SURVEY ARTICLE

WILEY

UAV-based simultaneous localization and mapping in outdoor environments: A systematic scoping review

Kaiwen Wang[1,2] | Lammmert Kooistra[3] | Ruoxi Pan[4] | Wensheng Wang[2] | João Valente[1] 🔟

[1]Information Technology Group, Wageningen University & Research, Wageningen, KN, The Netherlands

[2]Agricultural Information Institute, Chinese Academy of Agriculture Sciences, Beijing, China

[3]Laboratory of Geo-information Science and Remote Sensing, Wageningen University & Research, Wageningen, PB, The Netherlands

[4]Warwick Medical School, Warwick University, Coventry, UK

Correspondence
João Valente, Information Technology Group, Wageningen University & Research, Hollandseweg 1, 6706 KN, The Netherlands.
Email: joao.valente@wur.nl

Abstract

This study aims to investigate the current knowledge of unmanned aerial vehicle (UAV)-based simultaneous localization and mapping (SLAM) in outdoor environments and to discuss challenges and limitations in this field. A literature search was conducted in three online databases (Web of Science, Scopus, and IEEE) for articles published before October 2022 related to UAV-based SLAM. A scoping review was carried out to identify the key concepts and applications, and discover research gaps in the use of algorithm-oriented and task-oriented, open-source studies. A total of 97 studies met the criteria after conducting a two-step screening by a systematic method followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses. Among eligible studies, 97 were classified into two main categories: algorithm-oriented studies and task-oriented studies. The analysis of the literature revealed that the majority of the studies were focused on the development and implementation of new algorithms and algorithms. This review highlights the significance and diversity of sensors utilized in UAVs in different tasks and applications scenarios that employ different types of sensors. The evaluation method is able to show the real results and performance of the new algorithms in the target scenarios compared with the evaluation method by the public data set and simulation platform.

KEYWORDS
SLAM, UAVs

1 | INTRODUCTION

Robots of various shapes and sizes are available to assist people in their daily lives. From small drones and household cleaning robots to large self-driving cars, industrial robots, and Mars rovers, there is a wide range of robotic technologies available. Among these technologies, unmanned aerial vehicles (UAVs) have emerged as a versatile tool in industry, agriculture, geoscience, and remote sensing due to their flexibility and cost-effectiveness, making them a preferred choice over other platforms (Nex et al., 2022). UAVs can be classified into two types based on their usage scenarios: indoor UAV, such as search and rescue in indoor spaces, and outdoor UAV,

such as construction and infrastructure monitoring, as well as precision farming. However, compared with indoor UAV, the application of outdoor UAV involves more challenges due to changing variables, such as illumination conditions, dynamic environments, unexpected obstacles, changing weather conditions, and so forth (Shalal et al., 2013). These challenges increase the difficulty of UAVs' perception of their surroundings, making their operation and navigation more complex.

The robots' perception capabilities are primarily designed to support the robots' localization of itself in the environment and to map the observed environment. Normally, autonomous navigation of mobile robots is map-based. Although known maps are feasible for autonomous navigation and localization (Kim et al., 2007), the robot typically works in unknown environments (Li et al., 2012), especially for UAVs. In general, the navigation of UAVs uses simultaneous localization and mapping (SLAM) or beacon-based system for indoors and global positioning system (GPS)-based method for outdoors. However, obstacles such as tree canopies, walls, and buildings, as well as satellite availability can greatly affect the GPS positioning capability for outdoor UAVs' navigation. Therefore, SLAM has been identified as a crucial method for a UAV system to achieve localization, navigation, and mapping in unknown environments.

SLAM is a fundamental task in robotics, which allows a robot to navigate autonomously in an unknown or changing environment. SLAM algorithms typically use sensor data to create a map of the environment and estimate the robot's position and orientation relative to that map (Khairuddin et al., 2015). The theoretical basis for the SLAM algorithm is the paper published by Smith and Cheeseman in 1986 on the estimation of the position of objects in their environment (Smith & Cheeseman, 1986). This study demonstrated the mathematical feasibility of position estimation of objects in the environment, which provided a theoretical basis for using sensors to observe other objects in an unknown environment and to estimate their position and pose of themselves with the observed objects. On the basis of this theory, the field of SLAM has been further elaborated and different studies have developed SLAM algorithms that use a variation of sensors (e.g., radar, Light Detection and Ranging [LiDAR], monocular camera, binocular camera, RGB-D, sonar, inertial measurement unit [IMU], GPS, etc.) to provide observations for SLAM in an unknown environment. SLAM was initially used for the navigation and location of robots indoors due to the lack of GPS positioning information indoors. SLAM is also used in outdoor robots due to its complementary performance in robot localization and navigation work in GPS-denied environments. However, compared with general robots and mobile ground vehicles, UAV-based SLAM may have some challenges in outdoor environments such as the various application scenarios, different types of sensors, and the evaluation method due to the limited payload capacity and high flexibility of UAVs.

This study aims to investigate the current knowledge of UAV-based SLAM in outdoor environments and to discuss the challenges and limitations in this field. Although previous studies have already been summarized, compared, and analyzed on SLAM, most of the

SLAM studies have targeted a specific sensor, such as LiDAR (Smith & Cheeseman, 1986) or visual sensor (Siegwart et al., 2011), and have not been constrained to only specific types of robots. The findings in this review will be investigated by conducting a scoping review with a systematic method to the application of UAV-based SLAM in outdoor environments. Three main research questions (RQs) have been defined to extract interesting aspects of UAV-based SLAM methods applied to outdoor environments:

> **RQ1.** *What is the contribution of UAV-based SLAM within the current framework in outdoor environments?*

> **RQ2.** *What is the framework of UAV-based SLAM, including sensors, platforms, algorithms, and evaluation method, and their strengths and weaknesses?*

> **RQ3.** *What are the challenges, research gaps, and perspectives on the potential use of UAV-based SLAM in outdoor environments? The review analysis and discussion have been elaborated based on the formulated RQs. And this paper is organized as follows. Section 2 introduces the framework and background of SLAM. Section 3 describes the method for this review. Section 4 presents the quantitative results in the selected literature. Section 5 discusses the results according to the RQ of this study. Finally, Section 6 is the conclusion of this paper.*

2 | AN OVERVIEW OF SLAM

The SLAM method (Khairuddin et al., 2015) is employed by autonomous vehicles, such as ground mobile vehicles and UAVs, to concurrently create a map of their environment and determine their position within it. This allows the vehicle to navigate and avoid obstacles without relying on prior maps or external localization systems. SLAM algorithms integrate sensor information, such as LiDAR or visual data, and intricate mathematical calculations to generate the map and estimate the vehicle's location. SLAM is generally comprised of five distinct components (sensor, odometry, back-end optimization, loop closure, and map) that work together to form a sophisticated system (Figure 1). In the following paragraphs, these different components are explained in more detail.

2.1 | Sensors

The first step in implementing SLAM involves acquiring data from GPS and hardware devices, such as cameras, LiDAR, sonar sensors, and IMU. These sensors gather information on landmarks, including features and objects in the environment (Khairuddin et al., 2015). In general, the sensors can be categorized into three groups based on the information they acquire: range sensors that actively emit and receive energy and signals to measure distance to objects in the environment (e.g., LiDAR, sonar sensor, time-of-flight [ToF] camera,
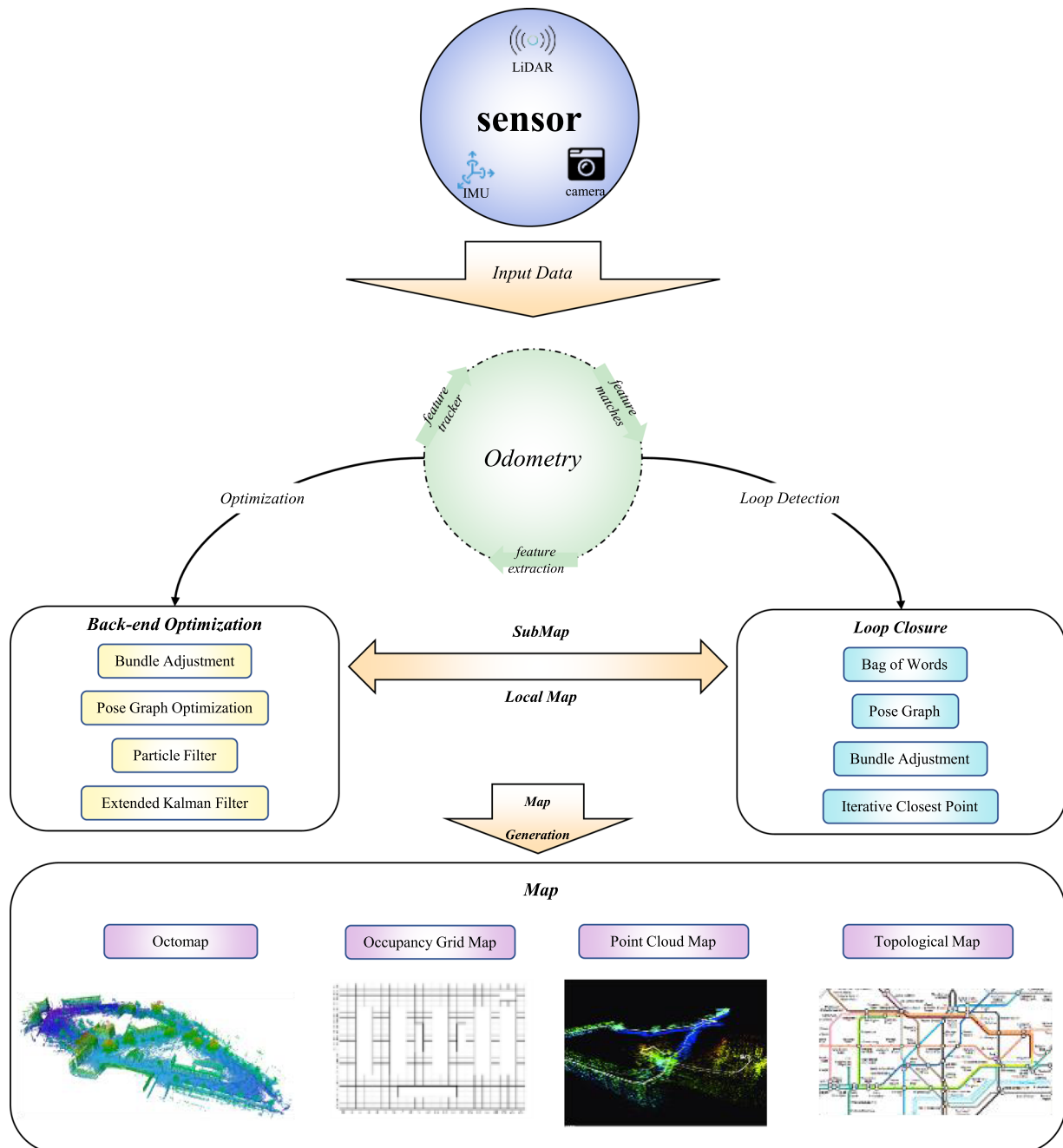
FIGURE 1 Overall SLAM framework consisting of five major components: sensors, odometry, back-end optimization, loop closure, and mapping. The sensors provide data to odometry, after data processing in odometry, the processed data would optimize and detect loop closure. Finally, the map of surroundings can be generated into different forms. SLAM, simultaneous localization and mapping. [Color figure can be viewed at wileyonlinelibrary.com]

and structured light sensor), passive sensors that passively receive and measure radiation emitted by natural or artificial sources instead of actively emitting energy or signals (e.g., monocular camera, stereo camera, and RGB-D camera) and orientation sensors which measure the orientation or rotation of the robot or vehicle (e.g., IMU) (Siegwart et al., 2011). Early SLAM methods primarily used range sensors, such as acoustic sensors or LiDAR, which can provide accurate depth data (Dissanayake, Newman, et al., 2001; Zaffar

et al., 2018). However, since 2005, research into visual SLAM has gained momentum, particularly in the use of RGB cameras due to the increasing ubiquity of cameras (Karlsson et al., 2005). Although visual sensors offer abundant feature information, they lack depth or distance information (Zaffar et al., 2018). However, depth information can be acquired from stereo cameras or RGB-D cameras. As a result, visual SLAM has become a prominent research topic and practical application area.

2.2 | Odometry

Sensors supply data with landmarks such as features and objects for SLAM (Khairuddin et al., 2015) which can be used to initialize the position of the robot (Cadena et al., 2016; Khairuddin et al., 2015). Odometry involves measuring the movement of the robot or sensor over time and estimating its distance traveled by extracting the features and landmarks and calculating the changes in them (Zhang & Singh, 2015). By monitoring changes in movements, the robot can determine its position within a known map of the environment or create a map from scratch.

Odometry, including visual odometry, and LiDAR odometry, can use various sensors to obtain features and estimate motions. Two common visual odometry methods are currently available: the feature-based method, and the appearance-based method (Aqel et al., 2016; Nistér et al., 2004; Scaramuzza & Fraundorfer, 2011). The camera captures individual images in a sequence, which can be created into video frames, and the feature-based method tracks distinctive features (e.g., points, lines, and planes) extracted from video frames to match and track the distinctive ones among the extracted features, which can estimate sensor motion (Naroditsky et al., 2011; Nistér et al., 2006). Currently, there are many different feature points (such as oriented fast and rotated brief [ORB] points, Rublee et al., 2011; scale-invariant feature transform points, Lowe, 2004; speeded up robust features points, Bay et al., 2006) which can be used in visual odometry. Instead of extracting and tracking features, the appearance-based method monitors changes in the appearance of acquired images and the intensity of pixel information within the images (Bellotto et al., 2008; Gonzalez et al., 2012, 2013). Normally, the optical flow-based method computes the displacement of brightness patterns from one image frame to another by using the intensity values of neighboring pixels (Barron et al., 1994; Campbell et al., 2004). Similar to visual odometry, LiDAR odometry acquires distance feature information (Point to Point, Point to Plane, or Point to Edge) from LiDAR point cloud data and estimates the motion and the position of the sensors (Ji & Singh, 2014).

2.3 | Back-end optimization and loop closure

Odometry can be quite accurate in predicting the robot's position and orientation over a short time interval but can be subject to drift and accumulated errors over longer distances and time. However, solutions like back-end optimization and loop closure are required to overcome the drift and accumulated error.

In the back end of SLAM, the mapping and localization process is performed recursively by computing the movements from the features and landmarks extracted from input data (Khairuddin et al., 2015). To optimize the estimation of movement, the extended Kalman filter (EKF) method was first introduced in SLAM as EKF-SLAM (Leonard & Durrant-Whyte, 1991). EKF can handle nonlinear

systems, reducing the impact of noise on the estimation and prediction of the SLAM system (Bar-Shalom et al., 2001). Later, Fast-SLAM was introduced by integrating particle filter (PF) and EKF (Montemerlo et al., 2002), further improving the accuracy of motion and position estimation. However, unavoidable noise (such as sensor observation noise and odometry noise) can still lead to inaccuracies in the estimation of SLAM over time. To reduce these errors, loop-closure detection was introduced (Ho & Newman, 2007). With loop closure, the robot can recognize when it returns to a previously visited location and make the necessary corrections to the global error. Currently, popular loop-closure detection methods include bag of words (Sivic & Zisserman, 2008), pose graph (Lu & Milios, 1997), and bundle adjust (Triggs et al., 2000).

2.4 | Mapping

When the SLAM is executed, many local maps and submaps (Figure 1) containing local coordinate systems are generated for the estimation of motion (Khairuddin et al., 2015). These maps are processed every few seconds or per frame. Eventually, various representations of maps are generated in SLAM, such as point cloud map (Rusu & Cousins, 2011), Octomap (Hornung et al., 2013), and occupancy grid map (Fankhauser & Hutter, 2016). The resulting maps can be used in different fields for path planning, obstacle avoidance, remote sensing, and 3D reconstruction.

2.5 | Milestones of SLAM

In the last three decades, the field of SLAM has made significant progress based on the theory of object position estimation (Smith & Cheeseman, 1986). Currently, there are various architectures of SLAM algorithms using different types of sensors. To clearly illustrate the differences between different SLAM algorithms and to show the evolution of SLAM, the flowchart in Figure 2 presents the SLAM milestones and main algorithms from the first complete SLAM method (EKF-SLAM) (Smith, 1988) to the most recent approaches. We divide SLAM methods into two categories according to the sensors: visual SLAM and LiDAR SLAM. Initially SLAM was conducted in ground vehicles, but as UAVs became vehicles for remote sensing-based monitoring, SLAM also started to be developed for UAV-based applications. The adoption of SLAM for UAV-based sensors with monocular camera started to emerge in 2003 (Kim & Sukkarieh, 2003), this was also the beginning of visual SLAM (Davison, 2003).

