Computers, Materials & Continua

Tech Science Press

https://doi.org/10.32604/cmc.2026.086270

REVIEW

A Systematic Review of Agentic AI for Autonomous Navigation: SLAM-Based Intelligent Agents

Mukesh Dalal[1,*], **Anterpreet Kaur Bedi**[1] and **Payal Mittal**[2]

[1]Department of Electrical and Instrumentation Engineering, Thapar Institute of Engineering & Technology, Patiala, Punjab, India
[2]Department of Computer Science and Engineering, Thapar Institute of Engineering & Technology, Patiala, Punjab, India

*Corresponding Author: Mukesh Dalal. Email: mukesh.dalal@thapar.edu

Received: 27 May 2026; Accepted: 01 July 2026; Published: 13 August 2026

ABSTRACT: Autonomous navigation poses a key challenge in Artificial Intelligence (AI), necessitating agents to plan and execute actions in complex, partially visible surroundings. Simultaneous Localization and Mapping (SLAM) facilitates autonomous navigation of robots and vehicle objects to construct an unfamiliar environment map while concurrently monitoring their inside position. This systematic review investigates the nascent convergence of agentic AI, defined by goal-oriented autonomy, with adaptive decision-making and reasoning, with SLAM-based navigation systems. This paper utilized Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) methodology, which concentrated on peer-reviewed articles published in (2017–2026), particularly in SLAM-based intelligent agents utilized for autonomous navigation. SLAM has transformed from a geometry-based localization framework into an advanced perceptual and reasoning paradigm for autonomous navigation. Recent advancements in agentic AI, semantic perception, multimodal learning, and embodied foundation models have facilitated autonomous agents in progressing from passive mapping to context-aware decision-making and goal-directed navigation. The emergence of agentic AI and embodied AI has revolutionized SLAM into a spatial world model that facilitates perception, memory, reasoning, and autonomous decision-making. Contemporary research emphasizes lifelong SLAM, collaborative multi-agent mapping, semantic world modelling, and the integration of Large Language Models (LLMs) and Vision-Language Models (VLMs) for intelligent autonomous agents. Consequently, SLAM has evolved from a localization instrument to an extensive cognitive framework facilitating advanced autonomous navigation systems.

KEYWORDS: SLAM; agentic AI; autonomous navigation; multi-agents; intelligent agents

1 Introduction

Autonomous navigation is the ability of a modular robotic system, autonomous vehicle, or intelligent agent to cautiously sense and locate its environment, leading to feasible route planning to execute its motion towards a defined target without any human intervention. Traditionally realized using a robotic framework comprising vision, SLAM, route planning, and low-level control, such conventional navigation stacks have facilitated reliable functionality in structured contexts, such as warehouses, factories, and road environments. Despite the widespread utilization of this architecture owing to its understandability and simplicity, it is limited by a significant amount of dependence between modules, resulting in a transmission of errors in perception or localization throughout the entire pipeline [1,2]. Further, these frameworks depend on static maps, predefined goals, and manual rules, which restrict their effectiveness in a real-time environment. Thus, there is a need for accurate autonomy beyond the classical navigation stack paradigm with the increasing complexity of environmental applications, such as in autonomous driving, disaster responses, or service robots.

Advanced autonomous agents in navigation stacks enhance the traditional perception-planning-control pipeline by combining adaptive decision-making, and reasoning. These agents integrate LLMs, VLMs, and reinforcement learning to dynamically pick operations, sense specific goals, and analyze semantic context instead of following predetermined navigation rules [3]. Compared to traditional navigation architectures, these systems can more successfully engage with unpredictable settings, simplify complicated tasks by breaking them down, and update plans online. A new generation of intelligent navigation systems that go beyond merely reactive navigation frameworks and display more autonomous, adaptive, and goal-oriented behavior has emerged as a result of this shift [4,5]. Agentic AI is one example of an intelligent system that can see its surroundings, determine goals, and make decisions without the need for human interaction [6]. Traditional autonomous navigation techniques can be effectively improved by agentic AI, which enables robots to reason about goals and modify their behavior in dynamic, unexpected surroundings rather than just following preset paths. A dynamic world model that provides spatial memory and environmental context for decision-making, SLAM serves as a localization and mapping tool in this setting. Agentic systems can use this representation to understand complex goals, decompose them into smaller tasks, and continuously optimize navigation techniques in response to changes and ambiguity in the environment [7]. Indeed, Agentic AI is becoming a paradigm shift in the autonomous navigation field, especially when combined with SLAM-based frameworks to build smart goal-oriented robotic systems. Recent progress in deep learning models such as VLMs, LLMs, etc., further enhances these attributes by enabling robots to understand instructions in natural languages and perform contextual reasoning [4]. This unification brings SLAM into the realm of agentic reasoning, linking low-level geometric mapping to high-level cognitive intelligence for optimal performance in real-world environments. SLAM is the imprinted memory of the agent, and provides a way for sustained situational awareness and interaction between symbolic reasoning and low-level control. SLAM-based agentic systems are actively researched for applications in autonomous driving, service robotics, and search-and-rescue, etc., where adaptability, autonomy and robustness are the key features. The current research has been focused on the integration of SLAM with world models and agentic reasoning, especially in systems where intelligent agents are actively asking questions, modifying, and utilizing maps to improve decisions related to navigation.

In the past three decades, research on SLAM-based intelligent agents has been substantially developed. Early probabilistic models of localization and mapping were proposed in the late 1980s and 1990s. Over the past few years, the progress has been concentrating on semantic SLAM, dynamic SLAM, and active SLAM, where robots not only build maps but also analyze objects, manage real-world environment conditions, and perform active exploration to reduce uncertainty [8–10]. In the current scenario, the combination of AI techniques such as deep learning and reinforcement learning has resulted in the development of SLAM-based intelligent agents that combine mapping with perception, learning and decision-making capabilities. Modern systems integrate semantic knowledge, multimodal perception and world-model representations, enabling robots to perform tasks in complex environments. Applications of these systems are evident in domains such as autonomous driving, service robotics, and exploratory tasks. However, open research questions remain on long-term autonomy, scalability, and robustness in dynamic situations. This paper presented a comprehensive systematic review of SLAM-based agentic autonomous systems following PRISMA guidelines. The main contributions of the paper are as follows:

1. The paper provides a systematic overview of how agentic AI-driven navigation systems have replaced traditional SLAM-based navigation stacks. It proposed a new taxonomy of agentic perception and world

modelling that combines multimodal representations with uncertainty-aware reasoning and memory-driven adaptation.

2. A detailed investigation of robust and active SLAM techniques is conducted, emphasizing how risk-aware decision-making and information-driven exploration can enhance navigation reliability.
3. A systematic viewpoint on experimental procedures and benchmarking on multi-agent SLAM, datasets, simulation environments, and evaluation metrics is presented.
4. The paper also discusses important domains of application and open research challenges, highlighting future directions for intelligent autonomous systems.

Furthermore, Fig. 1 shows the paper's overall layout and provides an overview of the key sections discussed in the review.

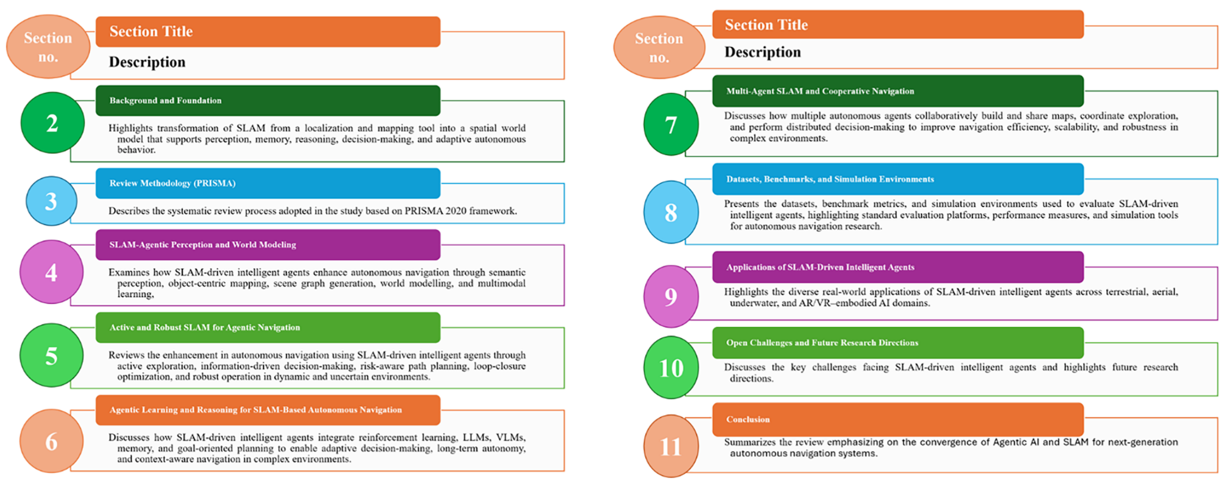

Figure 1: Overall layout of the review paper with the details of each section.

The rest of the paper is structured as follows: The background and foundations of agentic SLAM are given in Section 2. The review methodology is discussed in Section 3, and Section 4 introduces the proposed taxonomy of agentic perception and world modelling in SLAM. The function of robust and active SLAM for agentic navigation is examined in Section 5. Sections 6 and 7 discuss applications for Agentic learning and reasoning, with multi-agent SLAM and cooperative navigation. Details of datasets, benchmarks and simulation environments are discussed in Section 8, with applications described in Section 9. Section 10 highlights the open challenges and future research directions, and Section 11 concludes the systematic review.

2 Background and Foundations

For autonomous systems, SLAM is crucial, enabling agents to define their location and create maps of the environment. Classical SLAM guarantees precise localization by combining probabilistic approaches such as Extended Kalman Filters and graph-based optimization with geometric features. Traditional SLAM is embedded within a modular pipeline, consisting of perception, mapping, planning, and control, but those systems often assume static environments and have difficulties in dynamic, uncertain environments [11].

Recently, deep learning has been used for the development of semantic and object-aware SLAM, which improves the understanding of the environment with perception modules. These findings are integrated into a new paradigm for agentic AI systems, so that observation, action and learning can be carried out within an integrated loop. In agentic systems, based on SLAM perception, geometric and semantic information is exploited for planning and decision making with the aid of world models built by SLAM. Agentic SLAM is a system that allows an autonomous agent to make decisions in navigation, reasoning about goals, remembering past interactions, and selecting actions based on the environment. It learns to handle uncertainties and improve behavior from feedback. But adding semantic perception or learning policies does not in itself make a navigation system truly agentic. An agentic SLAM system is a complete system that integrates perception, world modeling, reasoning, planning, action execution, memory and adaptation into a single decision-making framework [12]. This method improves navigation reliability in dynamic environments and overcomes the limitations of traditional SLAM methods. The key elements of agentic SLAM are goal-directed behavior, reasoning and planning, memory and world-state representation, uncertainty-aware decision-making, closed-loop adaptation from feedback, and autonomous action selection to achieve tasks.

Fig. 2 shows the evolution of SLAM from classical geometric mapping methodologies to agentic SLAMs, showing the shift from passive localization to learning-oriented, intelligent decision-making autonomous agents. Traditional SLAM is about the robot pose estimation and geometric map construction [13]. Newer paradigms have made it more intelligent and autonomous. Visual SLAM has emerged as a way to perform localization and mapping through camera-based perception, i.e., the extraction, tracking and loop-closure detection of visual features [14]. Further, semantic SLAM enhances scene representations with object-level and contextual understanding, leading to richer scene representations [15]. Learning-based SLAM uses adaptive navigation policies based on reinforcement and imitation learning to increase adaptability in a broad set of environments [16]. Building on these ideas, Agentic SLAM integrates perception, semantic world modeling, reasoning, memory and goal-oriented decision making, transforming SLAM into an intelligent autonomous system [17]. This evolution allows it to perceive, reason, plan, act and adapt to dynamic environments. Semantic, active, learning-based and multi-agent SLAM are not independently form an Agentic SLAM, but they offer complementary capabilities that augment agentic autonomy. An agentic SLAM system integrates perception, world modeling, reasoning, memory, planning, autonomous action selection and continuous adaptation in a closed-loop decision-making framework.

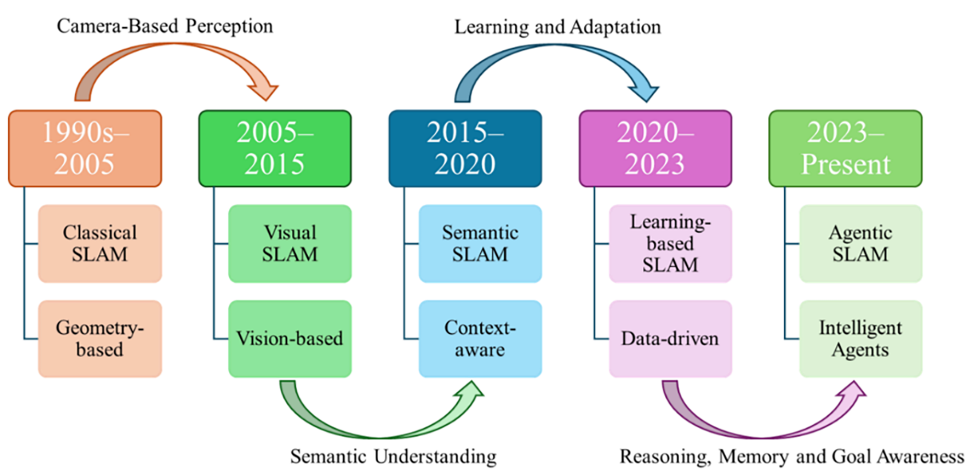

Figure 2: Evolution of SLAM-driven agentic autonomous navigation.

Further, Table 1 discusses the distinctions among SLAMs, ranging from classical to agentic.

Table 1: Comparison of the different SLAMs shift from perception to decision making.

Features	Classical SLAM	Visual SLAM	Semantic SLAM	Learning-Based SLAM	Agentic SLAM
Bases	Geometry based	Vision based	Context aware	Data-driven	Intelligent Agents
Key Focus	Localization mapping	Visual perception and mapping	Environment understanding	Data-driven models, learning & adaptation	Decision-making agents, autonomy & reasoning
Limitations /Challenges	No semantics, static assumptions [13]	Sensitive to light and texture [14]	No decision-making capability, limited reasoning [15]	Interpretability, data dependency [16]	Generalization, safety concerns [17]
Example	Fast-SLAM, grid mapping, probabilistic models	Mono-SLAM, Feature-based methods	Object SLAM, scene graphs, semantic maps, object detection	CNN-based SLAM, End-to-end models, policy learning	Active SLAM, LLM-based planning, World Models

3 Review Methodology (PRISMA)

This section delineates the systematic review technique employed in this study. The review adheres to the PRISMA 2020 guidelines to ensure reproducibility, transparency, and rigor in the selection and analysis of pertinent material [18].

The review paper is based on the following five research questions:

1. How is Agentic Perception and World Modeling utilized for SLAM-based autonomous navigation?
2. How do active and robust SLAM methods improve agentic autonomous navigation?
3. How do agentic learning and reasoning enhance decision-making, generalization, and long-term autonomy in SLAM-based navigation?
4. What are the datasets, benchmarks and simulation environments commonly used to assess the performance of SLAM-based agents?
5. What are the key applications of agentic SLAM for autonomous navigation, and what are the key challenges of the current agentic SLAM systems?

