Simposio CEA de Robótica, Bioingeniería, Visión Artificial y Automática Marina 2025

Agriculture Surface Edge Explorer (AgriSEE): Active reconstruction of occluded tomatoes in greenhouses

Cañadas-Aránega, Fernando[a,*], Border, Rowan[b], Blanco-Claraco, José L.[c], Moreno-Úbeda, José C.[a]

[a] *University of Almeria, Department of Informatics, CIESOL, ceia3, Ctra. Sacramento s/n, 04120, Almería, Spain. (fernando.ca, jcmoreno)@ual.es*
[b] *Vision for Robotics Lab (V4RL), University of Cyprus, Kallipoleos 75, Nicosia 1678, Cyprus. border.rowan@ucy.ac.cy*
[c] *University of Almeria, Department of Engineering, CIESOL, ceia3, Ctra. Sacramento s/n, 04120, Almería, Spain. jlblanco@ual.es*

Resumen

La agricultura intensiva bajo invernadero se ha convertido en uno de los pilares del crecimiento demográfico de la sociedad. Sin embargo, a medida que pasan los años, el aumento de la superpoblación tanto en los humanos como en el mundo animal, requiere que la agricultura existente deba ser más eficiente y sostenible. En esta búsqueda, la automatización en general, y la robótica en particular, juegan un papel fundamental, ya que son herramientas para la resolución óptima de algunos problemas claves en este campo, relacionados con el desarrollo de tareas tediosas, sucias y/o peligrosas, las tareas denominadas *DDD* (del inglés *Dull*, *Dirty*, *Dangerous*). Este trabajo se centra en la aplicación de un algoritmo de *Next Best View* (NBV), en concreto SEE+, para la obtención de un modelo 3D de nube de puntos haciendo uso de un brazo robot UR10 de Universal Robot. El trabajo describe como usar una cámara que cuenta con tecnología LiDAR para construir un modelo 3D de los frutos, que será usado posteriormente para la recolección automatizada en el interior de invernaderos. Para ello, este algoritmo NBV clasifica la nube de puntos para determinar la próxima mejor visión y obtener un rápido escaneo del fruto para su posterior recolección. Los resultados muestran como ha sido posible detectar un tomate para el que, previamente y con técnicas tradicionales, era imposible conocer con certeza su ubicación.

Palabras clave: Robótica en agricultura, Invernaderos, Siguiente mejor vista, ROS 2, SLAM

Agriculture Surface Edge Explorer (AgriSEE): Active reconstruction of occluded tomatoes in greenhouses

Abstract

Intensive greenhouse agriculture has become one of the pillars of society's population growth. However, increasing overpopulation in both the human and animal world requires existing agriculture to become more efficient and sustainable. Automation in general, and robotics in particular, play a fundamental role in addressing these challenges, particularly for 'dull', 'dirty' and 'dangerous' tasks. This work focuses on applying a *Next Best View* (NBV) algorithm, SEE+, to obtain a 3D point cloud model using a Universal Robot UR10 robot arm. The work describes how to use a camera with LiDAR technology to generate a 3D model for fruits, which will later be used for the automated harvesting inside greenhouses. The NBV algorithm classifies the point cloud to determine the next best view and can efficiently obtain a scan of the fruit for subsequent harvesting. The results show how it has been possible to determine with precision the location of a tomato, something difficult to do using traditional techniques.