Until the introduction of MonoSLAM (Davison, 2003) in 2003, LiDAR was the dominant sensor in SLAM. Furthermore, filtering methods such as EKF and PF were also the main optimization solution in SLAM. After that, both LiDAR and visual SLAM came to several eras and there were some milestone studies, which are explained in the following subsections.
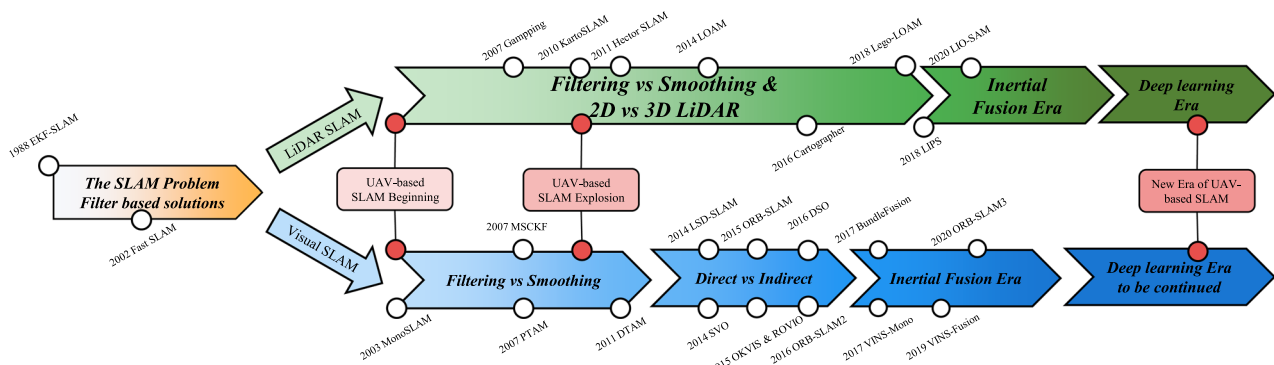
FIGURE 2   The history and milestones of general SLAM, the path to SLAM divided into LiDAR-based (green blocks) and Visual-based (blue blocks). The red blocks show the beginning of UAV-based SLAM in 2003, the explosion of UAV applications, and the new era of UAV-based SLAM. The path in LiDAR-based SLAM was through 2D versus 3D LiDAR era, and the path in visual SLAM was through Filtering versus Smoothing era, Direct versus Indirect era, and Inertial Fusion era. And the next stage for LiDAR-based and visual SLAM was deep learning era and sensor fusion. 2D, two dimensional; LiDAR, Light Detection and Ranging; SLAM, simultaneous localization and mapping; UAV, unmanned aerial vehicle. [Color figure can be viewed at wileyonlinelibrary.com]

2.5.1 | Filtering methods or smoothing methods

In SLAM, filtering methods are used to estimate the state of the system based on sensory information, while smoothing methods are used to improve the accuracy of these estimates over time. Filtering methods, such as Kalman filtering and particle filtering, use a mathematical model of the system and sensory information to estimate the state of the system at a particular point in time. Smoothing methods, on the other hand, use information about the past history of the system to improve the accuracy of these estimates. While filtering methods focus on estimating the state of the system at a single point in time, smoothing methods take into account the entire history of the system to produce more accurate estimates.

From 2003 to 2011, many visual SLAM algorithms, such as EKF (PTAM; Klein & Murray, 2007), multistate constraint Kalman filter (MSCKF; Mourikis & Roumeliotis, 2007), dense tracking and mapping (Newcombe et al., 2011) tried to compare which method is better or combine the advantages of filtering methods and smoothing methods among them. Meanwhile, LiDAR SLAM also experienced a similar period to determine whether filter-based or smoothing-based methods, such as Gmapping (Grisetti et al., 2005), Karto-SLAM (Konolige et al., 2010), and Cartographer (Hess et al., 2016).

2.5.2 | Direct visual odometry or indirect visual odometry

Direct visual odometry involves using visual sensors, such as cameras, to directly measure the motion of a vehicle or robot. This is typically done by tracking the features of the images from the cameras and using that information to estimate the movement of the vehicle. Indirect visual odometry, on the other hand, involves

using visual sensors to indirectly estimate the motion of a vehicle or robot by comparing the images from the cameras to a prebuilt map of the environment. This allows the vehicle or robot to localize itself within the map and use that information to estimate its motion over time.

There are several open-source SLAM algorithms using direct method or indirect method in visual odometry, such as large-scale direct (LSD)-SLAM (Engel et al., 2014), ORB-SLAM (Mur-Artal et al., 2015), semidirect visual odometry (SVO; Forster et al., 2014), and open keyframe-based visual-inertial SLAM (OKVIS, Leutenegger et al., 2015) and robust visual inertial odometry (ROVIO, Bloesch et al., 2015).

2.5.3 | Two-dimensional (2D) and 3D LiDAR

When LiDAR SLAM research argued whether filtering methods or smoothing methods were better, 2D and 3D LiDAR SLAM also started to be debated. The difference between 2D and 3D LiDAR is that 2D LiDAR measures the distance to objects in a single plane while 3D LiDAR measures in multiple planes. LiDAR odometry and mapping (LOAM; Ji & Singh, 2014) is one of the most important milestones in 3D LiDAR SLAM, which proposed a complete and pure LiDAR sensor SLAM framework. Lego-LOAM (Shan & Englot, 2018) was developed as a lightweight model and better optimization based on LOAM.

2.5.4 | Inertial fusion era

Inertial fusion is an approach that combines information from inertial sensors, such as gyroscopes, accelerometers, IMUs, with other types of sensors, such as cameras, to improve the accuracy of SLAM.

Previous studies in visual SLAM, such as monocular visual-inertial system (VINS-Mono; Qin et al., 2018) and VINS-Fusion (Qin et al., 2019), have reported that inertial sensors provide high-frequency information about the motion of robot or sensor, which can be used to improve the accuracy of the SLAM estimates over time. Inertial sensors also help correct drift and accumulate errors. Moreover, inertial fusion allows SLAM to be more robust to individual sensor errors or failures, which can improve the overall reliability of SLAM. A similar era also occurred in LiDAR SLAM, such as LiDAR inertial odometry via smoothing and mapping (LIO-SAM; Shan et al., 2020) and LiDAR-inertial 3D plane SLAM (Geneva et al., 2018), which proposed to make fusion with LiDAR and inertial sensor to improve the robustness and accuracy of SLAM.

3 | METHODOLOGY

The purpose of the scoping review approach (Arksey & O'Malley, 2005; Munn et al., 2018) is to survey and outline the key concepts, types of evidence, and research gaps in a specific subject area. It is typically performed at the start of a research project to pinpoint the most pertinent and significant studies and to determine areas in need of further investigation. Unlike a systematic review, which is a comprehensive and thorough review of the literature, a scoping review is broader and has more exploratory approach aimed at gaining an overview of the state of research in a particular field. With this in mind, a scoping review was carried out to identify the key concepts and applications, and discover research gaps in the use of SLAM for outdoor UAV-based applications.

This scoping review adopted a systematic approach in designing search strategies and selection criteria to ensure rigor, credibility, transparency, and reproducibility. The designing, searching, and implementation of the systemic method were conducted in October 2022. The screening workflow followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Moher et al., 2009).

3.1 | Search strategy

The search strategy for this scoping review was designed according to the guidelines of Peters et al. (2015) for conducting systematic scoping reviews. A literature search was conducted in three online databases (Web of Science, Scopus, and IEEE) for articles published before October 2022 related to UAVs, SLAM, and outdoors. Detailed literature search terms are described in Table 1 and the search expressions used in the databases can also be found in Appendix A.

The search strategy was specifically designed to identify literature related to UAV-based SLAM in outdoor environments, using a combination of keywords, synonyms, and abbreviations to maximize the retrieval of relevant studies. The resulting literature was then imported into EndNote 20.04 for further screening and literature management.

TABLE 1 The search terms and expressions used for searching the selected literature databases.

Facet A	SLAM
	SLAM OR simultaneous localization and mapping
Facet B	UAV platform
	UAV OR UAS OR MAV OR microaerial vehicle OR unmanned aerial vehicle OR unmanned aerial system OR drone
Facet C	Outdoor
	Outdoor OR field OR orchard OR forest OR vineyard OR park OR mountain OR grassland OR (GPS-denied NOT indoor)
Combined	A AND B AND C

Abbreviations: GPS, global positioning system; UAS, unmanned aircraft system; UAV, unmanned aerial vehicle.

3.2 | Inclusion/exclusion criteria for screening searched literature

This study employed a systematic approach to carry out a scoping literature review of the published literature in the field of UAV-based SLAM. Duplicates were removed, and a five-step literature screening strategy was implemented with defined inclusion and exclusion criteria, as detailed in Table 2. The screening process involved two stages (Figure 3), with the first being a rapid screening based on criteria A–E that only evaluates the title, abstract, and keywords. The second stage was a thorough screening based on criteria B–E, where the full text of the literature that passed the first stage was evaluated.

3.3 | Screening results and data extraction

After the removal of the duplicate literature (n = 205) and literature cannot be accessed (n = 12) from the searching result (n = 575), there are 113 literatures from Web of Science, 271 literature from Scopus, and 191 literature from IEEE (Figure 3). And 97 literature met the criteria after conducting a two-step screening. Some necessary information such as the title of the studies, publish year, author, journal and conference, country, sensors used in the studies, SLAM simulation platforms, UAV platforms, flight height, and so forth was extracted from the selected literature to an excel file, which can be found in Appendix A.

4 | RESULTS

4.1 | Overview of the reviewed papers

Overall, this review analyzed 97 studies on UAV-based outdoor applications with SLAM published in a range of journals (n = 58) (Figure A1) over a span of 19 years (2003–2022) (Figure 4). These

TABLE 2 Inclusion/exclusion criteria for a two-step screening of the searched literature.

A	The article is written in English.	If yes, proceed to the next criterion.
		If not, tag as "not English," retain butexclude.
B	The study is concerned with the SLAM algorithm.	If yes, proceed to the next criterion.
		If not, tag as "not SLAM," retain but exclude.
C	The study is concerned with new algorithms, including improvements or applications to existing algorithms, proposing completely new algorithms, excluding review, and survey.	If yes, proceed to the next criterion.
		If not, tag as "not new," retain but exclude.
D	The study is concerned with a single UAV, which means that the sensors will be deployed on UAV to implement SLAM algorithms.	If yes, proceed to the next criterion.
		If not, tag as "not UAV or not single UAV," retain but exclude.
E	The study is concerned with outdoor sceneries, including park, orchard, forest, university, factory, and simulated outdoor scenes, where SLAM will be evaluated and applied.	If yes, proceed to the next criterion.
		If not, tag as "not outdoor," retain but exclude.

Abbreviations: SLAM, simultaneous localization and mapping; UAV, unmanned aerial vehicle.
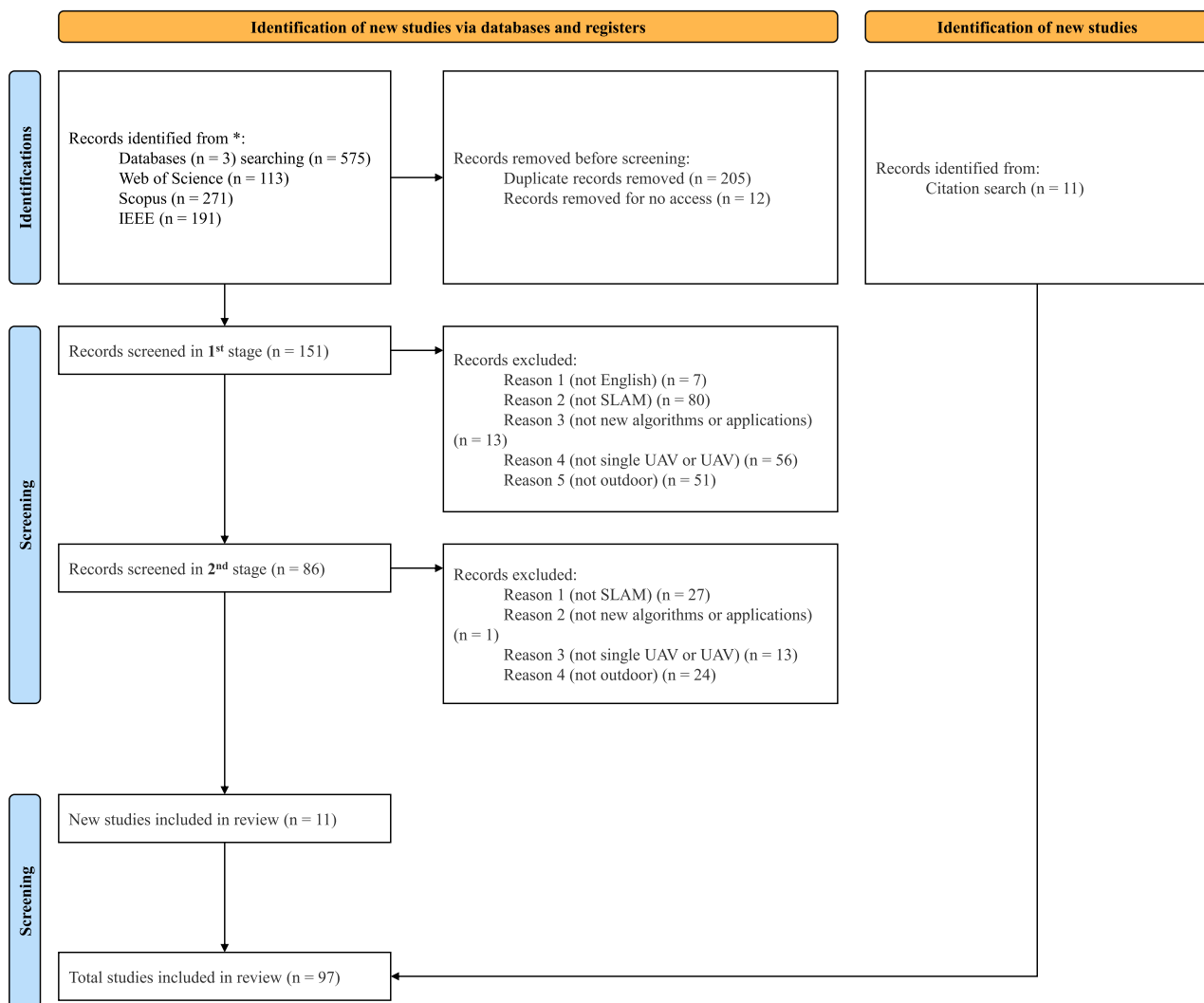
FIGURE 3 PRISMA 2020 flow diagram for updated systematic scoping reviews, which included searches of databases, registers, and other sources. PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses; SLAM, simultaneous localization and mapping; UAV, unmanned aerial vehicle. [Color figure can be viewed at wileyonlinelibrary.com]
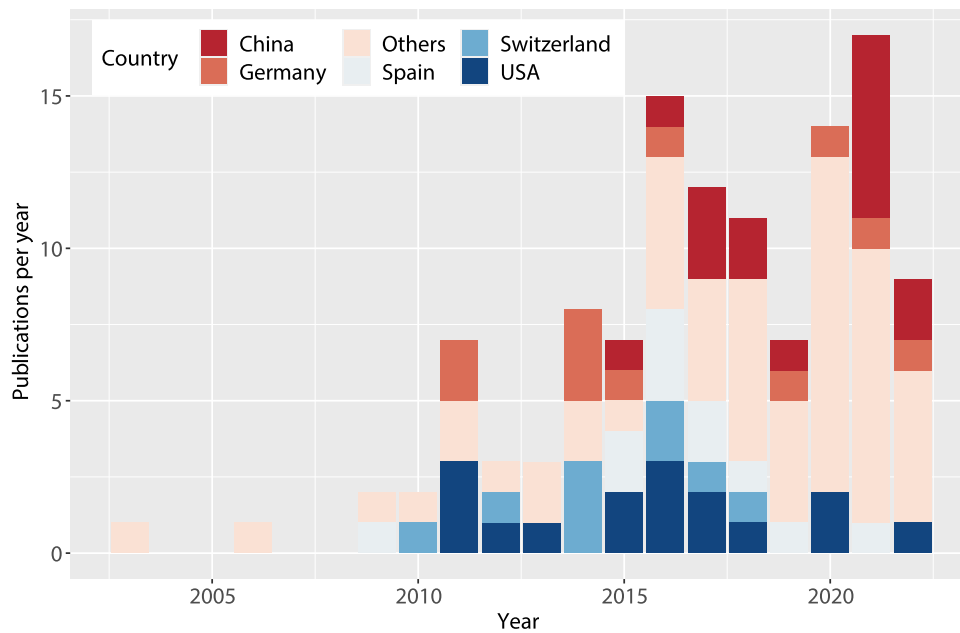
FIGURE 4 The publication trend in the field of UAV-based SLAM from 2003 to 2022. The publications per year of five main countries are shown in this figure. SLAM, simultaneous localization and mapping; UAV, unmanned aerial vehicle. [Color figure can be viewed at wileyonlinelibrary.com]

studies were conducted in 29 countries, with the largest number of studies coming from the USA ($n = 20$), China ($n = 19$), Germany ($n = 13$), Switzerland ($n = 12$), and Spain ($n = 11$).

The number of publications peaked in 2021 ($n = 13$) with fluctuations in the field, although there was a dip in 2022 due to the search period limitations (by October 8, 2022). Rapid growth after 2010 could probably be due to the improvement of UAV technologies and the increased availability of off-the-shelf UAV systems. Due to the sharply increasing interest in UAVs after 2010 (Nex et al., 2022), there was a definite upward trend in UAV-based SLAM research after 2010. Most studies were published in robotics and remote sensing-related journals and conferences (Figure A1), including IEEE International Conference on Robotics and Automation ($n = 8$), IEEE Transactions on Robotics ($n = 7$), Journal of Field Robotics ($n = 5$), Journal of Intelligent and Robotic Systems ($n = 5$), and IEEE/RSJ International Conference on Intelligent Robots and Systems ($n = 5$).

## 4.2 | Objectives of UAV-based outdoor SLAM studies

To illustrate the research objectives of UAV-based outdoor SLAM, 97 studies were divided into two main categories: algorithm-oriented studies and task-oriented studies (Figure 5). In this review, algorithm-oriented studies meant that the research objectives and results were the development and implementation of new SLAM algorithms. While the task-oriented studies meant the research objectives and results aimed to achieve specific tasks using SLAM and UAV. The task-oriented studies contained seven different tasks, including detection

task ($n = 13$), navigation task ($n = 21$), mapping task ($n = 7$), UAV control task ($n = 5$), forest assessment task ($n = 4$), and other tasks ($n = 6$). Specifically, the UAV detection tasks involve printed landmarks (Kim & Sukkarieh, 2003), smoke and fire (Campoy et al., 2009), civilian and industrial objects (Birk et al., 2011; Ellenberg et al., 2015; Jung et al., 2020; Ma et al., 2021; Medeiros et al., 2021; Pinto et al., 2021; Silva et al., 2020; Usenko et al., 2020; Yang, Li, et al., 2017), and scene changes (Nourani-Vatani & Pradalier, 2010) detection. Other tasks contained three 3D reconstruction studies (Gee et al., 2016; Kaul et al., 2016; Wendel et al., 2012), one calculation of 3D geodetic coordinates of the target study (Li et al., 2016), and two performance analysis studies of SLAM methods (Diep et al., 2022; Hansen et al., 2022). Some researchers (Hyyppä et al., 2020) evaluated the methods of data collection in forests by handheld devices, UAV flying under the canopies, and UAV flying above the canopies, concluding that features collected by UAV flying under the canopies are more accurate. As a result, the flight height of a UAV is employed as an important evaluation indicator when performing a UAV task. This topic can best be treated under five levels: low altitude (height <10 m), medium-low altitude (10–30 m), medium altitude (30–50 m), medium-high altitude (50–100 m), and high altitude (height >100 m). It was shown clearly that the proportion of flight heights that were less than 10 m accounted for the greatest number ($n = 52$) of tasks in every UAV mission.

