Computer Modeling in Engineering & Sciences

Tech Science Press

https://doi.org/10.32604/cmes.2026.085691

ARTICLE

GreenSeg: Ground Segmentation Algorithm for Agricultural Robots in Mediterranean Greenhouses Using RGB-D Point Clouds

Fernando Cañadas-Aránega[1,*], José C. Moreno[1] and José L. Blanco-Claraco[2]

[1]Department of Informatics, University of Almería, CIESOL, ceiA3, Almería, Spain
[2]Department of Engineering, University of Almería, CIESOL, ceiA3, Almería, Spain
*Corresponding Author: Fernando Cañadas-Aránega. Email: fernando.ca@ual.es

Received: 15 May 2026; Accepted: 09 September 2026; Published: 21 October 2026

ABSTRACT: Greenhouse agriculture in the Mediterranean region faces significant automation challenges due to its unique structural and environmental constraints. These environments are characterised by extremely narrow aisles, heterogeneous terrains—ranging from concrete to tilled soil—and severe optical interference caused by polyethylene covers, which induce specular reflections and 'ghost points' in depth sensors. While autonomous navigation is essential for digitising agricultural tasks, traditional solutions often rely on expensive 3D LiDAR systems that are economically unscalable for most facilities. To address this problem, this paper introduces GreenSeg, an architectural perception framework engineered for real-time robotic deployment within asynchronous ROS 2 pipelines using cost-effective RGB-D sensing. The proposed method introduces a dual-layer validation strategy: a robust global plane fitting combined with a surface curvature filter for terrain adaptability, and a seed-point-based Region Growing constraint to ensure the spatial continuity of the navigable plane. Experimental validation was conducted using the AGRICOBIOT I platform across four diurnal scenarios with varying solar elevations. The results show that GreenSeg consistently outperforms benchmark segmentation methods, achieving peak improvements of +11.72% in mean Recall and +19.21% in Mean Intersection over Union (mIoU) during critical rotational manoeuvres at the end of corridors. These findings confirm that the proposed algorithm enables stable and safe autonomous navigation in unstructured, dynamic agricultural environments that are subject to budget constraints and sensitive to lighting conditions. The open repository for this work is at this link: https://github.com/FerCanAra/ros2_greenseg.

KEYWORDS: Agricultural robotics; point cloud processing; autonomous navigation; visual navigation; region growing; ROS 2

1 Introduction

Over the past few decades, greenhouse agriculture has established itself as a fundamental strategy for increasing food production and optimising the management of water and energy resources [1]. These advancements are driven by the progressive integration of automated systems and advanced control mechanisms into critical processes, such as microclimate management. Within the realm of agricultural automation, mobile robots have emerged as essential agents in the transition toward a digitised agricultural framework, undertaking transport, inspection, and assistance tasks traditionally deemed labour-intensive or hazardous [2]. However, despite being relatively structured environments, greenhouses impose unique challenges for autonomous mobility that diverge significantly from those encountered in standard urban or industrial settings [3].

Copyright © 2026 The Authors. Published by Tech Science Press. This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Mediterranean-type greenhouses, which account for 92% of the global protected cultivation area, feature a structural configuration designed to maximise crop space at the expense of manoeuvreability [4]. These environments are characterised by extremely narrow aisles (ranging from 90 to 110 cm), which drastically constrains the kinematic margin of manoeuvre and demands centimetre-level precision for safe navigation [5,6]. Furthermore, the terrain exhibits complex, non-linear dynamics; the ground composition can alternate abruptly among compacted sand, concrete, gravel, and tilled soil. This heterogeneity, often compounded by crop residues and variable slopes (typically ranging from 1% to 2%), significantly increases the uncertainty in the robot's pose estimation, as we have already discussed in previous papers [7,8]. A critical bottleneck for robotic perception in such environments is the optical interference generated by polyethylene covers [9]. The white plastic materials and metallic frameworks induce specular reflections and glare phenomena that saturate depth sensors, generating "ghost points" or multipath reflection artifacts below the actual ground level. To ensure safe autonomous navigation, it is imperative to deploy robust ground segmentation algorithms capable of filtering these virtual artifacts and dynamically adapting to the irregular topography in real time [10].

In contrast to conventional navigation frameworks that rely on expensive omnidirectional 3D Light Detection and Ranging (LiDAR) sensors, the use of cameras as an alternative has proved particularly useful in recent years [11]. The implementation of adaptive segmentation algorithms on such sensors facilitates the extraction of dense point clouds for accurate navigable space identification, thereby significantly reducing technological deployment costs in agricultural facilities. By integrating specialized modules for multipath noise suppression and dynamic ground probability estimation, the proposed system ensures highly efficient segmentation, maintaining robustness even in the presence of severe optical saturation and the extreme terrain variability typical of greenhouse infrastructures [12].

The main objective of this research is to introduce and validate GreenSeg, an architectural perception framework optimised for real-time field-robotics deployment within asynchronous ROS 2 pipelines. Relying exclusively on low-cost depth vision systems via the AGRICOBIOT I mobile platform, this approach significantly improves the execution of autonomous navigation in dynamic and complex environments. This study proposes a novel dual-layer validation architecture for the camera-extracted point cloud: a surface curvature filter integrated with mathematical plane fitting to handle terrain irregularities, and a Region Growing connectivity constraint to ensure the spatial integrity of the navigable plane. System performance was rigorously evaluated across complex topological zones—including narrow crop rows, central corridors, and transitions between heterogeneous materials (concrete, gravel, and tilled soil)—while assessing its robustness against extreme diurnal lighting variations and platform kinematics. Experimental results demonstrate that this new approach, termed GreenSeg, achieves peak improvements of +11.72% in mean Recall and +19.21% in mIoU compared to the baseline algorithm during critical turning manouvre. These findings yield superior segmentation in dynamic environments characterised by severe shadows, depth discontinuities, and polyethylene-induced optical interference. Ultimately, this research confirms that precise and safe mobility can be achieved within Mediterranean greenhouse infrastructures utilizing affordable sensory suites, effectively overcoming traditional constraints associated with sensor "ghosting" and complex terrain irregularities without the need for high-cost LiDAR systems.

The principal challenges and problematics that this paper aims to address are summarised as follows:

- *Structural and Terrain Greenhouses Morphology*: Mediterranean greenhouses present an exceptionally complex operational space that combines severe canopy constraints, optical noise, and terrain heterogeneity. The dense growth pattern occludes the ground in narrow aisles (90 to 110 cm wide), while white polyethylene cladding generates massive specular reflections that create "ghost points" below the floor level. Contrasting directly with flat industrial environments, the ground exhibits complex non-linear

dynamics, transitioning among compacted sand, concrete, gravel, and tilled soil, with local elevations, variable slopes (1% to 2%), and operational obstacles such as irrigation hoses and crop residues that degrade pose estimation and dynamic stability.

- *Limitations of Traditional Navigation Frameworks*: Within greenhouse infrastructures, Global Positioning System (GPS)/Global Navigation Satellite System (GNSS) signal reception is severely attenuated or completely denied due to the metallic framework and dense vegetation canopy [13]. This forces an exclusive reliance on local exteroceptive sensors for autonomous navigation. Consequently, there is a critical need to develop cost-effective perception alternatives (such as Red, Green, Blue and Depth (RGB-D) cameras) that offer a scalable advantage over existing robust navigation systems that depend heavily on expensive 3D LiDARs.
- *Environmental Conditions*: The perceptual system must maintain its segmentation efficacy throughout the diurnal cycle, exhibiting resilience against continuous variations in solar elevation and azimuth. These dynamic lighting conditions directly exacerbate the glare on polyethylene surfaces and compromise the reliability of optical sensors.

The remainder of this paper is organized as follows. Section 2 reviews the state of the art in ground segmentation and environmental perception for agricultural robotics. Section 3 describes the experimental site, the AGRICOBIOT I robotic platform, and the baseline perception pipeline. Section 4 presents the proposed GreenSeg algorithm. Section 5 reports the experimental results under the four navigation trajectories, varying solar conditions, and behaviour in the presence of dynamic and static obstacles. Finally, Section 6 summarises the main conclusions and outlines directions for future work.

2 Related Works

Ground segmentation and environmental perception constitute the fundamental pillars of autonomous navigation in agriculture. Over the years, the dominant paradigm has evolved from conventional image processing techniques to advanced deep learning models tailored for unstructured scenes [14].

2.1 Rule-Based Techniques

Traditional ground segmentation techniques predominantly rely on the physical and geometric characteristics of point clouds [15]. Among the most widely adopted methods is plane fitting via RANdom SAmple Consensus (RANSAC) algorithm [16], which seeks the geometric plane maximise the number of inliers. However, these models frequently suffer from under-segmentation in non-planar terrains, such as greenhouses characterised by variable slopes and surface irregularities [10]. To enhance terrain representation, several studies have proposed the use of elevation maps, which discretize the space into a 2.5 D grid and facilitate obstacle identification through height variance analysis. Nevertheless, these methods are prone to failure in the presence of dense vegetation or overhanging structures, often yielding false positive detections [17].

In this context, algorithms such as Patchwork and its successor, Patchwork++, have represented a significant breakthrough by introducing region-based segmentation strategies and adaptive terrain analysis, achieving a robust balance between accuracy and computational efficiency in complex environments [18]. Similarly, techniques like Cloth Simulation Filtering (CSF) model the ground as a deformable surface, yielding satisfactory results in natural scenes; however, they remain highly sensitive to noise and require extensive manual parameter tuning [19]. Other probabilistic approaches, such as those based on Markov Random Fields (MRF), have successfully modelled the spatial coherence of the terrain, but their prohibitive computational cost restricts their applicability under real-time constraints [20]. Overall, these traditional methods struggle to generalise across highly dynamic and visually complex environments like greenhouses.

These geometric assumptions often collapse in Mediterranean greenhouses, where the high-reflectivity polyethylene covers generate multipath artifacts that RANSAC-based methods mistakenly incorporate into the ground plane, leading to critical errors in robot pose estimation.

2.2 Machine Learning Vision Model

The advent of machine learning has overcome the inherent limitations of hand-crafted descriptors through the automated learning of hierarchical representations [21]. Architectures such as PointNet and PointNet++ introduced the direct processing of unordered point clouds to the agricultural domain, achieving remarkable results in object classification and semantic segmentation in cabbage and potato crops, among others [12,22,23]. Nonetheless, these networks typically struggle to model local contextual geometries in large-scale, high-density agricultural environments [24].

To address these shortcomings, more advanced architectures like Kernel Point Convolution (KPConv), Randomized LAyer-aggregation Network (RandLA-Net), and Sparse Convolutional Networks (SCNs) have been developed across various industrial domains. These frameworks introduce efficient spatial convolutions and attention mechanisms to enhance the representation of geometric structures when operating on plants or fruits in open agricultural fields, where lighting and irradiance conditions are known [17]. RandLA-Net, in particular, has demonstrated outstanding real-time performance owing to its random sampling strategy and hierarchical feature aggregation when deployed in field environments using Unmanned Aerial Vehicle (UAV)-acquired point clouds over eggplant plantations [25]. Another highly relevant research avenue is deep supervised learning specifically designed for ground/non-ground classification, exemplified by Ground Estimation Net (GndNet). This network learns terrain distributions directly from LiDAR data, effectively bypassing the constraints of classical geometric methods [17].

To tackle computational efficiency, SCNs and voxel-based methods have been engineered to execute convolution operations exclusively in non-empty regions, thereby drastically reducing memory consumption [26]. Models like RandLA-Net have set benchmarks in efficiency by employing random sampling alongside lightweight attention mechanisms to process millions of points in real time across diverse plantations [27]. In recent investigations, the integration of Plane Attention modules has enabled the capture of critical vertical geometric variations at ground-plant interfaces, substantially improving accuracy in boundary regions where structural limits are ambiguous [18]. Furthermore, multi-view strategies such as SnapNet project the point cloud into multiple 2D snapshots to leverage the maturity of image-based Convolutional Neural Networks (CNNs), subsequently performing a back-projection into the 3D space for the final semantic labeling [28]. While state-of-the-art machine learning architectures offer high semantic accuracy, their deployment on lightweight agricultural platforms remains constrained by prohibitive computational demands and a lack of training datasets that specifically account for the "ghost point" phenomena prevalent in plastic-covered environments.

2.3 Semantic Segmentation in Precision Agriculture

In recent years, ground and vegetation segmentation within agricultural environments has gravitated toward multi-modal approaches combining RGB, RGB-D, and LiDAR sensors [29]. These strategies aim to enhance perceptual robustness under real-world conditions characterised by extreme visual and structural variability [30,31]. In the computer vision domain, several studies have deployed semantic segmentation networks such as U-Net, DeepLabv3+, and Segmentation Network (SegNet) to differentiate among soil, crop, and weed classes using RGB imagery. For instance, Milioto et al. [32] proposed a CNN-based approach for semantic segmentation in precision agriculture, demonstrating high accuracy in structured crop scenarios. Similarly, more recent works have incorporated RGB-D data to refine depth estimation and improve

the separation between ground and vegetation planes, enabling superior discrimination in scenes with partial occlusions.