Keywords: Agricultural robotics, Greenhouses, Next Best View, ROS 2, SLAM

1. Introduction

The global area of greenhouses now exceeds 490,000 ha, with an estimated annual growth rate of 20% since 1980. Around 20% of this area is located in the southeastern Iberian Peninsula of Spain, where the total area dedicated to vegetable crops exceeds 77,000 ha (Trenda, 2023). This sector has a high social relevance in the greenhouse farming industry of the peninsula, generating approximately 100,000 direct jobs and around 25,000 indirect ones. Spanish greenhouses represent the Mediterranean-type greenhouse, which account for 92% of the global greenhouse area and typically features a low or medium level of technological development (Moreno et al.,

[*]Corresponding author: fernando.ca@ual.es

2024; Cañadas-Aránega et al., 2024a). However, this sector faces competition from systems with highly advanced technologies in developed countries, such as the Netherlands, and low-tech systems with significantly lower associated costs, as in Morocco or Turkey. It is essential to improve productivity and quality in order for Spanish greenhouses to remain competitive. Moreover, the growing demand for food — both for humans and animals — combined with the labour shortage in rural areas and the increasing interest in autonomous systems, partially driven by the COVID-19 pandemic, further justifies the exploration of such technologies (Moreno Úbeda et al., 2022).

Greenhouses exhibit a certain degree of structural organization but differ significantly from controlled environments in industrial settings, such as automotive production lines. Highly automated machinery is crucial for addressing this challenge to ensure the development and advancement of greenhouse agriculture. One of the key considerations when deploying robots in this context is the use of suitable sensors that must be capable of detecting obstacles and localizing the robot within greenhouses, which are highly dynamic environments with limited communication (Bac et al., 2013; Ko et al., 2014; Aranega et al., 2024; Blanco-Claraco et al., 2023). These systems can perform various tasks beyond navigation, such as automated crop harvesting. Maximizing the efficiency of greenhouse surfaces is a fundamental strategy to ensure optimal productivity and it is essential to adapt the various task-specific algorithms to operate effectively under this constraint.

The use of robots designed specifically for greenhouse environments has been researched since 1987. During the crop season in these agricultural settings up to 80% of the total time is spent on monitoring and fruit harvesting, which has been one of the main tasks studied (Cañadas-Aránega et al., 2024b; Sánchez-Molina et al., 2024). The most successful developments have employed robotic arms to perform automated harvesting tasks. These robots are typically equipped with RGB-D cameras to detect shapes or depth levels in the scene and/or LiDAR sensors, which enable 3D scanning and mapping of the environment (Rong et al., 2022).

Environmental mapping is crucial in harvesting tasks as it determines the fruit's 3D position. This positioning is used by various algorithms that identify and label objects within the greenhouse. In Rong et al. (2022), the YOLOv5 algorithm (Ge et al., 2021) is used to detect a tomato cluster from a 2D image and estimate the position of the fruit. In Zheng et al. (2024), the YOLOv5_SE variant is employed to identify tomatoes using RGB and depth images. In Liu et al. (2024), SLAM techniques are used to determine the distance between fruits with measurements from a 2D LiDAR. However, accurately determining the number of fruits inside a greenhouse remains challenging as some are occluded and can not be detected by these technologies. This motivates the development of algorithms that can estimate such occlusions and ensure efficient fruit harvesting.

It is important to have a 3D model of the fruit as this allows the robotic arm to plan and execute appropriate trajectories. However, obtaining high-quality 3D observations is challenging regardless of the final application. A scene (i.e. a bounded region of space) is observed by combining individual 3D measurements of surfaces taken from different viewpoints. An observation is considered complete when sufficient measurement coverage is obtained from all visible surfaces. The final surface coverage depends on the sensor capabilities, scene structure and viewpoints from which the measurements are acquired Border et al. (2018). Algorithmic view selection reduces human uncertainty by intelligently choosing the most informative viewpoints. This challenge of planning a subsequent view that provides the most significant improvement in a scene observation is known as the *Next Best View* (NBV) problem.

This work presents an approach for scanning different views of a tomato cluster and reconstructing a 3D model from the resulting point cloud, based on the NBV algorithm presented in Border et al. (2018). The simulation setup in that work consists of a fixed platform where the target objects are placed, and an independent UR10 robotic arm equipped with an Intel RealSense L515 camera mounted on the end-effector. The object of study in the simulation experiments is a 3D model of a tomato plant, representative of the typical crops cultivated in Almería's greenhouses. The experimental results confirm that SEE is capable of acquiring high-quality observations, providing a solid foundation for the future development of algorithms aimed at detecting occluded fruits.