4.3 | Application scenarios of task-oriented studies

In both task-oriented and algorithms-oriented studies, the application scenarios can be classified into aquatic, agricultural, industrial, urban,
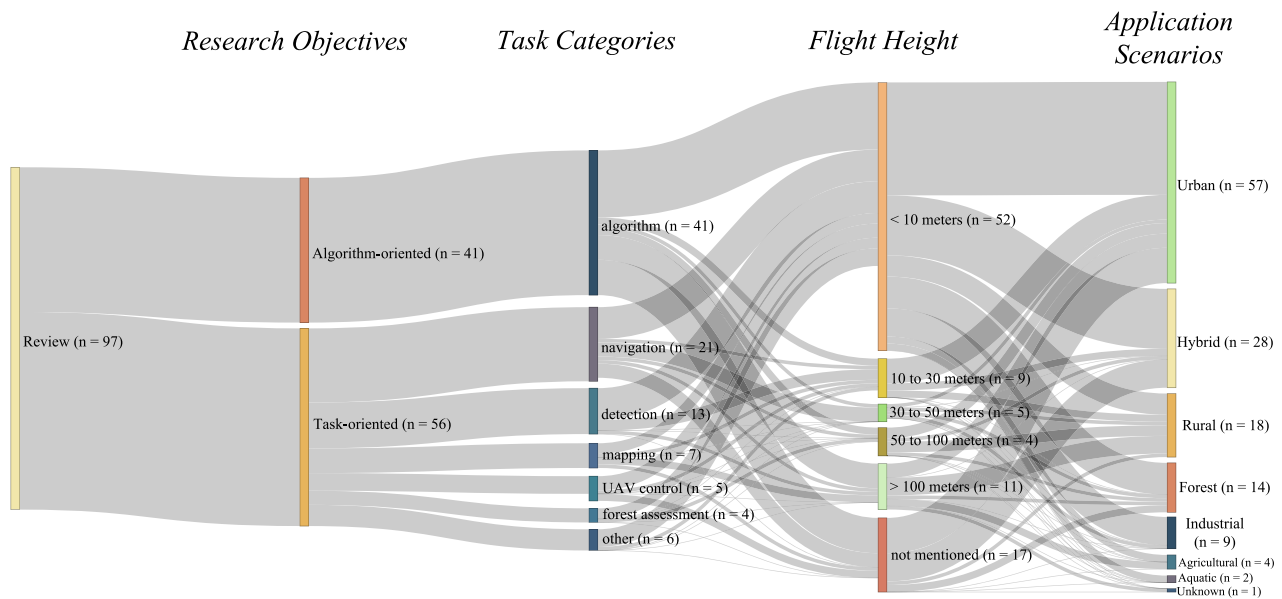
FIGURE 5 Overview of application characteristics in the eligible studies, flowing from Research Objectives, Task Categories, UAV flight height ("not mentioned" means the investigated study did not mention the flight height), and Application Scenarios ("urban" includes city, university, highway, "rural" includes village, suburb, "hybrid" includes indoor and outdoor, "forest" includes jungle, forest, rainforest, "industrial" includes wind power station, ship, factory, bridge, "agricultural" includes crop field, "aquatic" includes river, lake, ocean, "unknown" means the investigated studies did not mention the scenarios). UAV, unmanned aerial vehicle. [Color figure can be viewed at wileyonlinelibrary.com]

rural, forest, and hybrid (indoor and outdoor) environments on the basis of the experiment and application scenarios (Figure 5). Among the seven types of scenarios mentioned above, urban environments (n = 46) were the most used for evaluation, experiments, or applications. While the aquatic scenario was only used by two studies (Yang, Dani, et al., 2017; Yang et al., 2011). Although this review focused on SLAM in outdoor environments, there were many studies that evaluated SLAM in indoor environments or indoor open-source data sets first, and then conduct experiments in outdoor environments. Therefore, hybrid (indoor and outdoor) environment (n = 26) also accounted for a relatively large proportion of the application scenarios.
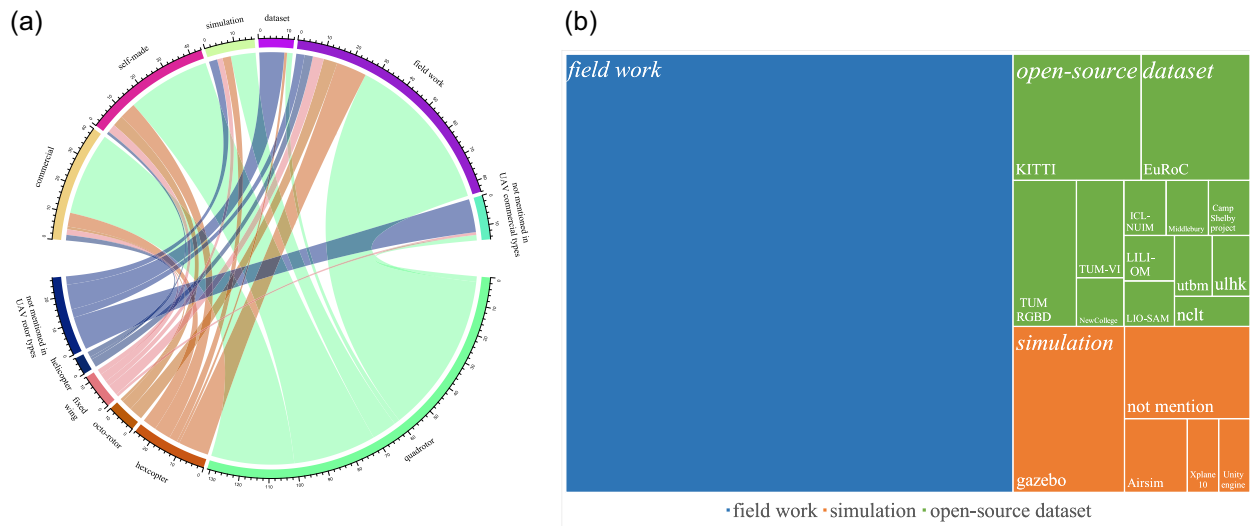
4.4 | Experimental methods in SLAM

To illustrate how the SLAM algorithms have been typically evaluated, the experimental methods are categorized in three ways: public data sets, simulation platforms experiments, and field experiments (Figure 6). Although some studies implemented both public data sets, simulation platforms, or field experiments, overall, field experiments (n = 66) for evaluating SLAM performance accounted for the vast majority in the selected papers compared with the public data set and simulation platform experiment methods. Generally, field experiments carry out SLAM with UAVs and some necessary equipment such as ground computing platform and sensors to targeted scenarios mentioned in Figure 6. This method is able to show the real results and performance of SLAM in the target scenarios compared with evaluating SLAM by the public data set and

simulation platform. However, the use of public data sets and simulation platforms is very efficient in developing new SLAM algorithms and is convenient to carry out benchmark to other SLAM algorithms.

There are some famous public data sets for SLAM evaluation shown in Table 3. Among them, evaluating SLAM with the KITTI data set (Geiger et al., 2012) accounts for the largest number in the selected papers (n = 7). And the EuRoC data sets (Burri et al., 2016) were used to evaluate SLAM in the selected papers in large numbers (n = 6) as well. In addition, only one study using the ICL-NUIM data set for evaluation (Forster et al., 2016) and one study using the NewCollege data set for evaluation (Mur-Artal et al., 2015). While LILI-OM (Li, Li, et al., 2021), LIO-SAM (Shan et al., 2020), utbm (Yan et al., 2020), ulhk (Wen, Zhou, et al., 2020), and nclt (Carlevaris-Bianco et al., 2016) were only used in one study (Xu et al., 2022).

To evaluate SLAM in a specific scenario, simulation platforms can also be used to perform SLAM experiments. The investigation of simulation platforms has shown that Gazebo (n = 8) (Koenig & Howard, 2004), which is an open-source software library for robotics developers was the most popular simulation platform in the eligible papers. It provides an essential robotic toolbox to simulate robots accurately and efficiently in complex indoor and outdoor environments. However, other simulation platforms accounted for a very small proportion. Airsim is a simulator developed by Microsoft for drones and cars based on Unreal Engine, which was used by two studies (Nguyen et al., 2021; Xie et al., 2021) in eligible papers. In addition, there was one study using Unity engine (Liu et al., 2022) and one study using Xplane 10 (Yang et al., 2019).

FIGURE 6 Experiments methods categories, the chord chart (a) indicated the relationship and proportion among experimental methods (fieldwork, simulation, and open-source data set), type of UAVs (commercial and self-developed) and the number of rotors of UAVs (quadrotor, hexacopter, octorotor, fixed wing, and helicopter), the tree chart (b) illustrated the detailed simulation platforms and open-source data sets in experimental methods as well as their proportion used in the literature by the size of the squares. UAV, unmanned aerial vehicle. [Color figure can be viewed at wileyonlinelibrary.com]

TABLE 3 Public data sets for UAV-based SLAM used in the selected papers, including the name of data sets, publication year, number of the data sets used in the selected papers, detailed sensor categories, original reference, and the link to the data sets.

Name	Year	Number	Sensor categories			Reference	Link
			Camera	IMU	LiDAR		
KITTI	2012	7	Color stereo	N/A	Velodyne-64	1	https://www.cvlibs.net/datasets/kitti/
			Gray stereo				
EuRoC	2016	6	Gary stereo	N/A	N/A	2	https://projects.asl.ethz.ch/datasets/doku.php?id=kmavvisualinertialdatasets
TUM-RGBD	2012	4	RGB-D	N/A	N/A	3	https://cvg.cit.tum.de/data/datasets/rgbd-dataset
TUM-VI	2018	2	Gray stereo	True	N/A	4	https://cvg.cit.tum.de/data/datasets/visual-inertial-dataset
ICL-NUIM	2014	1	RGB-D	N/A	N/A	5	https://www.doc.ic.ac.uk/~ahanda/VaFRIC/iclnuim.html
NewCollege	2009	1	Gray stereo	True	Sick-2D	6	https://ori-drs.github.io/newer-college-dataset/
LILI-OM	2021	1	N/A	True	Velodyne-64	7	https://isas-server.iar.kit.edu/lidardataset/
LIO-SAM	2020	1	N/A	True	VLP-16	8	https://github.com/TixiaoShan/LIO-SAM
utbm	2020	1	N/A	True	Velodyne-32	9	https://epan-utbm.github.io/utbm_robocar_dataset/
ulhk	2020	1	N/A	True	Velodyne-32	10	https://github.com/weisongwen/UrbanLoco
nclt	2016	1	Color omni	True	Velodyne-32	11	http://robots.engin.umich.edu/nclt/
					Hokuyo-2D		

Abbreviations: IMU, inertial measurement unit; LiDAR, Light Detection and Ranging; SLAM, simultaneous localization and mapping; UAV, unmanned aerial vehicle.

4.5 | Hardware for UAV-based outdoor SLAM

4.5.1 | UAV platforms

The UAV platforms adopted in the selected studies can be categorized based on commercial or self-made into quadrotor, hexacopter, octorotor, helicopter, and fixed wing. The ratio of commercial UAVs to self-made UAVs in the eligible papers was almost 1:1 (Figure 6). It was shown that quadrotor UAVs accounted for a very high proportion in both commercial ($n = 30$) and self-made ($n = 30$). Fixed-wing UAVs accounted for a small percentage of UAV-based SLAM studies compared with multirotor UAVs, with two
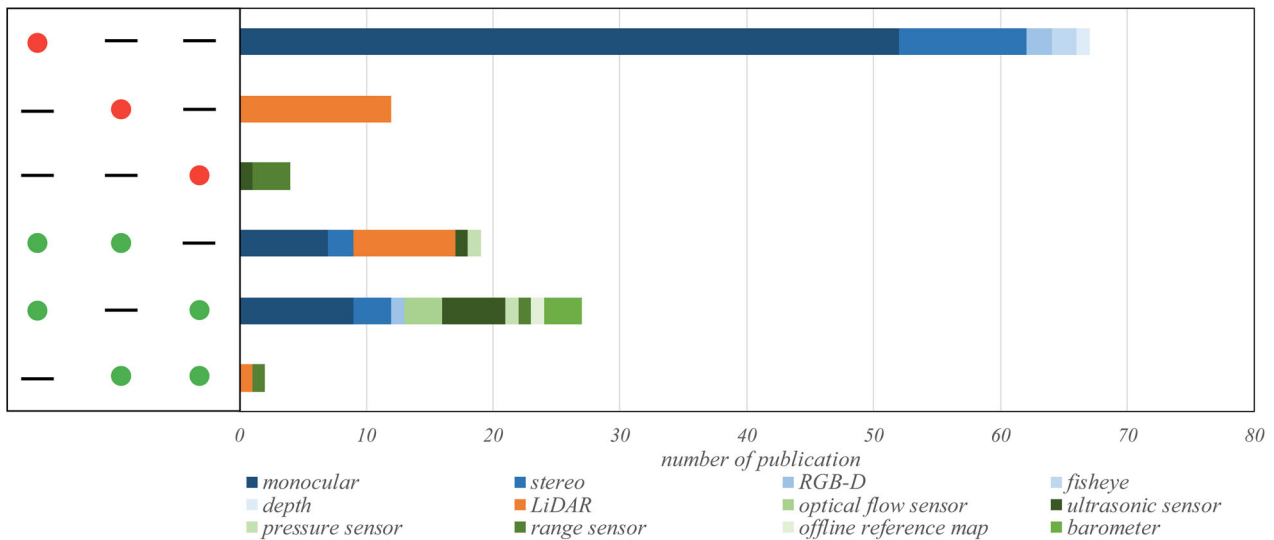
FIGURE 7 The distribution of sensors in the investigated studies. The sensors were divided into camera-based, LiDAR-based and others. The red dots on the left columns indicate the studies using a single type of sensor, and green dots indicate the studies using two types of sensors. LiDAR, Light Detection and Ranging. [Color figure can be viewed at wileyonlinelibrary.com]

studies (Bryson & Sukkarieh, 2006; Kim & Sukkarieh, 2003) using commercial fixed-wing UAVs and three studies (Hinzmann et al., 2016; Suzuki, Amano, & Hashizume, 2011; Suzuki, Amano, Hashizume, & Suzuki, 2011) using self-made fixed-wing UAVs.

4.5.2 | Sensor categories

The sensors used in the eligible studies can be classified into three categories based on their sensing principles: camera, LiDAR, and others (Figure 7). Monocular cameras ($n = 68$) were the most popular sensors used in SLAM. Both LiDAR ($n = 21$) and stereo camera ($n = 15$) accounted for 11.5%. Depth camera (Li-Chee-Ming & Armenakis, 2018) and offline reference map (Shao et al., 2021) were only used in one study each.

Monocular cameras accounted for a significant portion of UAV-based SLAM research using only cameras. In SLAM research involving sensor fusion, monocular cameras, and LiDAR account for the largest share. No studies for outdoor applications have been combining cameras, LiDAR, and other sensors yet.

Table 4 provides an overview of the major sensor types with sensing principle and basic parameters of the sensors.

4.6 | SLAM algorithms

SLAM could be classified according to the name of the open-source architecture into 16 categories in the selected studies (Figure 8). Among the existing SLAM methods, studies using EKF-SLAM ($n = 24$) and ORB-SLAM ($n = 23$) represented the largest proportion of the selected papers. Depending on the type of sensors, the highest

number of LiDAR-based SLAM studies used LOAM ($n = 5$) in the investigated studies, while ORB-SLAM ($n = 18$) accounted for the largest proportion in visual-based SLAM. Among the remaining SLAM algorithms in the selected studies, most were visual SLAM, such as SVO ($n = 5$), LSD-SLAM ($n = 5$), PTAM ($n = 10$), VINS-Mono ($n = 3$), VINS-Fusion ($n = 1$), MSCKF ($n = 2$), and FAST-LIO ($n = 1$) was based on LiDAR SLAM.

In addition to using efficient optimization algorithms to reduce processing times and improve accuracy, having a powerful processing unit such as a high-performance CPU or GPU is also crucial for the real-time performance of SLAM. Therefore, to better illustrate the computational performance of SLAM, the computational platform was classified in terms of the chip architecture and computation into CPU and GPU (Table 5). Approximately three-quarters of papers (Figure 8) in the eligible studies used only the CPU ($n = 76$) in their SLAM implementations. Only a very small proportion of the studies involved GPU ($n = 8$) in UAV-based SLAM.

In the field of SLAM, the availability of open-source code is significant for community development. Yet, only one in five of all investigated studies accounted for open-source studies. Both task-oriented and algorithm-oriented, open-source studies played a significant role in the eligible studies in this review. Although the proportion of open-source studies was relatively small to the total research, these open-source studies have become milestones in SLAM.

5 | DISCUSSION

There have been reputable studies pointing out some issues in current SLAM research (Cadena et al., 2016). For example, there are problems in robustness and scalability under long-term autonomy, such as the failure

TABLE 4 Overview of sensor in SLAM, including current main sensor types with their sensing principle, and basic parameters, including sensing distance, data types, price, power consumption, and the weight of the sensor from small to large.

