2026 IEEE International Conference on Robotics and Automation (ICRA 2026)
June 1-5, 2026. Vienna, Austria

dPWM: Autonomous Exploration via Diffusion-Based Map Prediction Guided Planning

Zemei Jia[1], Peng Qi[1], Xiaoxiang Liu[2], Zhihao Yao[1], Liang Li[1]

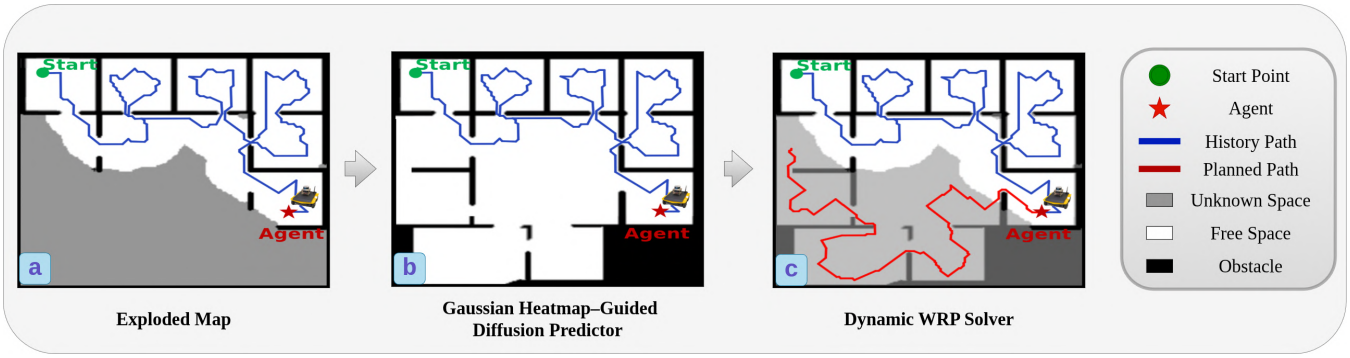


Fig. 1: **An overview of the proposed dPWM-based exploration method.** The method predicts the global structural characteristics **(b)** from the partially explored map **(a)** and employs the improved WRP solver to compute a closed-loop optimal exploration path. The robot is guided to move toward the direction with higher utility **(c)**, thereby achieving efficient exploration.

***Abstract*— Autonomous exploration aims to efficiently map unknown environments, yet utilizing limited environmental information to achieve efficient path planning remains challenging. In this work, we focus on leveraging latent information in partial observations to predict the complete environmental structure, thereby furnishing a proposed path planner with the necessary context to devise a long-term optimal exploration strategy. Most existing prediction approaches extract environment features through convolutional neural networks (CNN) and infer the characteristics of neighboring regions. This information then feeds into a value function that evaluates candidate frontiers and guides the robot's planning. Notwithstanding its advantages over traditional heuristic methods, this paradigm remains inherently constrained by its lack of long-term foresight. To this end, we propose dPWM, a diffusion model-based framework for global map prediction, consisting of two key components. The first employs a DDPM with a variable mask to estimate the probability distribution of unknown regions and thereby predict structural features of the global map. We incorporate Gaussian heatmap positional fields into the denoising process via a cross-attention mechanism to enhance regional awareness. This guides the model to focus on nearby areas that are most valuable for exploration. Once the global predictive map is obtained, the second component refers to a designed Watchman Route Problem (WRP) solver to generate an optimal path from the current exploration state. Extensive evaluations show that dPWM reduces exploration path length by 18.53% on HouseExpo and achieves a 16.37% improvement in cross-domain generalization on Dungeon over SOTA baselines. Real-world experiments further validate its effectiveness in physical environments.**

This work is supported by the National Key R&D Program of China (No. 2024YFB3909903), and the NSFC 62088101 Autonomous Intelligent Unmanned Systems. (*Corresponding author: Liang Li.*)

[1] College of Control Science and Engineering, Zhejiang University, Hangzhou, 310027, China.
[2] Beijing Institute of Control Engineering, Beijing, 100190, China.

I. Introduction

Research on autonomous exploration has gained increasing importance in robotics, which aims to construct maps of the environments with partial and noisy observations while minimizing path cost. Nevertheless, the growing complexity of real-world scenarios (e.g., disaster response [1] and autonomous navigation [2]) imposes increasingly higher demands on decision-making.

Traditional exploration strategies, including frontier-based methods [3]–[5] and reinforcement learning (RL)–based dynamic policies [2], [6], typically rely on evaluating frontier points within the robot's local observation range. While effective in near-term decision-making, such approaches are inherently short-sighted and struggle to support globally optimal exploration planning. To overcome this limitation, we draw inspiration from the Watchman Route Problem (WRP) [7], which aims to compute the shortest path that fully observes a given static environment. Unlike frontier-based strategies, WRP provides a principled way to derive globally optimal paths. Nevertheless, classical WRP solvers [8] assume a fully known and static environment and are inapplicable in dynamic exploration scenarios.

To this end, we propose dPWM, a method that first utilize a diffusion model to predict a complete environmental map and then employs a designed WRP solver to compute exploration paths, as illustrated in Fig. 1. To predict high-quality global maps, most existing works employ neural predictors to infer unexplored regions. However, due to their limited receptive fields and difficulty in modeling long-term dependencies, these predictors often fail to capture reliable structures in distant areas [9]–[13]. As a result, most prediction-based approaches remain confined to local predictions—such as next-view [9] and short-horizon regional prediction [11], [12]—whose effectiveness degrades substantially when ap-

plied to global map prediction. Motivated by advances in image inpainting [14], we introduce a diffusion-model-based map predictor that extrapolates structural information of unseen areas from partial observations. In exploration tasks, nearby regions carry higher exploration value than distant ones. To leverage this property, we incorporate Gaussian heatmaps centered at the robot's position into the cross-attention layers of the reverse diffusion process at specific resolutions. This positional prior enhances local dependencies and significantly improves prediction quality in high-impact regions.