This paper is organized as follows. Section 2 provides an overview of the project in which this study is framed and details of the simulation experiments. Section 3 presents the simulation results for the models derived from the experimental setup. Finally, Section 4 outlines the main conclusions of this research.

2. Materials and Methods

This section outlines the materials and methods used in the development of this work.

2.1. Materials

This subsection describes the materials employed during the execution of the project.

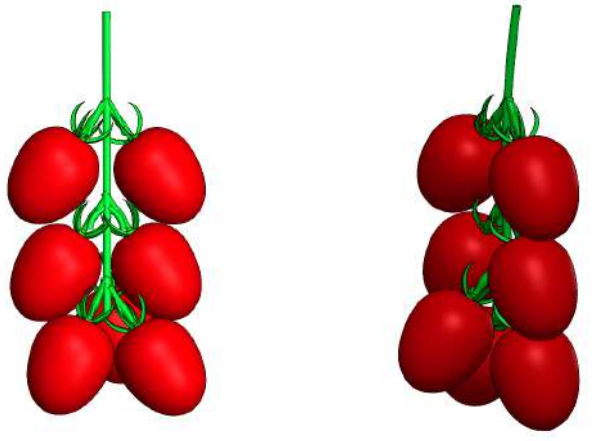
2.1.1. 3D Tomato Model

A 3D model of the tomato fruit cultivated in the experimental greenhouse of Agroconnect was used for the simulation. These facilities[1] are located in La Cañada de San Urbano, Almería. The greenhouse is part of the Institute for Research and Training in Agriculture, Fisheries, Food, and Organic Production of Andalusia (IFAPA), next to the University of Almería. It is a Mediterranean-type greenhouse, typical of the region, with an area of 1,850 square meters, a robust steel structure, and a polyethylene cover.

The accurate modeling of the greenhouse and its components is essential for programming the various tasks to be performed inside. The models required for the simulation were generated based on real data from the greenhouse, as presented in Cañadas-Aránega et al. (2024). This work aims to scan a tomato plant to identify occluded tomatoes. Figure 1 shows a

[1]Funded by the Ministry of Science, Innovation and Universities in collaboration with the European Regional Development Fund (ERDF) through the 2019 call for advanced scientific and technological equipment.

3D model of a tomato plant that meticulously reproduces all aspects of its structure, including the various shapes of its leaves and stems.



a Frontal view b Rear isometric view

Figure 1: 3D model of the tomato plant

2.1.2. *UR10 Robotic Arm*

The UR10 robotic manipulator from Universal Robots was used in the simulation. This 6 DoF robotic arm is designed for industrial applications that require precision and repeatability. The UR10 supports payloads up to 10 kg and has a maximum reach of 1300 mm, making it suitable for inspection, 3D scanning, and object manipulation. The robot was controlled using the UR Robot Operating System (ROS) interface within the simulated environment, allowing complete control and monitoring of its movements (Figure 2).

