\}. \quad (2)$$

Each region is assigned one of the following labels:

Explored: if it contains observed free or occupied cells without frontiers.

Boundary: if it contains frontier cells.

Unobserved: if no cells are observed. For every unobserved region, its centroid is added to the grid node set \mathcal{V}_g as a placeholder.

B. Region-Evaluation Global Graph Inference

1) Global Node Predictor: As outlined in Algorithm. 1, the predictor estimates potential free nodes in unknown areas. First, the known free and obstacle nodes are rasterized onto a lattice with resolution d_n to form a tri-valued image I (255 free, 0 occupied, 128 unknown). The unknown pixels serve as a mask, which is completed by the LaMa inpainting model [31], fine-tuned on our node dataset, yielding \hat{I} that predicts potential free nodes.

Candidate nodes are extracted from \hat{I} by binarization to identify free nodes, then refined by removing isolated nodes through component analysis, yielding the final set \mathcal{V}_p .

Algorithm 1: Global Node Predictor

Input: Known free nodes \mathcal{V}_f , obstacle nodes \mathcal{V}_o

Output: Predicted free nodes \mathcal{V}_p

- 1 $I \leftarrow \mathbf{Rasterize}(\mathcal{V}_f, \mathcal{V}_o, d_n)$;
 - 2 $\hat{I} \leftarrow \mathbf{LaMa}(I, \text{mask} = 128)$;
 - 3 $P \leftarrow \mathbf{Binarize}(\hat{I}, \tau)$;
 - 4 $\mathcal{V}_p \leftarrow \mathbf{Nodes}(P)$;
 - 5 **for** node $v \in \mathcal{V}_p$ **do**
 - 6 **if** $\mathbf{IsIsolated}(v)$ **then**
 - 7 remove v
 - 8 **return** \mathcal{V}_p ;
-

2) Region-Evaluation Mechanism: Based on the grid decomposition in Sec. IV-A, we design a region-evaluation mechanism to assess regions and filter the predicted free nodes \mathcal{V}_p for credible predictions (Fig. 3). Regions close to the robot and supported by frontiers are regarded as reliable for guiding exploration, while distant regions with limited information are less useful and represented sparsely.

Each region g is assigned a score based on its centroid:

$$r(g) = \frac{\omega_f}{d_f} + \frac{\omega_r}{d_r}, \quad (3)$$

where d_f is the average distance from the centroid to the Z nearest frontiers and d_r is the distance from the centroid to the robot. This scoring is computed only for regions labeled as *Boundary* or *Unobserved*.

We establish a region-evaluated unknown node set \mathcal{V}_u : for regions with higher scores, their predicted nodes from

\mathcal{V}_p are retained. In contrast, predictions in low-score regions are discarded. Specifically, in low-score *Unobserved* regions, the region centroid v_g is added to \mathcal{V}_u when the predicted nodes within the region are sufficiently dense and discarded otherwise, to ensure a compact global representation. The final node set is therefore

$$V = \mathcal{V}_f \cup \mathcal{V}_u, \quad (4)$$

combining known free nodes with region-evaluated unknown nodes.

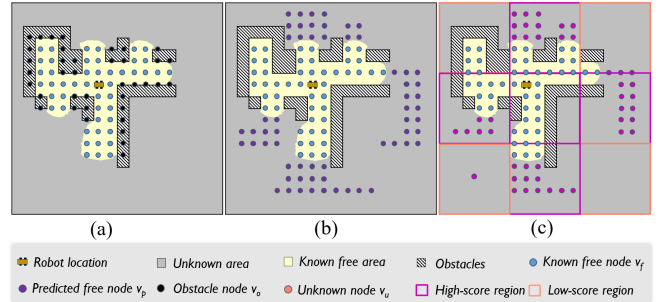


Fig. 3. Illustration of the process of node prediction with region-evaluation. (a) Observed nodes (free and obstacles) as input to the predictor. (b) Predicted free nodes \mathcal{V}_p from the Global Node Predictor. (c) Final unknown-node set \mathcal{V}_u after region-evaluation, where high-score regions retain predictions while others are replaced by centroids or discarded.