Conversely, the utilisation of LiDAR sensors has gained substantial traction due to their ability to capture highly accurate geometric information. In this context, Paigwar et al. [17] introduced GndNet, a neural network specifically tailored for ground segmentation in LiDAR point clouds, outperforming traditional geometric methods in complex environments. Similarly, Arafin et al. [33] developed deep learning-based approaches for semantic terrain classification in agricultural settings by integrating both geometric and contextual features. Moreover, works such as Kang and Wang [34] have explored the fusion of LiDAR data and RGB cameras to enhance crop and ground segmentation, exploiting both spectral and structural modalities. These multi-modal frameworks have proven particularly effective in scenarios subjected to highly variable illumination, where vision-only methods frequently degrade.

Within greenhouse environments, where structural density is higher and lighting conditions are significantly more volatile, recent studies have tackled segmentation through hybrid methodologies [35]. Despite these advancements, the majority of existing frameworks exhibit notable limitations concerning generalization capabilities, computational efficiency, or an over-reliance on controlled conditions. Specifically, purely vision-based methods remain highly susceptible to dynamic diurnal illumination shifts and extreme glare, whereas LiDAR-exclusive approaches struggle to discriminate between heterogeneous soil and low-lying vegetation [36]. Consequently, there remains a pressing need to develop robust, real-time segmentation algorithms capable of adapting to the idiosyncratic conditions of greenhouses, where heterogeneous surfaces, complex structures, and multiple dynamic agents invariably coexist. This work aligns with this imperative, proposing a framework specifically geared toward highly reliable ground segmentation for autonomous navigation in real-world agricultural settings.

3 Materials and Methods

This section details the experimental framework and methodological procedures employed to validate the proposed cost-effective RGB-D ground segmentation system within Mediterranean greenhouse environments. First, the experimental site located at the AgroConnect facilities is described; this setting features heterogeneous terrain (comprising sand, concrete, and gravel) alongside polyethylene plastic covers that induce significant optical artifacts. Subsequently, the AGRICOBIOT I mobile platform is introduced, a differential-drive robot specifically engineered for payload transportation within these confined spaces. The methodological core of this work lies in the development of a perception framework based on the Patchwork++ 3D point cloud segmentation algorithm, specifically adapted for integration with the Intel RealSense® D435 camera.

3.1 Materials

A description of the materials used is provided, based on the description of the greenhouse and the actual robot model.

3.1.1 AgroConnect Greenhouse Facilities

The experimental validation was conducted at the AgroConnect research facilities (Almería, Spain; 36°50′ N, 2°24′ W), a site co-funded by the Spanish Ministry of Science, Innovation, and Universities and the European Regional Development Fund (FEDER) (Fig. 1a). Situated at an elevation of 3 m above sea level, the terrain exhibits a consistent 1% northern gradient. The testbed comprises a 1850 m^2 Mediterranean-style greenhouse (traditionally classified as a *raspa y amagado* structure) featuring a robust steel framework

and high-density polyethylene cladding. Topologically, the facility is structured around a 2 m wide central corridor serving as the primary navigational artery. This axis connects to eleven lateral aisles on each side; the northern aisles measure 12.5 m in length, while the southern aisles extend to 22.5 m, both constrained to a width of 2 m without taking plants into account (Fig. 1b).



a Greenhouse outdoor



b Greenhouse indoor

Figure 1: AgroConnect greenhouse.

3.1.2 AGRICOBIOT I Robot

The experimental validation involves the deployment of the AGRICOBIOT I mobile platform, seamlessly integrated into a fog-based architecture to facilitate real-time data exchange across the facility's broader operational network. This robotic platform is multi-functional, engineered to assist greenhouse operators in crop supervision, the transport of raw materials and harvested vegetable crates, and targeted phytosanitary applications, thereby fostering a collaborative human-robot working environment. Built upon a commercial Husky chassis by Clearpath Robotics®, this platform employs a differential drive system capable of zero-radius turning, a kinematic requisite for navigating the severely constrained aisles of greenhouse environments. To optimise agricultural utility, the robot is outfitted with a custom-engineered superstructure featuring four central posts and a removable, rail-guided implement; this design facilitates the transport of standardised harvest boxes at an ergonomic height while maximising the available payload space for the perceptual suite. Fig. 2 shows various views of the AGRICOBIOT I robot.

To execute the experimental validation, the robotic platform is equipped with the following heterogeneous computational and sensory suite:

- *HITTSON® Unit Control*: This high-performance industrial onboard computer serves as the central processing hub for the robotic agent. It integrates a Wi-Fi 6 (802.11ax) wireless module, establishing high-bandwidth, low-latency communication with the local greenhouse network infrastructure.
- *Intel® RealSense™ D435 Camera*: An active stereo depth camera deployed for core visual perception and identification tasks. Within this architecture, it acquires the dense RGB-D data essential for the perceptual pipeline, enabling the system to extract precise spatial coordinates of dynamic agents (e.g.,

human operators and peer robots) while serving as the primary sensory input for the proposed ground segmentation framework. For the experimental test, the characteristics of the data measured by the robot's Intel RealSense sensor, as shown in Table 1, were used.



a AGRICOBIOT I in a corridor with high plants **b** AGRICOBIOT I in a corridor with low plants

Figure 2: AGRICOBIOT I robot.

Table 1: Technical specifications and data characteristics of the RGB-D sensor used in the experimental validation.

Data Parameter	Sensor Feature/Value	Data Stream Characteristic
Sensor Model	Intel RealSense D435	Active IR Stereo & RGB Camera
Depth Resolution	1280 × 720 pixels	High-density raw 3D Point Cloud
Frame Rate (Frequency)	30 fps (Hz)	Real-time data throughput
Field of View (FoV)	(87° ± 3°) × (58° ± 1°) × (95° ± 3°)	Horizontal × Vertical × Diagonal
Operational Range	0.28 to 3.0 m	Optimised for narrow greenhouse aisles
Z-Accuracy (Noise)	≤2% at 2 m	Subject to exponential depth degradation

3.2 Methods

This section describes the methods and algorithms used in the development of ground segmentation systems.

3.2.1 Robot Operating System 2 (ROS 2)

The robotic software architecture is fundamentally anchored in ROS 2 Humble Hawksbill, serving as the decentralized middleware framework for hardware abstraction and high-level task orchestration. Leveraging the Data Distribution Service (DDS) standard, this architecture implements rigorous Quality of Service (QoS) policies, which are critical for ensuring deterministic and reliable inter-process communication amidst the severe network degradation typical of greenhouse environments [37,38]. Within this multi-agent paradigm, ROS 2 guarantees seamless interoperability via standardised messaging protocols, enabling heterogeneous platforms to construct and share a unified world model. To ensure scalability and

mitigate data contention, each robotic unit operates within a strictly defined namespace. This architectural isolation permits a dedicated Message Queuing Telemetry Transport (MQTT) bridge to seamlessly aggregate local perceptual telemetry into a centralized fleet management layer devoid of cross-talk. Furthermore, the framework integrates the `tf2` transformation library to maintain rigorous spatial synchronization across all agents, dynamically projecting localized semantic detections (e.g., via YOLOx) into a globally consistent coordinate frame. Ultimately, the confluence of decentralized DDS communication and real-time spatial synchronization elevates individual autonomous nodes into a cohesive, perception-aware fleet. This synergy is fundamental for executing the complex cooperative navigation and collaborative human-robot interactions demanded by high-density agricultural settings [39].

3.2.2 Base Ground-floor Segmentation Framework

The perception pipeline is built upon a ground-floor segmentation algorithm developed by Intel® and distributed as part of the Robotics AI Suite [40]. The algorithm, based on Ground Plane Fitting (GPF) method, ingests raw 3-D sensor data produced by either Intel® RealSense™ active-stereo depth cameras or rotating 3D LiDAR scanners, and assigns one of four semantic labels to every valid measurement point, thereby partitioning the workspace into a navigable ground manifold, positive collision obstacles, overhanging structures, and sensor noise artefacts. This is achieved by retrieving the depth data published in ROS 2 by the Intel RealSense camera, along with the camera's intrinsic parameters as provided by the manufacturer, and extrinsic parameters entered manually.

The algorithm operates on a raw point cloud $\mathcal{P} = \{\mathbf{p}_i\}_{i=1}^{N}$ (where $i, N \in \mathbb{N}^+$) to transform it into the robot's coordinate frame, ensuring that each point $\mathbf{p}_i = (x_i, y_i, z_i)^\top \in \mathbb{R}^3$ is expressed relative to the robot's reference frame (`base_link`). Each point captured in the camera's optical frame is transformed via the ROS 2 TF2 tree according to the spatial relationship defined in Eq. (1).

$$\mathbf{p}_i^{\text{base}} = T_{\text{cam}}^{\text{base}} \cdot \mathbf{p}_i^{\text{cam}} = \mathbf{R}\,\mathbf{p}_i + \mathbf{t}, \tag{1}$$

where $T_{\text{cam}}^{\text{base}} \in SE(3)$ denotes the rigid body transformation (comprising a rotation matrix $\mathbf{R} \in SO(3)$ and a translation vector $\mathbf{t} \in \mathbb{R}^3$) between the sensor and the `base_link` frame. This transformation is obtained in real-time by querying the ROS 2 `TF2` transform tree at the specific timestamp of each frame. Subsequently, a secondary pre-filtering stage based on a minimum distance threshold is implemented to discard points in the immediate vicinity of the robot. This region is typically prone to self-specular reflections, occlusion shadows, and sensor saturation artifacts. To mitigate these effects, a radial distance threshold $d_{\min} \in \mathbb{R}^+$ is applied within the xy plane of the robot's reference frame, as formulated in Eq. (2).

$$\mathcal{P}_{\text{valid}} = \left\{ \mathbf{p}_i \in \mathcal{P} \;\middle|\; \sqrt{x_i^2 + y_i^2} \geq d_{\min} \right\}. \tag{2}$$

The subsequent phase focuses on identifying the dominant traversable surface through a Ground Fitting algorithm, which employs planar adjustment heuristics. Geometrically, the ground plane is parameterized by its unit normal vector $\mathbf{n} = (n_x, n_y, n_z)^\top \in \mathbb{R}^3$ ($\|\mathbf{n}\| = 1$) and a scalar offset $d \in \mathbb{R}$, such that any point p belonging to the plane satisfies $\mathbf{n}^\top p + d = 0$. The estimation process seeks the dominant horizontal manifold. To enforce this, a horizontal prior is implemented by constraining the normal vector to be approximately vertical, as defined in

$$n_z \gg n_x,\ n_y \qquad \Longrightarrow \qquad \mathbf{n} \approx (0,\ 0,\ 1)^\top. \tag{3}$$

The signed perpendicular distance $\delta_i = \hat{\mathbf{n}}^\top \mathbf{p}_i + \hat{d} \in \mathbb{R}$ represents the deviation from each candidate point to the estimated plane. Utilizing the estimated plane parameters $(\hat{\mathbf{n}}, \hat{d}) \in \mathbb{R}^4$ and a set of predefined heuristic thresholds, each point $\mathbf{p}_i \in \mathcal{P}_{\text{valid}}$ is assigned a semantic label ℓ_i according to the following classification logic (Eq. (4)).

$$\ell_i = \begin{cases} \textit{ground} & \text{if } |z_i| \le h_{\text{ground}} \text{ and } |\theta_i| \le \theta_{\max}, \\ \textit{obstacle} & \text{if } h_{\text{ground}} < z_i \le H_{\text{robot}} \text{ or } (|z_i| \le h_{\text{ground}} \text{ and } |\theta_i| > \theta_{\max}), \\ \textit{above} & \text{if } z_i > H_{\text{robot}}, \\ \textit{noise} & \text{otherwise}, \end{cases} \tag{4}$$

where $\theta = \arccos(\hat{n}_z) \in [0°, 180°]$ represents the tilt angle of the surface normal relative to the vertical axis, and $\theta_{\max} < 90°$. The parameters represent the following physical constraints:

- $h_{\text{ground}} \in \mathbb{R}^+$: The maximum perpendicular tolerance allowed for a point to be categorized as traversable ground.
- $\theta_{\max} \in \mathbb{R}^+$: The maximum permissible inclination of the fitted plane relative to the horizontal.
- $H_{\text{robot}} \in \mathbb{R}^+$: The physical height of the robot, acting as the decision boundary between the *obstacle* and *above* classes.

The plane parameters $(\hat{\mathbf{n}}, \hat{d})$ are refined by optimising the fit over the initial inlier set $\mathcal{P}^{(0)}_{\text{ground}} = \{\mathbf{p}_i \mid |\delta_i| \le h_{\text{ground}}\}$ through a constrained least-squares formulation:

$$(\hat{\mathbf{n}}, \hat{d}) = \arg\min_{\mathbf{n}, d} \sum_{\mathbf{p}_i \in \mathcal{P}^{(0)}_{\text{ground}}} \left(\mathbf{n}^\top \mathbf{p}_i + d\right)^2 \qquad \text{with} \quad \|\mathbf{n}\| = 1. \tag{5}$$

This optimization admits a closed-form solution: $\hat{\mathbf{n}}$ corresponds to the eigenvector associated with the smallest eigenvalue of the $\mathbf{C}_i$, satisfying $\mathcal{P}^{(0)}_{\text{ground}}$, while the offset is determined by $\hat{d} = -\hat{\mathbf{n}}^\top \bar{p}$, where $\bar{p}$ denotes the centroid of the inlier set. Finally, the framework publishes two distinct ROS 2 `sensor_msgs/PointCloud2` topics, one of them showing the floor `segmentation/labeled_points` and the other the obstacles `segmentation/obstacle_points`. The first constitutes the complete semantically labeled cloud (Eq. (6)) and the second is the dedicated obstacle cloud, which is directly integrated into the navigation stack's local cost-map (Eq. (7))

$$\mathcal{P}_{\text{ground}} = \left\{(\mathbf{p}_i, \ell_i)\right\}_{i=1}^{|\mathcal{P}_{\text{valid}}|}. \tag{6}$$

$$\mathcal{P}_{\text{obs}} = \{\mathbf{p}_i \in \mathcal{P}_{\text{valid}} \mid \ell_i = \textit{obstacle}\}. \tag{7}$$

These ROS 2 topics provide the essential perceptual feedback required for autonomous navigation and path planning within the greenhouse environment. Under such structured and static operating conditions, the absence of severe surface irregularities or variable gradients allows the plane-fitting heuristics to maintain systemic stability. However, this strict flat-world assumption exposes a significant algorithmic vulnerability when extrapolated to the unstructured, variable terrain inherent to Mediterranean greenhouses. The presence of uneven soil, arbitrary slopes, and agricultural waste often leads to under-segmentation, thereby necessitating the advanced algorithmic refinements and robust adaptations proposed within this study to ensure fail-safe autonomous navigation.