We validate our approach on the real-world indoor dataset HouseExpo [15]. Experimental results demonstrate that our method achieves a 18.53% improvement in path efficiency over the SOTA prediction-based approaches [12], an 20.49% improvement over representative reinforcement learning (RL)-based exploration methods [16], and a 28.00% improvement compared with classical frontier-based exploration schemes [3].

The main contributions of this work are summarized as follows:

- Positional prior–guided diffusion for global map prediction: We introduce a Gaussian heatmap–guided diffusion model that leverages positional weighting and cross-attention to enable agent-centric global map prediction. Extensive experiments on the HouseExpo [15] dataset show that our approach significantly outperforms frontier-based [3]–[5], RL-based [16], and map-prediction [12] baselines in exploration efficiency.
- Dynamic WRP solver for exploration: We design a dynamic WRP solver with exploration masks to compute adaptive and near-optimal exploration paths in large-scale environments.
- Extensive evaluation and real-world validation: We validate the effectiveness of dPWM on the HouseExpo [15] dataset, where it outperforms all baseline methods. On the Dungeon [17] dataset, dPWM demonstrates superior generalization, achieving a 26.48% improvement over the SOTA DRL-based method and a 16.37% gain over the SOTA map-prediction approach. Ablation studies further confirm the contributions of its key components, and real-world deployment experiments highlight its strong potential for practical applications.

II. Related Work

Exploration has long been a fundamental problem in robotics, and decades of research have led to a variety of representative approaches. Broadly, these methods can be classified into observation-driven and prediction-driven types.

A. Observation-Driven Exploration

Such strategies guide exploration by analyzing frontiers in the observed map. A typical representative is heuristic-based exploration [3]–[5], which employs explicitly designed value functions to evaluate candidate actions. Such approaches are often limited to reasoning about neighbourhoods. Classical examples include frontier-based exploration [3] and sampling-based methods such as RRT [18], MCTS [19], though their adaptability and performance are often constrained in practical applications.

With the rapid development of machine learning, reinforcement learning (RL) has been widely explored as an alternative to heuristic exploration strategies [6], [16]. RL-based methods learn value functions or policies from interactions with the environment, allowing robots to adaptively select actions for more efficient exploration. For example, Niroui et al. [20] combined deep reinforcement learning with classical frontier-based exploration, where an A3C network processes the known map, the robot's position, and frontier locations to output the next target frontier.

B. Prediction-Driven Exploration

A mainstream prediction-based exploration strategy [9], [10] is to train models with deep learning(DL) to predict maps, and then compute information gain through a utility function to determine the next action. MapEx [12] and PIPE [13] fine-tune LaMa [21] on the KTH floor dataset for large-scale map prediction, where the predicted mean and variance are further utilized to compute frontier or path information gain. However, map prediction performance remains limited in distant regions. Ericson et al. [22] addressed indoor wall prediction with an autoregressive line-segment model trained on a deep network to estimate frontier information gain. Yet, global maps often admit multiple plausible structures (e.g., a corridor may continue straight, turn left, or turn right). While CNNs typically yield a single deterministic output, diffusion models, as probabilistic generators, naturally capture such diversity and produce multiple plausible predictions, resulting in more robust map completions.

C. Diffusion Model in Exploration

Denoising Diffusion Probabilistic Models (DDPMs) [23] mitigate the myopic nature of conventional neural networks by capturing long-term dependencies and maintain global structural consistency. They generate predictions via iterative denoising to recover the global data distribution. Song et al. proposed DDIMs [24] to accelerate sampling through skip steps. Subsequent advances include variance prediction and scheduling [25], classifier guidance for large-scale models [26], and consistency models for efficient sampling [27].

Diffusion models have been widely applied to image inpainting [14], style transfer [28], and various multimodal tasks [29]. Beyond these fields, diffusion models have also found numerous applications in robotic exploration [12], [30]. For example, 4CNet [30] predicts unknown regions in multi-robot systems using a conditional consistency model and estimates prediction confidence to support exploration under resource constraints. DARE [31] leverages a diffusion policy trained on expert demonstrations to reason about potential structures in unknown areas and generate entire exploration paths in a single inference. Liang et al. [32] employ diffusion models to generate trajectories for outdoor navigation, while NoMaD [33] directly produces executable robot actions in unknown environments. Building on these ideas, we adopt a skip-resampled DDPM with variable-mask conditioning as our global map prediction model.