These questions define the structure of the review paper, and the following sections discuss them in detail for SLAM-based intelligent agents.

3.1 PRISMA Framework

This review paper utilized the PRISMA-2020 guidelines to conduct a comprehensive, systematic, and unbiased review [19]. The PRISMA framework is essential for offering a uniform methodology for several phases of the review, which encompass: Identification entails the aggregation of pertinent studies from various databases; Screening involves the elimination of duplicates and irrelevant records; Eligibility assesses studies according to established inclusion and exclusion criteria; and Inclusion culminates in the final selection of studies for qualitative analysis. The PRISMA flow diagram is illustrated in Fig. 3.

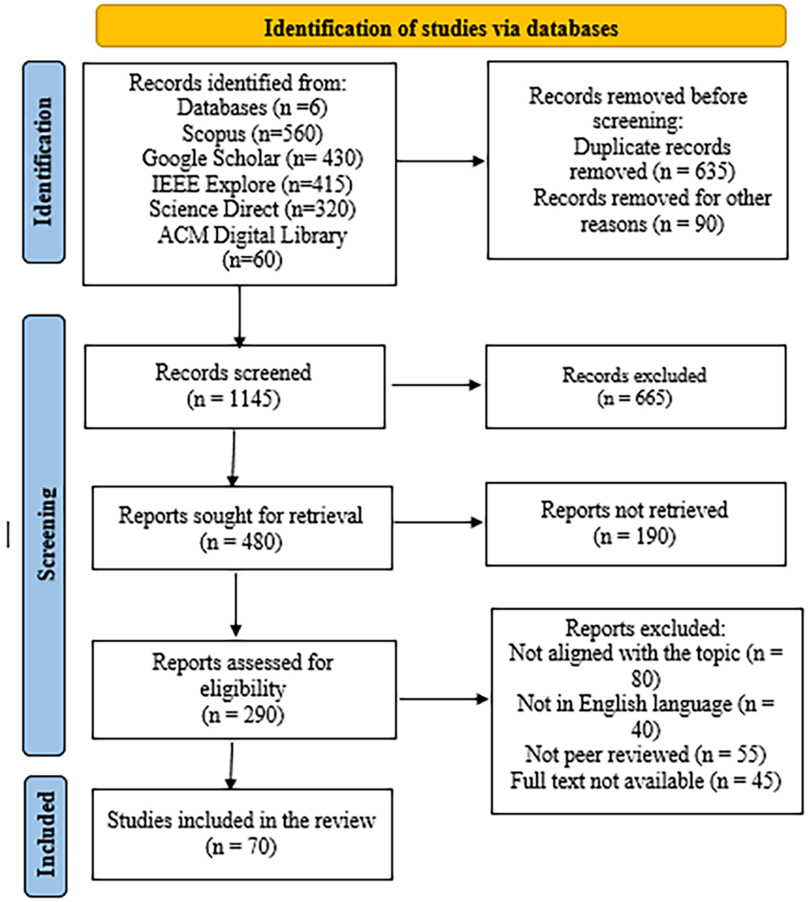

Figure 3: PRISMA 2020 flow diagram.

3.2 Search Strategy and Databases

The search was conducted across different databases, including IEEE Xplore, GoogleScholar, Scopus, Web of Science, ScienceDirect, arXiv, etc., to ensure extensive coverage of the pertinent studies. The total number of articles included in this systematic review is 141, among which 44 are conference articles, and 97 are journal articles, including books. Fig. 4 gives an overview of the number of articles chosen per year for this review paper. The search period was from 2017 to 2026, and the search execution was done from 10 February 2026, to 12 May 2026.

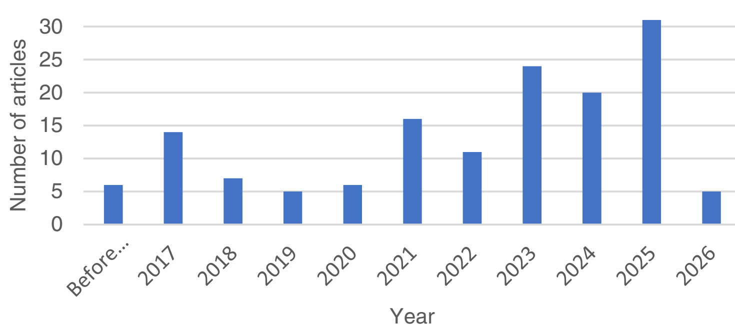

Figure 4: Number of articles per year chosen for the review.

The search was initiated with simple keywords SLAM, Agentic AI, Autonomous Navigation, etc. An extensive search was done by adding the synonyms of the keywords to enhance the coverage, and the search was performed using combinations of logical keywords such as "AND" and "OR". The search was maximized by adding the formal queries for database search engines and progressively enhancing them to optimize the retrieval efficiency. Further, boolean operators and filters were utilized to refine the obtained results based on the type of publication and relevance to the required field. Table 2 shows the search strings for different databases.

Table 2: Search strings for different databases for the literature review.

Database	Complete Search String
IEEE Xplore	("SLAM" OR "Agentic SLAM" OR "Simultaneous Localization and Mapping") AND ("Robot Navigation" OR "Autonomous Navigation") AND ("Active SLAM" OR "Semantic SLAM" OR "Object-aware SLAM") AND ("Agentic AI" OR "Intelligent Agents" OR "Embodied AI") AND ("Deep Learning" OR "Reinforcement Learning" OR "Vision-Language Models") AND ("CNN SLAM" OR "Transformer SLAM" OR "Multimodal Fusion SLAM" OR "Large Language Models").
Web of Science	TS = (("SLAM" OR "Agentic SLAM" OR "Simultaneous Localization and Mapping") AND ("Robot Navigation" OR "Autonomous Navigation") AND ("Active SLAM" OR "Semantic SLAM" OR "Object-aware SLAM") AND ("Agentic AI" OR "Intelligent Agents" OR "Embodied AI") AND ("Deep Learning" OR "Reinforcement Learning" OR "Vision-Language Models") AND ("CNN SLAM" OR "Transformer SLAM" OR "Multimodal Fusion SLAM" OR "Large Language Models")).
Scopus	TITLE-ABS-KEY (("SLAM" OR "Agentic SLAM" OR "Simultaneous Localization and Mapping") AND ("Robot Navigation" OR "Autonomous Navigation") AND ("Active SLAM" OR "Semantic SLAM" OR "Object-aware SLAM") AND ("Agentic AI" OR "Intelligent Agents" OR "Embodied AI") AND ("Deep Learning" OR "Reinforcement Learning" OR "Vision-Language Models") AND ("CNN SLAM" OR "Transformer SLAM" OR "Multimodal Fusion SLAM" OR "Large Language Models")).
Google Scholar	Keywords: "Agentic SLAM", "Agentic Navigation", "Semantic SLAM", "Active SLAM", "Learning-Based SLAM", "Autonomous Navigation", "Embodied AI", "Large Language Models for Robotics", "Vision-Language Navigation", "Multi-Agent SLAM".
ScienceDirect	("SLAM" OR "Agentic SLAM") AND ("Robot Navigation" OR "Autonomous Navigation") AND ("Active SLAM" OR "Semantic SLAM") AND ("Agentic AI" OR "Intelligent Agents" OR "Embodied AI") AND ("Deep Learning" OR "Reinforcement Learning" OR "Vision-Language Models") AND ("CNN SLAM" OR "Transformer SLAM" OR "Multimodal Fusion SLAM" OR "Large Language Models").
SpringerLink	("SLAM" OR "Agentic SLAM") AND ("Robot Navigation" OR "Autonomous Navigation") AND ("Active SLAM" OR "Semantic SLAM") AND ("Agentic AI" OR "Intelligent Agents" OR "Embodied AI") AND ("Deep Learning" OR "Reinforcement Learning" OR "Vision-Language Models") AND ("CNN SLAM" OR "Transformer SLAM" OR "Multimodal Fusion SLAM" OR "Large Language Models").

(Continued)

Table 2 (continued)

Database	Complete Search String
arXiv	(“Agentic SLAM” OR “Agentic Navigation” OR “Embodied AI”) AND (“SLAM” OR “Visual SLAM” OR “Semantic SLAM” OR “Navigation”) AND (“Transformer SLAM” OR “Large Language Models” OR “Vision Language Models” OR “Foundation Models” OR “Multi-Agent Systems”).

3.3 Inclusion and Exclusion Criteria

The inclusion and exclusion criteria were formulated to ensure the quality of the chosen literature as per the PRISMA guidelines, which directed the eligibility and screening stages of the review conducted. The inclusion criteria mandated that literature concentrate on SLAM-based autonomous navigation, integrating principles of agentic AI, encompassing LiDAR, optical, or multi-sensor SLAM systems, including decision-making and learning-oriented navigation. Moreover, research should focus on object-aware, semantic or deep learning-based perception inside SLAM frameworks, explore active SLAM in dynamic settings, and employ learning paradigms such as deep or reinforcement learning for diverse tasks. Furthermore, inclusion necessitated experimental validation in either simulated or real-world environments with established evaluation metrics. Additionally, eligible literature must be published in peer-reviewed journals or conferences, high-quality preprints, and must be accessible in English as full-text.

The exclusion criteria have been laid to ensure that the presented review focuses on agentic SLAM, autonomous navigation, and deep learning. Studies that solely investigated classical SLAM approaches without addressing navigation or decision-making, or those focused entirely on theoretical elements, were omitted. Furthermore, the exclusion literature focused primarily on low-level algorithmic advances, excluding navigation, which has no relevance to real-world applications. The review excludes works that have not been published in English and brief contributions, such as merely abstracts or a lack of full-text access. These criteria were developed to ensure that the chosen literature incorporates innovative approaches in agentic SLAM, emphasizing adaptability, robustness, and practical applicability.

3.4 Study Selection and Data Extraction

The study selection process followed the PRISMA guidelines to ensure transparency and reproducibility. The method consisted of several stages, such as identification, screening, eligibility and inclusion. A thorough search across multiple databases was conducted to eliminate duplicate records. The remaining research underwent a two-stage screening process. Initially, titles and abstracts have been examined for relevance to SLAM-based autonomous navigation, agentic AI, and learning-driven techniques, with irrelevant publications excluded. The second stage was a thorough evaluation of the text to determine compliance with established inclusion and exclusion criteria, with a focus on areas such as agentic perception, world modeling, SLAM methodologies, and decision making. The study selection process was conducted by the corresponding author and subsequently verified by the co-authors to ensure consistency. Disagreements over studies were resolved by discussion among the authors and the reviewers. Data extraction is done by using important attributes relevant to the presented systematic review to select the final literature. For better synthesis and comparative analysis relevant data was systematically extracted from each study in the literature. To ensure consistency throughout the literature, a consistent data extraction procedure was used by utilizing features as mentioned in Table 3.

Table 3: Major attributes used for data extraction process.

Attributes	Details
Bibliographic Information	Authors, publication year, venue (journal/conference)
SLAM Modality	Optical, Visual, LiDAR-based SLAM, or multi-sensor fusion approaches
Agentic Components	Presence of decision-making/reasoning, planning, or agent-based control mechanisms
Perception Strategy	Traditional feature-based approaches or deep learning methodologies (e.g., semantic segmentation, object detection, etc.)
Navigation Approach	Active SLAM, Passive SLAM, goal-oriented navigation, or exploration-centric methodologies
Learning frameworks	Reinforcement learning, Deep learning, constant learning, or hybrid methodologies
Evaluation Setup	Datasets used, simulation or real-world experiments, and assessment metrics
Robustness Characteristics	Management of dynamic environments, challenging situations (e.g., low illumination, inclement weather), and uncertainty modeling
Key Contributions and Challenges	Main innovations, enhancements in performance, and recognized problems and challenges

3.5 Quality Assessment and Bias Analysis

The presented review is based on PRISMA guidelines, and a rigorous bias and quality assessment was conducted to ensure the reliability and validity of the selected studies. The quality assessment evaluated each article using criteria such as objectives of the study, originality and technological contributions, dataset, experimental validation, and reproducibility [20]. These criteria were used to qualitatively grade each study.

Potential biases such as publication, dataset, methodological, model, and selection bias were identified by the bias research. For instance, dataset bias may develop from reliance on specific benchmarks, while publishing bias may lead to an emphasis on favorable results. The presented review combined publications from several databases, including modeling and empirical evaluations. It also conducted comparative analyses and critically assessed constraints to address these biases as per the predefined inclusion and exclusion criteria. This methodology facilitates an overview of the agentic SLAM techniques for autonomous navigation and ensures that the review conclusions are supported by reliable research.

Furthermore, the review protocol was not formally registered in a public repository. However, the review methodology, inclusion criteria, and screening process were predefined before conducting the literature search. Moreover, further details can be found in the Supplementary Materials.

4 SLAM-Agentic Perception and World Modeling

This section presents the proposed taxonomy of agentic perception and world modeling in Fig. 5 for SLAM-based intelligent navigation systems. In autonomous systems for SLAMs, agents perceive the information of the world model from the environment to develop and sustain themselves for action and decision-making.

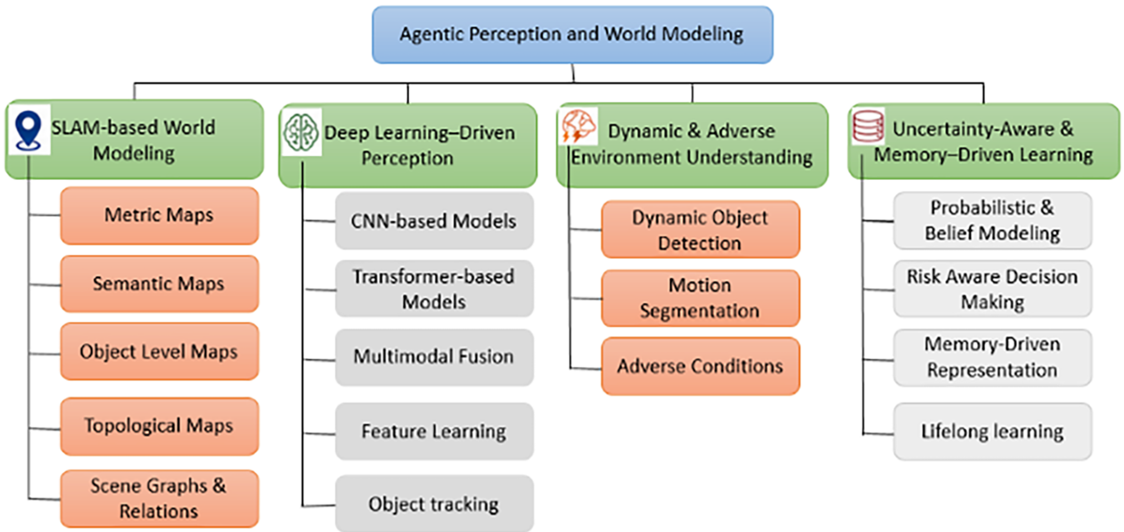

Figure 5: Proposed taxonomy of agentic perception and world modeling.

