Neurocomputing 700 (2026) 134497

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Survey paper

Advanced techniques and applications of LiDAR place recognition in agricultural environments: A comprehensive survey

Judith Vilella-Cantos*, Mónica Ballesta, David Valiente, María Flores, Luis Payá

University Institute for Engineering Research, Miguel Hernández University, Avda. de la Universidad S/N, Edificio Innova, Elche, 03202, Alicante, Spain

ARTICLE INFO

Communicated by M. Chen

Keywords:
LiDAR place recognition
Agricultural environments
Autonomous robotics
Precision agriculture
Localization

ABSTRACT

An optimal solution to the localization problem is essential for developing autonomous robotic systems. Apart from autonomous vehicles, precision agriculture is one of the fields that can benefit most from these systems. Although LiDAR Place Recognition (LPR) is a widely used technique in recent years to achieve accurate localization, it is mostly used in urban settings. However, the lack of distinctive features and the unstructured nature of agricultural environments make place recognition challenging. This work presents a comprehensive review of state-of-the-art deep learning applications for agricultural environments and LPR techniques. We focus on the challenges that arise in these environments. We analyze the existing approaches, datasets, and metrics used to evaluate LPR system performance and discuss the limitations and future directions of research in this field. This is the first survey that focuses on LiDAR-based localization in agricultural settings, with the aim of providing a thorough understanding and fostering further research in this specialized domain.

1. Introduction

The increasing global population has led to a corresponding rise in the demand for food production. In order to address this growing demand, the agricultural sector is increasingly turning to technological advancements and automation to enhance efficiency and productivity. Additionally, the rising average age of the agricultural workforce is a contributing factor to the demand for automation [1]. The integration of autonomous robotic systems into agricultural practices could induce a paradigm shift within the field, transforming the landscape of agriculture in the following ways: firstly, the implementation of such systems could lead to a reduction in labor costs; secondly, there is the possibility of increasing crop yields; and thirdly, the environmental impact could be minimized [2,3]. However, for these systems to operate effectively, they must possess the capability to accurately localize themselves within their environment.

Place recognition approaches have been extensively utilized in the domain of robotics to achieve precise localization. These techniques entail the robot's ability to recognize previously visited locations based on sensory data, thereby determining its position within a map [4,5]. Light Detection and Ranging (LiDAR) sensors are frequently employed for place recognition due to their capacity to provide precise three-dimensional information about the environment. Furthermore, while vision solutions have historically dominated the field, LiDAR sensors have demonstrated enhanced resilience to variations in lighting and weather conditions [6].

The performance of LiDAR place recognition (LPR) is highly dependent on the characteristics of the environment. A significant number of recent approaches emphasize the use of urban datasets for the training of Deep Learning (DL) models [7]. However, these models may not always exhibit optimal generalizability when applied to agricultural environments. The presence of characteristic elements, such as buildings, roads, and vehicles, in urban environments can be readily identified by sophisticated algorithms. Conversely, agricultural environments frequently exhibit characteristics that are unstructured and void of distinctive features, thereby compromising the efficacy of place recognition algorithms in accurately identifying locations [8].

This survey addresses agricultural environments as an emerging field for LPR, providing comprehensive review of the state-of-the-art (SOTA), with a special focus on the challenges that arise in agricultural environments. Prior to that, we delve into the evolution of agricultural applications with DL approaches, with a focus on the recent advances and needs within the localization problem.

The contributions of this work are as follows:

- A comprehensive review of the SOTA in LPR, concentrating on actual challenges within agricultural settings.

* Corresponding author.
Email address: jvilella@umh.es (J. Vilella-Cantos).

https://doi.org/10.1016/j.neucom.2026.134497
Received 23 February 2026; Received in revised form 31 March 2026; Accepted 12 July 2026
Available online 13 July 2026

- An analysis of the different existing approaches, datasets and metrics used to evaluate the performance of localization approaches in agricultural scenarios.
- A critical review of urban-specific LPR benchmarks versus agricultural needs, emphasizing the performance gaps caused by high temporal variability and environmental dynamism.

1.1. Scope of the survey

The present survey focuses on the use of LiDAR for place recognition in agricultural environments. While numerous surveys on place recognition in urban environments already exist, there is a lack of comprehensive reviews that specifically address the challenges and solutions for agricultural settings. The objective of this survey is to bridge this knowledge gap by offering a comprehensive analysis of the SOTA in DL applications in agricultural environments, with a particular focus on the LPR problem. In this study, particular emphasis is placed on methodologies employed from 2020 onward, as the field has witnessed substantial advancements in recent years, particularly with the emergence of DL techniques. In addition, relevant works from earlier years have been included to provide context and background for the current SOTA.

Other place recognition surveys, such as [9,10], take a general approach to the place recognition problem, reviewing the various vision, LiDAR, and cross-modal approaches developed in recent years. The goal is to find a robust solution to the place recognition problem using any given sensor. Meanwhile, surveys such as [6,7] focus specifically on a modality of LiDAR place recognition. While these surveys provide valuable insights into the general landscape of place recognition, they do not address the specific challenges and solutions associated with agricultural environments. Conversely, surveys such as [3,11–13] comprehensively review DL applications in agricultural environments and robotics applications in these types of scenarios. However, they do not specifically address the LPR problem for accurate localization. Our survey complements these works by providing a specialized review of LPR in agricultural settings and addressing its unique characteristics and requirements. We believe this area is still unexplored and deserves a comprehensive review because it has the potential to significantly impact the development of autonomous robotic systems for agriculture.

1.2. Organization of the survey

The remainder of the survey is structured as follows. Section 2 explores the main applications of DL in these settings over the past several years. A thorough examination of the existing DL solutions for localization problems, such as place recognition and SLAM, in agricultural environments is presented in Section 3. This section encompasses both vision-based and LiDAR-based methodologies in order to provide a broader view of how the localization problem has been approached in these types of settings. In Section 4, the first approaches (urban-tailored) to the LPR problem are discussed, establishing the difference between learning-based and geometry-based methods. Other approaches, such as point-cloud preprocessing and cross-modal solutions, are also discussed, and a comparison between the different methods is provided. Additionally, this section provides a detailed discussion of the specific localization challenges that arise in agricultural environments and a qualitative comparison of different structured and unstructured scenarios. Section 5 reviews the existing datasets collected in agricultural settings. These datasets are examined in terms of their traditional use in detection and segmentation, as well as in terms of the more recent localization-focused datasets. As outlined in Section 6, the evaluation metrics employed to assess the performance of LPR algorithms are discussed. Finally, the concluding Section 7, presents a thorough summary of the research findings and envisages several key insights for future studies in this field.

Fig. 1 provides a visual overview of the structure of the survey, illustrating the logical flow between the various sections and how they are interconnected to provide a comprehensive understanding of LPR in agricultural environments. The dashed arrows highlight important aspects of the survey within the connected sections. For example, Section 4 discusses the comparison between agricultural and urban environments. This figure serves as a roadmap for readers, guiding them through the survey's content and highlighting the relationships between different topics.

2. Agricultural deep learning applications

The application of DL in agricultural environments has expanded rapidly in recent years. While traditional computer vision methods often struggle with the variability of outdoor settings, supervised and unsupervised DL models have demonstrated remarkable performance in core perception tasks, such as semantic segmentation, object detection, crop classification, and fruit counting.

However, deploying a localization DL model in real-world agricultural scenarios remains non-trivial compared to structured industrial or urban environments. Agricultural settings are characterized by high

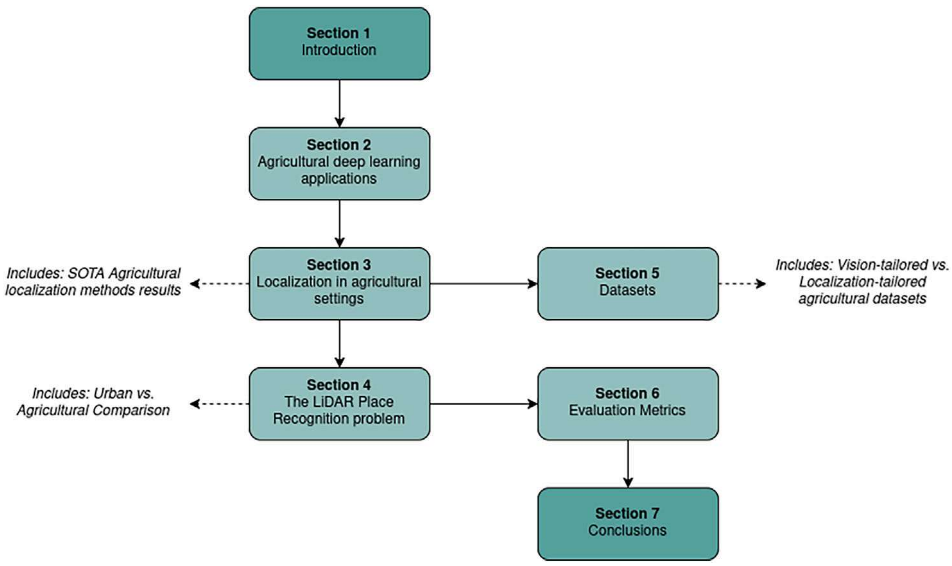

Fig. 1. Manuscript organization and logical flow of the survey.

unstructuredness, severe occlusions, variable lighting conditions, and seasonal appearance changes [14,15]. These factors demand robust architectures capable of generalizing across different crop stages and environmental conditions.

In this section, we review the primary applications of DL in agriculture, categorizing them by their specific utility. We analyze how SOTA approaches address the uniquely unstructured nature of these environments to achieve reliable perception.

2.1. Fields of application

According to the European Commission, the average age of farmers in the European Union (EU) is 57 years, with only 12% of farmers under the age of 40. This suggests that the agricultural sector is experiencing a shortage of labor [16]. The integration of robotics within agricultural settings is driven by the objective of enhancing efficiency, mitigating labor expenses, and augmenting crop yields [17]. These agricultural robots have the capacity to execute tasks such as planting, harvesting, and crop monitoring with a high degree of precision. To do so, a reliable sensor suite is imperative for the successful operation of these robotic systems. In this section, we will examine several primary applications of DL techniques in agricultural environments, with a particular emphasis on the past five years.

2.1.1. Precision agriculture

The prediction of food needs is a problem that has been studied for centuries as it is a key aspect in the preservation of the human species [18]. Following current trends, the constant growth of the human population requires an increase in food production and precision agriculture can help achieve this goal by improving the efficiency of farming practices [19,20]. The United Nations' world population prospects point out that the population will reach 8.6 billion people by 2030 [21]. Precision agriculture consists of the use of technology to optimize crop production and reduce waste. This is crucial in a context where climate change challenges agriculture and food security must be ensured for a larger population [22]. Accurate localization sustained by reliable sensing sources is crucial for precision agriculture, as it allows farmers to precisely apply fertilizers, pesticides, and other inputs to specific areas of the field [23]. Fig. 2 depicts the evolution of trends related to precision agriculture in a schema proposed by Karunathilake et al. [24].

Unmanned aerial vehicles (UAVs) provide high utility in precision agriculture, as they can be used to monitor crop health, identify areas of stress, and optimize irrigation. UAVs are more economical than manned aircraft, and provide higher quality images with respect to satellites [25–27]. UAV-based remote sensing for precision agriculture has a wide variety of benefits, such as improved yield prediction, disease detection, and resource management [28]. In what regards real-world applications, in [29], the authors propose an Internet-of-Things (IoT) UAV-based system for precision agriculture, focused on solving accurately the path planning problem in order to achieve optimal coverage of the field. In [30], a solution to the path planning problem for UAVs in precision agriculture is proposed, using a genetic algorithm approach to optimize the coverage of the field by a heterogeneous group of UAV vehicles.

Precision agriculture benefits from robust DL networks that accurately detect and classify the crops [31,32]. These networks can be trained on large datasets of agricultural images, enabling them to learn the unique characteristics of different crops and identify them with high accuracy. Reviews such as [33] analyze the performance of different DL architectures for crop detection and classification, highlighting the strengths and weaknesses of each approach.

2.1.2. Autonomous agricultural vehicles

With a robust solution to the localization problem, several applications can be unlocked in the agricultural field [34–36]. One of the main applications is autonomous agricultural vehicles, such as tractors and harvesters [37]. An autonomous agricultural vehicle can perform a variety of tasks, including planting, fertilizing, pruning [38,39], hill farming [40], and harvesting [41–43]. In orchard environments, autonomous robots can also be useful for transporting operators performing tree maintenance, as earlier studies from 2015 indicate [44]. For an autonomous agricultural vehicle to operate efficiently, it needs an optimal sensor setup that allows it to perceive the environment accurately [45] and a robust solution to the localization problem [46]. LPR can help these vehicles recognize previously visited locations and navigate the field more efficiently. Fig. 3 shows an example of an autonomous agricultural vehicle operating in a vineyard environment.

A reliable path planning solution is crucial for the correct functioning of an autonomous agricultural vehicle [47,48]. DeepWay [49] is an example of a DL approach that predicts waypoints in vineyard settings for global path generation taking an occupancy grid of the vineyard as an input. By accurately localizing itself within the field, the vehicle can optimize its path and reduce overlap, leading to more efficient operations and reduced fuel consumption [50]. As an extension to the DeepWay method, Salvetti et al. [51] propose a DL method that clusters waypoints

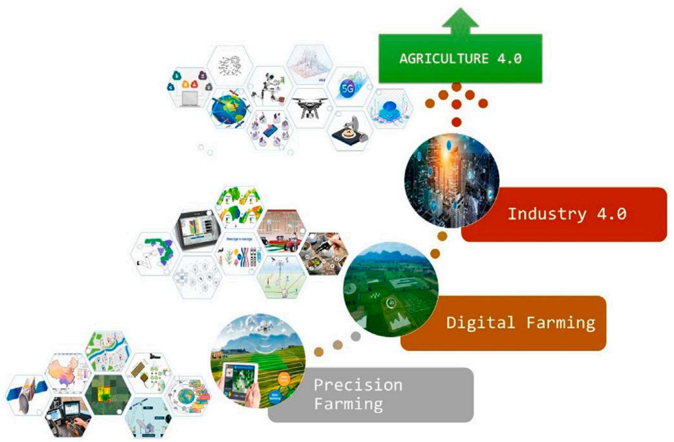

Fig. 2. Evolution of trends related to precision agriculture. Source: Karunathilake et al. [24].

Fig. 3. Autonomous agricultural vehicle operating in an orchard environment. Source: Washington State University [54].

and vineyard rows using a contrastive loss function. In [52], the authors pioneered a motion planning solution in vineyard environments, leveraging depth maps and convolutional neural networks (CNNs). In [53], Liu et al. present a solution for the autonomous navigation problem, solving the path planning problem using the data captured by a vision sensor. In order to achieve a robust solution, the authors use segmentation to identify the traversable zones of the environment from the heatmap of an RGB-D image.

