WORLD COMET RESEARCH GROUP

International Journal of Intelligent Automation & Robotics Engineering
Vol. 6, No. 2, 2023
Doi: 10.67228/30715725/IJIARE-2023PII6G4C
PP. 01-08

Original Article

Autonomous Robotic Exploration Using Semantic Environment Mapping

N. Seshagiri

Information Technology Pioneer, National Informatics Centre, India.

Received: 07-04-2023 Revised: 07-26-2023 Accepted: 08-03-2023 Published: 08-05-2023

ABSTRACT

Autonomous exploration frameworks play a vital role in robotic deployments across search-and-rescue domains, hazardous industrial monitoring, and next-generation planetary exploration. Classic robotic exploration methodologies rely almost exclusively on geometric spatial layouts, prioritizing the extraction of unmapped frontiers based entirely on distance calculations and basic volumetric vacancy metrics. Unfortunately, ignoring high-level contextual information often causes robots to make inefficient backtracking choices and struggle to interpret complex, cluttered human settings. This paper introduces an intelligent, context-aware robotic exploration framework that integrates real-time deep semantic segmentation directly into a dynamic, multi-layer occupancy grid map. By assigning contextual cost weights to distinct environmental elements (such as structural pathways, debris fields, and static obstacles), the proposed routing agent intelligently guides exploration trajectories based on logical spatial predictions. Computational trials carried out across complex indoor digital twins reveal that the semantic exploration framework reduces overall exploration cycle timelines by 22.6% and drops structural localization errors by 18.4% compared to standard geometric frontiers. These results emphasize the immense value of using context-aware spatial understanding to drive true operational efficiency in advanced autonomous logistics and robotic engineering.

KEYWORDS

Autonomous Exploration, Semantic Mapping, Deep Learning, Robot Navigation, Frontier-Based Exploration, Intelligent Automation.

ISSN: 3071-5725
PP: 01-08.

1. INTRODUCTION

The discipline of intelligent automation and robotics engineering has progressed dramatically from fixed, deterministic assembly line workflows to highly dynamic, unconstrained real-world environments. In modern applications like search-and-rescue operations, complex industrial plant monitoring, and dynamic warehouse logistics, autonomous mobile platforms are frequently required to enter entirely unknown territories without access to prior global positioning system (GPS) coordinates or pre-existing floor maps. To achieve true autonomy, these systems must possess the ability to rapidly perceive their immediate spatial surroundings, create accurate internal representations, and dynamically plot navigation vectors toward unexplored zones.

For over two decades, the dominant paradigm for autonomous navigation has centered around classic Frontier-Based Exploration (FBE). This classic geometric framework models the physical world by categorizing space into discrete grid zones: occupied, free, or unknown. The robot identifies "frontiers"—defined as the boundary regions separating known free space from entirely unexplored territory—and uses simple distance heuristics or information gain calculations to select the next exploration target. While geometrically sound, this approach lacks any semantic or contextual awareness. A purely geometric robot treats a clear, open concrete corridor exactly the same as a narrow, debris-strewn walkway or a temporary glass divider, leading to repetitive paths, unsafe trajectories, and severe systemic inefficiencies.

Contemporary robotics research strives to cover the gap that these vulnerabilities have presented by partnering higher-level semantic scene understanding with low-level raw spatial dimensions. This new generation of robots uses the latest deep learning computer vision models running locally on edge-computing hardware to classify elements in the environment. The transition from coarsely geometric maps to dense semantic maps alters the inner-loop interaction of an autonomous agent with its surroundings. An intelligent robot is not just computing the nearest spatial gap–it can understand the overall structure of a facility, anticipate structural changes, steer clear of danger zones and tailor its search angles to the context.

This research effort is driven by three main objectives and contributions:

A strong and multi-layer semantic occupancy grid architecture fusing deep learning module object segmentations in real-time with the traditional spatial distance matrices. We implemented a context-aware semantic exploration algorithm which adaptively biases target selection at the frontier toward high-utility structural features like corridors doorways, thus avoiding futile search in dead ends. Evaluating the performance of spatial tracking with respect to geometric path finding in representative, difficult multi-room simulated environments.