Sensor type	Principle	Sensing distance (m)	Output data types	Cost ($)	Power consumption (W)	Weight
Monocular camera	Monocular cameras capture images or videos with a single lens and convert the light into digital data that can be processed by computers. While it lacks depth perception, it may be integrated with other sensors or algorithms to estimate depth and provide extra environmental information.	NA	RGB images	2–100k	0.01–10	Small
Stereo/binocular camera	Stereo cameras capture images or videos from slightly varied perspectives, allowing it to discern depth and create 3D images. The principle of the sensor is that they mimic the human eyes perceive depth by using the differences in the two images to calculate the distance to objects in the scene.	0.2–30	RGB images and depth information	30–1500	2–15	Medium
RGB-D camera	RGB-D cameras captures both color and depth information using a combination of traditional RGB cameras and infrared or time-of-flight (ToF) sensor. It allows for more precise and comprehensive 3D modeling and object detection, which are widely utilized in robotics, virtual reality, and computer vision applications.	0.2–5	RGB images and depth information	35–700	2–5	Medium
Fisheye camera	Fisheye cameras capture hemispherical or panoramic images of the world using a wide-angle lens. The fisheye cameras use lens with a very short focal length and a highly curved front element to achieve a wider field of view than traditional camera lens.	NA	Distortional RGB image	5–4000	0.01–10	Medium
LiDAR	LiDAR is mainly composed of laser transmitter, receiver, scanner, and signal processing module. The current mainstream LiDAR ranging principle is to calculate ToF of the transmitted pulse signal and the reflected signal received by the receiver.	0.02–50	Point cloud data	5k–100k	50–300	Large
Acoustic Sensor	The acoustic sensor is similar to LiDAR and consists of a transmitter, receiver, and signal processing module. The transmitter will continuously emit an acoustic signal, and the receiver will receive a return signal when the acoustic wave encounters an obstacle. The distance between the sensor and the obstacle can be calculated by the time of the sound wave round trip.	2–5	Sound wave signal	10–500	0.01–1	Small

Note: The parameters of the sensors were based on Amazon and JingDong (JD) products search and description by the end of 2022.

Abbreviations: 3D, three dimensional; LiDAR, Light Detection and Ranging; SLAM, simultaneous localization and mapping.
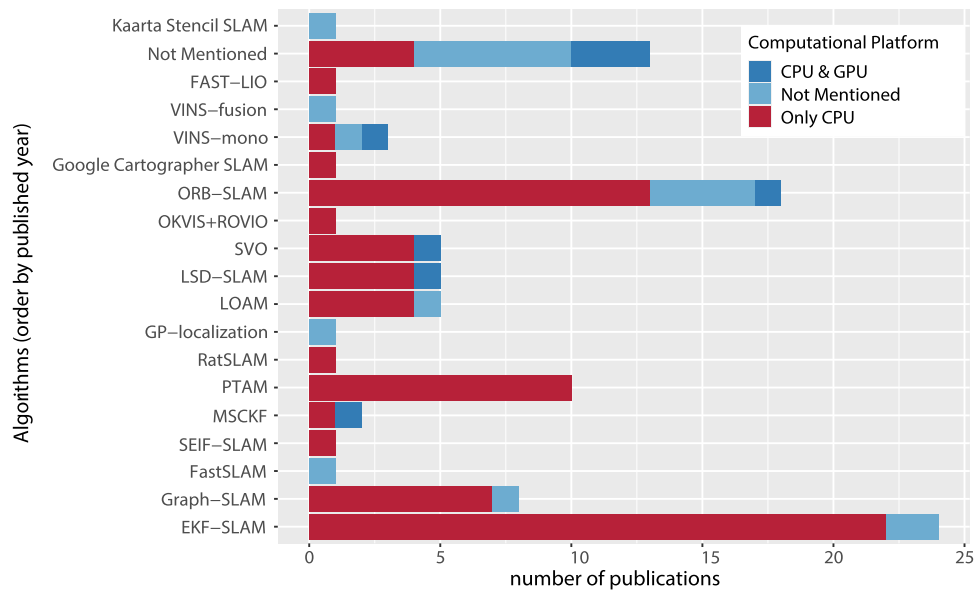
FIGURE 8 The distribution of the SLAM algorithms using UAVs. The y-axis is algorithms ordered by the published year (earlier year at the lower end of the y-axis), excepted for "Not Mention" and "Kaarta Stencil SLAM." For the computational platform, "Only CPU" means the studies only use CPU to run, test, and evaluate SLAM algorithms, "CPU and GPU" means the studies use both CPU and GPU to run, test, and evaluate SLAM algorithms, "Not mentioned" means the studies did not specify the computational platforms used. CPU, central processing unit; EKF, extended Kalman filter; FAST-LIO, fast direct lidar-inertial odometry; GPU, graphics processing unit; LOAM, LiDAR odometry and mapping; LSD, large-scale direct; MSCKF, multistate constraint Kalman filter; PTAM, parallel tracking and mapping; SLAM, simultaneous localization and mapping; SVO, semidirect visual odometry; UAV, unmanned aerial vehicle; VINS-mono, monocular visual-inertial system. [Color figure can be viewed at wileyonlinelibrary.com]

TABLE 5 CPU and GPU model examples used in UAV-based SLAM.

Computing Component	CPU-Based		GPU-Based	
	onboard	off-board	onboard	off-board
Model	Intel Atom	Intel Core i series	NVIDIA Jeston TX1	NVIDIA GTX 1080
Physical Figure				
Computational Performance	*	**	*	***
Power (W)	*	**	*	***
Price	*	**	**	***
Example	(Harmat et al., 2015)	(Silva et al., 2020)	(Forster et al., 2016)	(Yang et al., 2019)

Note: The number of symbols "*" in this table represents the level of computational performance, power consumption, and price cost.
Abbreviations: CPU, central processing unit; GPU, graphics processing unit; SLAM, simultaneous localization and mapping; UAV, unmanned aerial vehicle.

and recovery of SLAM and metric relocalization. Additionally, metric maps, semantic map models, SLAM combined with deep learning, and novel and unconventional sensors (e.g., range camera, light-field camera, and event camera) all have some limitations to be addressed in SLAM research. And when it comes to UAV-based SLAM, these issues will remain and may give rise to new challenges.

To enable the comparison and discussion among the selected studies with the same scenarios, a data synthesis for UAV-based SLAM was prepared (Figure 9). This data synthesis provides a blueprint and guidance to researchers new to the field of UAV-based SLAM. New researchers can refer to the setups to prepare their studies and experiments according to the synthesis. The categories of

Application Scenarios	Urban (n = 46)	Hybrid (n = 16)	Rural (n = 17)	Forest (n = 14)	Industrial (n = 9)	Agricultural (n = 4)	Aquatic (n = 2)
Requirements	Structured Complex Large-scale	High contrast lighting Limited field of view Repetitive patterns	Lack of features	Dense vegetation Limited GPS signal Challenging lighting Complex	Repetitive patterns Cluttered environment	Dense vegetation Limited GPS signal Dynamic objects Challenging lighting Large-scale	Lack of features Limited GPS signal
UAV Data Acquisition	Monocular CameraStereo Camera LiDAR Depth Camera Range Sensor Ultrasonic Sensor Barometer Offline Reference Map	Monocular CameraStereo Camera RGB-D Camera LiDARUltrasonic Sensor Barometer Optical Flow Sensor	Monocular Camera LiDAR Pressure Sensor	LiDAR Stereo Camera Monocular Camera	LiDAR Monocular Camera Stereo CameraRGB-D Camera Pressure Sensor Ultrasonic Sensor Range Sensor	Monocular Camera LiDAR	Monocular Camera Ultrasonic Sensor Barameter
Experimental Method	Field Work Simulation Open-Source Dataset	Field Work Open-Source Dataset Simulation	Field Work Simulation	Field Work Simulation Open-Source Dataset	Field Work Simulation	Field Work Simulation	Field Work
Visual SLAM	EKF-SLAM ORB-SLAM Graph SLAM SVO PTAM LSD-SLAM GP-localization VINS-Mono MSCKF RatSLAM	EKF-SLAM ORB-SLAM Graph SLAM SVO PTAM LSD-SLAM VINS-mono	EKF-SLAM ORB-SLAM SVO Graph SLAM	EKF-SLAM Graph SLAM ORB-SLAM LSD-SLAM SLOAM+S-MSCKF	EKF-SLAM ORB-SLAM Graph SLAM VINS-Mono VINS-Fusion	ORB-SLAM OKVIS+ROVIO	EKF-SLAM
LiDAR SLAM	EKF-SLAM LOAM Cartographer SLAM FAST-LIO SEIF-SLAM	EKF-SLAM	EKF-SLAM	LOAM VINS-Fusion Kaarta Stencil SLAM	LOAM VINS-Fusion	LOAM	No Studies
Flight Height	<10 meters 10 - 30 meters 30 - 50 meters >100 meters	<10 meters 10 - 30 meters 30 - 50 meters	<10 meters >100 meters 10 - 30 meters 50 - 100 meters	<10 meters 30 - 50 meters 50 - 100 meters >100 meters	<10 meters 30 - 50 meters	<10 meters > 100 meters 30 - 50 meters 50 - 100 meters	<10 meters
Results & Objectives	algorithms navigation mapping detection UAV control 3D reconstruction	algorithms navigation UAV control 3D reconstruction	algorithms detection navigation mapping UAV control	navigation forest assessment mapping detection algorithms	detection algorithms	algorithms detection	algorithms navigation

FIGURE 9 Synthesis for UAV-based SLAM studies, based on the selected 97 papers, the sequences of the terms in each block emphasize the number of studies, from top to bottom decreasing. 3D, three dimensional; EKF, extended Kalman filter; FAST-LIO, fast direct lidar-inertial odometry; GPS, global positioning system; LiDAR, Light Detection and Ranging; LOAM, LiDAR odometry and mapping; LSD, large-scale direct; MSCKF, multistate constraint Kalman filter; PTAM, parallel tracking and mapping; SLAM, simultaneous localization and mapping; SVO, semidirect visual odometry; UAV, unmanned aerial vehicle; VINS-mono, monocular visual-inertial system. [Color figure can be viewed at wileyonlinelibrary.com]

scenarios in the synthesis were based on the results of this review (Figure 5).

To better illustrate the challenges of UAV-based SLAM in different scenarios, the following sections of this discussion focused on the issues, including sensors, algorithms, and tasks, and application scenarios and factors impacting UAV-based SLAM explored in these studies.

5.1 | Sensor

This review highlights the significance and diversity of sensors utilized in UAV-based SLAM (Figure 7). Studies using LiDAR accounted for 20% of the total investigated studies. LiDAR accounted for a large proportion of all application scenarios except for the aquatic environments. This may be likely due to its extreme accuracy, simple error model, and robustness in various illumination conditions (Huang, 2021), especially for forest and industrial applications that may involve very high accuracy measurements tasks such as forest growth estimation and industrial scene modeling. Although LiDAR was currently accurate enough and can be applied to most tasks and application scenarios, it had limitations (Table 4) due to the high cost, large size, and single type of acquired information from LiDAR (Huang, 2021).

In comparison, visual sensors such as monocular camera (70.7%), stereo camera (15.5%), and RGB-D camera (3.1%) were widely used in all kinds of scenarios for UAV-based outdoor SLAM (Figure 7). RGB-D camera can acquire both depth information and visual information directly, which was very suitable for SLAM research.

However, some studies (Jin et al., 2019) pointed out that since the acquisition of depth information is based on structured light or ToF sensors, the measurement accuracy can be easily affected by environmental illumination changes, which meant that RGB-D camera was not reliable for outdoor environment. The monocular camera was more stable and robust than RGB-D cameras in complex, light-changing environment. Moreover, it was low cost, and had simple structure, small size and can be easily mounted on all kinds of robot platforms. Although monocular cameras cannot directly acquire depth information to determine the absolute scale from images in SLAM, many studies illustrated that the absolute scale issue can be solved from IMU data (Nützi et al., 2011) or estimated by detected objects (Sucar & Hayet, 2018) and self-induced oscillations in hover control of SLAM (Lee & de Croon, 2018). In addition, stereo camera was inspired from human eye, which can use the disparity of two camera images in the same scene to calculate depth information and did not suffer from the problem of scale-drift in monocular camera 4. However, this review found that only 20% of the investigated studies used stereo camera in UAV-based outdoor SLAM. This may be caused by the complexity and difficulty of initial configuration, the calibration of stereo camera and its large size compared with monocular camera for all types of robot platforms, especially for UAV.

Overall, UAV platforms equipped with visual sensors, especially monocular cameras accounted for the majority of the outdoor SLAM studies investigated. And compared with other robotic platforms, this may be due to the characteristics of UAVs, such as limited payload, specific flight motion modes, and limited power.

5.2 | Algorithm

5.2.1 | Conventional SLAM

Regardless of visual sensors or LiDAR sensors, almost all the investigated studies at the algorithmic level were based on geometric features (such as points, lines, and planes) for outdoor UAV-based SLAM applications (Figure 8). Among them, the ORB-SLAM series (18. 6%) were the most popular visual SLAM algorithms in UAV-based outdoor SLAM studies. ORB-SLAM (Mur-Artal et al., 2015) was a SLAM algorithm based on the PTAM framework (Klein & Murray, 2007), which used ORB features (Rublee et al., 2011) with better angular invariance instead of a fast corner detector and incorporates closure loop detection to eliminate drift error accumulation. Mur-Artal and Tardos also proposed ORB-SLAM2 (Mur-Artal & Tardós, 2017) and ORB-SLAM3 (Campos et al., 2021), which supported for stereo cameras, RGB-D cameras, and fisheye cameras. As for LiDAR SLAM, Cartographer (Hess et al., 2016) was the current mainstream algorithm of 2D LiDAR SLAM which was a graph-based optimization algorithm for indoor map construction. However, 2D LiDAR has limitations in handling complex outdoor environments with factors, such as smoke, dust, and rain. The introduction of LOAM (Ji & Singh, 2014) addressed these challenges with 3D LiDAR.

LOAM proposed a novel feature extraction (edge and planar features) and developed a motion compensation for SLAM using timestamps but did not achieve back-end optimization to improve mapping accuracy. To avoid trajectory drift and further improve the accuracy of pose estimation and map construction, LeGO-LOAM (Shan & Englot, 2018) utilized Georgia tech smoothing and mapping (gtsam) as back-end optimization. Although these traditional SLAM algorithms performed well on current tasks, SLAM still lacked intelligent perception capability (Figure 2), especially for some specific scenarios and tasks. For example, to solve the problem that the monocular camera cannot acquire depth information which leads to scale uncertainty (Table 4), a monocular depth estimation solution (Kuznietsov et al., 2017) was proposed by using a semisupervised deep learning model. In addition, in some tasks with special requirements (Figure 5) such as detection, recognition, classification, conventional geometric feature-based SLAM cannot handle these tasks. Thus, semantic information can be combined with SLAM by deep learning (Li et al., 2017; McCormac et al., 2017). Although there were some works reporting such solutions, they have been mostly used with ground vehicles or other handheld equipment (Lategahn et al., 2011; Quan et al., 2021; Roussillon et al., 2011). When addressing this problem with UAVs, none or few studies have been found due to the complexities and features of UAV such as the limited payload capacity, flight time, and energy consumption, as stated by some studies (Abeywickrama et al., 2018; Cabreira et al., 2019).

Dealing with high dynamic and unstructured surroundings is one of the most difficult issues in UAV-based SLAM. For example, UAVs are typically operated in open outdoor areas where illumination conditions may vary rapidly and wind can severely impair the UAVs' stability and control (Floreano & Wood, 2015). Thus, the robustness of SLAM is more critical for UAV-based SLAM than ground-based SLAM (Ravankar et al., 2018). However there are more factors affecting UAV-based SLAM compared with other platforms (e.g., ground vehicle, handheld equipment, and backpack device). For example, the relatively short flight time of UAVs limited by the batteries might restrict the quantity of data acquired by the sensors making it difficult to generate reliable maps of the area (Hardin & Jensen, 2011). As a result relocalization capability for UAVs from the previous SLAM results can also be a challenge, which was also reported in a previous study (Zaffar et al., 2018).

In comparison to UAV-based SLAM which can involve working in various environments (Figure 9) with complex features (Opromolla et al., 2016; Semsch et al., 2009) ground vehicles are often used in more controlled settings such as indoor or outdoor locations with well-defined paths and prominent landmarks (Bostelman et al., 2015; Kumar et al., 2021; Martínez-Barberá & Herrero-Pérez, 2010). This is because ground vehicles typically follow predetermined routes and rely on pre-existing maps or landmarks for navigation, which can limit their flexibility in unstructured environments (Lynch et al., 2018; Shahzadi et al., 2021). Although both UAVs and ground vehicles may encounter obstacles in their navigation leading to mission failure, the excellent maneuverability (Dvorak et al., 2015) and cross-terrain

capabilities (Sujiwo et al., 2016) of UAVs make it easier to deal with obstacles and uneven terrain. However these characteristics of UAVs also add many challenges at the algorithmic level of SLAM. For example, when performing highly maneuverable UAV tasks (e.g., rescue, military mission, and rapid navigation), the geometric feature-based SLAM algorithm mentioned above (e.g., ORB-SLAM) may fail to initialize or fail to extract features due to the interference with sensors when tracking features with excessive speed of the UAV (Esfahlani, 2019). Therefore SLAM algorithms should be optimized for different tasks in different application scenarios to increase their robustness.

5.2.2 | Deep learning SLAM

Compared with the conventional geometric feature-based SLAM, Deep Learning can extract feature from low level in shallow layer to high level in deep layer can be more robust for visual SLAM in the challenging environments (Taisho & Kanji, 2016). Except for improving feature extraction, deep learning such as Convolutional Neural Networks (CNNs) can also provide semantic information for both visual and LiDAR SLAM. Monocular depth estimation with CNN model solves the issue that SLAM with monocular camera cannot obtain depth information from images directly (Steenbeek & Nex, 2022). There are different ways to build the monocular depth estimation model, such as Unsupervised Learning (Zhan et al., 2018), Semisupervised Learning (Kuznietsov et al., 2017), and Self-supervised Learning (Godard et al., 2019). Moreover, some computer vision models can extract semantic information from environments for understanding surroundings, improving pose estimation accuracy, and supporting loop-closure detection (Qin et al., 2021; Tateno et al., 2017). There are also some deep learning models for LiDAR SLAM. For instance, Chen et al. introduced a novel semantic mapping method to detect dynamic object and improve pose estimation accuracy with LiDAR solely (Chen et al., 2019).