3) Global Graph Construction: We update node utilities within a utility update range $r = \frac{R_{\text{sensor}}}{2}$ to measure exploration potential. As illustrated in Fig. 4(a), frontier counts alone can be ambiguous—nodes with identical frontier numbers may correspond to vastly different unexplored extents. To resolve this, we incorporate predicted unknown nodes into the utility definition:

$$u(v) = \begin{cases} \min \left\{ F_m(v), F(v) \left(1 + \frac{N_u(v)}{N_{um}(v)} \right) \right\}, & v \in \mathcal{V}_f, \\ 0, & v \in \mathcal{V}_u, \end{cases} \quad (5)$$

where $F(v)$ is the number of frontier cells obtained by ray casting within the update range r , $N_u(v)$ is the number of predicted unknown nodes in this range, $N_{um}(v)$ is the maximum number of nodes that could occupy its unknown area at resolution d_n , and $F_m(v)$ is the upper bound on frontier count along the boundary of r . Thus, denser predicted nodes raise the utility while keeping it bounded by sensor visibility.

Edges are then constructed as shown in Fig. 4(b). Free nodes connect to their collision-free neighbors, while unknown nodes are linked to nearby known or unknown nodes within the adjacent grid range, provided the path does not intersect occupied cells, even if it traverses unobserved space.

Combining observed free nodes with region-evaluated unknown nodes and their connecting edges, we obtain the global graph $G = (\hat{V}, E)$, which stores dynamically updated utilities and encodes feasible and hypothesized connectivity for exploration.

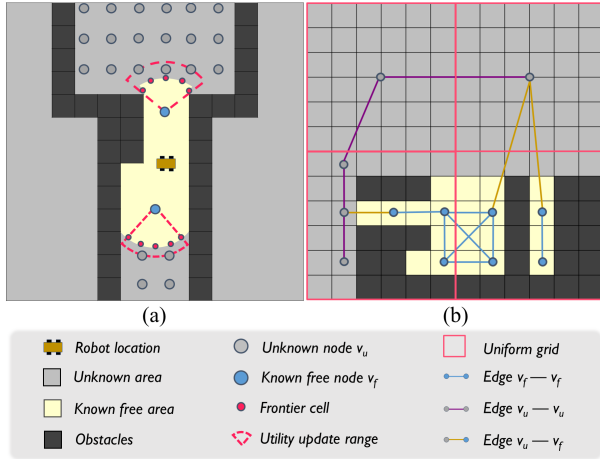


Fig. 4. Illustration of (a) utility update, where nodes with the same frontier count obtain different utilities depending on predicted nodes within the utility update range; and (b) edge construction, showing connections between different node types.

C. Diffusion-Based Decision Guided by Global Graph

We integrate the global graph from Sec. IV-B into the diffusion-policy framework. The graph G , which incorporates both observed information and predicted unexplored areas, is encoded by a Transformer encoder into a compact latent representation z_t . This graph embedding is then integrated into the observation to form an augmented observation O_t , which serves as the conditional input to the diffusion model.

At each timestep t , the action sequence $A_t^0 = [a_t, a_{t+1} \dots, a_{t+T_p}]$ represents planned exploration moves within the prediction horizon T_p . Following the diffusion policy framework, an initial noisy action sequence A_t^K is sampled from a Gaussian distribution and progressively denoised over K iterations by a noise-prediction network ϵ_θ , yielding the final action sequence A_t^0 . The denoising step is formulated as:

$$A_t^{k-1} = \alpha_k (A_t^k - \gamma_k \epsilon_\theta(O_t, z_t, A_t^k, k)) + \mathcal{N}(0, \sigma_k^2 I), \quad (6)$$

where $\alpha_k, \gamma_k, \sigma_k$ are schedule-dependent parameters.

Only the first T_a steps are executed in a receding-horizon manner, ensuring reactivity to new information.

The noise-prediction network ϵ_θ is trained with expert demonstrations obtained from ground-truth trajectories. The training objective minimizes the mean squared error (MSE) between the sampled Gaussian noise ϵ_k and the corresponding model prediction:

$$\mathcal{L} = \mathbb{E} [\|\epsilon_k - \epsilon_\theta(O_t, z_t, A_t^0 + \epsilon_k, k)\|^2]. \quad (7)$$

By leveraging the enriched global graph G in the diffusion process, the policy generates stable and foresighted action sequence A_t^0 while substantially reducing the number of denoising steps K compared to conventional diffusion policies.

Subsequently, the next waypoint is extracted from the generated action sequence A_t^0 and then directly provided to

the local planner [4], which generates executable, collision-free trajectories.