4 A Ground-Segmentation Approach for Greenhouses

This section describes the proposed method for floor segmentation, specifically for complex environments such as greenhouses.

4.1 Robot Setup

Initially, the camera was positioned beneath the black transport platform. This recessed position is intended to minimise interference from sunlight on the lens, which might otherwise cause unwanted reflections and, consequently, false positives in the perception algorithms. In this regard, Fig. 3 illustrates the camera configuration during one of the tests carried out in a real greenhouse. The sensor was recessed by 12 cm from the front overhang to minimise direct exposure to sunlight. Following its repositioning, it was lowered by 8 cm towards the floor and tilted at an angle of 30 degrees to reduce redundant or useless visual data.

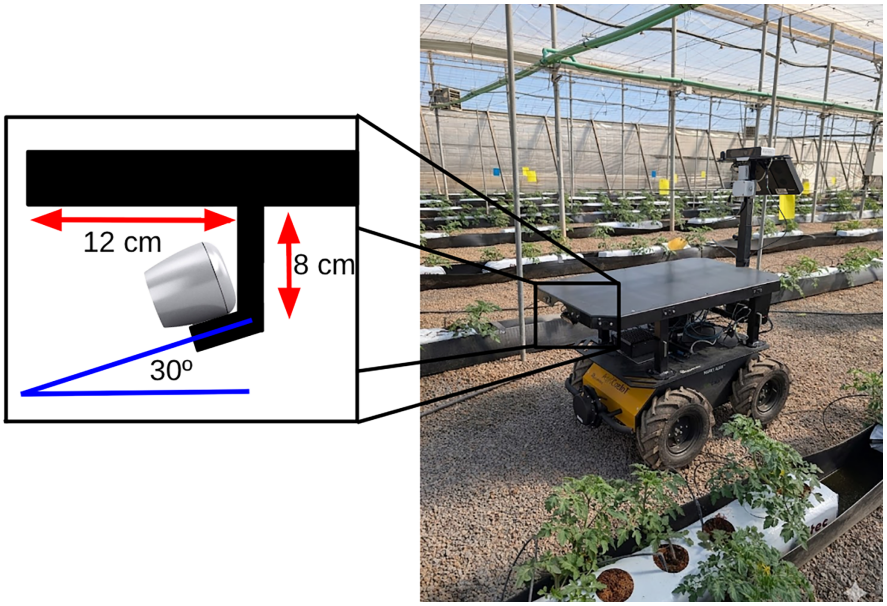


Figure 3: Camera location in robot.

4.2 Proposed Approach to Robust Ground Identification: GreenSeg

The base algorithm produces a map $\mathcal{P}_{\text{ground}}$ and an obstacle cloud $\mathcal{P}_{\text{obs}}$, which contains errors. In greenhouse environments, this map may be contaminated by specular reflections from the structure's polyethylene, highly irregular surfaces, or different types of flooring. The geometric verification stage acts as a second filter, ensuring that only geometrically consistent observations update the temporal model. The aim is to ensure future autonomous navigation in all possible scenarios a robot may encounter in a greenhouse, which will be analysed and studied in this section. The output of the base algorithm undergoes a second phase for soil classification, where $\mathcal{P}_{\text{ground}}$ and $\mathcal{P}_{\text{obs}}$ are evaluated using the *Region Growing* algorithm [41], with the new approach proposed in [42].

4.2.1 Local Surface Normal Estimation

For each measured point $\mathbf{p}_i \in \mathcal{P}^{(0)}_{\text{ground}}$, a local surface normal is estimated from its spherical neighbourhood $\mathcal{N}_i \subset \mathbb{R}^3$ of radius $r \in \mathbb{R}^+$:

$$\mathcal{N}_i = \left\{ \mathbf{p}_j \in \mathcal{P}_{\text{valid}} \;\middle|\; \left\| \mathbf{p}_j - \mathbf{p}_i \right\| \le r \right\}. \tag{8}$$

The local covariance matrix $\mathbf{C}_i \in \mathbb{R}^{3\times 3}$ of the neighbourhood is

$$\mathbf{C}_i = \frac{1}{|\mathcal{N}_i|} \sum_{\mathbf{p}_j \in \mathcal{N}_i} \left(\mathbf{p}_j - \bar{\mathbf{p}}_i^{\mathcal{N}} \right) \left(\mathbf{p}_j - \bar{\mathbf{p}}_i^{\mathcal{N}} \right)^{\top}, \tag{9}$$

where $\bar{\mathbf{p}}_i^{\mathcal{N}} = \frac{1}{|\mathcal{N}_i|} \sum_j \mathbf{p}_j$ is the estimated centroid of $\mathcal{N}_i$. The estimated local normal $\hat{\mathbf{m}}_i \in \mathbb{R}^3$ is defined as the unit eigenvector associated with the smallest eigenvalue $\lambda_{\min}$ of $\mathbf{C}_i$ as $\det(\mathbf{C}_i - \lambda \mathbf{I}) = 0$:

$$\mathbf{C}_i \, \hat{\mathbf{m}}_i = \lambda_{\min} \, \hat{\mathbf{m}}_i, \qquad \left\| \hat{\mathbf{m}}_i \right\| = 1. \tag{10}$$

Points with $|\mathcal{N}_i| < 30$ (value obtained from [42]) are immediately reclassified as *noise* and excluded from subsequent steps. A point in $\mathcal{P}^{(0)}_{\text{ground}}$ is geometrically consistent with the globally fitted plane $(\hat{\mathbf{n}}, \hat{d})$ (Eq. (5)) if its local normal is nearly collinear with $\hat{\mathbf{n}}$. The consistency score is defined as the absolute cosine similarity between the two normals:

$$\rho_i = \left| \hat{\mathbf{m}}_i^{\top} \hat{\mathbf{n}} \right| \in [0, 1]. \tag{11}$$

A value $\rho_i \approx 1$ indicates that the local surface at $\mathbf{p}_i$ is parallel to the globally fitted plane, while $\rho_i \approx 0$ signals a sharp local discontinuity incompatible with flat navigable terrain. The set of candidates passing the normal consistency check is

$$\mathcal{P}^{\rho}_{\text{ground}} = \left\{ \mathbf{p}_i \in \mathcal{P}^{(0)}_{\text{ground}} \;\middle|\; \rho_i \ge \rho_{\min} \right\}, \tag{12}$$

where $\rho_{\min} \in (0, 1]$ is a configurable threshold.

4.2.2 Local Curvature Filtering

The surface curvature at point $\mathbf{p}_i$ is estimated as the fraction of total neighbourhood variance explained by the normal direction, using the eigenvalues $\lambda_1 \ge \lambda_2 \ge \lambda_3 = \lambda_{\min}$ of $\mathbf{C}_i$:

$$\kappa_i = \frac{\lambda_{\min}}{\lambda_1 + \lambda_2 + \lambda_3} \in \left[0, \frac{1}{3} \right]. \tag{13}$$

Small values of κ_i indicate that the neighbours of $\mathbf{p}_i$ lie approximately on a plane (low curvature), whereas large values characterise edges, ridges, or scattered noise. Points with excessive curvature are removed from the candidate ground set:

$$\mathcal{P}^{\kappa}_{\text{ground}} = \left\{ \mathbf{p}_i \in \mathcal{P}^{\rho}_{\text{ground}} \;\middle|\; \kappa_i \le \kappa_{\max} \right\}. \tag{14}$$

The threshold $\kappa_{\max}$ controls the tolerance to local surface irregularities.

4.2.3 Region Growing Algorithm

The region growing algorithm verifies the spatial connectivity of the candidate ground set $\mathcal{P}^{\kappa}_{\text{ground}}$. The physical rationale is that a real navigable surface constitutes a single contiguous region; isolated clusters of points that pass the local geometric checks but are spatially disconnected from the main ground patch most likely correspond to sensor artefacts or small flat-topped debris.

The seed point $p_s \in \mathbb{R}^3$ is selected as the point in $\mathcal{P}^{\kappa}_{\text{ground}}$ closest to the robot origin in the xy plane of `base_link`, since the ground directly beneath the robot is the most reliable reference:

$$p_s = \underset{\mathbf{p}_i \in \mathcal{P}^{\kappa}_{\text{ground}}}{\arg\min} \sqrt{x_i^2 + y_i^2}. \tag{15}$$

Starting from $\mathcal{R} = \{p_s\}$, a candidate point $\mathbf{p}_j \in \mathcal{P}^{\kappa}_{\text{ground}} \backslash \mathcal{R}$ is added to the region if it simultaneously satisfies three conditions: (i) its signed distance to the globally fitted plane does not exceed the ground tolerance h_{ground} already defined in Eq. (4); (ii) its local normal is consistent with $\hat{\mathbf{n}}$ at level $\rho_{\min}$ (Eq. (12)); and (iii) it is spatially adjacent to at least one point already in $\mathcal{R}$ within a growing radius $r_g \in \mathbb{R}^+$:

$$\mathbf{p}_j \text{ is added to } \mathcal{R} \iff \begin{cases} \left|\delta_j\right| \le h_{\text{ground}}, \text{ and} \\ \rho_j \ge \rho_{\min}, \text{ and} \\ \exists\, \mathbf{p}_k \in \mathcal{R} : \left\|\mathbf{p}_j - \mathbf{p}_k\right\| \le r_g \end{cases} \tag{16}$$

where $\delta_j = \hat{\mathbf{n}}^\top \mathbf{p}_j + \hat{d}$ is the signed distance of $\mathbf{p}_j$ to the fitted plane, and r_g is the spatial adjacency radius.

4.2.4 Verified Ground and Obstacle Outputs

The verified ground set $\mathcal{P}^{\text{RG}}_{\text{ground}}$ is defined as the connected region produced by the growing algorithm:

$$\mathcal{P}^{\text{RG}}_{\text{ground}} = \mathcal{R}. \tag{17}$$

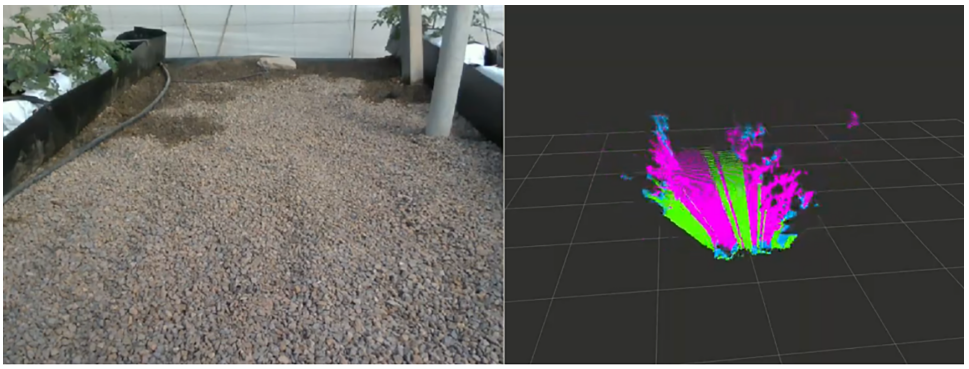
All points that were initially classified as ground by the base algorithm but did not survive the region growing validation are conservatively promoted to the *obstacle* or *noise* class. The verified obstacle cloud is therefore updated in the ROS 2 topic:

$$\mathcal{P}^{\text{RG}}_{\text{obs}} = \mathcal{P}_{\text{obs}} \cup \left(\mathcal{P}^{(0)}_{\text{ground}} \backslash \mathcal{P}^{\text{RG}}_{\text{ground}}\right), \tag{18}$$

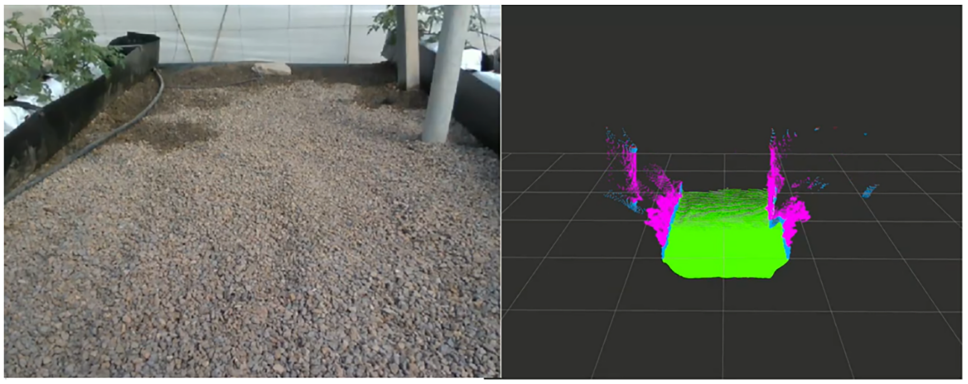
where $\mathcal{P}_{\text{obs}}$ is the obstacle cloud produced by the base algorithm (Eq. (7)). This conservative assignment ensures that any point whose ground membership cannot be spatially confirmed is treated as a potential collision threat, which is the appropriate behaviour for delicate agricultural environments where a false negative (missed obstacle) is more costly than a false positive (over-cautious path planning). The outputs $\mathcal{P}^{\text{RG}}_{\text{ground}}$ and $\mathcal{P}^{\text{RG}}_{\text{obs}}$ are published as `sensor_msgs/msg/PointCloud2` messages under two distinct ROS 2 topics: `/greenseg/labeled_points` for the navigable floor and `/greenseg/obstacle_points` for non-navigable entities. The first constitutes the complete semantically labeled cloud. The performance enhancement relative to the baseline algorithm is illustrated in Fig. 4, where Fig. 4a depicts the baseline representation in contrast to the proposed approach within the same experimental trial compared with the new approach shown in the Fig. 4b. Furthermore, Fig. 5 shows the flowchart, highlighting the inputs and outputs of each algorithm. In both cases, green indicates the ground, purple indicates obstacles, and cyan indicates points representing structures that are higher than the robot (such as the greenhouse roof, lamps or overhead cables).