2.1.3. Crop monitoring

Crop monitoring involves the continuous observation of plant growth, health, and productivity over time, often through remote sensing or autonomous field robots [55]. This task is of high importance in order to ensure food security and sustainable agricultural practices, as well as avoiding speculation in food prices [56,57]. Recent trends apply DL techniques to crop monitoring, enabling more accurate and efficient assessments [58,59]. In [60], the authors present a comprehensive review of remote sensing techniques for crop monitoring, discussing the challenges and opportunities in this field. The review covers various aspects of remote sensing for crop monitoring, including data acquisition, processing, and analysis. The authors also discuss the potential benefits of remote sensing for crop monitoring, such as improved yield prediction, disease detection, and resource management [61].

While primarily focused on assessing crop status, these activities are closely related to localization. Variations in crop height, density, and spectral appearance throughout the growing season introduce substantial temporal changes that can affect localization performance. For example, navigation systems relying on visual or geometric cues may face difficulties when the vegetation canopy evolves, occluding landmarks or altering the perceived structure of the environment [62]. In order to provide a solution for the inspection and monitoring of crops using DL techniques methods such as QSeedNet [63] or PTL-Inception [64] have been proposed. The use of DL techniques has been demonstrated to facilitate more precise and expeditious assessments of crop health and growth. This subject has been extensively examined in recent literature, as evidenced by the numerous reviews conducted on the subject [65–67].

Applying clustering algorithms to these unstructured environments is an effective solution in order to classify distinct regions of interest. In [68], the authors present an online machine learning-based sensor clustering system for efficient and cost-effective environmental monitoring in controlled environment agriculture. In the solution proposed by Swain et al. [59], the authors propose an unsupervised clustering approach for crop and weed identification in agricultural fields using LiDAR data.

In [69], BonnBot is presented, as an autonomous field robot designed for crop monitoring tasks in agricultural environments. The robot is equipped with a variety of sensors, including LiDAR, cameras, and GPS, to enable accurate localization and navigation within the field. BonnBot is capable of performing a range of crop monitoring tasks, including plant height measurement, leaf area index estimation, and disease detection.

2.1.4. Phenotyping

Phenotyping refers to the process of measuring and analyzing observable traits of plants, such as growth rate, stress, yield, and resistance to pests and diseases [70]. This information is crucial for plant breeding programs, as it allows researchers to identify desirable traits and develop new crop varieties that are better suited to specific environments [71–73]. Recent sensing technologies have improved the accuracy and efficiency of phenotyping, enabling high-throughput data collection and analysis [74].

Multi-temporal registration of LiDAR data is utilized in [75] to track plant growth and deformation over several months. By aligning different crop states, this approach enables a precise longitudinal analysis of phenotypic traits. The effectiveness of the method was validated using real-world datasets of maize and tomato, achieving high accuracy in complex registration tasks. The data used for the aforementioned experiments is published under the Pheno4D dataset [76]. Fig. 4 shows an example of the phenotyping process using LiDAR data from the Pheno4D dataset.

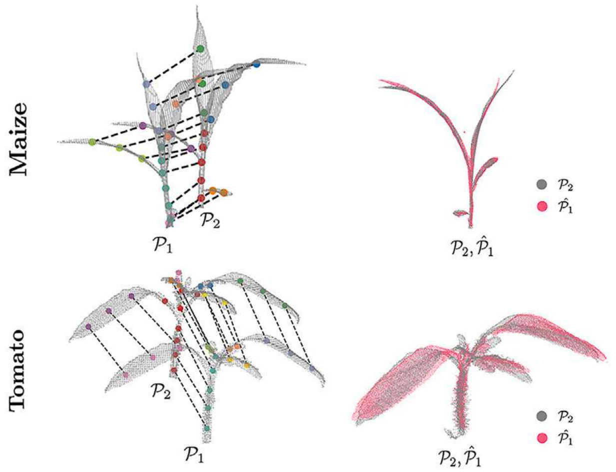

Fig. 4. Example of 4D registration in the 3D phenotyping method proposed in [75] using LiDAR data from the Pheno4D dataset [76].

To facilitate phenotyping tasks in outdoor environments, Zenkl et al. [77] leverage deep learning—specifically a ResNet-based backbone—to segment plants and extract accurate phenotypic traits from images. A similar goal is shared by the Maize-IAS software [78], which is tailored for maize plants and automates the extraction of height, leaf area, and biomass through computer vision.

2.1.5. Pruning and spraying

Pruning and spraying are routine agricultural operations that introduce significant variability in the environment, posing challenges for reliable localization and mapping [38,79]. Pruning modifies the geometry and appearance of vegetation by removing branches and leaves, which alters canopy density and shape. Such structural changes can disrupt the consistency of spatial features used by localization systems, particularly those relying on visual or geometric cues for map matching [80] (Fig. 5).

Spraying operations, on the other hand, affect environmental perception more transiently. The presence of mist, droplets, or wet surfaces can interfere with both optical and LiDAR-based sensing, leading to reflections, occlusions, or altered radiometric responses. Additionally, spraying schedules are often seasonal, introducing recurring but short-lived changes in the sensory data [12,13,82]. In [83] a spraying robot for tobacco fields is proposed. Fig. 6 shows an image of the proposed spraying robot. The robot is equipped with a variety of sensors, including a camera to accurately localize itself within the field and navigate between rows of crops.

3. Localization in agricultural settings

Historically, artificial intelligence (AI) research in agricultural environments has predominantly focused on crop monitoring, segmentation, and classification tasks [11,84]. This trend is evident in studies ranging from subterranean analysis, such as the multi-stacking ensemble learning for soil classification proposed by Padmapriya et al. [85], to aerial assessments like the UAV-based vineyard segmentation studied by Barros et al. [86]. While these works are essential for holistic farm management, they primarily rely on spectral or chemical data to describe “what” is in the field. Consequently, the “where”—specifically the problem of spatial awareness and place recognition—has remained comparatively under-explored.

Fig. 5. Autonomous pruning robot operating in a vineyard environment as proposed by Williams et al. [81].

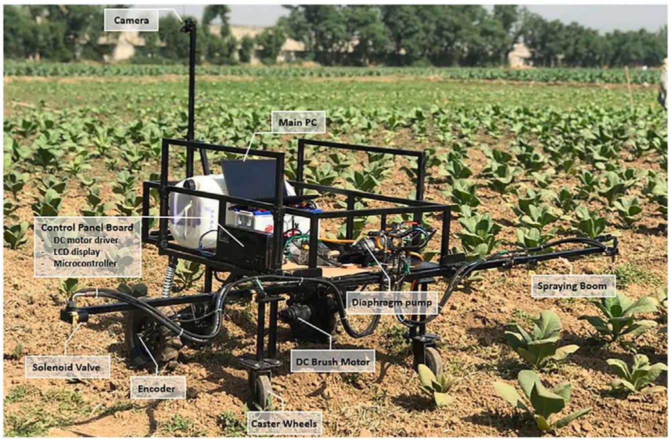

Fig. 6. Autonomous spraying robot operating in a tobacco field proposed by Nasir et al. [83].

In order to specifically address this type of environment in localization problems, several DL solutions have been released targeting agricultural settings. In this section, we review the existing DL solutions for localization problems (place recognition and SLAM) in agricultural environments. We divide the methods into two main categories: vision-based methods and LiDAR-based methods. Furthermore, we also review methods that incorporate semantic information to improve localization performance. Finally, we comment on the specific challenges that arise in agricultural environments for localization tasks.

3.1. Vision-based methods

Focusing on visual descriptors, Tanco et al. [87] developed a self-supervised CNN approach to extract robust features in agricultural terrains. Their method, which includes a modified MobileNet [88] for semantic classification, improves visual SLAM performance by adapting keypoint detection to the specific geometry of vineyards. Experimental results from datasets in Northern Italy underscore the effectiveness of this adaptive feature extraction for long-term navigation.

We find a variety of agricultural SLAM methods as this problem is widely studied as a standard in robotic localization [8,89]. While Tanco et al. rely on deep learning to adapt to crop geometry, other approaches prioritize the integration of geometric primitives to handle repetitive textures. For example, Agri-SLAM [90], a stereo visual SLAM system specifically designed for agricultural environments, was concretely tested on a macadamia field. The method introduces a local point and line feature recovery technique and combines it with inertial measurements to achieve accurate localization in challenging agricultural settings. In [91], the authors evaluate the performance of monocular visual SLAM in dynamic agricultural environments for the first time, as previous studies focus only on stereo cameras. The study highlights a shift away from the “stereo-standard” to address the challenges of moving vegetation and changing lighting with a lighter hardware footprint. However, while Shu et al. propose general strategies to mitigate the ”noise” of moving leaves, mask ORB label optimization SLAM (MOLO-SLAM) [92] introduces a more explicit, computationally intensive solution. This method combines ORBSLAM2 [93] with the Mask-RCNN (region-based convolutional neural network) instance segmentation algorithm to identify and mask dynamic objects in the environment, improving the accuracy of localization and mapping in these challenging settings. This semantic masking represents a much more targeted approach to dynamic environments than the point-line matching seen in contemporary stereo systems like those by Xu et al. [94]. Xu et al. [94] propose a robust visual-inertial SLAM algorithm which is based on point-line features detected on the input stereo images. These visual features are matched to perform a keypoint selection and, subsequently, the pose is estimated using a sliding window algorithm. This approach suggests that, for many agricultural tasks, multi-feature geometric tracking (points and lines) paired with inertial data remains the most reliable compromise between the heavy semantic overhead of MOLO-SLAM and the simplicity of monocular point-tracking.

3.2. LiDAR-based methods

ORCHNet [95] is a DL application that focuses on LPR in orchards. It proposes a new feature aggregation method for global descriptors that can be applied to different backbones. This proposed method for orchard environments has reported more than 90% of Recall@1% on same-season scenarios. However, the Recall@1 value remains lower, with values around 50%. On a similar note, TriLoc-NetVLAD [96] is another DL approach that enhances long-term place recognition in vineyards using a novel LiDAR-based method. The final approach integrates a handcrafted spatial context descriptor, which is derived from the density, height, and intensity of the point cloud, with a channel selection strategy that prioritizes the most successful layers of the backbone. These layers are more robust in facing the unstructured nature of vineyards. This method was tested both against same season and cross season cases, with Recall@1 values of 75% and 55%, respectively. Unlike the purely learned approach of ORCHNet, this hybrid strategy explicitly leverages the geometric “unstructuredness” of vineyards to improve cross-season performance.

The evolution of architectural representations is further seen in the works of Barros et al., who transitioned from voxelization to segment-level reasoning. While SPVSoAP3D [14] focuses on a novel Spherical Projection Voxelization (SPV) with second-order pooling, the authors’ subsequent work, PointNetPGAP-SLC [97], shifts the focus toward environmental structure. By incorporating a Segment-Level Consistency (SLC) model that accounts for row and waypoint data, PointNetPGAP-SLC achieves a superior 75% Recall@1 compared to their earlier voxel-based metrics, but was also only tested on similar seasonal campaigns. PointNet [98] functions as the foundational framework for the extraction of local features from point clouds. These features are then aggregated through the implementation of a novel Pairwise Feature Interactions and Global Average Pooling (PGAP) algorithm, thereby yielding a global descriptor.

A distinct methodological lineage exists in handcrafted descriptors, which often rival DL methods in large-scale settings without the need for training data. The approach introduced by Ou et al. [99] achieves robust place recognition by encoding point clouds into a Spatial Binary Pattern (SBP), followed by an entropy-based attention reweighting strategy. Although this method demonstrates SOTA performance in large-scale orchards and the KITTI dataset, it represents a specific handcrafted lineage within the field. While Ou et al. focus on binary encoding for scalability, the most recent DL method, MinkUNeXt-VINE [100] focuses on achieving robust performance with low-cost, low-resolution inputs. It simplifies the architecture of its parent method, MinkUNeXt [101], and emphasizes encoding repetitive environments with lower-dimensional global descriptors. Additionally, it uses a Matryoshka Representation Learning (MRL) [102] loss function for processing different dimensionalities of the global descriptor and enhancing efficiency. This approach achieves nearly 70% of Recall@1 with a Livox sensor. This represents a significant push toward democratization in agricultural robotics, providing high accuracy with a much lighter computational and financial footprint than the heavy backbones of the ORCHNet era.

In the context of SLAM methodologies, research has diverged between probabilistic filters and structural templates. Aguiar et al. [103] propose VineSLAM, a LiDAR-based SLAM system that is specifically designed for agricultural environments. The method uses point-feature extraction and semiplane segmentation to generate a comprehensive map of the environment. Subsequently, the information is integrated into a particle filter, thereby facilitating precise localization. This “segmentation-first” approach contrasts with the work of Fei and Vougioukas [104], who instead rely on “row templates” to estimate occupancy frequency. These templates are derived from point cloud data, with the associated ground truth values serving as a reference point. These values represent the anticipated measurement values of the sensor (voxel occupancy frequency) in a three-dimensional grid configuration, under the assumption that the sensor is precisely placed at the centerline of the row, with no lateral or heading errors present. Subsequently, the localization process is executed through the use of a conventional Monte Carlo approach. While VineSLAM maps the environment through geometric primitives, Fei’s Monte Carlo approach focuses on the centerline of rows, simplifying the localization problem into a template-matching task.

Bridging the gap between loop closure and odometry, the SG-ISBP framework [105] improves upon the binary patterns seen in Ou’s earlier work by integrating inertial measurement unit (IMU) data and a de-skewed version of the point cloud with ground-assisted odometry values to form an improved spatial binary pattern (ISBP) handcrafted loop closing algorithm. This methodology for closing general loops can be applied to distinct localization applications, such as place recognition or SLAM, in isolation. In contrast, in Purdue-AgSLAM (P-AgSLAM) [106], the authors present LiDAR SLAM system specifically designed for robot pose estimation and agricultural monitoring in maize fields. This SLAM approach is designed to address the unique morphological features of maize fields using LiDAR technology, combining a feature extraction mechanism with an Extended Kalman Filter (EKF) fusion of the input orientation sources (wheel odometry and IMU). The P-AgSLAM system is integrated into the Purdue Ag-Bot (P-AgBot) robotic setup, which is equipped with a monocular camera, two 3D LiDAR (Velodyne VLP-16), a GPS system and an IMU. Where SG-ISBP seeks a generalizable loop-closing pattern, P-AgSLAM provides a robust, platform-specific solution (the P-AgBot) that balances LiDAR features with traditional robotic odometry.

3.3. Semantic information

Methods that incorporate semantic information have been suggested as a way to overcome the challenges posed by agricultural environments. One of such methods is Keypoint Semantic Integration (KSI) [107], which integrates semantic information into the keypoint selection process for LPR in outdoor agricultural environments. KSI enriches keypoints with semantic labels, thereby improving the robustness of place recognition algorithms to environmental changes while accounting for the few distinctive features in these environments, such as pipes, poles, and trunks. The presented results demonstrate improved performance on various descriptors and sequences for place recognition and SLAM metrics. In Papadimitriou et al. [108], the authors propose a method that applies semantic segmentation of grapevines. This method uses a Mask-RCNN DL approach and uses the trunks as a reference for localizing vineyards. The method’s output is combined with a traditional Graph-SLAM localization algorithm to improve localization accuracy.

