Computers, Materials & Continua

Tech Science Press

https://doi.org/10.32604/cmc.2026.081556

ARTICLE

SSAG: Situational Semantic Augmented Graph for Active SLAM in Object-Goal Navigation

Shasha Tian[1,2], Zhengyang Chen[1,3], Kai Ren[1,2], Na Li[1,2], Chongwei Ruan[4], Zhijia Cui[1,3] and Mian Wu[4,*]

[1]School of Computer Science, South-Central Minzu University, Wuhan, China
[2]Hubei Provincial Engineering Research Center of Agricultural Blockchain and Intelligent Management, Wuhan, China
[3]Hubei Provincial Engineering Research Center for Intelligent Management of Manufacturing Enterprises, Wuhan, China
[4]DONG FENG Machine Tool PLANT Co., Ltd., Shiyan, China

*Corresponding Author: Mian Wu. Email: rljc-wumian@dfl.com.cn

Received: 04 March 2026; Accepted: 03 May 2026; Published: 23 July 2026

ABSTRACT: To address the issues of low exploration efficiency and "geometric myopia" caused by the lack of high-level environmental structure modeling for mobile robots in complex indoor environments, this paper proposes an active SLAM object navigation method based on Situational Semantic Augmented Graph (SSAG). Unlike methods that learn policies solely on pixel-level semantic maps or exploit only object-level relations for implicit association, this work elevates local observations online into a room-level topological graph and performs explicit semantic reasoning over unobserved regions. First, an online room segmentation algorithm is employed to transform unstructured sensory data into a structured graph representation that characterizes room-level functional attributes and topological associations. Subsequently, a Graph Attention Network (GAT) is utilized to perform explicit reasoning on the semantic attributes of unobserved regions, providing decision-making priors for long-range exploration. Building upon this, we introduce a Product of Experts (PoE) mechanism within the Proximal Policy Optimization (PPO) framework to deeply fuse semantic reasoning heatmaps with geometric hard constraints. This integration optimizes the target selection strategy, generating long-term navigation goals with superior semantic consistency and physical reachability. Experimental results on the Habitat simulator and Gibson dataset demonstrate that the proposed SSAG achieves a Success weighted by Path Length (SPL) of 0.340, outperforming SemExp and SemGO by 15.3% and 4.9%, respectively, with a total success rate of 68.4%. Notably, for object search tasks with strong spatial correlations (e.g., "bed" and "toilet"), the success rates reach 73.2% and 72.5%, representing a performance gain of over 20% compared to baseline methods. These results validate the effectiveness of room-level semantic reasoning for object navigation and demonstrate the capability of the proposed method to achieve efficient autonomous navigation in unknown, complex scenarios.

KEYWORDS: Active SLAM; object navigation; situational semantic augmented graph; graph attention network; deep reinforcement learning

1 Introduction

Simultaneous Location and Mapping (SLAM) serves as the fundamental technology for mobile robots to achieve autonomous navigation and long-term operation in unknown environments. Its performance directly dictates the adaptability of the robot's task in complex scenarios [1]. Traditional SLAM methods, which rely on prescribed trajectories and fuse multi-source sensor observations to complete pose estimation and environmental modeling, have made significant strides in controlled settings and short-duration

Copyright © 2026 The Authors. Published by Tech Science Press. This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

tasks [2]. However, these methods struggle to satisfy the intricate demands of indoor environments in the real-world, where robots must operate continuously in unknown, dynamic, and explicitly goal-oriented scenes [3]. Relying solely on passive perception and fixed motion strategies, traditional frameworks fail to concurrently address the multi-objective requirements of exploration efficiency, localization robustness, and task success rates [4].

To overcome these limitations, the Active SLAM paradigm has been proposed. Its core involves explicitly incorporating robotic motion decision-making into the SLAM framework, allowing the robot to actively select observation positions with high informational value based on its current environmental perception, thus achieving synergy between efficient exploration and precise mapping [5]. Early Active SLAM research relied primarily on geometric heuristics [6], such as frontier points and information gain [7], to guide exploration by maximizing unknown space coverage and minimizing map uncertainty [8]. However, these geometry-centric methods are inherently weakly correlated with specific task objectives. In complex indoor scenarios, they are prone to "shortsighted" behaviors—such as redundant local back-and-forth exploration and a lack of global judgment regarding key functional regions—which severely constrains overall navigation efficiency.

In recent years, Deep Reinforcement Learning (DRL) has paved new avenues for sequential decision-making challenges in Active SLAM [9]. Specifically, several studies formulate Active SLAM as a Partially Observable Markov Decision Process (POMDP), where exploration policies are learned through continuous agent-environment interactions [10]. Although these approaches exhibit superior environmental adaptability compared to traditional heuristics, they remain heavily reliant on low-level geometric observations or dense grid representations [11]. The lack of explicit modeling for high-level environmental structures makes these policies susceptible to sensor noise and limits their generalization capabilities across diverse scenes [12]. Currently, advances in semantic perception have introduced promising optimization directions for Active SLAM [13]. Statistical correlations between object distributions, room functions, and spatial layouts have been proven to significantly improve the rationality of exploration decisions. When navigating unknown environments, humans do not rely solely on geometric cues; instead, they infer potential target regions based on the underlying semantic structure of the scene [14]. For instance, when searching for a "toilet," humans prioritize searching the bathroom over the living room; similarly, the observation of a "bed" prompts an inference of a nearby ensuite bathroom. Such common-sense-based semantic reasoning is precisely what current geometry-driven methods lack.

However, existing semantic SLAM approaches predominantly treat semantic information as auxiliary attributes within the perception layer, confined to tasks such as object detection or obstacle discrimination. Consequently, they do not integrate semantics into the core of active decision-making, leaving a loose coupling state between semantic perception and action selection [15]. We argue that the crux of Active SLAM lies in the explicit modeling of the environment's situational semantic structure and its transformation into actionable decision-making priors for the policy network, rather than merely increasing the density of semantic labels. More importantly, the methodological novelty of SSAG does not lie in using semantic mapping, graph learning, or PPO as isolated modules, since these components have all appeared in prior literature. Instead, the novelty lies in coupling them through a task-oriented hierarchy: online room segmentation converts dense semantic maps into room instances, an edge-aware GAT explicitly reasons about the semantic attributes of unobserved rooms, and a PoE-based decision layer separates policy intention, semantic plausibility, and geometric reachability during long-term goal selection. To this end, this paper proposes an active SLAM method based on the Situational Semantic Augmented Graph (SSAG), as conceptually illustrated in Fig. 1. The method discretizes the continuous space into semantic regions with distinct functional attributes and constructs a structured graph representation that characterizes the

topological and semantic associations between these regions. Using graph attention networks (GAT) [16], the framework performs semantic reasoning on unobserved regions and fuses these semantic priors with geometric constraints within a reinforcement learning framework to generate globally consistent exploration policies. This approach utilizes room-level semantic constraints to effectively mitigate the issues of "geometric myopia" and "policy instability" inherent in traditional methods, thereby enhancing the efficiency and robustness of goal-oriented exploration in complex indoor environments.

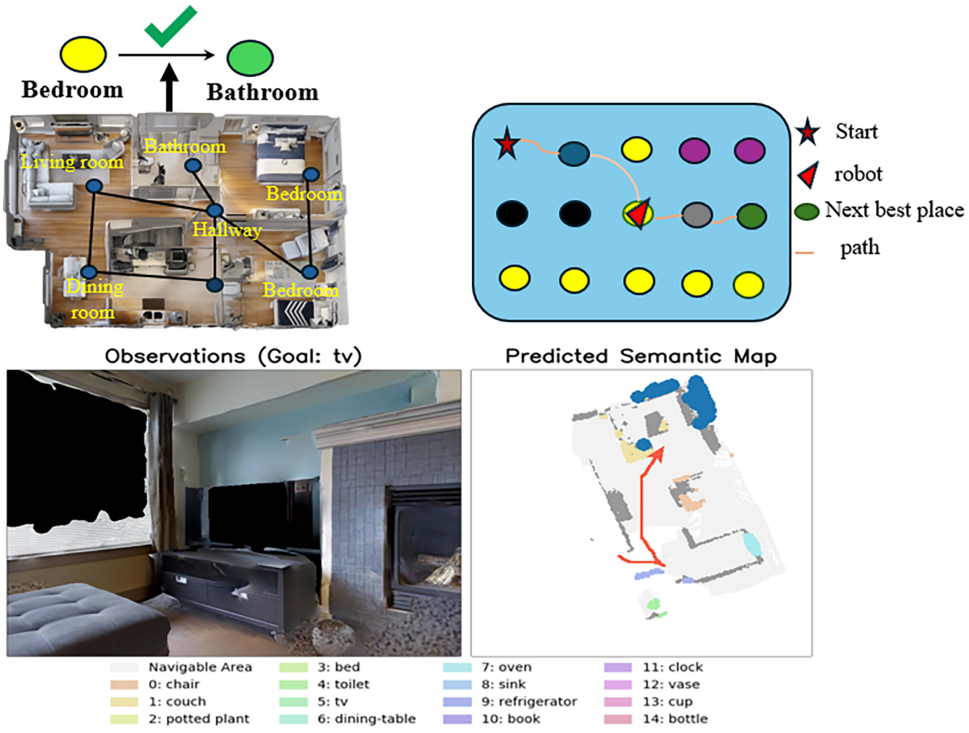

Figure 1: Example of a target navigation task. Navigation process with "TV" as the target: The top left is a scene top-down view and room annotations; the top right is a room topological graph, where the agent utilizes the "bedroom-bathroom" adjacency relationship to plan the exploration path; the bottom left is the current first-person observation; the bottom right is the real-time constructed semantic map, with the red trajectory representing the traveled path and the blue area indicating the detected target object.