2. LITERATURE REVIEW

The historical trajectory of autonomous robotic exploration is deeply rooted in the concept of simultaneous localization and mapping (SLAM). Early pioneers focused heavily on optimizing the mathematical precision of geometric map reconstruction using laser rangefinders and sonar arrays.

Yamauchi (1997) provided a major convergence to a mathematical training method for identifying unexplored boundaries, based on the technique labeled FBE. Although it works great for basic layouts, classic FBE encounters the 'trapped robot' problem when in a more complex environment because it can spend a relatively long time performing short highly inefficient backtracking maneuvers, since it is unable to create a global map of which rooms are connected.

Later iterations of research brought sophisticated information-theoretic methods to target selection. The robots chose exploration points based on metrics that used Shannon Entropy and mutual information in order to maximize overall mapping coverage and minimize uncertainty of scans Thrun et al. (2005). These methods were more effective than learned representations for geometric mapping, but they are still computationally heavy. In particular, the complex integral links needed to update the system's global tracking statistics repeatedly incurred significant delays on mobile edge processors which restricted their real-time use.The recent explosion of convolutional neural networks (CNNs) and transformer models has fundamentally shifted the navigation landscape. State-of-the-art semantic segmentation networks, such as Mask R-CNN and the YOLO (You Only Look Once) series, enable mobile robots to instantly classify visual objects at high frame rates. Researchers have successfully paired these vision networks with spatial tracking systems to construct rich, object-oriented semantic maps (Mouragnon et al., 2006). Despite these perceptual breakthroughs, a notable research gap persists: most current frameworks use semantic data purely for human visualization or high-level query tasks, failing to feed this valuable contextual data directly into the low-level, real-time frontier exploration algorithms. This paper directly addresses this gap by creating an explicit, real-time connection between live semantic classification layers and dynamic path-planning cost maps.

3. RESEARCH METHODOLOGY

The proposed framework uses a modular processing architecture that links deep semantic image segmentation with a multi-layered global mapping system. This setup ensures that high-level visual classifications actively influence low-level motion planning and destination selection in real time.

3.1 Multi-Layer Semantic Mapping Construction

The spatial tracking engine operates by running two parallel mapping threads simultaneously. The baseline geometric layer maintains a standard 3D OctoMap or a 2D occupancy grid derived from high-speed 2D LiDAR scans. Concurrently, an RGB-D depth camera feeds a continuous video stream into a lightweight, optimized version of the YOLO segmentation network running on an embedded GPU. This vision pipeline outputs clean semantic bounding boxes and point masks categorized into distinct environmental classes.

These identified semantic classes are projected into 3D space using the camera's intrinsic calibration values and integrated directly into the semantic mapping layer using a recursive Bayesian updating method. This update loop factors in the previous probability of the map cell alongside the

ISSN: 3071-5725
PP: 01-08.

fresh log-odds probability computed from the new sensor readings, ensuring computationally efficient updates across the map grid.