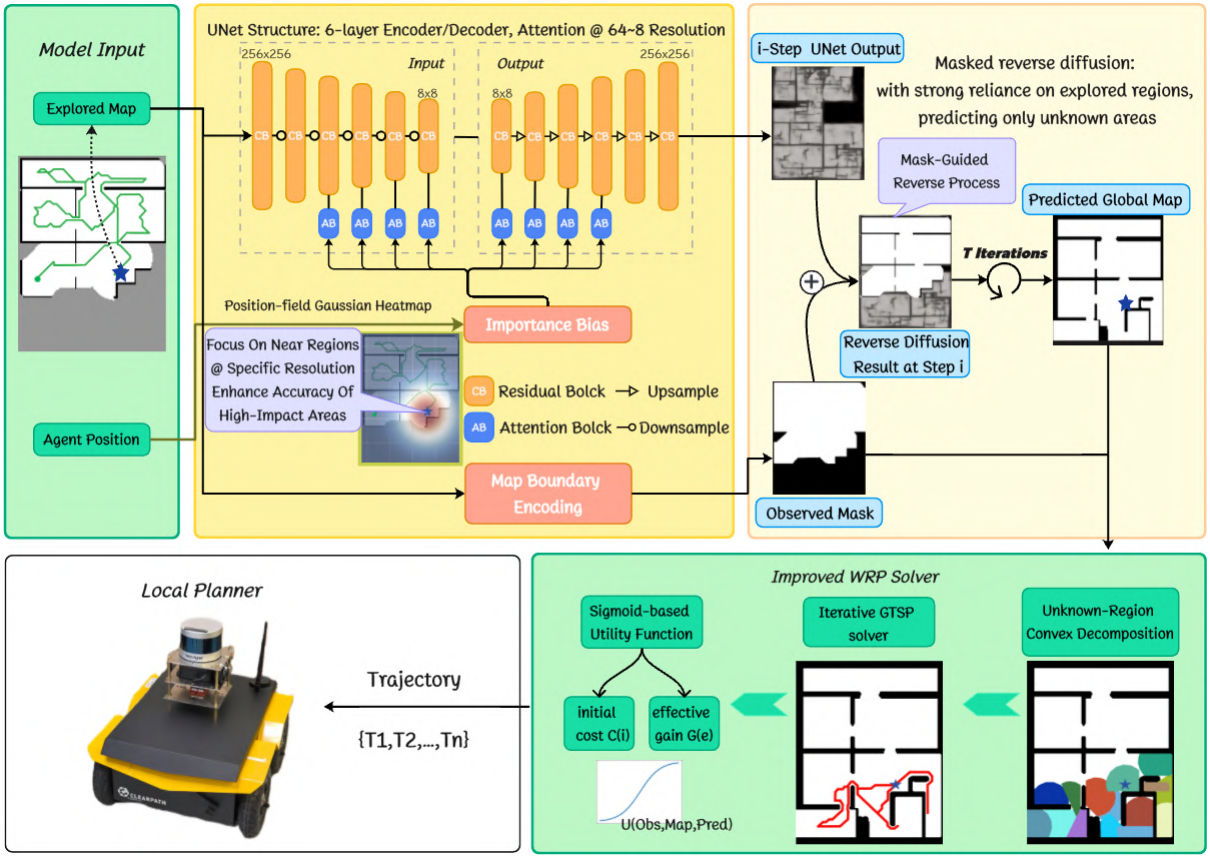


Fig. 2: **Framework of dPWM.** Our exploration framework is implemented by solving a globally optimal path. The model takes as input the exploration map and a Gaussian heatmap of the robot's position. These inputs are encoded with a variable mask, and during each denoising step, attention mechanisms are applied at four resolution levels (64, 32, 16, and 8) to enhance long-term structural dependencies. After the predicted map is obtained, a dynamic WRP solver performs convex decomposition of the potential traversable regions and iteratively solves the GTSP to compute the optimal exploration path.

III. Methodology

We formulate the indoor environment map prediction task as image inpainting on a 2D occupancy grid map. An improved DDPM combined with variable masks is employed to predict the potential features of unknown regions. After obtaining the globally predicted map, the environment is partitioned into observed traversable areas and predicted potential traversable areas. The path planning objective is then to traverse all potential traversable areas with the shortest possible path. The overall framework of dPWM is illustrated in Fig. 2.

A. Exploration Problem Formulation

The environment is represented as a 2D occupancy grid, where each cell $m_{i,j}$ is either occupied, free, or unknown. The state at time t is defined as $s_t = (x_t, m_t)$, with x_t denoting the robot pose, m_t the occupancy grid and a_t the action. The state transition can be factorized as

$$p(s_{t+1} \mid s_t, a_t) = p(x_{t+1} \mid x_t, a_t), p(m_{t+1} \mid m_t, x_{t+1}), \tag{1}$$

where the first term models the motion dynamics and the second term updates the map using new sensor measurements.

Based on the observed map and the robot's current state, a predictive map model $\mathcal{F}$ is employed to infer the latent global traversability features, yielding a predicted map P_i for downstream path planning:

$$P_i = \mathcal{F}(O_i, S_i). \tag{2}$$

A new round of prediction is triggered once the deviation between the predicted map and the newly observed map exceeds a predefined threshold. Based on the predicted map, a globally optimal path is then generated, while the utility function U evaluates alternative exploration directions by jointly considering information gain and traversal cost. This value-driven strategy enables the robot to balance local exploration with long-term planning objectives.

B. Improved DDPM

Forward diffusion process: In the forward (noising) process, we construct a Markov chain in which the data distribution at step i depends only on the previous step $i-1$ and the Gaussian noise injected at step i:

$$\begin{aligned} q(x_i \mid x_{i-1}) &= \mathcal{N}\left(x_i; \sqrt{1-\beta_i}\, x_{i-1}, \beta_i \mathbf{I}\right), \\ x_t &= \sqrt{1-\beta_t} x_{t-1} + \sqrt{\beta_t} \cdot \epsilon, \end{aligned} \tag{3}$$

where β_i denotes the variance schedule that controls the noise level at step i. Nichol et al. [25] demonstrated

that employing a cosine variance schedule and learning the variance parameters in the forward process can enhance the generation quality and accelerate the sampling procedure.

$$f(t) = \cos\left(\frac{t/T + s}{1+s} \cdot \frac{\pi}{2}\right), \\ \sum_{\theta}(x_t, t) = \exp\left(\nu \log \beta_t + (1-\nu)\log \hat{\beta}_t\right), \tag{4}$$

Reverse diffusion process: In the reverse process, the model progressively denoises the noisy sample x_T back to the clean data x_0 through an iterative denoising procedure.

$$q(x_{t-1}|x_t, x_0) = q(x_t|x_{t-1}, x_0) \cdot \frac{q(x_{t-1}|x_0)}{q(x_t|x_0)}, \tag{5}$$

The data distribution at each step can be modeled as a Gaussian

$$x_{t-1} \sim \mathcal{N}\left(\tilde{\mu}_t, \tilde{\beta}_t \mathbf{I}\right), \tag{6}$$

where $\tilde{\mu}_t$ and $\tilde{\beta}_t$ denote the mean and variance of the reverse transition. A U-Net $\epsilon_\theta(x_t, t)$ is employed to take the current sample x_t and the timestep t as inputs, and to predict the parameters of this Gaussian distribution.