4.1 SLAM-Based World Modeling and Representations

Agentic SLAM-based world modeling and representation involves better mapping of the world as compared to traditional systems [7]. Many researchers started working in this field; one such example is RS-SLAM proposed by Ran et al. [21], a robust semantic localization in dynamic settings by the application of semantic filtering and robust feature association, leading to enhanced localization stability and less trajectory inconsistency. Further, an object-centric semantic SO-SLAM framework was presented by Liao et al. [22], which incorporated object-level constraints into the mapping procedure. The experimental results showed an 18% decrease in localization drift relative to conventional visual SLAM, resulting in improved robustness and interpretability by enabling agents to better understand their environments. Moreover, Tian et al. [23] enhanced semantic world modeling for collaborative multi-agent settings, achieving a 15%–20% improvement in distributed mapping consistency among robots via shared representations and collaborative semantic fusion. Additionally, Hughes et al. [24] proposed a hierarchical framework by employing spatial scene graphs and persistent semantic memory for world modelling. The results showed that the proposed framework attained less than 5 cm of localization drift and improved long-term comprehension of environments.

4.2 Deep Learning–Based Perception for Agentic SLAM

In agentic perception, deep learning is at the core for enhancing the capabilities of SLAM-driven agents by utilizing learning-based perception modules for understanding the semantic scene, object detection, depth estimation, and multimodal sensor fusion. Perception outputs from deep learning networks such as Convolutional Neural Networks (CNNs) and transformer-based models can be used with SLAM backends to develop object-aware and semantically enriched maps [25]. The integration provides robustness in dynamic, unstructured environments where geometric features may be insufficient. The major progress is achieved on increased functionalities in different perspectives and lighting conditions. Focusing on multimodal thinking for autonomous agents, Driess et al. [26] presented the PaLM-E approach that combines visual information with linguistic understanding and robotic input for contextual reasoning and navigation. The experimental results demonstrated robust performance with improved task generalization. Martins et al. [27] presented OVO-SLAM, an open-vocabulary semantic perception method that enhanced semantic localization by approximately 19% in novel environments. Furthermore, Teed and Deng [28] proposed DROID-SLAM, a deep learning-based SLAM system with recurrent optimization and deep optical flow estimates, which achieved an impressive absolute trajectory error of 1.1 cm on the EuRoC MAV benchmark. Finally, Hong et al. [29] proposed a semantic augmentation for object SLAM by combining multimodal semantic reasoning with language-guided object association. The experimental results showed that the proposed system was able to achieve an 18% enhancement in object association accuracy with a better grasp of the context.

4.3 Dynamic and Adverse Environment Understanding

Real environments change, and have moving objects and structural modifications as well. Agentic perception systems must be able to identify which part belongs to the static map and which one belongs to the dynamic interaction. SLAM-driven agents improve these capabilities with learning-based dynamic object detection techniques, motion consistency, object tracking, and scene change detection methods. However, classical SLAM suffers from adverse conditions, such as poor illumination and motion blur [25]. Significant implementations such as DynaSLAM [30,31] and StaticFusion [32] improve localization through semantic segmentation and deformation modeling. Resilient learning-based approaches improve the perception by training on compromised datasets, integrating multi-sensor data, for example, LiDAR, radar, etc., and responding to adverse environmental settings [27]. In addition, event-based cameras and learning-based improvement methods have demonstrated their effectiveness in dealing with motion blur and low-light conditions, enabling reliable perception in situations where conventional cameras fail.

4.4 Uncertainty, Memory, and Lifelong SLAM

In agentic systems, uncertainty is essential for mapping and perception. An agentic system simulates uncertainty to make faster decisions and recover from them. Probabilistic SLAM techniques enable the evaluation of map reliability, localization confidence, and perception uncertainty, which can guide agents in choosing safer navigation strategies, re-localization or exploratory triggers or adapting to degraded environments [24]. This integration of uncertainty is fundamental for sustainable autonomy and reliability. Memory-augmented SLAM systems also provide persistent semantic mapping, adaptability to the environment, and continuous learning, as mentioned by [26]. Further, to strengthen the details of the considered literature, these methods, along with the other considered literature, are shown in Table 4.

The results shown in Table 4 demonstrate relative improvements for each method within its evaluation framework. Due to variations in datasets, benchmarks, metrics, baseline methods, and experimental conditions across the studies, these numerical values are not directly comparable but serve to illustrate each approach's effectiveness in its context.

Table 4: Comparative analysis of slam-based agentic perception and world modeling.

Study & Year	Perception Method	World Modelling	Dataset	Semantic Capability	Uncertainty /Memory	Results/Findings	Key Contribution
Rosinol et al. [33]; 2021	Visual-Inertial + Scene Graphs	Dynamic scene graph	EuRoC MAV, uHumans2	Strong	Moderate	2–4 cm RMSE; ~10 Hz real-time optimization	Multi-layer semantic world representation
Ran et al. [21]; 2021	Semantic segmentation + geometry	Dynamic semantic map	Dynamic SLAM benchmarks	Strong	Limited	Improved localization stability in dynamic scenes	Real-time SLAM in dynamic scenes
Teed and Deng [28]; 2021	Deep optical flow + recurrent networks	Dense geometric map	EuRoC MAV	Moderate	Moderate	~1.1 cm ATE	End-to-end learned SLAM
Liao et al. [22]; 2022	Object-level semantic perception	Object-centric mapping	TUM Dynamic	Strong	Partial	~18% drift reduction	Robust object-level SLAM using spatial constraints
Hughes et al. [24]; 2022	Multi-modal semantic perception	Scene graph memory	Habitat, Replica	Strong	Strong	<5 cm localization drift	Hierarchical world modeling
Tian et al. [23]; 2022	Visual-Inertial + Semantic Mesh	Dense metric-semantic map	Multi-robot simulation datasets	Strong	Moderate	~15%–20% mapping consistency improvement	Distributed semantic SLAM for multi-robot systems

(Continued)

Table 4 (continued)

Study & Year	Perception Method	World Modelling	Dataset	Semantic Capability	Uncertainty /Memory	Results/Findings	Key Contribution
Yang and Wang [25]; 2022	CNN-based semantic filtering	Dense outdoor semantic map	KITTI, Cityscapes	Strong	Moderate	~10%–15% semantic consistency improvement	Large-scale outdoor semantic SLAM
Driess et al. [26]; 2023	Vision-language robotics	Embodied multimodal world representation	Embodied robotics benchmarks	Strong	Strong	Improved multimodal task generalization	Embodied LLM for autonomous navigation and robotic reasoning
Martins et al. [27]; 2024	Vision-language grounding	Semantic scene graph	ScanNet, Matterport3D	Very Strong	Strong	~19% localization improvement	Open-world semantic localization
Hong et al. [29]; 2025	Multimodal LLM agents	Semantic object graph	SUN RGB-D	Very Strong	Strong	~18% object association improvement	Heterogeneous multimodal reasoning

5 Active and Robust SLAM for Agentic Navigation

Active and robust SLAMs have abilities such as decision making, which help them in better mapping and perception as compared to the passive and classical SLAMs. The integration, as shown in Fig. 6, highlights that active exploration, robustness, and task-oriented mapping are essential for autonomous systems, highlighting the importance of action selection to enhance localization precision and overall task efficiency during navigation.

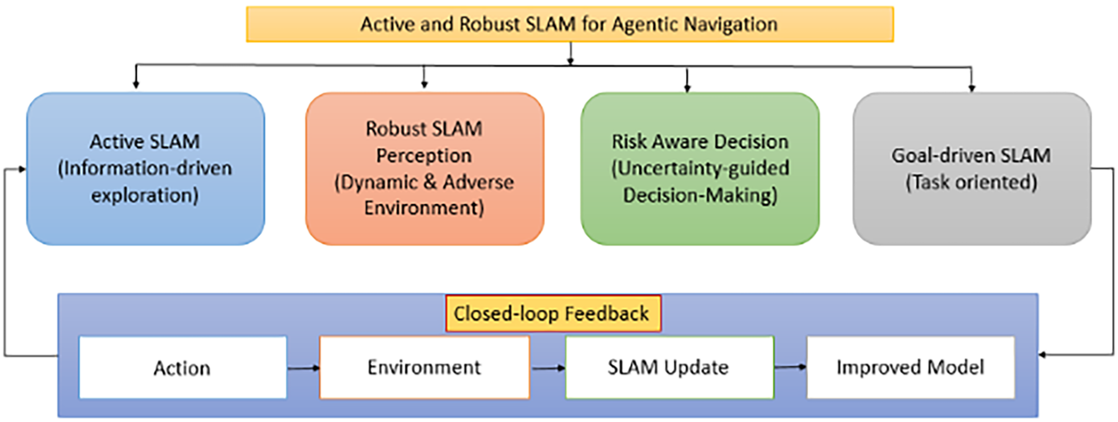

Figure 6: Autonomous navigation with active and robust SLAM.

5.1 Active SLAM and Information-Driven Exploration

The core principles of active SLAM are based on information-theoretic exploration strategies like entropy reduction and mutual information maximization that help agents navigate environments while minimizing uncertainty [10]. In passive SLAM, map estimation is based on the fact that the system moves while remaining stationary. In active SLAM motion, the map and uncertainty can be optimized simultaneously. It integrates decision-making into the agent's behaviors, allowing for the selection of techniques that improve mapping efficiency and localization accuracy. Active SLAM is different from the classical SLAM in that the motion planning is incorporated into the SLAM framework, which allows agents to improve information gathering, localization and task goals in a goal-oriented fashion. Active SLAM improves navigation and localization in unfamiliar environments with goals such as coverage-driven exploration and uncertainty awareness. Chaplot et al. [34] presented an active neural SLAM system that enhances autonomous exploration with reinforcement learning and a neural occupancy memory to make exploration more efficient and informative. Further, Ahmed et al. [35] proposed an efficient multi-robot active SLAM method that focuses on decentralized planning and information sharing to improve the coverage of the map and minimize the redundancy in exploration. Other advanced techniques, including learning-based principles, utilizing reinforcement learning to formulate navigation algorithms that adeptly balance exploration and exploitation within complex, partially observable environments.

5.2 Robust SLAM in Dynamic and Adverse Environments

The real-world environment is dynamic with continuous lighting changes, moving objects, motion blur and adverse weather, which need to be considered while building any SLAM system. The SLAM systems must be robust against these real-world challenges. Some researchers focused on the development of robust SLAM systems [36,37] that incorporate mechanisms to detect, filter, and model these challenges. Bescos et al. [30] used semantic segmentation to remove dynamic objects to improve localization precision in adverse environmental conditions. The experimental results showed that the proposed method reduce 35% trajectory error, resulting in improved pose estimation and map optimization. Further, performance was improved by aiding multi-object tracking in extremely dynamic environments [31]. Perception is susceptible to environmental factors such as low lighting and bad weather, which can be mitigated by deep learning-based perception modules that adapt to diverse situations and multi-sensor fusion for improved data collection. You et al. [38] proposed MISD-SLAM, a multimodal semantic SLAM framework that integrates instance segmentation, multiview geometry, and semantic mapping to improve localization robustness and mapping accuracy in highly dynamic environments. Moreover, deployment under adverse conditions requires robust strategies to maintain performance, such as dynamic object identification and filtering. Agentic SLAM systems consider environmental dynamics instead of ignoring them, resulting in making more robust systems. In challenging settings, the adaptation of the perspective selection, multi-sensor integration and real-time re-localization are important strategies [39]. These systems attempt to maintain accurate maps despite compromised sensing by proactively changing their behavior such as going slower or moving in the wrong direction.

5.3 Risk-Aware Navigation and Failure Recovery

Agentic SLAM systems improve robustness by leveraging uncertainty estimates to direct navigation and enable risk-aware decisions. Probabilistic SLAM frameworks, as emphasized by [10] and further extended by [40], offer essential uncertainty metrics that enhance planning safety. Key strategies of risk-aware navigation include monitoring localization confidence and map consistency, initiating re-localization or loop-closure detection, adapting motion strategies to risk-prone areas, and altering sensing modalities when the confidence falls [41]. Such strategies help agents to avoid catastrophic failures and to keep operations stable in unpredictable situations. Critical failure recovery solutions in the context of extended deployments include re-initialization, sensor redundancy and map rectification. Campos et al. [42] utilized visual-inertial relocalization and loop closure recovery, attaining less than 3 cm trajectory errors on EuRoC MAV benchmarks. Furthermore, advanced graph optimization techniques such as DRG-SLAM [43] improve the navigation robustness under bad weather conditions with adaptive recovery procedures. Robust agentic SLAM prioritizes risk awareness, allowing agents to evaluate the failure probability and take precautionary measures. During failures, the robust SLAM systems provide continuous localization and monitor the map integrity with re-exploration and smooth transit between sensing or navigation modes.

5.4 Task-Oriented and Goal-Driven SLAM

In autonomous systems, task-oriented SLAM has evolved from an independent mapping procedure into a core component of the navigation framework. This method highlights the need to prioritize perception and to allocate resources according to the goals of the agents, hence encouraging efficient and goal-oriented behaviors. Unlike classical SLAM, where the environment is systematically mapped, task-oriented SLAM allows the agent to focus on relevant locations such as traversable paths and target objects, to decrease the mapping resolution in less relevant areas, and focus on the mission goals [44]. This paradigm aligns with goal-conditioned exploration, enhancing mapping and navigation for completing assigned tasks instead of simply exhaustive mapping of the environment. In real-world applications such as search and rescue operations, where the critical elements of the map for navigation or manipulation are prioritized, while the rest of the environment is ignored to reduce the computational load, the autonomous SLAM systems play an important role. Additionally, Table 5 also gives highlights of the comparative analysis of active and robust SLAM methods discussed in this section. Further, the results mentioned in the table are as per the respective methods and are presented to highlight the performance in their original experimental settings and they are not directly comparable due to different benchmarks and metrics used by the respective studies.

Table 5: Comparison of active and robust SLAM methods for agentic navigation.