The main contributions of this paper are as follows.

(1) Proposing a Situational Semantic Augmented Graph (SSAG) modeling method for active SLAM, which explicitly incorporates room-level semantic structures and topological relationships into the decision-making process through online room segmentation and feature encoding, thereby achieving a deep coupling between high-level semantics and active exploration;

(2) Designing a semantic reasoning mechanism based on Graph Attention Networks (GAT), which aggregates semantic information from adjacent rooms via a message-passing mechanism to accurately predict the semantic attributes of unexplored regions, providing effective decision-making priors for long-range exploration;

(3) Constructing a reinforcement learning exploration strategy that fuses semantic priors with geometric constraints, and validating the effectiveness of the proposed method through comparative experiments on the Habitat simulation environment and Gibson dataset.

2 Related Work

The goal-oriented indoor navigation problem investigated in this paper intersects two core technical domains: active SLAM and semantic perception. In recent years, the development of deep reinforcement learning (DRL) and semantic understanding technologies has provided new paradigms for addressing this challenge. Related research has primarily evolved along two technical trajectories: first, RL-based active SLAM methods, which focus on enhancing exploration efficiency through end-to-end or hierarchical policy learning; second, active SLAM integrated with semantic information, aimed at utilizing environmental semantic structures to guide navigation decisions. This section reviews the research progress and limitations of related works from these two perspectives and elucidates the distinctions and connections between the proposed method and existing approaches.

2.1 Reinforcement Learning-Based Active SLAM

With breakthroughs in deep reinforcement learning (DRL) for sequential decision-making, researchers have increasingly modeled active SLAM as a Partially Observable Markov Decision Process (POMDP), enabling robots to autonomously learn exploration policies through environmental interactions, which significantly enhances environmental adaptability. Neural SLAM [9] introduced external memory modules to explicitly maintain spatial representations within a learning framework, laying the foundation for active SLAM research based on learning. Chaplot et al. [10] employed a hierarchical reinforcement learning (HRL) architecture to decouple high-level goal selection from low-level motion control, thus improving the stability of long-horizon exploration policies. Although these methods mitigate the sample efficiency issues of end-to-end learning through structured memory, the content of such memory remains predominantly focused on geometric information.

Regarding optimization of exploration policies, Niroui et al. [11] utilized reinforcement learning to achieve adaptive selection of candidate exploration targets, catering to complex unknown scenarios such as search and rescue. The ARIADNE method [13] further improved exploration efficiency in simulated environments by ranking candidate viewpoints through an attention mechanism. Although these approaches improve the ability to focus on key regions, the learning of attention weights still relies on geometric features and lacks semantic-level guidance. Furthermore, Zhang et al. [14] utilized the Proximal Policy Optimization (PPO) algorithm to address the challenges of the design of reward functions in exploration policies, proposing an adaptive reward adjustment mechanism that allows robots to reasonably evaluate their behavior value in different environments. Li et al. [15] combined deep reinforcement learning with particle swarm optimization (PSO) to accelerate the search for global optimal routes.

Despite the various innovations in policy learning mechanisms, the aforementioned learning-based methods share common limitations: they rely on low-level geometric features or dense grid maps as state inputs and lack explicit modeling of high-level environmental semantic structures. This makes the policies sensitive to sensor noise and results in weak cross-scene generalization. Simultaneously, these exploration strategies exhibit "geometry-driven" characteristics that fail to replicate the human cognitive reasoning process based on semantic common sense, making them difficult to adapt to complex goal-oriented tasks. This paper addresses the deficiency in high-level semantic modeling at the state representation level by introducing the Situational Semantic Augmented Graph (SSAG) and the GAT reasoning mechanism [17].

2.2 Semantic-Integrated Active SLAM

As robotic tasks extend toward high-level objectives such as object navigation and situational understanding, the limitations of active SLAM methods relying solely on geometric information—specifically in decision-making efficiency and task relevance—have become increasingly pronounced. Thus, integrating

semantic information has emerged as a crucial direction for improving active SLAM performance. Early semantic SLAM research focused mainly on semantic fusion during the mapping stage, treating object detection and semantic segmentation results as auxiliary attributes of geometric maps to optimize map representation or assist in subsequent planning [18]. However, in these approaches, semantic information does not directly participate in exploration decision-making, representing a "posterior annotation" mode where semantics and decision-making remain loosely coupled.

In object navigation tasks, semantic information has proven to be effective in narrowing the search space and improving the rationality of decision-making. Chaplot et al. [12] utilized semantic segmentation to construct multi-channel semantic maps, guiding robots to prioritize unexplored semantic regions related to the target, which significantly improved the success rate of object navigation. Yu et al. [19] combined semantic information with frontier selection strategies to propose a semantic frontier exploration method adapted for visual object navigation scenarios. The SemGO method proposed by Wu et al. [20] introduced a multi-head self-attention mechanism to model global and local semantic dependencies, addressing the issue of insufficient long-range correlation in semantic policies and improving both path efficiency and task success rates. Although these methods integrate semantic information into the input to decision-making, the semantic representations remain at the pixel or object level, not capturing room-level spatial structures.

Recently, researchers have begun to focus on the role of room-level high-level semantic constraints and graph structure modeling. Luo and Gu [21] proposed the RGN algorithm, which introduces the constraints of the room category and the mechanisms of clustering of boundaries to distinguish functional differences between similar objects in different rooms, validating the guiding value of high-level semantics for decision-making. Asgharivaskasi and Atanasov [22] constructed multi-class semantic OctoMaps and designed exploration criteria at the semantic level based on Shannon mutual information; however, this method entails high computational complexity and is difficult to integrate deeply with reinforcement learning. The RASLS method [23] maps object semantics and spatial distributions into an augmented graph structure and combines it with reinforcement learning to enhance exploration stability, yet lacks explicit semantic reasoning for unobserved regions.

Table 1 summarizes the technical characteristics of representative methods. Despite their advances, existing approaches exhibit the following limitations: (1) Methods such as SemExp introduce semantic maps, but fail to model the topological associations between rooms. (2) SemGO [20] enhances semantic modeling through an attention mechanism but lacks the ability to express reasoning about unobserved regions. (3) RGN [21] introduces room-level constraints, but room categories are determined by predefined rules rather than learned through end-to-end semantic reasoning with Graph Neural Networks.

Table 1: Comparison of representative goal-oriented navigation methods.

Method	Semantic Repr	Topo Modeling	Semantic Reasoning	Decision Fusion	Main Limitations
SemExp [12]	Pixel-level map	×	×	CNN concatenation	Lacks high-level structure
SemGO [20]	Object-level attention	×	Implicit (MHSA)	Attention weighting	No explicit topo. reasoning
RGN [21]	Room category label	Partial (Clusters)	Rule-based	Constraint filtering	Rule-based (not learned)
RASLS [23]	Object-space graph	✓	Implicit (GNN)	Reinforcement learning	Unfused geometric constraints
SSAG (Ours)	Room-level graph	✓	Explicit (GAT)	PoE Mechanism	—

In contrast, the SSAG method proposed in this paper constructs a room-level topological graph and leverages Graph Attention Networks (GAT) for message passing, enabling explicit reasoning of semantic attributes for unexplored regions. Furthermore, through the Product of Experts (PoE) mechanism, SSAG deeply integrates semantic priors with geometric constraints, effectively addressing the issue of loose coupling between semantic perception and active decision-making.

Another research direction highly relevant to real-world deployment is SLAM in dynamic environments. Representative methods such as DynaSLAM [24] explicitly detect and remove moving objects by combining multi-view geometry with learning-based segmentation, thereby maintaining a static map representation. DS-SLAM [25] further integrates semantic segmentation, motion-consistency checking, and dense semantic mapping to improve localization robustness in dynamic scenes. More recently, STDyn-SLAM [26] combines stereo vision and semantic segmentation to enhance VSLAM stability in outdoor dynamic environments. These methods mainly strengthen the SLAM front-end by filtering or modeling dynamic elements, whereas SSAG focuses on room-level semantic reasoning for long-horizon task-driven exploration. Therefore, the two directions are complementary: future work can incorporate dynamic-object filtering into the semantic mapping front-end of SSAG to improve robustness in realistic indoor scenes.

3 Method

The proposed active navigation system based on the Situational Semantic Augmented Graph (SSAG) establishes a comprehensive closed-loop from perception to execution. This framework enables the agent to achieve efficient navigation in unknown indoor environments without global priors, relying solely on noisy egocentric observations.