4.3 Analysis of the Most Adverse Conditions in the Mediterranean Greenhouse

The performance of the optical system is intrinsically linked to the geometric orientation of the incident light, defined by the sun's elevation and azimuth angles [43]. Furthermore, in a greenhouse, this reflection is increased as the polyethylene in the roof sheeting deflects the light, scattering it and making it more diffuse.



a Greenhouse floor segmentation using the base algorithm



b GreenSeg greenhouse floor segmentation

Figure 4: Comparison of algorithms. Green points indicate the ground, purple indicates obstacles, and cyan indicates points representing structures that are higher than the robot.

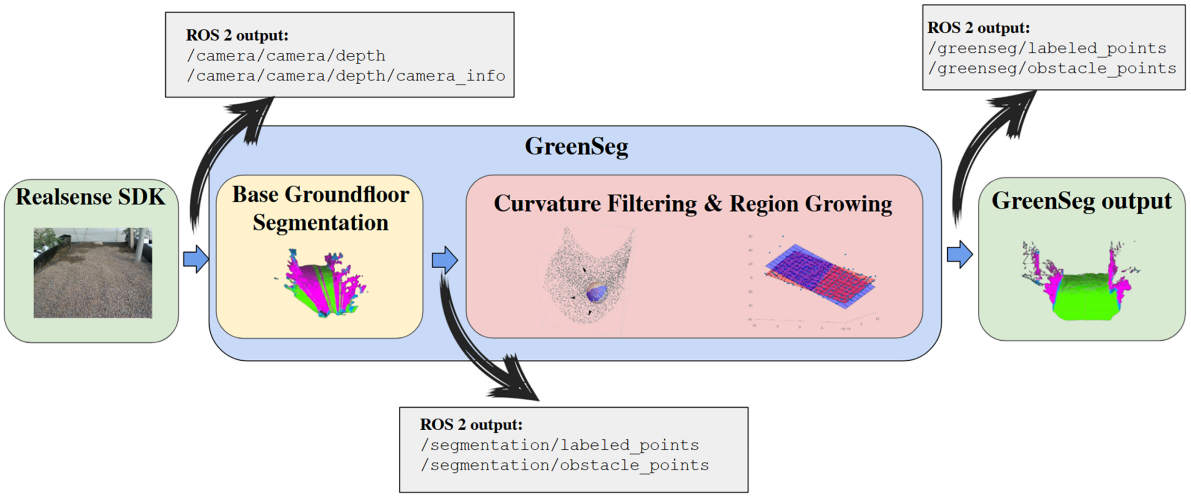


Figure 5: Flowchart of the proposed approach, showing the input and output topics for each layer.

4.3.1 Analysis of Solar Lighting in the Greenhouse

Whilst both determine the dynamic range of the scene, their physical impact on the camera lens elements follows distinct phenomenological trajectories:

- *Sun Elevation (γ)*: This parameter has the most direct and disruptive influence on the lens's internal optics, though its effects differ across its range. At low elevation angles (e.g., during dawn or dusk), light rays strike the lens at grazing angles, significantly increasing the probability of lens flare, ghosting, and stray light reflections within the internal lens barrel [44]. Conversely, at high elevation angles, the increased solar irradiance intensifies specular reflections on the polyethylene covers of the greenhouse, which often leads to sensor saturation in the depth perceptual pipeline.
- *Solar Azimuth (α)*: While azimuth does not typically cause the same level of internal optical aberration as elevation, it fundamentally governs the spatial distribution of shadows and the orientation of high-contrast boundaries. In the narrow aisles of a greenhouse, the azimuth angle determines whether a crop row is illuminated frontally or back-lit. Back-lighting (high relative azimuth) creates "halo" effects around leaves and structures, challenging the edge-detection capabilities of the segmentation algorithm and potentially distorting the perceived geometry of the ground plane [43,44].

The irradiance phenomena and solar geometry affecting the perception system are inherently dictated by the geographical coordinates of the testbed (Fig. 6).

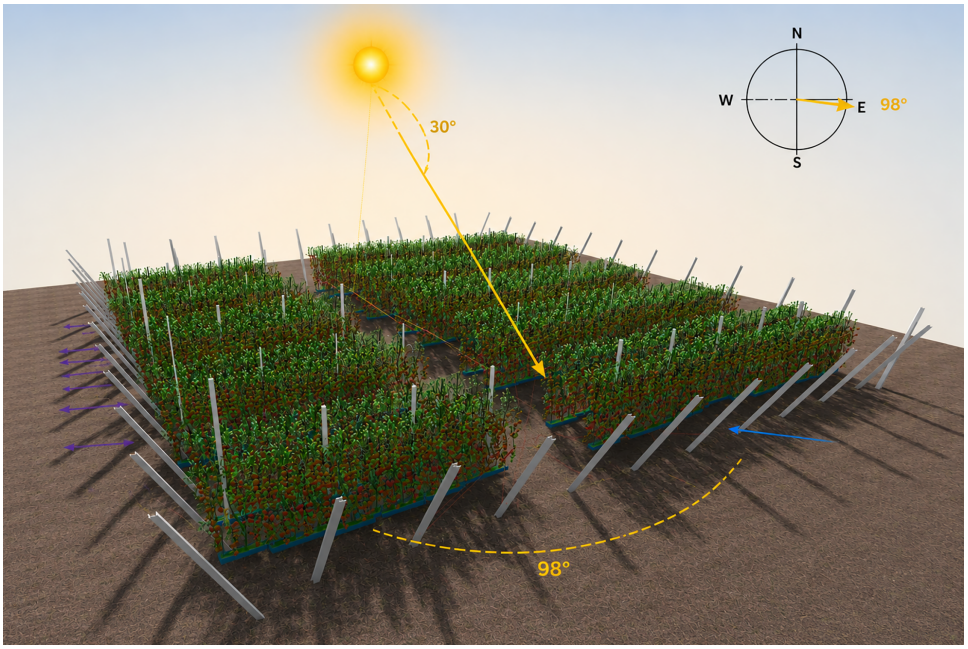


Figure 6: Example of solar elevation and azimuth in the greenhouse.

For this study, the AgroConnect greenhouse is located in Almería, Spain (36°50′ N, 2°24′ W). The experiments were carried out on 21st April 2026. To identify the critical operational windows, solar intensity profiles were retrieved from *TheSkyLive* platform[1], while precise solar elevation and azimuth angles were computed using the *SunEarthTools* ephemeris engine[2]. Based on this spatial-temporal analysis, the performance of the proposed algorithm was evaluated under the experimental scenarios presented in Table 2:

[1] TheSkyLive: https://theskylive.com.
[2] SunEarthTools: https://www.sunearthtools.com.

Table 2: Solar geometry parameters at specific operational windows on 21/04/2026 in Almería, Spain. Local time (Central European Summer Time, UTC+2).

Scenario	Time (UTC+2)	Sun Elevation (γ)	Solar Azimuth (α)
1	08:00 h	5.4°	79.2°
2	10:00 h	29.3°	97.2°
3	12:00 h	51.9°	122.5°
4	14:00 h	65.1°	175.0°

4.3.2 Analysis of Critical Areas in the Greenhouse

The greenhouse environment presents unique structural challenges that directly impact the reliability of autonomous navigation systems. These critical areas are primarily defined by the interaction between the crop architecture and the incident solar radiation, which directly affects the performance of the proposed algorithm. First, the narrow crop aisles function as occlusion tunnels that depend on the solar azimuth and elevation discussed previously. Second, the greenhouse's plastic cladding induces non-uniform light scattering. Under high solar elevation conditions, specular saturation points occur on the polyethylene surfaces, adversely affecting depth sensors and creating "blind spots" within the point cloud. Furthermore, the areas adjacent to the lateral and frontal greenhouse openings are particularly critical due to the ingress of unfiltered direct sunlight, which induces optical artifacts (lens flare) in the camera sensors. Analyzing these zones is essential to validate the robustness of the proposed algorithm, as the operational performance is highly dependent on the specific area where the robot is operating and how these occlusions impact the system. The various scenarios typically encountered in a Mediterranean greenhouse, all of which serve as the basis for evaluating the algorithm's robustness, are: i) Moving along the central corridor; ii) Moving between plants; iii) Turning at the end of the corridor; and iv) Changing to the next corridor.

4.4 Evaluation of the Proposed Approach

The performance of the proposed perception pipeline is assessed through a set of quantitative metrics computed over four field trials conducted in a Mediterranean greenhouse environment. Each test is compared with the baseline framework and the Ground Truth (GT); the methodology is detailed below.

4.4.1 Ground-Truth Identification

In dynamic agricultural environments, static capture-based ground-truth generation is not representative of real operating conditions, since the robot navigates continuously over uneven soil, exposed roots, and scattered debris. Instead, a fully automatic ground-truth generation procedure is proposed, exploiting the temporal consistency of the raw point cloud stream to identify high-confidence labels without manual intervention. This ensures that the proposed algorithm is not uniquely restricted to this specific scenario but is fully scalable to any greenhouse agricultural environment.

A semantic class label is considered reliable for point $\tilde{\mathbf{p}}_i$ at time t if and only if the geometric evidence supporting that classification remains consistent across a temporal window of $W = 10$ consecutive frames (which has a voxel resolution of 0.05 m) centered around t, defined by the frame index set $\mathcal{W}_t = \{k \in \mathbb{Z} \mid t - 4 \leq k \leq t + 5\}$. Transient or ambiguous classifications caused by specular reflections, motion blur, or sensor noise generate inconsistent labels across frames and are automatically excluded from the ground-truth reference dataset. To enable point-to-point correspondence across frames despite the robot's motion, the 3D space is discretized into a voxel grid with a resolution defined in the robot's `base_link` reference

frame. Each voxel accumulates the semantic class labels assigned to all points falling within it during the temporal window $\mathcal{W}_t$. Voxels labeled as *undefined* are excluded from all metric computations. For $W = 10$ frames, a semantic class label is accepted into the ground-truth reference dataset if that specific label is assigned to the voxel in at least 90% of the frames ($\psi_{\min} = 0.9$) within $\mathcal{W}_t$, where $\psi_{\min}$ denotes the temporal consistency threshold.

4.4.2 Evaluation Metrics

For each semantic class $k \in \mathcal{K} = \{ground, obstacle, above, noise\}$, the standard confusion-matrix quantities are defined over $\mathcal{D}_{\mathrm{GT}}$ as:

- TP_k (True Positives): Points predicted as class k that actually belong to class k in $\mathcal{D}_{\mathrm{GT}}$.
- FP_k (False Positives): Points predicted as class k that actually belong to a different class $j \neq k$ in $\mathcal{D}_{\mathrm{GT}}$.
- FN_k (False Negatives): Points belonging to class k in $\mathcal{D}_{\mathrm{GT}}$ that were incorrectly predicted as a different class $j \neq k$.
- TN_k (True Negatives): Points belonging to a different class $j \neq k$ in $\mathcal{D}_{\mathrm{GT}}$ that were correctly predicted as non-k.