D. Training Details

We adopt a two-stage (hierarchical) training scheme, where the first stage fine-tunes an inpainting network for node prediction and the second stage trains diffusion planner.

Stage I: The model is fine-tuned on 10,000 (observed, ground-truth) pairs collected from robot rollouts in randomly generated maze environments with sizes ranging from 100 m \times 100 m to 180 m \times 180 m. At multiple exploration steps, free and obstacle nodes are rasterized with a resolution of $d_n = 4$ m, unknown masks are derived from observed information, and inputs are padded to sizes 32 \times 32, 40 \times 40, or 48 \times 48.

Stage II: We keep the fine-tuned node prediction model fixed and train the diffusion planner conditioned on the global graph. We train the diffusion policy on 4,000 expert trajectories generated by a ground-truth-map coverage planner, where shortest coverage paths are computed via a Traveling Salesman Problem (TSP). The diffusion network plans with horizon $T_p = 8$ given an observation horizon $T_o = 2$, executes within action horizon $T_a = 2$, and employs the Square Cosine Noise Scheduler [34] with $K = 30$ denoising steps.

Both stages are trained on a single RTX 4090, taking about 13 and 40 hours, respectively.

V. EXPERIMENTS

A. Validation in Maze Datasets

All experiments are conducted on a workstation equipped with an Intel i9-14900K CPU. We first evaluate the structural inference capability of our method, followed by comparisons with representative exploration baselines.

To evaluate the structural inference precision of our Global Graph inference method, we conduct tests in 100 unseen maze environments of size 100 m \times 100 m. Precision is evaluated every 15 exploration steps as the proportion of region-evaluated unknown nodes \mathcal{V}_u that fall within the ground-truth free space:

$$\text{Pre}(k) = \frac{|\mathcal{V}_u^k \cap \mathcal{M}_{free}|}{|\mathcal{V}_u^k|}. \quad (8)$$

As shown in Table I, our method can reliably predict potential free node locations. The high efficiency of the inference process is attributed to its lightweight representation, with each prediction completed in under 10 ms.

TABLE I
AVERAGE STRUCTURAL INFERENCE PRECISION AND PREDICTION TIME OVER 100 MAZES.

Steps (k)	15	30	45	60
Precision (%)	75.92	79.73	79.24	84.61
Time (ms)	5.42	5.58	6.29	5.71

To further evaluate the exploration efficiency of our method, we compare it against several representative exploration planners, including NBVP [5], the DRL planner [9],

MapEx [12], the DARE planner [10], and the TARE local planner [4]. For reference, we also provide an approximate optimal solution obtained from ground-truth coverage.

TABLE II
COMPARISON WITH BASELINE PLANNERS

	NBVP	DRL	MapEx	DARE	TARE	Ours	Optimal
Distance (m)	640(\pm 105)	575(\pm 80)	571(\pm 82)	565(\pm 71)	560(\pm 65)	545(\pm 65)	501(\pm 60)
Gap to Optimal (%)	27.7%	14.8%	14.0%	12.8%	11.8%	8.8%	0%

As shown in Table II, we compare the average path lengths required by different methods to complete exploration. Compared with baseline methods, our approach achieves significantly shorter exploration paths.

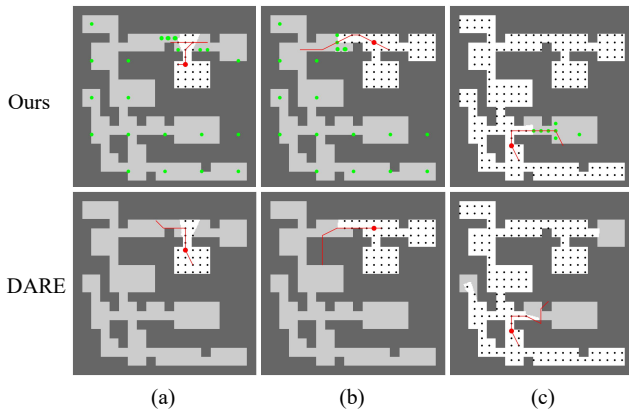


Fig. 5. Action sequence on the same map at approximately the same locations (a)–(c). In each pair, the bottom shows diffusion based only on observed information and the top our method, with red indicating the generated action sequence and green the region-evaluated unknown nodes. By leveraging the global graph representation, our method produces more reasonable trajectories.