Other approaches such as [109] have proposed novel particle filter-based localization methods that incorporate semantic information to enhance the robustness of the methods in agricultural environments. By leveraging semantic information, these methods can more effectively address the challenges posed by unstructured and dynamic agricultural settings. This method uses semantic information to enhance the particle filter likelihood estimation process. The proposed approach integrates semantic labels into the localization framework, enhancing the robustness and accuracy of SLAM in agricultural environments where traditional methods may encounter challenges due to the absence of distinctive features and dynamic changes in the environment.

As a summary, Table 1 presents the commented localization methods designed within an agricultural context. This table includes the method name, type of sensor taken as input, year of publication, the approach followed (whether it is a place recognition or a SLAM method), and key features of each approach.

Table 1
Summary of the main localization methods specifically designed for agricultural environments.

Type	Method	Year	Approach	Base architecture
Vision	Shu et al. [91]	2021	SLAM	ORB-SLAM2 [93]
	Papadimitriou et al. [108]	2022	SLAM	Mask R-CNN + Graph-SLAM
	Agri-SLAM [90]	2023	SLAM	Pose graph optimization (PGO)
	Xu et al. [94]	2024	SLAM	Sliding window
	Lv et al. [92]	2024	SLAM	Mask-RCNN
	KSI [107]	2025	Localization	YOLO + SuperGlue (matching)
LiDAR	VineSLAM [103]	2022	SLAM	Particle filter
	Fei and Vougioukas [104]	2022	SLAM	Monte Carlo
	ORCHNet [95]	2023	PR	Global feature aggregation
	Ou et al. [99]	2023	PR	Attention score map + Spatial Binary Pattern
	TriLoc-NetVLAD [96]	2024	PR	CNN + NetVLAD
	SPVSoAP3D [14]	2024	PR	Spherical Projection Voxelization representation and Second Order Average Pooling
	PointNetPGAP-SLC [97]	2024	PR	Pairwise Feature Interactions and Segment-Level Consistency
	SG-ISBP [105]	2024	PR & SLAM	Improved Spatial Binary Pattern
	P-AgSLAM [106]	2024	SLAM	Feature Matching + Extended Kalman Filter
	De et al. [109]	2025	SLAM	Semantics + Particle Filter Likelihood Estimation

Table 2
Summary of the results of the mentioned SLAM and place recognition methods specifically designed for agricultural environments.

Type	Method	Year	Dataset	Recall@1	Recall@1%	RMSE [cm]
Vision	Agri-SLAM [90]	2023	ROSARIO (06) [110]	–	–	1.62
	Xu et al. [94]	2024	Banana (inhouse)	–	–	9.1
	Lv et al. [92]	2024	Sequence2 (inhouse)	–	–	1.2
	KSI [107]	2025	BLT [111]	0.99	–	–
LiDAR	VineSLAM [103]	2022	Sequence2 (inhouse)	–	–	44
	Ou et al. [99]	2023	Orchard3 (inhouse)	–	–	18
	ORCHNet [95]	2023	HORTO-3DLM [14]	0.52	0.92	–
	TriLoc-NetVLAD (same season) [96]	2024	Vineyard (inhouse)	0.75	–	–
	TriLoc-NetVLAD (cross season) [96]	2024	Vineyard (inhouse)	0.55	–	–
	SPVSoAP3D [14]	2024	HORTO-3DLM [14]	0.63	0.96	–
	PointNetPGAP-SLC [97]	2024	HORTO-3DLM [14]	0.75	–	–
	P-AgSLAM [106]	2024	real1 (inhouse)	–	–	26
	SG-ISBP-SLAM [105]	2024	Forest2 (inhouse)	–	–	18
	MinkUNeXt-VINE (same season) [100]	2026	TEMPO-VINE [112]	0.68	0.97	–
	MinkUNeXt-VINE (cross season) [100]	2026	TEMPO-VINE [112]	0.41	0.70	–

To provide an overview of the reported results in these environments, Table 2 summarizes the results of the mentioned SLAM and place recognition methods specifically designed for agricultural environments that are discussed in this section. The table includes the method name, year of publication, dataset used for evaluation, and the reported metrics such as Recall@1, Recall@1%, and Root Mean Square Error (RMSE) in centimeters. For brevity, only best results per method are listed, omitting redundant data (e.g., KSI [107]). This summary enables a quick comparison of the performance of different methods in agricultural settings. As can be observed, the reported results vary significantly across methods and datasets, highlighting the challenges of localization in agricultural environments and the need for standardized evaluation protocols. While the proposed vision approaches reach RMSE values around 1 cm, the LiDAR methods report higher RMSE values, with the best-performing method reaching 18 cm (Ou et al. [99]). This shows that more work is needed to adapt current LiDAR-based methods to the specific challenges of agricultural environments, such as their unstructured nature. Additionally, it can be observed that the reported Recall@1 values for place recognition methods are generally below 80%, indicating that there is still room for improvement in terms of robustness and generalization of these methods in this context. For this reason, this survey focuses on the current state of the LiDAR place recognition problem in these types of scenarios. In Section 6, we will provide a detailed analysis of these metrics and, in Section 5 we will provide more insights on the public localization datasets used for evaluation. In the next Section 4, we will review the different approaches to the LPR problem in urban environments. This is the problem that has been studied the most in the literature, and we will compare it to unstructured settings.

4. The LiDAR place recognition problem

LPR constitutes a crucial component of long-term localization and mapping systems in autonomous robotics. The objective is to determine whether a 3D LiDAR scan corresponds to a previously visited location, enabling loop closure detection and global position correction. Compared to vision-based approaches, LiDAR offers high geometric accuracy and invariance to lighting and appearance changes, which is essential for operation under varying environmental conditions. Nevertheless, the unstructured nature of point clouds and the presence of dynamic or seasonal changes make the extraction of robust and distinctive representations a challenging task. Recent advances in DL have significantly improved LiDAR-based place recognition by learning compact and invariant descriptors capable of generalizing across different scenes and sensor configurations [113–115].

Fig. 7 shows a typical LPR pipeline. The input point cloud can first be preprocessed to remove noise and outliers or to select a region of interest, and then a feature extraction network is used to extract local features from the point cloud. These local features are then aggregated into a global descriptor using a feature aggregation method. Finally, the global descriptor is compared to a database of previously stored descriptors to recognize the place. This resulting descriptor can then be used for localization within a pre-built map of the environment or for other applications, such as clustering the environment.

In this section, we will first focus on urban environments and review the various approaches to the LPR problem in this context, as it was originally studied here. First, we review the various DL methods proposed since the initial mention of LPR in the literature. Second, we will provide an in-depth commentary on the existing geometry-based approaches in the literature. Then, we discuss the impact of preprocessing LiDAR scans on the performance of LPR approaches. Next, we discuss the newest cross-modal approaches that aim to improve LPR robustness by extracting information from various sensors. Finally, we analyze the specific challenges that arise in agricultural environments for LPR tasks, making a qualitative comparison with urban environments.

4.1. LiDAR deep learning methods

The objective of using trainable methods to find a solution to the problem is to learn the embeddings of the input data, in this case, point clouds, into a lower-dimensional space where similar places are closer together. This is typically achieved using deep neural networks, which can learn complex representations of the input data. Deep learning methods can be supervised or unsupervised. Supervised methods require a labeled dataset, where each point cloud is associated with a specific location. The network is trained to minimize the distance between the embeddings of point clouds from the same location and to maximize the distance between embeddings from different locations. Unsupervised methods, on the other hand, do not require labeled data and instead rely on techniques such as clustering or autoencoders to learn the embeddings.

PointNetVLAD [116] remains a cornerstone of DL-based LPR, establishing the paradigm of combining PointNet [98] for feature extraction with a NetVLAD layer [117] for aggregation. While PointNetVLAD effectively handles unordered point clouds via global descriptors and triplet loss, it often struggles with fine-grained local geometry—a limitation addressed by LPD-Net [118]. Unlike the purely global approach of PointNetVLAD, LPD-Net introduces a graph-based aggregation method to better capture the spatial distribution of local features.

A distinct methodological departure is found in OverlapNet [119], which moves away from traditional global descriptors in favor of a Siamese architecture. Instead of seeking a single embedding, it directly predicts the overlap between point cloud pairs, offering a more geometric perspective on similarity. However, the industry has seen a significant shift toward efficiency through sparse 3D convolutions.

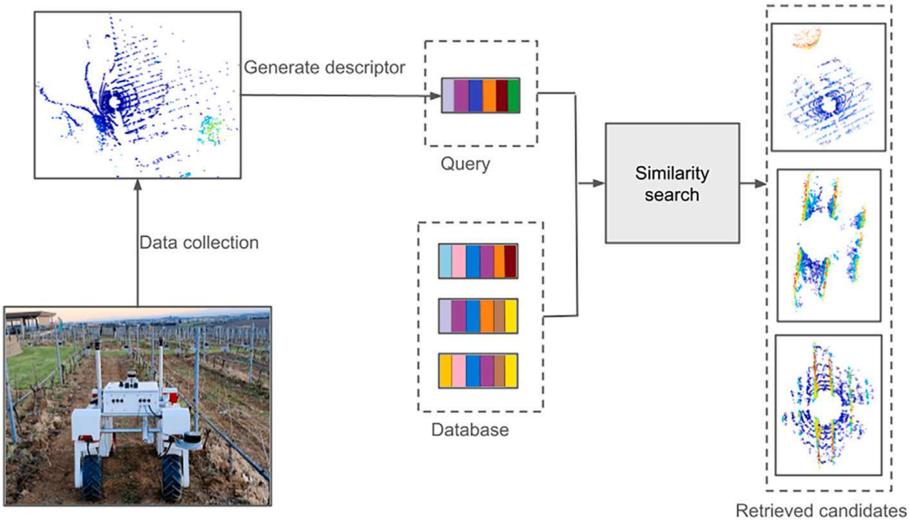

Fig. 7. Diagram of a typical LPR pipeline. For each point cloud, a global descriptor is generated. This descriptor can then be used for place recognition by comparing it to a database of previously stored descriptors.

MinkLoc3D [120] exemplifies this trend, offering high accuracy at a lower computational cost compared to PointNet-based architectures.

The evolution of the MinkLoc family further highlights the importance of training objectives. While the original MinkLoc3D relied on triplet loss, MinkLoc3Dv2 [121] introduced structural modifications and a ranking-based loss function, yielding a 3.4% improvement in recall. This transition from triplet-based to ranking-based optimization appears to be a key driver in recent performance gains. This is further evidenced by MinkUNeXt [101], which integrates a UNet-like structure with ranking-based loss to push Recall@1 results above 90% in complex urban environments, representing the current state-of-the-art compared to the 80% baseline established by PointNetVLAD.

Recent approaches have also studied the use of additional information, such as intensity or color, to improve the performance of LPR. For example, the variants MinkLoc3D-SI [122] and MinkUNeXt-SI [123] have also been proven to present superior results to their original backbones by using the spherical projection of the points and including the intensity information, with recall improvements of around 5%. While these SI-variants optimize raw sensor data, Cgis-Net [124] takes a more heterogeneous approach by fusing color and geometric features directly with semantic labels to learn its embeddings.

Beyond raw attribute fusion, a sub-field has emerged that prioritizes environmental “meaning” through semantic segmentation. SegMap [125] is a method that uses a 3D convolutional neural network to learn the embeddings of segmented point clouds. The network is trained to predict the semantic labels of the segments, which can be used for place recognition. In contrast to SegMap’s 3D approach, Coral [126] achieves semantic enrichment through a 2D projection, using the bird’s eye view (BEV) image fused with vision information. Bridging the gap between these methods, Locus [127] provides a more holistic descriptor by explicitly aggregating high-level semantic features with low-level geometric ones, ensuring robustness in scenes where semantic labels alone might be ambiguous.

While the aforementioned methods rely on static, pre-trained models, a new frontier of LPR focuses on the reality of long-term deployment. Novel approaches consider incremental learning techniques to adapt the model to new environments or conditions without forgetting previously learned information. This marks a departure from the “fixed-weights” paradigm of MinkLoc or PointNetVLAD. For example, InCloud [128] is an incremental learning method that uses a memory buffer to store previously seen data. The model is updated with new data while retaining the knowledge from the memory buffer.

4.2. LiDAR handcrafted methods

Handcrafted methods for LPR typically involve the extraction of geometric features from point clouds, which are then used to create descriptors that can be compared to recognize places [129]. Unlike the “black-box” nature of the DL methods discussed in Section 4.1, these approaches rely on geometric primitives and explicit rules for feature extraction. There are different types of handcrafted descriptors, but the most prominent ones are signature-based and histogram-based methods.

Firstly, signature-based methods are characterized by encoding the prominent features or repetitive structures of the point cloud in a signature. Specifically, signature-based methods compress the spatial information of a point cloud into distinctive descriptors, often taking the form of range images, polar projections, or BEV representations. By doing so, they transform unstructured 3D data into a structured format that facilitates efficient matching. The most popular signature-based approach is Scan Context [130], which creates a global descriptor by projecting the point cloud onto a 2D polar grid and encoding the height information. While the original Scan Context provides a strong structural baseline, it remains sensitive to the sensor’s modality and viewpoint. This led to the development of Intensity Scan Context [131], which addresses the limitations of purely geometric data by incorporating reflectance information for greater robustness. Similarly, Scan Context++ [132] further evolves this lineage by introducing lateral and rotational invariance, specifically targeting the viewpoint-sensitivity that often causes standard Scan Context to fail in reverse-loop scenarios. The results of applying Scan Context to agricultural datasets are presented in the TEMPO-VINE dataset work [112]. Martini et al. demonstrated that, although promising performance can be achieved, performance significantly decreases when seasonal changes are considered, dropping recall values from 90% to 40%.

In contrast to the structural “signatures” of Scan Context, histogram-based methods prioritize statistical distributions of local or global measurements, such as distances, normals, curvatures or intensities. Instances of this type of method include the Fast Point Feature Histogram (FPFH) [133], which computes histograms of local geometric features around each point in the point cloud. The histograms are then aggregated into a global descriptor that can be used for place recognition. While FPFH excels at capturing fine-grained local topology, it can struggle in sparse or repetitive environments—a challenge mitigated by the Ensemble of Shape Functions (ESF) [134], which shifts the focus toward global shape primitives. These shape functions are then integrated into a global descriptor that can be used for place recognition. In recent works such as Omni Point [135], the authors propose a novel histogram-based descriptor that captures the omni-directional geometry of the point cloud. The method computes histograms of local geometric features in multiple directions and aggregates them into a global descriptor that can be used for place recognition. By computing multi-directional histograms of local features, Omni Point achieves a more comprehensive omni-directional geometry than FPFH or ESF, which can be particularly beneficial in complex environments with varying viewpoints.

4.3. Point cloud preprocessing

Data preprocessing is a crucial step in LPR, as it can significantly impact the performance of the algorithms. Preprocessing techniques can include filtering, downsampling, and normalization of the point cloud data.

While basic filtering techniques remove sensor noise and outliers, more strategic filtering focuses on "feature saliency" by removing non-discriminative data. For instance, removing ground points or points outside a specific radius [136] is a common practice to focus the descriptor on vertical structures like building facades or poles, which are more stable for recognition than the ground plane. This contrasts with earlier handcrafted methods that often required the ground for orientation alignment, whereas modern DL backbones often find ground points redundant or even detrimental to the triplet loss convergence. Downsampling techniques can be used to reduce the size of the point cloud, which can improve the computational efficiency of the algorithms and also reduce the memory footprint. This is particularly important for real-time applications, where computational resources are limited. The choice of downsampling, such as voxel grid filtering versus random sampling, often depends on the backbone architecture; for example, voxel grids are preferred for sparse convolution methods like MinkLoc3D [120] to maintain a consistent spatial resolution.