Ref. & Year	Exploration Strategy	Dataset	Dynamic Handling	Adverse Conditions	Risk Awareness	Results /Findings	Key Contribution
Labbe and Michaud [36]; 2018	Graph optimization exploration	KITTI	Moderate	Moderate	Partial	Loop closure precision >95% in long-term mapping	Long-term robust localization
Bescos et al. [30]; 2018	Motion-aware SLAM	TUM Dynamic	Strong	Limited	Partial	~35% reduction in trajectory error in dynamic scenes	Dynamic object removal
Chaplot et al. [34]; 2020	Reinforcement learning exploration	Gibson, Matterport3D	Moderate	Partial	Strong	~25% increase in navigation success rate	Learning-driven active SLAM
Campos et al. [42]; 2021	Passive exploration	EuRoC MAV	Limited	Moderate	Limited	ATE < 3 cm in visual-inertial sequences	Accurate visual-inertial localization
Chang et al. [37]; 2022	Collaborative exploration	DARPA SubT	Strong	Strong	Moderate	>90% mapping accuracy in underground environments	Robust mapping in underground environments
Wang et al. [43]; 2022	Deep robust mapping	Adverse weather datasets	Strong	Strong	Moderate	~28% reduction in localization failure rate	Adverse-environment SLAM
Denniston et al. [41]; 2022	Loop closure optimization	Large-scale SLAM benchmarks	Moderate	Partial	Moderate	~40% reduction in optimization overhead	Efficient large-scale SLAM
Ahmed et al. [35]; 2023	Cooperative active exploration	Multi-agent simulations	Strong	Moderate	Strong	~30% faster exploration coverage	Efficient multi-agent exploration
Abate et al. [39]; 2024	Semantic exploration	EuRoC MAV	Strong	Strong	Moderate	~20% robustness improvement in low-visibility environments	Robust semantic SLAM
You et al. [38]; 2024	Multi-modal semantic exploration	TUM RGB-D Dynamic Dataset and real-world dynamic indoor environments	Strong	Strong	Moderate	improvement in localization in real-world indoor dynamic sequence	Robust navigation under adverse conditions

6 Agentic Learning and Reasoning for SLAM-Based Autonomous Navigation

The learning of AI agents is enhanced by transforming systems into goal-oriented, reasoning capable as compared to the traditional passive estimators SLAM. Classical SLAM emphasizes precise localization and mapping, whereas agentic SLAM incorporates decision-making, planning, and adaptive learning, making it ideal for autonomous robots in dynamic and partially observable environments. Fig. 7 shows an agentic learning, reasoning and memory framework for SLAM-based navigation environment.

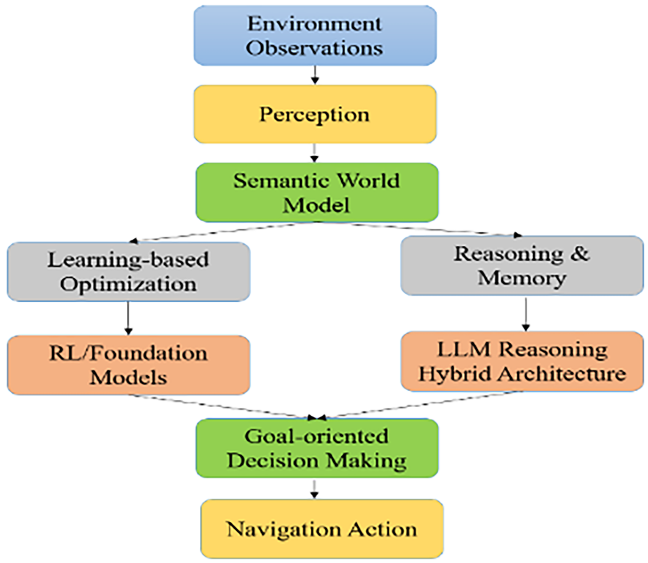

Figure 7: Agentic learning and reasoning for SLAM-based navigation environment.

Recently, researchers started utilizing neural architectures and Deep Reinforcement Learning (DRL) to learn navigation policies from environmental interaction. Additionally, world models and latent representations enable the agents to simulate future trajectories, resulting in model-based planning and reduction in real-world exploration costs [45]. In addition, techniques such as Deep Q-Networks (DQN) and Proximal Policy Optimization (PPO) allow agents to optimize long-term rewards, including exploration efficiency and safety [46].

6.1 Learning-Based Navigation and Policy Optimization

Learning-based SLAM incorporates end-to-end perceptual pipelines using differentiable architectures with control-theoretic formulations by mapping raw sensory inputs such as RGB-D imagery, LiDAR point clouds, and inertial measurements directly to navigation actions. These architectures use CNNs for hierarchical spatial feature extraction and Vision Transformers (ViTs) for long-range dependency modeling. This allows to derive value functions and stochastic policies in a common latent space [47]. Savva et al. [48] created navigation agents utilizing reinforcement learning and semantic exploration, attaining impressive metrics in Success weighted by Path Length (SPL) and navigation success. Further, multi-scale feature pyramids augmented with skip connections (e.g., U-Net-inspired encoders) improve metric map resolution while preserving semantic context [49]. Chiun et al. [50] presented the improved multi-agent policy for collaborative exploration and demonstrated 27% improvement in efficiency in simulated environments. However, these approaches are data-hungry and require millions of samples generated by simulation with domain randomization to bridge the sim-to-real transfer gap caused by photometric and dynamic discrepancies [51]. Nowadays, researchers have widely utilized PPO and Soft Actor-Critic (SAC) algorithms for navigation in continuous action space and include entropy regularization to balance exploration-exploitation in partially observable environments [52]. Moreover, meta-learning paradigms such as Model-Agnostic Meta-Learning (MAML) [53] are trained to quickly adapt the policy to unseen environments by using few-shot gradient updates to improve generalization abilities. Simultaneously, Explainable AI (XAI) techniques such as SHapley Additive exPlanations (SHAP) and Gradient-weighted Class Activation Mapping (Grad-CAM) saliency maps rollout are gaining popularity for auditing decision-making procedures and verifying spatial reasoning in safety-critical navigation domains [54]. Formal verification of the learned policies using a reachability analysis is still a research challenge, especially for environments that are dense with obstacles for which collision-avoidance must be guaranteed [55].

6.2 Reasoning, Memory, and Hybrid Architectures

A huge number of agentic SLAM systems are now using neuro-symbolic hybrid architectures that combine sub-symbolic neural perception with symbolic reasoning engines. This combination allows hierarchical task decomposition and goal-directed navigation in semantically rich environments. High-level planning on abstract semantic graphs is done with nodes representing objects/landmarks and edges representing spatial relations. Neural controllers regulate low-level trajectory execution of the raw sensory streams [56]. These architectures separate representational learning from logical reasoning, enhancing interpretability while maintaining perceptual integrity [57]. Neural Turing Machines (NTMs) and Differentiable Neural Computers (DNCs) extend standard recurrent architectures with addressable external memory banks, enabling agents to store and retrieve metric-semantic map embedding across arbitrary time spans [58]. The temporal credit assignment problem inherent in long-horizon navigation tasks is being tackled by memory-augmented architectures [44,59]. Transformer-based memory tokens implemented as persistent key-value pairs in cross-attention layers facilitate episodic spatial recall without the vanishing gradient limitations [60]. Hybrid topological-metric maps encode spatial knowledge at different granularities, i.e., coarse topological graphs for navigable connectivity and dense metric submaps for accurate geometric constraints for localization [61]. These architectures significantly enhance navigation scalability in large-scale environments at the cost of non-trivial computational overhead due to memory consolidation, graph maintenance, and inter-level consistency enforcement [62,63]. Lifelong learning mechanisms are used to mitigate catastrophic forgetting during continual map updates, e.g., Elastic Weight Consolidation (EWC) and Progressive Neural Networks (PNN) [64]. Furthermore, Table 6 shows the comparison of the studies with their original results obtained using different benchmarks, metrics and experimental settings for agentic learning and reasoning in SLAM-based autonomous navigation systems.

Table 6: Comparison of learning and reasoning mechanisms in agentic SLAM.

Ref. & Year	Learning Paradigm	Reasoning Capability	Dataset	Foundation Models	Adaptation	Results /Findings	Key Contribution
Haarnoja et al. [52]; 2018	Deep reinforcement learning (maximum entropy RL)	Moderate	Continuous control bench-marks	No	Strong	Improved sample efficiency and policy stability compared with conventional RL methods	Robust policy optimization for autonomous decision-making and navigation
Chen et al. [47]; 2019	Deep imitation learning	Moderate	Urban driving scenarios	No	Strong	Improved navigation safety and stable autonomous driving in complex urban scenarios	Safe policy learning for autonomous urban navigation

(Continued)

Table 6 (continued)

Ref. & Year	Learning Paradigm	Reasoning Capability	Dataset	Foundation Models	Adaptation	Results /Findings	Key Contribution
Savva et al. [48]; 2019	RL + semantic navigation	Moderate	Habitat Challenge	Partial	Strong	Top-tier SPL and navigation success metrics	Benchmark embodied navigation
Ahn et al. [59]; 2022	LLM-guided planning	Strong	Real-world robotics tasks	Yes	Moderate	~84% task completion accuracy	Language-driven embodied reasoning
Shah et al. [44]; 2023	Vision-language learning	High-level reasoning	Habitat	Yes	Strong	~80% navigation success rate in language-guided tasks	Navigation using language and visual priors
Huang et al. [65]; 2023	LLM + affordance reasoning	Strong	RLBench	Yes	Moderate	~81% success rate in language-conditioned planning	Task-aware robotic planning
Brohan et al. [62]; 2023	Foundation action models	Strong	PaLM robotic bench-mark	Yes	Strong	Generalized across >6 unseen robotic tasks	Web-scale robotic control
Mu et al. [63]; 2023	Vision-language-action learning	Strong	Embodied AI bench-marks	Yes	Strong	~18% improvement in contextual decision-making accuracy	Embodied autonomous reasoning
Chiun et al. [50]; 2025	Multi-agent RL	Cooperative reasoning	Multi-agent simulation	Partial	Strong	~27% improvement in cooperative exploration efficiency	Cooperative autonomous exploration

7 Multi-Agent SLAM and Cooperative Navigation

Single-agent probabilistic mapping can be extended using Multi-Agent SLAM (MA-SLAM) for distributed systems in which a heterogeneous team of robots collaboratively build and maintains a globally consistent representation of the environment. Each agent maintains a local factor graph that encodes pose constraints from loop closures, odometers, and relative pose measurements between pairs of agents from common landmark observations [66]. Distributed Bayesian inference on these coupled factor graphs results in a fused global map that improves coverage, robustness to individual sensor failures, and exploration efficiency. All these properties are critical for time-sensitive applications such as swarm robotics, autonomous vehicle platoons and search-and-rescue operations [67]. Further, distributed Pose Graph Optimization (PGO) algorithms like Distributed Gauss-Seidel (DGS) and Riemannian stochastic gradient descent on the SE (3) manifold provide scalable solutions for globally consistent map fusion without centralizing raw sensor data [68]. The inter-agent data association problem of establishing correspondences between independent observations obtained by different agents under varying viewpoints and temporal offsets is solved using invariant descriptor matching, such as Signature of Histograms of OrienTations (SHOT) and Fast Point Feature Histograms (FPFH), combined with RANdom SAmple Consensus (RANSAC)-based outlier rejection [69]. Byzantine fault-tolerant consensus protocols provide additional robustness of these systems to compromised or malfunctioning agents in adversarial environments [70]. Fig. 8 shows framework of cooperative multi-agent SLAM.

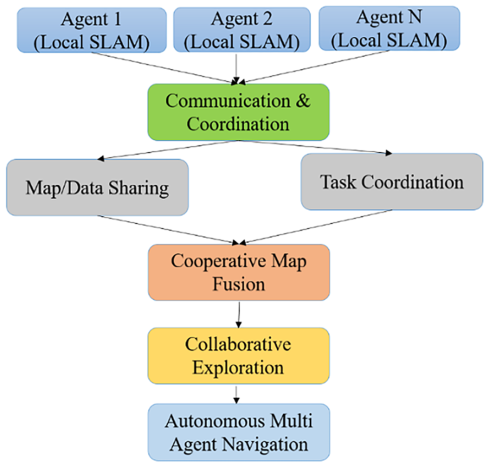

Figure 8: Framework of cooperative multi-agent SLAM.

7.1 Cooperative Mapping and Exploration

For MA-SLAM systems, cooperative exploration approaches are formulated as Decentralized Partially Observable Markov Decision Processes (Dec-POMDPs) in which each agent selects frontier-based or information-theoretic waypoints to reduce joint map entropy related to communication and energy constraints [71]. Further, frameworks utilizing centralized value function decomposition methods like Q-Mixing Networks (QMIX), Value Decomposition Network (VDN), etc., are developed using Multi-Agent Reinforcement Learning (MARL). These frameworks enable agents to learn coordinated exploration policies through shared global reward signals encoding mutual information gain over the occupancy grid [72]. The Centralized Training with Decentralized Execution (CTDE) paradigm addresses the non-stationarity problem in concurrent multi-agent learning. It utilized centralized critics that focus on global state information during the training phase, and execution-time policies are limited to local observations and messages received from communication [73]. Mahboob et al. [74] reduced the communication overhead by emphasising decentralized map fusion. Graph Neural Networks (GNNs) applied over dynamic agent communication graphs enable permutation-invariant policy representations that generalize across variable team sizes and topologies [75]. Schmuck et al. [76] proposed a large-scale collaborative visual SLAM framework based on shared optimization and distributed loop closure detection. The proposed framework was able to achieve a loop closure accuracy of 92% by improving collaborative visual SLAM. Active neural SLAM integrates semantic segmentation outputs with occupancy predictions to prioritize exploration of semantically novel regions, improving task-relevant coverage metrics over purely geometric approaches [77,78]. Lajoie et al. [79] provided a distributed SLAM system, robust to outliers, decreasing localization failures by around 30%. Further, Zuo et al. [80] used centralized optimization to propose Communication-Efficient Centralized Multi-Robot Dense-SLAM (CCMD-SLAM) for attaining a real-time mapping at 15 frames per second.

7.2 Communication and Coordination Challenges

In MA-SLAM, communication constraints are a fundamental bottleneck because the raw sensor data volumes, e.g., 3D LiDAR point clouds at 10–20 Hz per agent, are not compatible with the limited bandwidth wireless channels in real-world deployments. Learned map compression techniques with Variational Autoencoders (VAEs) and discrete tokenization such as Vector Quantized Variational Autoencoder (VQ-VAE), reduce transmitted payloads by up to two orders of magnitude while preserving localization-relevant geometric structure [81]. Event-triggered communication protocols based on threshold-gated information gain metrics are used to suppress redundant transmissions to reduce channel utilization without significant map accuracy degradation [82]. Asynchronous distributed optimization algorithms such as Alternating Direction Method of Multipliers (ADMM) with proximal updates can handle delayed gradient information and still guarantee convergence under bounded communication latency [83]. Further, stochastic communication graphs with probabilistic edge weights are used for network topology dynamics such as packet loss, intermittent connectivity, and variable inter-agent distances [84]. Proper allocation of tasks is done using combinatorial optimization frameworks such as Nash equilibrium-seeking game-theoretic protocols, Auction-Based Task Assignment (Consensus-Based Bundle Algorithm (CBBA)), and Hungarian algorithm variants. These frameworks together confirm deadlock-free mission execution and load-balanced role distribution among heterogeneous multi-robot teams [85]. Furthermore, due to the rise in the number of autonomous agents, another issue is scalability in MA-SLAM systems, which complicates the maintenance of coherent maps and synchronization. To address this issue, Swarm-SLAM [86] and D2SLAM [87] utilized lightweight decentralized communication protocols. Furthermore, to improve semantic consistency using shared scene graphs and distributed semantic optimization, Hydra-Multi [88] and multi-robot decentralized collaborative SLAM [89] are proposed by the researchers. Moreover, Table 7 shows the comparative analysis of the considered multi-agent and cooperative SALM systems with their obtained results in the original experimental context.

Table 7: Comparison of multi-agent SLAM and cooperative navigation systems.