In particular, to highlight the contribution of the global graph representation to the diffusion model, we provide a comparison in Fig. 5: unlike the diffusion model which relies on observed information, incorporating the global graph representation enables our approach to generate foresighted and efficient exploration paths across diverse scenarios.

B. Ablation Experiments

To examine the contribution of each major component in our framework, we conduct ablation studies focusing on the **Region-Evaluation Mechanism** and the **Diffusion-Based Decision framework**. To this end, we design the following variants for comparison:

(i) **Without Region-Evaluation Mechanism:** This variant makes decisions based on all predictions of unexplored areas without applying the region-evaluation mechanism.

(ii) **Without Diffusion:** This variant samples nodes from the global graph representation, and then generates exploration paths via TSP.

As shown in Table III, both variants result in noticeable performance drops compared to our full method, confirming

that both the region-evaluation mechanism and diffusion-based decision-making framework are crucial for improving exploration efficiency.

TABLE III
ABLATION EXPERIMENT

Method	Distance (m)	Gap to Optimal (%)
Ours	545(\pm65)	8.8%
Ours(w/o Region-Evaluation)	561(\pm 70)	11.9%
Ours(w/o Diffusion)	580(\pm 79)	15.7%
Optimal	501(\pm 60)	0%

C. Validation in Gazebo Environments

To evaluate the exploration efficiency metric of our method in high-fidelity simulation, we conduct experiments in two structured Gazebo environments provided by [35], with map sizes of (64 m \times 85 m) and (92 m \times 74 m). The robot’s maximum velocity is 2 m/s, the map resolution to 0.4 m, the node resolution to 1.6 m, and the sensing range to 16 m.

TABLE IV
RESULTS OF SIMULATIONS IN TWO INDOOR ENVIRONMENTS(5 RUNS EACH)

scene & size(m)	method	Distance (m)	Time (s)	Efficiency (m^3/m)	Efficiency (m^3/s)
Scene 1 64 \times 85	TARE	1064.72	585.53	5.11	9.29
	DRL	741.15	536.77	7.34	10.13
	DARE	681.41	482.82	7.98	11.27
	Ours	537.18	432.11	10.13	12.59
Scene 2 92 \times 74	TARE	1033.72	615.37	6.58	11.06
	DRL	759.88	635.36	8.96	10.72
	DARE	797.38	623.27	8.54	10.93
	Ours	531.37	513.09	12.81	13.27

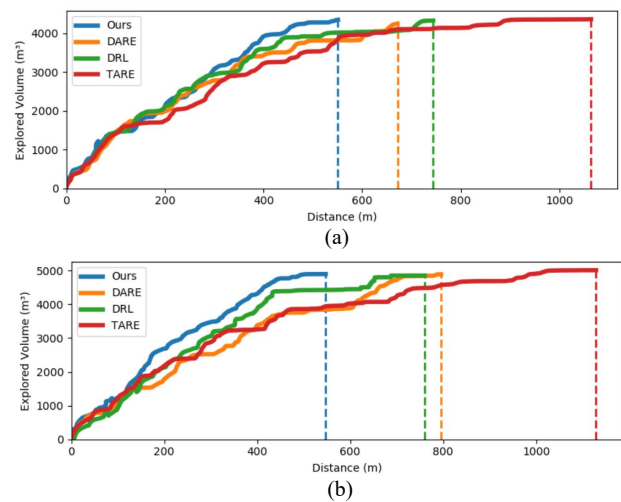


Fig. 6. Comparison of exploration trajectories between our method and baseline planners in the Gazebo simulation: (a) Scene 1 and (b) Scene 2.

Efficiency is defined as the explored free volume divided by the travel distance or the makespan. As shown in Table IV

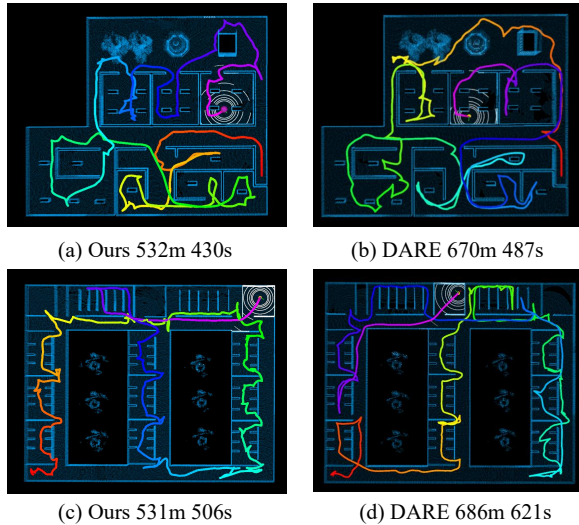


Fig. 7. Comparison of exploration paths in the Gazebo simulation: (a)–(b) Scene 1 and (c)–(d) Scene 2.