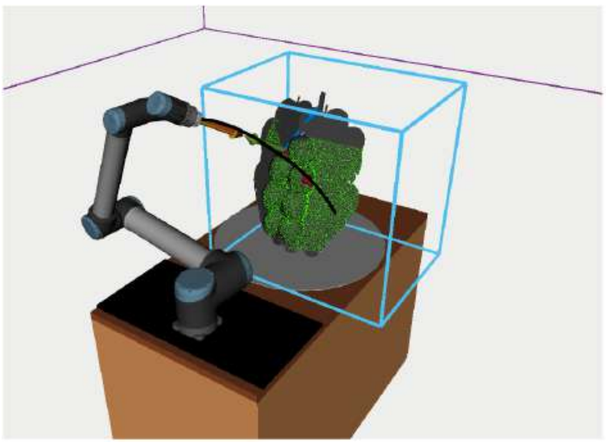


Figure 2: UR10 with Intel RealSense L515 mounted on the end-effector

2.1.3. *Intel RealSense L515*

An Intel RealSense L515 camera was employed to capture depth data. This depth-sensing device is based on time-of-flight (ToF) technology and is especially suitable for indoor environments. It offers high accuracy when measuring distances from reflective and dark surfaces. The sensor provides a depth resolution of up to 1280×720 pixels at 30 fps and an RGB resolution of 1920×1080 pixels. The LiDAR technology integrated into the camera was used for this experiment (Figure 3). The camera was mounted on the end-effector of the robot arm, enabling dynamic scanning of 3D scenes during motion. Data acquisition and synchronization with the UR10 were managed using the Intel RealSense SDK 2.0 and ROS.

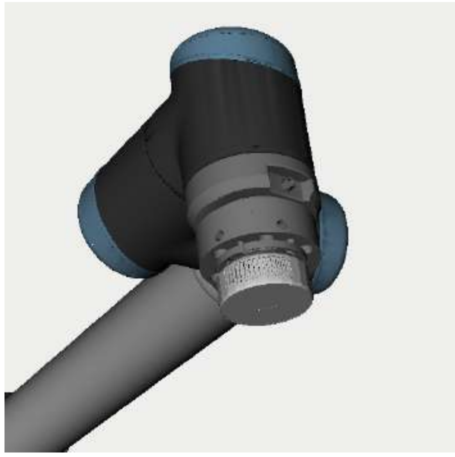


Figure 3: Intel RealSense L515

2.2. *Methods*

This subsection describes the methods employed during the execution of the project.

2.2.1. *Surface Edge Explorer (SEE)*

SEE is a measurement-direct NBV approach that makes viewpoint planning decisions directly from sensor measurements to achieve a minimum required point density. It performs surface scanning by classifying each point individually based on the density of neighboring measurements (Figure 4). Points located at the boundary between regions are classified as frontier points. The algorithm proposes viewpoints to acquire new measurements around these frontier points. Initial viewpoints are selected based on the local surface geometry but can be refined to proactively avoid known occlusions. Observation efficiency is maximized by selecting views that cover the most frontier points while minimizing motion. If a viewpoint fails, it is reactively adjusted to avoid unknown occlusions or discontinuities not taken into account in previous iterations. The observation process concludes once all frontier points are observed or deemed unobservable (Border and Gammell, 2024).

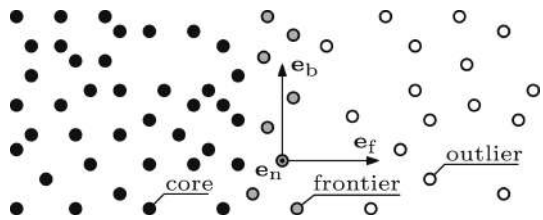


Figure 4: Illustration of the orthogonal vectors used by SEE to represent the local surface geometry. The normal vector, $\mathbf{e}_n$ is perpendicular to the local plane (out of page), the frontier vector, $\mathbf{e}_f$, points toward the partially observed region, and the boundary vector, $\mathbf{e}_b$, lies along the boundary between fully and partially observed regions (Border and Gammell, 2024).

The experiment evaluates the reconstructed model obtained by applying SEE to scan a tomato plant. The tested approaches were tuned to maximize surface coverage using the fewest viewpoints and the shortest time. Proper scanning of a tomato plant requires adapting SEE to the specific geometry of the fruit. The proposed viewpoints for observing frontier points on the tomato are based on the cylindrical shape of the surface normals, so accurate model calibration before the experiment is essential. Each LiDAR measurement is classified as core (c), frontier (f), or outlier (o), using the algorithm from Border et al. (2018). For frontier points (f), their surface normal, $\mathbf{e}_\mathrm{n}$, must point outwards from the surface to be considered valid viewpoints, as shown in Figure 5. After multiple scans, SEE stores the points and merges them using the Iterative Closest Point (ICP) algorithm to reconstruct the final 3D object (Border et al., 2024).