Normalization techniques can be used to standardize the point cloud data, which can improve the robustness of the algorithms to changes in scale and orientation. This is particularly vital for methods that lack inherent rotational invariance, acting as a manual substitute for the learned invariance found in architectures like Scan Context++ [132].

Another preprocessing technique that has been proven successful in LPR is the use of other information present in the LiDAR data, mainly the intensity value of each coordinate. The intensity value represents the intensity of the reflected laser beam and can provide additional information about the surface properties of the scanned objects. While traditional geometry-only methods like PointNetVLAD ignore these values, integrating intensity during the preprocessing stage has been the primary driver for the performance jumps seen in "SI" (Spherical Representation + Intensity) variants. Several works have shown that including this information in the input data can improve the performance of LPR algorithms [122,123], effectively bridging the gap between purely geometric 3D processing and multi-modal scene understanding.

4.4. Cross-modal approaches

Cross-modal approaches for LPR involve the integration of data from multiple sensors to improve the accuracy and robustness of place recognition algorithms. These approaches can leverage the complementary information provided by different sensors to create more robust descriptors for place recognition. Depending on the point of the network where the features from both modalities are fused, we can identify three different fusion strategies: early fusion [126,137], mid fusion [138] and late fusion [139]. In early fusion, the raw data from both sensors are combined before feature extraction. In mid fusion (which is also referred to as feature-level fusion), features are extracted separately from each modality and then combined at a later stage in the network. Finally, in late fusion, the features from each modality are processed independently and then combined at the decision level.

Visual place recognition has the disadvantage of being affected by changes in lighting and appearance, which can significantly degrade the performance of place recognition algorithms. On the other hand, LiDAR-based place recognition is more robust to these changes, but it can be affected by changes in the environment, such as seasonal changes or dynamic objects. By combining data from both sensors, cross-modal approaches can leverage the strengths of each modality to create more robust descriptors for place recognition. However, MinkLoc++ treats both modalities as equally important, a static approach that can be suboptimal in extreme conditions (e.g., total darkness). To address this, AdaFusion [140] introduces an adaptive weighting strategy. Unlike the fixed concatenation of MinkLoc++, AdaFusion dynamically adjusts the contribution of visual vs. LiDAR features based on the environmental context. Taking this interaction a step further, BdFusion [141] utilizes a bi-directional attention mechanism. Whereas AdaFusion weights the outputs, BdFusion's mid-fusion strategy allows the two modalities to actively "inform" one another during the feature extraction process, resulting in a more deeply integrated descriptor.

A significant shift in recent research is the move beyond traditional visual sensors toward unconventional modality pairings. One such method is MambaPlace [142], which combines text and point-cloud information for place recognition. The method uses a coarse-to-fine approach. First, the text input is encoded with attention mechanisms and MLP layers to create a global semantic descriptor. Then, the point cloud data is processed with a dual-path forward scanning approach that considers spatial relationships. In the second, fine stage, features from both modalities are fused with a cascaded cross-attention module to create a more robust descriptor for place recognition. By encoding global semantic text descriptors, MambaPlace provides a high-level context that geometric sensors alone often lack. On another note, RADAR and LiDAR fusion has also been explored for place recognition. For example, LRFusionPR [143] is a method that combines RADAR and LiDAR data for place recognition. It uses cross attention mechanism to fuse the features from both sensors, which were previously both projected to a polar unified BEV representation. While MambaPlace focuses on semantic enrichment through language, LRFusionPR focuses on geometric redundancy, ensuring localization in conditions where even the most advanced vision-based systems would be blinded.

4.5. Specific challenges in agricultural environments

We have now reviewed the SOTA in generalistic place recognition approaches. These approaches have made significant progress with urban datasets. However, while urban environments are characterized by distinct and stable features such as buildings and roads, agricultural environments present unique challenges due to their unstructured and dynamic nature. Table 3 summarizes the main differences between urban and agricultural environments. These differences pose significant challenges for LPR algorithms, which must be able to adapt to the specific characteristics of agricultural settings.

4.5.1. Lack of distinctive features

Agricultural environments often lack distinctive features that can be easily recognized by place recognition algorithms. For example, in a

Table 3
Comparison between urban and agricultural environments for LPR.

Aspect	Urban Environments	Agricultural Environments
Distinctive features	Presence of buildings, roads	Lack of distinctive features; similar-looking plants
Repetitive patterns	Less common; unique structures	Common; rows of crops or trees
Seasonal changes	Less pronounced; stable structures	Significant changes; vegetation growth and decay
Environmental variability	More structured and predictable	Unstructured and dynamic

Fig. 8. Example of lack of distinctive features in an orchard environment. The plants look identical and lack unique characteristics that can be used for recognition. Image taken from the MAgro dataset [15].

Fig. 9. Example of repetitive patterns in a vineyard environment. The rows create a repetitive pattern that can be challenging for place recognition algorithms to accurately identify locations. Image obtained from the BLT dataset [111].

field of crops, the plants may look similar and lack unique characteristics that can be used for recognition. This can lead to confusion and misidentification of locations (Fig. 8).

4.5.2. Repetitive patterns

Agricultural environments often contain repetitive patterns, such as rows of crops or trees, which can make it difficult for place recognition algorithms to distinguish between different locations. This can lead to false positives, where the algorithm incorrectly identifies a location as being the same as another location.

Fig. 9 shows an example of repetitive patterns in an orchard environment. The rows of trees create a repetitive pattern that can be challenging for place recognition algorithms to accurately identify locations.

4.5.3. Seasonal changes

Agricultural environments are subject to seasonal changes, which can significantly alter the appearance of the environment. For example, a field of crops may look very different in the summer compared to its winter, when it loses all of the leaves. This can make it challenging for place recognition algorithms to accurately identify locations, as the features used for recognition may no longer be present. Even to an external observer, the place may look completely different depending on the season. Fig. 10 shows an example of seasonal changes in a vineyard environment (the TEMPO-VINE dataset [112]). It illustrates how the same location can appear drastically different between winter and summer seasons, posing a significant challenge for LPR systems.

4.5.4. Irregular terrain

In contrast to urban environments, agricultural settings are characterized by irregular terrain, a factor that can introduce navigational challenges for robotic systems. The presence of variations in elevation, uneven ground surfaces, and the existence of natural obstacles, such as rocks or ditches, can collectively contribute to a more complex scenario for the implementation of path planning applications. The odometry of robots operating in such environments, where the terrain may be slippery or unstable, can be adversely affected, leading to inaccuracies in position estimation. These irregularities can result in deviations from the intended path, thereby complicating the localization process.

4.5.5. Geo-referencing difficulties

Agricultural environments frequently present challenges for geo-referencing due to the absence of reliable GPS signals. The obstruction of satellite signals by factors such as dense vegetation or the presence of large structures, such as barns or silos, can be attributed to multipath effects. Consequently, obtaining a reliable ground truth for localization can be challenging in such environments.

4.6. Urban vs. agricultural environments for LPR

In order to address the aforementioned challenges in the design of LPR algorithms for agricultural environments, it is crucial to understand the specific differences between urban and agricultural settings. These differences will affect how the methods extract significant features from the environment and how they evaluate the performance of the algorithms.

Traditionally, the evaluation of LPR algorithms is conducted through a retrieval-based approach, where the algorithm is assessed based on its ability to retrieve the correct place from a database of previously stored descriptors. Different distance thresholds can be used to determine whether a retrieved place is considered a correct match or not. While in urban environments, a common threshold is to consider a match as correct if the retrieved place is within 25 meters of the ground truth location [101,116], in agricultural environments, where the scale of the environment can be smaller, a threshold of 10 meters is more commonly used [14,97]. However, in row-crop environments, a more dynamic approach should be employed, where the threshold is determined based on the distance between the rows.

The differences between urban and agricultural environments also affect the design of the algorithms. In urban environments, the presence of distinctive features such as buildings and roads can be leveraged to create more robust descriptors for place recognition. In contrast, in agricultural environments, where there is a lack of distinctive features and the presence of repetitive patterns, algorithms need to be designed to extract more subtle features from the environment, such as the shape of the plants or the texture of the soil. Additionally, algorithms need to be designed to be robust to seasonal changes, which can significantly alter the appearance of the environment.

In summary, while LPR algorithms have made significant progress in urban environments, the unique challenges posed by agricultural settings require a different approach to both the design of the algorithms and the evaluation of their performance. Understanding these differences is crucial for developing effective LPR solutions for agricultural applications. Fig. 11 illustrates the performance gap between urban and agricultural environments for methods presented in both Sections 3 and 4, showing how the same algorithm can achieve high recall values in urban datasets while struggling to achieve similar performance in agricultural datasets. In the urban methods, the best performing algorithms achieve recall values above 90%, while in agricultural environments, the best performing algorithms struggle to achieve recall values above 75%. This performance gap highlights the need for further research and development of LPR algorithms specifically designed for agricultural environments. Furthermore, it can be seen that the agricultural-specific methods are much more recent, with the first method of the graph being

(a) Winter season

(b) Summer season

Fig. 10. Example of seasonal changes from a same location in a pergola-type vineyard environment extracted from the TEMPO-VINE dataset [112]. (a) Winter season with bare vines. (b) Summer season with abundant vegetation covering the robot.

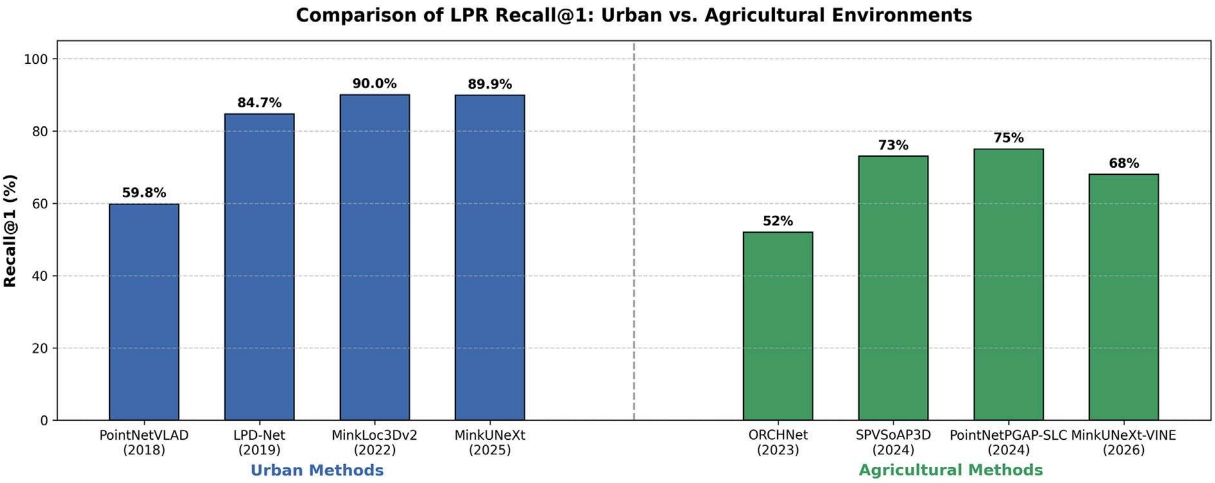

Fig. 11. Comparison between Recall@1 values obtained for SOTA LPR methods in generalist or agriculture-specific approaches.

published in 2024, while the urban methods have been published since 2018. This also highlights the fact that LPR in agricultural environments is a relatively new research area that is still in its early stages of development.

As can be seen from Fig. 11, while urban methods have been tested on large datasets and with standardized evaluation protocols, the need for agricultural-specific, large-scale and long-term datasets is evident. The following section will review existing datasets that enable localization tasks in agricultural environments. We will analyze these datasets' characteristics and compare them with agricultural datasets focused on vision applications, such as detection, segmentation, and fruit counting.

5. Datasets

A variety of datasets have been used to evaluate the performance of LPR algorithms. However, most of these datasets are collected in urban environments, which may not be representative of the challenges faced in agricultural settings. In this section, we analyze some of the most commonly used datasets for LPR and discuss their relevance to agricultural environments.

5.1. Transitional datasets

Although many localization datasets are collected in urban environments, some of these environments include vegetation-rich sections. The KITTI dataset [144] is one of the most widely used datasets for LPR. It includes a variety of urban environments, including some sections with vegetation, such as parks and tree-lined streets. However, the majority of the dataset is collected in urban settings, which may not be representative of the challenges faced in agricultural environments. While a variety of different trajectories has been captured within this dataset, it does not contain long-term information, as all the sessions were collected within a scope of 1 month. Another popular dataset is the Oxford Robotcar Dataset [145], which includes a variety of urban environments, including some sections with vegetation. The total extent of the data captured by this dataset is 10 km. Although it contains long-term information as the sessions were recorded over a year, the vegetation sections are limited and may not be representative of agricultural settings.

Several datasets capture off-road environments with the presence of natural elements. The ORFD dataset [146] is a dataset collected in a mixed urban and off-road (forest) environment. It includes a variety of vegetation types, such as trees, bushes, and grass. The dataset contains long-term information, as the sessions were recorded covering seasons from spring to winter. However, the extent of the recorded data in each session is very limited and the off-road sections are restricted and may not be representative of agricultural settings. The Wild-Places dataset [147] is a dataset proposed for LPR and SLAM specifically, and was collected in a variety of outdoor environments, including forests, parks, and rural areas. The dataset contains long-term information, as the sessions were recorded over a year. However, the dataset does not include agricultural environments specifically. Finally, the Rellis-3D dataset [148] is a dataset collected in rural environments, including fields and forests. The dataset contains a variety of vegetation types, including crops and trees. However, the dataset does not include long-term information, as all the sessions were collected within a short period of time. In these datasets, both 3D information from terrestrial LiDAR sensors and accurate ground truth annotations are provided, which are essential for developing and evaluating LPR algorithms.

5.2. Datasets for agricultural environments

Several datasets have been collected in agricultural environments to address the unique challenges posed by these settings. Table 4 summarizes the main characteristics of these datasets, including the year of publication, the crop type, the sensor used, the number of sequences, and whether they include long-term variations.

The majority of agricultural datasets focus on collecting visual information with RGB cameras. However, there has been a growing interest in collecting LiDAR data in agricultural settings in recent years. The GREENBOT dataset [159] is one of the first datasets to include LiDAR data collected in a tomato greenhouse environment. The MAgro dataset [15] includes LiDAR data collected in apple and pear orchards, while the HORTO-3DLM dataset [14] includes data from multiple crop types, including apple, strawberry, cherry, and tomato and it was collected using a different sensor setup in each sequence and location. The BLT dataset [111] focuses on grapevine environments, while the ROSARIO dataset [110] includes soybean fields. The ARD-VO dataset [160] also focuses on vineyards. Finally, the TEMPO-VINE dataset [112] includes long-term LiDAR data collected in vineyards using both Velodyne VLP-16 and Livox Mid360 sensors, providing heterogeneous LiDAR information.

