Doorway-Aware Frontier Scheduling for Indoor Mobile Robots Using RGB-D Geometric Screening and Temporal Semantic Anchors

Xiangyi Hu[1], Chunfang Liu[1,*], Hengyu Shi[2]
[1]School of Information Science and Technology, Beijing University of Technology, Beijing, China
[2]School of Computer Science, Beijing University of Technology, Beijing, China

Corresponding author: Chunfang Liu (email: cfliu1985@bjut.edu.cn).

ABSTRACT Indoor mobile robots must select navigation goals while operating with incomplete maps, but raw visual semantic detections can be too unstable to influence planning directly. This paper presents a doorway-aware frontier-scheduling framework that combines RGB-D geometric screening, temporal semantic-anchor maintenance, an incremental semantic-support layer, and anchor-dependent frontier-utility reweighting. Doorway candidates enter the planning loop only after depth-validity, local-plane, and geometric-consistency checks; accepted observations are then fused into persistent anchors whose confidence decays after missed updates. In ten paired same-condition simulations, the proposed configuration increased the mean semantic-support activation ratio from 5.11×10^{-4} to 1.447×10^{-3} and was higher in nine of the ten paired runs (two-sided exact sign test, p = 0.0215). The mean number of completed navigation goals increased from 0.10 to 2.50 per run, and all episodes terminated normally under the fixed evaluation protocol. Single-run demonstrations in two auxiliary layouts confirmed closed-loop executability but were not treated as statistical generalization evidence. The results support the feasibility of using geometrically screened, temporally maintained doorway cues to modify frontier scheduling; direct traversable-area coverage, detector accuracy, runtime, and real-robot performance remain outside the present evidence.

INDEX TERMS Indoor mobile robot, Frontier exploration; RGB-D perception, Doorway detection, Temporal semantic anchor, Semantic frontier scheduling

I. INTRODUCTION

Indoor mobile robots are increasingly deployed for service, inspection, search, and autonomous mapping in buildings whose maps are incomplete or unavailable. In such settings, the robot must choose navigation goals while simultaneously reducing geometric uncertainty. Doorways are especially important because they often mark room transitions and guide access to unexplored regions. If these structures are ignored, a robot may repeatedly expand local frontiers while delaying cross-room exploration. This makes doorway-aware scheduling a useful intermediate problem between purely geometric exploration and broad semantic navigation.

Frontier-based exploration remains a practical foundation because it is reproducible, compatible with occupancy-grid mapping, and easy to connect with standard navigation stacks [1-5]. Recent active SLAM, semantic mapping, and object-goal navigation studies further show that high-level cues can help robots prioritize more informative goals [6-14]. However, indoor semantic detections are noisy: a single RGB frame may produce false positives, missing detections, or detections that are geometrically inconsistent with depth and map evidence. Passing such raw outputs directly to the planner can therefore destabilize navigation, especially when a detector mistake is converted into a goal or a strong map prior.

Existing research has already explored information gain, active mapping, semantic maps, object-guided navigation, and vision-language frontier selection, providing important tools for robot decision making [10-14]. Nevertheless, many approaches pursue richer scene representations or target-object discovery, while fewer works examine the narrower mechanism of how a structural cue such as a doorway should enter a conventional frontier scheduler. The unresolved issue is not whether semantics are useful in general, but how to restrict semantic influence so that useful transition cues are retained and unsupported detections are suppressed.

To address this problem, this paper uses a conservative doorway-aware strategy. Doorway candidates are first generated from RGB observations, then screened by depth validity, local plane support, and geometric consistency. Only accepted observations are fused into temporal semantic anchors. These anchors are not treated as direct goals; instead, their confidence values update a semantic-support layer and

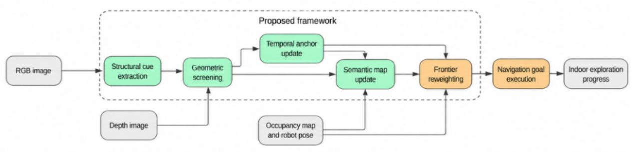

FIGURE 1. Doorway-Aware Frontier-Scheduling Framework

reweight nearby frontier utilities. In this way, the original frontier planner remains the geometric backbone, while doorway semantics provide bounded scheduling guidance.

The study evaluates the framework through ten paired same-condition simulations in a doorway-rich primary layout and two auxiliary single-run demonstrations. The paired design keeps the simulator reset, observation window, perception checkpoint, geometry gate, map update logic, and Nav2 execution stack consistent, while changing whether anchor utility contributes to frontier ranking. The evidence is therefore limited to the scheduling effect of screened temporal anchors rather than to detector training, physical floor-area coverage, or real-robot generalization.

The contributions are threefold. First, the paper proposes a closed-loop frontier-scheduling framework that introduces doorway cues only after RGB-D geometric screening. Second, it formulates temporal anchor-confidence maintenance, semantic-support updating, and anchor-dependent frontier reweighting within one planning loop. Third, it reports paired mechanism-level evidence and explicitly separates supported conclusions from claims that require additional coverage, runtime, detector-accuracy, and real-robot validation.

As shown in Figure 1, RGB-D observations are geometrically screened, maintained as temporal anchors, integrated into a semantic-support layer, and used to reweight frontier utilities before navigation-goal execution. This organization ensures that raw visual detections never act as navigation goals directly; unsupported detections are filtered, while stable doorway cues influence planning through a bounded utility term.

II. RELATED WORK

A. FRONTIER-BASED AND INFORMATION-GAIN EXPLORATION

Frontier-based exploration is the classical formulation for unknown-environment expansion [1]. Later work improved frontier detection efficiency, information-gain estimation, and map-coupled exploration policies [2–5]. These methods provide a strong foundation for indoor exploration, but they primarily evaluate candidate goals from geometric uncertainty and traversal cost. They identify where unknown space exists, yet do not explicitly distinguish a doorway-linked frontier from an ordinary local boundary.

