Orchard digital twin: A prototype for smart agricultural monitoring

PROCEEDINGS BOOK

ISBN: 978-618-82194-1-0

Published by:

Hellenic Society of Agricultural Engineers

Website: https://egme.gr/

Editor:
Nikolaos Katsoulas, University of Thessaly, Greece

Editorial Supervision:
CONVIN S. A.
29 Kosta Varnali Str., 15233 Halandri, Athens, Greece
Tel.: +30 210 6833600
E-mail: info@convin.gr
Website: http://www.convin.gr
ISBN: 978-618-82194-1-0

Publication: February 2025

Agricultural Engineering challenges in existing and new agroecosystems

1-4 July 2024, Agricultural University of Athens, Greece

Organizing Committee:

General Chair: Nikolaos Katsoulas, University of Thessaly, Greece

Technical Chair: Thomas Bartzanas, Agricultural University of Athens, Greece

Spyros Fountas, Agricultural University of Athens, Greece
George Papadakis, Agricultural University of Athens, Greece
Thomas Kotsopoulos, Aristotle University of Thessaloniki, Greece
Pantazis Georgiou, Aristotle University of Thessaloniki, Greece
George Xanthopoulos, Agricultural University of Athens, Greece
Konstantinos Ferentinos, Hellenic Agricultural Organization "Dimitra", Greece
Chryssoula Papaioannou, University of Thessaly, Greece

Table of Contents

Orchard digital twin: A prototype for smart agricultural monitoring

Kyra Smith[a*], Andrea Botta[b], Giovanni Colucci[b], Marco Piras[a], Giuseppe Quaglia[b], Elena Belcore[a]

[a] Department of Environment, Land and Infrastructure Engineering, Politecnico di Torino, 10129 Torino, Italy

[b] Department of Mechanical and Aerospace Engineering, Politecnico di Torino, 10129 Torino, Italy

* Corresponding author. Emails: name.surname@polito.it

Abstract

This study presents a method for creating detailed digital twins (DT) of apple orchards using an unmanned ground vehicle (UGV) and a combination of sensors, including a handheld laser scanner (HLS), a low-cost depth camera, and a Global Navigation Satellite Systems (GNSS) receiver. We aim to provide an updatable 3D geospatial database to automate agricultural processes, such as apple monitoring and collection. The workflow was tested in a portion of an apple orchard in northwest Italy and is comprised of four phases. Phase 1 involves sensor-integration on the vehicle and data acquisition. Phase 2 defines and stores the orchard's geometries in a georeferenced 3D model; this phase also includes the segmentation of individual trees within each row. Phase 3 detects and segments apples using an artificial intelligence (AI) algorithm applied to RGB images captured by the depth camera; segmented apples are then projected onto the 3D model. Our work demonstrates how unmanned ground vehicles (UGVs) integrated with sensors can be applied to create a detailed, updatable orchard DT, a tool which can inform and automate agricultural tasks, ultimately increasing efficiency and reducing waste.

Keywords: smart farming, agricultural digital twin, mobile robotics, artificial intelligence, crop monitoring

1. Introduction

The optimisation of agricultural systems is vital as they face growing demands and pressures. One widespread approach to achieving this is precision agriculture (PA). PA uses spatiotemporal data from agricultural systems to guide management decisions and improve productivity, sustainability, and resource-use efficiency (ISPA, 2024). In the 1990s, PA was hailed as “an information technology revolution”. Today, agriculture is experiencing another revolution, called “Agriculture 4.0”, which includes the rise of smart farming (Javaid et al., 2022). Smart farming goes a step further than PA, utilising information and communications technologies (ICT), such as artificial intelligence (AI) and Internet of Things (IoT), to automate and optimise not only agricultural operations but overall farm management, leading to even greater efficiency and productivity (Kamilaris et al., 2016; Moysiadis et al., 2021).

The employment of unmanned ground vehicles (UGVs) in agriculture is on the rise thanks to its synergy with the smart farming paradigm, which depends heavily on extensive and repetitive data collection to perform informed, precise agricultural tasks. In this context, a UGV can assist farmers by effortlessly performing repetitive tasks without losing accuracy throughout the day.