As illustrated in Fig. 2, the system consists of four primary components. Initially, the semantic map mapping module transforms RGB-D images and pose estimates into a multi-channel semantic map that encompasses obstacle distribution, exploration status, and semantic category probabilities. Subsequently, the situational semantic reasoning module dynamically constructs the SSAG based on this map and employs a Graph Attention Network (GAT) to infer the semantic attributes of room nodes. This module outputs a semantic heatmap that represents the probability that target objects appear in various rooms. Building upon this, the goal-oriented semantic search strategy module integrates long-term goal generation to formulate a global navigation policy. By predicting the posterior probability of the target object's presence in each room, it provides high-level guidance to narrow the search space. Finally, the local policy and action execution module generates local movement commands based on global guidance and real-time observations. New observations are fed back into the mapping module, forming a dynamic perception-reasoning-decision-execution loop that continues until the agent localizes the target or reaches the maximum time steps. Through the dynamic construction and reasoning of the SSAG, the system effectively overcomes pose noise and semantic uncertainty in unexplored regions, facilitating robust active object navigation.

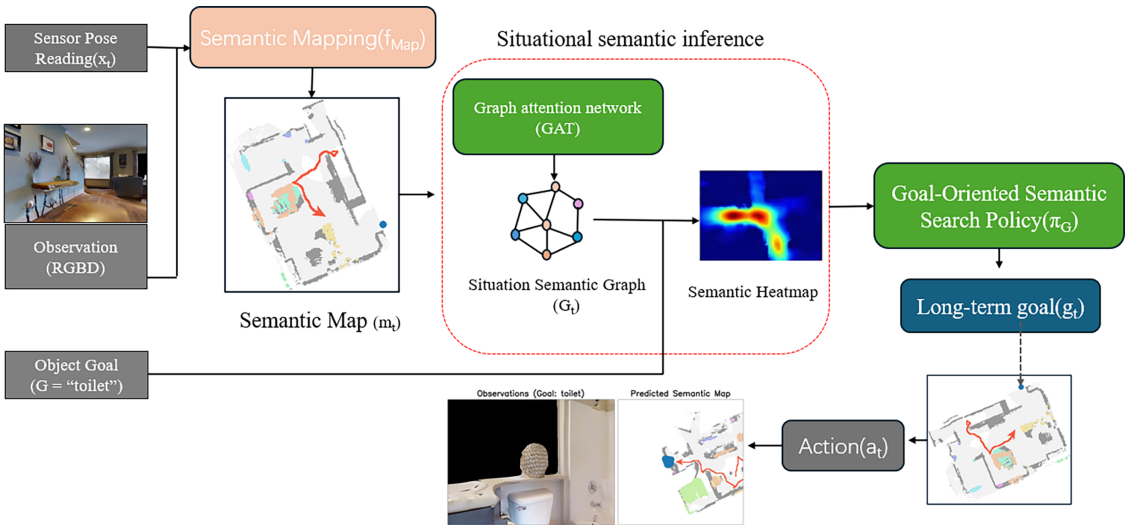

Figure 2: Schematic diagram of the overall process of the SSAG method.

3.1 Problem Formulation

In an unknown environment, the agent is required to navigate to a target object specified by its category name O_{goal}, where $O_{\text{goal}} \in \mathcal{O}$. The set of target objects is defined as $\mathcal{O} = \{\text{bed, toilet, chair, couch, potted plant, tv}\}$. This set includes objects with strong room-level associations, such as "bed" and "toilet," as well as objects with more dispersed distribution characteristics, such as "chair," "couch," "tv," and "potted plant." The set of potential room categories in the environment is defined as $\mathcal{C} = \{\text{bedroom, bathroom, dining room, kitchen, living room, corridor, None}\}$. At the start of each task ($t = 0$), the agent is randomly initialized at a navigable starting state S_0 within the environment. Equipped only with egocentric sensors, the agent acquires observation data $o_t = \{I_t^{RGB}, I_t^{Depth}, \hat{\xi}_t\}$ at each time step t, where I_t^{RGB} denotes the RGB image, I_t^{Depth} is the depth image, and $\hat{\xi}_t$ represents the estimation of noisy poses. The system does not rely on global pose sensors or pre-generated environmental maps. Based on these sensor observations, the agent must construct and maintain a multi-channel semantic map M_t in real-time. This map contains core information such as obstacle distribution, explored regions, and probabilities of semantic categories. Currently, a situational semantic augmented graph G_t is dynamically generated to infer the semantic attributes of unexplored regions through graph-structured reasoning. Specifically, the agent predicts the posterior probability that the target object O_{goal} appears within each room node $R_i \in V_t$. The set of executable actions is defined as $\mathcal{A} = \{\text{turn left, turn right, move forward, stop}\}$. We adopt this standard four-action discrete setting to maintain fairness with Habitat-based baselines. It should be noted that the absence of an explicit backward action is not the main reason for failed episodes, because the agent can reorient itself by approximately 180° through consecutive turn actions and then move forward to realize an effective retreat. Thus, the agent is not strictly trapped in confined corners or dead-end regions under the current action design, although this indirect maneuver may incur slightly higher local motion cost in some cases. The objective is to learn a policy function $\pi(a_t \mid M_t, G_t, O_{\text{goal}})$ to select the optimal action within a maximum number of time steps $T_{\max}$. The task is successfully completed if the agent stops within a predefined distance threshold from the target object. The task ends when the agent executes the stop action or reaches $T_{\max}$.

3.2 Situational Semantic Augmented Graph

Situational semantics encompass not only the category information of objects within an environment but also the implicit functional attributes of rooms, spatial geometric features, and co-occurrence relationships between objects. To endow the agent with a human-like understanding of high-level environmental topological structures, this paper constructs the Situational Semantic Augmented Graph (SSAG). This framework transforms unstructured sensory data into structured graph representations, providing a foundation for subsequent semantic reasoning and decision-making. Unlike traditional occupancy grid maps or viewpoint-based topological graphs, SSAG is defined as a dynamic undirected graph $G_t = (V_t, E_t)$, where t denotes the current time step, V_t is the set of identified room nodes at time t, and E_t represents the set of edges that characterize physical connectivity between rooms. The construction of the SSAG primarily involves four stages: online room segmentation, node feature modeling, edge feature modeling, and inference via Graph Attention Networks (GAT).

3.2.1 Online Heuristic Room Segmentation

To discretize the continuous free navigation space into individual "room instances" with distinct semantic attributes (e.g., bedrooms or kitchens), this paper constructs an online heuristic room segmentation algorithm. Centered on morphological operations and connected component analysis, this algorithm generates robust room segmentation results in real-time without requiring a pre-generated environment map. The operational workflow is visualized in Fig. 3.

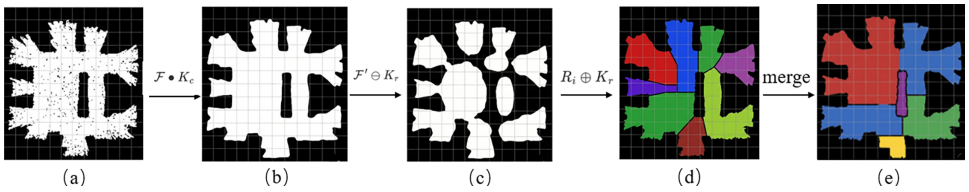

Figure 3: The complete processing flow of the online heuristic room segmentation algorithm.

Free Space Extraction and Noise Suppression: The traversable area $\mathcal{T}$ is intersected with the explored area $\mathcal{E}_{\text{exp}}$ to obtain the free space $\mathcal{F} = \mathcal{T} \cap \mathcal{E}_{\text{exp}}$ that can currently be safely segmented. A 3×3 circular structuring element K_c is used to perform a morphological closing operation $\mathcal{F} \bullet K_c$ on the free space $\mathcal{F}$. This step eliminates small holes and sensor noise, resulting in denoised free space $\mathcal{F}'$. The specific operations are illustrated in Fig. 3b.

Channel Cutting and Initial Partitioning: A circular structure element K_r, with a radius equal to half the width of a doorway, is selected to carry out a morphological erosion operation $\mathcal{F}'' = \mathcal{F}' \ominus K_r$ on the denoised $\mathcal{F}'$. This causes narrow channel areas to shrink until they disappear, thereby geometrically disconnecting adjacent rooms. The specific operations are illustrated in Fig. 3c.

Room Identification and Boundary Recovery: A 4-connected component labeling algorithm is applied to the erosion result to obtain a set of n initially segmented independent room regions $\mathcal{R} = \{R_1, R_2, \ldots, R_n\} = \text{CC}_4(\mathcal{F}'')$. Subsequently, a morphological dilation operation $R_i' = R_i \oplus K_r$ is performed in each room region R_i to restore the room boundaries to their original dimensions. The specific operations are illustrated in Fig. 3d.

Redundant Area Optimization: By reducing the area threshold $A_{\min}$ ($A_{\min} = S_r/\delta^2$, where S_r is the preset minimum room area), redundant small regions significantly below this threshold are filtered out. Gleichzeitig, small functional areas with independent semantic significance (e.g., storage rooms, foyers) are

preserved, enhancing the detail retention and robustness of the segmentation results. The specific operations are illustrated in Fig. 3e. Note that these morphological parameters are specified in metric space and then discretized according to the map resolution δ. In particular, the doorway-related erosion radius scales with $1/\delta$, while the minimum room area threshold is normalized by $A_{\min} = S_r/\delta^2$. Therefore, the segmentation is not tied to a single pixel resolution, although highly irregular open-plan layouts may still require retuning of the doorway-width prior.