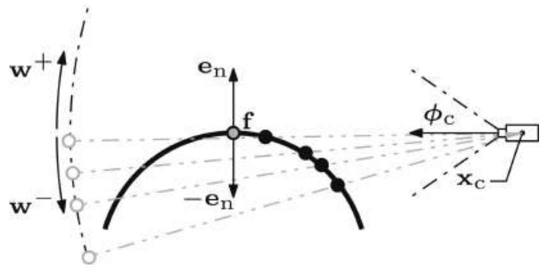


Figure 5: Illustration of how the correct normal direction is determined. Normal vectors $\mathbf{e}_\mathrm{n}$ and $-\mathbf{e}_\mathrm{n}$ are evaluated for visibility from the current viewpoint $\mathbf{v}_\mathrm{c} = (\mathbf{x}_\mathrm{c}, \phi_\mathrm{c})$. The correct vector is the one not occluded by nearby surface measurements (black dots). It is obtained by projecting occluding points onto a sphere and searching along both projected directions, $\mathbf{w}^+$ and $\mathbf{w}^-$, until free space is found (Border and Gammell, 2024).

2.2.2. *ROS and RViz*

The robotic system was developed using ROS Noetic, the latest distribution compatible with Ubuntu 20.04 LTS and aimed at x86_64 architectures. ROS Noetic provided the communication infrastructure between nodes, enabling the integration of the UR10 robotic arm, the Intel RealSense L515 camera, and the data processing modules within the simulation. The project architecture relied on topics, services, and custom messages to synchronize the robot's kinematics with visual and depth data.

RViz, the official visualization tool for ROS, was used to visualize the environment, the robot's 3D model, and sensor data. The URDF models of the UR10 and the point cloud streams from the RealSense camera were loaded, allowing real-time monitoring of the robotic arm's movements and environment scanning. Additionally, visual markers were used to represent trajectories, key positions, and points of interest.

3. Results

Several experiments were conducted on the tomato plant model developed in this study and here we present the one that yielded the best results. The model was observed within a robotic arm simulation environment. The setup consists of a UR10 robotic arm with an RGB-D camera mounted on the end-effector and a table on which the plant branch was placed (Figure 6). The center of the table and the base of the UR10 are separated by 0.75 m. The table offers a usable volume of 0.8 m; therefore, smaller models were adjusted to fit within a bounding box of $0.8 \times 0.8 \times 0.6$ m. The maximum model height is 0.6 m to allow the end-effector to access top-down views of the model.

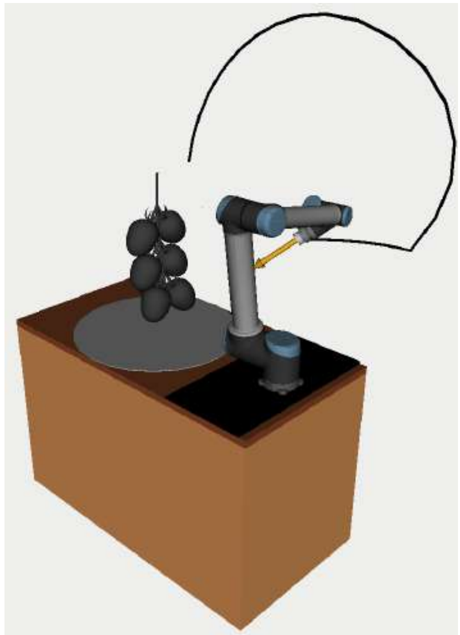


Figure 6: Simulation environment

For this scanning process, the parameters defined in (Border and Gammell, 2024) are maintained, as they have given the best results. These values are averaged over the total number of independent experiments using the tomato plant model. In this context, a total of 20 experiments were carried out, showing in this paper the one with the best results. The simulated camera is defined by its field of view in degrees, θ_x and θ_y, and its resolution in pixels, ω_x and ω_y (Table 1). The sensor measurements are obtained by projecting rays onto the triangulated surface of the model and adding Gaussian noise ($\mu = 0$ m, $\sigma = 0.01$ m) to the ray intersections. This noise magnitude was selected to represent the measurement noise associated with depth cameras (e.g. Intel RealSense L515) (Border and Gammell, 2024).