To assess the precision $P_k \in \mathbb{R}^+$, the fraction of predicted positives for class k that are correct is calculated:

$$P_k = \frac{TP_k}{TP_k + FP_k}. \tag{19}$$

Furthermore, the recall for class k, denoted as $R_k \in [0,1]$, quantifies the sensitivity of the model by representing the proportion of ground truth instances that are correctly identified:

$$R_k = \frac{TP_k}{TP_k + FN_k}. \tag{20}$$

The harmonic mean of precision and recall, $F_{1,k} \in \mathbb{R}^+$, providing a single balanced measure, is given by:

$$F_{1,k} = 2 \cdot \frac{P_k \cdot R_k}{P_k + R_k} = \frac{2TP_k}{2TP_k + FP_k + FN_k}. \tag{21}$$

The standard overlap metric for semantic segmentation is the Intersection over Union (IoU) $\mathrm{IoU}_k \in \mathbb{R}^+$, measuring the ratio of correctly classified points to the total points involved for class k:

$$\mathrm{IoU}_k = \frac{TP_k}{TP_k + FP_k + FN_k}. \tag{22}$$

5 Results and Discussion

Prior to the field deployment in the greenhouses, a series of preliminary tests was conducted using the most common computer vision algorithms previously discussed, all of them tailored for ground segmentation in dynamic and complex environments. As previously noted, the target application requires the algorithm to operate at a high frame rate on computers with low computational specifications. In this regard, most of these frameworks failed to meet the baseline requirements necessary for a fair benchmarking against our proposed method. Consequently, Table 3 provides an overview of the applications, objectives and operational issues raised by these algorithms, which justifies limiting the quantitative evaluation exclusively to a single comparative framework. The open repository for this work is at this link: https://github.com/FerCanAra/ros2_greenseg.

Table 3: Comparison of representative ground segmentation methods.

Method	Paradigm	Target Sensor/ Application Domain	Data	Embedded RT
PointNet++/ KPConv	Deep Learning	General 3D point clouds/Dense scans	High	Medium–Low
RandLA-Net/GndNet	Deep Learning	Large-scale 3D point clouds (mainly LiDAR)/Autonomous driving	High	Medium
Patchwork++	Geometric	360° rotating LiDAR/Autonomous driving	None	High
CSF	Geometric	Airborne and terrestrial LiDAR/Forestry and topography	None	Low
GreenSeg (Proposed)	**Geometric**	**Low-cost RGB-D/Greenhouse environments**	**None**	**High**

A comparison will therefore be carried out with the base algorithm, which has already been compared in numerous articles [9,35,36]. This section describes the simulations and experiments designed to validate the proposed approach, as well as the results obtained.

5.1 Performance under Critical Zones and Solar Conditions

No dynamic obstacles were present during this trial. The terrain consisted of dry, homogeneous concrete, and the robot navigated at a constant speed of 0.5 m/s, while the ROS 2 perception node processed data at 10 Hz. Figs. 7 and 8 show the different paths followed by the robot and the visual results for the solar scenarios considered in Table 2, respectively. The different tracks for the robot are:

- *Moving along the central corridor*: The robot navigates through the central concrete aisle, as indicated by the pink path in Fig. 7. These trials were conducted at 08:00, 10:00, 12:00, and 14:00 h to assess the impact of the solar elevation and azimuth parameters defined in Table 2.
- *Moving between plants*: The robot traverses the tomato plants, following the red path shown in Fig. 7. As in the previous case, all solar scenarios from Table 2 were evaluated.
- *Turning at the end of the aisle*: The robot navigates through the crop rows and executes a turn at the end of the corridors to go back along the same corridor, represented by the blue path in Fig. 7. All solar scenarios from Table 2 were tested.
- *Changing to the adjacent corridor*: The robot moves across the crop rows and performs a transition to the next aisle, indicated by the green path in Fig. 7. Similarly, all lighting conditions defined in Table 2 were evaluated.

The algorithm processes the depth information stream from the `/camera/camera/depth` topic, alongside the intrinsic and extrinsic parameters provided by the `/camera/camera/depth/camera_info` topic. The intrinsic parameters utilised were those calibrated by the camera manufacturer, whereas the extrinsic parameters defining the spatial transformation relative to the robot's `base_link` were manually configured at $(0.4, 0, 0.5)$ m. The specific configuration parameters used for the segmentation process are detailed in Table 4.

The results of the trials are shown in the following.

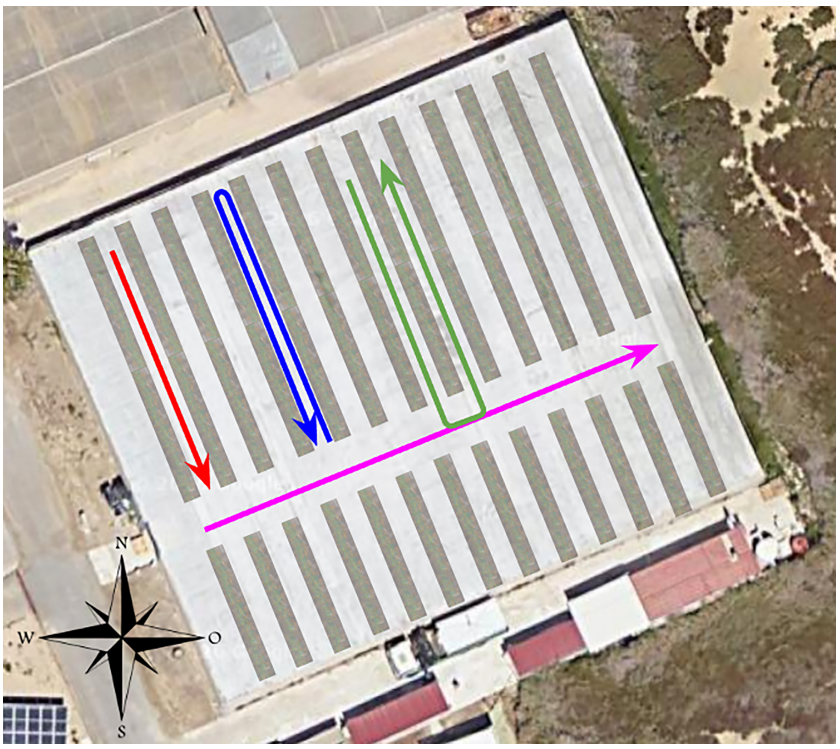
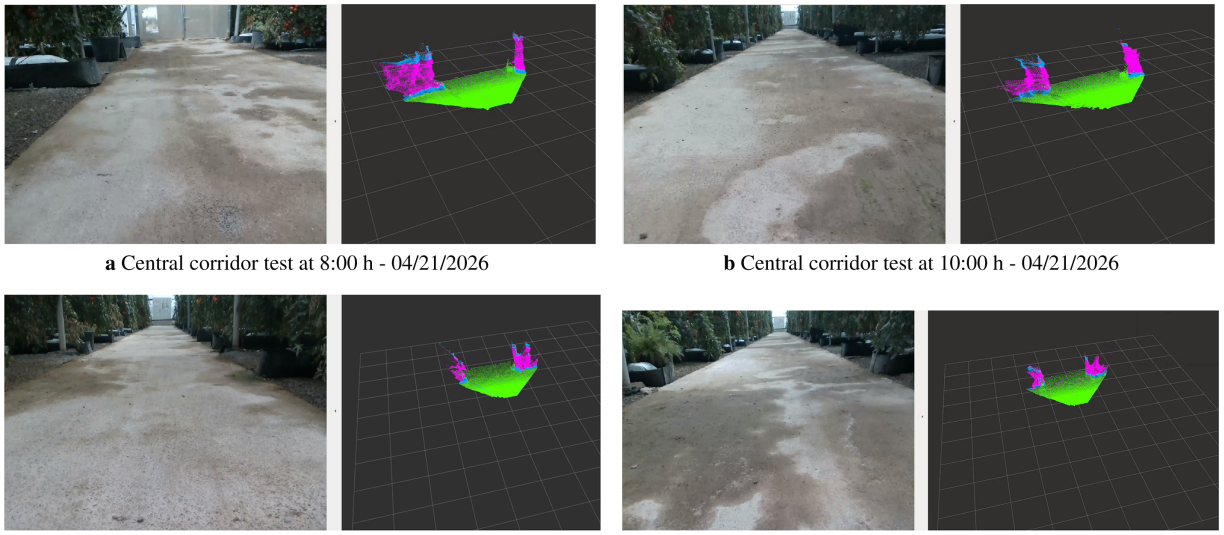


Figure 7: Tests paths. The purple line represents the path moving through the central corridor; the red line represents the path moving through the plants; the blue line represents the path turning at the end of the corridor; the green line represents the path changing to the next corridor.



a Central corridor test at 8:00 h - 04/21/2026

b Central corridor test at 10:00 h - 04/21/2026

c Central corridor test at 12:00 h - 04/21/2026

d Central corridor test at 14:00 h - 04/21/2026

Figure 8: Central corridor test at four solar conditions. Green points indicate the ground, purple indicates obstacles, and cyan indicates points are higher than the robot.

Table 4: GreenSeg algorithm parameters.

Parameters	Value	Description
`max_surface_height` (h_{ground})	0.12 m	Maximum perpendicular tolerance for ground plane
`max_incline`	30.0°	Maximum permissible inclination of the terrain
`robot_height` (h_{robot})	0.5 m	Physical height of the AGRICOBIOT I platform
`n_neighbors`	30	Minimum points for local normal estimation
`r_neighbors` (r)	0.05 m	Neighborhood search radius
`rho_min` (ρ_{min})	0.90	Minimum cosine similarity for normal consistency
`kappa_max` (κ_{max})	0.05	Maximum surface curvature threshold
`r_growing` (r_g)	0.05 m	Spatial adjacency radius for Region Growing
`max_distance_filtered`	3.0 m	Maximum effective range for ground segmentation
`min_distance_filtered`	0.3 m	Minimum depth cutoff to avoid self-occlusion
`Frames per second (fps)`	20 fps	Frames per second that the algorithm can process

5.1.1 Moving along the Central Corridor

No dynamic obstacles were present, the terrain was dry and homogeneous, and the robot navigated at constant speed of 0.5 m/s. The ROS 2 node publishes the images at a rate of 10 Hz. The results at different times from this test are shown in the Fig. 8.

As observed, the central corridor exhibits the most uniform terrain, composed of concrete to facilitate heavy machinery transit. Consequently, the perception framework demonstrates peak robustness in this area. The lateral aisles feature tomato plants in their 20th week of growth (reaching approximately a height of 2 meters) and robot is moving at constant speed of 0.5 m/s. The diurnal experimental analysis is detailed below:

- *Scenario 1 (08:00 h)-Fig. 8a*: Low solar elevation ($\gamma = 30°$) combined with the dense crop canopy prevents direct irradiance on the lens. GreenSeg correctly identifies the navigable plane, effectively filtering minor multipath noise near the corridor edges.
- *Scenario 2 (10:00 h)-Fig. 8b*: Although localized glare begins to emerge in the RGB stream, the normal consistency check ($\rho_{\min}$) prevents these optical artifacts from being misclassified as ground-level obstacles.
- *Scenario 3 (12:00 h)-Fig. 8c*: Despite increased solar irradiance amplified by the polyethylene cladding, the proposed second-layer validation maintains a stable IoU, significantly outperforming the baseline which struggles with floor-level specular reflections.
- *Scenario 4 (14:00 h)-Fig. 8d*: At peak solar intensity ($\gamma = 60°$), the strategic placement of the camera beneath the AGRICOBIOT I transport platform provides passive optical shielding. This physical protection, combined with the Region Growing connectivity constraint, ensures that isolated 'ghost points' are successfully rejected.

Table 5 reports all metrics defined in Section 4.4. These results correspond to the average value of the experiment, with each data point classified within its respective class. This process is repeated for each subsequent experiment.

Table 5: Results between baseline and GreenSeg for Test 1: Moving along the central corridor. Bold values indicate the best performance for each metric and class.

Class k	P_k		R_k		$F_{1,k}$		IoU_k	
	Base	**GreenSeg**	**Base**	**GreenSeg**	**Base**	**GreenSeg**	**Base**	**GreenSeg**
Ground	0.956	**0.970**	0.948	**0.960**	0.952	**0.965**	0.908	**0.933**
Obstacle	0.891	**0.910**	0.915	**0.970**	0.903	**0.939**	0.823	**0.885**
Above	0.935	**0.940**	0.926	**0.930**	0.930	**0.935**	0.870	**0.878**
Noise	0.856	**0.880**	0.815	**0.820**	0.835	**0.849**	0.717	**0.737**
Mean	0.910	**0.925**	0.901	**0.920**	0.905	**0.922**	0.830	**0.858**
Improvement	+1.65%		+2.11%		+1.87%		+3.44%	

The central corridor represents the most structured and least challenging scenario, characterised by flat terrain and minimal overhanging obstacles. Consequently, the baseline algorithm performs optimally, establishing a highly competitive benchmark. Furthermore, this is the scenario in which sunlight causes the most interference due to the refraction of sunlight by the plastic; however, as can be seen, this does not affect the image to any significant extent thanks to the camera's position. Nevertheless, GreenSeg still provides a measurable improvement, enhancing the mean IoU by +3.44% and the obstacle-class Recall by +3.44%. This demonstrates that even in nominal conditions where geometric heuristics are generally sufficient, the proposed normal consistency and region-growing checks refine the segmentation boundaries without degrading the baseline's inherent stability.