Since the complete sampling procedure of diffusion models requires iterative denoising from steps $\{T, T-1, \ldots, 2, 1\}$, the generation process is inherently slow and unsuitable for real-time exploration. A common strategy is to adopt skip sampling during the reverse process [24]; however, increasing the skip interval often leads to blurred generations. To address this issue, we employ skip resampling [14], which enhances long-range dependencies through resampling operations while preserving generation quality.

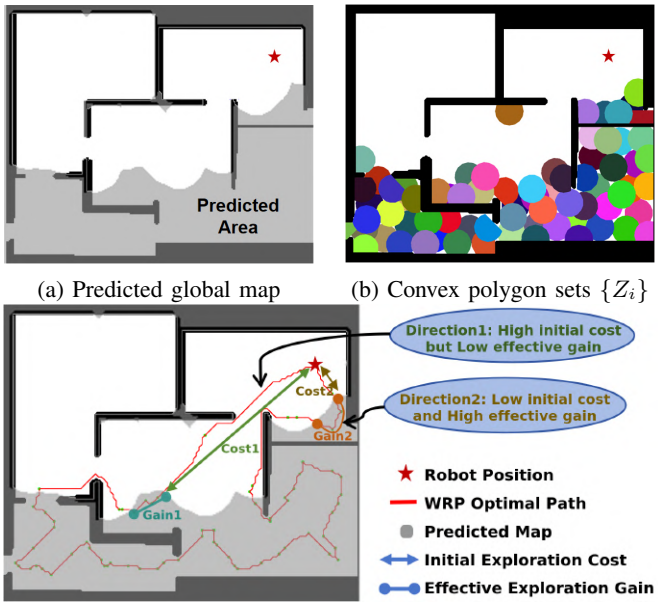
Position-based Gaussian Heatmap: In unexplored regions, areas near the robot are more critical for the next decision, as accurate predictions in these regions help ensure the optimality of each action. To emphasize these regions, we incorporate a Gaussian heatmap–based positional field into the reverse process:

$$H(x, y) = A \exp\left(-\frac{(x-x_0)^2 + (y-y_0)^2}{2\sigma^2}\right), \tag{7}$$

where (x_0, y_0) denotes the keypoint location. The standard deviation σ is defined as $\sigma = \alpha \cdot s$, with α set to 0.04 and s denoting the image scale ($s = 256$). The amplitude A controls the peak value of the heatmap, and we set $A = 1$ in all experiments.

Specifically, this field is injected at the 64, 32, 16, and 8 resolution layers of the U-Net, where different spatial regions are assigned varying weights, enabling the model to focus more on areas that are crucial for decision-making. By assigning higher importance to regions closer to the robot, the model prioritizes accurate predictions in areas that directly influence the immediate exploration decisions, while still maintaining coverage of distant regions. This strategy effectively balances local precision and global awareness, which is particularly beneficial in partially observable environments where errors in nearby regions can propagate through planning.

Introducing such a distance-aware positional weighting in the diffusion process has several advantages: **i)** it guides the model to focus on regions most critical for the next action, while down-weighting uncertain distant areas, thereby enhancing prediction quality and stabilizing the sampling process; **ii)** it provides a natural inductive bias that encodes spatial importance relative to the robot's current position, enhancing generalization across different maps and exploration scenarios. Since diffusion models learn the underlying data distribution rather than task-specific features, this positional guidance allows the model to leverage global structural patterns while emphasizing locally relevant regions, leading to more robust and efficient planning.



(a) Predicted global map (b) Convex polygon sets $\{Z_i\}$

(c) WRP optimal path

Fig. 3: Path planned by the improved dynamic WRP solver, which traverses the unexplored regions while treating the explored map as a traversable area.

C. Dynamic WRP-Solver

The Watchman Route Problem (WRP) is a classical graph-theoretic formulation that seeks the shortest path from which an agent can observe an entire environment. However, traditional WRP solvers, designed for static maps, are not directly applicable to dynamic exploration tasks.Based on the predicted map P, we compute the optimal exploration path using an improved Watchman Route Problem (WRP) solver, as illustrated in Fig. 3. The algorithm first samples the predicted region and generates a set of convex polygons, ensuring that the robot can fully observe each subregion from any position within a polygon. Next, points are sampled along the edges of each convex polygon at fixed intervals to construct the node sets for the Generalized Traveling Salesman Problem (GTSP). Finally, an iterative optimization method is applied to solve for the optimal GTSP path. To guide exploration, we employ the exploration mask O_m to identify the boundaries of unexplored regions.

$$O_m(y,x) = \begin{cases} 0, & P(y,x)\ is\ unexplored\ or\ occupied, \\ 1, & P(y,x)\ is\ free. \end{cases} \tag{8}$$

During convex polygon decomposition [7], O_m is employed to constrain the partitioning process, yielding a set of convex polygons Z_i that collectively cover all unexplored regions in the predicted map, as illustrated in Figs. 3a and 3b. Target points are then sampled within each convex polygon, and a distance matrix $A[i,j]$ is constructed between every pair of target points. Finally, the problem is formulated as a GTSP, which is solved iteratively to obtain the optimal target path sequence $\{T_i\}$ that visits each convex polygon. Since each polygon is geometrically convex, traversing the set of convex polygons ensures that all unexplored regions are covered within the specified observation radius. By leveraging the dynamically updated observation map O_m, this approach enables the computation of an optimal exploration path that adapts in real time to newly observed areas.