B. ACTIVE SIMULTANEOUS LOCALIZATION AND MAPPING, SEMANTIC MAPPING, AND STRUCTURAL REPRESENTATION

Active simultaneous localization and mapping (SLAM) and semantic mapping connect environment modeling with decision making [10–12,15–31]. Representative directions include object-level SLAM [16], dense semantic mapping [17], multi-hierarchical semantic maps [26], and metric-semantic frameworks such as Kimera [30,31]. These studies show that semantics can be integrated into map estimation and planning, but many of them emphasize global consistency, reconstruction quality, or generalized semantic representation rather than doorway-centered indoor transition behavior. The present work is narrower and focuses on whether geometrically screened structural cues can improve exploration choices in a reproducible indoor loop.

C. SEMANTIC EXPLORATION AND MULTI-SENSOR PERCEPTION

Semantic exploration and object-guided navigation use higher-level cues to bias exploration toward more meaningful targets [6–9,13,14]. Door-specific RGB-D perception has also shown that two-dimensional appearance cues can be combined with depth information for real-time door detection and state classification on resource-constrained platforms [32]. Recent modular ObjectNav systems predict frontier potentials from partial semantic maps [33], combine open-vocabulary localization with classical exploration [34], construct language-grounded frontier maps [35], infer unobserved semantic context through generative maps [36], prompt online three-dimensional scene graphs to score frontiers and reject false positives [37], fuse three-dimensional structure for simultaneous exploration and target identification [38], or use multimodal goal embeddings for zero-shot navigation [39]. At the same time, multi-sensor perception studies highlight the need for consistency across image, depth, and map evidence [17,21,30]. The remaining system-level gap is how to introduce structural semantics into frontier scheduling without letting unstable single-frame detections dominate navigation. Compared with broader semantic-topometric exploration systems such as that of Fredriksson et al. [8], the present work is deliberately narrower: it focuses on doorway-specific structural cues, RGB-D geometric verification, temporal anchor persistence, and frontier-utility modification rather than a broader topometric semantic representation. This paper therefore addresses the gap by using only geometrically screened and temporally stabilized doorway observations as decision-relevant anchors, with a deliberately narrow focus on screened structural guidance in indoor exploration rather than broad semantic navigation.

III. MATERIALS AND METHODS

A. FRAMEWORK OVERVIEW AND PROBLEM SETTING

We consider a mobile robot exploring a partially unknown indoor environment. At time step t, the robot receives an

RGB image, a depth observation, an occupancy map, and its current pose:

$$X_t = \{I_t, Z_t, M_t^{\mathrm{occ}}, x_t\} \tag{1}$$

where X_t is the complete input set at time step t; I_t is the RGB image; Z_t is the depth observation; M_t^{occ} is the occupancy map; x_t is the robot pose. The system outputs verified semantic observations, a set of anchor states, a temporary navigation goal, and an updated semantic map:

$$Y_t = \{O_t^{\mathrm{ver}}, A_t, g_t, M_t^{\mathrm{sem}}\} \tag{2}$$

where Y_t is the system output set at time step t; O_t^{ver} is the set of geometrically verified semantic observations; A_t is the active temporal-anchor set; g_t is the temporary navigation goal; M_t^{sem} is the updated semantic-support map. The design objective is conservative. A navigation goal is never produced from a raw visual detection alone. Instead, the system lets structural semantics influence exploration only after geometric screening, temporal stabilization, and map-level integration.

B. GEOMETRICALLY SCREENED STRUCTURAL OBSERVATION GENERATION

Each visual candidate is represented as

$$d_i = (c_i, b_i, s_i^{\mathrm{det}}) \tag{3}$$

where d_i is the i-th visual detection candidate; c_i is its semantic class label; b_i is its image-space bounding box; s_i^{det} is its detector confidence.

The experiments focus on doorway observations. For each candidate, the corresponding depth region is checked by valid-depth ratio, local planar support, and geometric consistency. A candidate is promoted to a verified observation only when

$$r_i^{\mathrm{depth}} \geq \eta_d, \quad r_i^{\mathrm{plane}} \geq \eta_p, \quad s_i^{\mathrm{geo}} \geq \eta_g \tag{4}$$

where r_i^{depth} is the valid-depth ratio; r_i^{plane} is the local plane-fit inlier ratio; s_i^{geo} is the geometric-consistency score; η_d, η_p, and η_g are the minimum acceptance thresholds for depth validity, planar support, and geometric consistency, respectively.

The verified observation is written as

$$o_i^{\mathrm{ver}} = (c_i, p_i, s_i^{\mathrm{det}}, s_i^{\mathrm{geo}}) \tag{5}$$

where o_i^{ver} is the i-th verified structural observation; p_i is the recovered planar position in the map frame; c_i, s_i^{det}, and s_i^{geo} retain the class label, detector confidence, and geometric-consistency score defined above.

This step suppresses isolated image detections that are not supported by local geometry.

In the reported configuration, doorway candidates are generated by an Ultralytics YOLOv8 detector [40] with a YOLOv8n-scale weight file, an input size of 640 pixels, a confidence threshold of 0.36, and an intersection-over-union threshold of 0.45. Only the doorway class is used in the main comparison. To reduce visually implausible detections before geometric verification, the detector output is filtered by doorway-specific image heuristics, including aspect-ratio, width-ratio, and edge-touch constraints. After detection, a square depth patch is extracted around the box center using a scale factor of 0.68, clamped to a 9-81 pixel support window. Depth values are accepted only within 0.35-10.0 m, and a candidate must retain at least four valid depth samples with a valid-depth ratio not smaller than η_d=0.08.

In this manuscript, “verified” is an operational label meaning that a candidate passed the implemented RGB-D geometry gate; it does not imply ground-truth validation of the detector. The YOLOv8n-scale checkpoint was treated as a fixed project component and was held constant in every compared run. The records available for manuscript preparation do not include the checkpoint’s original training/validation split, epoch history, or standalone precision, recall, and mAP. Detector training and cross-dataset detector generalization are therefore outside the scope of the reported claims.