Ref. & Year	Collaboration Type	Communication	Dataset /Benchmark	Map Fusion	Scalability	Results/Findings	Key Contribution
Lajoie et al. [79]; 2020	Distributed SLAM	Multi-robot communication	Multi-robot SLAM datasets	Moderate	Strong	~30% reduction in outlier-induced failures	Outlier-resilient cooperative SLAM
Schmuck et al. [76]; 2021	Collaborative VSLAM	Shared optimization	KITTI	Strong	Moderate	>92% loop closure precision	Large-scale collaborative mapping
Mahboob [74]; 2023	Distributed graph optimization	Distributed fusion	Distributed SLAM benchmarks	Moderate	Strong	~35% lower communication overhead	Distributed map optimization
Wang et al. [78]; 2023	Neural collaborative mapping	Shared latent representation	Replica	Strong	Moderate	Improved neural map reconstruction accuracy by ~18%	Neural implicit collaborative SLAM
Lajoie and Bel-trame [86]; 2023	Swarm intelligence	Lightweight communication	UAV simulations	Moderate	Strong	~26% improvement in distributed exploration coverage	Swarm-based distributed mapping

(Continued)

Table 7 (continued)

Ref. & Year	Collaboration Type	Communication	Dataset /Benchmark	Map Fusion	Scalability	Results/Findings	Key Contribution
Chang et al. [88]; 2023	Multi-agent scene understanding	Shared semantic graph	Habitat	Strong	Strong	~20% improvement in semantic consistency	Semantic cooperative world modeling
Zuo et al. [80]; 2024	Centralized SLAM	Shared server	Indoor robotics datasets	Strong	Moderate	Real-time collaborative mapping at ~15 FPS	Real-time collaborative SLAM
Xu et al. [87]; 2024	Decentralized VIO	Distributed	UAV swarm benchmarks	Strong	Strong	<5 cm relative localization error across UAV swarm	Collaborative aerial swarm SLAM
Lajoie et al. [89]; 2025	Cooperative semantic SLAM	Distributed semantic fusion	Cooperative semantic datasets	Strong	Strong	~24% improvement in semantic map consistency	Cooperative semantic autonomy

8 Datasets, Benchmarks, and Simulation Environments

Standardized datasets and simulation environments constitute the empirical backbone of SLAM research, providing reproducible evaluation conditions across perception, localization, and navigation modalities. The fidelity of a benchmark is determined by sensor diversity, ground truth precision, environmental variety, and trajectory complexity. As algorithmic sophistication increases, benchmark requirements have expanded from simple monocular odometry sequences to large-scale, multi-modal, multi-agent scenarios incorporating dynamic obstacles and challenging illumination conditions [90]. Dataset diversity enables robust evaluation of generalization, while well-defined metric protocols ensure comparability across heterogeneous algorithmic families, including feature-based, direct, and learning-based SLAM pipelines [91].

8.1 Benchmark Datasets and Metrics

The Karlsruhe Institute of Technology and Toyota Technological Institute (KITTI) Benchmark Suite [90] remains one of the most widely adopted datasets for evaluating outdoor autonomous driving systems, providing synchronized stereo imagery, 64-beam Velodyne HDL-64E LiDAR point clouds, and RTK-GPS ground truth at ±2 cm accuracy across 22 odometry and 43 tracking sequences spanning 39.2 km of urban and rural driving scenarios. The European Robotics Challenge (EuRoC) Micro Aerial Vehicle (MAV) dataset [92] supports numerous stereo-inertial sequences of varying difficulty (Easy/Medium/Hard) recorded in an industrial hall and machine room environments with a Leica MS50 laser tracker providing sub-millimetre pose ground truth for rigorous Visual Inertial Odometry (VIO) evaluation. Further, the Technical University of Munich (TUM) RGB-D dataset [91] targets indoor visual SLAM with 39 sequences captured by a Microsoft Kinect v1 at 640 × 480 resolution, with sub-millimetre ground truth provided by a Vicon motion capture system operating at 100 Hz. More recent benchmarks address open challenges in SLAM research: the Replica dataset [93] provides 18 photo-realistic synthetic indoor scenes with dense semantic mesh annotations and zero-noise ground truth, enabling evaluation of semantic SLAM and Neural Radiance Field (NeRF)-based mapping methods. ScanNet [94] offers 1513 RGB-D scans with instance-level semantic labels across 707 distinct indoor spaces, supporting 3D reconstruction and semantic segmentation benchmarking. In contrast, Oxford RobotCar [95] uniquely addresses long-term localization across seasonal and lighting variations with over 100 repeated traversals of a 10 km urban route, enabling evaluation of place recognition and appearance change robustness. Fig. 9 shows a comparison of multi-metric benchmark datasets based on sensor diversity, scalability, ground truth quality, environmental diversity, and community acceptance. In addition, Table 8 provides a structured comparison of these datasets across key technical dimensions.

Figure 9: Multi-metric benchmark dataset comparison across sensing diversity, scale, ground truth quality, environment diversity, and community adoption.

Table 8: Comparative analysis of benchmark SLAM datasets.

Dataset	Sensor Modalities	Environment	Sequences	GT Source	Best ATE (m)	Primary Use
KITTI [90]	Stereo RGB, LiDAR (64-beam Velodyne), GPS/IMU	Outdoor urban	22 odometry + 43 tracking	GPS/RTK, ±2 cm	0.53	Autonomous driving, odometry
TUM RGB-D [91]	RGB-D (Kinect v1, 640 × 480), IMU	Indoor structured	39 sequences	Motion capture (Vicon), ±1 mm	0.008	Visual SLAM, depth estimation
EuRoC MAV [92]	Stereo greyscale, 6-DOF IMU (ADIS16448)	Indoor (hall/room)	11 sequences	Leica MS50 laser tracker, ±1 mm	0.009	Aerial SLAM, VIO
Replica [93]	Synthetic RGB-D, semantic meshes	Indoor photo-realistic	18 scenes	Mesh GT (zero noise)	—	Embodied AI, scene understanding
ScanNet [94]	RGB-D (Structure Sensor), IMU	Indoor large-scale	1513 scans	BundleFusion 3D reconstruction	0.081	3D reconstruction, semantic seg.
Oxford Robot-Car [95]	Stereo RGB, 3× 2D LiDAR, GPS, IMU	Outdoor long-term	100 + traversals	GPS/RTK + HD map	0.71	Long-term localization, mapping

Quantitative SLAM evaluation is based on a standard set of metrics for local accuracy and global consistency. One of the metrics, known as Absolute Trajectory Error (ATE), is used to compare the absolute distances between the estimated trajectory, described by $P_1, P_2, \ldots, P_n \in SE(3)$ and the ground truth trajectory, described by $Q_1, Q_2, \ldots, Q_n \in SE(3)$ [91]. In addition, Relative Pose Error (RPE) measures the local accuracy of the trajectory over a fixed time interval Δ, thus corresponding to the drift of the trajectory which is in particular useful for the evaluation of visual odometry systems [91]. Additional metrics include loop closure precision/recall (F1-score), volumetric map completeness, and computational throughput in frames-per-second (FPS), which together characterize the trade-off space between accuracy, map fidelity, and real-time feasibility. Table 9 formalizes these metrics with mathematical definitions and an evaluation context.

Table 9: SLAM evaluation metrics: definitions and relevance.

Metric	Definition	Mathematical Formulation	Units/Range	Relevance
ATE	Absolute Trajectory Error: compares the absolute distances between the estimated and the ground truth (GT) trajectory for a rigid-body transformation S	ATE at time step i: $F_i = Q_i^{-1} S P_i$	Meters (↓ better)	Global consistency, loop closure quality
RPE	Relative Pose Error: measures the local accuracy of the trajectory over a fixed time interval Δ	RPE at time step i: $E_i = \left(Q_i^{-1} Q_{i+\Delta}\right)^{-1} \left(P_i^{-1} P_{i+\Delta}\right)$	m/m or °/m (↓ better)	Local odometry accuracy, incremental drift
LC Precision/Recall	Loop Closure Detection accuracy: used to evaluate classification models and loop closure detection systems, based on precision and recall.	$P = \frac{TP}{TP+FP}$; $R = \frac{TP}{TP+FN}$ True Positives (TP): The system correctly identifies a true revisited location. False Positives (FP): The system incorrectly claims a loop closure (perceptual aliasing), which can catastrophically warp the map. False Negatives (FN): The system fails to recognize a revisited location, missing an opportunity to correct accumulated drift.	[0, 1] (↑ better)	Topological map consistency, re-visitation
Map Completeness	Fraction of ground-truth voxels (v) reconstructed above occupancy threshold τ	$\lvert v : occupancy\,(v) > \tau \rvert \cap GT$	[0, 1] (↑ better)	3D reconstruction fidelity
Runtime/FPS	Processing throughput of front-end and back-end pipelines	Frames per second (or ms/frame)	Hz (↑ better)	Real-time capability on target hardware

8.2 Simulation Platforms and Emerging Trends

Physics-based simulation environments provide controlled, repeatable testing infrastructure that supports large-scale algorithmic iteration without the logistical cost of physical hardware experiments. Car Learning to Act (CARLA) [96] used Unreal Engine 4 to generate photorealistic urban driving scenarios with configurable weather, dynamic pedestrians, and a comprehensive sensor suite (RGB cameras, depth sensors, 3D LiDAR, radar), making it the dominant platform for perception research and autonomous vehicle SLAM. Another platform based on Unreal Engine 4 and Unity is AirSim [97], which enhances simulation fidelity to aerial and ground vehicles with physically accurate aerodynamic models so that VIO systems can be evaluated in unconstrained 3D environments. The sim-to-real fidelity is taken to the next level with

NVIDIA Isaac Sim [98], which uses PhysX for rigid body dynamics and RTX ray tracing for Physically-Based Rendering (PBR) with complete USD-based digital twin integration to directly import CAD and BIM models from industrial environments. The Habitat simulation platform [48] is designed for embodied AI research and provides a highly optimized renderer that can achieve thousands of fps on Replica and Matterport3D scene meshes to support large-scale training of major Point and Object-Goal navigation policies. Table 10 shows the comparison of features of simulation platforms, and Fig. 10 shows the architecture of the complete sim-to-real transfer pipeline. Further, Table 11 provides a simulation feature matrix showing platform capabilities.

Table 10: Simulation platform feature comparison.

Platform	Engine	Env. Type	Sensor Suite	Multi-Agent	ROS	Digital Twin	Key Strength
CARLA [96]	Unreal 4	Outdoor urban	RGB, Depth, LiDAR, Radar, GNSS	✓	✓ (ROS bridge)	Partial	Photo-realistic urban AV sim
AirSim [97]	Unreal 4/Unity	Outdoor + indoor	RGB, Depth, IMU, LiDAR, sonar	✓	✓	Partial	Aerial & ground vehicle sim
Habitat [48]	Custom renderer	Indoor (Replica, MP3D)	RGB-D, semantic, panoramic	✓	✗	✗	Embodied AI/PointNav research
Isaac Sim [98]	PhysX/RTX	Outdoor + industrial	Full sensor suite + force/torque	✓	✓	✓ (USD)	High-fidelity physics & digital twin
iGibson [99]	PBR renderer	Interactive indoor	RGB-D, LiDAR, haptic	✓	✓	Partial	Object interaction & re-arrangement
MORSE [100]	Blender/Bullet	Outdoor + indoor	Modular: any sensor	✓	✓	✗	Modular, scriptable robotics sim

Note: ✓—Supported, ✗—Not Supported.

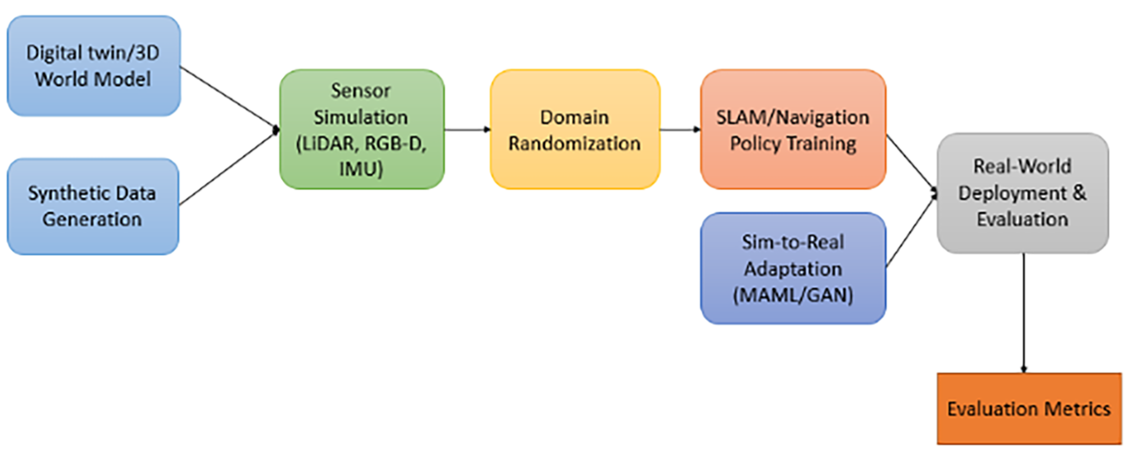

Figure 10: End-to-end sim-to-real transfer pipeline.

Table 11: Simulation platform feature matrix across key capabilities required for SLAM evaluation.

	Outdoor	Indoor	LiDAR	RGB-D	IMU	Multi-Agent	Digital Twin	ROS
CARLA	✓	✗	✓	✓	✓	✓	~	✓
AirSim	✓	~	✓	✓	✓	✓	~	✓
Habitat	✗	✓	✗	✓	~	✓	✗	✗
Issac Sim	✓	✓	✓	✓	✓	✓	✓	✓
iGibson	✗	✓	✗	✓	✓	✓	~	✓
MORSE	~	~	~	✓	✓	~	✗	✓

Note: ✓—Supported, ~—Partial, ✗—Not Supported.

The emerging trends in SLAM benchmarking and simulation are mainly focused on three directions. (i) Digital twin technology leverages high-fidelity 3D scanning, including terrestrial LiDAR, photogrammetry, etc., and physics simulation to generate dynamically consistent virtual replicas of target deployment environments [98]. (ii) Procedural generation of scalable training data with perfect ground-truth annotations can be achieved using Generative Adversarial Networks (GANs) and neural rendering, including NeRF, 3D Gaussian Splatting, enabling infinite variation in scene appearance, lighting, and dynamic agent behaviour without manual scene authoring [101]. (iii) Domain randomization with curriculum learning adapts simulation parameters, including sensor noise models, material reflectance, and agent dynamics, to the current policy performance, accelerating convergence when transferring to reality while maintaining sample efficiency [102].

9 Applications of SLAM-Driven Intelligent Agents

SLAM-driven intelligent agents have enabled a broad spectrum of real-world applications via the integration of spatial SLAM-driven intelligent agents have facilitated a wide range of real-world applications via the combination of spatial mapping with perception, reasoning, and self-directed decision-making. Through their integration with sensing, learning, and reasoning modules, these agents have established themselves as key elements of advanced autonomous systems across several application domains.