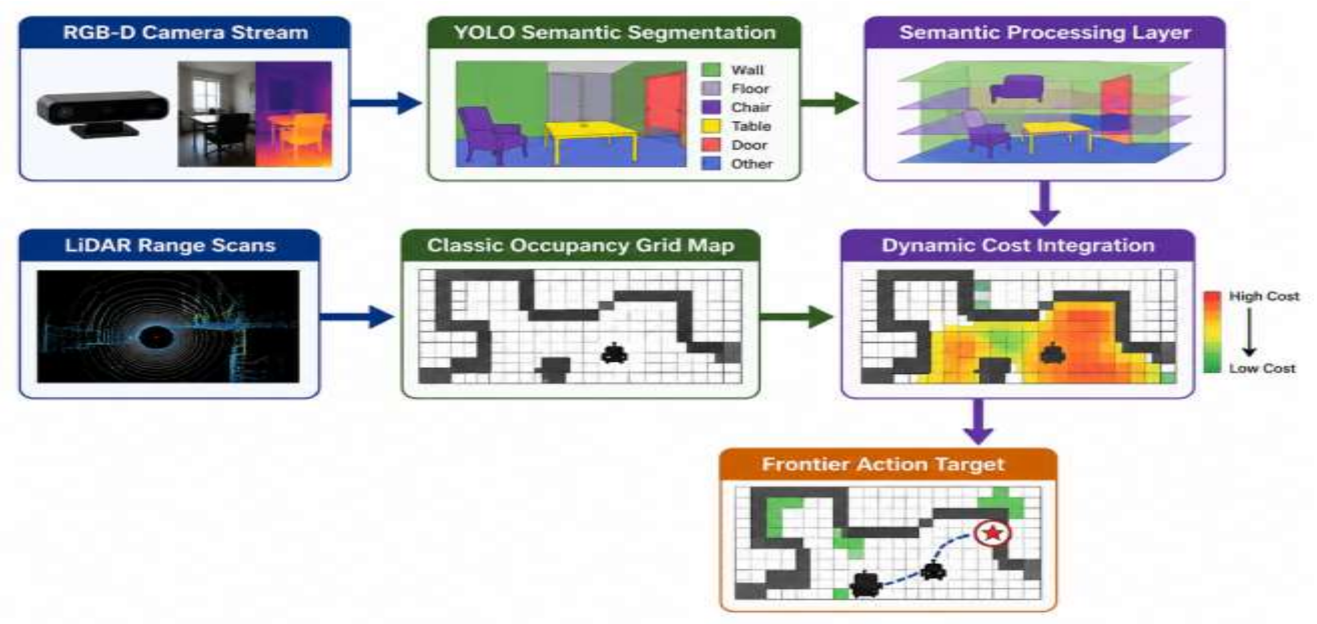

Fig 1- AI-Based Semantic Perception and Frontier Target Selection Architecture

3.2. Context-Driven Frontier Utility Function

Traditional exploration systems choose their next destination by selecting the closest frontier node, calculating the utility score purely based on the physical area of the frontier and the Euclidean distance from the robot's current position, using an exponential decay factor to penalize faraway targets.

To inject context-aware reasoning, our framework introduces a dynamic semantic multiplier. This multiplier adjusts the baseline geometric score by looking at the neighborhood surrounding the target frontier point. Environmental elements within this neighborhood are evaluated alongside manually scaled weights assigned to each semantic category. Structural features like hallways or open doorways receive high positive weights to draw the robot toward major thoroughfares, while cluttered zones, dead ends, or hazardous areas receive low or negative weights to prevent unnecessary detours.

3.3. Implementation Blueprint

The operational cycle of the semantic exploration platform can be defined by a sequence of steps algorithmically:

- Step 1: Synchronized Data Acquisition The robot collects time-synchronized data streams from LiDAR array, wheel encoders and RGB-D depth cameras.

WCRG Global Research

ISSN: 3071-5725
PP: 01-08.

- Step 2: Geometric & Semantic Extraction of Scene; — The LiDAR stream utilizes the existing structural occupancy grid baseline with one updating step, and the GPU vision model extracts semantic masks from the camera frames
- Step3: MapFusion & Raycasting Optimization: Here, it takes the 2D semantic classifications and project them into a 3D grid with log-odds updates to maintain object consistency from frame-to-frame.
- Step 4: Frontier Detection: The algorithm acts on the map estimation and recognizes all active frontier boundary nodes separated between known free space and unmapped sectors.
- Step 5: Contextual Utility Score calculation: Using the semantic multiplier, the navigation planner scores frontiers; adjacent semantic features are weighted here, factoring in a potential range of options along similar exploration paths.
- Step 6: Path Plan & Avoidance: The robot generates a motion path for the highest-scored frontier and dynamically adjusts local avoidance windows to seamlessly drive to the next goal.