3.2.2 Room Node Feature Encoding

In the Scene Semantic Augmentation Graph, to fully represent the environmental state, we encode each room node as a 31-dimensional comprehensive feature vector $v = [v_{geo}, v_{obj}, v_{exp}, \mathbf{p}_i^{\mathrm{h}}] \in \mathbb{R}^{31}$. This vector integrates four types of information: geometric morphology, semantic content, exploration status, and room type priors. The specific definitions of each dimension are as follows.

Geometric Morphology Features: To reflect the spatial structural characteristics of the room, we extract a 6-dimensional geometric feature vector $\mathbf{v}_{geo} \in \mathbb{R}^6$. This set of features includes normalized room area, perimeter, aspect ratio, convexity, compactness, and shape complexity. These geometric descriptors effectively distinguish the physical morphology differences of various functional areas, such as corridors (long and narrow) and living rooms (square and large).

Semantic Object Features: The semantic features $\mathbf{v}_{obj} \in \mathbb{R}^{15}$ are designed to represent the key target objects contained within the room (e.g., beds, couches). For the 15 categories of identifiable indoor objects, the normalized cumulative confidence for each category within the room region R_i is calculated and defined as:

$$S_{obj}^k = \min\left(1, \frac{\sum_{(x,y)\in R} P_{sem}^k[x,y]}{|R_i|}\right), \quad k = 0, 1, \ldots, 14 \tag{1}$$

where R represents the set of pixel coordinates for the current room, $|R_i|$ denotes the area of the current room (number of pixels), and $P_{sem}^k[x,y]$ is the confidence level of the semantic segmentation network for the k-th class of objects at coordinates (x,y). When $S_{obj}^k \approx 0$, it indicates that the k-th class of objects is nearly absent in the room; when $S_{obj}^k \approx 1$, it indicates a high probability of the presence of the object.

Exploration Status Features: To guide the robot's active exploration behavior, we define a 3-dimensional exploration feature vector $\mathbf{v}_{exp} = [\rho_{exp}, \rho_{bnd}, \rho_{frt}]^\top$, which, respectively, characterizes the exploration progress, the degree of boundary confirmation, and the potential exploration value of the room. This vector aims to quantify the agent's perceptual completeness and potential exploration value for the current room R. Let $M \in \{0,1\}^{H\times W}$ be the global binary exploration map (1 for explored, 0 for unexplored), where the explored area is defined as the set $\mathcal{E}_{exp} = \{(x,y) \mid M[x,y] = 1\}$. The indices are defined as follows.

Exploration Rate (ρ_{exp}): This metric measures the completeness of the coverage of the room's internal area, defined as the ratio of the number of explored pixels to the total area of the room:

$$\rho_{exp} = \frac{\sum_{(x,y)\in R_i} M[x,y]}{|R_i|} \tag{2}$$

where $|R_i|$ represents the cardinality of pixels (i.e., area) of room R_i. A higher ρ_{exp} indicates that most of the room has been explored.

Boundary Coverage (ρ_{bnd}): To evaluate the degree of reconstruction of the geometric structure of the room, the boundary coverage is introduced. Let ∂R_i be the set of pixels for the physical boundary of the room (i.e., the contours of the wall). This metric calculates the proportion of the boundary covered by sensors:

$$\rho_{bnd} = \frac{|\partial R_i \cap \mathcal{E}_{exp}|}{|\partial R_i|} \tag{3}$$

A value closer to 1 indicates that the room's contour structure has been fully confirmed, with no unknown geometric boundaries remaining.

Frontier Ratio (ρ_{frt}): This metric extracts exploration frontiers through morphological operations to quantify the remaining exploration potential. The frontier area is defined as the boundary zone between explored and unexplored regions, calculated as follows:

$$\rho_{frt} = \frac{|R_i \cap \bar{\mathcal{E}}_{exp} \cap (\mathcal{E}_{exp} \oplus K)|}{|R_i|} \tag{4}$$

where $\bar{\mathcal{E}}_{exp}$ represents the unexplored area, $\oplus$ denotes the morphological dilation operation, and K is the structuring element. The numerator captures unknown regions located at the exploration edge; a non-zero value typically implies that navigable exploration targets still exist within the room.

Heuristic Room Type Prior: Considering the significant spatial regularity in the distribution of indoor objects (e.g., "beds" frequently appearing in "bedrooms"), we introduce a heuristic algorithm to estimate room types. We construct an object-room co-occurrence matrix $\mathbf{M} \in \mathbb{R}^{15\times 7}$, where the element $M_{k,c}$ represents the conditional probability that a room belongs to category c given that the object k is observed. This matrix is based on scene annotations from the Gibson dataset, statistical analysis of the frequency of various objects in different room types, and applying Laplace smoothing for conditional probability estimation. Table 2 shows some typical values of the co-occurrence matrix. Based on the cumulative object confidence W_k in the current room (derived from Eq. (1)), we calculate the weighted voting score $\mathbf{u}$ for each type of room and generate a 7-dimensional probability distribution vector $p^{\mathrm{h}}_{i,c}$ using the Softmax function:

$$\mathbf{u} = \sum_{k=0}^{14} W_k \cdot \mathbf{M}[k,:] \tag{5}$$

$$p^{\mathrm{h}}_{i,c} = \frac{\exp(u_c)}{\sum_{c=0}^{6}\exp(u_c)}, \quad c = 0,\ldots,6 \tag{6}$$

where i is the index of the room, representing the probability of the i-th node of the room, $c \in \{0,1,\ldots,6\}$ corresponds to the seven types of rooms, and h denotes the prior probability obtained by the heuristic method. This distribution $\mathbf{p}^{\mathrm{h}}_i = [p^{\mathrm{h}}_{i,0},\ldots,p^{\mathrm{h}}_{i,6}]^\top$ provides an initial classification guess based on local observations, which serves as one of the input node features for the Graph Neural Network (GNN) in subsequent stages. This heuristic prior relies solely on observed objects and does not consider inter-room topological relationships, which will be further optimized by the GNN.

Table 2: Ablation study results on the Gibson validation set (mean ± standard deviation, 5 runs).

Configuration	SPL ↑	Success ↑	DTS (m) ↓
SSAG (Full Model)	0.340 ± 0.012	0.684 ± 0.018	1.529 ± 0.045
w/o GAT (Heuristic Prior Only)	0.298 ± 0.015	0.621 ± 0.022	1.698 ± 0.058
w/o PoE (Policy Network Only)	0.312 ± 0.014	0.645 ± 0.020	1.612 ± 0.052
w/o Room Reward	0.305 ± 0.016	0.632 ± 0.024	1.658 ± 0.061
w/o Edge Features (Node Features Only)	0.325 ± 0.013	0.662 ± 0.019	1.573 ± 0.048

3.2.3 Edge Feature Encoding

In the Scene Semantic Augmentation Graph, the edge $e_{ij} \in E_t$ not only represents the physical adjacency between room R_i and room R_j but also carries geometric and semantic association information regarding spatial transitions. For any edge $(i, j) \in \mathcal{E}_t$, it is encoded as an 8-dimensional feature vector $\mathbf{e}_{ij} \in \mathbb{R}^8$, used to describe the connection attributes between room R_i and room R_j.

Adjacency Detection: The adjacency of the rooms is determined by using a morphological dilation operator. Let K be a disk-shaped structuring element with a radius of 1 pixel. If the dilated room region overlaps with an adjacent room, i.e., $(R_i \oplus K) \cap R_j \neq \varnothing$, room R_i and room R_j are considered adjacent, and an edge (i, j) is added to the graph. The contact area between the two rooms is defined as $\mathcal{A}_{ij} = (R_i \oplus K) \cap R_j$, which corresponds to doors or passages in the actual environment.

Feature Vector Construction: The edge feature vector $\mathbf{e}_{ij} \in \mathbb{R}^8$ encompasses information in four core dimensions. Doorway features (3 dimensions) include the detected door width and the distances from the center of the door to the centroids of the rooms on both sides, used to measure the ease of passage between rooms. The spatial relationship features (2 dimensions) represent the straight-line distance and the relative azimuth angle between the centroids of the two rooms, quantifying their spatial positional correlation. The relative azimuth angle is computed in the global map coordinate frame from the centroid of R_i to the centroid of R_j, rather than in the robot's egocentric frame, so that edge attributes remain invariant to instantaneous heading changes. Semantic similarity features (2 dimensions) calculate the cosine similarity between adjacent rooms in terms of room type probability distributions and object distributions, capturing strong semantic associations such as "master bedroom and master bathroom." The strength of connectivity (1 dimension) is measured by the number of pixels on the contact boundary between the two rooms, quantitatively characterizing the physical tightness of the connection between them.