Table 1: Intel RealSense L515 properties

Property	**RGB camera**	**Units**
θ_x	70	degrees
θ_y	43	degrees
ω_x	848	pixels
ω_y	480	pixels

Figure 7 shows the simulation environment in operation, where the arm traces its path with a black line, shows the view normals, displays the blue bounding cube, and classifies the scanned points in green. SEE successfully scans a high percentage of the plant, including occluded surfaces. The result of the scan is shown in Figure 8.

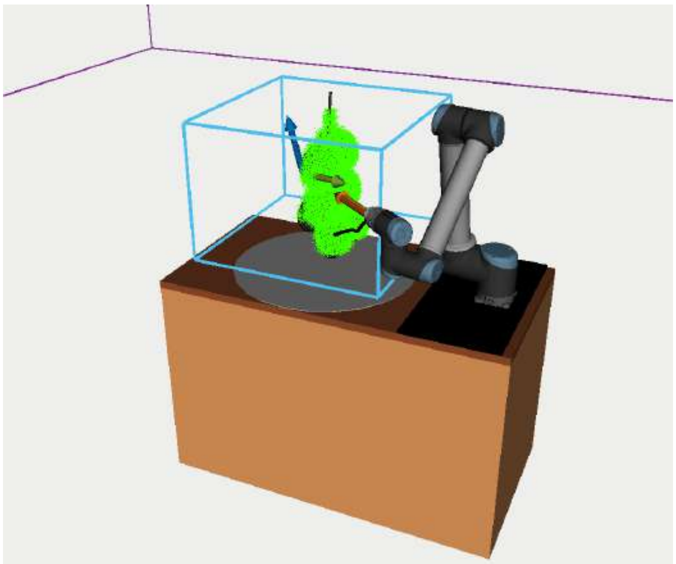
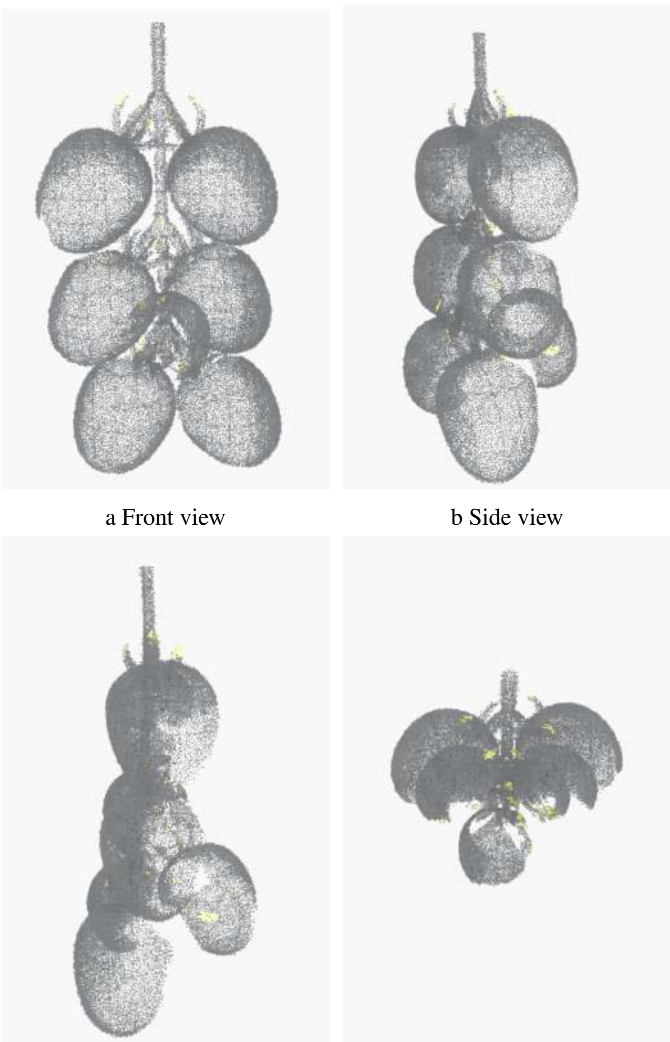