5.3 | Tasks and application scenarios

As mentioned in this review, most of the tasks in the analyzed studies were focused on navigation and detection, and most of the application scenarios were also focused on urban, rural, and forest (Figure 5). Our analysis revealed some interesting patterns in the use of UAV-based SLAM in different scenarios. Specifically, we found that algorithm-oriented research is often carried out in urban and rural scenarios, such as university campuses or parks. These scenarios are usually not chosen for specific applications in these scenarios for SLAM algorithm implementation but only for algorithm evaluation in outdoor open spaces. For other scenarios such as forest, agriculture, industry, aquatic, the selection of these scenarios aimed to address some challenges found in the specific scenarios (Cui et al., 2014; Schultz et al., 2016; Yang, Dani, et al., 2017). However, the number of selected studies worked in forest, agriculture, industry, aquatic is
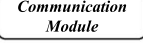
much less than in urban and rural. This suggests that more work is needed to explore the potential applications of UAV-based SLAM in real-world and task-oriented settings in these environments, such as mapping and monitoring of infrastructure, precision agriculture, and disaster response scenarios. And since algorithm-oriented tasks account for 40% of all tasks (Figure 5), most algorithm-oriented tasks are evaluated in public data sets rather than in field experiments. The benefit of conducting SLAM evaluations on public data sets was a uniform SLAM algorithms benchmark (Bujanca et al., 2019) can be established to illustrate the differences in algorithm performance. But excluding algorithm-oriented tasks, most other studies tended to conduct field experiments (60%). A possible explanation for this could be that these experimented SLAM algorithms may perform well in only a few specific scenarios and do not achieve comparable performance in other scenarios. Another possible explanation is that UAV has high experimental requirements, such as pilot specialists to operate the UAVs, ground equipment to assist the SLAM in collecting and processing data, and safety issues of the UAV. And some studies (Kim & Eustice, 2013; Porav et al., 2019) have also conducted specific field experiments for different tasks (inspection, detection), and scenarios (rainy weather, underwater). Therefore, the possible explanation that field experiments of specific scenarios will be determinant in task-oriented SLAM research is further supported.
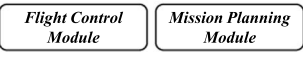
In addition, this review found significant differences in the flight height of UAVs in different tasks or application scenarios (Figure 5). Another important finding was that the flight height of UAVs under 10 m accounted for almost 50% of both task-oriented and algorithm-oriented studies. Some studies have found that low-flying UAVs perform well in tasks, such as identifying ruts and potholes (Saad & Tahar, 2019), precision weed management (Huang et al., 2018), and crop characterization (Uto et al., 2013). It seems possible that these results are due to the fact that UAVs with low-altitude flight height are able to acquire more precise and detailed information. These results matched those observed in earlier studies. However, some studies have also pointed out that high altitudes are more beneficial in other UAV tasks, such as digital elevation model mapping (Ageed & Abulrahman, 2020), georeferenced mosaics of wheat (Gómez-Candón et al., 2014), and single tree height estimation in broadleaf forests (Sadeghi & Sohrabi, 2019). Thus, a low UAV flight altitude may be more beneficial in some tasks with high accuracy requirements while a high UAV flight altitude may be more beneficial in some tasks with high efficiency requirements and a large-scale range.

5.4 | Exploring experimental factors affecting UAV-based SLAM

Experimental Factors that influence the performance of SLAM in UAV applications can be divided into hardware, software, and external factors based on the components of UAV-based SLAM (Figure 10).

In terms of hardware, there are several experimental factors to consider for UAV-based SLAM. The sensor is one of the most

WANG ET AL.

WILEY | 17

15564967, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/rob.22325 by Wageningen University And Research Facilitair Bedrijf, Wiley Online Library on [22/05/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Experimental Factors Impacting UAV-based SLAM		
Hardware Factors	*Software Factors*	*External Factors*
Sensor (type, quality, placement)	Sensor calibration	Flight environment
Processing power	UAV control system	UAV flight setup (speed, height)
UAV types	Communication Module / Flight Control Module / Mission Planning Module	
Example (Huh et al., 2013) (Torres-González et al., 2015) (Harmat et al., 2015)	Example (Bryson and Sukkarieh, 2006) (Campoy et al., 2009) (Zhang et al., 2015)	Example (Nourani-Vatani and Pradalier, 2010) (Cui et al., 2014) (Yang et al., 2017)

FIGURE 10 Experimental factors impacting on UAV-based SLAM, the first column (blue) shows the hardware factors, including sensor, processing power, and UAV types, the second column (orange) shows software factors, including the sensor calibration, SLAM control system, and the third column (green) shows the external factor which includes flight environment (e.g., wind and weathers), and UAV flight setup (e.g., speed and flight height). Some examples mentioned these factors in the selected literature. LiDAR, Light Detection and Ranging; SLAM, simultaneous localization and mapping; UAV, unmanned aerial vehicle. [Color figure can be viewed at wileyonlinelibrary.com]

important hardware components that can affect the performance of SLAM. All investigated studies clearly mentioned the sensors (Figure 7) in their experiment setup. SLAM performance relies on the data from sensors to map the environment and estimate the UAV's location and orientation. The quality of the raw data dictates the upper limit of SLAM performance (Aguilar-Moreno & Graña, 2022; Zaffar et al., 2018). If the raw data is noisy, inaccurate, or missing, the result of SLAM including the map and pose estimations will be correspondingly less accurate (Moosmann & Stiller, 2011). The computational platform (Figure 8) can also be a relatively crucial factor for SLAM in UAV applications. 81% of studies stated the detailed experiment setup of computational platform including the type of CPU or GPU (such as Intel Atom, Intel i7-4900MQ, Nvidia Jeston TX1, Nvidia GeForce GTX-650, etc.) which can satisfy the requirements of real-time processing capacity for massive sensor data. Nearly 1% of the studies involved both GPU and CPU to process SLAM due to some SLAM algorithms combining deep learning and computer vision models. High-performance computational platforms can enhance the processing efficiency of SLAM (Tuna et al., 2012), especially for some deep learning-based SLAM (McCormac et al., 2017; Rosinol et al., 2020). In addition, different types of UAVs can affect the performance of SLAM due to their different flight styles.

In terms of software, sensor calibration and the UAV control system could be considered in the SLAM experiment setup. Only one investigated study focused on sensor calibration, which developed a self-calibration system for visual SLAM (Heng et al., 2015). Sensor calibration is a preprocessing procedure for conducting SLAM, which is very crucial for obtaining accurate data from sensors (Sturm et al., 2012; Trejos et al., 2022). However, sensor calibration is ignored in some SLAM studies for UAV applications. This is mainly because most of the UAVs used in the studies were commercial products (Figure 6a) that did not require secondary development and whose sensors were calibrated at the factory in advance. Another software factor can be the control system of the UAV, which contains communication modules, flight control modules, mission planning modules, and so forth (Figure 10). For example, software that controls the UAV's flight path and the timing of sensor measurements can affect consistency and quality of sensor data (He et al., 2006) which can in turn affect SLAM performance.

For external experimental factors, flight environment conditions and UAV flight setup can impact SLAM in UAV applications. All investigated studies illustrated scenarios during flying UAV with SLAM (Figure 5), but none or few showed detailed information about environment conditions. The flight environment, such as terrain, wind, weather, and illumination conditions can impact sensor data quality and consistency (Mohammed et al., 2020; Vargas et al., 2021). For example, in poor light environments, camera sensors may produce low-quality images leading to inaccuracies in feature identification and matching in erroneous pose estimations and poor

map quality (Laible et al., 2012). Furthermore, flying UAVs in windy or turbulent environments might cause them to move unpredictably resulting in errors in estimating UAV's position and orientation (Kothari et al., 2014). In addition, terrain can also have an impact on the performance of the SLAM algorithm. For example, in environments with numerous identical or repetitive features such as forests or agricultural fields, the SLAM algorithm may struggle to distinguish between different features leading to errors in feature matching and pose estimation (Zhu et al., 2021). Flight setup including speed and flight height of UAV can also have an impact on SLAM (Figure 5). Motion of UAV especially speed has a considerable influence on SLAM (Henein et al., 2020; Zhou et al., 2021). If UAV is moving too quickly it may not be able to acquire enough sensor measurements to accurately estimate its position and orientation which can result in errors and drift in the estimates leading to poor navigation and control performance (Dissanayake, Sukkarieh, et al., 2001). Also in case of cameras motion blur can occur if the UAV is moving too quickly to blurry or distorted images that are more difficult to process and use for SLAM (Hayakawa & Ishikawa, 2016). Accuracy of data acquired by sensors may be affected when the UAV is flying at different altitudes, which can also influence SLAM (Hyyppä et al., 2020).

5.5 | Strengths and limitations

There are several strengths and weaknesses of UAV-based SLAM limited by the advantages and disadvantages of the UAV platform. UAVs are high-speed and have excellent capability to cross terrain and obstacles compared with ground vehicles. Moreover, UAVs are able to have an advantageous view from high altitude which cannot be available to ground-based equipment (Mohr & Fitzpatrick, 2008). More importantly, the use of UAVs has less damage to the ecosystem (Kushleyev et al., 2013). However, current technologies of UAVs have some limitations. Due to the vibrations caused by the wind at high flight altitudes, the center of gravity will shift rapidly, and there is a high risk of losing control (Coppola et al., 2020). Also, due to limited flight time, limited remote controlling distance, and limited power consumption, UAVs cannot be used in critical missions, particularly for UAVs weighing less than 4 kg (Nieto et al., 2003).

In terms of algorithms, there are plenty of SLAM algorithms with high performance and robustness in UAV applications currently (Figure 8). However, there are some challenges for UAV-based SLAM in specific scenes. For example, limitations appear when UAVs implement long-term autonomy SLAM which may cover a large scale and span a long period, such as environmental monitoring in ocean or land exploration and large-scale precision agriculture. Due to the dynamic and harsh environment outdoors, data association in SLAM may fail easily (Martinez-Cantin & Castellanos, 2005; Sossalla et al., 2022). And there are many parameters tuned when initializing the SLAM, but unchanged parameters may cause failures of SLAM when environments change (Saeedi et al., 2019). Besides, hardware

including sensors and UAV platforms may have some errors when implementing tasks. This can cause some disasters, such as falling UAV or the lost UAVs. In addition, computational time and memory footprint are constrained by resources of UAV due to the extremely limited payload which may consume lots of resources when generating maps.

At the hardware level, the main sensors in SLAM are LiDAR and cameras (Figure 7). Limitations of LiDAR sensors are lack of variety of information acquired while for visual sensors limitation is small field of view leading to incomplete information acquisition. There are some SLAM studies trying to address challenges in sensors such as solving the lack of information from LiDAR by fusion of different sensors (Xu et al., 2022) or developing new sensors such as event cameras (Gehrig et al., 2020) or smell sensor to decrease computational requirement in algorithmic level. However, most current attempts do not focus on SLAM in UAV applications so there are still many research gaps on UAV-based SLAM. Also payload capacity of UAVs is not comparable to that of ground vehicles and handheld devices. Larger LiDARs and high-performance mobile computing devices are difficult to mount on UAV platforms so development of miniaturized sensor devices, real-time computing platforms, and refinement of SLAM algorithms are current research challenges in field of UAV-based SLAM.

There were other challenges on evaluation method in UAV-based SLAM. Currently most public data sets for evaluating the performance of SLAM were collected by handheld equipment or ground vehicles. Compared with other robot platforms (e.g., autonomous cars and mobile ground robots), few public data sets available for UAVs. From now on there are only seven public data sets available for evaluating UAV-based SLAM, including ICL (Delmerico et al., 2019), EuRoC (Burri et al., 2016), UZH–FPV Drone Racing (Majdik et al., 2017), Zurich Urban MAV (Oettershagen et al., 2016), Solar-powered UAV Sensing and Mapping data set (Oettershagen et al., 2016), Kagaru Airborne Stereo data set (Warren et al., 2014), and Event-Camera Data set (Mueggler et al., 2017). Lack of UAV data sets has led to lack of benchmark evaluations of UAV-based SLAM. However, most algorithm-oriented studies have heavily relied on public data sets for implementation and evaluation. And some public data sets are overexposed which may lead to biased evaluation (Liu, Fu, et al., 2021). In addition most public data sets cannot adequately reflect real-world circumstances that a UAV might face in a new environment because environments in public data sets are frequently preselected. Therefore how to avoid evaluation bias to improve robustness and generalizability of SLAM algorithms is a challenge. Moreover, most public data sets only include a single type of basic sensors, which is not the same as a real-world UAV would have access to. For example Middlebury (Seitz et al., 2006), EPFL (Strecha et al., 2008), TUM MonoVO data sets (Engel et al., 2016) only provide video data captured by cameras, and KinectFusion (Meister et al., 2012), TUM RGB-D (Sturm et al., 2012), and ETH RGB-D (Oleynikova et al., 2017) data sets only equip with RGB-D sensors.

5.6 | Future study

On the basis of the results and discussions, there are some research points for UAV-based SLAM in future studies. In terms of sensors, multisensor fusion can better support the robustness, stability, and accuracy of SLAM. For UAV-based SLAM, sensor miniaturization and the development of new sensors are major challenges for future research. For example, wide field-of-view cameras such as fisheye cameras or omnidirectional cameras (Caruso et al., 2015) allow for observing and reconstructing a wider scene. Event cameras (Vidal et al., 2018) are robust to motion blur and have very high dynamic range. With the miniaturization of sensors, it is possible to carry a wider variety of sensors in a SLAM with limited payload. The fusion of multiple sensors will not only improve the accuracy of SLAM but will also provide diverse feature that may be suitable for different specific tasks. Thus, SLAM with fusion of multiple sensors (LiDAR, cameras, IMU, GPS, sonar) will likely be the next research points in UAV-based SLAM (De Pazzi et al., 2022; Zhang et al., 2012).

For UAV platforms, although current UAV devices are sufficiently stable and have a wide range of exploration, fast battery consumption, lack of durability, and limited communication range for SLAM are issues that need to be addressed in future research (Galkin et al., 2019; Olsson et al., 2010). Furthermore, UAV swarm collaboration and collaboration between UAVs and ground vehicles will also address individual UAV shortcomings (Zhou et al., 2022). For example, UAV swarm can increase coverage area, enhance fault tolerance, and provide higher mission efficiency than single UAV (Zhou et al., 2020, 2022). The collaboration between UAVs and ground vehicle can supply energy to UAVs which ensures long-term flight (Li, Cheng, et al., 2021; Papachristos & Tzes, 2014).

Regarding SLAM algorithms, on small mobile platforms like UAVs, lightweight and real-time algorithms need to be developed to ensure a reasonable allocation of computational resources. Recent advances in computer vision and deep learning have shown great promise for improving the accuracy and efficiency of UAV-based SLAM. For example, semantic segmentation and object detection algorithms such as swin transformer (Liu, Lin, et al., 2021), deeplab-V3 (Chen et al., 2017), and YOLO-V7 (Wang et al., 2022) can address the issues of identification and classification of objects in the environment, enabling more effective localization and mapping (e.g., dynamic feature removal and object-based feature) as well as fruitful semantic information (Wu et al., 2022; Zhang et al., 2018). As mentioned above the lack of absolute scale for monocular camera SLAM can also be effectively addressed by the recent monocular depth estimation algorithm (Cordts et al., 2016). Similarly, 3D reconstruction methods can be used to generate more detailed and accurate maps of the environment, for example, Nerf (Mildenhall et al., 2021). By integrating the techniques such as computer vision, deep learning, and reinforcement learning, UAV-based SLAM can achieve more advanced and complex tasks, such as autonomous navigation, object manipulation, and so forth. Several studies (Botteghi et al., 2020; Li et al., 2018; Wen, Zhao, et al., 2020) have demonstrated the effectiveness of these techniques in improving the performance of SLAM. Moreover, the current SLAM algorithm based on UAV is applicable to fewer scenarios, and future research should collect and establish relevant open-source data sets for different scenarios to prepare for SLAM algorithms and benchmark tests.

Finally, the current research area of UAV-based SLAM is mainly focused on urban and rural environments, while the proportion of research in agriculture and industrial areas is low. The high mobility and autonomy of UAVs combined with SLAM offer significant potential for a wide range of applications, such as smart agriculture, infrastructure, and natural hazard monitoring. For instance, UAVs can collect high-resolution imagery and generate comprehensive maps of crop growth, health conditions, and yield estimation, enabling precision agriculture practices and more effective resource utilization with SLAM (Aslan et al., 2022; Cheein et al., 2011; Krul et al., 2021). UAV with SLAM can also be used to inspect and monitor structures, such as bridges, ships, pipelines, and buildings, to identify potential hazards and carry out maintenance and repairs more effectively (Bian et al., 2018; Sato & Anezaki, 2017). Similarly, UAV-based SLAM can also be applied in scenarios involving disaster management, allowing rapid mapping and assessment of affected areas to direct response activities (Lee et al., 2016). These are just a few examples of the many potential applications of UAV-based SLAM. As technology develops, it is likely that many new opportunities will arise.

The autonomy of UAVs is dependent on its ability to perceive and navigate their environment independently, and UAV-based SLAM is a key enabling technology to achieve this. Several recent studies have highlighted the potential of UAV-based SLAM for improving autonomy and increasing the range of tasks that UAVs can perform. For example, several studies (Naveed et al., 2022; Wang et al., 2019) showed that by integrating visual SLAM with reinforcement learning, robots could learn to navigate complex indoor environments with high efficiency and autonomy. Another study (Azpúrua et al., 2021) demonstrated the potential of UAV for performing complex inspection tasks in hazardous environments.

In future, the development of UAV-based SLAM can be combined with deep learning, sensor, battery, and communication technologies, leading to an unmanned, autonomous UAV system for every scenario in city, agriculture, forest, industry, river, and ocean.

6 | CONCLUSION

This scoping review provides an overview of different tasks and applications scenarios employing SLAM and UAVs with various sensors. Currently, the navigation of UAVs relies on GPS/global navigation satellite systems or expert operation within visible range. However, when it comes to GPS-denied outdoor environment, SLAM could be an important technology for UAV navigation and generate maps of surroundings.