3.2.4 Graph Attention Network Inference

Based on the aforementioned room node features $\mathbf{v}_i \in \mathbb{R}^{31}$ and edge features $\mathbf{e}_{ij} \in \mathbb{R}^8$, a Graph Attention Network (GATv2) is adopted for message passing and room type inference. Let $\mathbf{h}_i^{(l)} \in \mathbb{R}^d$ denote the hidden representation of node i in the l-th layer, and $\mathcal{N}(i)$ denote the set of neighboring nodes for node i. The message passing update rule is defined as follows:

$$\mathbf{h}_i^{(l+1)} = \mathrm{LN}\left(\mathbf{h}_i^{(l)} + \sum_{j \in \mathcal{N}(i)} \alpha_{ij} \mathbf{W}^{(l)} \mathbf{h}_j^{(l)}\right) \tag{7}$$

where $\mathrm{LN}(\cdot)$ represents the Layer Normalization function, $\mathbf{W}^{(l)} \in \mathbb{R}^{d \times d}$ is the learnable weight matrix of the l-th layer, and α_{ij} is the attention weight modulated by the edge features $\mathbf{e}_{ij}$. After L layers of message passing, the classifier outputs the room type probability predicted by the GNN, denoted as $\hat{\mathbf{p}}_i^{\mathrm{gnn}} \in \mathbb{R}^7$. The

final room type probability is obtained through a weighted fusion of the GNN prediction and the heuristic prior $\mathbf{p}_i^{\mathrm{h}}$:

$$\mathbf{p}_i^{*} = \lambda \hat{\mathbf{p}}_i^{\mathrm{gnn}} + (1-\lambda)\mathbf{p}_i^{\mathrm{h}}, \quad \lambda = 0.6 \tag{8}$$

where λ controls the relative contribution of the GNN inference vs. the previous heuristic. When λ is too small, the model relies excessively on local object observations and fails to exploit topological information; conversely, if λ is too large, the GNN inference may become unstable in the early stages of exploration when the graph structure is incomplete. The parameter λ is set to 0.6 according to the experimental results on the validation set. The number of layers L determines the range of hops for the propagation of information. Although $L = 1$ only aggregates direct neighbor information, an excessively large L may lead to the over-smoothing problem. Our experiments indicate that $L = 2$ is the optimal configuration, allowing the model to effectively capture two-hop adjacency relationships such as "bedroom-corridor-bathroom." Accordingly, the choice $L = 2$ should be understood as the best trade-off on the Gibson validation set rather than a universally optimal depth. In larger or more topologically complex environments, deeper, hierarchical, or adaptive graph propagation may be required to capture longer-range dependencies without severe over-smoothing.

Through the aforementioned feature modeling of nodes and edges, the SSAG transforms the originally discrete and noisy semantic maps into compact, robust, and structured graph data. This not only preserves the geometric and semantic core information of the environment, but also provides a rich information foundation for subsequent use of the Graph Attention Network to capture long-horizon room-type dependencies and achieve goal-oriented decision-making.

3.3 Reward Function Design

In long-horizon goal navigation tasks, traditional reward functions based on geometric information, such as target distance and exploration coverage, exhibit two critical limitations. First, reward signals are often sparse because the agent only receives positive feedback when approaching the target or discovering new regions, which leads to difficulties in policy convergence within large-scale indoor environments. Second, these functions lack semantic guidance. Geometric rewards cannot distinguish between the functional attributes of different rooms, which may cause an agent searching for a "bed" to wander in a kitchen or bathroom for extended periods, resulting in significant inefficient exploration behavior. To address these issues, this paper proposes a room-level semantic reward r_{room}. The core idea is to leverage the room-type probabilities inferred from the Situational Semantic Augmented Graph (SSAG) to provide the agent with dense and semantically consistent intermediate reward signals. When the agent enters a room R_i for the first time at time step t, a reward is granted based on the semantic correlation between that room and the target object:

$$r_{\mathrm{room}} = \begin{cases} \lambda_{\mathrm{room}} \cdot \mathbf{m}_{o_{\mathrm{goal}}}^{\top} \mathbf{p}_i^{*}, & \text{if } R_i \text{ first visit} \\ 0, & \text{otherwise} \end{cases} \tag{9}$$

where $\mathbf{m}_{o_{\mathrm{goal}}} = \mathbf{M}[o_{\mathrm{goal}}, :] \in \mathbb{R}^7$ represents the room-type preference vector corresponding to the target object, $\mathbf{p}_i^{*}$ denotes the fused room-type probability distribution of room R_i, and λ_{room} is the room reward coefficient. This design ensures that the agent receives a higher reward when entering a "bedroom" while searching for a "bed," where as entering semantically irrelevant rooms yields a lower reward. Importantly, r_{room} acts as a bounded shaping reward rather than an oracle supervision signal. It is triggered only when the agent enters a room for the first time and is computed from the fused probability $\mathbf{p}_i^{*}$, which combines

the heuristic prior with the GNN prediction. This design reduces the risk that unstable early-stage graph inference directly misleads policy updates.

Building upon the room-level semantic reward, this paper designs a composite reward function to balance global exploration efficiency with target search precision. The overall definition is given by:

$$r_t = r_{\text{dist}} + r_{\text{cov}} + r_{\text{room}} + r_{\text{time}} \tag{10}$$

where r_{dist} represents the target proximity reward, which guides the agent to move toward the search target. Let D_t denote the distance from the agent's position x_t to the target object at time t. This term is defined as a linear function of the distance difference between adjacent time steps: $r_{\text{dist}} = \lambda_{\text{dist}}(D_{t-1} - D_t)$, where λ_{dist} is the distance reward weight. The reward is positive when the agent advances toward the target and negative otherwise. The term r_{cov} is the exploration coverage reward, which encourages the agent to actively explore unknown regions. It is defined as $r_{\text{cov}} = \lambda_{\text{cov}} \cdot \Delta A_{\text{exp}}$, where λ_{cov} is the exploration reward coefficient and A_{exp} denotes the explored area. By rewarding the geometric information gain, this term effectively prevents the agent from wandering repeatedly within local areas. Finally, r_{time} is the time penalty, which constrains the length of the navigation trajectory and improves overall execution efficiency. It applies a fixed penalty at each time step, defined as $r_{\text{time}} = -\lambda_{\text{time}}$.

3.4 Goal-Oriented Semantic Search Strategy

To enable the robot to understand the spatial structure and layout of the environment based on semantic map information for efficient searching, this paper proposes a goal-oriented semantic search strategy. This strategy is built upon the Proximal Policy Optimization (PPO) framework and integrated with the Situational Semantic Augmented Graph (SSAG) for joint training, as illustrated in Fig. 4.

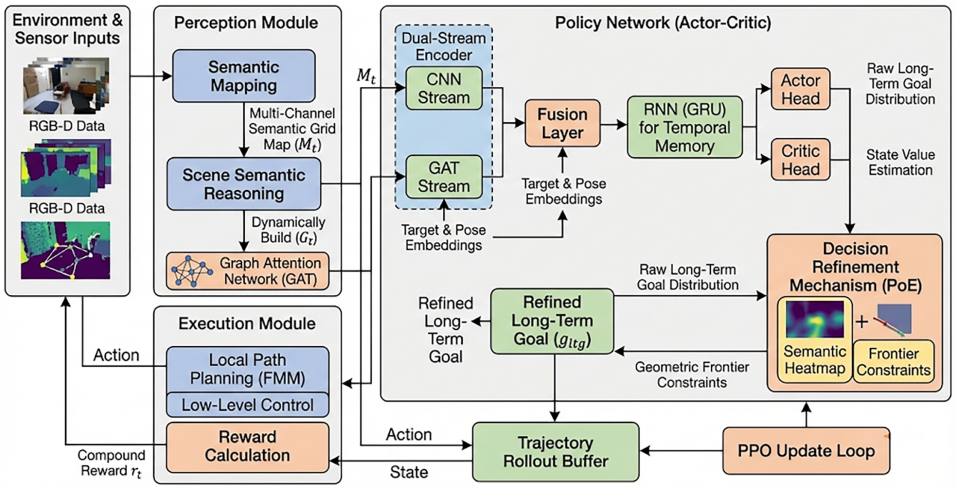

Figure 4: SSAG goal-oriented semantic search strategy system architecture diagram.

The system takes RGB-D observations as input and constructs the semantic map (M_t) and the SSAG (G_t) through the perception module. The policy network adopts a dual-stream encoder architecture to simultaneously utilize local geometric details and global topological semantics. Let the semantic grid map at time t be $M_t \in \mathbb{R}^{C\times H\times W}$ and the situational semantic augmented graph be $G_t = (\mathcal{V}_t, \mathcal{E}_t)$. The CNN stream processes the semantic grid map to extract local geometric features, while the GAT stream processes the SSAG to aggregate semantic correlations and topological structures between nodes. These two streams of

features are concatenated with the target category embedding and the robot pose embedding, then fed into a Gated Recurrent Unit (GRU) to maintain temporal memory. The output layer of the network consists of two branches: the Actor layer, which outputs the probability distribution of long-term goals on the global map $\pi_\theta(g \mid s_t)$, and the Critic layer, which outputs the value estimation $V(s_t)$ of the current state for calculating the advantage function. The raw long-term goal distribution generated by the policy network then enters a decision refinement mechanism based on the Product of Experts (PoE). This mechanism fuses the distribution with the semantic heatmap and geometric boundary constraints to generate the final refined long-term goal g_{ltg}. This goal is subsequently passed to a local path planner using the Fast Marching Method (FMM) to execute low-level motion control. The states and reward signals generated after the agent executes an action are stored in a trajectory buffer for the PPO policy update loop.