Planar support is then estimated by back-projecting the valid patch pixels into 3D and fitting a local plane by singular-value decomposition. The plane-fit inlier test uses a 0.05 m residual threshold, and the inlier ratio must satisfy η_p=0.20. The geometric-consistency score used in the implementation is a bounded weighted combination of the plane inlier ratio, the plane-fit error, the valid-depth ratio, and the plane orientation:

$$s_i^{\mathrm{geo}} = 0.4 r_i^{\mathrm{plane}} + 0.3 s_i^{\mathrm{rmse}} + 0.2 s_i^{\mathrm{depth}} + 0.1 s_i^{\mathrm{orient}} \tag{6}$$

where s_i^{geo} is the weighted geometric-consistency score; r_i^{plane} is the plane-fit inlier ratio; s_i^{rmse} is the normalized plane-fit residual score; s_i^{depth} is the normalized valid-depth score; s_i^{orient} is the bounded plane-orientation compatibility score; RMSE_i is the plane-fitting root-mean-square error; ϵ_p is the plane residual scale.

Here, the normalized residual and depth terms are computed using RMSE_i and the valid-depth ratio, while the orientation term increases with compatibility to a vertical doorway-support plane. A candidate is accepted only when the geometric-consistency score reaches the threshold η_g = 0.18.

C. TEMPORAL ANCHOR STATE UPDATE

To avoid treating every verified observation as an independent event, the system maintains a persistent anchor state for each matched structural cue. For anchor k, the maintained state is

$$a_k^t = (l_k, p_k^t, q_k^t, v_k^t, m_k^t) \tag{7}$$

where a_k^t is the state of anchor k at time step t; l_k is its semantic class label; p_k^t is its planar position; q_k^t is its confidence; v_k^t is its current visibility state; m_k^t is its consecutive missed-update count. The confidence of an incoming verified observation is first fused as

$$s_i = w_d s_i^{\text{det}} + w_g s_i^{\text{geo}}, \quad w_d + w_g = 1 \tag{8}$$

where s_i is the fused confidence of verified observation i; w_d and w_g are the detector and geometry fusion weights; s_i^{det} and s_i^{geo} are the detector confidence and geometric-consistency score; $w_d + w_g = 1$ constrains the fusion to a convex weighted combination.

In the reported implementation, the detector and geometry fusion weights are 0.6 and 0.4, respectively. Therefore, the fused anchor-input score remains within [0, 1] when both component scores are bounded within [0, 1].

A verified observation is matched to an existing anchor when both class and spatial consistency are satisfied:

$$l_k = c_i, \quad \|p_k^{t-1} - p_i\|_2 < \tau_m \tag{9}$$

where l_k and c_i are the anchor and observation class labels; p_k^{t-1} is the previous planar position of anchor k; p_i is the verified observation position; $\|\cdot\|_2$ denotes Euclidean distance; τ_m is the spatial matching radius.

If a match exists, anchor confidence is updated by exponential smoothing:

$$q_k^t = \lambda q_k^{t-1} + \beta s_i \tag{10}$$

where q_k^t is the updated anchor confidence; q_k^{t-1} is the previous confidence; λ is the temporal-memory coefficient; β is the new-observation weight; s_i is the fused confidence of the matched observation.

If no match exists, a new anchor is initialized as

$$q_k^0 = \beta s_i \tag{11}$$

where q_k^0 is the initial confidence of a newly created anchor; β is the initialization weight; s_i is the fused confidence of the unmatched verified observation.

When an anchor is not observed, confidence is softly decayed after a short delay:

$$q_k^t = \gamma q_k^{t-1}, \quad \text{for } m_k^t \geq m_0 \tag{12}$$

where γ is the missed-observation decay factor; q_k^{t-1} is the previous anchor confidence; m_k^t is the current missed-update count; m_0 is the delay threshold after which decay begins. For the main experiments, the matching radius is $\tau_m = 0.5$ m, the exponential-memory coefficient is $\lambda = 0.7$, the observation weight is $\beta = 0.3$, and the miss-decay factor is $\gamma = 0.98$. The implementation starts miss-based decay after $m_0 = 5$ unmatched update cycles, keeps an anchor marked as recently seen for up to 20 missed updates, activates an anchor when $q_k^t \geq 0.6$, and removes stale anchors after 80 missed updates. When a verified observation matches an existing anchor, the anchor position is overwritten by the latest verified planar position. This design keeps doorway cues persistent enough to influence frontier ranking, but not so persistent that they pollute later decisions.

D. SEMANTIC FRONTIER REWEIGHTING

Let f denote a frontier candidate with information gain I(f) and path cost C(f). The proposed scheduler augments the frontier utility with an anchor-dependent bonus:

$$U(f) = \lambda_I I(f) - \lambda_C C(f) + \lambda_A A(f) \tag{13}$$

where U is the utility assigned to frontier candidate f; $I(f)$ is its expected information gain; $C(f)$ is its path cost; $A(f)$ is the doorway-anchor semantic bonus; λ_I, λ_C, and λ_A are the corresponding gain, cost, and anchor weights. The semantic bonus is defined as

$$A(f) = \max_k \left(q_k^t \exp\left(-\frac{\|p_f - p_k^t\|_2}{\tau_d} \right) \right) \tag{14}$$

where $A(f)$ is the maximum semantic bonus for frontier f over all active anchors k; q_k^t is anchor confidence; p_f is the frontier position; p_k^t is the anchor position; $\|\cdot\|_2$ is Euclidean distance; τ_d is the spatial decay constant.

In the main comparison, the information-gain, path-cost, and anchor weights are 1.0, 0.5, and 3.0, respectively, and the anchor-distance decay constant is 2.0 m. The frontier selector retains a semantic-anchor target only when its utility exceeds the frontier-side reference by a ratio threshold of 1.2. A frontier closer to a high-confidence doorway anchor therefore receives a larger bonus. The target remains a frontier-like exploration goal rather than a raw semantic goal.