The UAV-based SLAM is still in its infancy compared with other robotic platforms, such as autonomous cars, warehouse robots, and industrial robots. UAV applications with conventional geometric feature-based SLAM algorithms (e.g., ORB-SLAM, LSD-SLAM, and

VINS-Mono) are relatively mature and have some practical applications, such as fire detection, forest growth assessment, and infrastructure inspection. However, there are still relatively few other applications for UAV-based SLAM in other fields, and compared with SLAM for other robotic platforms, the technology of SLAM is still not widely combined with current technologies, such as deep learning. Lack of related data sets, real-time processing performance, and robustness for long-term autonomy are still challenges. As a next step, UAV-based SLAM will be developed in specific application scenarios and related data sets will be built.

At the moment, the monocular camera is the most widely used sensor in UAV-based SLAM due to its small size, lightweight, easy-to-use, and robust performance. Different types of sensors have their pros and cons in different tasks and application scenarios. Therefore, some studies have focused on multisensor fusion to compensate for the deficiencies of different sensors that can improve the accuracy of SLAM. However, except for IMU and GPS sensor data, there were none or few studies adopting sensor fusion for more than three different types of sensors. This may be due to the limitations of the UAV's payload, so it is equally important to balance the weight of the battery, computing platform, and sensors within the limited UAV payload.

Due to the current limitations of UAVs, sensors, and computational power, UAV-based SLAM experiments should be carefully considered. And some challenging influences should be considered in different experimental scenarios. Thus, we provide three factors from hardware, software, and external level in UAV-based SLAM experiment setup. Researchers can design their UAV-based SLAM experiments referring to our experimental factors that may influence the SLAM results (Harmat et al., 2015).

AUTHOR CONTRIBUTIONS

Kaiwen Wang: Conceptualization, methodology, formal analysis, investigation, writing—original draft, writing—review and editing, visualization, project administration. **Lammert Kooistra**: Conceptualization, methodology, writing—review and editing, visualization, and supervision. **Ruoxi Pan**: Methodology, formal analysis, writing—original draft, writing—review and editing, and visualization. **Wensheng Wang**: Supervision, funding, and project administration. **João Valente**: Conceptualization, methodology, writing—review and editing, visualization, supervision, and project administration.

ACKNOWLEDGMENTS

DATA AVAILABILITY STATEMENT

The data in this review paper, including the selected literature and extracted information, can be accessed at the link: https://docs.google.com/spreadsheets/d/1pLYAU2cFl8-QIfy5mx3DTFvJeRwSMGaA/edit?usp=sharing&ouid=111844158354137128412&rtpof=true&sd=true.

The data that supports the findings of this study are available in the supplementary material of this article.

ORCID

João Valente http://orcid.org/0000-0002-6241-4124

REFERENCES

Abeywickrama, H.V., Jayawickrama, B.A., He, Y. & Dutkiewicz, E. (2018) Comprehensive energy consumption model for unmanned aerial vehicles, based on empirical studies of battery performance. *IEEE Access*, 6, 58383–58394.

Ageed, I.M. & Abulrahman, F.H. (2020) Evaluation of UAV-based DEM for volume calculation. *Journal of Duhok University*, 23(1), 11–24.

Aguilar-Moreno, M. & Graña, M. (2022) On registration methods for SLAM with low resolution lidar sensor. *Logic Journal of the IGPL*, 31(4), 751–761.

Aqel, M.O., Marhaban, M.H., Saripan, M.I. & Ismail, N.B. (2016) Review of visual odometry: types, approaches, challenges, and applications. *SpringerPlus*, 5, 1–26.

Arksey, H. & O'Malley, L. (2005) Scoping studies: towards a methodological framework. *International Journal of Social Research Methodology*, 8(1), 19–32.

Aslan, M.F., Durdu, A., Sabanci, K., Ropelewska, E. & Gültekin, S.S. (2022) A comprehensive survey of the recent studies with UAV for precision agriculture in open fields and greenhouses. *Applied Sciences*, 12(3), 1047.

Azpúrua, H., Rezende, A., Potje, G., Júnior, G.P.d.C., Fernandes, R., Miranda, V. et al. (2021) Towards semi-autonomous robotic inspection and mapping in confined spaces with the espeleorobô. *Journal of Intelligent & Robotic Systems*, 101, 1–27.

Barron, J.L., Fleet, D.J. & Beauchemin, S.S. (1994) Performance of optical flow techniques. *International Journal of Computer Vision*, 12, 43–77.

Bar-Shalom, Y., Li, X.R. & Kirubarajan, T. (2001) *Estimation with applications to tracking and navigation: theory algorithms and software*. John Wiley & Sons.

Bay, H., Tuytelaars, T. & Van Gool, L. (2006) SURF: speeded up robust features. *Lecture Notes in Computer Science*, 3951, 404–417.

Bellotto, N., Burn, K., Fletcher, E. & Wermter, S. (2008) Appearance-based localization for mobile robots using digital zoom and visual compass. *Robotics and Autonomous Systems*, 56(2), 143–156.

Bian, J., Hui, X., Zhao, X. & Tan, M. (2018) A point-line-based SLAM framework for UAV close proximity transmission tower inspection. In: *2018 IEEE International Conference on Robotics and Biomimetics (ROBIO)*. IEEE, pp. 1016–1021.

Birk, A., Wiggerich, B., Bülow, H., Pfingsthorn, M. & Schwertfeger, S. (2011) Safety, security, and rescue missions with an unmanned aerial vehicle (UAV) aerial mosaicking and autonomous flight at the 2009 European land robots trials (ELROB) and the 2010 response robot evaluation exercises (RREE). *Journal of Intelligent & Robotic Systems*, 64, 57–76.

Bloesch, M., Omari, S., Hutter, M. & Siegwart, R. (2015). Robust visual inertial odometry using a direct EKF-based approach. In: *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. https://doi.org/10.1109/iros.2015.7353389

Bostelman, R., Hong, T. & Cheok, G. (2015) Navigation performance evaluation for automatic guided vehicles. In: *2015 IEEE International Conference on Technologies for Practical Robot Applications (TePRA)*. IEEE, pp. 1–6.

Botteghi, N., Sirmacek, B., Schulte, R., Poel, M. & Brune, C. (2020) Reinforcement learning helps SLAM: learning to build maps. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, 43, 329–335.

Bryson, M. & Sukkarieh, S. (2006) Active airborne localisation and exploration in unknown environments using inertial SLAM. In: *2006 IEEE Aerospace Conference*. IEEE, pp. 13.

Bujanca, M., Gafton, P., Saeedi, S., Nisbet, A., Bodin, B., O'Boyle, M.F. et al. (2019) SLAMBench 3.0: systematic automated reproducible

evaluation of SLAM systems for robot vision challenges and scene understanding. In: *2019 International Conference on Robotics and Automation (ICRA)*. IEEE, pp. 6351–6358.

Burri, M., Nikolic, J., Gohl, P., Schneider, T., Rehder, J., Omari, S. et al. (2016) The EuRoC micro aerial vehicle datasets. *The International Journal of Robotics Research*, 35(10), 1157–1163.

Cabreira, T.M., Brisolara, L.B. & Ferreira Jr., P.R. (2019) Survey on coverage path planning with unmanned aerial vehicles. *Drones*, 3(1), 4.

Cadena, C., Carlone, L., Carrillo, H., Latif, Y., Scaramuzza, D., Neira, J. et al. (2016) Past, present, and future of simultaneous localization and mapping: toward the robust-perception age. *IEEE Transactions on Robotics*, 32(6), 1309–1332.

Campbell, J., Sukthankar, R. & Nourbakhsh, I. (2004) Techniques for evaluating optical flow for visual odometry in extreme terrain. In: *2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (IEEE Cat. No. 04CH37566)*. Vol. 4. IEEE, pp. 3704–3711.

Campos, C., Elvira, R., Rodríguez, J.J.G., Montiel, J.M. & Tardós, J.D. (2021) ORB-SLAM3: an accurate open-source library for visual, visual–inertial, and multimap SLAM. *IEEE Transactions on Robotics*, 37(6), 1874–1890.

Campoy, P., Correa, J.F., Mondragón, I., Martínez, C., Olivares, M., Mejías, L. et al. (2009) Computer vision onboard UAVs for civilian tasks. *Journal of Intelligent and Robotic Systems*, 54, 105–135.

Carlevaris-Bianco, N., Ushani, A.K. & Eustice, R.M. (2016) University of Michigan north campus long-term vision and lidar dataset. *The International Journal of Robotics Research*, 35(9), 1023–1035.

Caruso, D., Engel, J. & Cremers, D. (2015) Large-scale direct SLAM for omnidirectional cameras. In: *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, pp. 141–148.

Cheein, F.A., Steiner, G., Paina, G.P. & Carelli, R. (2011) Optimized EIF-SLAM algorithm for precision agriculture mapping based on stems detection. *Computers and Electronics in Agriculture*, 78(2), 195–207.

Chen, X., Milioto, A., Palazzolo, E., Giguere, P., Behley, J. & Stachniss, C. (2019) SuMa++: efficient LiDAR-based semantic SLAM. In: *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, pp. 4530–4537.

Chen, L.-C., Papandreou, G., Schroff, F. & Adam, H. (2017) Rethinking atrous convolution for semantic image segmentation. arXiv preprint arXiv:1706.05587.

Coppola, M., McGuire, K.N., De Wagter, C. & De Croon, G.C. (2020) A survey on swarming with micro air vehicles: fundamental challenges and constraints. *Frontiers in Robotics and AI*, 7, 18.

Cordts, M., Omran, M., Ramos, S., Rehfeld, T., Enzweiler, M., Benenson, R. et al. (2016) The cityscapes dataset for semantic urban scene understanding. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3213–3223.

Cui, J.Q., Lai, S., Dong, X., Liu, P., Chen, B.M. & Lee, T.H. (2014) Autonomous navigation of UAV in forest. In: *2014 International Conference on Unmanned Aircraft Systems (ICUAS)*. IEEE, pp. 726–733.

Davison, A.J. (2003) Real-time simultaneous localisation and mapping with a single camera. In: *IEEE International Conference on Computer Vision*. Vol. 3. IEEE Computer Society, p. 1403.

Delmerico, J., Cieslewski, T., Rebecq, H., Faessler, M. & Scaramuzza, D. (2019) Are we ready for autonomous drone racing? The UZH-FPV drone racing dataset. In: *2019 International Conference on Robotics and Automation (ICRA)*. IEEE, pp. 6713–6719.

De Pazzi, D., Pertile, M. & Chiodini, S. (2022) 3D radiometric mapping by means of lidar SLAM and thermal camera data fusion. *Sensors*, 22(21), 8512.

Diep, J., Gómez-Casco, D., Villamide, X.O., Swinden, R.D. & Crosta, P. (2022) Performance analysis of mass-market GNSS receivers in UAV applications. In: *2022 10th Workshop on Satellite Navigation Technology (NAVITEC)*. IEEE, pp. 1–8.

Dissanayake, M.G., Newman, P., Clark, S., Durrant-Whyte, H.F. & Csorba, M. (2001) A solution to the simultaneous localization and map building (SLAM) problem. *IEEE Transactions on Robotics and Automation*, 17(3), 229–241.

Dissanayake, G., Sukkarieh, S., Nebot, E. & Durrant-Whyte, H. (2001) The aiding of a low-cost strapdown inertial measurement unit using vehicle model constraints for land vehicle applications. *IEEE Transactions on Robotics and Automation*, 17(5), 731–747.

Dvorak, P., Mullerova, J., Bartalos, T. & Bruna, J. (2015) Unmanned aerial vehicles for alien plant species detection and monitoring. In: *ISPRS International Conference on Unmanned Aerial Vehicles in Geomatics*.

Ellenberg, A., Branco, L., Krick, A., Bartoli, I. & Kontsos, A. (2015) Use of unmanned aerial vehicle for quantitative infrastructure evaluation. *Journal of Infrastructure Systems*, 21(3), 04014054.

Engel, J., Schöps, T. & Cremers, D. (2014) LSD-SLAM: large-scale direct monocular SLAM. In: *Proceedings of the Computer Vision-ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6–12, 2014, Part II 13*. Springer, pp. 834–849.

Engel, J., Usenko, V. & Cremers, D. (2016) A photometrically calibrated benchmark for monocular visual odometry. arXiv preprint arXiv:1607.02555.

Esfahlani, S.S. (2019) Mixed reality and remote sensing application of unmanned aerial vehicle in fire and smoke detection. *Journal of Industrial Information Integration*, 15, 42–49.

Fankhauser, P. & Hutter, M. (2016) A universal grid map library: implementation and use case for rough terrain navigation. *Robot Operating System (ROS)—The Complete Reference*, 1, 99–120.

Floreano, D. & Wood, R.J. (2015) Science, technology and the future of small autonomous drones. *Nature*, 521(7553), 460–466.

Forster, C., Pizzoli, M. & Scaramuzza, D. (2014) SVO: fast semi-direct monocular visual odometry. In: *2014 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, pp. 15–22.

Forster, C., Zhang, Z., Gassner, M., Werlberger, M. & Scaramuzza, D. (2016) SVO: semidirect visual odometry for monocular and multi-camera systems. *IEEE Transactions on Robotics*, 33(2), 249–265.

Galkin, B., Kibilda, J. & DaSilva, L.A. (2019) UAVs as mobile infrastructure: addressing battery lifetime. *IEEE Communications Magazine*, 57(6), 132–137.

Gee, T., James, J., Van Der Mark, W., Delmas, P. & Gimel'farb, G. (2016) Lidar guided stereo simultaneous localization and mapping (SLAM) for UAV outdoor 3-D scene reconstruction. In: *2016 International Conference on Image and Vision Computing New Zealand (IVCNZ)*. IEEE, pp. 1–6.

Gehrig, D., Gehrig, M., Hidalgo-Carrió, J. & Scaramuzza, D. (2020) Video to events: recycling video datasets for event cameras. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 3586–3595.

Geiger, A., Lenz, P. & Urtasun, R. (2012) Are we ready for autonomous driving? The KITTI vision benchmark suite. In: *2012 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, pp. 3354–3361.

Geneva, P., Eckenhoff, K., Yang, Y. & Huang, G. (2018) LIPS: LiDAR-inertial 3D plane SLAM. In: *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, pp. 123–130.

Godard, C., Mac Aodha, O., Firman, M. & Brostow, G.J. (2019) Digging into self-supervised monocular depth estimation. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 3828–3838.

Gómez-Candón, D., De Castro, A. & López-Granados, F. (2014) Assessing the accuracy of mosaics from unmanned aerial vehicle (UAV) imagery for precision agriculture purposes in wheat. *Precision Agriculture*, 15, 44–56.

Gonzalez, R., Rodriguez, F., Guzman, J.L., Pradalier, C. & Siegwart, R. (2012) Combined visual odometry and visual compass for off-road mobile robots localization. *Robotica*, 30(6), 865–878.

Gonzalez, R., Rodriguez, F., Guzman, J.L., Pradalier, C. & Siegwart, R. (2013) Control of off-road mobile robots using visual odometry and slip compensation. *Advanced Robotics*, 27(11), 893–906.

Grisetti, G., Stachniss, C. & Burgard, W. (2005) Improving grid-based SLAM with rao-blackwellized particle filters by adaptive proposals and selective resampling. In: *Proceedings of the 2005 IEEE International Conference on Robotics and Automation*. IEEE, pp. 2432–2437.

Hansen, J.G., Heiß, M., Kozlowski, M. & Kayacan, E. (2022) UAV trajectory evaluation in large industrial environments: a cost-effective solution. In: *2022 European Control Conference (ECC)*. IEEE, pp. 1336–1341.

Hardin, P.J. & Jensen, R.R. (2011) Small-scale unmanned aerial vehicles in environmental remote sensing: challenges and opportunities. *GIScience & Remote Sensing*, 48(1), 99–111.

Harmat, A., Trentini, M. & Sharf, I. (2015) Multi-camera tracking and mapping for unmanned aerial vehicles in unstructured environments. *Journal of Intelligent & Robotic Systems*, 78, 291–317.

Hayakawa, T. & Ishikawa, M. (2016) Development of motion-blur-compensated high-speed moving visual inspection vehicle for tunnels. *International Journal of Civil and Structural Engineering Research*, 5, 151–155.

He, Z., Iyer, R.V. & Chandler, P.R. (2006) Vision-based UAV flight control and obstacle avoidance. In: *2006 American Control Conference*. IEEE, pp. 5.

Henein, M., Zhang, J., Mahony, R. & Ila, V. (2020) Dynamic SLAM: the need for speed. In: *2020 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, pp. 2123–2129.

Heng, L., Lee, G.H. & Pollefeys, M. (2015) Self-calibration and visual SLAM with a multi-camera system on a micro aerial vehicle. *Autonomous Robots*, 39, 259–277.

Hess, W., Kohler, D., Rapp, H. & Andor, D. (2016) Real-time loop closure in 2D lidar SLAM. In: *2016 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, pp. 1271–1278.

Hinzmann, T., Schneider, T., Dymczyk, M., Schaffner, A., Lynen, S., Siegwart, R. et al. (2016) Monocular visual–inertial SLAM for fixed-wing UAVs using sliding window based nonlinear optimization. In: *Proceedings of the Advances in Visual Computing: 12th International Symposium, ISVC 2016, Las Vegas, NV, USA, December 12–14, 2016, Part I 12*. Springer, pp. 569–581.

Ho, K. & Newman, P. (2007) Detecting loop closure with scene sequences. *International Journal of Computer Vision*, 74(3), 261–286.

Hornung, A., Wurm, K.M., Bennewitz, M., Stachniss, C. & Burgard, W. (2013) Octomap: an efficient probabilistic 3D mapping framework based on octrees. *Autonomous robots*, 34, 189–206.

Huang, L. (2021) Review on lidar-based SLAM techniques. In: *2021 International Conference on Signal Processing and Machine Learning (CONF-SPML)*. IEEE, pp. 163–168.