To address the randomness issues inherent in pure end-to-end policies during the early stages of training, this paper proposes a targeted optimization strategy based on the Product of Experts (PoE) concept. This method treats three information sources as independent "evaluation experts": the policy network provides the macro exploration direction, the situational semantic graph identifies high-probability regions for the target, and the occupancy grid map defines the physically reachable boundaries.

Let the robot state at time t be s_t, and the candidate long-term goal position be $\mathbf{g} = (g_x, g_y)$, representing 2D coordinates on the global map. The final goal probability distribution is defined as the normalized product of the three expert distributions:

$$P(\mathbf{g} \mid s_t) = \frac{1}{Z} \cdot \pi_\theta(\mathbf{g} \mid s_t) \cdot P(o_{\text{goal}} \mid \mathrm{R}_g) \cdot M_{\text{geo}}(\mathbf{g}) \tag{11}$$

where Z is a normalization constant.

Policy Constraint Field $\pi_\theta(\mathbf{g} \mid s_t)$ is the long-term goal probability distribution output by the Actor layer of the PPO policy network, reflecting the exploration preferences and macro planning intentions learned by the agent through historical experience.

Semantic Constraint Field $P(o_{\text{goal}} \mid \mathrm{R}_g)$ is the target semantic heatmap generated via situational semantic augmented graph inference. Given a target object category o_{goal}, its corresponding room-type preference vector is $\mathbf{M}[o_{\text{goal}}, :] \in \mathbb{R}^7$, representing the prior probability of the object appearing in various room types. The probability of the target object appearing in room $R_{\mathbf{g}}$ is:

$$P(o_{\text{goal}} \mid \mathrm{R}_g) = \sum_{c=0}^{6} P(o_{\text{goal}} \mid \text{room}_c) \cdot P(\text{room}_c \mid R_{\mathbf{g}}) \tag{12}$$

The physical significance of this formula is that the target appearance probability equals the inner product of the "probability of the target appearing in each room type" and the "probability that the target's current room belongs to each room type."

Geometric Constraint Field $M_{\text{geo}}(\mathbf{g})$ is a binary mask generated based on current map boundary points, acting as a hard constraint in physical space to filter out invalid target points inside obstacles or within explored regions:

$$M_{\text{geo}}(\mathbf{g}) = \begin{cases} 1, & \text{if } \mathbf{g} \in \mathcal{F}_{\text{frt}} \\ \epsilon, & \text{otherwise} \end{cases} \tag{13}$$

where $\mathcal{F}_{\text{frt}} \subset \mathbb{R}^2$ is the navigable boundary region and ϵ is a minimal positive constant to ensure that the selected goal falls within the reachable area of the exploration boundary. In implementation, each expert distribution is further lower-bounded by a small ϵ before the final normalization. In this way, the PoE module

acts as a conservative reweighting mechanism rather than a brittle hard veto, alleviating the risk that one overconfident expert suppresses a feasible target region.

4 Experiment and Result Analysis

4.1 Experimental Environment and Dataset

The experimental platform for this study is built on the Ubuntu 24.04 LTS operating system, using an NVIDIA GeForce RTX 5070 Ti GPU for accelerated training and inference of deep learning models. The Habitat high-performance 3D indoor navigation simulator is selected as the simulation environment, as it provides a robust physics engine and realistic visual rendering.

This paper evaluates the proposed method using the Gibson tiny data set. This data set contains 25 training scenes and 5 evaluation scenes. Scene models are reconstructed from real-world buildings through professional scanning, preserving authentic features such as lighting variations, blurred textures, cluttered object stacking, and irregular room layouts. These characteristics significantly increase the difficulty of visual perception and motion control, effectively validating the robustness of the agent in complex environments. The average area of the scene is 127.3 m^2 for the training set and 134.8 m^2 for the validation set, while the average number of rooms is 8.2 for the training scenes and 7.6 for the validation scenes. Six representative indoor object categories are selected as navigation targets: bed, toilet, chair, sofa, television, and potted plant.

The agent is equipped with an RGB-D camera mounted at a height of 0.88 m, with a horizontal field of view (HFOV) of 79° and a visual observation resolution of 640 × 480. The semantic mapping module employs Mask R-CNN as the object detection and semantic segmentation network, using ResNet-FPN as the backbone. To suppress visual noise, the confidence threshold for semantic prediction is set at 0.9. This relatively high threshold favors precision over recall: it suppresses false positives that could contaminate room priors, graph inference, and long-term goal selection, but it may also increase false negatives under occlusion or poor illumination. We therefore regard 0.9 as a conservative operating point rather than a universally optimal choice. The resolution of the semantic occupancy grid map is set to 5 cm/pixel, with a coverage area of 24 m × 24 m. To simulate the motion constraints of real-world robots, the system limits the maximum trajectory to 500 steps. A trial is considered successful if the agent reaches within 1.0 m of the target object and executes the stop action within the allowed steps.

4.2 Experimental Setting

In the policy network training phase, the agent is trained using the Proximal Policy Optimization (PPO) algorithm, with network parameters updated via the Adam optimizer. The learning rate is set to 2.5×10^{-5}, the discount factor is 0.99, and the PPO clipping coefficient is 0.2. Each update iteration consists of 4 epochs to ensure the stability of the policy update process. The policy network adopts a GRU structure to support temporal memory, with a hidden layer dimension of 256. To encourage a larger exploration, the entropy regularization coefficient is set to 1×10^{-4}.

The reward function uses a composite mechanism to balance the proximity of the target and efficient exploration. This mechanism consists of four components: a target distance reward (coefficient 0.1) to reinforce the agent's goal-approaching behavior; an intrinsic exploration reward (coefficient 0.02) to incentivize active exploration of unknown regions through entropy reduction of map coverage; a room-level reward (coefficient 0.5) to provide high-level semantic constraints and guidance; and a time penalty (coefficient 0.01) applied at each time step to constrain trajectory length and enhance execution efficiency, thereby promoting stable learning of long-horizon policies.

During the training stage, the agent is trained for a total of 25,000 episodes across 25 training scenes, with 1000 episodes per scene. In the validation stage, the agent is deployed in 5 completely unseen scenes for testing, with 200 episodes per scene, totaling 1000 episodes. At the initialization of each episode, the agent is randomly placed in a navigable area and receives a target object category, such as "bed" or "sofa," as a navigation command. The agent is required to navigate to a position within 1 m of the target object within a maximum of 500 steps. The action space includes four discrete actions: move forward, turn left 30°, turn right 30°, and stop. This action setting is kept consistent across all reproduced baselines for fair comparison.

4.3 Ablation Experiment and Analysis

To validate the effectiveness of each component in the SSAG model, we first conduct ablation studies, quantitatively analyzing the contributions of the Graph Attention Network (GAT), the Product of Experts (PoE) mechanism, the Room Reward and Edge Features to the model performance. The results are presented in Table 2.

The experimental results indicate that the Graph Attention Network (GAT) has the most significant impact on model performance. Removing the GAT module leads to a 12.4% drop in the SPL metric, demonstrating that the capability for room-level topological reasoning is a critical factor allowing SSAG to achieve efficient navigation in complex environments. Secondly, removing the room-level reward r_{room} causes a SPL decrease of 10.3%, which shows that introducing an explicit semantic guidance reward plays an irreplaceable role in helping the agent learn stable long-horizon search strategies. Furthermore, the removal of the Product of Experts (PoE) fusion mechanism results in an 8.2% decline in SPL, verifying the necessity of effectively fusing semantic priors with geometric constraints to generate accurate long-term goal predictions. Finally, even the sole removal of Edge Features leads to a consistent 4.4% performance drop in SPL, indicating that explicitly modeling connectivity information between rooms provides additional performance gains for graph reasoning.

Based on the above, this article further investigates the impact of key hyperparameters on model performance, with a focus on the weight of the fusion λ and the number of layers of the neural network of the graph L. Regarding the influence of the weight of the fusion λ (as shown in Table 3), experiments reveal that the model achieves optimal performance when $\lambda = 0.6$. When λ is set too small, the policy relies overly on heuristic priors, leading to a limited generalization capability. Conversely, when λ is set too large, it can easily cause instability in reasoning during the initial stages when the graph structure construction is not yet complete.

Table 3: Impact of the fusion weight λ on model performance.

λ	0.2	0.4	0.6	0.8	1.0
SPL	0.305	0.328	0.340	0.332	0.318
Success	0.638	0.665	0.684	0.671	0.652

Regarding the impact of the number of layers in the graph neural network L (as shown in Table 4), experiments indicate that $L = 2$ is the optimal configuration. In this setting, the model can effectively capture two-hop adjacency relationships (e.g., the topological connection of "bedroom–hallway–bathroom"), thereby enhancing global perception capabilities. However, when the number of layers increases further, the model exhibits the common over-smoothing phenomenon of graph neural networks, leading to homogenized node features and consequently a degradation in performance.

Table 4: Impact of the number of graph neural network layers L on model performance.