Besides a limited number of robotic UGVs developed to autonomously perform agricultural tasks such as weeding, pruning, spraying, and harvesting (Visentin et al., 2023), most agricultural UGVs are entirely dedicated to sensing and monitoring of the crop or the soil. In this sense, UGVs navigate through the field as mobile sensor carriers to collect data such as soil structure and composition, water content, plant and fruits condition, presence of pests or diseases, crop yield, and more (Vulpi et al., 2022). Different combinations of sensors can be mounted on the UGV at different times to collect all relevant data, making modular robotic platforms very flexible and effective (Tiozzo Fasiolo et al., 2023). This ability to carry multiple sensors or robotic components gives UGVs an advantage over unmanned aerial vehicles (UAVs), which are more mobile but limited by the weight of their payload (Tardaguila et al., 2021; Apeinans et al., 2023).

The synergy between smart farming and robotics enables the collection of vast amounts of crop data, which must be organised, processed, and stored. When the collected information describes the system exhaustively, it can be organised and stored in a Digital Twin (DT).

A DT is comprised of three parts: (1) the physical component, (2) the virtual component, and (3) the flow of information that connects them (Grieves, 2015). The European Commission’s Destination Earth program (Destination Earth, 2024) aims to expand the use of DTs to describe the processes of our planet and the related human activities, including agricultural systems. Agricultural DTs are accurate digital replicas of real agricultural systems, continuously updated by smart farming techniques and sensors. They have been proposed to simulate machinery performance (Tsolakis et al., 2019), provide a decision support system for aquaponics (Ghandar et al., 2021), and to improve water efficiency by automating irrigation systems (Alves et al., 2023), among other applications.

Nonetheless, adoption of DTs in agriculture has been found to be lagging behind other sectors, with most agricultural DTs still in the prototype stage (Pylianidis et al., 2021). Furthermore, most of the agricultural DTs reviewed were of non-living systems (such as machinery or structures) rather than living systems, likely due to their complexity (Pylianidis et al., 2021). For the few DTs of living systems, the reported benefits included early detection of disease and other threats, lower costs, and higher product quality (Pylianidis et al., 2021).

Orchards are an ideal agricultural setting for testing methods of generating DTs of living systems, as the objects of interest (principally trees) are sessile, can be differentiated from one another, and are often organised in recognisable patterns. These qualities allow for repeatable data collection and data storage at the tree-level. Potential applications of orchard DTs include yield estimation, monitoring fruit development, detecting pests or diseases, autonomous harvesting, and more. To date, there are few examples of orchard DTs and, to the knowledge of these authors, none that propose an exhaustive workflow combining UGV technology and sensors to build a georeferenced DT of an apple orchard.

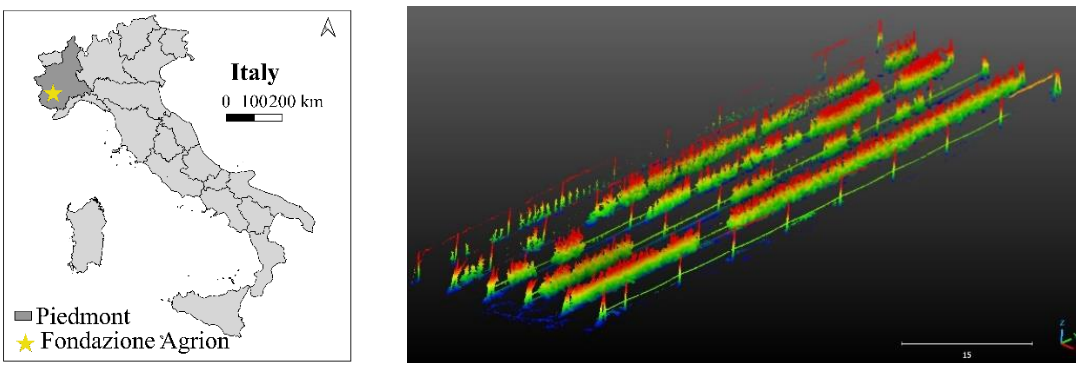
This work seeks to develop a workflow for the creation of a high-quality DT of an apple orchard. This requires the achievement of several sub-objectives, which are: (1) to test the performance of the UGV 'Agri.Q' in an orchard environment and the arrangement of sensors on Agri.Q; (2) to create a 3D model of the orchard and segment individual trees in order to increase our resolution from a row- to tree-level; (3) to train a deep learning model to segment apples.