Huang, Y., Reddy, K.N., Fletcher, R.S. & Pennington, D. (2018) UAV low-altitude remote sensing for precision weed management. *Weed Technology*, 32(1), 2–6.

Hyyppä, E., Yu, X., Kaartinen, H., Hakala, T., Kukko, A., Vastaranta, M. et al. (2020) Comparison of backpack, handheld, under-canopy UAV, and above-canopy UAV laser scanning for field reference data collection in boreal forests. *Remote Sensing*, 12(20), 3327.

Ji, Z. & Singh, S. (2014) LOAM: lidar odometry and mapping in real-time. In: *Robotics: Science and Systems Conference*.

Jin, Q., Liu, Y., Man, Y. & Li, F. (2019) Visual SLAM with RGB-D cameras. In: *2019 Chinese Control Conference(CCC)*. IEEE, pp. 4072–4077.

Jung, S., Choi, D., Song, S. & Myung, H. (2020) Bridge inspection using unmanned aerial vehicle based on HG-SLAM: hierarchical graph-based SLAM. *Remote Sensing*, 12(18), 3022.

Karlsson, N., Di Bernardo, E., Ostrowski, J., Goncalves, L., Pirjanian, P. & Munich, M.E. (2005) The vSLAM algorithm for robust localization and mapping. In: *Proceedings of the 2005 IEEE International Conference on Robotics and Automation*. IEEE, pp. 24–29.

Kaul, L., Zlot, R. & Bosse, M. (2016) Continuous-time three-dimensional mapping for micro aerial vehicles with a passively actuated rotating laser scanner. *Journal of Field Robotics*, 33(1), 103–132.

Khairuddin, A.R., Talib, M.S. & Haron, H. (2015) Review on simultaneous localization and mapping (SLAM). In: *2015 IEEE International Conference on Control System, Computing and Engineering (ICCSCE)*. IEEE, pp. 85–90.

Kim, A. & Eustice, R.M. (2013) Real-time visual SLAM for autonomous underwater hull inspection using visual saliency. *IEEE Transactions on Robotics*, 29(3), 719–733.

Kim, J.-H. & Sukkarieh, S. (2003) Airborne simultaneous localisation and map building. In: *2003 IEEE International Conference on Robotics and Automation (Cat. No. 03CH37422)*. Vol. 1. IEEE, pp. 406–411.

Kim, Y.-J., Won, C.-H., Pak, J.-M. & Lim, M.-T. (2007) Fuzzy adaptive particle filter for localization of a mobile robot. In: *Proceedings of the Knowledge-Based Intelligent Information and Engineering Systems: 11th International Conference, KES 2007, XVII Italian Workshop on Neural Networks, Vietri sul Mare, Italy, September 12–14, 2007, Part III 11*. Springer, pp. 41–48.

Klein, G. & Murray, D. (2007) Parallel tracking and mapping for small ar workspaces. In: *2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality*. IEEE, pp. 225–234.

Koenig, N. & Howard, A. (2004) Design and use paradigms for gazebo, an open-source multi-robot simulator. In: *2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (IEEE Cat. No. 04CH37566)*. Vol. 3. IEEE, pp. 2149–2154.

Konolige, K., Grisetti, G., Kümmerle, R., Burgard, W., Limketkai, B. & Vincent, R. (2010) Efficient sparse pose adjustment for 2D mapping. In: *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, pp. 22–29.

Kothari, M., Postlethwaite, I. & Gu, D.-W. (2014) UAV path following in windy urban environments. *Journal of Intelligent & Robotic Systems*, 74, 1013–1028.

Krul, S., Pantos, C., Frangulea, M. & Valente, J. (2021) Visual SLAM for indoor livestock and farming using a small drone with a monocular camera: a feasibility study. *Drones*, 5(2), 41.

Kumar, S.A., Sharma, B., Vanualailai, J. & Prasad, A. (2021) Stable switched controllers for a swarm of UGVs for hierarchal landmark navigation. *Swarm and Evolutionary Computation*, 65, 100926.

Kushleyev, A., Mellinger, D., Powers, C. & Kumar, V. (2013) Towards a swarm of agile micro quadrotors. *Autonomous Robots*, 35(4), 287–300.

Kuznietsov, Y., Stuckler, J. & Leibe, B. (2017) Semi-supervised deep learning for monocular depth map prediction. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 6647–6655.

Laible, S., Khan, Y.N., Bohlmann, K. & Zell, A. (2012) 3D lidar- and camera-based terrain classification under different lighting conditions. In: *Autonomous Mobile Systems 2012: 22. Fachgespräch Stuttgart, 26. bis 28. September 2012*. Springer, pp. 21–29.

Lategahn, H., Geiger, A. & Kitt, B. (2011) Visual SLAM for autonomous ground vehicles. In: *2011 IEEE International Conference on Robotics and Automation*. IEEE, pp. 1732–1737.

Lee, S., Har, D. & Kum, D. (2016) Drone-assisted disaster management: finding victims via infrared camera and lidar sensor fusion. In: *2016 3rd Asia-Pacific World Congress on Computer Science and Engineering (APWC on CSE)*. IEEE, pp. 84–89.

Lee, S.H. & de Croon, G. (2018) Stability-based scale estimation for monocular SLAM. *IEEE Robotics and Automation Letters*, 3(2), 780–787.

Leonard, J.J. & Durrant-Whyte, H.F. (1991) Mobile robot localization by tracking geometric beacons. *IEEE Transactions on Robotics and Automation*, 7(3), 376–382.

Leutenegger, S., Lynen, S., Bosse, M., Siegwart, R. & Furgale, P. (2015) Keyframe-based visual–inertial odometry using nonlinear optimization. *The International Journal of Robotics Research*, 34(3), 314–334.

Li, X., Ao, H., Belaroussi, R. & Gruyer, D. (2017) Fast semi-dense 3D semantic mapping with monocular visual SLAM. In: *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, pp. 385–390.

Li, J., Cheng, L., Wu, H., Xiong, L. & Wang, D. (2012) An overview of the simultaneous localization and mapping on mobile robot. In: *2012 Proceedings of International Conference on Modelling, Identification and Control*. IEEE, pp. 358–364.

Li, J., Cheng, Y., Zhou, J., Chen, J., Liu, Z., Hu, S. et al. (2021) Energy-efficient ground traversability mapping based on UAV–UGV collaborative system. *IEEE Transactions on Green Communications and Networking*, 6(1), 69–78.

Li, K., Li, M. & Hanebeck, U.D. (2021) Towards high-performance solid-state-lidar-inertial odometry and mapping. *IEEE Robotics and Automation Letters*, 6(3), 5167–5174.

Li, Y., Xie, C., Lu, H., Chen, X., Xiao, J. & Zhang, H. (2018) Scale-aware monocular SLAM based on convolutional neural network. In: *2018 IEEE International Conference on Information and Automation (ICIA)*. IEEE, pp. 51–56.

Li, Z., Yang, T., Li, G., Li, J. & Zhang, Y. (2016) Geodetic coordinate calculation based on monocular vision on UAV platform. In: *2016 IEEE 13th International Conference on Signal Processing (ICSP)*. IEEE, pp. 310–314.

Li-Chee-Ming, J. & Armenakis, C. (2018) UAV navigation system using line-based sensor pose estimation. *Geo-spatial Information Science*, 21(1), 2–11.

Liu, X., Nardari, G.V., Ojeda, F.C., Tao, Y., Zhou, A., Donnelly, T. et al. (2022) Large-scale autonomous flight with real-time semantic SLAM under dense forest canopy. *IEEE Robotics and Automation Letters*, 7(2), 5512–5519.

Liu, Y., Fu, Y., Chen, F., Goossens, B., Tao, W. & Zhao, H. (2021) Simultaneous localization and mapping related datasets: A comprehensive survey. arXiv preprint arXiv:2102.04036.

Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z. et al. (2021) Swin transformer: hierarchical vision transformer using shifted windows. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*. pp. 10012–10022.

Lowe, D.G. (2004) Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60, 91–110.

Lu, F. & Milios, E. (1997) Globally consistent range scan alignment for environment mapping. *Autonomous Robots*, 4, 333–349.

Lynch, L., Newe, T., Clifford, J., Coleman, J., Walsh, J. & Toal, D. (2018) Automated ground vehicle (AGV) and sensor technologies—a review. In: *2018 12th International Conference on Sensing Technology (ICST)*. IEEE, pp. 347–352.

Ma, Y., Wang, S., Yu, D. & Zhu, K. (2021) Robust visual–inertial odometry with point and line features for blade inspection UAV. *Industrial Robot: The International Journal of Robotics Research and Application*, 48(2), 179–188.

Majdik, A.L., Till, C. & Scaramuzza, D. (2017) The Zurich urban micro aerial vehicle dataset. *The International Journal of Robotics Research*, 36(3), 269–273.

Martínez-Barberá, H. & Herrero-Pérez, D. (2010) Autonomous navigation of an automated guided vehicle in industrial environments. *Robotics and Computer-Integrated Manufacturing*, 26(4), 296–311.

Martinez-Cantin, R. & Castellanos, J.A. (2005) Unscented SLAM for large-scale outdoor environments. In: *2005 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, pp. 3427–3432.

McCormac, J., Handa, A., Davison, A. & Leutenegger, S. (2017) Semanticfusion: dense 3D semantic mapping with convolutional neural networks. In: *2017 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, pp. 4628–4635.

Medeiros, A.C., Ratsamee, P., Orlosky, J., Uranishi, Y., Higashida, M. & Takemura, H. (2021) UAV target-selection: 3D pointing interface system for large-scale environment. In: *2021 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, pp. 3963–3969.

Meister, S., Izadi, S., Kohli, P., Hämmerle, M., Rother, C. & Kondermann, D. (2012) When can we use kinectfusion for ground truth acquisition. In: *Proceedings of the Workshop on Color-Depth Camera Fusion in Robotics*. Vol. 2. IEEE, p. 3.

Mildenhall, B., Srinivasan, P.P., Tancik, M., Barron, J.T., Ramamoorthi, R. & Ng, R. (2021) Nerf: representing scenes as neural radiance fields for view synthesis. *Communications of the ACM*, 65(1), 99–106.

Mohammed, A.S., Amamou, A., Ayevide, F.K., Kelouwani, S., Agbossou, K. & Zioui, N. (2020) The perception system of intelligent ground vehicles in all weather conditions: A systematic literature review. *Sensors*, 20(22), 6532.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G. & The PRISMA Group. (2009) Preferred reporting items for systematic reviews and meta-analyses: the prisma statement. *Annals of Internal Medicine*, 151(4), 264–269.

Mohr, B.B. & Fitzpatrick, D.L. (2008) Micro air vehicle navigation system. *IEEE Aerospace and Electronic Systems Magazine*, 23(4), 19–24.

Montemerlo, M., Thrun, S., Koller, D. & Wegbreit, B. (2002) *FastSLAM: a factored solution to the simultaneous localization and mapping problem*. AAAI/IAAI, p. 593598.

Moosmann, F. & Stiller, C. (2011) Velodyne SLAM. In: *2011 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, pp. 393–398.

Mourikis, A.I. & Roumeliotis, S.I. (2007) A multi-state constraint Kalman filter for vision-aided inertial navigation. In: *Proceedings of the 2007 IEEE International Conference on Robotics and Automation*. IEEE, pp. 3565–3572.

Mueggler, E., Rebecq, H., Gallego, G., Delbruck, T. & Scaramuzza, D. (2017) The event-camera dataset and simulator: event-based data for pose estimation, visual odometry, and SLAM. *The International Journal of Robotics Research*, 36(2), 142–149.

Munn, Z., Peters, M.D., Stern, C., Tufanaru, C., McArthur, A. & Aromataris, E. (2018) Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. *BMC Medical Research Methodology*, 18, 1–7.

Mur-Artal, R., Montiel, J.M.M. & Tardos, J.D. (2015) ORB-SLAM: a versatile and accurate monocular SLAM system. *IEEE Transactions on Robotics*, 31(5), 1147–1163.

Mur-Artal, R. & Tardós, J.D. (2017) ORB-SLAM2: an open-source SLAM system for monocular, stereo, and RGB-D cameras. *IEEE Transactions on Robotics*, 33(5), 1255–1262.

Naroditsky, O., Zhou, X.S., Gallier, J., Roumeliotis, S.I. & Daniilidis, K. (2011) Two efficient solutions for visual odometry using directional correspondence. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(4), 818–824.

Naveed, K., Anjum, M.L., Hussain, W. & Lee, D. (2022) Deep introspective SLAM: deep reinforcement learning based approach to avoid tracking failure in visual SLAM. *Autonomous Robots*, 46(6), 705–724.

Newcombe, R.A., Lovegrove, S.J. & Davison, A.J. (2011) DTAM: dense tracking and mapping in real-time. In: *2011 International Conference on Computer Vision*. IEEE, pp. 2320–2327.

Nex, F., Armenakis, C., Cramer, M., Cucci, D.A., Gerke, M., Honkavaara, E. et al. (2022) UAV in the advent of the twenties: where we stand and what is next. *ISPRS Journal of Photogrammetry and Remote Sensing*, 184, 215–242.

Nguyen, T.-M., Cao, M., Yuan, S., Lyu, Y., Nguyen, T.H. & Xie, L. (2021) Viral-fusion: a visual–inertial-ranging-lidar sensor fusion approach. *IEEE Transactions on Robotics*, 38(2), 958–977.

Nieto, J., Guivant, J., Nebot, E. & Thrun, S. (2003) Real time data association for FastSLAM. In: *2003 IEEE International Conference on Robotics and Automation (Cat. No. 03CH37422)*. Vol. 1. IEEE, pp. 412–418.

Nistér, D., Naroditsky, O. & Bergen, J. (2004) Visual odometry. In: *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004*, Vol. 1. IEEE, p. I.

Nistér, D., Naroditsky, O. & Bergen, J. (2006) Visual odometry for ground vehicle applications. *Journal of Field Robotics*, 23(1), 3–20.

Nourani-Vatani, N. & Pradalier, C. (2010) Scene change detection for vision-based topological mapping and localization. In: *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, pp. 3792–3797.

Nützi, G., Weiss, S., Scaramuzza, D. & Siegwart, R. (2011) Fusion of IMU and vision for absolute scale estimation in monocular SLAM. *Journal of Intelligent & Robotic Systems*, 61(1–4), 287–299.

Oettershagen, P., Stastny, T., Mantel, T., Melzer, A., Rudin, K., Gohl, P. et al. (2016) Long-endurance sensing and mapping using a hand-launchable solar-powered UAV. In: *Field and Service Robotics: Results of the 10th International Conference*. Springer, pp. 441–454.

Oleynikova, H., Taylor, Z., Fehr, M., Siegwart, R. & Nieto, J. (2017) Voxblox: Incremental 3D Euclidean signed distance fields for on-board MAV planning. In: *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, pp. 1366–1373.

Olsson, P.-M., Kvarnström, J., Doherty, P., Burdakov, O. & Holmberg, K. (2010) Generating UAV communication networks for monitoring and surveillance. In: *2010 11th International Conference on Control Automation Robotics & Vision*. IEEE, pp. 1070–1077.

Opromolla, R., Fasano, G., Rufino, G., Grassi, M. & Savvaris, A. (2016) Lidar-inertial integration for UAV localization and mapping in complex environments. In: *2016 International Conference on Unmanned Aircraft Systems (ICUAS)*. IEEE, pp. 649–656.

Papachristos, C. & Tzes, A. (2014) The power-tethered UAV–UGV team: a collaborative strategy for navigation in partially-mapped environments. In: *22nd Mediterranean Conference on Control and Automation*. IEEE, pp. 1153–1158.

Peters, M.D., Godfrey, C.M., Khalil, H., McInerney, P., Parker, D. & Soares, C.B. (2015) Guidance for conducting systematic scoping reviews. *JBI Evidence Implementation*, 13(3), 141–146.

Pinto, L.R., Vale, A., Brouwer, Y., Borbinha, J., Corisco, J., Ventura, R. et al. (2021) Radiological scouting, monitoring and inspection using drones. *Sensors*, 21(9), 3143.

Porav, H., Bruls, T. & Newman, P. (2019) Don't worry about the weather: unsupervised condition-dependent domain adaptation. In: *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*. IEEE, pp. 33–40.

Qin, C., Zhang, Y., Liu, Y. & Lv, G. (2021) Semantic loop closure detection based on graph matching in multi-objects scenes. *Journal of Visual Communication and Image Representation*, 76, 103072.

Qin, T., Cao, S., Pan, J. & Shen, S. (2019) A general optimization-based framework for global pose estimation with multiple sensors. arXiv preprint arXiv:1901.03642.

Qin, T., Li, P. & Shen, S. (2018) VINS-Mono: a robust and versatile monocular visual–inertial state estimator. *IEEE Transactions on Robotics*, 34(4), 1004–1020.

Quan, M., Piao, S., He, Y., Liu, X. & Qadir, M.Z. (2021) Monocular visual SLAM with points and lines for ground robots in particular scenes: parameterization for lines on ground. *Journal of Intelligent & Robotic Systems*, 101, 1–19.

Ravankar, A.A., Ravankar, A., Kobayashi, Y. & Emaru, T. (2018) Autonomous mapping and exploration with unmanned aerial vehicles using low cost sensors. *Multidisciplinary Digital Publishing Institute Proceedings*, 4(1), 44.

Rosinol, A., Abate, M., Chang, Y. & Carlone, L. (2020) Kimera: an open-source library for real-time metric-semantic localization and mapping. In: *2020 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, pp. 1689–1696.

Roussillon, C., Gonzalez, A., Solà, J., Codol, J.-M., Mansard, N., Lacroix, S. et al. (2011) RT-SLAM: a generic and real-time visual SLAM implementation. In: *Proceedings of the Computer Vision Systems: 8th International Conference, ICVS 2011, Sophia Antipolis, France, September 20-22, 2011. Vol. 8*. Springer, pp. 31–40.

Rublee, E., Rabaud, V., Konolige, K. & Bradski, G. (2011) ORB: an efficient alternative to SIFT or SURF. In: *2011 International Conference on Computer Vision*. IEEE, pp. 2564–2571.