and Fig. 6, our method consistently achieves up to 18.3% faster coverage and 34.9% reduction in redundant movements compared with TARE local planner [4], DRL planner [9], and DARE planner [10]. In addition, the comparative exploration paths illustrated in Fig. 7 demonstrate that our method generates foresighted and stable trajectories which maximize coverage efficiency while minimizing redundant movements. In environments with complex intersections and room layouts, the ability to predict the structure of unexplored areas effectively improves overall efficiency.

TABLE V

COMPARISON OF AVERAGE COMPUTATION TIME AND EXPLORATION EFFICIENCY UNDER DIFFERENT DENOISING STEPS.

Method	$K = 10$		$K = 30$	
	Time (s)	Path (m)	Time (s)	Path (m)
Ours	0.177	578.71	0.227	533.18
DARE	0.163	762.47	0.211	747.21
Method	$K = 50$		$K = 100$	
	Time (s)	Path (m)	Time (s)	Path (m)
Ours	0.663	531.71	1.283	530.05
DARE	0.647	723.92	1.271	708.31

Furthermore, to demonstrate GUIDE’s ability to generate stable, foresighted action sequences with fewer denoising steps, we compare it with DARE [10] under varying values of K . The experiments are conducted in the two structured Gazebo environments described above, and the results are reported as the averages across these two scenarios. As shown in Table V, DARE requires up to $K = 100$ steps with the Scheduler [34] to produce stable trajectories, resulting in an average planning time of approximately 1.3 s. GUIDE leverages the enriched global graph representation to achieve

reliable outputs with only $K = 30$ denoising steps. This reduces the planning time to around 0.2 s while also yielding shorter overall exploration paths.

D. Validation in Real World

Finally, we validate our method in a structured real-world environment of size 40 m \times 20 m, as shown in Fig. 8. The experiments are carried out using an Agilex Scout Mini robot equipped with a Velodyne VLP-16 LiDAR and an onboard NVIDIA AGX Xavier for computation. The robot operates with a maximum speed of 0.5 m/s, a map resolution of 0.2 m, a node resolution of 1 m, and a sensor range of 10 m during experiments. As shown in Fig. 1, the robot leverages the global graph representation to generate foresighted exploration trajectories. Results demonstrate that anticipating the structure of unknown areas effectively reduces redundant exploration and significantly improves overall efficiency, confirming the stability and practical applicability of GUIDE in real-world deployments.

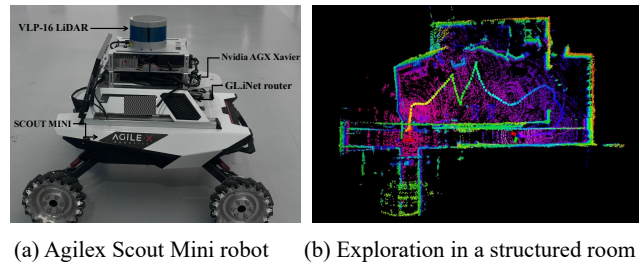


Fig. 8. Real world experiment in a 40 m \times 20 m structured environment.

VI. CONCLUSIONS

In this work, we propose GUIDE, a novel autonomous exploration framework that integrates region-evaluation global graph inference with diffusion-based decision-making. By constructing a unified environmental representation that combines observed information with predictions of unexplored areas, GUIDE generates stable and foresighted exploration trajectories while requiring significantly fewer denoising steps. Extensive simulations and real-world deployments demonstrate GUIDE’s superior performance in structural reasoning, coverage efficiency, and path optimality. Future work will focus on extending GUIDE to 3D environments and multi-robot systems to enable cooperative exploration and enhance scalability in larger and more complex scenarios.

ACKNOWLEDGMENT

We would like to express our deepest gratitude to the Red Bird MPhil Program at The Hong Kong University of Science and Technology (Guangzhou) for providing the generous support, resources, and funding, which have been instrumental in the successful completion of our research.