2. Materials and Methods

This work was carried out in three phases that correspond to our sub-objectives.

Phase 1: Data acquisition

Data acquisition was carried out on October 10, 2023, in an apple orchard belonging to the Research Centre for Fruit Cultivation (Centro Ricerche per la Frutticoltura) of Fondazione Agrion, located in Manta, Piedmont, Italy (Fig.1(a)). The study area was comprised of flat, grassy terrain with three parallel rows of mature apple trees, each row about 90m long and 4m apart (Fig.1(b)). At the time of acquisition, the trees were fruit-bearing. Several apple varieties were represented. 16 black and white geometric markers were placed at varying heights throughout the rows to be used as references during point cloud construction.



(a) (b)

Figure 1. (a) The location of Fondazione Agrion in Piedmont, Italy and (b) orchard HLS point cloud.

Agri.Q represents an innovative UGV specifically engineered for PA applications within vineyards and orchards. This UGV demonstrates adaptability to unstructured environments and uneven terrain thanks to its unique architecture. As depicted in Fig. 2, the UGV consists of two skid-steering modules, each equipped with two locomotion units that drive two tires each via a chain drive system. This architectural design ensures efficiency comparable to traditional wheeled systems while effectively distributing pressure on the ground, akin to tracked vehicles. Consequently, this design feature, coupled with the UGV's relatively modest weight of approximately 110kg, reduces the tires sinkage and mitigates soil degradation in comparison to conventional agricultural machinery.

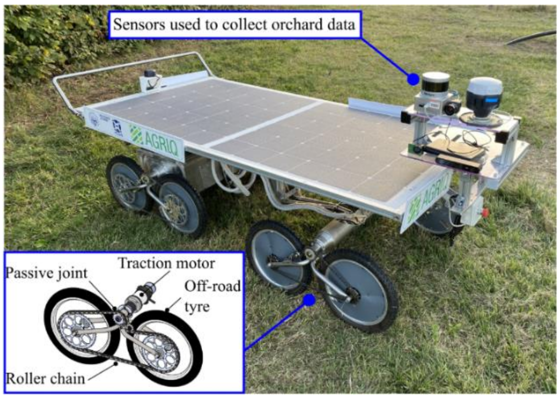


Figure 2. The Agri.Q UGV has four skid-steer modules, each characterised by a single drive motor that transmits power to two off-road wheels via a chain transmission. The passive joint allows Agri.Q to adapt to uneven and complex terrain.

Agri.Q can be equipped with an array of instruments and sensors tailored to specific tasks such as field mapping and crop monitoring. Notably, it can feature a redundant 7 degrees of freedom (DOF) collaborative robot arm, facilitating interactions with the environment, including soil, leaf, and fruit sample collection (Quaglia et al., 2024). Moreover, it incorporates a 2 DOF photovoltaic (PV) panel capable of self-orientation to optimise solar energy collection for battery recharge, but, also, to provide an always level landing platform for drones to seamlessly collaborate with UAVs when required (Visconte et al., 2021; Botta and Cavallone, 2022). Additionally, the robotic arm and the sensor unit are mounted on the frame of the orientable platform, enabling dynamic adaptation of the robot's manipulation workspace (or the sensors' field of view) to various tasks and scenarios (Colucci et al., 2023), as showed in Fig. 3. Regarding the specific application in apple orchards, Agri.Q has been equipped with a sensor support plate positioned on the rear module of the UGV so that its pitch adjustment mechanism can raise the sensors from a height of 700 mm to 1600 mm. It's worth noting that a hinge joint has been incorporated to keep the plate parallel to the ground, even when the Agri.Q panel is tilted.