Figure 7: Working simulation environment



a Front view

b Side view

c Isometric view

d Top view

Figure 8: .PLY model with σ= 0.01 m

Observation performance is quantified in terms of the surface coverage obtained, the travel distance required and the time taken to capture a complete observation. This model achieved a surface coverage of approximately 60% compared to the ground truth, with a total travelled distance of 0.5 m and a scanning time of 1 minute. As shown, the model includes information about occluded tomatoes due to the multiple views obtained by SEE, some examples are front view (Figure 8a), side view (Figure 8b), isometric view (Figure 8c), and top view (Figure 8d). This information is crucial for planning efficient harvesting tasks, as it enables the detection of hidden objects that would otherwise be impossible to perceive using traditional 2D sensor-based techniques.

In this *link*, you can observe the movements of the robot arm with SEE in Rviz, from its initial position until the algorithm finishes, resulting in the complete 3D model.

4. Conclusion

This work contributes to the advancement of agricultural robotics by implementing an innovative NBV approach aimed at obtaining more complete and accurate 3D models of tomato clusters. The simulation experiments, employing a UR10 robotic arm and an Intel RealSense L515 camera, validate the feasibility of capturing multiple perspectives of a 3D tomato plant model to generating a dense and detailed point cloud. This type of observation is essential for detecting occluded fruits, which can not be perceived using conventional RGB imaging or 2D sensing techniques.

The SEE algorithm has proven particularly effective in planning views that maximize surface coverage and reconstruction quality, thereby reducing reliance on predefined camera positions or manual interventions. These richer 3D observations pave the way for future research on automated trajectory planning for robotic arms during harvesting tasks, optimizing execution time and process accuracy.

In summary, these results lay the groundwork for developing autonomous robotic systems capable of operating efficiently in complex agricultural environments. This work represents a significant first step towards improved precision agriculture in greenhouses, leveraging robust, adaptive, and intelligent technologies focused on fruit harvesting — a task that accounts for approximately 80% of the time spent in crop production.

Acknowledgements

This work has been carried out within the framework of the LIFE-ACCLIAMTE project (LIFE23-CCA-ES-LIFE-ACCLIMATE/101157315), co-funded by the LIFE Programme of the European Union, CyberGreen Project, PID2021-122560OB-I00, and the AgroConnect.es facilities, grant EQC2019-006658-P, both funded by MCIN/AEI/ 10.13039/501100011033 and by 'ERDF A way to make Europe'. The first author, Fernando Cañadas-Aránega, holds an FPI grant (PRE2022-102415) from the Spanish Ministry of Science, Innovation, and Universities. The authors would like to thank Professor Dr. Margarita Chli, who was the supervising host for Fernando's visiting stay at the University of Cyprus.

References

Aranega, F. C., Claraco, J. L. B., Alvarez, F. J. M., Úbeda, J. C. M., 2024. Navegación de un robot ackermann para tareas de transporte en invernaderos mediterráneos con multivehicle simulator (mvsim). Jornadas de Automática (45).

Bac, C., Hemming, J., Van Henten, E., 2013. Robust pixel-based classification of obstacles for robotic harvesting of sweet-pepper. Computers and electronics in agriculture 96, 148–162.

Blanco-Claraco, J.-L., Tymchenko, B., Mañas-Alvarez, F. J., Cañadas-Aránega, F., López-Gázquez, Á., Moreno, J. C., 2023. Multivehicle simulator (mvsim): Lightweight dynamics simulator for multiagents and mobile robotics research. SoftwareX 23, 101443.