If we compare these datasets with the most popular urban datasets, such as the Ford Campus dataset [161] or the University of Michigan's North Campus Long-Term dataset (NCLT) [162], we can notice that the agricultural datasets are still limited in terms of the number of sequences and the variety of environments. There is a wider niche of urban datasets that ease the task of training and evaluating LPR algorithms. However, the unique challenges posed by agricultural environments require specialized datasets that can effectively capture the unstructured nature of these settings.

It is also important to take into account that the presence of the necessary information for developing LPR applications is a more recent approach, as there are many agricultural environments that have not yet been utilized for this precise purpose. The most popular approach is to publish data collected from vision sensors, which facilitates classification algorithms, useful for different tasks such as disease identification [163] or fruit detection [164]. This trend has been explored since the mid 2010s decade, with published data such as [150], where the authors propose data taken from different vision systems in different orchards (apple, mango and almond plantations) and they use this data in their proposed Faster RCNN DL detection method. In a similar manner, Liu et al. [151] propose an oranges and apples monocular images dataset for fruit counting through their pipeline, and Stein et al. [149] publish data taken in a mango orchard for fruit detection. CropDeep [152] is an annotated dataset created specifically for classification and detection in agricultural environments collected using mobile cameras, autonomous robots and IoT cameras.

More recent works such as the Low-Light dataset [90] present data collected in macadamia fields, but it does not contain either LiDAR or GPS information. The GrapeNet dataset [156] provides RGB and RGB-D information in a vineyard environment in order to support vision applications. The MinneApple dataset [153] provides annotated visual information taken in an apple orchard, which also supports DL semantic and classification algorithms. The use of UAV LiDAR has also been an extended practice when it comes to collecting data in agricultural fields [165]. The VineLiDAR dataset [166] provides high-resolution UAV-LiDAR data collected over two years in northern Spain. Alessandrini

Table 4
Summary of available agricultural environment datasets.

Task	Year	Dataset	Crop type	LiDAR	Vision	GPS	Sequences	Long-term
Detect. Segment. Count.	2016	Stein et al. [149]	Mango	Velodyne HDL64E	Prosilica GT3300c Kowa LM8CX	Yes	3	No
	2017	Bargoti et al. [150]	Apple Mango Almond	–	PointGrey LadyBug Prosilica GT3300c Canon EOS60D	No	3	No
	2018	Liu et al. [151]	Apple Orange	–	Samsung Galaxy S4	No	2	No
	2019	CropDeep [152]	Diverse (greenhouse)	–	IoT camera Autonomous robots Smartphone	No	Unspecified	No
	2020	MinneApple [153]	Apple	–	Samsung Galaxy S4	No	Unspecified	Yes
	2021	Alessandrini et al. [154]	Grapevine	–	iPadPro Samsung J7 iPhone 8	No	Unspecified	No
	2021	Abdelghafour et al. [155]	Grapevine	–	Basler Ace (acA2500-14gc GigE)	No	Unspecified	No
	2023	Low-Light [90]	Macadamia	–	Zed Stereo	No	1	No
	2023	GrapeNet [156]	Grapevine	–	One Plus 7 Mobile Intel Real-sense D435I	No	18	Yes
	2025	Mohammed et al. [157]	Pomegranate	–	Sony ILCE-7RM4	No	1	No
Localiz.	2016	Sugarbeets [158]	Sugarbeet	Velodyne VLP16 Puck	JAI AD-130GE Kinect One	Yes	30	No
	2022	GREENBOT [159]	Tomato	Velodyne VLP-16	Bumblebee BB2-08S2	Yes	9	No
	2022–2023	MAgro [15]	Apple; Pear	Velodyne Puck 3D	StereoLabs Zed2 (2)	Yes	9	Yes
	2022–2023	HORTO-3DLM [14]	Apple; Strawberry; Cherry; Tomato	Velodyne VLP-16	–	Yes	6	No
	2022–2023	BLT [111]	Grapevine	Ouster OS1-16	StereoLabs Zed2	Yes	15	Yes
	2023	ROSARIO [110]	Soybean	Velodyne VLP-16	Intel Realsense D435i	Yes	6	Yes
	2023	ARD-VO [160]	Vineyards	Velodyne VLP-16	RedEdge MX Blackfly S	Yes	11	Yes
	2025	TEMPO-VINE [112]	Grapevine	Velodyne VLP-16 Livox Mid360	RGB-D Intel Realsense D435	Yes	11	Yes

et al. [154] propose a vineyard dataset with RGB-D data captured using smartphone devices for the detection of esca disease in leaves with DL applications. Abdelghafour et al. [155] also provide an annotated imagery dataset taken in a vineyard environment that facilitates the task of identifying diseases in the Merlot grape variety. In their more recent work, Mohammed et al. [157] presented a public dataset comprising pomegranate images. These images, which depict both healthy and unhealthy crops, are designed to support the development of advanced learning applications focused on disease detection. Further efforts are needed to collect and curate datasets that can support the development of robust LPR algorithms for agricultural applications, providing data from terrestrial LiDAR sensors and accurate ground truth annotations.

6. Evaluation metrics

The performance of LPR algorithms is evaluated by measuring the ability of the algorithm to correctly recognize previously visited places and to accurately localize the robot within a pre-built map of the environment. In this section, an examination of the most commonly used metrics for evaluating the global descriptors produced by LPR algorithms in both SLAM and localization tasks is conducted. Furthermore, the computational latency and the memory requirements of the algorithms are also analyzed, as they are crucial factors for real-time applications.

6.1. SLAM

Localization performance is primarily evaluated using the Relative Pose Error (RPE) for local accuracy and Absolute Pose Error (APE) for global consistency [167]. As defined in Eq. (1), RPE averages the drift over short intervals, while APE compares the full trajectory against ground truth (T_{gt}):

$$\mathrm{RPE} = \frac{1}{n}\sum_{i=1}^{n}\left\|\left(T_{est}^{-1}T_{gt}\right)_i\right\|, \quad \mathrm{APE} = \frac{1}{n}\sum_{i=1}^{n}\left\|T_{est} - T_{gt}\right\| \tag{1}$$

To provide statistical depth, recent studies [168–170] advocate for the Root Mean Square Error (RMSE), defined as RMSE $= \sqrt{\frac{1}{n}\sum(y_i - \hat{y}_i)^2}$, which is sensitive to large errors [171,172], and the Mean Absolute Error (MAE $= \frac{1}{n}\sum|y_i - \hat{y}_i|$), which treats errors uniformly.

Geometric accuracy is further dissected into translational and rotational components. The Mean Translation Error (MTE) and L2 norm quantify Euclidean distance deviations ($\|p_{est} - p_{gt}\|$), often reported alongside the Angular Error (AE) to assess orientation drift [173]. The AE is crucial for visual-inertial systems and is computed as:

$$\mathrm{AE} = \arccos\left(\frac{tr(R_{est}R_{gt}^{T}) - 1}{2}\right) \tag{2}$$

where R_{est} and R_{gt} are the estimated and ground truth rotation matrices, respectively.

6.2. Localization

The performance of LPR algorithms is typically evaluated using metrics such as precision, recall, and F1-score [174]. Precision measures the proportion of correctly recognized plates among all recognized plates, while recall measures the proportion of correctly recognized plates among all actual plates. Eqs. (3) and (4) present the formulas for the precision and the recall, respectively.

$$\mathrm{P} = \frac{TP}{TP + FP} \tag{3}$$

$$\mathrm{R} = \frac{TP}{TP + FN} \tag{4}$$

Where TP is the number of true positives, FP is the number of false positives, and FN is the number of false negatives.

The F1-score is the harmonic mean of precision and recall, providing a single metric that balances both aspects. Another commonly used metric is the area under the precision-recall curve (AUC-PR), which summarizes the trade-off between precision and recall at different thresholds. Additionally, some studies also report the average recall at top-K retrievals, which measures the proportion of correctly recognized places among the top K retrieved places [175–177]. Eq. (5) shows the formula for the AUC-PR.

$$\mathrm{AUC} = \int_0^1 \mathrm{P}(r)dr \tag{5}$$

where P(r) is the precision at recall r.

Furthermore, both the F1 and F2 scores are also used in some works to evaluate the performance of LPR algorithms. The F1-score gives equal weight to precision and recall, while the F2-score gives more weight to recall. Eqs. (6) and (7) show the formulas for the F1 and F2 scores, respectively.

$$\mathrm{F1} = 2 \cdot \frac{\mathrm{P} \cdot \mathrm{R}}{\mathrm{P} + \mathrm{R}} \tag{6}$$

$$\mathrm{F2} = 5 \cdot \frac{\mathrm{P} \cdot \mathrm{R}}{4 \cdot \mathrm{P} + \mathrm{R}} \tag{7}$$

Finally, Max F1-Score is also used in recent works [178,179] to evaluate the performance of LPR algorithms. While a standard F1-Score is often reported at a fixed similarity threshold, the Max F1-Score is determined by identifying the peak value across the entire Precision-Recall curve. This provides a threshold-independent summary of the best possible balance between precision and recall, which is particularly vital for loop closure detection where preventing false positives is as important as identifying matches. Eq. (8) shows the formula for the Max F1-Score.

$$\text{Max F1} = \max_{\tau}\left(2 \cdot \frac{\mathrm{P}(\tau) \cdot \mathrm{R}(\tau)}{\mathrm{P}(\tau) + \mathrm{R}(\tau)}\right) \tag{8}$$

Where τ is the threshold used to determine whether a place is recognized or not, and $P(\tau)$ and $R(\tau)$ are the precision and recall at that threshold, respectively.

6.3. Clustering

Global descriptors generated by LPR algorithms can be used for clustering. The performance on this application is typically evaluated using metrics such as the silhouette score [180,181] and the Davies-Bouldin index [182]. The silhouette score measures the similarity of each point to its own cluster compared to other clusters, while the Davies-Bouldin index measures the average similarity between each cluster and its most similar cluster. Eq. (9) shows the formula for the silhouette score, while Eq. (10) shows the formula for the Davies-Bouldin index.

$$\mathrm{s(i)} = \frac{b(i) - a(i)}{\max(a(i), b(i))} \tag{9}$$

Where $a(i)$ is the average distance between point i and all other points in its own cluster, and $b(i)$ is the average distance between point i and all other points in the nearest cluster.

$$\mathrm{DB} = \frac{1}{k}\sum_{i=1}^{k}\max_{j\neq i}\left(\frac{s_i + s_j}{d_{ij}}\right) \tag{10}$$

Where k is the number of clusters, s_i is the average distance between points in cluster i, s_j is the average distance between points in cluster j, and d_{ij} is the distance between the centroids of clusters i and j.

To provide a more rigorous evaluation, several works utilize the Adjusted Rand Index (ARI) [183]. The ARI compensates for the expected overlap due to pure chance, providing a more reliable indicator of clustering quality as expressed in Eq. (11).

$$\mathrm{ARI} = \frac{\sum_{ij}\binom{n_{ij}}{2} - \left[\sum_i \binom{a_i}{2}\sum_j \binom{b_j}{2}\right]/\binom{n}{2}}{\frac{1}{2}\left[\sum_i \binom{a_i}{2} + \sum_j \binom{b_j}{2}\right] - \left[\sum_i \binom{a_i}{2}\sum_j \binom{b_j}{2}\right]/\binom{n}{2}} \tag{11}$$

where n_{ij} is the number of points in both cluster i and cluster j, a_i is the number of points in cluster i, b_j is the number of points in cluster j, and n is the total number of points.

6.4. Computational latency

In addition to the aforementioned metrics, computational latency is a critical factor in evaluating the performance of LPR algorithms, especially for real-time applications in agricultural robotics. Latency can be measured in terms of the average time taken to process a single query or the total time taken to process a batch of queries. Commonly, this time is reported in milliseconds (ms). The evaluation of latency is essential for ensuring that the LPR system can operate effectively within the constraints of real-time processing requirements [184].

Several papers present latency results for their proposed methods through the inference time per query [185], which is the time taken to process a single query and generate a place recognition result. Furthermore, the size of the point cloud data can also impact the computational load and processing time, which makes it necessary to compute the inference time for different point cloud sizes. This metric is crucial for applications that require real-time performance, as it directly impacts the responsiveness of the system. Additionally, some of the most recent works report the total processing time for a batch of queries [186], which provides insight into the scalability of the algorithm when handling larger datasets. This last approach is not usually reported in the LPR literature, but it is particularly relevant for agricultural environments where the robot may need to process a large number of queries in a short amount of time.

6.5. Memory usage

In the context of resource-constrained agricultural robots, the memory usage is a critical consideration when developing LPR systems. Memory usage can be measured in terms of the average memory consumed during the execution of the algorithm, as well as the peak memory usage. Evaluating memory usage is crucial for ensuring that the LPR system can operate within the hardware limitations of agricultural robots, which often have limited computational resources. Vilella-Cantos et al. [100] delve into the necessity of using low-dimensional global descriptors to encode the information of the places while maintaining a low memory footprint. Furthermore, while urban environments often require the storage of large maps for effective localization, agricultural environments can benefit from more compact representations due to their repetitive and structured nature. This allows for the development of LPR algorithms that are not only accurate but also efficient in terms of memory usage, which is essential for the practical deployment of autonomous agricultural robots.

7. Conclusions

The achievement of robust LPR stands as a critical milestone for the operational success of autonomous mobile robotics. Unlike static urban scenarios, agricultural environments present unique challenges—such as unstructured terrain and severe seasonal variability—that render standard urban-centric solutions ineffective. Consequently, developing specialized LPR approaches is not merely a technical necessity but a strategic one. Reliable localization in GNSS-denied environments is the key enabler for precision agriculture, allowing for optimized food production and mitigating the impact of the growing global agricultural workforce shortage.

To support this goal, this work presented a comprehensive review of the SOTA in agricultural localization, focusing on general agricultural DL literature from 2020 onwards and the existing specific LPR works in these settings. Moreover, we provided an analysis of the current results found in the scarce agricultural LPR literature. Although there have been improvements in recent years, cross-season validation remains a significant challenge, with recall values often dropping below 50%. We also observed significant performance differences between winter and summer campaigns, primarily due to severe environmental occlusions, with a difference in the Recall@1 metric of more than 20%. These findings underscore the necessity for additional research to develop robust algorithms that can handle the dynamic nature of agricultural environments.

We discussed the distinct environmental constraints and provided a critical analysis of current datasets and metrics. Notably, our comparison highlights a disparity between the abundance of computer vision datasets and the scarcity of data specifically curated for localization tasks. It is our hope that this survey clarifies the current landscape and guides future research toward bridging the gap between theoretical algorithms and real-world agricultural autonomy.

CRediT authorship contribution statement

Judith Vilella-Cantos: Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis, Conceptualization. **Mónica Ballesta:** Writing – review & editing, Visualization, Supervision, Resources, Formal analysis. **David Valiente:** Writing – review & editing, Visualization, Supervision, Resources, Formal analysis. **María Flores:** Visualization, Resources, Funding acquisition. **Luis Payá:** Visualization, Project administration, Funding acquisition.

Author statement

We would like to make the following declarations: All authors confirm that this work is original and has not been published elsewhere, nor is it currently under consideration for publication elsewhere.