E. INCREMENTAL MAP UPDATE AND CLOSED-LOOP INTEGRATION

The map representation combines an occupancy layer, an anchor layer, and a semantic-support layer. When a verified observation falls into cell u, the cell-wise semantic support is updated as

$$M_t^{\text{sem}}(u) = (1 - \rho) M_{t-1}^{\text{sem}}(u) + \rho \tag{15}$$

where M_t^{sem} is the semantic-support layer at time step t; u is the map cell receiving the verified observation; M_{t-1}^{sem} is the previous semantic-support value at that cell; ρ is the hit-update coefficient that moves the cell value toward one. After each cycle, the semantic-support layer is softly decayed:

$$M_t^{\text{sem}} \leftarrow \delta M_t^{\text{sem}} \tag{16}$$

where δ is the global semantic-support decay factor; M_t^{sem} is the semantic-support layer after the current update; $\leftarrow$ indicates that every semantic-support cell is multiplied by δ once per cycle.

In the reported implementation, the hit-update coefficient is 0.2 and the map-wide decay factor is 0.99. If geometric occupancy and semantic support become inconsistent, the semantic value is reset toward a neutral state when the cell prior exceeds 0.7; the conflicting cell is then reassigned to 0.5 rather than preserved as a strong semantic prior. Figure 2 summarizes how verified observation generation, anchor

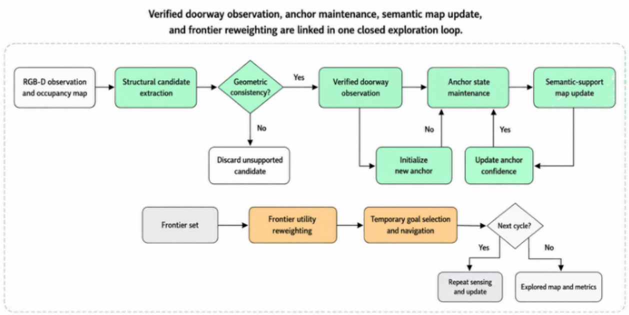

FIGURE 2. Closed-Loop Workflow of the Proposed Method

maintenance, map update, and frontier scheduling are coupled in the full loop.

Architecturally, the semantic branch remains localized: geometric screening is applied only to current detection candidates, anchor maintenance is restricted to the active anchor set, and semantic influence enters planning through frontier reweighting rather than through a separate global semantic planner. Runtime and computational-load measurements were not retained in the present experiment logs, so no claim is made about real-time speed or computational superiority.

As shown in Figure 2, doorway-candidate extraction, RGB-D geometric verification, anchor maintenance, semantic-support updating, frontier-utility reweighting, and navigation-goal execution are linked in one closed loop. The organization limits semantic influence to verified observations, persistent anchors, and a bounded utility bonus, so conventional frontier exploration remains available when candidate detections are rejected.

The parameter configuration in Table 1 emphasizes three properties: selective geometric acceptance, temporary rather than permanent anchor memory, and bounded semantic influence on frontier utility. Setting the baseline anchor weight to zero preserves the same perception and logging pipeline, so the paired comparison isolates the scheduling contribution of semantic-anchor reweighting.

F. REPRODUCIBILITY SCOPE AND CONTROLLED FACTORS

The retained experiment configuration identifies ROS 2 Humble, Gazebo, Navigation2, the TurtleBot3 Waffle simulation platform, the fixed YOLOv8n-scale checkpoint, and the principal parameters summarized in Table 1. The full machine-readable settings remain in the archived configuration files. For the paired comparison, the baseline and proposed configurations used the same simulator reset condition, evaluation window, perception checkpoint, sensor streams, geometry gate, semantic-map update logic, and Nav2 execution stack; the planning difference was that the anchor-utility weight was set to zero in the baseline. The records used for manuscript preparation do not preserve the exact camera calibration, Gazebo subversion, execution hardware, or original detector-training history. Reproducibility claims are therefore restricted to the archived algorithmic parameters, reported protocol, configuration files, and processed experiment summaries.

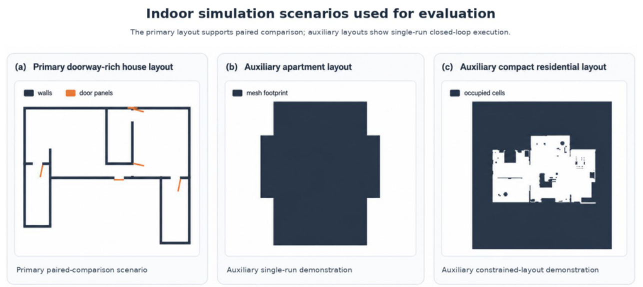

FIGURE 3. Indoor Simulation Scenarios Used for Evaluation(a) the Primary Doorway-Rich House Layout Used for Paired Comparisons; (b) an Auxiliary Apartment Layout; and (c) an Auxiliary Compact Residential Layout Used for Single-Run Closed-Loop Demonstrations. Panels (a) and (b) Are Rendered from the Underlying Wall-Collision Geometry, Whereas Panel (c) Is Rendered from the Released Navigation Map

IV. RESULTS

A. SIMULATION PLATFORM AND VALIDATION SCENARIOS

The experiments were conducted in a controlled indoor RGB-D mobile-robot simulation stack built on ROS 2 Humble, Gazebo, and Navigation2 (Nav2) [41–44]. All trials used a TurtleBot3 Waffle platform equipped with the doorway-oriented perception module. The evaluation includes a primary doorway-rich house layout and two auxiliary indoor layouts. The primary scenario supports the paired same-condition comparison because it contains repeated room transitions and doorway structures. The apartment and compact residential layouts are used only to demonstrate closed-loop execution in different geometries. Accordingly, the quantitative claims are restricted to the primary scenario, while the auxiliary cases are reported as single-run feasibility evidence. Figure 3 shows the three layouts.