REFERENCES

- [1] P. Petráček, V. Krátký, M. Petrлік, T. Báča, R. Kratochvíl, and M. Saska, "Large-scale exploration of cave environments by unmanned aerial vehicles," *IEEE Robotics and Automation Letters*, vol. 6, no. 4, pp. 7596–7603, 2021.
- [2] P. De Petris, H. Nguyen, M. Dharmadhikari, M. Kulkarni, N. Khedekar, F. Mascarich, and K. Alexis, "Rmf-owl: A collision-tolerant flying robot for autonomous subterranean exploration," in *2022 International Conference on Unmanned Aircraft Systems (ICUAS)*, pp. 536–543, IEEE, 2022.
- [3] B. Zhou, Y. Zhang, X. Chen, and S. Shen, "FUEL: Fast UAV exploration using incremental frontier structure and hierarchical planning," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 779–786, 2021.
- [4] C. Cao, H. Zhu, H. Choset, and J. Zhang, "TARE: A hierarchical framework for efficiently exploring complex 3D environments," in *Robotics: Science and Systems*, vol. 5, p. 2, 2021.
- [5] A. Bircher, M. Kamel, K. Alexis, H. Oleynikova, and R. Siegwart, "Receding horizon" next-best-view" planner for 3D exploration," in *2016 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 1462–1468, 2016.
- [6] X. Kan, H. Teng, and K. Karydis, "Online exploration and coverage planning in unknown obstacle-cluttered environments," *IEEE Robotics and Automation Letters*, vol. 5, no. 4, pp. 5969–5976, 2020.
- [7] Y. Zhang, X. Chen, C. Feng, B. Zhou, and S. Shen, "FALCON: Fast autonomous aerial exploration using coverage path guidance," *IEEE Transactions on Robotics*, 2024.
- [8] Y. Cao, T. Hou, Y. Wang, X. Yi, and G. Sartoretti, "ARiADNE: A reinforcement learning approach using attention-based deep networks for exploration," in *2023 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 10219–10225, 2023.
- [9] Y. Cao, R. Zhao, Y. Wang, B. Xiang, and G. Sartoretti, "Deep reinforcement learning-based large-scale robot exploration," *IEEE Robotics and Automation Letters*, vol. 9, no. 5, pp. 4631–4638, 2024.
- [10] Y. Cao, J. Lew, J. Liang, J. Cheng, and G. Sartoretti, "DARE: Diffusion policy for autonomous robot exploration," in *2025 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 11987–11993, 2025.
- [11] Y. Tao, Y. Wu, B. Li, F. Cladera, A. Zhou, D. Thakur, and V. Kumar, "SEER: Safe efficient exploration for aerial robots using learning to predict information gain," in *2023 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 1235–1241, 2023.
- [12] C. Ho, S. Kim, B. Moon, A. Parandekar, N. Harutyunyan, C. Wang, K. Sycara, G. Best, and S. Scherer, "MapEx: Indoor structure exploration with probabilistic information gain from global map predictions," in *2025 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 13074–13080, 2025.
- [13] B. Yamauchi, "A frontier-based approach for autonomous exploration," in *Proceedings 1997 IEEE International Symposium on Computational Intelligence in Robotics and Automation CIRA'97. Towards New Computational Principles for Robotics and Automation*, pp. 146–151, 1997.
- [14] M. Dharmadhikari, T. Dang, L. Solanka, J. Loje, H. Nguyen, N. Khedekar, and K. Alexis, "Motion primitives-based path planning for fast and agile exploration using aerial robots," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 179–185, 2020.
- [15] T. Dang, M. Tranzatto, S. Khattak, F. Mascarich, K. Alexis, and M. Hutter, "Graph-based subterranean exploration path planning using aerial and legged robots," *Journal of Field Robotics*, vol. 37, no. 8, pp. 1363–1388, 2020.
- [16] Y. Zhao, L. Yan, H. Xie, J. Dai, and P. Wei, "Autonomous exploration method for fast unknown environment mapping by using uav equipped with limited fov sensor," *IEEE Transactions on Industrial Electronics*, pp. 1–10, 2023.
- [17] J. Yu, H. Shen, J. Xu, and T. Zhang, "ECHO: An efficient heuristic viewpoint determination method on frontier-based autonomous exploration for quadrotors," *IEEE Robotics and Automation Letters*, vol. 8, no. 8, pp. 5047–5054, 2023.