We design a sigmoid-based utility function to evaluate the expected exploration gain in both directions along the computed optimal closed-loop path. The core objective of our exploration strategy is to guide the robot toward unexplored regions, thereby minimizing redundant traversal of already explored areas. To this end, we define two key metrics: the initial exploration cost C_{Initial}, which denotes the path length from the robot's current position to the entry point of an unexplored region, and the effective exploration gain $G_{\text{Effective}}$, which measures the path length traveled by the robot within unexplored regions. The overall exploration gain is then evaluated by introducing a weighting parameter λ, which balances these two terms. The utility function is defined as

$$U = \frac{1}{1 + e^{-(G_{\text{Effective}} - \lambda C_{\text{Initial}})}}, \tag{9}$$

The Fig. 3c highlights how the utility function is employed to evaluate closed-loop WRP paths across different exploration directions, yielding comprehensive scores that enable the robot to select globally optimal actions.

IV. Experiments

We begin by evaluating dPWM on the HouseExpo [15] dataset, comparing its performance against multiple baselines. To assess cross-domain generalization, we further test learning-based methods on the Dungeon [17] dataset. Ablation studies are conducted to quantify the contributions of individual components and their combinations. Finally, real-world experiments are performed to validate the practical feasibility of the proposed framework.

A. Experimental Setup

We train our model on 30,000 HouseExpo indoor maps and evaluate it on 3,000 test maps, all with a resolution of 1 m per pixel, in both simulation and real-world environments. In each trial, the robot's initial position is randomly set within free space; it is equipped with a multi-line LiDAR (12 m range) and constrained to a linear velocity of 1.5 m/s and an angular velocity of 1.2 rad/s. The diffusion model is trained for 650k iterations with a batch size of 128, a learning rate of 1×10^{-5}, 1000 diffusion steps, using a cosine noise schedule, and optimized with KL divergence loss.

Model training and simulation are conducted on an Ubuntu 18.04 system equipped with an Intel Xeon Gold 6326 CPU and four NVIDIA RTX 3090 GPUs, while real-world experiments are performed on an Ubuntu 20.04 platform with an Intel i7-13620H CPU and an NVIDIA RTX 4060 GPU.

TABLE I: Exploration distances (m) of each method over 3 runs with average

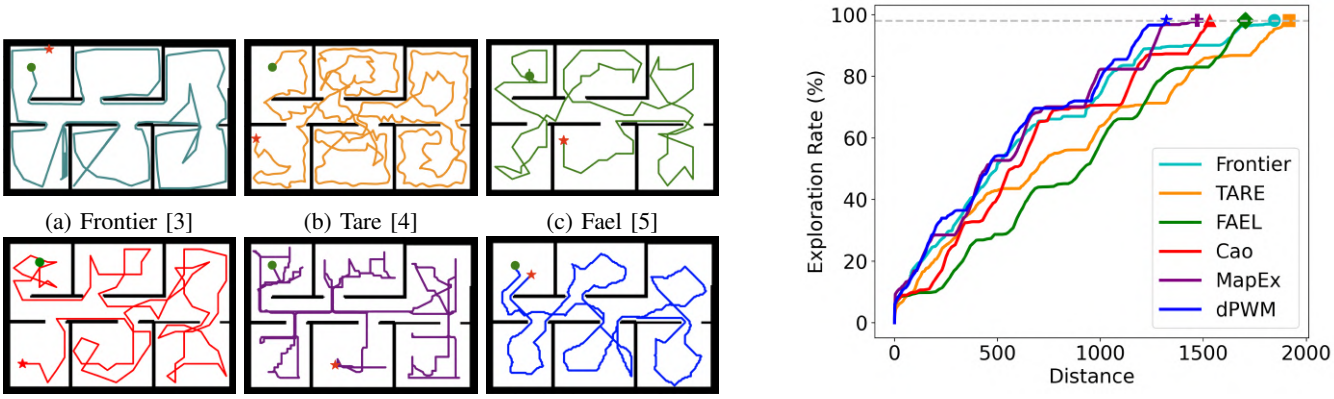
Method	Average Path Length	Variance
Frontier [3]	2017.4	604.6
TARE [4]	2756.2	607.5
FAEL [5]	1960.2	582.1
Cao et al. [16]	1826.7	530.4
MapEx [12]	1782.8	631.5
dPWM(ours)	**1452.5**	**438.7**

B. Performance

During evaluation, we consider the exploration is finished when 98% of the area is explored. The exploration process is terminated once the explored map coverage exceeds the threshold. We benchmark our proposed dPWM method against state-of-the-art approaches: **i)** heuristic methods (Frontier [3], TARE [4], and FAEL [5]), **ii)** a reinforcement learning-based method [16], and **iii)** the map-prediction-based approach MapEx [12]. To further evaluate the generalization capability of our method, we compare dPWM with two learning-based approaches.

Evaluation: For evaluation, we randomly select images from the HouseExpo test set and perform three repeated trials per image, resulting in a total of 1,000 evaluation samples. The mean of the actual executed path length is reported as the performance metric, while the variance reflects the stability of each method. Frontier [3], FAEL [5], and TARE [4] are tested with their default parameter configurations. For TARE, we additionally generate the required Gazebo environments by converting 2D maps into the corresponding simulation scenarios. For the RL-based method by Cao [16], the training data is restricted to 640×480 inputs. Therefore, we preprocess the HouseExpo dataset by resizing and padding the maps to the required dimensions, setting the resolution to 0.5 m/pixel, and further adapting the map style to match the model requirements. For MapEx [12], we fine-tune the LaMa-Place2 model on a self-collected dataset of partially explored maps. Each training sample is represented as a (map, mask) pair, where the partial exploration maps and corresponding masks (unknown regions) are generated using Cao's RL-based exploration policy.