(i) Terrestrial: SLAM-driven intelligent agents are being widely used for a variety of terrestrial applications, especially in ever-changing, semi-structured environments that require reliable localization, mapping, and smart and adaptive decision-making for their functioning. These applications primarily integrate LiDAR, vision, and odometry to manage dynamic challenges, human interaction, and real-time decision-making [103].

  a. Self-directed Driving: SLAM can enable autonomous driving by facilitating accurate vehicle localization and environmental perception, particularly in GPS-denied environments. SLAM also helps vehicles in comprehending lanes, traffic signs, and dynamic objects by enhancing safe navigation and real-time decision-making [40,104,105].
  b. Mining and Underground Exploration: SLAM-based autonomous navigation agents are essential for tunnel mapping, supporting safe navigation, and monitoring structural integrity in GPS-denied and dangerous environmental conditions [106]. The SLAM-based agentic systems are capable of identifying obstacles (e.g., debris in post-blast inspection), structural instabilities, monitoring deformation in real-time and exploring semantic mapping in mines [37,107].
  c. Agriculture: In precision agriculture, SLAM-driven intelligent agents are being extensively adopted to facilitate autonomous navigation, crop monitoring, field mapping, and intelligent farm management in complex outdoor environments [108,109]. By integrating multispectral imaging

and AI-based perception with SLAM, intelligent agents can identify stressed crops, nutrient deficiencies, and pest infestations at an early stage [110,111].

d. Storage automation: Autonomous robots in warehouses use SLAM to avoid obstacles, optimize logistics, and navigate dynamic layouts without fixed infrastructure. These intelligent agents increase the efficiency and scalability of logistics by optimizing routes and adapting to layout changes [19,112].

e. Service Robots: SLAM-based intelligent agents are used to enable effective indoor navigation and human interaction for service robots used in hotels, hospitals, and households. These robots exhibit agentic characteristics that enable them to interpret user commands, find the most efficient routes for delivery, and adapt to complex scenarios such as crowded environments [113–115]. They also allow task prioritization and adaptive routing in dynamic clinical environments [116,117]. Further, SLAM-based intelligent agents can actively build semantic maps of indoor home environments. By combining active exploration with semantic understanding, the system improves household robot efficiency in tasks such as navigation, object search, and human–robot interaction [118–120]. In addition, these service robots are also widely used in hotels for indoor autonomous navigation with real-time mapping and localization, with hospitality tasks such as cleaning rooms, delivering and collecting food orders in diners to improve operational efficiency and enable contactless services [121,122].

f. Self-directed delivery Robots: SLAM-based intelligent agents play a vital role in autonomous delivery robots, assisting with efficient navigation in sophisticated environments such as urban locales and indoor spaces. By employing SLAM, these robots can generate and modify maps while identifying their exact location, thus alleviating the necessity for GPS or pre-installed infrastructure [123,124]. Further, these sophisticated agents are capable of augmenting task scheduling and route optimization for effective last-mile delivery [125,126].

(ii) Aerial (UAV): SLAM-enabled intelligent agents are being substantially exploited for various aerial applications, such as aerial robots, smart city monitoring, etc. Aerial SLAM emphasizes efficient, lightweight computing, resilience to quick movements, and real-time mapping [127–129].

a. Autonomous Drones (UAVs): SLAM-driven intelligent agents play a critical role in aerial robotics, specifically in the field of unmanned aerial vehicles (UAVs), where reliable and precise navigation must be implemented in three-dimensional, GPS-denied, and highly dynamic environments [130,131]. Visual and LiDAR-based SLAM techniques facilitate UAVs to estimate pose and generate real-time maps in complex environments, such as dense forests, canyons, etc. Enhanced by SLAM-based intelligent agent capabilities, these systems enable high-level reasoning, adaptive path planning, and decision-making to allow drones to replan trajectory autonomously, avoid dynamic obstacles, and optimize exploration strategies [132].

b. Smart City Monitoring: SLAM-based UAVs are employed in surveillance and smart city monitoring for various tasks such as traffic analysis, crowd monitoring, and infrastructure inspection. They have the capability to generate georeferenced maps and improve situational awareness across wide areas. These aerial robots are efficient in minimizing the requirement for manual inspections in dangerous or difficult-to-reach locations by automatically inspecting infrastructure including wind turbines, bridges, power lines, and high-rise buildings [133].

c. Space and Extreme Environment Applications: In space and extreme environments, SLAM-driven intelligent agents are essential in various fields by allowing autonomous operations, especially in situations where human accessibility is restricted [134]. Using optical, LiDAR, and inertial sensing to navigate barriers and identify important scientific areas, rovers with SLAM systems may autonomously locate and survey uncharted territory on celestial planets such as Mars and the Moon [135].

(iii) Augmented Reality (AR) and Virtual Reality (VR): For AR and VR, SLAM-based intelligent agents are essential embodied AI systems by enabling precise spatial understanding, real-time environment mapping, and context-aware interactions between digital agents and the physical world [136]. SLAM is a key component for AR applications, where it is used to simultaneously track the position and orientation of devices and build a map of the surrounding environment so that virtual objects can be stably anchored in real-world scenes [137]. Whereas in VR systems, SLAM enables spatial interaction, room-scale tracking, and dynamic scene reconstruction, improving user immersion and interaction accuracy [138]. These systems are widely studied in autonomous robotics, virtual simulation platforms and studies in Human-Robot Interaction [139].

(iv) Underwater Applications: SLAM-based intelligent agents are essential for underwater applications, such as underwater robots, due to their capability to function in challenging real-world environments such as with no GPS and poor visibility [140,141].

  a. Gas and oil industries: SLAM-based agents have potential for use in the offshore oil and gas industry for inspection of pipelines and subsea infrastructure, where they move along pipelines autonomously, look for anomalies such as corrosion or leaks, and produce complex 3D reconstructions [140]. They take high-resolution 3D photos of unknown areas for deep-sea exploration despite the extreme pressure and poor visibility [142].

  b. AUVs and ROVs: Robots such as Autonomous Underwater Vehicles (AUVs) and Remotely Operated Vehicles (ROVs) use SLAM techniques based on sonar, acoustic sensors, Doppler Velocity Logs (DVL), and Inertial Measurement Units (IMUs) for generating maps in underwater conditions [141,143]. Recent advancements include multi-robot cooperative SLAM, where multiple AUVs can cooperate to share mapping information, leading to improved coverage and efficiency over large underwater spaces [86,144,145].

10 Open Challenges and Future Research Directions

Despite the advances in SLAM-based intelligent agents, for autonomous navigation, there are still many challenges in this area. The challenges include integration of semantic reasoning, geometric mapping and limitations of the current frameworks. Some of the other prominent challenges in the current SLAM-based autonomous navigation agentic systems are as follows:

(i) Robustness is one of the key challenges in ever-changing, unstructured scenarios. Moving objects, variations in lighting conditions, sensor noise and weather can cause high degradation of localization and mapping accuracy. Classical SLAM systems often assume static environments and are therefore susceptible to errors in disaster-response scenarios, indoor public spaces, and crowded city centers.

(ii) The computational efficiency and scalability of SLAM-based intelligent agents are another important challenge. Systems nowadays rely increasingly on deep learning, multimodal sensing and large-scale semantic mapping. These techniques require significant computational resources and energy consumption. Collaboration of multiple agents also leads to challenges in synchronization between robots, distributed mapping, and communication constraints.

(iii) Semantic understanding and contextual reasoning also pose significant challenges in current SLAM-driven methods. Although recent semantic SLAM approaches incorporate object recognition and scene comprehension, most of the systems continue to face problems in reconciling geometric mapping with advanced cognitive reasoning. Addressing these challenges requires enhanced integration among SLAM, LLMs, VLMs, and embodied AI frameworks.

(iv) Moreover, ethical and security issues such as data privacy, secure human-robot interaction, and adversarial attacks are gaining significance as intelligent agents function inside shared human discourse.

In future, SLAM-based intelligent agents are expected to evolve to the creation of lifelong SLAM systems that can continuously learn, update the maps, and correct the drift in dynamic settings. One major trend is to combine SLAM with foundation models, such as VLMS, LLMs, etc., that enable agents to understand natural language instructions and perform goal-oriented navigation. Furthermore, multi-robot collaboration, uncertainty-aware mapping, and secure and explainable decision-making are becoming major areas to ensure scalability, robustness, and reliable deployment in real-world scenarios.

11 Conclusion

This paper provided a comprehensive systematic review based on PRISMA 2020 guidelines for SLAM-based autonomous navigation agentic systems. The review was conducted with a total of 1870 articles explored for consideration, resulting in 141 chosen articles with 70 studies included in the review process. The review highlighted the background and evolution of agentic SLAM, starting from classical SLAM. The paper proposed a taxonomy for agentic perception and world modeling with details of active and robust SLAM for agentic navigation. Agentic SLAM functions as an intelligent decision-maker in an adverse and real-world environment. The paper also discusses dynamic and effective SLAM approaches with multi-agent SLAM, demonstrating the benefit of risk-aware planning, multi-sensor fusion and information-driven exploration for increased reliability in real-world applications. The paper also includes benchmark datasets and simulation platforms used for SLAM-based intelligent agents for autonomous navigation. Furthermore, different application areas of agentic SLAM in autonomous navigation, including agriculture, robotics, UAV, etc., were discussed. The paper also highlighted the open challenges and future directions for researchers who intend to work in this field. Future research should focus on integrated frameworks that include perception, reasoning and learning, as well as advances in multimodal and lifelong learning.

Acknowledgement: The authors would like to thank Thapar Institute of Engineering & Technology, Patiala, for their support.

Funding Statement: The authors received no specific funding for this study.

Author Contributions: All the authors contributed equally in writing the manuscript. Mukesh Dalal: Initial draft, data curation, comparative analysis, and thorough proofreading. Anterpreet Kaur Bedi: Write the Introduction, background, applications, draw figures and Payal Mittal: Wrote equations, critical analysis, and thorough proofreading. All authors reviewed and approved the final version of the manuscript.

Availability of Data and Materials: The data will be available on reasonable request.

Ethics Approval: Not applicable.

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

Supplementary Materials: The supplementary material is available online at https://www.techscience.com/doi/10.32604/cmc.2026.086270/s1.

References

1. Chen W, Chi W, Ji S, Ye H, Liu J, Jia Y, et al. A survey of autonomous robots and multi-robot navigation: perception, planning and collaboration. Biomim Intell Robot. 2025;5(2):100203. doi:10.1016/j.birob.2024.100203.
2. Wei Z, Wang S, Chen K, Wang F. ROS-based navigation and obstacle avoidance: a study of architectures, methods, and trends. Sensors. 2025;25(14):4306. doi:10.3390/s25144306.

3. Liu Y, Sun Q, Kapadia DR. Integrating large language models into robotic autonomy: a review of motion, voice, and training pipelines. AI. 2025;6(7):158. doi:10.3390/ai6070158.
4. Espada JP, Qiu SY, Crespo RG, Carús JL. Leveraging large language models for autonomous robotic mapping and navigation. Int J Adv Rob Syst. 2025;22(2):17298806251325965. doi:10.1177/17298806251325965.
5. Kim Y, Kim D, Choi J, Park J, Oh N, Park D. A survey on integration of large language models with intelligent robots. Intell Serv Robot. 2024;17(5):1091–107. doi:10.1007/s11370-024-00550-5.
6. Russell SJ, Norvig P. Artificial intelligence: a modern approach. 3rd ed. Boston, MA, USA: Pearson; 2010. 22 p.
7. 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.
8. Xia L, Cui J, Shen R, Xu X, Gao Y, Li X. A survey of image semantics-based visual simultaneous localization and mapping: application-oriented solutions to autonomous navigation of mobile robots. Int J Adv Rob Syst. 2020;17(3):1729881420919185. doi:10.1177/1729881420919185.
9. Xu Z, Rong Z, Wu Y. A survey: which features are required for dynamic visual simultaneous localization and mapping? Vis Comput Ind Biomed Art. 2021;4(1):20. doi:10.1186/s42492-021-00086-w.
10. 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.2023.3248510.
11. Taheri H, Xia ZC. SLAM; definition and evolution. Eng Appl Artif Intell. 2021;97(5):104032. doi:10.1016/j.engappai.2020.104032.
12. Malathi V, Sreedharan P, Rthuraj PR, Kumar VA, Sadasivan AL, Udupa G, et al. Decision-making for path planning of mobile robots under uncertainty: a review of belief-space planning simplifications. Robotics. 2025;14(9):127. doi:10.3390/robotics14090127.
13. Ohno K, Nomura T, Tadokoro S. Real-time robot trajectory estimation and 3D map construction using 3D camera. In: Proceedings of the 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems; 2006 Oct 9–15; Beijing, China. p. 5279–85. doi:10.1109/IROS.2006.282027.
14. Tsintotas KA, Bampis L, Gasteratos A. The revisiting problem in simultaneous localization and mapping: a survey on visual loop closure detection. IEEE Trans Intell Transp Syst. 2022;23(11):19929–53. doi:10.1109/TITS.2022.3175656.
15. Miao L, Liu W, Deng Z. A frontier review of semantic SLAM technologies applied to the open world. Sensors. 2025;25(16):4994. doi:10.3390/s25164994.
16. Tezerjani MD, Khoshnazar M, Tangestanizadeh M, Kiani A, Yang Q. A survey on reinforcement learning applications in SLAM. J Mach Learn Deep Learn. 2024;1(1):20–31. doi:10.64820/aepjmldl.11.20.31.122024.
17. Chaplot DS. Building intelligent autonomous navigation agents. arXiv:2106.13415. 2021.
18. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372:n71.
19. Zimmer D, Jurišić M, Plaščak I, Barač Ž, Glavaš H, Radočaj D, et al. A systematic review of intelligent navigation in smart warehouses using prisma: integrating AI, SLAM, and sensor fusion for mobile robots. Eng. 2025;6(12):339. doi:10.3390/eng6120339.
20. Luchini C, Veronese N, Nottegar A, Shin JI, Gentile G, Granziol U, et al. Assessing the quality of studies in meta-research: review/guidelines on the most important quality assessment tools. Pharm Stat. 2021;20(1):185–95. doi:10.1002/pst.2068.
21. Ran T, Yuan L, Zhang J, Tang D, He L. RS-SLAM: a robust semantic SLAM in dynamic environments based on RGB-D sensor. IEEE Sens J. 2021;21(18):20657–64. doi:10.1109/JSEN.2021.3099511.
22. Liao Z, Hu Y, Zhang J, Qi X, Zhang X, Wang W. SO-SLAM: semantic object SLAM with scale proportional and symmetrical texture constraints. IEEE Robot Autom Lett. 2022;7(2):4008–15. doi:10.1109/LRA.2022.3148465.
23. Tian Y, Chang Y, Arias FH, Nieto-Granda C, How JP, Carlone L. Kimera-multi: robust, distributed, dense metric-semantic SLAM for multi-robot systems. IEEE Trans Robot. 2022;38(4):2022–38. doi:10.1109/tro.2021.3137751.