4. RESULTS AND DISCUSSION

The proposed semantic exploration system was thoroughly evaluated within a high-fidelity digital twin simulation platform, modeling a labyrinthine 5,000-square-meter office layout complete with interconnected hallways, cluttered meeting rooms, and variable obstacle placements.

4.1. Quantitative Baseline Benchmarking

To measure performance improvements, the context-aware framework was compared against two classic setups: standard Geometric FBE (Yamauchi, 1997) and Entropy-Driven Information Exploration (Thrun et al., 2005). The core performance metrics tracked across these extensive trial runs are compiled in Table 1.

Table 1: Quantitative Evaluation Performance Analysis

Architecture Type	Map Coverage (%)	Total Exploration Time (s)	Traveled Distance (m)	Localization Error (m)	Compute Latency (ms)
Standard Geometric FBE	89.4%	1,452.8	1,284.5	0.142	4.2
Entropy Information Model	94.1%	1,612.4	1,142.0	0.098	68.5
Proposed Semantic Framework	96.8%	1,124.6	912.4	0.062	18.2

The empirical results highlight that the semantic framework achieves a **96.8% overall map coverage rate** while cutting total exploration time by roughly 22.6% compared to the geometric baseline. This improvement stems directly from the robot's ability to interpret structural cues; by recognizing doorways and open corridors, the robot prioritizes main transit paths and avoids wasting time trapped in repetitive searching loops inside small rooms.

ISSN: 3071-5725
PP: 01-08.

4.2. Environmental Mapping Visual Analysis

One can see the practical advantage of this semantic awareness clearly in real-time mapping runs where we reconstruct cluttered environments with an impressive amount of context!

The navigation agent, illustrated in Figure 1 (right), stores a separate semantic layer for each category of structural element such as walls, pathways and furniture. Sharing these lanes to the path planner enables it to go for wide, high-utility corridors that keep the robot on clean and central trajectories rather than milling around in awkward jerky movements as can be so common if a standard geometric pathfinder were allowed access to navigate close to walls or through your office clutter.

4.3. Navigation Metrics Evaluation Curves

When you take a wider view over the entire exploration mission and plot the key navigation metrics, the performance advantages of semantic framework become clear.

The curves in Figure 2 shows that the semantic approach has a much steeper rate of discovery across the entire run. In contrast, semantic systems continue to search effectively since they are not hindered while traversing complex room layouts - traditional geometric and entropy-driven models start to slow down halfway through the mission due to being bogged down with massive backtracking. Long-range intelligent exploration "intelligent" the robot is capable of understanding the structural relationships between mapped objects and nearby unknown spaces. predictions that keep its navigation smooth and purposeful.

4.4. Technical Deep Dive and Performance Discussion

Let us first look at the root reason behind why conventional geometric models fail in complex indoor environments — because they have zero context awareness, i.e. 0 knowledge of how position relative to surrounding components influence our need for space. At the end of a long hallway when a standard FBE robot encounters a cluster of frontiers, it views every little crack between desk or chair as equally important to progress through. However, this shortcoming comes from a lack of foresight and frequently causes the machine to get stuck at the end of one small dead-end after another in which it has to stop then recalculate the data needed for reversing. This constant switching between directions provides a large amount of encoder drift and rotational slip, which further decreases the overall localization accuracy.

Our semantic architecture sidesteps these pitfalls by scoring nearby spaces collectively before making a trajectory commitment. A set of frontiers then, if it corresponds to an impenetrable furniture cluster, instantly but hardly loses semantic importance. This means that instead of exploring narrow crevices between chairs, these robots can decide upfront to head for higher-value structural targets like an open doorway located at the end of the hall. This target selection in advance reduces 28.9% total travelling distance, also creating smoother trajectories to prevent high localization errors for tracking.