As shown in Table I, the proposed dPWM method outperforms all baseline approaches in both exploration path efficiency and stability. Specifically, dPWM reduces exploration path length by 18.53% compared to the SOTA map-prediction method, and improves stability by 20.49% relative to the SOTA DRL-based approach. These results demonstrate that dPWM achieves more globally optimal and long-sighted exploration compared to traditional prediction-based strategies. Moreover, since diffusion models inherently capture the diversity of unexplored regions [34], dPWM also



(a) Frontier [3] (b) Tare [4] (c) Fael [5]

(d) Cao et al. [16] (e) MapEx [12] (f) dPWM(ours)

(g) Exploration rate comparison

Fig. 4: Comparison of results. (a)-(f) show six cases, and (g) shows the comparison of exploration rate curves across different methods.

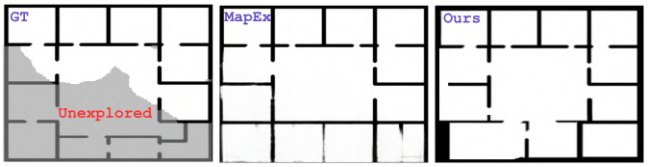
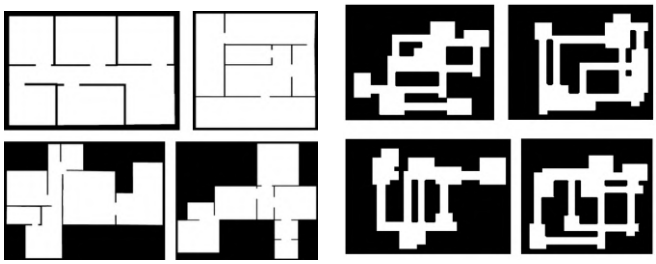


Fig. 5: Comparison of global map prediction between our dPWM and baseline method.

achieves higher exploration stability. As illustrated in Fig. 4, we compare the exploration rate curves of different methods within the same scenario. As the explored area increases, the input information for dPWM becomes progressively more complete, leading its predicted maps to converge toward the ground-truth map. Consequently, even in the later stages of exploration, our prediction-based approach is able to sustain a higher exploration rate advantage.

To evaluate prediction performance, we compare our framework against baseline map-prediction model [12]. As shown in Fig. 5, our diffusion-based framework achieves superior preservation of global structural consistency compared to the baseline.



(a) HouseExpo(Train) (b) Dungeon(Test)

Fig. 6: Datasets for generalization experiments

Generalization: We evaluate generalization on the Dungeon dataset by comparing our method with two learning-based baselines, while excluding heuristic approaches whose rule-based nature precludes learnable generalization. All models are trained on the HouseExpo [15] dataset and subsequently tested on a stylistically different 2D grid-map dataset, Dungeon [17], to assess their cross-dataset generalization performance (Fig. 6). HouseExpo provides large-scale residential layouts with relatively regular room and corridor structures, whereas Dungeon comprises artificially generated maze-like environments with irregular topologies, long corridors, and dead ends, making it a suitable out-of-distribution benchmark for evaluating the generalization of learning-based exploration methods. As shown in Table II, the proposed dPWM framework demonstrates superior generalization performance. Specifically, it outperforms the SOTA DRL-based method by 26.48%, and achieves an improvement of 16.37% over the SOTA map-prediction-based method. These results highlight that diffusion models, by learning data distributions beyond explicit structural features, are able to capture a deeper understanding of spatial layouts, thereby enabling more robust cross-domain generalization.

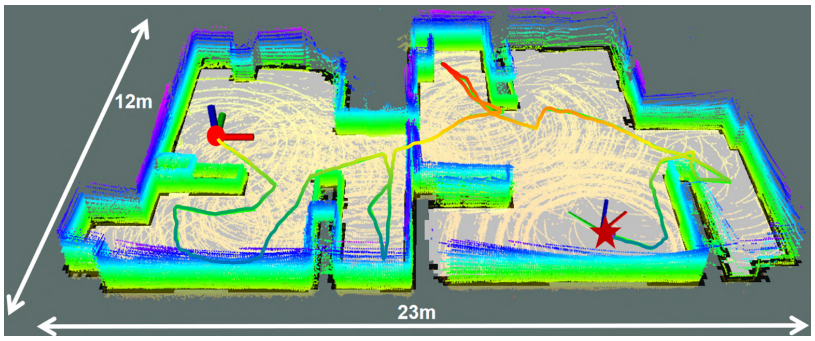
TABLE II: Generalization Performance of Different Method

Method	Cao.	MapEx	dPWM(ours)
Average Path Length	1714.5	1507.3	**1260.5**
Variance	373.5	274.2	**258.3**

C. Ablation Study

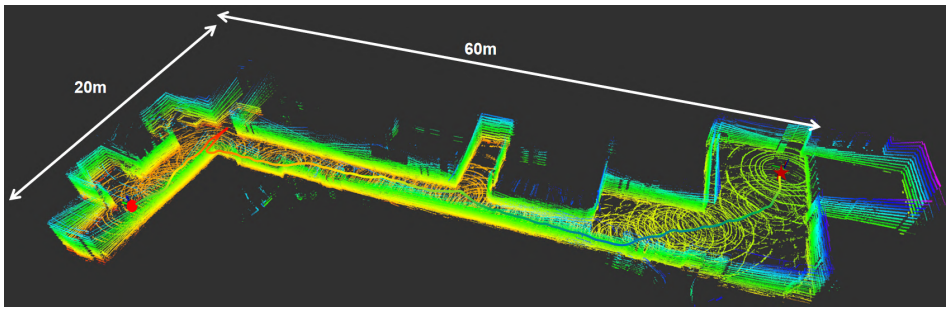
We conduct ablation experiments on two key components of our framework: the Gaussian heatmap–based positional prior and the dynamic WRP solver. Four variants are evaluated: **None** denotes the baseline model without any additional modules; **Prior** introduces a keypoint-based Gaussian heatmap prior to enhance local dependency; **D-WRP** incorporates the improved dynamic WRP solver; and **dPWM** represents the full model.