24. Hughes N, Chang Y, Carlone L. Hydra: a real-time spatial perception system for 3D scene graph construction and optimization. arXiv:2201.13360. 2022.
25. Yang L, Wang L. A semantic SLAM-based dense mapping approach for large-scale dynamic outdoor environment. Measurement. 2022;204(5):112001. doi:10.1016/j.measurement.2022.112001.
26. Driess D, Xia F, Sajjadi MS, Lynch C, Chowdhery A, Ichter B, et al. Palm-E: an embodied multimodal language model. arXiv:2303.03378. 2023.
27. Martins TB, Oswald MR, Civera J. OVO-SLAM: open-vocabulary online simultaneous localization and mapping. arXiv:2411.15043. 2024.
28. Teed Z, Deng J. DROID-SLAM: deep visual SLAM for monocular, stereo, and RGB-D cameras. Adv Neural Inf Process Syst. 2021;34:16558–69.
29. Hong J, Choi R, Leonard JJ. Semantic enhancement for object SLAM with heterogeneous multimodal large language model agents. In: Proceedings of the 2025 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS); 2025 Oct 19–25; Hangzhou, China.
30. 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.
31. Bescos B, Campos C, Tardos JD, Neira J. DynaSLAM II: tightly-coupled multi-object tracking and SLAM. IEEE Robot Autom Lett. 2021;6(3):5191–8. doi:10.1109/lra.2021.3068640.
32. Scona R, Jaimez M, Petillot YR, Fallon M, Cremers D. StaticFusion: background reconstruction for dense RGB-D SLAM in dynamic environments. In: Proceedings of the 2018 IEEE International Conference on Robotics and Automation (ICRA); 2018 May 21–25; Brisbane, Australia. p. 3849–56. doi:10.1109/ICRA.2018.8460681.
33. Rosinol A, Violette A, Abate M, Hughes N, Chang Y, Shi J, et al. Kimera: from SLAM to spatial perception with 3D dynamic scene graphs. Int J Robot Res. 2021;40(12–14):1510–46. doi:10.1177/02783649211056674.
34. Chaplot DS, Gandhi D, Gupta S, Gupta A, Salakhutdinov R. Learning to explore using active neural SLAM. arXiv:2004.05155. 2020.
35. Ahmed MF, Maragliano M, Frémont V, Recchiuto CT. Efficient multi-robot active slam. J Intell Robot Syst. 2025;111(2):64. doi:10.1007/s10846-025-02275-8.
36. Labbe M, Michaud F. RTAB-map as an open-source lidar and visual simultaneous localization and mapping library for large-scale and long-term online operation. J Field Robot. 2019;36(2):416–46. doi:10.1002/rob.21831.
37. Chang Y, Ebadi K, Denniston CE, Ginting MF, Rosinol A, Reinke A, et al. LAMP 2.0: a robust multi-robot SLAM system for operation in challenging large-scale underground environments. IEEE Robot Autom Lett. 2022;7(4):9175–82. doi:10.1109/lra.2022.3191204.
38. You Y, Wei P, Cai J, Huang W, Kang R, Liu H. MISD-SLAM: multimodal semantic SLAM for dynamic environments. Wirel Commun Mob Comput. 2022;2022(1):7600669. doi:10.1155/2022/7600669.
39. Abate M, Chang Y, Hughes N, Carlone L. Kimera2: robust and accurate metric-semantic SLAM in the real world. arXiv:2401.06323. 2024.
40. Zheng S, Wang J, Rizos C, Ding W, El-Mowafy A. Simultaneous localization and mapping (SLAM) for autonomous driving: concept and analysis. Remote Sens. 2023;15(4):1156. doi:10.3390/rs15041156.
41. Denniston CE, Chang Y, Reinke A, Ebadi K, Sukhatme GS, Carlone L, et al. Loop closure prioritization for efficient and scalable multi-robot SLAM. arXiv:2205.12402. 2022.
42. Campos C, Elvira R, Rodriguez JJG, Montiel JMM, Tardos JD. ORB-SLAM3: an accurate open-source library for visual, visual-inertial, and multimap SLAM. IEEE Trans Robot. 2021;37(6):1874–90. doi:10.1109/tro.2021.3075644.
43. Wang Y, Xu K, Tian Y, Ding X. DRG-SLAM: a semantic RGB-D SLAM using geometric features for indoor dynamic scene. In: Proceedings of the 2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS); 2022 Oct 23–27; Kyoto, Japan.
44. Shah D, Osiński BZ, Ichter B, Levine S. LM-nav robotic navigation with large pre-trained models of language, vision, and action. In: Proceedings of the Conference on Robot Learning (CoRL); 2022 Dec 14–18; Auckland, New Zealand. p. 492–504.
45. Ha D, Schmidhuber J. World models. arXiv:1803.10122. 2018.

46. Schulman J, Wolski F, Dhariwal P, Radford A, Klimov O. Proximal policy optimization algorithms. arXiv:1707.06347. 2017.
47. Chen J, Yuan B, Tomizuka M. Deep imitation learning for autonomous driving in generic urban scenarios with enhanced safety. In: Proceedings of the 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS); 2019 Nov 3–8; Macau, China. p. 2884–90. doi:10.1109/IROS40897.2019.8968225.
48. Savva M, Kadian A, Maksymets O, Zhao Y, Wijmans E, Jain B, et al. Habitat: a platform for embodied AI research. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV); 2019 Oct 27–Nov 2; Seoul, Republic of Korea. p. 9339–47.
49. Younesi A, Ansari M, Fazli M, Ejlali A, Shafique M, Henkel J. A comprehensive survey of convolutions in deep learning: applications, challenges, and future trends. IEEE Access. 2024;12(2):41180–218. doi:10.1109/ACCESS.2024.3376441.
50. Chiun J, Zhang S, Wang Y, Cao Y, Sartoretti G. MARVEL: multi-agent reinforcement learning for constrained field-of-view multi-robot exploration in large-scale environments. In: Proceedings of the 2025 IEEE International Conference on Robotics and Automation (ICRA); 2025 May 19–23; Atlanta, GA, USA. p. 11392–8. doi:10.1109/ICRA55743.2025.11127700.
51. Zhu AZ, Atanasov N, Daniilidis K. Event-based visual inertial odometry. In: Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 2017 Jul 21–26; Honolulu, HI, USA. p. 5816–24. doi:10.1109/cvpr.2017.616.
52. Haarnoja T, Zhou A, Abbeel P, Levine S. Soft actor-critic: off-policy maximum entropy deep reinforcement learning with a stochastic actor. In: Proceedings of the International Conference on Machine Learning; 2018 Jul 10–15; Stockholm, Sweden. p. 1856–65.
53. Finn C, Abbeel P, Levine S. Model-agnostic meta-learning for fast adaptation of deep networks. In: Proceedings of the 34th International Conference on Machine Learning (ICML); 2017 Aug 6–11; Sydney, Australia. p. 1126–35.
54. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. In: Proceedings of the 31st International Conference on Neural Information Processing Systems (NeurIPS); 2017 Dec 4–9; Long Beach, CA, USA. p. 4765–74.
55. Katz G, Barrett C, Dill DL, Julian K, Kochenderfer MJ. Reluplex: an efficient SMT solver for verifying deep neural networks. In: Computer aided verification. Cham, Switzerland: Springer; 2017. p. 97–117. doi:10.1007/978-3-319-63387-9_5.
56. Amado L, Pereira RF, Meneguzzi F. Robust neuro-symbolic goal and plan recognition. In: Proceedings of the 37th AAAI Conference on Artificial Intelligence (AAAI-23); 2023 Feb 7–14; Washington, DC, USA. p. 11937–44.
57. Vezhnevets AS, Osindero S, Schaul T, Heess N, Jaderberg M, Silver D, et al. FeUdal networks for hierarchical reinforcement learning. In: Proceedings of the 34th International Conference on Machine Learning (ICML); 2017 Aug 6–11; Sydney, Australia. p. 3540–9.
58. Chen Z, Zhang J, Shen S, Fu W, Cao L. Memory-augmented continuous-time neural policy for vision-guided embodied navigation. In: Proceedings of the 7th ACM International Conference on Multimedia in Asia; 2025 Dec 9–12; Kuala Lumpur, Malaysia. p. 1–8. doi:10.1145/3769748.3773343.
59. Ahn M, Brohan A, Brown N, Chebotar Y, Cortes O, David B, et al. Do as I can, not as I say: grounding language in robotic affordances. In: Proceedings of the Conference on Robot Learning (CoRL); 2022 Dec 14–18; Auckland, New Zealand.
60. Esser P, Rombach R, Ommer B. Taming transformers for high-resolution image synthesis. In: Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR); 2021 Jun 20–25; Nashville, TN, USA. p. 12868–78. doi:10.1109/cvpr46437.2021.01268.
61. Bloesch M, Czarnowski J, Clark R, Leutenegger S, Davison AJ. CodeSLAM—learning a compact, optimisable representation for dense visual SLAM. In: Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition; 2018 Jun 18–23, Salt Lake City, UT, USA. p. 2560–8. doi:10.1109/CVPR.2018.00271.
62. Brohan A, Brown N, Carbajal J, Chebotar Y, Dabis J, Finn C, et al. RT-2: vision-language-action models transfer web knowledge to robotic control. In: Proceedings of the Conference on Robot Learning (CoRL); 2023 Nov 6–9; Atlanta, GA, USA. p. 2165–83.

63. Mu Y, Zhang Q, Hu M, Wang W, Ding M, Jin J, et al. EmbodiedGPT: vision-language pre-training via embodied chain of thought. In: Proceedings of the Advances in Neural Information Processing Systems 36; 2023 Dec 10–16; New Orleans, LA, USA. p. 25081–94. doi:10.52202/075280-1090.
64. Kirkpatrick J, Pascanu R, Rabinowitz N, Veness J, Desjardins G, Rusu AA, et al. Overcoming catastrophic forgetting in neural networks. Proc Natl Acad Sci U S A. 2017;114(13):3521–6. doi:10.1073/pnas.1611835114.
65. Huang W, Wang C, Zhang R, Li Y, Wu J, Li FF. VoxPoser: composable 3D value maps for robotic manipulation with language models. In: Proceedings of the Conference on Robot Learning (CoRL); 2023 Nov 6–9; Atlanta, GA, USA.
66. Zhang H, Song Z, Wen F. Probabilistic relative pose calibration for object-level multi-agent cooperative perception. In: Proceedings of the 2024 IEEE Intelligent Vehicles Symposium (IV); 2024 Jun 2–5; Jeju Island, Republic of Korea. p. 496–501. doi:10.1109/IV55156.2024.10588531.
67. Tian Y, Liu K, Ok K, Tran L, Allen D, Roy N, et al. Search and rescue under the forest canopy using multiple UAVs. Int J Robot Res. 2020;39(10–11):1201–21. doi:10.1177/0278364920929398.
68. Rosen DM, Carlone L, Bandeira AS, Leonard JJ. SE-Sync: a certifiably correct algorithm for synchronization over the special Euclidean group. Int J Robot Res. 2019;38(2–3):95–125. doi:10.1177/0278364918784361.
69. Rusu RB, Blodow N, Beetz M. Fast point feature histograms (FPFH) for 3D registration. In: Proceedings of the 2009 IEEE International Conference on Robotics and Automation; 2009 May 12–17; Kobe, Japan. p. 3212–7.
70. Lamport L, Shostak R, Pease M. The byzantine generals problem. In: Concurrency: the works of Leslie Lamport. New York, NY, USA: Association for Computing Machinery; 2019. doi:10.1145/3335772.3335936.
71. Lv B, Duan S. Learning-based multi-robot active SLAM: a conceptual framework and survey. Appl Sci. 2026;16(3):1412. doi:10.3390/app16031412.
72. Rashid T, Samvelyan M, Schroeder de Witt C, Farquhar G, Foerster J, Whiteson S. Monotonic value function factorisation for deep multi-agent reinforcement learning. J Mach Learn Res. 2020;21(178):1–51.
73. Foerster J, Farquhar G, Afouras T, Nardelli N, Whiteson S. Counterfactual multi-agent policy gradients. In: Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence; 2018 Feb 2–7; New Orleans, LA, USA. doi:10.1609/aaai.v32i1.11794.
74. Mahboob H, Yasin JN, Jokinen S, Haghbayan MH, Plosila J, Yasin MM. DCP-SLAM: distributed collaborative partial swarm SLAM for efficient navigation of autonomous robots. Sensors. 2023;23(2):1025. doi:10.3390/s23021025.
75. Lu Y, Zhang S, Liu C, Zhang R, Ai B, Niyato D, et al. Agentic graph neural networks for wireless communications and networking toward edge general intelligence: a survey. IEEE Commun Surv Tutorials. 2026;28(127):4519–54. doi:10.1109/comst.2026.3651990.
76. Schmuck P, Ziegler T, Karrer M, Perraudin J, Chli M. COVINS: visual-inertial SLAM for centralized collaboration. In: Proceedings of the 2021 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct); 2021 Oct 4–8; Bari, Italy. p. 171–6. doi:10.1109/ismar-adjunct54149.2021.00043.
77. Chaplot DS, Gandhi DP, Gupta A, Salakhutdinov RR. Object goal navigation using goal-oriented semantic exploration. Adv Neural Inf Process Syst. 2020;33:4247–58.
78. Wang H, Wang J, Agapito L. Co-SLAM: joint coordinate and sparse parametric encodings for neural real-time SLAM. In: Proceedings of the 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR); 2023 Jun 17–24; Vancouver, BC, Canada. p. 13293–302. doi:10.1109/CVPR52729.2023.01277.
79. Lajoie PY, Ramtoula B, Chang Y, Carlone L, Beltrame G. DOOR-SLAM: distributed, online, and outlier resilient SLAM for robotic teams. IEEE Robot Autom Lett. 2020;5(2):1656–63. doi:10.1109/LRA.2020.2967681.
80. Zuo C, Feng Z, Xiao X. CCMD-SLAM: communication-efficient centralized multirobot dense SLAM with real-time point cloud maintenance. IEEE Trans Instrum Meas. 2024;73:7504812. doi:10.1109/TIM.2024.3398100.
81. van den Oord A, Vinyals O, Kavukcuoglu K. Neural discrete representation learning. In: Proceedings of the Advances in Neural Information Processing Systems; 2017 Dec 4–9; Long Beach, CA, USA.
82. Hu G, Zhu Y, Zhao D, Zhao M, Hao J. Event-triggered communication network with limited-bandwidth constraint for multi-agent reinforcement learning. IEEE Trans Neural Netw Learn Syst. 2023;34(8):3966–78. doi:10.1109/TNNLS.2021.3121546.