We confirm that the manuscript has been read and approved by all named authors and that there are no other persons who satisfy the criteria for authorship but are not listed. We further confirm that the order of authors listed in the manuscript has been approved by all of us.

We understand that the Corresponding Author is the sole contact for the editorial process. He is responsible for communicating with the other authors about progress, submission of revisions and final approval of proofs.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research work is part of the project PID2023-149575OB-I00 funded by MICIU/AEI/10.13039/501100011033 and by FEDER, UE. It is also part of the project CIPROM/2024/8, funded by Generalitat Valenciana, Conselleria de Educación, Cultura, Universidades y Empleo (program PROMETEO 2025).

Data availability

No data were used for the research described in the article.

References

[1] F.K. Shaikh, S. Karim, S. Zeadally, J. Nebhen, Recent trends in internet-of-things-enabled sensor technologies for smart agriculture, IEEE Internet Things J. 9 (23) (2022) 23583–23598.

[2] V. Dhanya, A. Subeesh, N. Kushwaha, D.K. Vishwakarma, T.N. Kumar, G. Ritika, A. Singh, Deep learning based computer vision approaches for smart agricultural applications, Artif. Intell. Agric. 6 (2022) 211–229.

[3] I. Attri, L.K. Awasthi, T.P. Sharma, P. Rathee, A review of deep learning techniques used in agriculture, Ecol. Inform. 77 (2023) 102217.
[4] C. Masone, B. Caputo, A survey on deep visual place recognition, IEEE Access 9 (2021) 19516–19547.
[5] S. Garg, T. Fischer, M. Milford, Where is your place, visual place recognition? arXiv preprint arXiv:2103.06443, 2021.
[6] Y. Zhang, P. Shi, J. Li, LiDAR-based place recognition for autonomous driving: a survey, ACM Comput. Surv. 57 (4) (2024) 1–36.
[7] Z. Du, S. Ji, K. Khoshelham, 3-D LiDAR-based place recognition techniques: a review of the past ten years, IEEE Trans. Instrum. Meas. 73 (2024) 1–24.
[8] H. Ding, B. Zhang, J. Zhou, Y. Yan, G. Tian, B. Gu, Recent developments and applications of simultaneous localization and mapping in agriculture, J. Field Robot. 39 (6) (2022) 956–983.
[9] Z. Li, T. Shang, P. Xu, Z. Deng, Place recognition meet multiple modalities: a comprehensive review, current challenges and future development, Artif. Intell. Rev. 58 (11) (2025) 363.
[10] P. Yin, J. Jiao, S. Zhao, L. Xu, G. Huang, H. Choset, S. Scherer, J. Han, General place recognition survey: towards real-world autonomy, IEEE Trans. Robot. 41 (2025).
[11] S.I. Moazzam, U.S. Khan, M.I. Tiwana, J. Iqbal, W.S. Qureshi, S.I. Shah, A review of application of deep learning for weeds and crops classification in agriculture, in: 2019 International Conference on Robotics and Automation in Industry (ICRAI), IEEE, 2019, pp. 1–6.
[12] K. Lochan, A. Khan, I. Elsayed, B. Suthar, L. Seneviratne, I. Hussain, Advancements in precision spraying of agricultural robots: a comprehensive review, IEEE Access 12 (2024).
[13] A.T. Meshram, A.V. Vanalkar, K.B. Kalambe, A.M. Badar, Pesticide spraying robot for precision agriculture: a categorical literature review and future trends, J. Field Robot. 39 (2) (2022) 153–171.
[14] T. Barros, C. Premebida, S. Aravecchia, C. Pradalier, U. Nunes, SPVSoAP3D: a second-order average pooling approach to enhance 3D place recognition in horticultural environments, in: 2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), IEEE, 2024, pp. 9–15.
[15] M. Marzoa Tanco, G. Trinidad Barnech, F. Andrade, J. Baliosian, M. LLofriu, J.M. Di Martino, G. Tejera, MAgro dataset: a dataset for simultaneous localization and mapping in agricultural environments, Int. J. Robot. Res. 43 (5) (2024) 591–601.
[16] E. Commission, Building the future of EU agriculture with young farmers in focus, 2025, https://commission.europa.eu/news-and-media/news/building-future-eu-agriculture-young-farmers-focus-2025-10-21_en.
[17] C. Cheng, J. Fu, H. Su, L. Ren, Recent advancements in agriculture robots: benefits and challenges, Machines 11 (1) (2023) 48.
[18] T. Malthus, An essay on the principle of population, as it affects the future improvement of society, with remarks on the speculations of Mr. Godwin, M. Condorcet, and other writers, in: Charity and Philanthropy in Nineteenth-Century Britain, Routledge, 1798, pp. 5–7.
[19] N. Yahya, Agricultural 4.0: its implementation toward future sustainability, in: Green Urea: For Future Sustainability, Springer, Singapore, 2018, pp. 125–145, https://link.springer.com/chapter/10.1007/978-981-10-7578-0_5.
[20] A. Corallo, M.E. Latino, M. Menegoli, From Industry 4.0 to agriculture 4.0: a framework to Manage product data in agri-food supply chain for voluntary traceability, Int. J. Nutr. Food Eng. 12 (5) (2018) 146–150.
[21] D. o. E. United Nations, P. D. Social Affairs, World population prospects 2030, 2025, https://www.un.org/en/development/desa/population/events/pdf/other/21/21June_FINAL%20PRESS%20RELEASE_WPP17.pdf.
[22] P. Tamasiga, H. Onyeaka, M. Bakwena, A. Happonen, M. Molala, et al., Forecasting disruptions in global food value chains to tackle food insecurity: the role of AI and big data analytics–a bibliometric and scientometric analysis, J. Agric. Food Res. 14 (2023) 100819.
[23] A. Soussi, E. Zero, R. Sacile, D. Trinchero, M. Fossa, Smart sensors and smart data for precision agriculture: a review, Sensors 24 (8) (2024) 2647.
[24] E. Karunathilake, A.T. Le, S. Heo, Y.S. Chung, S. Mansoor, The path to smart farming: innovations and opportunities in precision agriculture, Agriculture 13 (8) (2023) 1593.
[25] S.K. Phang, T.H.A. Chiang, A. Happonen, M.M.L. Chang, From satellite to UAV-based remote sensing: a review on precision agriculture, IEEE Access 11 (2023) 127057–127076.
[26] F. Toscano, C. Fiorentino, N. Capece, U. Erra, D. Travascia, A. Scopa, M. Drosos, P. D'Antonio, Unmanned aerial vehicle for precision agriculture: a review, IEEE Access 12 (2024) 69188–69205.
[27] P. Radoglou-Grammatikis, P. Sarigiannidis, T. Lagkas, I. Moscholios, A compilation of UAV applications for precision agriculture, Comput. Netw. 172 (2020) 107148.
[28] M. Awais, W. Li, M. Cheema, Q. Zaman, A. Shaheen, B. Aslam, W. Zhu, M. Ajmal, M. Faheem, S. Hussain, et al., UAV-based remote sensing in plant stress imagine using high-resolution thermal sensor for digital agriculture practices: a meta-review, Int. J. Environ. Sci. Technol. 20 (1) (2023) 1135–1152.
[29] P.K. Singh, A. Sharma, An intelligent WSN-UAV-based IoT framework for precision agriculture application, Comput. Electr. Eng. 100 (2022) 107912.
[30] R.I. Mukhamediev, K. Yakunin, M. Aubakirov, I. Assanov, Y. Kuchin, A. Symagulov, V. Levashenko, E. Zaitseva, D. Sokolov, Y. Amirgaliyev, Coverage path planning optimization of heterogeneous UAVs group for precision agriculture, IEEE Access 11 (2023) 5789–5803.
[31] S. Coulibaly, B. Kamsu-Foguem, D. Kamissoko, D. Traore, Deep learning for precision agriculture: a bibliometric analysis, Intell. Syst. Appl. 16 (2022) 200102.
[32] K.R. Kalla, S. Saluja, R.R. Swetha, K. Sravanthi, D. Gupta, Precision agriculture redefined: combining machine learning and deep learning for enhanced crop yield prediction, in: Recent Trends in Intelligent Computing and Communication, CRC Press, 2025, pp. 635–641.
[33] T. Saranya, C. Deisy, S. Sridevi, K.S.M. Anbananthen, A comparative study of deep learning and internet of things for precision agriculture, Eng. Appl. Artif. Intell. 122 (2023) 106034.
[34] W. Winterhalter, F. Fleckenstein, C. Dornhege, W. Burgard, Localization for precision navigation in agricultural fields—beyond crop row following, J. Field Robot. 38 (3) (2021) 429–451.
[35] M. Martini, S. Cerrato, F. Salvetti, S. Angarano, M. Chiaberge, Position-agnostic autonomous navigation in vineyards with deep reinforcement learning, in: 2022 IEEE 18th International Conference on Automation Science and Engineering (CASE), IEEE, 2022, pp. 477–484.
[36] Z. Ma, X. Wang, X. Chen, B. Hu, J. Li, Advances in crop row detection for agricultural robots: methods, performance indicators, and scene adaptability, Agriculture 15 (20) (2025) 2151.
[37] V. Rondelli, B. Franceschetti, D. Mengoli, A review of current and historical research contributions to the development of ground autonomous vehicles for agriculture, Sustainability 14 (15) (2022) 9221.
[38] A. Silwal, F. Yandun, A.K. Nellithimaru, T. Bates, G. Kantor, Bumblebee: a path towards fully autonomous robotic vine pruning, Field Robot. 2 (1) (2022) 1661–1696.
[39] V.S. Bhat, Y. Wang, Revisiting the control systems of autonomous vehicles in the agricultural sector: a systematic literature review, IEEE Access 13 (2025).
[40] M. Padhiary, L.N. Sethi, A. Kumar, Enhancing hill farming efficiency using unmanned agricultural vehicles: a comprehensive review, Trans. Indian Natl. Acad. Eng. 9 (2) (2024) 253–268.
[41] X. Xiao, Y. Wang, Y. Jiang, Review of research advances in fruit and vegetable harvesting robots, J. Electr. Eng. Technol. 19 (1) (2024) 773–789.
[42] H. Zhou, X. Wang, W. Au, H. Kang, C. Chen, Intelligent robots for fruit harvesting: recent developments and future challenges, Precis. Agric. 23 (5) (2022) 1856–1907.
[43] T. Yoshida, Y. Onishi, T. Kawahara, T. Fukao, Automated harvesting by a dual-arm fruit harvesting robot, ROBOMECH J. 9 (1) (2022) 19.
[44] M. Bergerman, S.M. Maeta, J. Zhang, G.M. Freitas, B. Hamner, S. Singh, G. Kantor, Robot farmers: autonomous orchard vehicles help tree fruit production, IEEE Robot. Autom. Mag. 22 (1) (2015) 54–63.
[45] H. Etezadi, S. Eshkabilov, A comprehensive overview of control algorithms, sensors, actuators, and communication tools of autonomous all-terrain vehicles in agriculture, Agriculture 14 (2) (2024) 163.
[46] W. Hua, Z. Zhang, W. Zhang, X. Liu, C. Hu, Y. He, M. Mhamed, X. Li, H. Dong, C.K. Saha, et al., Key technologies in apple harvesting robot for standardized Orchards: a comprehensive review of innovations, challenges, and future directions, Comput. Electron. Agric. 235 (2025) 110343.
[47] A. Navone, M. Martini, M. Ambrosio, A. Ostuni, S. Angarano, M. Chiaberge, GPS-free autonomous navigation in cluttered tree rows with deep semantic segmentation, Robot. Auton. Syst. 183 (2025) 104854.
[48] Y. Luo, G. Gao, P. Jiang, C. Sun, M. Xu, Z. Zhou, W. Hu, Y. Tan, A navigation line extraction method based on semantic segmentation and centerline fitting, Smart Agricultural Technology (2025) 101492.
[49] V. Mazzia, F. Salvetti, D. Aghi, M. Chiaberge, DeepWay: a deep learning waypoint estimator for global path generation, Comput. Electron. Agric. 184 (2021) 106091.
[50] Q. Liu, R. Yu, H. Suo, Y. Cai, L. Chen, H. Jiang, Autonomous driving in agricultural machinery: advancing the frontier of precision agriculture, in: Actuators, vol. 14, MDPI, 2025, pp. 464.
[51] F. Salvetti, S. Angarano, M. Martini, S. Cerrato, M. Chiaberge, Waypoint generation in row-based crops with deep learning and contrastive clustering, in: Joint European Conference on Machine Learning and Knowledge Discovery in Databases, Springer, 2022, pp. 203–218.
[52] D. Aghi, V. Mazzia, M. Chiaberge, Local motion planner for autonomous navigation in vineyards with a RGB-D camera-based algorithm and deep learning synergy, Machines 8 (2) (2020) 27.
[53] E. Liu, J. Monica, K. Gold, L. Cadle-Davidson, D. Combs, Y. Jiang, Vision-based vineyard navigation solution with automatic annotation, in: 2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), IEEE, 2023, pp. 4234–4241.
[54] D. Biehler, Using autonomous pathing orchard robots, 2023, https://csanr.wsu.edu/using-autonomous-pathing-orchard-robots/.
[55] B.K. Mishra, B. Singh, R. Panwar, Smart agriculture using deep learning and FDM for the relation of soil health, fertilizer and ground water, Int. J. Syst. Assur. Eng. Manag. (2025) 1–9.
[56] B. Wu, M. Zhang, H. Zeng, F. Tian, A.B. Potgieter, X. Qin, N. Yan, S. Chang, Y. Zhao, Q. Dong, et al., Challenges and opportunities in remote sensing-based crop monitoring: a review, Natl. Sci. Rev. 10 (4) (2023) nwac290.
[57] P. Karmakar, S.W. Teng, M. Murshed, S. Pang, Y. Li, H. Lin, Crop monitoring by multimodal remote sensing: a review, Remote Sens. Appl. Soc. Environ. 33 (2024) 101093.
[58] S. Khaki, L. Wang, Crop yield prediction using deep neural networks, Front. Plant Sci. 10 (2019) 621.
[59] K. Swain, S.R. Nayak, V. Ravi, S. Mishra, T.J. Alahmadi, P. Singh, M. Diwakar, Empowering crop selection with ensemble learning and K-means clustering: a modern agricultural perspective, Open Agric. J. 18 (1) (2024).
[60] E. Omia, H. Bae, E. Park, M.S. Kim, I. Baek, I. Kabenge, B.-K. Cho, Remote sensing in field crop monitoring: a comprehensive review of sensor systems, data analyses and recent advances, Remote Sens. 15 (2) (2023) 354.
[61] J. Kotwal, R. Kashyap, S. Pathan, Agricultural plant diseases identification: from traditional approach to deep learning, Mater. Today Proc. 80 (2023) 344–356.