Rusu, R.B. & Cousins, S. (2011) 3D is here: point cloud library (PCL). In: *2011 IEEE International Conference on Robotics and Automation*. IEEE, pp. 1–4.

Saad, A.M. & Tahar, K.N. (2019) Identification of rut and pothole by using multirotor unmanned aerial vehicle (UAV). *Measurement*, 137, 647–654.

Sadeghi, S. & Sohrabi, H. (2019) The effect of UAV flight altitude on the accuracy of individual tree height extraction in a broad-leaved forest. *The International Archives of the Photogrammetry, Remote Sensing, and Spatial Information Sciences*, 42(4), W18.

Saeedi, S., Carvalho, E.D., Li, W., Tzoumanikas, D., Leutenegger, S., Kelly, P.H. et al. (2019) Characterizing visual localization and mapping datasets. In: *2019 International Conference on Robotics and Automation (ICRA)*. IEEE, pp. 6699–6705.

Sato, S. & Anezaki, T. (2017) Autonomous flight drone for infrastructure (transmission line) inspection (2). In: *2017 International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS)*. IEEE, pp. 294–296.

Scaramuzza, D. & Fraundorfer, F. (2011) Visual odometry [tutorial]. *IEEE Robotics & Automation Magazine*, 18(4), 80–92.

Schultz, A., Gilabert, R., Bharadwaj, A., de Haag, M.U. & Zhu, Z. (2016) A navigation and mapping method for UAS during under-the-canopy forest operations. In: *2016 IEEE/ION Position, Location and Navigation Symposium (PLANS)*. IEEE, pp. 739–746.

Seitz, S.M., Curless, B., Diebel, J., Scharstein, D. & Szeliski, R. (2006) A comparison and evaluation of multi-view stereo reconstruction algorithms. In: *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06). Vol. 1*. IEEE, pp. 519–528.

Semsch, E., Jakob, M., Pavlicek, D. & Pechoucek, M. (2009) Autonomous UAV surveillance in complex urban environments. In: *2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology. Vol. 2*. IEEE, pp. 82–85.

Shahzadi, R., Ali, M., Khan, H.Z. & Naeem, M. (2021) UAV assisted 5G and beyond wireless networks: a survey. *Journal of Network and Computer Applications*, 189, 103114.

Shalal, N., Low, T., McCarthy, C. & Hancock, N. (2013) A review of autonomous navigation systems in agricultural environments. In: *SEAg 2013: innovative agricultural technologies for a sustainable future*.

Shan, T. & Englot, B. (2018) LeGO-LOAM: lightweight and ground-optimized lidar odometry and mapping on variable terrain. In: *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, pp. 4758–4765.

Shan, T., Englot, B., Meyers, D., Wang, W., Ratti, C. & Rus, D. (2020) LIO-SAM: tightly-coupled lidar inertial odometry via smoothing and mapping. In: *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, pp. 5135–5142.

Shao, P., Mo, F., Chen, Y., Ding, N. & Huang, R. (2021) Monocular object SLAM using quadrics and landmark reference map for outdoor UAV applications. In: *2021 IEEE International Conference on Real-time Computing and Robotics (RCAR)*. IEEE, pp. 1195–1201.

Siegwart, R., Nourbakhsh, I.R. & Scaramuzza, D. (2011) *Introduction to autonomous mobile robots*. MIT Press.

Silva, A., Mendonça, R. & Santana, P. (2020) Monocular trail detection and tracking aided by visual SLAM for small unmanned aerial vehicles. *Journal of Intelligent & Robotic Systems*, 97(3–4), 531–551.

Sivic, J. & Zisserman, A. (2008) Efficient visual search of videos cast as text retrieval. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(4), 591–606.

Smith, R. (1988) A stochastic map for uncertain spatial relationships. In: *Proceedings of the International Symposium on Robotics Research*, pp. 467–474.

Smith, R.C. & Cheeseman, P. (1986) On the representation and estimation of spatial uncertainty. *The International Journal of Robotics Research*, 5(4), 56–68.

Sossalla, P., Hofer, J., Rischke, J., Busch, J., Nguyen, G.T., Reisslein, M. et al. (2022) Optimizing edge SLAM: judicious parameter settings and parallelized map updates. In: *GLOBECOM 2022–2022 IEEE Global Communications Conference*. IEEE, pp. 1954–1959.

Steenbeek, A. & Nex, F. (2022) CNN-based dense monocular visual SLAM for real-time UAV exploration in emergency conditions. *Drones*, 6(3), 79.

Strecha, C., Von Hansen, W., Van Gool, L., Fua, P. & Thoennessen, U. (2008) On benchmarking camera calibration and multi-view stereo for high resolution imagery. In: *2008 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, pp. 1–8.

Sturm, J., Engelhard, N., Endres, F., Burgard, W. & Cremers, D. (2012) A benchmark for the evaluation of RGB-D SLAM systems. In: *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, pp. 573–580.

Sucar, E. & Hayet, J.-B. (2018) Bayesian scale estimation for monocular SLAM based on generic object detection for correcting scale drift. In: *2018 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, pp. 5152–5158.

Sujiwo, A., Ando, T., Takeuchi, E., Ninomiya, Y. & Edahiro, M. (2016) Monocular vision-based localization using ORB-SLAM with lidar-aided mapping in real-world robot challenge. *Journal of Robotics and Mechatronics*, 28(4), 479–490.

Suzuki, T., Amano, Y. & Hashizume, T. (2011) Development of a SIFT based monocular EKF-SLAM algorithm for a small unmanned aerial vehicle. In: *SICE Annual Conference 2011*. IEEE, pp. 1656–1659.

Suzuki, T., Amano, Y., Hashizume, T. & Suzuki, S. (2011) 3D terrain reconstruction by small unmanned aerial vehicle using SIFT-based monocular SLAM. *Journal of Robotics and Mechatronics*, 23(2), 292–301.

Taisho, T. & Kanji, T. (2016) Mining DCNN landmarks for long-term visual SLAM. In: *2016 IEEE International Conference on Robotics and Biomimetics (ROBIO)*. IEEE, pp. 570–576.

Tateno, K., Tombari, F., Laina, I. & Navab, N. (2017) CNN-SLAM: real-time dense monocular SLAM with learned depth prediction. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. pp. 6243–6252.

Trejos, K., Rincón, L., Bolaños, M., Fallas, J. & Marín, L. (2022) 2D SLAM algorithms characterization, calibration, and comparison considering pose error, map accuracy as well as CPU and memory usage. *Sensors*, 22(18), 6903.

Triggs, B., McLauchlan, P.F., Hartley, R.I. & Fitzgibbon, A.W. (2000) Bundle adjustment—a modern synthesis. In: *Proceedings of the Vision Algorithms: Theory and Practice: International Workshop on Vision Algorithms Corfu, Greece, September 21–22, 1999*. Springer, pp. 298–372.

Tuna, G., Gulez, K., Gungor, V.C. & Mumcu, T.V. (2012) Evaluations of different simultaneous localization and mapping (SLAM) algorithms. In: *IECON 2012—38th Annual Conference on IEEE Industrial Electronics Society*. IEEE, pp. 2693–2698.

Usenko, V., von Stumberg, L., Stückler, J. & Cremers, D. (2020) TUM flyers: vision–based MAV navigation for systematic inspection of structures. In: *Bringing innovative robotic technologies from research labs to industrial end-users: the experience of the European Robotics Challenges*, pp. 189–209.

Uto, K., Seki, H., Saito, G. & Kosugi, Y. (2013) Characterization of rice paddies by a UAV-mounted miniature hyperspectral sensor system. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 6(2), 851–860.

Vargas, J., Alsweiss, S., Toker, O., Razdan, R. & Santos, J. (2021) An overview of autonomous vehicles sensors and their vulnerability to weather conditions. *Sensors*, 21(16), 5397.

Vidal, A.R., Rebecq, H., Horstschaefer, T. & Scaramuzza, D. (2018) Ultimate SLAM? Combining events, images, and IMU for robust visual SLAM in HDR and high-speed scenarios. *IEEE Robotics and Automation Letters*, 3(2), 994–1001.

Wang, C.-Y., Bochkovskiy, A. & Liao, H.-Y.M. (2022) YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. arXiv preprint arXiv:2207.02696.

Wang, C., Wang, J., Shen, Y. & Zhang, X. (2019) Autonomous navigation of UAVs in large-scale complex environments: a deep reinforcement learning approach. *IEEE Transactions on Vehicular Technology*, 68(3), 2124–2136.

Warren, M., McKinnon, D., He, H., Glover, A., Shiel, M. & Upcroft, B. (2014) Large scale monocular vision-only mapping from a fixed-wing sUAS. In: *Field and Service Robotics: Results of the 8th International Conference*. Springer, pp. 495–509.

Wen, S., Zhao, Y., Yuan, X., Wang, Z., Zhang, D. & Manfredi, L. (2020) Path planning for active SLAM based on deep reinforcement learning under unknown environments. *Intelligent Service Robotics*, 13, 263–272.

Wen, W., Zhou, Y., Zhang, G., Fahandezh-Saadi, S., Bai, X., Zhan, W. et al. (2020) Urbanloco: a full sensor suite dataset for mapping and localization in urban scenes. In: *2020 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, pp. 2310–2316.

Wendel, A., Maurer, M., Graber, G., Pock, T. & Bischof, H. (2012) Dense reconstruction on-the-fly. In: *2012 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, pp. 1450–1457.

Wu, W., Guo, L., Gao, H., You, Z., Liu, Y. & Chen, Z. (2022) YOLO-SLAM: a semantic SLAM system towards dynamic environment with geometric constraint. *Neural Computing and Applications*, 34(8), 1–16.

Xie, X., Yang, T., Ning, Y., Zhang, F. & Zhang, Y. (2021) A monocular visual odometry method based on virtual-real hybrid map in low-texture outdoor environment. *Sensors*, 21(10), 3394.

Xu, W., Cai, Y., He, D., Lin, J. & Zhang, F. (2022) FAST-LIO2: fast direct lidar-inertial odometry. *IEEE Transactions on Robotics*, 38(4), 2053–2073.

Yan, Z., Sun, L., Krajník, T. & Ruichek, Y. (2020) EU long-term dataset with multiple sensors for autonomous driving. In: *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, pp. 10697–10704.

Yang, J., Dani, A., Chung, S.-J. & Hutchinson, S. (2017) Vision-based localization and robot-centric mapping in riverine environments. *Journal of Field Robotics*, 34(3), 429–450.

Yang, L., Li, B., Li, W., Liu, Z., Yang, G. & Xiao, J. (2017) A robotic system towards concrete structure spalling and crack database. In: *2017 IEEE International Conference on Robotics and Biomimetics (ROBIO)*. IEEE, pp. 1276–1281.

Yang, J., Rao, D., Chung, S.J. & Hutchinson, S. (2011) Monocular vision based navigation in gps-denied riverine environments. In: *AIAA Infotech@Aerospace Conference and Exhibit*.

Yang, Y., Xiong, J., She, X., Liu, C., Yang, C. & Li, J. (2019) Passive initialization method based on motion characteristics for monocular SLAM. *Complexity*, 2019, 1–11.

Zaffar, M., Ehsan, S., Stolkin, R. & Maier, K.M. (2018) Sensors, SLAM and long-term autonomy: a review. In: *2018 NASA/ESA Conference on Adaptive Hardware and Systems (AHS)*. IEEE, pp. 285–290.

Zhan, H., Garg, R., Weerasekera, C.S., Li, K., Agarwal, H. & Reid, I. (2018) Unsupervised learning of monocular depth estimation and visual odometry with deep feature reconstruction. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. pp. 340–349.

Zhang, X., Rad, A.B. & Wong, Y.-K. (2012) Sensor fusion of monocular cameras and laser rangefinders for line-based simultaneous localization and mapping (SLAM) tasks in autonomous mobile robots. *Sensors*, 12(1), 429–452.

Zhang, J. & Singh, S. (2015) Visual-lidar odometry and mapping: low-drift, robust, and fast. In: *2015 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, pp. 2174–2181.

Zhang, L., Wei, L., Shen, P., Wei, W., Zhu, G. & Song, J. (2018) Semantic SLAM based on object detection and improved octomap. *IEEE Access*, 6, 75545–75559.

Zhou, B., He, Y., Qian, K., Ma, X. & Li, X. (2021) S4-SLAM: a real-time 3D LiDAR SLAM system for ground/watersurface multi-scene outdoor applications. *Autonomous Robots*, 45, 77–98.

Zhou, Y., Rao, B. & Wang, W. (2020) UAV swarm intelligence: recent advances and future trends. *IEEE Access*, 8, 183856–183878.

Zhou, X., Wen, X., Wang, Z., Gao, Y., Li, H., Wang, Q., et al. (2022) Swarm of micro flying robots in the wild. *Science Robotics*, 7(66), eabm5954.

Zhu, R., Guo, Z. & Zhang, X. (2021) Forest 3D reconstruction and individual tree parameter extraction combining close-range photo enhancement and feature matching. *Remote Sensing*, 13(9), 1633.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

APPENDIX A

Search formula

Web of Science:

(((TS = (SLAM) OR TS = (simultaneous localization and mapping)) AND (TS = (UAV) OR TS = (UAS) OR TS = (MAV) OR TS = (unmanned aerial vehicle) OR TS = (unmanned aerial system) OR TS = (microaerial vehicle) OR TS = (drone)) AND (TS = (outdoor) OR TS = (field) OR TS = (orchard) OR TS = (forest) OR TS = (vineyard) OR TS = (park) OR TS = (mountain) OR TS = (grassland) OR (TS = (GPS-denied) NOT TS = (indoor))))

Scopus:

(((TITLE-ABS-KEY(SLAM) OR TITLE-ABS-KEY(simultaneous localization and mapping)) AND (TITLE-ABS-KEY(UAV) OR TITLE-ABS-KEY(UAS) OR TITLE-ABS-KEY(MAV) OR TITLE-ABS-KEY(unmanned aerial vehicle) OR TITLE-ABS-KEY(unmanned aerial system) OR TITLE-ABS-KEY(microaerial vehicle) OR TITLE-ABS-KEY(drone)) AND (TITLE-ABS-KEY(outdoor) OR TITLE-ABS-KEY(field) OR TITLE-ABS-KEY(orchard) OR TITLE-ABS-KEY(forest) OR TITLE-ABS-KEY(vineyard) OR TITLE-ABS-KEY(park) OR TITLE-ABS-KEY(mountain) OR TITLE-ABS-KEY(grassland) OR (TITLE-ABS-KEY(GPS-denied) AND NOT TITLE-ABS-KEY(indoor))))

IEEE database:

(((("All Metadata": SLAM) OR ("All Metadata": simultaneous localization and mapping)) AND (("All Metadata": UAV) OR ("All Metadata": UAS) OR ("All Metadata": MAV) OR ("All Metadata": unmanned aerial vehicle) OR ("All Metadata": unmanned aerial system) OR ("All Metadata": microaerial vehicle) OR ("All Metadata": drone)) AND (("All Metadata": outdoor) OR ("All Metadata": field) OR ("All Metadata": orchard) OR ("All Metadata": forest) OR ("All Metadata": vineyard) OR ("All Metadata": park) OR ("All Metadata": mountain) OR ("All Metadata": grassland) OR (("All Metadata": GPS-denied) NOT ("All Metadata": indoor))))

See Figure A1.
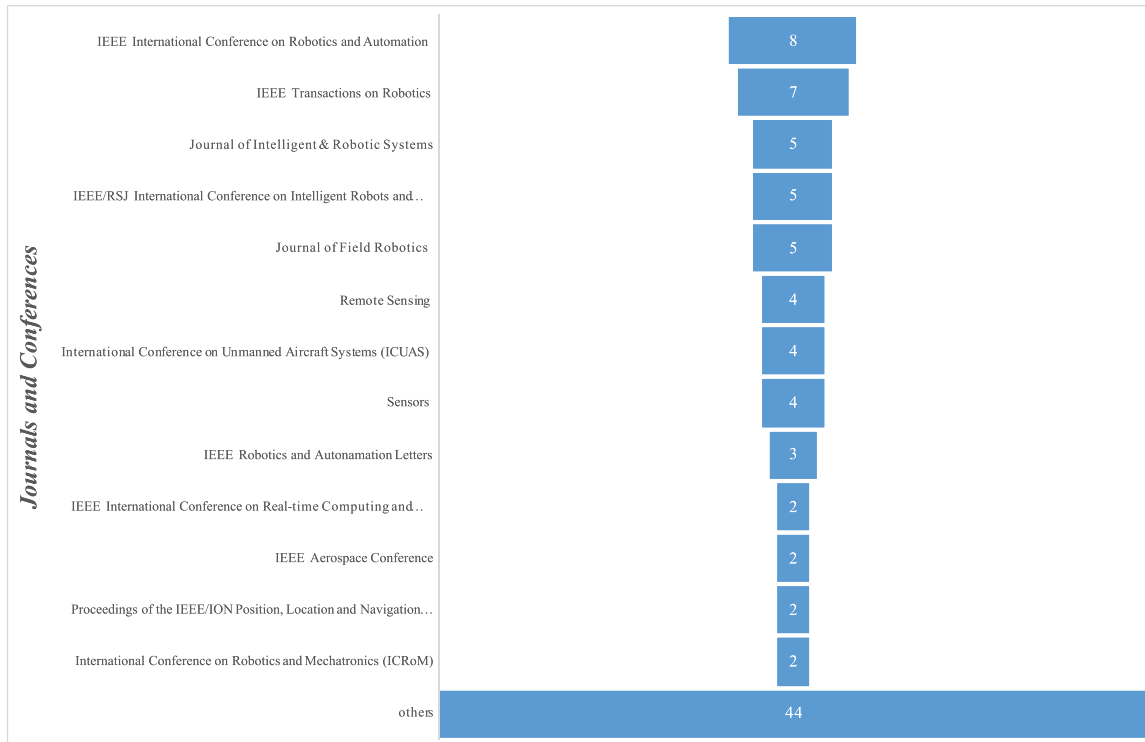
FIGURE A1 General Information on the included studies illustrating the distribution of the journals and conferences.