Gaussian heatmap: We primarily investigate the benefits of incorporating a Gaussian heatmap–based positional prior into the diffusion model. While the inherent sampling diversity of diffusion models enables the consideration of multiple possible global map configurations, it may also lead to "hallucinations," which can reduce the short-term utility of robot actions. To mitigate this, we employ a cross-attention mechanism guided by a Gaussian heatmap centered



(a) Point Cloud map and 2D Grid map of Office Environment

(b) Large-scale indoor corridor environment



(c) Point Cloud map of Large-scale 2D floor plan environment

(d) Jackal Robot

Fig. 7: Real-world validation. Experiments in two real-world environments: a 23×12 m office scene and a 60×24 m corridor scene.

at the robot's current position, encouraging the model to focus more on local regions around the robot. This design enhances local spatial dependencies without compromising global consistency.

To quantitatively assess the impact of the local positional prior, we evaluate the same model trained for 650k iterations under two settings (with and without the prior). For the partially explored map, we generate 100 predicted maps and compute the average PSNR and SSIM with respect to the corresponding ground-truth maps to evaluate the pixel-level accuracy and structural similarity of the predictions, thereby assessing both the fidelity and reliability of the predicted maps. Table IIIa demonstrates that incorporating the positional prior yields better PSNR and SSIM, with the improvement being particularly significant within the robot's local observation window. Table IIIb shows that introducing the Gaussian heatmap–based positional prior(Group **Prior**) improves exploration efficiency by 10.04% compared to the baseline model without any additional modules (**None**).

Dynamic WRP Solver: The traditional graph-based WRP solver [7] aims to compute a path that covers the entire environment given a fixed observation range. For incremental exploration maps, only the unexplored regions need to be traversed, while the known areas are treated as freely traversable. Consequently, the paths generated by the original WRP solver do not necessarily correspond to the optimal exploration route. We evaluated the performance of the original WRP solver and the dynamic WRP solver on the exploration task using the same test set. For each method, we record the mean final exploration path length. Table IIIb shows that incorporating the improved dynamic WRP solver reduces exploration path length by 5.5% compared to the **None** baseline. When combined (Group **dPWM**), the two modules achieve an 18.32% improvement, exceeding the sum of their individual contributions and highlighting their complementary effect.

TABLE III: Ablation Study

(a) Gaussian heatmap ablation results

Model	Origin	Positional Prior
PSNR	7.31	**8.07**
SSIM	0.66	**0.73**

(b) Ablation Analysis

Group	Exploration Cost
dPWM	**1452.5** ($\pm$ **438.7**)
Prior	1561.7 ($\pm$ 486.5)
D-WRP	1629.0 ($\pm$ 450.4)
None	1718.6 ($\pm$ 611.5)

D. Real-world Experiments

To validate the effectiveness of dPWM in real-world environments, we conduct experiments in two different physical scenarios: an office environment of 23×12 m and a large indoor corridor of 60×24 m. We use a Clearpath Jackal robot as the ground platform, equipped with an onboard mini-PC (Intel Core i7-13620H CPU and NVIDIA GeForce RTX 4060 GPU), an IMU, and a Velodyne VLP-16 LiDAR, as shown in Fig 7. During mapping, we set the map resolution to 0.2 m/pixel. The 3D LiDAR observations are first converted

into a 3D voxel map using UFOmap [35], which is then projected into a 2D occupancy grid for planning purposes. Our results show that the robot successfully completes the exploration tasks in both environments, demonstrating the applicability of dPWM to real-world scenarios.

V. Conclusion

In this paper, we presented **dPWM**, a diffusion-based predictive with WRP solver for autonomous map exploration. A key innovation of our framework is the introduction of a positional Gaussian heatmap, which enhances local dependencies while maintaining global stability in the diffusion sampling process. Combined with a dynamic WRP solver, our method enables more efficient and stable exploration compared with both heuristic and learning-based baselines. Experimental results in simulation and real-world scenarios demonstrated that dPWM achieves lower exploration cost, higher stability, and stronger generalization to unseen environments.

Future work will focus on reducing the computational overhead of diffusion sampling, extending the positional bias mechanism to multi-scale attention modules, and exploring multi-robot collaboration as well as tighter integration with real-time SLAM systems for large-scale exploration tasks.

References

[1] K. N. McGuire, C. De Wagter, K. Tuyls, H. J. Kappen, and G. C. de Croon, "Minimal navigation solution for a swarm of tiny flying robots to explore an unknown environment," *Science Robotics*, vol. 4, no. 35, p. eaaw9710, 2019.

[2] S. Nahavandi, R. Alizadehsani, D. Nahavandi, S. Mohamed, N. Mohajer, M. Rokonuzzaman, and I. Hossain, "A comprehensive review on autonomous navigation," *ACM Computing Surveys*, vol. 57, no. 9, pp. 1–67, 2025.

[3] 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'*. IEEE, 1997, pp. 146–151.