- [18] C. Feng, H. Li, M. Zhang, X. Chen, B. Zhou, and S. Shen, "FC-planner: A skeleton-guided planning framework for fast aerial coverage of complex 3D scenes," in *2024 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 8686–8692, 2024.
- [19] B. Zhou, H. Xu, and S. Shen, "RACER: Rapid collaborative exploration with a decentralized multi-uav system," *IEEE Transactions on Robotics*, vol. 39, no. 3, pp. 1816–1835, 2023.
- [20] Q. Dong, X. Zhang, S. Zhang, Z. Wang, Z. Ma, and H. Xi, "EDEN: Efficient dual-layer exploration planning for fast uav autonomous exploration in large 3-D environments," *arXiv preprint arXiv:2506.05106*, 2025.
- [21] L. Tai and M. Liu, "A robot exploration strategy based on q-learning network," in *2016 IEEE International Conference on Real-Time Computing and Robotics (RACR)*, pp. 57–62, 2016.
- [22] D. Zhu, T. Li, D. Ho, C. Wang, and M. Q.-H. Meng, "Deep reinforcement learning supervised autonomous exploration in office environments," in *2018 IEEE international conference on robotics and automation (ICRA)*, pp. 7548–7555, 2018.
- [23] F. Chen, J. D. Martin, Y. Huang, J. Wang, and B. Englot, "Autonomous exploration under uncertainty via deep reinforcement learning on graphs," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 6140–6147, 2020.
- [24] J. Hu, H. Niu, J. Carrasco, B. Lennox, and F. Arvin, "Voronoi-based multi-robot autonomous exploration in unknown environments via deep reinforcement learning," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 12, pp. 14413–14423, 2020.
- [25] J. Liang, Y. Cao, Y. Ma, H. Zhao, and G. Sartoretti, "HDPlanner: Advancing autonomous deployments in unknown environments through hierarchical decision networks," *IEEE Robotics and Automation Letters*, 2024.
- [26] A. Vashisth, J. Rückin, F. Magistri, C. Stachniss, and M. Popović, "Deep reinforcement learning with dynamic graphs for adaptive informative path planning," *IEEE Robotics and Automation Letters*, vol. 9, no. 9, pp. 7747–7754, 2024.
- [27] S. Nam, T. A. Nguyen, E. Choi, and D. Min, "FH-DRL: Exponential-hyperbolic frontier heuristics with drl for accelerated exploration in unknown environments," *arXiv preprint arXiv:2407.18892*, 2024.
- [28] J. Chiun, S. Zhang, Y. Wang, Y. Cao, and G. Sartoretti, "MARVEL: Multi-agent reinforcement learning for constrained field-of-view multi-robot exploration in large-scale environments," *arXiv preprint arXiv:2502.20217*, 2025.
- [29] R. Shrestha, F.-P. Tian, W. Feng, P. Tan, and R. Vaughan, "Learned map prediction for enhanced mobile robot exploration," in *2019 International Conference on Robotics and Automation (ICRA)*, pp. 1197–1204, 2019.
- [30] L. Ericson, D. Duberg, and P. Jensfelt, "Understanding greediness in map-mobile exploration planning," in *2021 European Conference on Mobile Robots (ECMR)*, pp. 1–7, 2021.
- [31] R. Suvorov, E. Logacheva, A. Mashikhin, A. Remizova, A. Ashukha, A. Silvestrov, N. Kong, H. Goka, K. Park, and V. Lempitsky, "Resolution-robust large mask inpainting with fourier convolutions," in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pp. 2149–2159, 2022.
- [32] A. Aydemir, P. Jensfelt, and J. Folkesson, "What can we learn from 38,000 rooms? reasoning about unexplored space in indoor environments," in *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 4675–4682, 2012.
- [33] S. Baek, B. Moon, S. Kim, M. Cao, C. Ho, S. Scherer, *et al.*, "PIPE Planner: Pathwise information gain with map predictions for indoor robot exploration," *arXiv preprint arXiv:2503.07504*, 2025.
- [34] A. Q. Nichol and P. Dhariwal, "Improved denoising diffusion probabilistic models," in *International Conference on Machine Learning*, pp. 8162–8171, 2021.
- [35] J. Huang, B. Zhou, Z. Fan, Y. Zhu, Y. Jie, L. Li, and H. Cheng, "FAEL: Fast autonomous exploration for large-scale environments with a mobile robot," *IEEE Robotics and Automation Letters*, vol. 8, no. 3, pp. 1667–1674, 2023.