83. Zhu Y, Wan Hasan WZ, Harun Ramli HR, Norsahperi NMH, Mohd Kassim MS, Yao Y. Deep reinforcement learning of mobile robot navigation in dynamic environment: a review. Sensors. 2025;25(11):3394. doi:10.3390/s25113394.
84. Zhang S, Li Z, Yin Y, Xu S, Chaudhary V, Xu H. A survey on multi-robot collaboration systems: architectures, performances, and applications. Preprint. 2025. doi:10.36227/techrxiv.176045766.60277537/v2.
85. Yu X, Queralta JP, Heikkonen J, Westerlund T. Federated learning in robotic and autonomous systems. Procedia Comput Sci. 2021;191(6):135–42. doi:10.1016/j.procs.2021.07.041.
86. Lajoie PY, Beltrame G. Swarm-SLAM: sparse decentralized collaborative simultaneous localization and mapping framework for multi-robot systems. IEEE Robot Autom Lett. 2024;9(1):475–82. doi:10.1109/LRA.2023.3333742.
87. Xu H, Liu P, Chen X, Shen S. D^2SLAM: decentralized and distributed collaborative visual-inertial SLAM system for aerial swarm. IEEE Trans Robot. 2024;40:3445–64. doi:10.1109/TRO.2024.3422003.
88. Chang Y, Hughes N, Ray A, Carlone L. Hydra-multi: collaborative online construction of 3D scene graphs with multi-robot teams. In: Proceedings of the 2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS); 2023 Oct 1–5; Detroit, MI, USA. p. 10995–1002. doi:10.1109/IROS55552.2023.10341838.
89. Lajoie PY, Soma K, Bong HM, Lemieux-Bourque A, Zhang R, Varadharajan VS, et al. Multirobot decentralized collaborative SLAM in planetary analogue environments: dataset, challenges, and lessons learned. IEEE Trans Field Robot. 2025;2:516–28. doi:10.1109/tfr.2025.3586209.
90. Geiger A, Lenz P, Stiller C, Urtasun R. Vision meets robotics: the KITTI dataset. Int J Robot Res. 2013;32(11):1231–7. doi:10.1177/0278364913491297.
91. Sturm J, Engelhard N, Endres F, Burgard W, Cremers D. A benchmark for the evaluation of RGB-D SLAM systems. In: Proceedings of the 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems; 2012 Oct 7–12; Vilamoura-Algarve, Portugal. p. 573–80. doi:10.1109/IROS.2012.6385773.
92. Burri M, Nikolic J, Gohl P, Schneider T, Rehder J, Omari S, et al. The EuRoC micro aerial vehicle datasets. Int J Robot Res. 2016;35(10):1157–63. doi:10.1177/0278364915620033.
93. Straub J, Whelan T, Ma L, Chen Y, Wijmans E, Green S, et al. The replica dataset: a digital replica of indoor spaces. arXiv:1906.05797. 2019.
94. Dai A, Chang AX, Savva M, Halber M, Funkhouser T, Niessner M. ScanNet: richly-annotated 3D reconstructions of indoor scenes. In: Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 2017 Jul 21–26; Honolulu, HI, USA. p. 2432–43. doi:10.1109/cvpr.2017.261.
95. Maddern W, Pascoe G, Linegar C, Newman P. 1 year, 1000 km: the Oxford RobotCar dataset. Int J Robot Res. 2017;36(1):3–15. doi:10.1177/0278364916679498.
96. Dosovitskiy A, Ros G, Codevilla F, Lopez A, Koltun V. CARLA: an open urban driving simulator. In: Proceedings of the 1st Annual Conference on Robot Learning; 2017 Nov 13–15; Mountain View, CA, USA. p. 1–16.
97. Shah S, Dey D, Lovett C, Kapoor A. AirSim: high-fidelity visual and physical simulation for autonomous vehicles. In: Proceedings of the Field and Service Robotics: Results of the 11th International Conference; 2017 Sep 12–15; Zurich, Switzerland. p. 621–35.
98. Bonetto E, Xu C, Ahmad A. GRADE: generating realistic and dynamic environments for robotics research with Isaac Sim. Int J Robot Res. 2026;45(2):204–32. doi:10.1177/02783649251346211.
99. Li C, Xia F, Martín-Martín R, Lingelbach M, Srivastava S, Shen B, et al. iGibson 2.0: object-centric simulation for robot learning of everyday household tasks. In: Proceedings of the 5th Conference on Robot Learning; 2021 Nov 8–11; London, UK. p. 455–65.
100. Echeverria G, Lemaignan S, Degroote A, Lacroix S, Karg M, Koch P, et al. Simulating complex robotic scenarios with MORSE. In: Simulation, modeling, and programming for autonomous robots. Berlin/Heidelberg, Germany: Springer; 2012. p. 197–208. doi:10.1007/978-3-642-34327-8_20.
101. Mildenhall B, Srinivasan PP, Tancik M, Barron JT, Ramamoorthi R, Ng R. Nerf: representing scenes as neural radiance fields for view synthesis. Commun ACM. 2021;65(1):99–106. doi:10.1007/978-3-030-58452-8_24.
102. Tobin J, Fong R, Ray A, Schneider J, Zaremba W, Abbeel P. Domain randomization for transferring deep neural networks from simulation to the real world. In: Proceedings of the 2017 IEEE/RSJ International Conference on

Intelligent Robots and Systems (IROS); 2017 Sep 24–28; Vancouver, BC, Canada. p. 23–30. doi:10.1109/IROS.2017.8202133.
103. Zou Q, Sun Q, Chen L, Nie B, Li Q. A comparative analysis of LiDAR SLAM-based indoor navigation for autonomous vehicles. IEEE Trans Intell Transp Syst. 2022;23(7):6907–21. doi:10.1109/TITS.2021.3063477.
104. Wang K, Guo J, Chen K, Lu J. An in-depth examination of SLAM methods: challenges, advancements, and applications in complex scenes for autonomous driving. IEEE Trans Intell Transp Syst. 2025;26(7):11066–87. doi:10.1109/TITS.2025.3545479.
105. Shen Y, Wang S. Application of slam technology for autonomous mobile robots in complex environments. Appl Comput Eng. 2025;128(1):108–12. doi:10.54254/2755-2721/2025.20212.
106. Konieczna-Fuławka M, Koval A, Nikolakopoulos G, Fumagalli M, Santas-Moreu L, Vigara-Puche V, et al. Autonomous mobile inspection robots in deep underground mining—the current state of the art and future perspectives. Sensors. 2025;25(12):3598. doi:10.3390/s25123598.
107. Hao M, Ren J, Ji X, Bi Y, Zhao S, Wu M. Research on multimodal data enhanced SLAM algorithm for global mapping of underground coal mines. Sci Rep. 2025;15(1):37124. doi:10.1038/s41598-025-21053-y.
108. Islam R, Habibullah H, Hossain T. AGRI-SLAM: a real-time stereo visual SLAM for agricultural environment. Auton Rob. 2023;47(6):649–68. doi:10.1007/s10514-023-10110-y.
109. Lin JH, Lin TT. Development of a visual SLAM-based autonomous UAV system for greenhouse plant monitoring. Drones. 2026;10(3):205. doi:10.3390/drones10030205.
110. Huang Z, Cai J, Yang N, Li Z, He Y, Wei Y, et al. AELVI-SLAM: LiDAR-visual–inertial SLAM for autonomous exploration in multiple agricultural scenarios. Artif Intell Agric. 2026;16(2):940–57. doi:10.1016/j.aiia.2026.03.013.
111. Kim K, Deb A, Cappelleri DJ. P-AgSLAM: in-row and under-canopy SLAM for agricultural monitoring in cornfields. IEEE Robot Autom Lett. 2024;9(6):4982–9. doi:10.1109/LRA.2024.3386466.
112. Wu C, Gong Z, Tao B, Tan K, Gu Z, Yin ZP. RF-SLAM: UHF-RFID based simultaneous tags mapping and robot localization algorithm for smart warehouse position service. IEEE Trans Ind Inform. 2023;19(12):11765–75. doi:10.1109/TII.2023.3252405.
113. Ismail H, Roy R, Sheu LJ, Chieng WH, Tang LC. Exploration-based SLAM (e-SLAM) for the indoor mobile robot using lidar. Sensors. 2022;22(4):1689. doi:10.3390/s22041689.
114. Alhmiedat T, Marei AM, Messoudi W, Albelwi S, Bushnag A, Bassfar Z, et al. A SLAM-based localization and navigation system for social robots: the pepper robot case. Machines. 2023;11(2):158. doi:10.3390/machines11020158.
115. Yue X, Zhang Y, Chen J, Chen J, Zhou X, He M. LiDAR-based SLAM for robotic mapping: state of the art and new frontiers. Ind Robot Int J Robot Res Appl. 2024;51(2):196–205. doi:10.1108/ir-09-2023-0225.
116. Xiao F, Fang J, Guo X, Zhang Y, Huang R. Dynamic SLAM system for hospital logistics robots based on nonlinear optimal filtering and deep learning. Discover Comput. 2025;28(1):85. doi:10.1007/s10791-025-09607-0.
117. Zheng C, Kou Y, Xu D. Research on semantic-based simultaneous localization and mapping technology for medical and nursing robots. In: Proceedings of the Third International Computing Imaging Conference (CITA 2023); 2023 Jun 1–3; Sydney, Australia. p. 1085–90.
118. Ishikawa T, Taniguchi A, Hagiwara Y, Taniguchi T. Active semantic mapping for household robots: rapid indoor adaptation and reduced user burden. In: Proceedings of the 2023 IEEE International Conference on Systems, Man, and Cybernetics (SMC); 2023 Oct 1–4; Honolulu, Oahu, HI, USA. p. 3116–23. doi:10.1109/smc53992.2023.10394143.
119. Yan Y. The navigation for home service robot with the least power consumption. Sci Rep. 2025;15(1):35894. doi:10.1038/s41598-025-19760-7.
120. Noh D, Lee J, Park J, Kim HR, Baek S. Semantic navigation with embodied AI for humanoid robots in personalized household environments. In: Proceedings of the 2025 IEEE-RAS 24th International Conference on Humanoid Robots (Humanoids); 2025 Sep 30–Oct 2; Seoul, Republic of Korea. p. 1–8. doi:10.1109/humanoids65713.2025.11203027.
121. Patil LN. Precision mapping and navigation: a robotic restaurant management system using SLAM and ROS. Sigma J Eng Nat Sci Sigma Müh Fen Bil Derg. 2025:1233–47. doi:10.14744/sigma.2025.00117.

122. Maqbool M, Ashfaq MH, Nawaz NA, Ishaq K. A SLAM algorithm: a superior approach for room cleaning robot. In: Proceedings of the 2024 5th International Conference on Innovative Computing (ICIC); 2024 Nov 20–21; Lahore, Pakistan. p. 1–11. doi:10.1109/ICIC63915.2024.11116501.
123. Nakamura R, Kambe T, Hashimoto M, Takahashi K. SLAM-based mapping in truck-and-robot system for last-mile delivery automation. In: Proceedings of the 2023 IARIA Annual Congress on Frontiers in Science, Technology, Services, and Applications; 2023 Nov 13–17; Valencia, Spain.
124. Zhang J, Zhang X, Shen X, Wu J, Li Y. A lidar slam based on improved particle filter and scan matching for unmanned delivery robot. J Phys Conf Ser. 2023;2506(1):012009. doi:10.1088/1742-6596/2506/1/012009.
125. Gan Y, Xie Z. SLAM and deep reinforcement learning-based autonomous navigation delivery robot. In: Proceedings of the International Conference on Computer Vision and Augmented Reality (CVAR 2025); 2025 May 16–18; Xi'an, China.
126. Kim T, Kang G, Lee D, Shim DH. Development of an indoor delivery mobile robot for a multi-floor environment. IEEE Access. 2024;12:45202–15. doi:10.1109/ACCESS.2024.3381489.
127. Sapkota R, Roumeliotis KI, Karkee M. UAVs meet agentic AI: a multidomain survey of autonomous aerial intelligence and agentic UAVs. arXiv:2506.08045. 2025.
128. Colelough BC. Advancing frontiers in SLAM: a survey of symbolic representation and human-machine teaming in environmental mapping. arXiv:2405.01398. 2024.
129. Fang R, He P, Gao Y. A review of SLAM techniques and applications in unmanned aerial vehicles. J Phys Conf Ser. 2024;2798(1):012033. doi:10.1088/1742-6596/2798/1/012033.
130. Küçükerdem H, Yilmaz C, Kahraman HT, Sönmez Y. Autonomous control of unmanned aerial vehicles: applications, requirements. challenges Cluster Comput. 2025;28(11):734. doi:10.1007/s10586-025-05418-6.
131. Wang Z, Zhang X. Artificial intelligence-driven innovations in SLAM for robotics and autonomous navigation. In: Proceedings of the 2025 9th International Conference on Robotics, Control and Automation (ICRCA); 2025 Mar 7–9; Shanghai, China. p. 1–6.
132. Anbarasu B, Seralathan S, Muthuram A. Autonomous mapping and ai-based navigation using deep learning, slam, and optical flow for micro aerial vehicle. In: Artificial intelligence applications in aeronautical and aerospace engineering. Beverly, MA, USA: Scrivener Publishing LLC; 2025. p. 277–301.
133. Jharko E, Mamchenko M, Khripunov SP. Robot/UAV indoor visual SLAM in smart cities based on remote data processing. In: Proceedings of the 2023 International Russian Smart Industry Conference (SmartIndustryCon); 2023 Mar 27–31; Sochi, Russia.
134. Yadav R, Kala R. Fusion of visual odometry and place recognition for SLAM in extreme conditions. Appl Intell. 2022;52(10):5567–80. doi:10.1007/s10489-021-03050-6.
135. Zhang Z, Cheng Y, Bu L, Ye J. Rapid SLAM method for star surface rover in unstructured space environments. Aerospace. 2024;11(9):768. doi:10.3390/aerospace11090768.
136. Obaideen K, AlShabi MA. SLAM in XR: a bibliometric analysis of its application in virtual, augmented, and mixed reality. Artif Intell Mach Learn Multi Domain Oper Appl VI. 2024;13051:565–71.
137. Sheng X, Mao S, Yan Y, Yang X. Review on SLAM algorithms for Augmented Reality. Displays. 2024;84(2):102806. doi:10.1016/j.displa.2024.102806.
138. de Sousa CAP, Hamann H, Deussen O. Mesh2SLAM in VR: a fast geometry-based SLAM framework for rapid prototyping in virtual reality applications. In: Proceedings of the 2025 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW); 2025 Mar 8–12; Saint Malo, France. p. 57–62. doi:10.1109/VRW66409.2025.00021.
139. Raychaudhuri S, Chang AX. Semantic mapping in indoor embodied AI—a survey on advances, challenges, and future directions. arXiv:2501.05750. 2025. doi:10.48550/arXiv.2501.05750.
140. Wang X, Fan X, Shi P, Ni J, Zhou Z. An overview of key SLAM technologies for underwater scenes. Remote Sens. 2023;15(10):2496. doi:10.3390/rs15102496.
141. Merveille FFR, Jia B, Xu Z, Fred B. Advancements in sensor fusion for underwater SLAM: a review on enhanced navigation and environmental perception. Sensors. 2024;24(23):7490. doi:10.3390/s24237490.

142. Yang K, Zhang Z, Cui R, Yan W. Acoustic-optic assisted multisensor navigation for autonomous underwater vehicles. Ocean Eng. 2024;297(6):117139. doi:10.1016/j.oceaneng.2024.117139.
143. Chen L, Liu Y, Dong P, Liang J, Wang A. An intelligent navigation control approach for autonomous unmanned vehicles via deep learning-enhanced visual SLAM framework. IEEE Access. 2023;11:119067–77. doi:10.1109/ACCESS.2023.3326754.
144. McConnell J, Shariati A, Szenher P, Li Y. Above and below: heterogeneous multi-robot slam across surface and underwater domains. IEEE Robot Autom Lett. 2025;11(1):129–36.
145. Yang Q, Zheng J, Wang C, Xiong M, Xie G. Electro-SLAM: distributed underwater multi-robot SLAM via bio-inspired active and passive electro-sensing. Int J Robot Res. 2026:02783649251414426. doi:10.1177/02783649251414426.