Figure 3 separates the evidence roles of the layouts. The doorway-rich layout supports the paired comparison, whereas the two auxiliary layouts test only whether the same closed loop remains executable under different geometry; they are not used to claim cross-layout statistical superiority. The evaluation protocol and metric roles are summarized in Table 2.

Table 2 therefore keeps the evaluation focused on mechanism-level evidence. The same-condition design reduces confounding from start state, observation window, and execution stack, allowing the observed differences to be attributed primarily to the anchor-dependent scheduling term. The semantic-support activation ratio is defined as the maximum number of map cells with nonzero semantic-support values divided by the full occupancy-grid canvas during an episode. It is a mechanism-level indicator of how strongly the semantic branch is activated; it is not a direct percentage of traversable floor area explored. Completed navigation goals denote goal executions that reached the completion

TABLE 1. Principal Implementation Parameters Used in the Reported Configuration

Group	Parameter	Value	Meaning
Detector	confidence threshold	0.36	minimum retained detector confidence
Detector	intersection-over-union threshold	0.45	nonmaximum-suppression overlap threshold
Geometry	depth range	0.35–10.0 m	accepted RGB-D depth interval
Geometry	η_d	0.08	minimum valid-depth ratio
Geometry	η_p	0.20	minimum plane-inlier ratio
Geometry	ϵ_p	0.05 m	plane-fit residual threshold
Geometry	η_g	0.18	minimum geometry-consistency score
Belief	τ_m	0.5 m	anchor matching distance
Belief	activation threshold	0.6	minimum belief value for an active anchor
Scheduler	$\lambda_I, \lambda_C, \lambda_A$	1.0, 0.5, 3.0	frontier gain, cost, and anchor weights
Scheduler	τ_d	2.0 m	anchor-distance decay constant
Mapping	δ	0.99	semantic-support decay factor
Protocol	baseline anchor weight	0	semantic outputs logged but excluded from frontier utility

TABLE 2. Experimental Protocol and Roles of the Reported Metrics

Item	Setting	Purpose
Robot platform	TurtleBot3 Waffle with RGB-D sensing	Reproducible indoor mobile exploration
Simulation stack	ROS 2 + Gazebo + Nav2	Closed-loop exploration and execution
Primary scenario	Doorway-rich house layout	Main same-condition comparison with doorway transitions
Auxiliary scenario A	Apartment layout	Related-layout closed-loop demonstration
Auxiliary scenario B	Compact residential layout	Constrained-layout closed-loop demonstration
Comparison protocol	Same-condition baseline vs. proposed method	Same start, reset rule, and observation window
Parameter policy	Fixed within each compared configuration in the primary protocol	Reduces episode-specific tuning bias
Main metrics	Semantic-support activation ratio, completed navigation goals, verified observations, and semantic-cell updates	Mechanism-level activation, goal execution, and semantic-branch activity
Additional analysis	Parameter-sensitivity variants	Examines trade-offs when geometry gating or anchor influence is reduced

state within a run. Verified observations count doorway candidates that passed the geometric gate, and semantic-cell updates count cumulative updates to the semantic-support layer. A direct occupancy-grid area-coverage measure was not available in the current run logs, so the reported activation ratio is not used to claim physical exploration coverage.

B. EXPERIMENTAL PROTOCOL AND QUANTITATIVE RESULTS

The main comparison follows a paired same-condition protocol: the baseline and proposed configurations start from the same reset condition and run for the same evaluation window in the primary scenario. The baseline uses the same perception, screening, and logging chain for diagnostic comparability, but semantic outputs do not modify frontier utilities because the anchor weight is set to zero. This design isolates the effect of anchor-dependent frontier reweighting from differences in initialization, perception execution, or stopping conditions. Parameters are fixed across runs rather than retuned per episode. Pairing refers to matched simulator reset and evaluation conditions; the retained records do not establish identical random-number seeds beyond that reset protocol. Because the sample size is moderate and paired outcomes are not assumed to be normally distributed, the analysis reports means, standard deviations, medians, paired differences, pairwise direction, and a two-sided exact sign test. Table 3 reports one closed-loop demonstration in each auxiliary layout; these runs establish executability only and are not treated as statistical transfer or generalization evidence. Table 4 summarizes the paired quantitative comparison.

Table 3 confirms that semantic-support updates and completed goals were obtained in both auxiliary layouts. Because each auxiliary case contains only one run, these results support system executability but do not establish that the proposed method is statistically better in those layouts.

Table 4 provides the strongest evidence for the proposed method: the mean activation ratio was 2.83-fold the baseline mean, the proposed configuration won 9 of 10 matched pairs, and completed goals increased from 0.10 to 2.50 per run. These gains show stronger and more consistent coupling between screened doorway cues and goal execution, although they should not be interpreted as direct floor-area coverage.

Figure 4 shows the paired semantic-support activation results in the primary scenario and shows both the run-wise values and proposed-minus-baseline paired differences.

TABLE 3. Auxiliary Single-Run Closed-Loop Demonstrations

Scenario	Demonstration mode	Activation ratio	Verified observations	Updated semantic cells	Completed goals	Reading
Auxiliary layout A	single-run closed-loop demonstration	0.000952	63	71	2	semantic support and navigation-goal execution were both observed
Auxiliary compact residential layout	single-run constrained-layout demonstration	0.000501	11	12	2	closed-loop execution observed under limited repeatability

TABLE 4. Main Paired Same-Condition Comparison

Metric	Definition	Baseline (n = 10)	Proposed method (n = 10)
Normal episode termination	Episodes that terminated normally under the fixed evaluation window	10/10	10/10
Activation ratio (mean ± SD)	Mean and standard deviation of the maximum semantic-support activation ratio	0.000511 ± 0.000726	0.001447 ± 0.000717
Median activation ratio	Median maximum semantic-support activation ratio	0.000227	0.001598
Verified observations/run	Mean number of doorway observations that passed geometric screening	30.1	81.2
Semantic-cell updates/run	Mean cumulative number of semantic-support map-cell updates	33.7	93.2
Completed navigation goals/run	Mean number of navigation goals that reached completed status	0.100	2.500
Paired mean activation difference	Mean proposed-minus-baseline paired activation difference	–	0.000936
Activation-winning pairs	Number of matched pairs with higher activation	1/10	9/10
Exact sign test	Two-sided exact sign test on paired activation direction	—	p = 0.0215