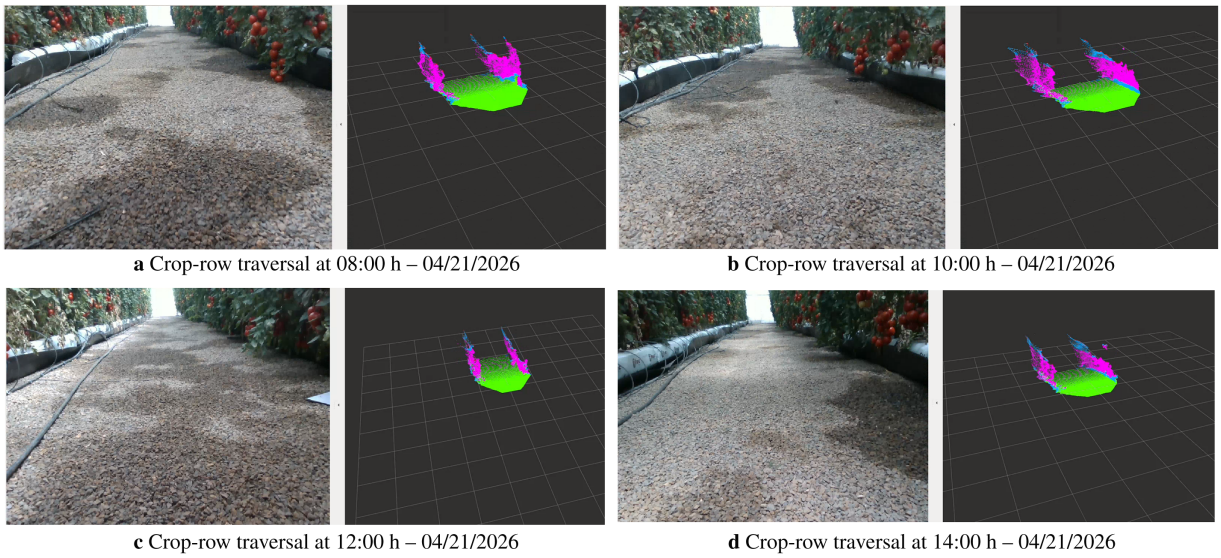
5.1.2 Moving between Plants

This trajectory imposes greater perceptual complexity than the central corridor, as the ground transitions from smooth concrete to irregular gravel and tilled soil. Additionally, the navigable path is partially occluded by the lower canopy of the tomato plant. During this trial, the AGRICOBIOT I platform maintained a constant speed of 0.5 m/s to navigate the narrow 1 m width aisles. The visual results across different solar conditions are illustrated in Fig. 9.

In the case of the reference algorithm, the greater heterogeneity of the terrain on gravel results in localised variations in height that the reference algorithm often misclassifies as obstacles. However, the integration of the surface curvature filter (κ_{max}) into GreenSeg allows for a more flexible, yet geometrically consistent, identification of the ground. The daytime analysis is as follows:

- *Scenario 1 (08:00 h)-Fig. 9a*: The low sun elevation and shadowing from the 2 m high plants provide an ideal contrast. GreenSeg successfully identifies the gravel plane despite the irregular texture.
- *Scenario 2 (10:00 h)-Fig. 9b*: Despite emerging glare, the spatial connectivity constraint of the Region Growing algorithm prevents isolated reflections on the plastic mulch from being segmented as navigable ground.
- *Scenario 3 (12:00 h)-Fig. 9c*: Peak light scattering through the polyethylene cover increases depth noise. GreenSeg compensates for this by using the local normal consistency check (ρ_{min}), filtering artefacts that do not align with the estimated ground plane.
- *Scenario 4 (14:00 h)-Fig. 9d*: The overhead sun ($\gamma = 60°$) creates harsh shadows between plants.

Quantitative results are summarised in Table 6.



a Crop-row traversal at 08:00 h – 04/21/2026 **b** Crop-row traversal at 10:00 h – 04/21/2026

c Crop-row traversal at 12:00 h – 04/21/2026 **d** Crop-row traversal at 14:00 h – 04/21/2026

Figure 9: Crop-row traversal test at four solar conditions. Green points indicate the ground, purple indicates obstacles, and cyan indicates points are higher than the robot.

Table 6: Results between baseline and GreenSeg for Test 2: Moving between crop plants. Bold values indicate the best performance for each metric and class.

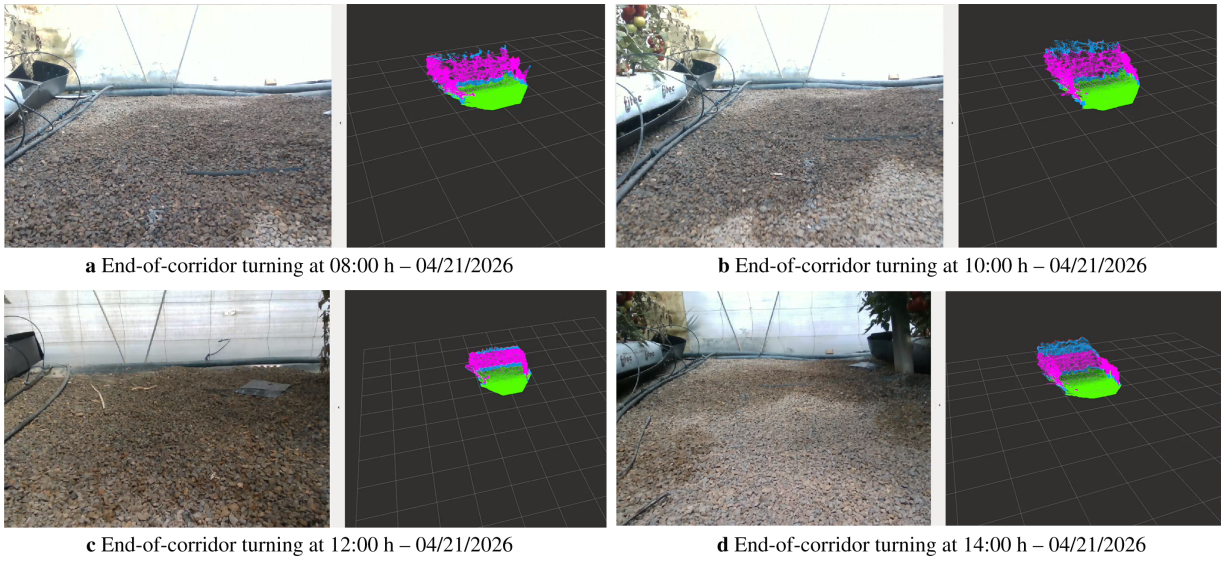
Class k	P_k		R_k		$F_{1,k}$		IoU_k	
	Base	**GreenSeg**	**Base**	**GreenSeg**	**Base**	**GreenSeg**	**Base**	**GreenSeg**
Ground	0.885	**0.935**	0.870	**0.925**	0.877	**0.930**	0.781	**0.869**
Obstacle	0.815	**0.885**	0.830	**0.935**	0.822	**0.909**	0.698	**0.833**
Above	0.860	**0.910**	0.855	**0.895**	0.857	**0.902**	0.750	**0.821**
Noise	0.780	**0.845**	0.740	**0.805**	0.759	**0.824**	0.612	**0.701**
Mean	0.835	**0.894**	0.824	**0.890**	0.829	**0.891**	0.710	**0.806**
Improvement	+7.06%		+8.00%		+7.47%		+13.52%	

In this environment, the baseline's column-wise scanning struggles with depth ambiguities caused by overhanging leaves, dynamic shadows, and irregular soil. By contrast, GreenSeg leverages region-growing connectivity to group leaf masses and continuous ground patches effectively. This architectural advantage is reflected in a substantial +13.52% increase in mean IoU and an +8.00% surge in overall Recall. Most importantly, the obstacle recall peaks at 0.935, proving that the proposed algorithm significantly reduces false negatives in dense vegetation, a critical factor for safe autonomous navigation.

5.1.3 Turning at the End of the Corridor

The end-of-aisle turning manoeuvre constitutes one of the most challenging scenarios for ground segmentation. During this phase, the robot reorients within a confined space while facing the frontal greenhouse openings, which admit unfiltered direct sunlight. This configuration frequently induces severe

lens flare and abrupt depth discontinuities. Additionally, the rotational motion introduces transient noise that challenges temporal consistency. The visual results for this trial are shown in Fig. 10.



a End-of-corridor turning at 08:00 h – 04/21/2026 **b** End-of-corridor turning at 10:00 h – 04/21/2026

c End-of-corridor turning at 12:00 h – 04/21/2026 **d** End-of-corridor turning at 14:00 h – 04/21/2026

Figure 10: End-of-corridor turning test at four solar conditions. Green points indicate the ground, purple indicates obstacles, and cyan indicates points are higher than the robot.

The experimental results for this manoeuvre are highly positive, as the proposed GreenSeg algorithm demonstrates superior capacity to maintain structural coherence compared to the baseline algorithm. By linking the ground plane to the robot's immediate vicinity via the seed point p_s, the system effectively overcomes the transition between shaded and overexposed areas. The diurnal analysis is as follows:

- *Scenario 1 (08:00 h)-Fig. 10a*: Despite the low solar angle, the crop rows act as a natural blind. GreenSeg correctly identifies the transition to the turning area, maintaining high geometric stability with an $F_{1,ground}$.
- *Scenario 2 (10:00 h)-Fig. 10b*: As the robot faces the greenhouse opening, glare artifacts appear. The normal consistency score (ρ_{min}) effectively rejects these virtual points.
- *Scenario 3 (12:00 h)-Fig. 10c*: Under maximum irradiance, the proposed second-layer validation outperforms the baseline, maintaining the segmentation of the navigable surface even when the polyethylene cover amplifies the scattering.
- *Scenario 4 (14:00 h)-Fig. 10d*: The sun is at its highest angle, which will have a direct impact when the robot turns at the end of the corridor.

Quantitative results reported in Table 7 confirm the effectiveness of the approach:

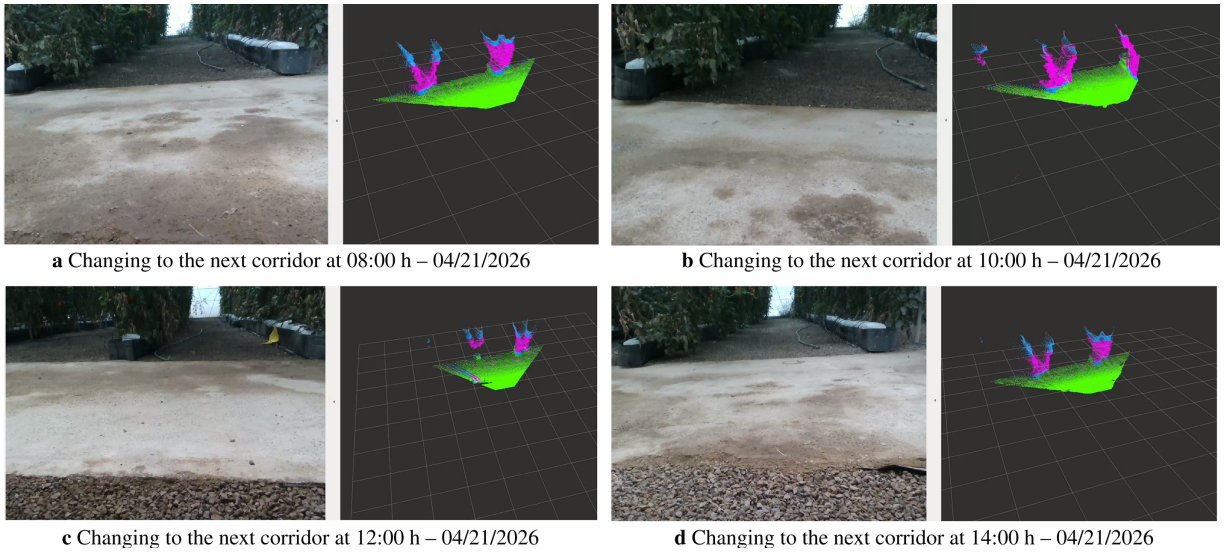
During end-of-aisle turns, the camera's rotation introduces severe motion artifacts and disrupts the assumption of a continuous forward-facing slope, causing the baseline's column-wise logic to fail and fragment the navigable space. In this case, as in the previous one, the refraction of light is not significantly affected by the presence of the crop. GreenSeg mitigates this through global plane fitting, which maintains a mathematical reference of the ground regardless of the camera's yaw rate. This robustness yields the most dramatic performance gains in the study: a +11.72% improvement in mean Recall and a remarkable +19.21% boost in mean IoU, ensuring the local costmap remains coherent precisely when the robot is most susceptible to collisions.

Table 7: Results between baseline and GreenSeg for Test 3: Turning at the end of the corridor. Bold values indicate the best performance for each metric and class.

Class k	P_k		R_k		$F_{1,k}$		IoU_k	
	Base	**GreenSeg**	**Base**	**GreenSeg**	**Base**	**GreenSeg**	**Base**	**GreenSeg**
Ground	0.810	**0.890**	0.805	**0.885**	0.807	**0.887**	0.677	**0.796**
Obstacle	0.740	**0.840**	0.765	**0.890**	0.752	**0.864**	0.602	**0.760**
Above	0.795	**0.855**	0.780	**0.845**	0.787	**0.849**	0.649	**0.738**
Noise	0.690	**0.780**	0.655	**0.735**	0.672	**0.756**	0.506	**0.608**
Mean	0.759	**0.841**	0.751	**0.839**	0.755	**0.839**	0.609	**0.726**
Improvement	+10.81%		+11.72%		+11.13%		+19.21%	

5.1.4 Changing to the Next Corridor

The corridor transition represents a hybrid challenge, as the robot must cross from one crop row to an adjacent aisle via the central corridor. This manoeuvre requires high perceptual flexibility to handle abrupt terrain transitions—moving from irregular sand/gravel to smooth concrete and back to sand—while simultaneously managing the shift from the shaded canopy to the briefly exposed central path. The visual results for this transition across the four solar scenarios are shown in Fig. 11.



a Changing to the next corridor at 08:00 h – 04/21/2026 **b** Changing to the next corridor at 10:00 h – 04/21/2026

c Changing to the next corridor at 12:00 h – 04/21/2026 **d** Changing to the next corridor at 14:00 h – 04/21/2026

Figure 11: Changing to the next corridor test at four solar conditions. Green points indicate the ground, purple indicates obstacles, and cyan indicates points are higher than the robot.