L	1	2	3	4
SPL	0.318	0.340	0.335	0.322
Success	0.652	0.684	0.678	0.661

4.4 Comparative Experiments and Analysis

This paper compares SSAG with six representative baseline methods, including Active Neural SLAM (ANS) [10], SemExp [12], Frontier [19], SemGO [20], RGN [21] and PONI [27]. All methods adopt a Mask R-CNN model pre-trained on COCO2014 as the semantic perception backbone and use consistent simulation environment settings and success criteria. For fairness, the reproduced baselines use the same RGB-D input, HFOV of 79°, image resolution of 640 × 480, semantic map resolution of 5 cm/pixel, map coverage of 24 m × 24 m, maximum episode length of 500 steps, and stop-distance success criterion as SSAG. To comprehensively and objectively quantify the overall performance of the agent in the goal-oriented navigation task, this document selects the Success Rate (SR), Success weighted by Path Length (SPL) and Distance to Success (DTS) as the core evaluation metrics.

The general comparative results for the goal-driven navigation task are presented in Table 5. The experimental results in the Gibson data set demonstrate that the proposed SSAG method achieves optimal performance in all three core metrics: SPL, Success, and DTS. Specifically, SSAG achieves an SPL of 0.340, representing improvements of approximately 4.9% and 15.3% compared to SemGO (0.324) and SemExp (0.295), respectively, indicating higher path efficiency. In terms of task reliability, SSAG achieves a success rate of 0.684, significantly outperforming RGN (0.633). This shows that explicitly modeling room topological relationships through the Graph Attention Network effectively enhances the agent's ability to locate targets in complex scenes. SSAG simultaneously reduces DTS to 1.529 m, the lowest among all methods, demonstrating a more precise target localization accuracy compared to modular methods such as PONI (1.865 m). In contrast, methods relying solely on geometric information, Frontier and Active Neural SLAM, are significantly limited in performance due to their lack of high-level semantic reasoning capabilities. Their SPLs are only 0.124 and 0.145, respectively, with success rates below 45%. In general, the performance gain of SSAG stems primarily from its ability to exploit environmental topological structures and effectively fuse semantic heatmaps with geometric accessibility constraints via the Product of Experts (PoE) mechanism. This integration generates long-term goal points that are both semantically consistent and physically reachable, significantly mitigating the aimless search problem common in traditional methods when dealing with complex indoor layouts.

Table 5: Overall performance comparison on the Gibson dataset.

Method	SPL ↑	Success ↑	DTS (m) ↓
Frontier	0.124	0.403	2.432
Active Neural SLAM (ANS)	0.145	0.446	2.275
PONI	0.296	0.590	1.865
SemExp	0.295	0.623	1.653
RGN	0.302	0.633	1.623
SemGO	0.324	0.632	1.601
SSAG (Ours)	0.340	0.684	1.529

In the overall performance comparison, the performance of SemExp and SemGO is significantly superior to other baseline methods. Therefore, this paper conducts a more extensive comparative analysis between SemExp, SemGO, and SSAG along the dimensions of the target categories. The results are shown in Table 6. It can be seen that SSAG performs better in the vast majority of target categories, its advantage being most pronounced in target search tasks that involve strong spatial semantic associations.

Table 6: Performance comparison across different target categories on the Gibson dataset.

Category	SemGO		RGN		SSAG (Ours)	
	SPL	Success	SPL	Success	SPL	Success
TV	0.349	0.671	0.316	0.629	0.327	0.652
Chair	0.366	0.633	0.368	0.664	0.334	0.686
Couch	0.342	0.638	0.325	0.638	0.351	0.712
Bed	0.278	0.593	0.231	0.645	0.362	0.732
Toilet	0.295	0.596	0.287	0.632	0.357	0.725
Potted Plant	0.318	0.668	0.289	0.598	0.302	0.622

To further validate the advantages of SSAG in exploration efficiency and topological reasoning capability, this article reproduces SemExp and performs a comparative visual analysis with SSAG-PPO, with the results shown in Fig. 5. In a representative navigation episode targeting a Toilet, SSAG-PPO successfully completes the task and correctly executes the stop action at $t = 126$, whereas SemExp does not end up at $t = 153$. Similarly, the SemExp movement trajectory (solid red line) in the predicted map exhibits significant redundancy of the path and inefficient exploration, reflecting a decision-making process that remains substantially blind. In contrast, the SSAG exploration trajectory is shorter and exhibits stronger semantic consistency, demonstrating a more efficient search strategy.

This efficiency improvement primarily benefits from the commonsense topological reasoning capability provided by SSAG. When the agent observes a "bed" during exploration and infers that the current context is a "bedroom," it can leverage the spatial adjacency priors encoded by SSAG to further deduce that the target "toilet" is more likely located in the "bathroom" area adjacent to the bedroom. Subsequently, the Product of Experts (PoE) fusion mechanism jointly models semantic plausibility and geometric accessibility constraints. This guides the long-term goal points to more accurately target potential regions, effectively avoiding the exhaustive and inefficient search behavior exhibited by SemExp, which stems from its lack of room-level topological relationship modeling and its primary reliance on local features. Although SSAG consistently improves SPL, Success, and DTS over representative baselines, the current validation remains confined to simulation. The gains over the strongest semantic baselines are moderate, and several failure modes still persist, including dead-end oscillation under the discrete action space, target misses caused by semantic perception uncertainty, and inaccurate room inference during very early exploration when the graph is incomplete. Evaluating SSAG under dynamic-scene settings, stronger embodied navigation baselines, and real-robot deployment is therefore an important direction for future work.

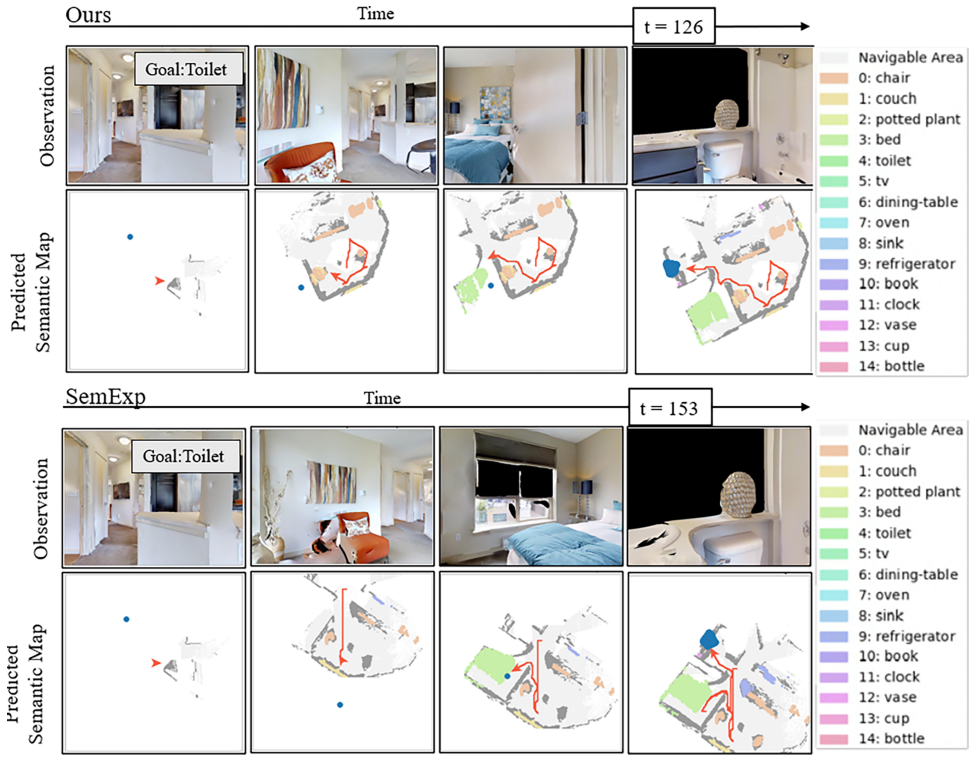

Figure 5: The figure shows an example comparing the experimental process of the proposed method with that of the SemExp method. Starting from the same location with the same target object "toilet", the proposed method finds the target object faster than the SemExp method.

5 Conclusion

This paper addresses the "geometric myopia" problem in mobile robots operating in complex indoor environments, which stems from a lack of high-level semantic understanding, by proposing and implementing a Situational Semantic Augmented Graph (SSAG) based active SLAM target navigation framework. The core idea of this method is to transform unstructured sensor observations into a structured graph representation rich in semantic and topological associations, and perform explicit semantic reasoning via Graph Neural Networks (GNNs), thereby providing the robot with effective priors for long-horizon exploration decisions.

Regarding the construction of the Situational Semantic Augmented Graph, this paper proposes an online heuristic room segmentation algorithm tailored for active SLAM, which discretizes the continuous navigation space into room instances with independent semantic attributes. By designing 31-dimensional node features (encompassing geometric morphology, semantic objects, exploration state, and room-type priors) and 8-dimensional edge features (encompassing doorway characteristics, spatial relationships, semantic similarity, and connectivity strength), explicit encoding of room-level semantic structure and topological relationships is achieved, laying the foundation for subsequent GNN reasoning. In terms of the semantic reasoning mechanism, this paper uses a Graph Attention Network (GATv2) to perform message passing and room-type inference on the room topology graph. By aggregating semantic information from adjacent rooms, the model can effectively capture two-hop adjacency relationships such as "bedroom–hallway–bathroom," enabling explicit prediction of room types in unexplored areas. Ablation studies show that removing the GAT module leads to a 12.4% drop in SPL, verifying the core value of the graph reasoning

mechanism for improving navigation performance. Regarding the fusion of the exploration strategy, this paper constructs a target selection strategy based on a Product of Experts (PoE) mechanism. The policy network output, semantic heatmap, and geometric accessibility constraints are treated as independent "experts" for probabilistic fusion. This ensures that the generated long-term goal points are both semantically plausible and physically reachable. Experiments indicate that this fusion mechanism improves the SPL metric by 8.2% compared to using the output of the policy network alone, effectively alleviating the issue of randomness of end-to-end policies during the initial training phase.