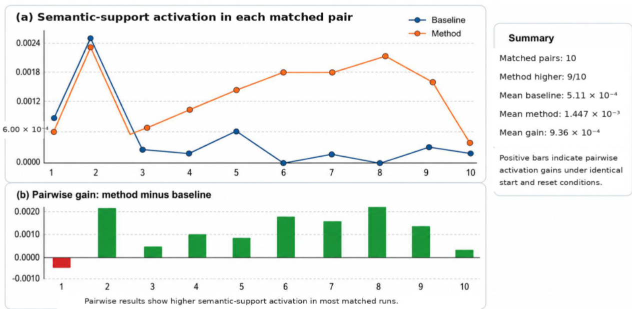

FIGURE 4. Paired Same-Condition Comparison of Semantic-Support Activation in the Primary Scenario: (a) Run-Wise Activation Ratios and (b) Proposed-Minus-Baseline Paired Differences. Positive Bars Indicate Higher Activation under the Proposed Configuration

The run-wise pattern in Figure 4 supports the aggregate result: positive paired differences occur in nine runs rather than being driven by a single outlier. This consistency is the main advantage demonstrated by the present experiment.

C. RESULT ANALYSIS

The primary quantitative result concerns semantic-layer activation and navigation-goal completion rather than physical floor-area coverage. In Table 4, the mean semantic-support activation ratio increased from 0.000511 to 0.001447, an absolute paired mean increase of 0.000936 and a 2.83-fold ratio relative to the baseline mean. The median increased from 0.000227 to 0.001598. Figure 4 shows that the effect is not confined to the aggregate mean. Verified doorway observations and semantic-cell updates increased by approximately 2.70-fold and 2.77-fold, respectively, showing that anchor-dependent reweighting changed closed-loop semantic activity rather than acting only as a passive annotation layer.

The paired direction is also consistent: the proposed configuration produced a higher activation ratio in 9 of 10 matched runs, yielding a two-sided exact sign-test result of p = 0.0215. Completed navigation goals increased from 0.100 to 2.500 per run, which is consistent with stronger coupling between screened doorway cues and goal execution. However, these metrics do not reveal how much traversable floor area was mapped, how far the robot travelled, or whether the detector produced false anchors. The two auxiliary runs in Table 3 only demonstrate that the closed loop executed in additional layouts; they are not used as a second statistical benchmark.

Table 5 presents a parameter-sensitivity analysis rather than a strict module-removal ablation. The variants preserve closed-loop operation while reducing either geometry-gate selectivity or anchor influence. The comparison is therefore diagnostic: it shows how the reported mechanism-level metrics shift under weaker gating or weaker semantic utility, but it does not establish that the full configuration is uniformly optimal for every objective.

TABLE 5. Parameter-Sensitivity Summary (N = 8)

Setting	Activation ratio ↑ (mean ± SD)	Semantic-cell updates ↑ (mean ± SD)	Completed goals ↑ (mean ± SD)	Reading
Baseline	0.000965 ± 0.001235	18.375 ± 41.012	0.250 ± 0.707	anchor weight set to zero; high variance
Proposed method	0.001180 ± 0.000692	34.750 ± 38.254	1.750 ± 1.581	reference full configuration
Reduced geometry gate	0.001565 ± 0.000272	60.125 ± 45.599	1.000 ± 1.195	higher semantic activation under weaker gating
Reduced anchor utility	0.001064 ± 0.000921	25.875 ± 28.109	2.250 ± 2.188	more completed goals but lower and less stable activation

Note: For completed goals, a higher value is preferable. For activation ratio and semantic-cell updates, a higher value indicates stronger semantic involvement, but not necessarily better overall navigation performance.

Table 5 shows a clear trade-off rather than a single uniformly best setting. Relative to the baseline, the proposed method increases semantic activation, semantic-cell updates, and completed goals, indicating stronger coupling between doorway cues and closed-loop execution. Reducing the geometry gate yields the highest activation ratio and the largest number of semantic-cell updates, which indicates more aggressive but less selective semantic propagation; because completed goals fall below the full method, this increase should not be interpreted as an unqualified improvement. Reducing anchor utility produces the highest mean number of completed goals, which is favorable for goal execution, but its activation ratio is lower and more variable, indicating weaker and less stable semantic influence. The full configuration should therefore be viewed as a reference trade-off among semantic involvement, selectivity, and goal completion, rather than as a setting that dominates every metric. A strict remove-one-module study with direct area coverage, path length, navigation time, goal-success rate, and false-anchor statistics is required for a stronger causal claim.

D. SUMMARY OF FINDINGS

Across the primary paired comparison, the proposed configuration produced higher semantic-support activation, more completed navigation goals, and normal termination in all evaluated episodes. The paired-win pattern and exact sign-test result indicate a consistent directional effect under the present protocol. The evidence establishes that screened doorway anchors influence frontier scheduling and closed-loop behavior; it does not yet establish improved traversable-area coverage or broad generalization across indoor layouts.

V. CONCLUSIONS

This study proposed a doorway-aware frontier-scheduling framework for indoor mobile robots. The framework combines RGB-D geometric screening, temporal semantic-anchor maintenance, semantic-support updating, and anchor-dependent frontier reweighting. Its central design is conservative: doorway observations do not replace the original frontier planner and do not become navigation goals directly. Instead, only geometrically verified and temporally maintained doorway cues provide a bounded utility bonus for nearby frontiers, so conventional geometry-driven exploration remains the backbone of the system.