The experimental results for this trajectory are highly positive, demonstrating GreenSeg's superior ability to maintain a stable navigable plane during sudden changes in surface reflectance and local geometry. The diurnal analysis highlights the following strengths:

- *Scenario 1 (08:00 h)-Fig. 11a*: The transition occurs under stable lighting. GreenSeg effectively segments the boundary between the aisle and the corridor, ensuring that the height difference at the concrete edge is not misclassified as an obstacle.

- *Scenario 2 (10:00 h)-Fig. 11b*: Despite glare artifacts emerging as the robot enters the central corridor, the consistency check (ρ_{min}) allows the system to ignore these specularities.
- *Scenario 3 (12:00 h)-Fig. 11c*: At peak irradiance, our method maintains an IoU for the ground class, outperforming the baseline by preventing the "fragmentation" of the ground plane that typically occurs during the material shift.
- *Scenario 4 (14:00 h)-Fig. 11d*: When the sun is at its highest point, this will have a direct effect on the illumination in the corridor due to the absence of plants or other elements that provide shade.

Quantitative results summarised in Table 8 confirm the robustness of the proposed framework:

Table 8: Results between baseline and GreenSeg for Test 4: Changing to the next corridor. Bold values indicate the best performance for each metric and class.

Class k	P_k		R_k		$F_{1,k}$		IoU_k	
	Base	**GreenSeg**	**Base**	**GreenSeg**	**Base**	**GreenSeg**	**Base**	**GreenSeg**
Ground	0.895	**0.925**	0.880	**0.910**	0.887	**0.917**	0.797	**0.846**
Obstacle	0.840	**0.880**	0.855	**0.915**	0.847	**0.897**	0.734	**0.813**
Above	0.880	**0.915**	0.875	**0.905**	0.877	**0.909**	0.781	**0.833**
Noise	0.805	**0.845**	0.765	**0.800**	0.784	**0.821**	0.645	**0.696**
Mean	0.855	**0.891**	0.843	**0.882**	0.848	**0.886**	0.739	**0.797**
Improvement	+4.21%		+4.62%		+4.48%		+7.84%	

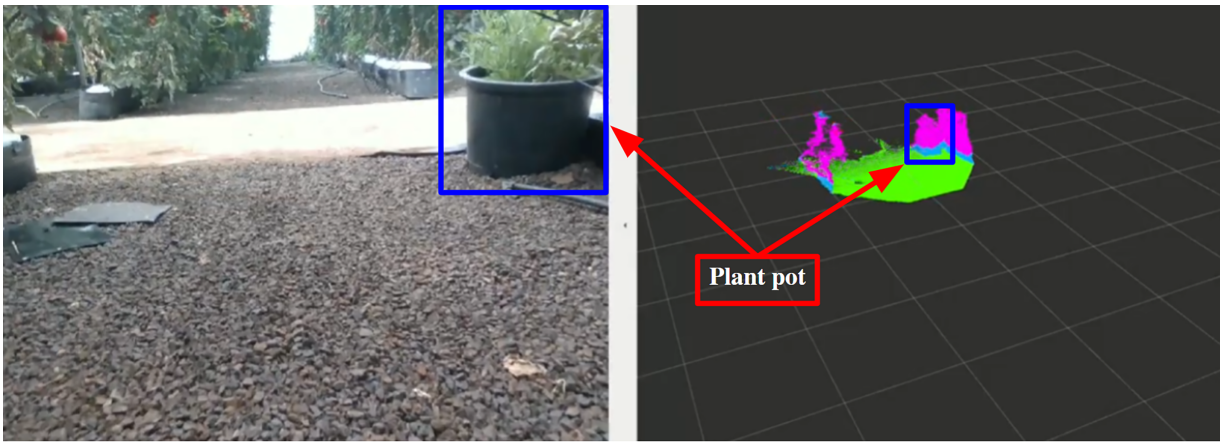
This scenario represents a hybrid linear movement but heterogeneous terrain, often including heating pipes or changes in lighting near the greenhouse headers. Furthermore, the algorithm also takes into account the effect of solar refraction, and it performs well. While the baseline recovers from the turning manoeuvre, GreenSeg maintains a clear superiority, delivering a +4.62% increase in mean Recall and a +7.84% improvement in mean IoU. These results confirm that the algorithm swiftly adapts to abrupt material and structural shifts, providing a resilient semantic classification of the environment as the robot re-enters the dense crop rows.

5.2 Visual Performance under Static and Dynamic Obstacles

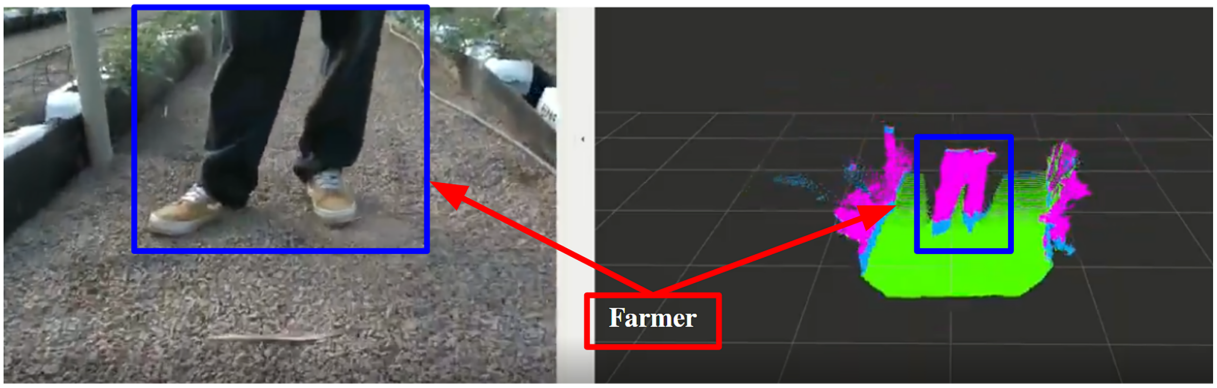
Having quantitatively verified the algorithm's performance across the terrain states reported in Section 5 (Tables 5–7), a complementary qualitative analysis was conducted to illustrate the algorithm's behaviour in the presence of static and dynamic obstacles. To this end, two additional trials were conducted: the first involved placing a plant pot directly in the robot's path to ensure it fell within its field of view, while the second trial featured the appearance and movement of a farmer within the greenhouse. Fig. 12a,b illustrates the point cloud outputs produced by the algorithm in the presence of these obstacles.

It should be noted that these two scenarios were not subjected to the quantitative protocol. This follows directly from the automatic ground-truth generation procedure itself: a label is only accepted into $\mathcal{D}_{\text{GT}}$ if it remains consistent in at least ψ_{min} of the $W = 10$-frame temporal window within a spatially registered voxel. This criterion is well suited to a static scene under continuous robot motion, but it is not designed to accommodate independently moving objects, whose occupied voxels change from frame to frame regardless of the robot's own ego-motion. Applying the automatic labeling procedure to these obstacle scenarios would therefore predominantly discard points as undefined rather than yield a statistically meaningful ground-truth reference. A dedicated, motion-aware ground-truth strategy for obstacle scenarios (e.g., object-level

tracking or manual annotation) falls outside the scope of the automatic procedure proposed in this work and is left as future work. These two trials are accordingly presented as a qualitative illustration of point cloud behaviour under obstacle conditions, complementary to the quantitative terrain evaluation.



a Static obstacles – Plant pot



b Dynamic obstacle – farmer

Figure 12: Performance under encountering obstacles.

As observed in the figures, the point clouds qualitatively suggest that the system correctly segments both obstacles as non-navigable terrain, even while the farmer is actively moving. Taken together with the quantitative terrain trials, which cover transitions between wet tilled soil and smooth concrete, these results support the framework's applicability across the terrain and obstacle conditions encountered in agricultural settings, while a fully quantitative treatment of obstacle scenarios, supported by a motion-aware ground-truth strategy, is identified as future work.

The following results were drawn within the scope of the test studies:

- The proposed GreenSeg algorithm consistently outperformed the baseline segmentation framework across all experimental trajectories, maintaining stable real-time execution at 20 FPS on resource-constrained embedded hardware. The most significant gains were observed during rotational manoeuvres at the end of the corridors, where the system achieved improvements of +11.72% in mean Recall and +19.21% in mIoU relative to the baseline. These results indicate that combining RANSAC global plane fitting with local region-growing helps define the boundary between ground and obstacles with higher geometric fidelity, which is expected to reduce collision risk in narrow-aisle environments.

- A key contribution of this work is the mitigation of optical and structural artifacts inherent in greenhouse sensing. The integration of seed-point-based geometric validation (p_s) combined with curvature filtering (κ_{max}) was effective in reducing "ghost points" and false-positive obstacle detections. These artifacts, typically caused by direct solar glare through frontal openings, specular reflections from the polyethylene cladding, or overhanging leaves, are largely filtered by requiring each segmented point to maintain spatial and normal consistency with the robot's localized footprint.
- The algorithm showed consistent adaptability to the heterogeneous terrains typical of Mediterranean greenhouses, particularly when navigating between crop rows (+13.52% mIoU improvement over the baseline). By utilising local surface normal estimation and region-growing connectivity, the system maintained structural consistency across the smooth concrete, gravel, and tilled soil surfaces evaluated in this work. This supports the continuity of the navigable plane across the material transitions tested, and suggests that the selected hyperparameters ($\kappa_{\max}$, neighborhood radius) generalise reasonably well across these conditions without environment-specific recalibration, though this has only been verified within the tested scenarios.
- During end-of-aisle turning manoeuvres, the algorithm handled the transient noise and depth discontinuities introduced by the platform's rotational motion. Unlike column-wise scanning methods, which are more sensitive to angular velocity, the spatial connectivity constraint used here restricts ground classification to points physically linked to the robot's base. This helped maintain a usable local costmap in the tested scenarios, including areas where unfiltered sunlight introduced depth saturation near the greenhouse exits.
- Across the four evaluated phases, GreenSeg maintained a stable navigable-plane estimate under the agricultural sensing conditions tested, including scenarios where the baseline method degraded due to dynamic shadows, irregular soil, and overhanging leaves, as well as during end-of-aisle turns. These results suggest that the algorithm is comparatively resilient to the lighting interference and terrain variability present in the evaluated greenhouse, providing a reasonably noise-resilient terrain representation across the phases tested.
- The robustness tests conducted in this work cover several agricultural conditions relevant to the target application—including terrain transitions, working farmers, and static/dynamic obstacles—and suggest that the algorithm could extend reasonably well to related settings with irregular terrain, curved crop rows, or variable inter-row spacing. Validating this generalisation for less structured environments beyond the greenhouse layout tested here remains an open question for future work.
- The results indicate that a low-cost RGB-D camera is adequate for the ground segmentation task addressed in this work, within the operating conditions and layout of the tested greenhouse. The primary engineering contribution of GreenSeg lies in the real-time integration of these spatial-geometric processing steps within an asynchronous ROS 2 node architecture, running on resource-constrained embedded hardware. This supports the feasibility of low-cost RGB-D-based perception for the specific navigation task evaluated, though no direct experimental comparison against LiDAR-based systems was performed in this work; establishing whether GreenSeg's approach can match or substitute LiDAR-based perception in this domain would require such a comparison and is left as future work.

6 Conclusion

This study has addressed the challenge of ground segmentation in Mediterranean greenhouse environments, where variable lighting, high vegetation density, and terrain heterogeneity typically degrade conventional robotic perception systems. The proposed GreenSeg algorithm demonstrated a significant improvement in mitigating the negative impact of lens flare and "ghost points" resulting from direct solar irradiance and reflections on polyethylene cladding. A fundamental aspect of this robustness is the system's

effectiveness when operating between crop rows, where lighting conditions shift constantly due to dynamic shadows projected by the plants and irregular light filtration.

By implementing a dual-consistency framework—combining RANSAC plane fitting with a normal-verified region-growing algorithm—the system successfully maintained the structural integrity of the navigable plane even under critical high-dynamic-range conditions and abrupt rotational movements. Experimental trials conducted on diverse surfaces, including smooth concrete, gravel, and wet tilled soil, validated that the inclusion of local curvature analysis (κ) allows the AGRICOBIOT I robot to adapt to irregular floor textures without triggering false positives. Furthermore, the framework's overall robustness was successfully validated against both static and dynamic obstacles within the field. The system exhibited a highly stable behaviour when encountering typical agricultural elements such as crates, farming machinery, and plant pots placed directly in the field of view, as well as during the active movement of working farmers, correctly classifying these entities as non-navigable terrain.