Experimental results in the Habitat simulation environment and on the Gibson dataset provide strong evidence for the effectiveness of the proposed method in simulation. SSAG achieves a Success weighted by Path Length (SPL) of 0.340, representing improvements of 15.3% and 4.9% over SemExp and SemGO, respectively. The overall task success rate reaches 68.4%. In particular, in search tasks for targets with strong spatial associations (e.g., "bed" and "toilet"), the success rates reach 73.2% and 72.5%, respectively, exceeding baseline methods by more than 20%. Visual comparative analysis further demonstrates that SSAG enables the robot to shift from blind exploration to structured semantic reasoning, significantly improving search efficiency and effectively validating the guiding value of semantic reasoning for target navigation. Nevertheless, the current study is still limited to simulation and largely static-scene assumptions. Future work will integrate dynamic-object filtering, broaden comparisons with more recent semantic navigation methods, and further evaluate SSAG on real robots and larger-scale indoor environments.

Acknowledgement: The authors are highly thankful to the Hubei Province Key Research and Development Special Project of Science and Technology Innovation Plan, to the Wuhan Key Research and Development Projects, to the open competition project for selecting the best candidates, Wuhan East Lake High-Tech Development Zone, to the Hubei Provincial Natural Science Foundation Guiding ProgramProjects, to the Fund for Research Platform of South-Central Minzu University.

Funding Statement: The authors are highly thankful to the Hubei Province Key Research and Development Special Project of Science and Technology Innovation Plan (No. 2023BAB087), to the Wuhan Key Research and Development Projects (No. 2023010402010614), to the open competition project for selecting the best candidates, Wuhan East Lake High-Tech Development Zone (No. 2024KJB328), to the Hubei Provincial Natural Science Foundation Guiding ProgramProjects [No. 2024AFC033], to the the Fund for Research Platform of South-Central Minzu University (No. PTZ25003).

Author Contributions: The authors confirm contribution to the paper as follows: Conceptualization, Shasha Tian, Zhengyang Chen; methodology, Shasha Tian, Zhengyang Chen; software, Chongwei Ruan, Mian Wu; validation, Shasha Tian, Zhengyang Chen; formal analysis, Shasha Tian, Kai Ren; resources, Kai Ren, Na Li; data curation, Zhijia Cui; writing—review and editing, Shasha Tian, Zhengyang Chen; visualization, Zhengyang Chen; supervision, Kai Ren, Na Li; project administration, Shasha Tian. All authors reviewed and approved the final version of the manuscript.

Availability of Data and Materials: The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

Ethics Approval: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Placed JA, Strader J, Carrillo H, Atanasov N, Indelman V, Carlone L, et al. A survey on active simultaneous localization and mapping: state of the art and new frontiers. IEEE Trans Robot. 2023;39(3):1686–705. doi:10.1109/TRO.2022.3224845.
2. Ahmed MF, Masood K, Fremont V, Fantoni I. Active SLAM: a review on last decade. Sensors. 2023;23(19):8097. doi:10.3390/s23198097.
3. Cadena C, Carlone L, Carrillo H, Latif Y, Scaramuzza D, Neira J, et al. Past, present, and future of simultaneous localization and mapping: toward the robust-perception age. IEEE Trans Robot. 2016;32(6):1309–32. doi:10.1109/TRO.2016.2624754.
4. Duan J, Yu S, Tan HL, Zhu H, Tan C. A survey of embodied AI: from simulators to research tasks. IEEE Trans Emerg Top Comput Intell. 2022;6(2):230–44. doi:10.1109/TETCI.2021.3132499.
5. Bourgault V, Makarenko A, Williams SB, Grocholsky B, Durrant-Whyte HF. Information-based adaptive robotic exploration. In: Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems. Piscataway, NJ, USA: IEEE; 2002. p. 540–5. doi:10.1109/IRDS.2002.1041456.
6. Yamauchi B. A frontier-based approach for autonomous exploration. In: Proceedings of the IEEE International Symposium on Computational Intelligence in Robotics and Automation. Piscataway, NJ, USA: IEEE; 1997. p. 146–51. doi:10.1109/CIRA.1997.613851.
7. Kulich M, Faigl J, Přeučil L. On distance utility in the exploration task. In: Proceedings of the IEEE International Conference on Robotics and Automation. Piscataway, NJ, USA: IEEE; 2011. p. 4455–60. doi:10.1109/ICRA.2011.5980200.
8. Bircher A, Kamel M, Alexis K, Oleynikova H, Siegwart R. Receding horizon next-best-view planner for 3D exploration. In: Proceedings of the IEEE International Conference on Robotics and Automation. Piscataway, NJ, USA: IEEE; 2016. p. 1462–8. doi:10.1109/ICRA.2016.7487277.
9. Zhang J, Tai L, Liu M, Boedecker J, Burgard W. Neural SLAM: learning to explore with external memory. arXiv:1706.09520. 2017.
10. Chaplot DS, Gandhi DP, Gupta S, Gupta A, Salakhutdinov R. Learning to explore using active neural SLAM. arXiv:2004.05155. 2020.
11. Niroui F, Zhang K, Kashino Z, Nejat G. Deep reinforcement learning robot for search and rescue applications: exploration in unknown cluttered environments. IEEE Robot Autom Lett. 2019;4(2):610–7. doi:10.1109/LRA.2019.2891972.
12. Chaplot DS, Gandhi D, Gupta A, Salakhutdinov R. Object goal navigation using goal-oriented semantic exploration. arXiv:2007.00643. 2020.
13. Cao Y, Hou T, Wang Y, Yi X, Sartoretti G. ARIADNE: a reinforcement learning approach using attention-based deep networks for exploration. arXiv:2301.11575. 2023.
14. Zhang F, Lin R, Liu L, Zhao M, Wu Q. Implementation of an autonomous exploration system in unknown environments based on transfer learning. Int J Adv Robot Syst. 2024;21(6):1–16. doi:10.1177/17298806241288669.
15. Li X, Tian B, Hou S, Li X, Li Y, Liu C, et al. Path planning for mount robot based on improved particle swarm optimization algorithm. Electronics. 2023;12(15):3289. doi:10.3390/electronics12153289.
16. Veličković P, Cucurull G, Casanova A, Romero A, Liò P, Bengio Y. Graph attention networks. arXiv:1710.10903. 2018.
17. Brody S, Alon U, Yahav E. How attentive are graph attention networks? arXiv:2105.14491. 2022.
18. Zhang L, Wei L, Shen P, Wei W, Zhu G, Song J. Semantic SLAM based on object detection and improved OctoMap. IEEE Access. 2018;6:75545–59. doi:10.1109/ACCESS.2018.2876798.
19. Yu B, Kasaei H, Cao M. Frontier semantic exploration for visual target navigation. arXiv:2304.05506. 2023.
20. Wu Y, Chen N, Rao L, Fan G, Yang D, Cheng S, et al. SemGO: goal-oriented semantic policy based on multi-headed self-attention for object goal navigation. In: 2024 27th International Conference on Computer Supported Cooperative Work in Design (CSCWD). Piscataway, NJ, USA: IEEE; 2024. p. 2931–6.
21. Luo JY, Gu Y. Target-driven navigation algorithm embedded with room category and boundary constraints. Comput Eng. 2025;51(4):85–96. doi:10.19678/j.issn.1000-3428.0069313.

22. Asgharivaskasi A, Atanasov N. Semantic octree mapping and Shannon mutual information computation for robot exploration. IEEE Trans Robot. 2023;39(3):1910–28. doi:10.1109/TRO.2022.3227357.
23. Tian C, Tian S, Kang Y, Wang H, Tie J, Xu S. RASLS: reinforcement learning active SLAM approach with layout semantic. In: 2024 International Joint Conference on Neural Networks (IJCNN). Piscataway, NJ, USA: IEEE; 2024. p. 1–8. doi:10.1109/IJCNN60899.2024.10650250.
24. Bescos B, Fácil JM, Civera J, Neira J. DynaSLAM: tracking, mapping, and inpainting in dynamic scenes. IEEE Robot Autom Lett. 2018;3(4):4076–83. doi:10.1109/LRA.2018.2860039.
25. Yu C, Liu Z, Liu X, Xie F, Yang Y, Wei Q, et al. DS-SLAM: a semantic visual SLAM towards dynamic environments. In: Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems. Piscataway, NJ, USA: IEEE; 2018. p. 1168–74. doi:10.1109/IROS.2018.8593691.
26. Esparza D, Flores G. The STDyn-SLAM: a stereo vision and semantic segmentation approach for VSLAM in dynamic outdoor environments. IEEE Access. 2022;10:18201–9. doi:10.1109/ACCESS.2022.3149885.
27. Ramakrishnan SK, Chaplot DS, Al-Halah Z, Malik J, Grauman K. PONI: potential functions for object-goal navigation with interaction-free learning. arXiv:2201.10029. 2022.