The paired same-condition experiments show that this mechanism changed closed-loop behavior under the tested protocol. In ten paired simulations in the doorway-rich layout, the proposed configuration increased the mean semantic-support activation ratio from 5.11×10^{-4} to 1.447×10^{-3}, achieved a higher activation value in 9 of 10 pairs, and increased completed navigation goals from 0.10 to 2.50 per run. The auxiliary single-run layouts further confirmed closed-loop executability under different geometries, but they were not treated as statistical generalization evidence. Overall, the results support the mechanism-level feasibility of using screened temporal anchors to influence frontier scheduling.

The findings should be interpreted within clear limitations. The reported activation ratio is a semantic-branch activity indicator rather than direct traversable-area coverage. The strongest quantitative evidence is concentrated in one simulated doorway-rich scenario; detector precision, recall, false-anchor rate, path length, navigation time, goal-success rate, runtime cost, camera calibration effects, and real-robot robustness were not directly measured. In addition, the parameter study is diagnostic rather than a strict remove-one-module ablation. These limitations mean that the study demonstrates altered scheduling behavior, not complete superiority in exploration efficiency.

Future research should therefore extend the evaluation in several directions. First, repeated trials in multiple indoor layouts should be conducted with direct occupancy-grid area coverage, path-efficiency, recovery, and goal-success metrics. Second, the doorway detector and anchor layer should be evaluated with annotated ground truth so that false positives, missed detections, anchor drift, and temporal persistence errors can be quantified. Third, strict ablation studies should remove or replace the geometry gate, anchor memory, semantic-support map, and frontier bonus separately. Finally, real-robot experiments should test sensor calibration, illumination variation, computational load, and navigation safety in physical indoor environments.

DECLARATION OF CONFLICTING INTERESTS

The author(s) declared no potential conflicts of interest with respect to the research, author-ship, and/or publication of this article.

DATA SHARING AGREEMENT

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

FUNDING

The work was supported by the horizontal scientific research cooperation project with enterprise cooperation (400580002025B2).

REFERENCES

[1] D. Brugali, L. Muratore, and A. De Luca, “Mobile robots exploration strategies and requirements: A systematic mapping study,” The International Journal of Robotics Research, vol. 44, no. 9, pp. 1461–1506, 2025.

[2] R. Wang, J. Zhang, M. Lyu, et al., “An improved frontier-based robot exploration strategy combined with deep reinforcement learning,” Robotics and Autonomous Systems, vol. 181, p. 104783, 2024.

[3] A. Feng, Y. Xie, Y. Sun, et al., “Efficient autonomous exploration and mapping in unknown environments,” Sensors, vol. 23, no. 10, p. 4766, 2023.

[4] M. Luperto, M. M. Ferrara, M. Princisgh, et al., “Estimating map completeness in robot exploration,” Autonomous Robots, vol. 50, p. 6, 2026.

[5] C. Liu, D. Zhang, W. Liu, et al., “Enhancing autonomous exploration for robotics via real time map optimization and improved frontier costs,” Scientific Reports, vol. 15, p. 12261, 2025.

[6] S. Fredriksson, A. Saradagi, and G. Nikolakopoulos, “Robotic exploration through semantic topometric mapping,” in Proceedings of the 2024 IEEE International Conference on Robotics and Automation (ICRA). Yokohama, Japan: IEEE, 2024, pp. 9404–9410.

[7] L. Lu, Y. Zhang, P. Zhou, et al., “Semantics-aware receding horizon planner for object-centric active mapping,” IEEE Robotics and Automation Letters, vol. 9, no. 4, pp. 3838–3845, 2024.

[8] R. Zhang, H. M. Bong, and G. Beltrame, “Active semantic mapping and pose graph spectral analysis for robot exploration,” in Proceedings of the 2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Abu Dhabi, United Arab Emirates: IEEE, 2024, pp. 13787–13794.

[9] X. Liu, A. Prabhu, F. Cladera, et al., “Active metric-semantic mapping by multiple aerial robots,” in Proceedings of the 2023 IEEE International Conference on Robotics and Automation (ICRA). London, UK: IEEE, 2023, pp. 3282–3288.

[10] Y. Tao, X. Liu, I. Spasojevic, et al., “3D active metric-semantic SLAM,” IEEE Robotics and Automation Letters, vol. 9, no. 3, pp. 2989–2996, 2024.

[11] O. Alama, A. Bhattacharya, H. He, et al., “RayFronts: Open-set semantic ray frontiers for online scene understanding and exploration,” in Proceedings of the 2025 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Hangzhou, China: IEEE, 2025, pp. 5930–5937.

[12] Z. Yang, A. W. Y. Sang, M. A. V. J. Muthugala, et al., “Mutual information-based hierarchical NBV decision for active semantic visual SLAM under dynamic environments,” Scientific Reports, vol. 16, p. 5847, 2026.

[13] S. K. Ramakrishnan, D. S. Chaplot, Z. Al-Halah, et al., “PONI: Potential functions for ObjectGoal navigation with interaction-free learning,” in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). New Orleans, LA, USA: IEEE, 2022, pp. 18890–18900.

[14] N. Yokoyama, S. Ha, D. Batra, et al., “VLFM: Vision-language frontier maps for zero-shot semantic navigation,” in Proceedings of the 2024 IEEE International Conference on Robotics and Automation (ICRA). Yokohama, Japan: IEEE, 2024, pp. 42–48.

[15] J. A. Placed, J. Strader, H. Carrillo, et al., “A survey on active simultaneous localization and mapping: State of the art and new frontiers,” IEEE Transactions on Robotics, vol. 39, no. 3, pp. 1686–1705, 2023.

[16] H. Kanso, A. Singh, E. El Zarif, et al., “Semantic SLAM: A comprehensive survey of methods and applications,” Intelligent Systems with Applications, vol. 28, p. 200591, 2025.

[17] L. Miao, W. Liu, and Z. Deng, “A frontier review of semantic SLAM technologies applied to the open world,” Sensors, vol. 25, no. 16, p. 4994, 2025.