[62] H. Teng, Y. Wang, X. Song, K. Karydis, Multimodal dataset for localization, mapping and crop monitoring in citrus tree farms, in: International Symposium on Visual Computing, Springer, 2023, pp. 571–582.
[63] J. Qiang, X. Qiao, Y. Yu, Z. Liu, Z. Tian, S. Zhang, QseedNet: a lightweight deep-learning framework for detection of quarantine weed seeds in imported soybean inspection, Ind. Crops Prod. 238 (2025) 122314.
[64] Y. Gulzar, Z. Ünal, K. Şahbaz, M. Alkanan, PTL-inception: integrating deep learning and taxonomy for desert plant classification, Diversity 17 (11) (2025) 806.
[65] A. Gadicha, K. Bijwe, Comparative analysis of deep learning models for fruit quality detection, in: EPJ Web of Conferences, vol. 341, EDP Sciences, 2025, p. 01002.
[66] B. Sambu, V. Mbandu, T. Anondo, Deep learning approach for detection and prediction of Pest infections on plants in greenhouses, International Journal of Professional Practice 13 (4) (2025) 1–15.
[67] A. Revathi, A.A. Agash, Pomegranate disease diagnosis with severity estimation and treatment remedies using deep learning and RAG-based LLM, Sci. Rep. 15 (1) (2025) 41278.
[68] D.D. Uyeh, A. Akinsoji, S. Asem-Hiablie, B.I. Bassey, A. Osinuga, R. Mallipeddi, M. Amaizu, Y. Ha, T. Park, An online machine learning-based sensors clustering system for efficient and cost-effective environmental monitoring in controlled environment agriculture, Comput. Electron. Agric. 199 (2022) 107139.
[69] A. Ahmadi, M. Halstead, C. McCool, Bonnbot-i: a precise weed management and crop monitoring platform, in: 2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), IEEE, 2022, pp. 9202–9209.
[70] T. Gill, S.K. Gill, D.K. Saini, Y. Chopra, J.P. de Koff, K.S. Sandhu, A comprehensive review of high throughput phenotyping and machine learning for plant stress phenotyping, Phenomics 2 (3) (2022) 156–183.
[71] S. Arya, K.S. Sandhu, J. Singh, S. Kumar, Deep learning: as the new frontier in high-throughput plant phenotyping, Euphytica 218 (4) (2022) 47.
[72] A. Atefi, Y. Ge, S. Pitla, J. Schnable, Robotic technologies for high-throughput plant phenotyping: contemporary reviews and future perspectives, Front. Plant Sci. 12 (2021) 611940.
[73] M.S. Akhtar, Z. Zafar, R. Nawaz, M.M. Fraz, Unlocking plant secrets: a systematic review of 3D imaging in plant phenotyping techniques, Comput. Electron. Agric. 222 (2024) 109033.
[74] R.-F. Wang, H.-R. Qu, W.-H. Su, From sensors to insights: technological trends in image-based high-throughput plant phenotyping, Smart Agric. Technol. (2025) 101257.
[75] N. Chebrolu, F. Magistri, T. Läbe, C. Stachniss, Registration of spatio-temporal point clouds of plants for phenotyping, PLoS One 16 (2) (2021) e0247243.
[76] D. Schunck, F. Magistri, R.A. Rosu, A. Cornelißen, N. Chebrolu, S. Paulus, J. Léon, S. Behnke, C. Stachniss, H. Kuhlmann, et al., Pheno4D: a spatio-temporal dataset of maize and tomato plant point clouds for phenotyping and advanced plant analysis, PLoS One 16 (8) (2021) e0256340.
[77] R. Zenkl, R. Timofte, N. Kirchgessner, L. Roth, A. Hund, L. Van Gool, A. Walter, H. Aasen, Outdoor plant segmentation with deep learning for high-throughput field phenotyping on a diverse wheat dataset, Front. Plant Sci. 12 (2022) 774068.
[78] S. Zhou, X. Chai, Z. Yang, H. Wang, C. Yang, T. Sun, Maize-IAS: a maize image analysis software using deep learning for high-throughput plant phenotyping, Plant Methods 17 (1) (2021) 48.
[79] M. Gao, T.-F. Lu, Image processing and analysis for autonomous Grapevine pruning, in: 2006 International Conference on Mechatronics and Automation, IEEE, 2006, pp. 922–927.
[80] A. Navone, M. Martini, M. Chiaberge, Autonomous robotic pruning in Orchards and vineyards: a review, arXiv preprint arXiv:2505.07318, 2025.
[81] H. Williams, D. Smith, J. Shahabi, T. Gee, M. Nejati, B. McGuinness, K. Black, J. Tobias, R. Jangali, H. Lim, et al., Modelling wine grapevines for autonomous robotic cane pruning, Biosyst. Eng. 235 (2023) 31–49.
[82] S. Jiang, P. Qi, L. Han, L. Liu, Y. Li, Z. Huang, Y. Liu, X. He, Navigation system for orchard spraying robot based on 3D LiDAR SLAM with NDT_ICP point cloud registration, Comput. Electron. Agric. 220 (2024) 108870.
[83] F.E. Nasir, M. Tufail, M. Haris, J. Iqbal, S. Khan, M.T. Khan, Precision agricultural robotic sprayer with real-time tobacco recognition and spraying system based on deep learning, PLoS One 18 (3) (2023) e0283801.
[84] A. Milioto, P. Lottes, C. Stachniss, Real-time semantic segmentation of crop and weed for precision agriculture robots leveraging background knowledge in CNNs, in: 2018 IEEE International Conference on Robotics and Automation (ICRA), IEEE, 2018, pp. 2229–2235.
[85] J. Padmapriya, T. Sasilatha, Deep learning based multi-labelled soil classification and empirical estimation toward sustainable agriculture, Eng. Appl. Artif. Intell. 119 (2023) 105690.
[86] T. Barros, P. Conde, G. Gonçalves, C. Premebida, M. Monteiro, C.S.S. Ferreira, U.J. Nunes, Multispectral vineyard segmentation: a deep learning comparison study, Comput. Electron. Agric. 195 (2022) 106782.
[87] M.M. Tanco, G. Tejera, J.M. Di Martino, Learning agriculture keypoint descriptors with triplet loss for visual SLAM, J. Ambient Intell. Humaniz. Comput. (2023) 1–10.
[88] A.G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, H. Adam, MobileNets: efficient convolutional neural networks for mobile vision applications, arXiv preprint arXiv:1704.04861, 2017.
[89] M.S.U. Khan, T. Ahmed, H. Chen, A review of SLAM techniques for agricultural robotics: challenges, architectures, and future directions, in: 2025 IEEE 15th International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER), IEEE, 2025, pp. 598–603.
[90] R. Islam, H. Habibullah, T. Hossain, AGRI-SLAM: a real-time stereo visual SLAM for agricultural environment, Auton. Robots 47 (6) (2023) 649–668.
[91] F. Shu, P. Lesur, Y. Xie, A. Pagani, D. Stricker, SLAM in the field: an evaluation of monocular mapping and localization on challenging dynamic agricultural environment, in: Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, 2021, pp. 1761–1771.
[92] J. Lv, B. Yao, H. Guo, C. Gao, W. Wu, J. Li, S. Sun, Q. Luo, MOLO-SLAM: a semantic SLAM for accurate removal of dynamic objects in agricultural environments, Agriculture 14 (6) (2024) 819.
[93] R. Mur-Artal, J.D. Tardós, ORB-SLAM2: an open-source SLAM system for monocular, stereo, and RGB-D cameras, IEEE Trans. Robot. 33 (5) (2017) 1255–1262.
[94] X. Xu, J. Liang, J. Li, G. Wu, J. Duan, M. Jin, H. Fu, Stereo visual-inertial localization algorithm for orchard robots based on point-line features, Comput. Electron. Agric. 224 (2024) 109202.
[95] T. Barros, L. Garrote, P. Conde, M. Coombes, C. Liu, C. Premebida, U.J. Nunes, ORCHNET: a robust global feature aggregation approach for 3D LiDAR-based place recognition in Orchards, arXiv preprint arXiv:2303.00477, 2023.
[96] N. Sun, Z. Fan, Q. Qiu, T. Li, Q. Feng, C. Ji, C. Zhao, TriLoc-NetVLAD: enhancing long-term place recognition in Orchards with a novel LiDAR-Based approach, in: 2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), IEEE, 2024, pp. 16–22.
[97] T. Barros, L. Garrote, P. Conde, M. Coombes, C. Liu, C. Premebida, U. Nunes, PointNetPGAP-SLC: a 3D LiDAR-Based place recognition approach with segment-level consistency training for mobile robots in horticulture, IEEE Robot. Autom. Lett. 9 (2024).
[98] C.R. Qi, H. Su, K. Mo, L.J. Guibas, PointNet: deep learning on point sets for 3D classification and segmentation, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 652–660.
[99] F. Ou, Y. Li, Z. Miao, Place recognition of large-scale unstructured Orchards with attention score maps, IEEE Robot. Autom. Lett. 8 (2) (2023) 958–965.
[100] J. Vilella-Cantos, M. Martini, M. Chiaberge, M. Ballesta, D. Valiente, Low cost, high efficiency: LiDAR place recognition in vineyards with Matryoshka representation learning, arXiv preprint, 2026, https://doi.org/10.48550/arXiv.2601.18714
[101] J.J. Cabrera, A. Santo, A. Gil, C. Viegas, L. Payá, MinkUNeXt: point cloud-based large-scale place recognition using 3D sparse convolutions, Array 28 (2025) 100569, https://doi.org/10.1016/j.array.2025.100569
[102] A. Kusupati, G. Bhatt, A. Rege, M. Wallingford, A. Sinha, V. Ramanujan, W. Howard-Snyder, K. Chen, S. Kakade, P. Jain, et al., Matryoshka representation learning, Adv. Neural Inf. Process. Syst. 35 (2022) 30233–30249.
[103] A.S. Aguiar, F. Neves dos Santos, H. Sobreira, J. Boaventura-Cunha, A.J. Sousa, Localization and mapping on agriculture based on point-feature extraction and semiplanes segmentation from 3D LiDAR data, Front. Robot. AI 9 (2022) 832165.
[104] Z. Fei, S. Vougioukas, Row-sensing templates: a generic 3D sensor-based approach to robot localization with respect to orchard row centerlines, J. Field Robot. 39 (6) (2022) 712–738.
[105] F. Ou, Y. Li, N. Li, J. Zhou, W. Zhang, Z. Miao, SG-ISBP: orchard robots localization and mapping with ground optimization and loop closure detection integration, IEEE Sens. J. 24 (8) (2024) 13164–13174.
[106] K. Kim, A. Deb, D.J. Cappelleri, P-AgSLAM: in-row and under-canopy SLAM for agricultural monitoring in cornfields, IEEE Robot. Autom. Lett. 9 (6) (2024) 4982–4989.
[107] R. de Silva, J. Cox, M. Popovic, C. Cadena, C. Stachniss, R. Polvara, Keypoint semantic integration for improved feature matching in outdoor agricultural environments, arXiv preprint arXiv:2503.08843, 2025.
[108] A. Papadimitriou, I. Kleitsiotis, I. Kostavelis, I. Mariolis, D. Giakoumis, S. Likothanassis, D. Tzovaras, Loop closure detection and SLAM in vineyards with deep semantic cues, in: 2022 International Conference on Robotics and Automation (ICRA), IEEE, 2022, pp. 2251–2258.
[109] R. de Silva, J. Cox, J.R. Heselden, M. Popovic, C. Cadena, R. Polvara, Semantic-aware particle filter for reliable vineyard robot localisation, arXiv preprint arXiv:2509.18342, 2025.
[110] N. Soncini, J. Cremona, E. Vidal, M. García, G. Castro, T. Pire, The Rosario dataset v2: multi-modal dataset for agricultural robotics, Int. J. Robot. Res. (2025) 02783649251368909.
[111] R. Polvara, S. Molina, I. Hroob, A. Papadimitriou, K. Tsiolis, D. Giakoumis, S. Likothanassis, D. Tzovaras, G. Cielniak, M. Hanheide, Bacchus long-term (BLT) data set: acquisition of the agricultural multimodal BLT data set with automated robot deployment, J. Field Robot. 41 (7) (2024) 2280–2298.
[112] M. Martini, M. Ambrosio, J. Vilella-Cantos, A. Navone, M. Chiaberge, TEMPO-VINE: a multi-temporal sensor fusion dataset for localization and mapping in vineyards, arXiv preprint arXiv:2512.04772, 2025.
[113] L. Luo, S. Zheng, Y. Li, Y. Fan, B. Yu, S.-Y. Cao, J. Li, H.-L. Shen, BEVPlace: learning LiDAR-based place recognition using bird's eye view images, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2023, pp. 8700–8709.
[114] D. Kong, X. Li, Q. Xu, Y. Hu, P. Ni, SC_LPR: semantically consistent LiDAR place recognition based on chained cascade network in long-term dynamic environments, IEEE Trans. Image Process. 33 (2024) 2145–2157.
[115] Y. Xia, M. Gladkova, R. Wang, Q. Li, U. Stilla, J.F. Henriques, D. Cremers, CASSPR: cross attention single scan place recognition, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2023, pp. 8461–8472.
[116] M.A. Uy, G.H. Lee, Pointnetvlad: deep point cloud based retrieval for large-scale place recognition, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 4470–4479.
[117] R. Arandjelovic, P. Gronat, A. Torii, T. Pajdla, J. Sivic, NetVLAD: CNN architecture for weakly supervised place recognition, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 5297–5307.