This advancement is essential for ensuring fluid mobility during complex tasks, such as navigating between dense plants and executing end-of-aisle turns. Ultimately, reaching peak improvements of +19.21% in mIoU and +11.72% in mean Recall during critical turning manoeuvres over baseline methods underscores that the proposed framework provides a resilient and economically viable perception foundation for Mediterranean agriculture. Crucially, the architectural value of GreenSeg lies in its ability to operate effectively under a single, static parametric configuration across all these adverse conditions. This ensures that the algorithm is not uniquely restricted to this specific scenario but remains fully scalable to other less-structured agricultural environments—such as fields with highly irregular terrain, curved crop rows, or variable inter-row spacing—since the experimental validation in this work successfully addresses the most challenging agricultural and sensing conditions a mobile robot can experience. Therefore, the use of this type of low-cost hardware, when paired with robust geometric filtering, constitutes a pivotal contribution to mobile robotics, providing high scalability, ease of deployment, and a reliability that rivals expensive 3D LiDARs. Despite its robustness, one limitation of this study is that it relies on the explicit availability of an area of ground near the robot to successfully initialise the starting points (p_s) for region growth. Consequently, the algorithm's performance could be reduced if the camera's immediate field of view were completely obstructed by dense, low-lying vegetation resting directly on the ground. Furthermore, given the dusty and muddy environment, the sensor lens may, under conditions of intensive greenhouse operation, present a hardware-dependent vulnerability that is not addressed by geometric filtering.

As a primary avenue for future research, the integration of the GreenSeg segmentation output directly into an autonomous navigation framework as a dynamic local costmap is envisioned. This approach will facilitate the establishment of a complete and robust autonomous navigation system in Mediterranean greenhouses—a solution based on low-cost hardware that is currently unavailable due to the extreme structural and lighting complexities of these facilities. The ability to transform noisy point clouds between crops into reliable spatial representations for path planning will represent a milestone in the digitisation of agricultural tasks, enabling fully autonomous operations in environments where traditional systems consistently fail.

Acknowledgement: None.

Funding Statement: This work has been carried out within the framework of the LIFE-ACCLIMATE project (LIFE23-CCAES-LIFE-ACCLIMATE/101157315), cofunded by the European Union under LIFE program Grant Agreement No. LIFE23-CCA-ES-LIFE-ACCLIMATE/101157315. The first author, Fernando Cañadas-Aránega, holds an FPI grant (PRE2022-102415) from the Spanish Ministry of Science, Innovation, and Universities.

Author Contributions: The authors confirm contribution to the paper as follows: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Original draft preparation, Visualisation, Project administration and Formal writing: Fernando Cañadas-Aránega; Methodology, Investigation, Formal analysis, Writing—review and editing, Project administration, Supervision: José C. Moreno; Supervision: José L. Blanco-Claraco. All authors reviewed and approved the final version of the manuscript.

Availability of Data and Materials: The open repository for this work is at this link: https://github.com/FerCanAra/ros2_greenseg.

Ethics Approval: Not applicable.

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

References

1. Duckett T, Pearson S, Blackmore S, Grieve B, Chen WH, Cielniak G, et al. Agricultural robotics: the future of robotic agriculture. arXiv:1806.06762. 2018.
2. Vougioukas SG. Agricultural robotics. Annu Rev Control Robot Auton Syst. 2019;2(1):365–92.
3. Cañadas-Aránega F, Mañas-Álvarez FJ, L-Guzmán J, Moreno JC, Blanco-Claraco JL. A ROS2 benchmarking framework for hierarchical control strategies in mobile robots for Mediterranean greenhouses. arXiv:2602.15162. 2026.
4. FAO. How to feed the world in 2050. 2024 [cited 2026 Sep 8]. Available from: https://www.fao.org/fileadmin/templates/wsfs/docs/expert_paper/How_to_Feed_the_World_in_2050.pdf.
5. Cañadas-Aránega F, Moreno JC, Blanco-Claraco JL, Giménez A, Rodríguez F, Sánchez-Hermosilla J. Autonomous collaborative mobile robot for greenhouses: design, development, and validation tests. Smart Agric Technol. 2024;9:100606.
6. Lytridis C, Kaburlasos VG, Pachidis T, Manios M, Vrochidou E, Kalampokas T, et al. An overview of cooperative robotics in agriculture. Agronomy. 2021;11(9):1818. doi:10.3390/agronomy11091818.
7. Cañadas-Aránega F, Moreno JC, Blanco-Claraco JL. A PID-based control architecture for mobile robot path planning in greenhouses. IFAC-PapersOnLine. 2024;58(7):503–8. doi:10.1016/j.ifacol.2024.08.112.
8. Cañadas-Aránega F, Wollherr D, Guzmán JL, Moreno JC, Blanco JL. An adaptive control architecture for slope and terrain compensation in autonomous navigation in Mediterranean greenhouses. arXiv:2609.02487. 2026.
9. Firkat E, An F, Peng B, Zhang J, Mijit T, Ahat A, et al. FGSeg: field-ground segmentation for agricultural robot based on LiDAR. Comput Electron Agric. 2023;211:107965.
10. Lee S, Lim H, Myung H. Patchwork++: fast and robust ground segmentation solving partial under-segmentation using 3D point cloud. In: Proceedings of the 2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS); 2022 Oct 23–27; Kyoto, Japan. p. 13276–83.
11. Song Y, Xu F, Yao Q, Liu J, Yang S. Navigation algorithm based on semantic segmentation in wheat fields using an RGB-D camera. Inform Process Agric. 2023;10(4):475–90. doi:10.1016/j.inpa.2022.05.002.
12. Luo L, Yin W, Ning Z, Wang J, Wei H, Chen W, et al. In-field pose estimation of grape clusters with combined point cloud segmentation and geometric analysis. Comput Electron Agric. 2022;200:107197.
13. Jin S, Camps A, Jia Y, Wang F, Martin-Neira M, Huang F, et al. Remote sensing and its applications using GNSS reflected signals: advances and prospects. Satellite Navig. 2024;5(1):19. doi:10.1186/s43020-024-00139-4.
14. Archana R, Jeevaraj PE. Deep learning models for digital image processing: a review: R. Archana, PS Eliahim Jeevaraj. Artif Intell Rev. 2024;57(1):11. doi:10.1007/s10462-023-10631-z.
15. Mijit T, Firkat E, Yuan X, Liang Y, Zhu J, Hamdulla A. LR-Seg: a ground segmentation method for low-resolution LiDAR point clouds. IEEE Trans Intell Veh. 2023;9(1):347–56.
16. Tarsha-Kurdi F, Landes T, Grussenmeyer P. Hough-transform and extended RANSAC algorithms for automatic detection of 3D building roof planes from lidar data. In: ISPRS Workshop on Laser Scanning 2007 and SilviLaser 2007; 2007 Sep 12–14; Espoo, Finland. p. 407–12.

17. Paigwar A, Erkent Ö, Sierra-Gonzalez D, Laugier C. GndNet: fast ground plane estimation and point cloud segmentation for autonomous vehicles. In: Proceedings of the 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS); 2020 Oct 24–Jan 25; Las Vegas, NV, USA. p. 2150–6.
18. Lim H, Oh M, Myung H. Patchwork: concentric zone-based region-wise ground segmentation with ground likelihood estimation using a 3D LiDAR sensor. IEEE Robot Autom Lett. 2021;6(4):6458–65. doi:10.1109/lra.2021.3093009.
19. Zhang W, Qi J, Wan P, Wang H, Xie D, Wang X, et al. An easy-to-use airborne LiDAR data filtering method based on cloth simulation. Remote Sens. 2016;8(6):501. doi:10.3390/rs8060501.
20. Galin E, Guérin E, Peytavie A, Cordonnier G, Cani MP, Benes B, et al. A review of digital terrain modeling. Comput Graph Forum. 2019;38(2):553–77.
21. Upadhyay A, Chandel NS, Singh KP, Chakraborty SK, Nandede BM, Kumar M, et al. Deep learning and computer vision in plant disease detection: a comprehensive review of techniques, models, and trends in precision agriculture. Artif Intell Rev. 2025;58(3):92. doi:10.1007/s10462-024-11100-x.
22. Yang C, Sun L, Zhang J, Fan X, Zhang D, Ren T, et al. Research on the quantification and automatic classification method of Chinese cabbage plant type based on point cloud data and PointNet++. Front Plant Sci. 2025;15:1458962. doi:10.3389/fpls.2024.1458962.
23. Yang Y, Wijewardane NK, Slonecki TJ, Wadl PA, Andreason SA, Chen J, et al. Nondestructive detection of sweet potato leaf curl virus using 3D laser imaging combined with deep learning. Smart Agric Technol. 2025;11:101004. doi:10.1016/j.atech.2025.101004.
24. Cañadas-Aránega F, Muñoz M, Moreno J, Blanco-Claraco J. Integrated cloud-based architecture for robot-robot and human-robot collaboration using ROS 2–MQTT in Mediterranean Greenhouses. arXiv:2606.22682. 2026.
25. Wang H, Shan Y, Chen L, Liu M, Wang L, Meng Z. Multi-scale feature learning for 3D semantic mapping of agricultural fields using UAV point clouds. Int J Appl Earth Obs Geoinf. 2025;141:104626. doi:10.1016/j.jag.2025.104626.
26. Yang X, Xu C, Wang Y, Feng R, Yu J, Su Z, et al. PlaneSegNet: a deep learning network with plane attention for plant point cloud segmentation in agricultural environments. Artif Intell Agric. 2026;16(1):284–99.
27. Hu Q, Yang B, Xie L, Rosa S, Guo Y, Wang Z, et al. Learning semantic segmentation of large-scale point clouds with random sampling. IEEE Trans Pattern Anal Mach Intell. 2021;44(11):8338–54. doi:10.1109/tpami.2021.3083288.
28. Boulch A, Guerry J, Le Saux B, Audebert N. SnapNet: 3D point cloud semantic labeling with 2D deep segmentation networks. Comput Graph. 2018;71:189–98. doi:10.1016/j.cag.2017.11.010.
29. Cañadas-Aránega F, Blanco-Claraco JL, Moreno JC, Rodriguez-Diaz F. Multimodal mobile robotic dataset for a typical Mediterranean greenhouse: the GREENBOT dataset. Sensors. 2024;24(6):1874. doi:10.3390/s24061874.
30. Qian X, Ye C. NCC-RANSAC: a fast plane extraction method for 3-D range data segmentation. IEEE Trans Cybern. 2014;44(12):2771–83.
31. Lentsch T, Montiel-Marín S, Caesar H, Gavrila DM. TerraSeg: self-supervised ground segmentation for any LiDAR. arXiv:2603.27344. 2026.
32. Milioto A, Lottes P, Stachniss C. Real-time semantic segmentation of crop and weed for precision agriculture robots leveraging background knowledge in CNNs. In: Proceedings of the 2018 IEEE International Conference on Robotics and Automation (ICRA); 2018 May 21–25; Brisbane, QLD, Australia. p. 2229–35.
33. Arafin T, Hosen A, Najdovski Z, Wei L, Rokonuzzaman M, Johnstone M. Advances and trends in terrain classification methods for off-road perception. J Field Robotics. 2025;42(7):3515–44. doi:10.1002/rob.22586.
34. Kang H, Wang X. Semantic segmentation of fruits on multi-sensor fused data in natural orchards. Comput Electron Agric. 2023;204:107569. doi:10.1016/j.compag.2022.107569.
35. He Y, Zhang X, Zhang Z, Fang H. Automated detection of boundary line in paddy field using MobileV2-UNet and RANSAC. Comput Electron Agric. 2022;194:106697. doi:10.1016/j.compag.2022.106697.
36. De Silva R, Cielniak G, Wang G, Gao J. Deep learning-based crop row detection for infield navigation of agri-robots. J Field Robot. 2024;41(7):2299–321. doi:10.1002/rob.22238.
37. Macenski S, Foote T, Gerkey B, Lalancette C, Woodall W. Robot operating system 2: design, architecture, and uses in the wild. Sci Robot. 2022;7(66):eabm6074.

38. Blanco-Claraco JL, Tymchenko B, Mañas-Alvarez FJ, Cañadas-Aránega F, López-Gázquez Á, Moreno JC. MultiVehicle simulator (MVSim): lightweight dynamics simulator for multiagents and mobile robotics research. SoftwareX. 2023;23:101443.
39. Macenski S, Martín F, White R, Clavero JG. The marathon 2: a navigation system. In: Proceedings of the 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS); 2020 Oct 24–2021 Jan 21; Las Vegas, NV, USA. p. 2718–25.
40. Intel Corporation. 3D Pointcloud Groundfloor Segmentation for RealSense™ Camera and 3D LiDAR; Open Edge Platform Documentation. 2024 [cited 2026 Apr 30]. Available from: https://docs.openedgeplatform.intel.com/dev/edge-ai-suites/robotics-ai-suite/robotics/dev_guide/tutorials_amr/perception/pointcloud-groundfloor-segmentation.html.
41. Adams R, Bischof L. Seeded region growing. IEEE Trans Pattern Anal Mach Intell. 1994;16(6):641–7. doi:10.1109/34.295913.
42. Xu Y, Ghamisi P. Consistency-regularized region-growing network for semantic segmentation of urban scenes with point-level annotations. IEEE Trans Image Process. 2022;31:5038–51. doi:10.1109/tip.2022.3189825.
43. Jafarbiglu H, Pourreza A. Impact of sun-view geometry on canopy spectral reflectance variability. ISPRS J Photogramm Remote Sens. 2023;196:270–86. doi:10.1016/j.isprsjprs.2022.12.002.
44. Honkavaara E, Markelin L, Rosnell T, Nurminen K. Influence of solar elevation in radiometric and geometric performance of multispectral photogrammetry. ISPRS J Photogramm Remote Sens. 2012;67:13–26. doi:10.1016/j.isprsjprs.2011.10.001.