[18] B. Al-Tawil, T. Hempel, A. Abdelrahman, et al., “A review of visual SLAM for robotics: Evolution, properties, and future applications,” Frontiers in Robotics and AI, vol. 11, p. 1347985, 2024.

[19] B. Lv, et al., “Learning-based multi-robot active SLAM: A conceptual survey,” Applied Sciences, vol. 16, no. 3, p. 1412, 2026.

[20] C. Campos, R. Elvira, J. J. G. Rodriguez, et al., “ORB-SLAM3: An accurate open-source library for visual, visual-inertial, and multi-map SLAM,” IEEE Transactions on Robotics, vol. 37, no. 6, pp. 1874–1890, 2021.

[21] A. Rosinol, A. Violette, M. Abate, et al., “Kimera: From SLAM to spatial perception with 3D dynamic scene graphs,” The International Journal of Robotics Research, vol. 40, no. 12-14, pp. 1510–1546, 2021.

[22] R. Alqobali, R. Alnasser, A. Rashidi, et al., “A real-time semantic map production system for indoor robot navigation,” Sensors, vol. 24, no. 20, p. 6691, 2024.

[23] S. H. Allu, I. Kadosh, T. Summers, et al., “Autonomous exploration and semantic updating of large-scale indoor environments with mobile robots,” arXiv:2409.15493, 2024.

[24] T. S. Nguyen, H. N. Cao, and M. T. Pham, “Semantic potential field for mobile robot navigation using grid maps,” ETRI Journal, vol. 47, no. 3, pp. 422–432, 2025.

[25] J. G. Ramôa, V. Lopes, L. A. Alexandre, et al., “Real-time 2D-3D door detection and state classification on a low-power device,” SN Applied Sciences, vol. 3, p. 590, 2021.

[26] S. Y. Gadre, M. Wortsman, G. Ilharco, et al., “CoWs on pasture: Baselines and benchmarks for language-driven zero-shot object navigation,” in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Vancouver, BC, Canada: IEEE, 2023, pp. 23171–23181.

[27] S. Zhang, X. Yu, X. Song, et al., “Imagine before go: Self-supervised generative map for object goal navigation,” in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, WA, USA: IEEE, 2024, pp. 16414–16425.

[28] H. Yin, X. Xu, Z. Wu, et al., “SG-Nav: Online 3D scene graph prompting for LLM-based zero-shot object navigation,” in Advances in Neural Information Processing Systems. Vancouver, Canada: NeurIPS, 2024, pp. 37: 5285–5307.

[29] J. Zhang, L. Dai, F. Meng, et al., “3D-aware object goal navigation via simultaneous exploration and identification,” in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Vancouver, BC, Canada: IEEE, 2023, pp. 6672–6682.

[30] A. Majumdar, G. Aggarwal, B. Devnani, et al., “ZSON: Zero-shot object-goal navigation using multimodal goal embeddings,” in Advances in Neural Information Processing Systems. New Orleans, LA, USA: NeurIPS, 2022, pp. 35: 32340–32352.

[31] C. Wen, Y. Huang, H. Huang, et al., “Zero-shot object navigation with vision-language models reasoning,” arXiv:2410.18570, 2024.

[32] T. Chabal, S. Chen, J. Ponce, et al., “FOM-Nav: Frontier-object maps for object goal navigation,” arXiv:2512.01009, 2025.

[33] M. Dharmadhikari and K. Alexis, “Semantics-aware exploration and inspection path planning,” in Proceedings of the 2023 IEEE International Conference on Robotics and Automation (ICRA). London, UK: IEEE, 2023, pp. 3360–3367.

[34] M. Luperto, M. Tellaroli, M. Antonazzi, et al., “Multi-robot rendezvous in communication-restricted unknown environments via backtracking and semantic frontier-based exploration,” Robotics and Autonomous Systems, vol. 194, p. 105137, 2025.

[35] M. H. Sarfi and M. Bisheban, “Risk-sensitive autonomous exploration of unknown environments: A deep reinforcement learning perspective,” Journal of Intelligent & Robotic Systems, vol. 111, p. 36, 2025.

[36] J. Ding, Y. Zhou, H. Xia, et al., “An improved RRT* algorithm for robot path planning based on path expansion heuristic sampling,” Journal of Computational Science, vol. 67, p. 101937, 2023.

[37] S. Mukhopadhyay, H. Umari, and K. Koirala, “Multi-robot map exploration based on multiple rapidly-exploring randomized trees,” SN Computer Science, vol. 5, p. 31, 2024.

[38] B. Lindqvist, A. A. Agha-Mohammadi, and G. Nikolakopoulos, “Exploration-RRT: A multi-objective path planning and exploration framework for unknown and unstructured environments,” in Proceedings of the 2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Prague, Czech Republic: IEEE, 2021, pp. 3429–3435.

[39] M. Dobiš, J. Ivan, M. Dekan, et al., “A next-best-view method for complex 3D environment exploration using robotic arm with hand-eye system,” Applied Sciences, vol. 15, no. 14, p. 7757, 2025.

[40] Ultralytics, “Ultralytics YOLO,” [2026-07- 06]. [Online]. Available: https://github.com/ultralytics/ultralytics.

[41] Open Robotics, “ROS 2 documentation: Humble Hawksbill,” [2026-07-06]. [Online]. Available: https://docs.ros.org/en/humble/.

[42] Gazebo Project, “Gazebo documentation,” [2026-07-06]. [Online]. Available: https://gazebosim.org/docs.

[43] S. Macenski, T. Moore, D. V. Lu, et al., “From the desks of ROS maintainers: A survey of modern and capable mobile robotics algorithms in the robot operating system 2,” Robotics and Autonomous Systems, vol. 168, p. 104493, 2023.

[44] Navigation2 Project, “Navigation2 documentation,” [2026-07-06]. [Online]. Available: https://docs.nav2.org/.