Border, R., Chebrolu, N., Tao, Y., Gammell, J. D., Fallon, M., 2024. Osprey: Multi-session autonomous aerial mapping with lidar-based slam and next best view planning. IEEE Transactions on Field Robotics.

Border, R., Gammell, J. D., 2024. The surface edge explorer (see): A measurement-direct approach to next best view planning. The International Journal of Robotics Research 43 (10), 1506–1532.

Border, R., Gammell, J. D., Newman, P., 2018. Surface edge explorer (see): Planning next best views directly from 3d observations. In: 2018 IEEE International Conference on Robotics and Automation (ICRA). IEEE, pp. 6116–6123.

Cañadas-Aránega, F., Moreno, J. C., Blanco-Claraco, J. L., 2024a. A pid-based control architecture for mobile robot path planning in greenhouses. IFAC-PapersOnLine 58 (7), 503–508.

Cañadas-Aránega, F., Moreno, J. C., Blanco-Claraco, J. L., Giménez, A., Rodríguez, F., Sánchez-Hermosilla, J., 2024b. Autonomous collaborative mobile robot for greenhouses: Design, development, and validation tests. Smart Agricultural Technology, 100606.

Cañadas-Aránega, F., Blanco-Claraco, J. L., Moreno, J. C., Rodriguez-Diaz, F., 2024. Multimodal mobile robotic dataset for a typical mediterranean greenhouse: The greenbot dataset. Sensors 24 (6).
URL: `https://www.mdpi.com/1424-8220/24/6/1874`
DOI: `10.3390/s24061874`

Ge, Z., Liu, S., Wang, F., Li, Z., Sun, J., 2021. Yolox: Exceeding yolo series in 2021. arXiv preprint arXiv:2107.08430.

Ko, M. H., Ryuh, B.-S., Kim, K. C., Suprem, A., Mahalik, N. P., 2014. Autonomous greenhouse mobile robot driving strategies from system integration perspective: Review and application. IEEE/ASME Transactions On Mechatronics 20 (4), 1705–1716.

Liu, K., Yu, J., Huang, Z., Liu, L., Shi, Y., 2024. Autonomous navigation system for greenhouse tomato picking robots based on laser slam. Alexandria engineering journal 100, 208–219.

Moreno, J., Rodríguez, F., Sánchez-Hermosilla, J., Giménez, A., Sánchez-Molina, J., 2024. Feasibility analysis of robots in greenhouses. a case study in european mediterranean countries. Smart Agricultural Technology 9, 100638.

Moreno Úbeda, J. C., Cañadas-Aránega, F., Rodríguez, F., Sánchez-Hermosilla, J., Giménez, A., 2022. Modelado 3d y diseño de un robot colaborativo para tareas de transporte en invernaderos. In: XLIII Jornadas de Automática. Universidade da Coruña. Servizo de Publicacións, pp. 785–791.

Rong, J., Wang, P., Wang, T., Hu, L., Yuan, T., 2022. Fruit pose recognition and directional orderly grasping strategies for tomato harvesting robots. Computers and Electronics in Agriculture 202, 107430.

Sánchez-Molina, J., Rodríguez, F., Moreno, J., Sánchez-Hermosilla, J., Giménez, A., 2024. Robotics in greenhouses. scoping review. Computers and Electronics in Agriculture 219, 108750.

Trenda, E., 2023. Greenhouse agricultural area in spain in 2022, by type of crop. `https://www.statista.com/statistics/1218871/greenhouse-area-spain-by-crop/`, accessed: 22 April 2023.

Zheng, X., Rong, J., Zhang, Z., Yang, Y., Li, W., Yuan, T., 2024. Fruit growing direction recognition and nesting grasping strategies for tomato harvesting robots. Journal of Field Robotics 41 (2), 300–313.