[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, 2021, p. 2.

[5] 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.

[6] Y. Cao, T. Hou, Y. Wang, X. Yi, and G. Sartoretti, "Ariadne: A reinforcement learning approach using attention-based deep networks for exploration," *arXiv preprint arXiv:2301.11575*, 2023.

[7] W.-P. Chin and S. Ntafos, "Optimum watchman routes," in *Proceedings of the second annual symposium on Computational geometry*, 1986, pp. 24–33.

[8] J. Mikula and M. Kulich, "Towards a continuous solution of the d-visibility watchman route problem in a polygon with holes," *IEEE Robotics and Automation Letters*, vol. 7, no. 3, pp. 5934–5941, 2022.

[9] Y. Wang and A. Del Bue, "Where to explore next? exhistcnn for history-aware autonomous 3d exploration," in *European Conference on Computer Vision*. Springer, 2020, pp. 125–140.

[10] Z. Xing, J. Wang, and X. Zhu, "An efficient learning based autonomous exploration algorithm for mobile robots," in *2022 IEEE International Conference on Real-time Computing and Robotics (RCAR)*. IEEE, 2022, pp. 551–556.

[11] H. Gao, H. Que, K. Li, W. Shan, M. Liu, R. Zhao, L. Mu, X. Yang, Q. Wei, and F. Qiao, "Mapping at first sense: A lightweight neural network-based indoor structures prediction method for robot autonomous exploration," *arXiv preprint arXiv:2504.04061*, 2025.

[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," *arXiv preprint arXiv:2409.15590*, 2024.

[13] 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.

[14] A. Lugmayr, M. Danelljan, A. Romero, F. Yu, R. Timofte, and L. Van Gool, "Repaint: Inpainting using denoising diffusion probabilistic models," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2022, pp. 11 461–11 471.

[15] T. Li, D. Ho, C. Li, D. Zhu, C. Wang, and M. Q.-H. Meng, "Houseexpo: A large-scale 2d indoor layout dataset for learning-based algorithms on mobile robots," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2020, pp. 5839–5846.

[16] 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.

[17] F. Chen, S. Bai, T. Shan, and B. Englot, "Self-learning exploration and mapping for mobile robots via deep reinforcement learning," in *Aiaa scitech 2019 forum*, 2019, p. 0396.

[18] S. LaValle, "Rapidly-exploring random trees: A new tool for path planning," *Research Report 9811*, 1998.

[19] F. Bourgault, A. A. Makarenko, S. B. Williams, B. Grocholsky, and H. F. Durrant-Whyte, "Information based adaptive robotic exploration," in *IEEE/RSJ international conference on intelligent robots and systems*, vol. 1. IEEE, 2002, pp. 540–545.

[20] F. Niroui, K. Zhang, Z. Kashino, and G. Nejat, "Deep reinforcement learning robot for search and rescue applications: Exploration in unknown cluttered environments," *IEEE Robotics and Automation Letters*, vol. 4, no. 2, pp. 610–617, 2019.

[21] 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*, 2022, pp. 2149–2159.

[22] L. Ericson and P. Jensfelt, "Beyond the frontier: Predicting unseen walls from occupancy grids by learning from floor plans," *IEEE Robotics and Automation Letters*, vol. 9, no. 8, pp. 6832–6839, 2024.

[23] J. Ho, A. Jain, and P. Abbeel, "Denoising diffusion probabilistic models," *Advances in neural information processing systems*, vol. 33, pp. 6840–6851, 2020.

[24] J. Song, C. Meng, and S. Ermon, "Denoising diffusion implicit models," *arXiv preprint arXiv:2010.02502*, 2020.

[25] A. Q. Nichol and P. Dhariwal, "Improved denoising diffusion probabilistic models," in *International conference on machine learning*. PMLR, 2021, pp. 8162–8171.

[26] P. Dhariwal and A. Nichol, "Diffusion models beat gans on image synthesis," *Advances in neural information processing systems*, vol. 34, pp. 8780–8794, 2021.

[27] Y. Song, P. Dhariwal, M. Chen, and I. Sutskever, "Consistency models," 2023.

[28] J. Chung, S. Hyun, and J.-P. Heo, "Style injection in diffusion: A training-free approach for adapting large-scale diffusion models for style transfer," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2024, pp. 8795–8805.

[29] X. Xu, Z. Wang, G. Zhang, K. Wang, and H. Shi, "Versatile diffusion: Text, images and variations all in one diffusion model," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2023, pp. 7754–7765.

[30] A. H. Tan, S. Narasimhan, and G. Nejat, "4cnet: A diffusion approach to map prediction for decentralized multi-robot exploration," *arXiv preprint arXiv:2402.17904*, 2024.

[31] 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)*. IEEE, 2025, pp. 11 987–11 993.

[32] J. Liang, A. Payandeh, D. Song, X. Xiao, and D. Manocha, "Dtg: Diffusion-based trajectory generation for mapless global navigation. in 2024 ieee," in *RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 5340–5347.

[33] A. Sridhar, D. Shah, C. Glossop, and S. Levine, "Nomad: Goal masked diffusion policies for navigation and exploration," in *2024 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2024, pp. 63–70.

[34] E. Fahnestock, E. Fuentes, P. R. Osteen, S. Ancha, and N. Roy, "Learning semantic traversability priors using diffusion models for uncertainty-aware global path planning," in *ICRA 2024 Workshop on Resilient Off-road Autonomy*, 2024.

[35] D. Duberg and P. Jensfelt, "Ufomap: An efficient probabilistic 3d mapping framework that embraces the unknown," *IEEE Robotics and Automation Letters*, vol. 5, no. 4, pp. 6411–6418, 2020.