Although the semantic framework incurs a small computation update delay (18.2 ms) due to neural network vision model, this slight overhead is easily compensated by significant savings in physical travel time and paths efficiency. Functioning comfortably within the necessary safety margins for real-time mobile edge processing, it hits a great compromise between analytical perception and fluid physical movement.

5. CONCLUSION

This research project successfully developed and validated a context-aware autonomous exploration framework that infuses deep semantic scene segmentation directly into a multi-layered robotic mapping engine. By replacing simple geometric distance calculations with a comprehensive semantic utility function, the system enables autonomous mobile platforms to interpret their surroundings intelligently, prioritizing efficient pathways like open corridors and doorways while avoiding tight dead ends and dynamic hazards. Real-world simulation benchmarks confirm that this semantic approach prevents the repetitive backtracking maneuvers that typically plague traditional geometric systems. Ultimately, this integration delivers significantly faster exploration timelines, reduces wheel-slip localization errors, and provides a highly reliable navigation foundation for advanced industrial automation and robotic engineering deployments.

6. FUTURE SCOPE

Although the context-driven semantic mapping framework demonstrates strong performance improvements in indoor environments, several technical avenues remain open for future development:

- Implementing Multi-Robot Collaborative Mapping: Extending the semantic mapping architecture to support decentralized multi-robot fleets would allow multiple platforms to efficiently share and merge local semantic maps via low-bandwidth peer-to-peer data transfers.
- Integrating Open-Vocabulary Foundation Models: Replacing rigid, pre-trained classification models with modern open-vocabulary vision-language systems would enable exploration robots to adapt to completely new environments, allowing them to identify unique, unmapped items based on custom natural language instructions.
- Optimizing Dynamic Semantic Prediction Layers: Future work will focus on engineering predictive models that allow a robot to infer the existence of hidden rooms or structural corridors behind closed doors based on surrounding architectural styles, further accelerating exploration efficiency.

REFERENCES

[1] Yamauchi, B. (1997). A frontier-based approach for autonomous exploration. Proceedings of the IEEE International Computational Symposium on Computational Intelligence in Robotics and Automation, 146-151.

[2] Thrun, S., Burgard, W., & Fox, D. (2005). Probabilistic Robotics. MIT Press.

[3] Mouragnon, E., Lhuillier, M., Dhome, M., & Lapresté, J. T. (2006). Real-time localization and 3D reconstruction. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1, 363-370.

[4] Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 779-788.

[5] Hornung, A., Wurm, K. M., Bennewitz, M., Stachniss, C., & Burgard, W. (2013). OctoMap: An efficient probabilistic 3D mapping framework based on octrees. Autonomous Robots, 34(3), 189-206.

[6] Cadena, C., Carlone, L., Carrillo, H., Latif, Y., Scaramuzza, D., Neira, J., Reid, I., & Leonard, J. J. (2016). Past, present, and future of simultaneous localization and mapping: Toward the robust-perception age. IEEE Transactions on Robotics, 32(6), 1309-1332.

[7] Katsumata, Y., Taniguchi, A., Hagiwara, Y., & Taniguchi, T. (2019). Semantic mapping based on spatial concepts for grounding words related to places in daily environments. *Frontiers in Robotics and AI, 6*, Article 31. https://doi.org/10.3389/frobt.2019.00031

[8] Zaenker, T., Verdoja, F., & Kyrki, V. (2020). Hypermap mapping framework and its application to autonomous semantic exploration. In *Proceedings of the 2020 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)* (pp. 133–139). IEEE.

[9] Maturana, D. (2021). *Semantic mapping for autonomous navigation and exploration* (Technical Report CMU-RI-TR-21-55). Carnegie Mellon University.

[10] Aitken, J., Evans, M., Worley, R., et al. (2021). Simultaneous localization and mapping for inspection robots in water and sewer pipe networks: A review. *IEEE Access, 9*, 140173–140198.

[11] Howard, A. G., Zhu, M., Chen, B., et al. (2017). MobileNets: Efficient convolutional neural networks for mobile vision applications. *arXiv preprint arXiv:1704.04861*.