[118] Z. Liu, S. Zhou, C. Suo, P. Yin, W. Chen, H. Wang, H. Li, Y.-H. Liu, LPD-Net: 3D point cloud learning for large-scale place recognition and environment analysis, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019, pp. 2831–2840.
[119] X. Chen, T. Läbe, A. Milioto, T. Röhling, J. Behley, C. Stachniss, OverlapNet: a Siamese network for computing LiDAR scan similarity with applications to loop closing and localization, Auton. Robots 46 (1) (2022) 61–81.
[120] J. Komorowski, MinkLoc3D: point cloud based large-scale place recognition, in: Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, 2021, pp. 1790–1799.
[121] J. Komorowski, Improving point cloud based place recognition with ranking-based loss and large batch training, in: 2022 26th International Conference on Pattern Recognition (ICPR), IEEE, 2022, pp. 3699–3705.
[122] K. Żywanowski, A. Banaszczyk, M.R. Nowicki, J. Komorowski, MinkLoc3D-SI: 3D LiDAR place recognition with sparse convolutions, spherical coordinates, and intensity, IEEE Robot. Autom. Lett. 7 (2) (2021) 1079–1086.
[123] J. Vilella-Cantos, J.J. Cabrera, L. Payá, M. Ballesta, D. Valiente, MinkUNeXt-SI: improving point cloud-based place recognition including spherical coordinates and LiDAR intensity, arXiv preprint arXiv:2505.17591, 2025.
[124] Y. Ming, X. Yang, G. Zhang, A. Calway, CGIS-net: aggregating colour, geometry and implicit semantic features for indoor place recognition, in: 2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), IEEE, 2022, pp. 6991–6997.
[125] R. Dube, A. Cramariuc, D. Dugas, H. Sommer, M. Dymczyk, J. Nieto, R. Siegwart, C. Cadena, SegMap: segment-based mapping and localization using data-driven descriptors, Int. J. Robot. Res. 39 (2–3) (2020) 339–355.
[126] Y. Pan, X. Xu, W. Li, Y. Cui, Y. Wang, R. Xiong, Coral: colored structural representation for bi-modal place recognition, in: 2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), IEEE, 2021, pp. 2084–2091.
[127] K. Vidanapathirana, P. Moghadam, B. Harwood, M. Zhao, S. Sridharan, C. Fookes, Locus: LiDAR-based place recognition using spatiotemporal higher-order pooling, in: 2021 IEEE International Conference on Robotics and Automation (ICRA), IEEE, 2021, pp. 5075–5081.
[128] J. Knights, P. Moghadam, M. Ramezani, S. Sridharan, C. Fookes, Incloud: incremental learning for point cloud place recognition, in: 2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), IEEE, 2022, pp. 8559–8566.
[129] F. Tombari, S. Salti, L. Di Stefano, A combined texture-shape descriptor for enhanced 3D feature matching, in: 2011 18th IEEE International Conference on Image Processing, IEEE, 2011, pp. 809–812.
[130] G. Kim, A. Kim, Scan context: egocentric spatial descriptor for place recognition within 3D point cloud map, in: 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), IEEE, 2018, pp. 4802–4809.
[131] H. Wang, C. Wang, L. Xie, Intensity scan context: coding intensity and geometry relations for loop closure detection, in: 2020 IEEE International Conference on Robotics and Automation (ICRA), IEEE, 2020, pp. 2095–2101.
[132] G. Kim, S. Choi, A. Kim, Scan context++: structural place recognition robust to rotation and lateral variations in urban environments, IEEE Trans. Robot. 38 (3) (2021) 1856–1874.
[133] R.B. Rusu, N. Blodow, M. Beetz, Fast point feature histograms (FPFH) for 3D registration, in: 2009 IEEE International Conference on Robotics and Automation, IEEE, 2009, pp. 3212–3217.
[134] W. Wohlkinger, M. Vincze, Ensemble of shape functions for 3D object classification, in: 2011 IEEE International Conference on Robotics and Biomimetics, IEEE, 2011, pp. 2987–2992.
[135] J.-U. Im, S.-W. Ki, J.-H. Won, Omni point: 3D LiDAR-based feature extraction method for place recognition and point registration, IEEE Trans. Intell. Veh. 9 (8) (2024) 5255–5271.
[136] A. Santo, E. Heredia, C. Viegas, D. Valiente, A. Gil, Ground segmentation for LiDAR point clouds in structured and unstructured environments using a hybrid neural–geometric approach, Technologies 13 (4) (2025) 162.
[137] K. Żywanowski, A. Banaszczyk, M.R. Nowicki, Comparison of camera-based and 3D LiDAR-based place recognition across weather conditions, in: 2020 16th International Conference on Control, Automation, Robotics and Vision (ICARCV), IEEE, 2020, pp. 886–891.
[138] Z. Qi, J. Ma, J. Xu, Z. Zhou, L. Cheng, G. Xiong, Gspr: multimodal place recognition using 3D Gaussian splatting for autonomous driving, in: 2025 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), IEEE, 2025, pp. 8864–8871.
[139] J. Komorowski, M. Wysoczańska, T. Trzcinski, Minkloc++: LiDAR and monocular image fusion for place recognition, in: 2021 International Joint Conference on Neural Networks (IJCNN), IEEE, 2021, pp. 1–8.
[140] H. Lai, P. Yin, S. Scherer, Adafusion: visual-lidar fusion with adaptive weights for place recognition, IEEE Robot. Autom. Lett. 7 (4) (2022) 12038–12045.
[141] A. Liang, Z. Chen, W. Xiong, F. Meng, Y. Yin, D. Zhang, Q. Li, BdFusion: bi-directional visual-LiDAR fusion for resilient place recognition, ISPRS J. Photogramm. Remote Sens. 228 (2025) 408–419.
[142] T. Shang, Z. Li, P. Xu, J. Qiao, Mambaplace: text-to-point-cloud cross-modal place recognition with attention Mamba mechanisms, in: 2025 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), IEEE, 2025, pp. 16985–16992.
[143] Z. Qi, L. Cheng, Z. Zhou, G. Xiong, LRFusionPR: a polar BEV-based LiDAR-Radar fusion network for place recognition, IEEE Robot. Autom. Lett. 10 (2025).
[144] A. Geiger, P. Lenz, C. Stiller, R. Urtasun, Vision meets robotics: the KITTI dataset, Int. J. Robot. Res. 32 (11) (2013) 1231–1237.
[145] W. Maddern, G. Pascoe, C. Linegar, P. Newman, 1 year, 1000 Km: the Oxford robotcar dataset, Int. J. Robot. Res. 36 (1) (2017) 3–15.
[146] C. Min, W. Jiang, D. Zhao, J. Xu, L. Xiao, Y. Nie, B. Dai, ORFD: a dataset and benchmark for off-road freespace detection, in: 2022 International Conference on Robotics and Automation (ICRA), IEEE, 2022, pp. 2532–2538.
[147] J. Knights, K. Vidanapathirana, M. Ramezani, S. Sridharan, C. Fookes, P. Moghadam, Wild-places: a large-scale dataset for LiDAR place recognition in unstructured natural environments, in: 2023 IEEE International Conference on Robotics and Automation (ICRA), IEEE, 2023, pp. 11322–11328.
[148] P. Jiang, P. Osteen, M. Wigness, S. Saripalli, Rellis-3D dataset: data, benchmarks and analysis, in: 2021 IEEE International Conference on Robotics and Automation (ICRA), IEEE, 2021, pp. 1110–1116.
[149] M. Stein, S. Bargoti, J. Underwood, Image based Mango fruit detection, localisation and yield estimation using multiple view geometry, Sensors 16 (11) (2016) 1915.
[150] S. Bargoti, J. Underwood, Deep fruit detection in Orchards, in: 2017 IEEE International Conference on Robotics and Automation (ICRA), IEEE, 2017, pp. 3626–3633.
[151] X. Liu, S.W. Chen, S. Aditya, N. Sivakumar, S. Dcunha, C. Qu, C.J. Taylor, J. Das, V. Kumar, Robust fruit counting: combining deep learning, tracking, and structure from motion, in: 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), IEEE, 2018, pp. 1045–1052.
[152] Y.-Y. Zheng, J.-L. Kong, X.-B. Jin, X.-Y. Wang, T.-L. Su, M. Zuo, CropDeep: the crop vision dataset for deep-learning-based classification and detection in precision agriculture, Sensors 19 (5) (2019) 1058.
[153] N. Häni, P. Roy, V. Isler, MinneApple: a benchmark dataset for apple detection and segmentation, IEEE Robot. Autom. Lett. 5 (2) (2020) 852–858.
[154] M. Alessandrini, R.C.F. Rivera, L. Falaschetti, D. Pau, V. Tomaselli, C. Turchetti, A Grapevine leaves dataset for early detection and classification of esca disease in vineyards through machine learning, Data Br. 35 (2021) 106809.
[155] F. Abdelghafour, B. Keresztes, A. Deshayes, C. Germain, J.-P. Da Costa, An annotated image dataset of downy mildew symptoms on Merlot grape variety, Data Br. 37 (2021) 107250.
[156] D.K. Barbole, P.M. Jadhav, GrapesNet: Indian RGB & RGB-D vineyard image datasets for deep learning applications, Data Br. 48 (2023) 109100.
[157] B.A. Mohammed, P.A. Abdalla, S.M. Aziz, H.O. Hassan, Pomegranate disease detection and classification dataset for deep learning applications: a case study from Halabja city, Data Br. (2025) 112298.
[158] N. Chebrolu, P. Lottes, A. Schaefer, W. Winterhalter, W. Burgard, C. Stachniss, Agricultural robot dataset for plant classification, localization and mapping on sugar beet fields, Int. J. Robot. Res. 36 (10) (2017) 1045–1052.
[159] F. Cañadas-Aránega, J.L. Blanco-Claraco, J.C. Moreno, F. Rodriguez-Diaz, Multimodal mobile robotic dataset for a typical Mediterranean greenhouse: the GREENBOT dataset, Sensors 24 (6) (2024) 1874.
[160] F. Crocetti, E. Bellocchio, A. Dionigi, S. Felicioni, G. Costante, M.L. Fravolini, P. Valigi, ARD-VO: agricultural robot data set of vineyards and olive Groves, J. Field Robot. 40 (6) (2023) 1678–1696.
[161] G. Pandey, J.R. McBride, R.M. Eustice, Ford campus vision and LiDAR data set, Int. J. Robot. Res. 30 (13) (2011) 1543–1552.
[162] N. Carlevaris-Bianco, A.K. Ushani, R.M. Eustice, University of Michigan north campus long-term vision and LiDAR dataset, Int. J. Robot. Res. 35 (9) (2016) 1023–1035.
[163] P. Gatou, X. Tsiara, A. Spitalas, S. Sioutas, G. Vonitsanos, Artificial intelligence techniques in Grapevine research: a comparative study with an extensive review of datasets, diseases, and techniques evaluation, Sensors 24 (19) (2024) 6211.
[164] I. Sa, Z. Ge, F. Dayoub, B. Upcroft, T. Perez, C. McCool, Deepfruits: a fruit detection system using deep neural networks, Sensors 16 (8) (2016) 1222.
[165] A. Bouguettaya, H. Zarzour, A. Kechida, A.M. Taberkit, Deep learning techniques to classify agricultural crops through UAV imagery: a review, Neural Comput. Appl. 34 (12) (2022) 9511–9536.
[166] S. Vélez, M. Ariza-Sentís, J. Valente, VineLiDAR: high-resolution UAV-LiDAR vineyard dataset acquired over two years in Northern Spain, Data Br. 51 (2023) 109686.
[167] J. Sturm, N. Engelhard, F. Endres, W. Burgard, D. Cremers, A benchmark for the evaluation of RGB-D SLAM systems, in: 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems, IEEE, 2012, pp. 573–580.
[168] T.O. Hodson, Root mean square error (RMSE) or mean absolute error (MAE): when to use them or not, Geosci. Model Dev. Discuss. 2022 (2022) 1–10.
[169] P.Z. Lodetti, E.A.A. Neto, M.A.I. Martins, G.H.D.S. Costa, M.A. Ludwig, MAE and RMSE analysis of K-means predictive algorithm for photovoltaic generation, in: 2022 International Conference on Electrical, Computer and Energy Technologies (ICECET), IEEE, 2022, pp. 1–6.
[170] D.S.K. Karunasingha, Root mean square error or mean absolute error? Use their ratio as well, Inf. Sci. 585 (2022) 609–629.
[171] C.J. Willmott, K. Matsuura, Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance, Clim. Res. 30 (1) (2005) 79–82.
[172] T. Chai, R.R. Draxler, Root mean square error (RMSE) or mean absolute error (MAE)?–Arguments against avoiding RMSE in the literature, Geosci. Model Dev. 7 (3) (2014) 1247–1250.
[173] T. Pire, T. Fischer, G. Castro, P. De Cristóforis, J. Civera, J.J. Berlles, S-PTAM: stereo parallel tracking and mapping, Robot. Auton. Syst. 93 (2017) 27–42.
[174] J. Terven, D.-M. Cordova-Esparza, J.-A. Romero-González, A. Ramírez-Pedraza, E. Chávez-Urbiola, A comprehensive survey of loss functions and metrics in deep learning, Artif. Intell. Rev. 58 (7) (2025) 195.
[175] A. Nouri, B.M. Merzah, S. Mosayyebpour, R. Mousa, S. Hesaraki, Evaluation metrics in learning systems: a survey, arXiv preprint, 2025.

[176] S. Schubert, P. Neubert, S. Garg, M. Milford, T. Fischer, Visual place recognition: a tutorial [tutorial], IEEE Robot. Autom. Mag. 31 (3) (2023) 139–153.
[177] N. Keetha, A. Mishra, J. Karhade, K.M. Jatavallabhula, S. Scherer, M. Krishna, S. Garg, AnyLoc: towards universal visual place recognition, IEEE Robot. Autom. Lett. 9 (2) (2023) 1286–1293.
[178] K. Tang, Q. Wang, C. Yan, Y. Sun, S. Liu, DTD: density triangle descriptor for 3D LiDAR loop closure detection, Sensors 26 (1) (2025) 201.
[179] L. Yang, R. Mascaro, I. Alzugaray, S.M. Prakhya, M. Karrer, Z. Liu, M. Chli, LiDAR loop closure detection using semantic graphs with graph attention networks, J. Intell. Robot. Syst. 111 (1) (2025) 13.
[180] P.J. Rousseeuw, Silhouettes: a graphical aid to the interpretation and validation of cluster analysis, J. Comput. Appl. Math. 20 (1987) 53–65.
[181] K.R. Shahapure, C. Nicholas, Cluster quality analysis using silhouette score, in: 2020 IEEE 7th International Conference on Data Science and Advanced Analytics (DSAA), IEEE, 2020, pp. 747–748.
[182] D.L. Davies, D.W. Bouldin, A cluster separation measure, IEEE Trans. Pattern Anal. Mach. Intell. PAMI-1 (2) (1979) 224–227, https://doi.org/10.1109/TPAMI.1979.4766909
[183] J.M. Santos, M. Embrechts, On the use of the adjusted rand index as a metric for evaluating supervised classification, in: International Conference on Artificial Neural Networks, Springer, 2009, pp. 175–184.
[184] T. Maweni, P. Amayo, Adapting SegMap: a LiDAR place recognition framework for standalone use in C++ applications, in: MATEC Web of Conferences, vol. 417, EDP Sciences, 2025, p. 04004.
[185] Z. Shi, X. Zhang, W. Xu, Y. Xia, Y. Zang, S. Shen, C. Wang, L2RSI: cross-view LiDAR-based place recognition for large-scale urban scenes via remote sensing imagery, arXiv preprint arXiv:2503.11245, 2025.
[186] W. Wang, X. Hua, C. Li, P. Tian, Y. Wang, L. Liu, LSSCC-net: integrating spatial-feature aggregation and adaptive attention for large-scale point cloud semantic segmentation, Symmetry 18 (1) (2026) 124.

Author biography

Judith Vilella-Cantos received her B.Sc. degree in Computer Science and her M.Sc. degree in Automation and Robotics from the University of Alicante, Spain. She is currently a Ph.D. candidate with the Automation, Robotics and Computer Vision Lab at Miguel Hernández University of Elche. Her research interests include LiDAR sensors and place recognition in unstructured environments.

Mónica Ballesta received her B.Sc. degree in Industrial Engineering and her M.Sc. degree in Robotics and Automation from Miguel Hernández University of Elche. She is currently an Associate Professor in System Engineering and Automation Department at the same university.

David Valiente received his B.Sc. and M.Sc. degrees in Telecommunications Engineering and his Ph.D. in Industrial Technologies from Miguel Hernández University of Elche. He is currently an Associate Professor in the Electronics Department at the same institution.

María Flores received her B.Sc. degree in Industrial Engineering and her M.Sc. and Ph.D. degrees in Robotics from Miguel Hernández University of Elche. She currently serves as an Associate Professor in the Electronics Department at the same university.

Luis Payá received his B.Sc. and M.Sc. degrees in Industrial Engineering and his Ph.D. in Industrial Technologies from Miguel Hernández University of Elche. He is currently a Full Professor in System Engineering and Automation Department at the same university.