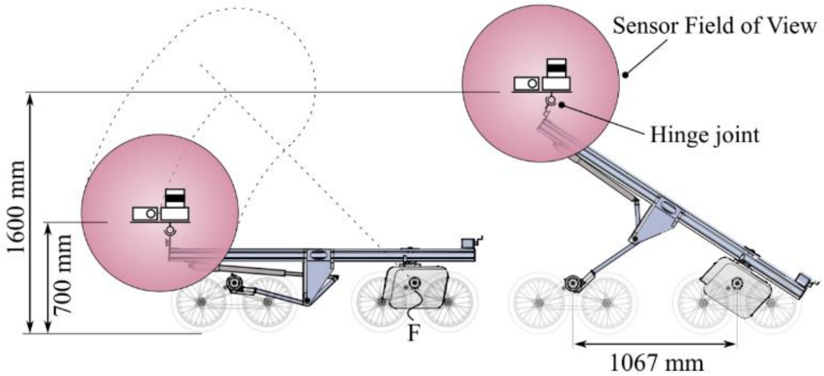


Figure 3. The sensors carried on the rear plate can be raised by means of the pitch mechanism adjustment, which imposes to the Agri.Q panel a rigid rotation around point F. The hinge joint between the plate and the panel allows to keep the plate parallel to the ground.

Three optical sensors were used, in addition to a dual frequency GNSS receiver used to georeference the DT (Table 1). The Kaarta Stencil 2 is a HLS that uses Velodyne LiDAR and advanced simultaneous localisation and mapping (SLAM) algorithms that integrate visual odometry (Velodyne Lidar, 2023). The Stencil 2 was integrated with an Emlid Reach 2 GNSS receiver in Network Real Time Kinematic (NRTK) acquisition to georeference the trajectory estimated by Stencil 2. The ZED 2 is a stereo camera produced by StereoLabs. It collected images with a resolution of 1920 x 1080 at a rate of 30 frames per second (fps). The Mapir 3N collects images in the red (R), green (G) and near-infrared (NIR) bands with 1 Hz frequency and it has a field of view of 41°.

Sensor Name	Sensor Type	Data Collected	Information	CRS
Kaarta Stencil 2	Handheld Laser Scanner (HLS)	Point cloud of orchard Point cloud of HLS trajectory	X, Y, Z Intensity values Time Confidence value (trajectory only)	Local
Emlid Reach 2	GNSS Receiver	HLS position	East, North, elevation above the Ellipsoid Root Mean Squared Errors	Global, EPSG: 32632
ZED 2	Stereo Camera	RGB images Depth images Colourised point cloud of orchard	Left RGB (stereo pairs) Right RGB (stereo pairs) Depth images	Local
Mapir 3N	Multispectral Camera	Multispectral images (R, G, NIR)	3-bands images: Near Infrared 850nm, Red 660nm, and Green 550nm	Local

Table 1. Sensors used to collect orchard data. CRS= Coordinate Reference System

All sensors were affixed to Agri.Q's sensor support plate using screws. The HLS, stereo camera, and multispectral camera were oriented such that the image plane was parallel to the row of trees (Fig. 4). Horizontal and vertical offsets between the centres of the sensors were recorded. The GNSS receiver was connected to the HLS via USB. All sensors were activated before the UGV began to move and they collected data continuously throughout the acquisition. During the test, the UGV was commanded to navigate between three rows at a constant speed ($0.2 m \cdot s^{-1}$).

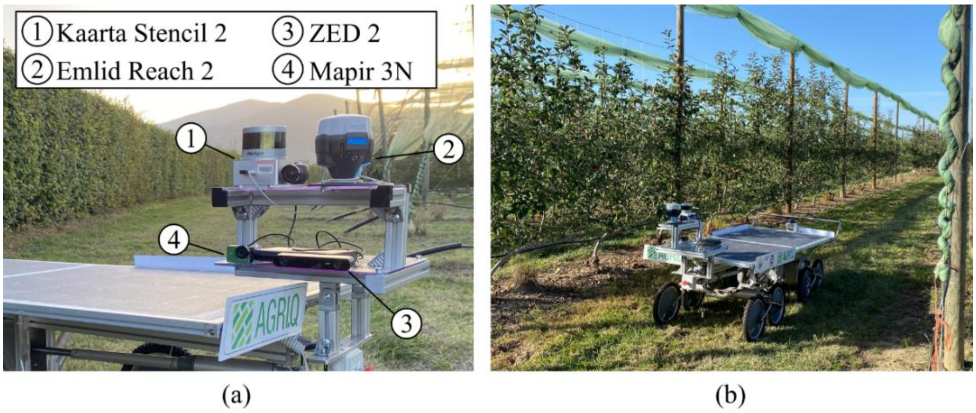


Figure 4. (a) The arrangement of the sensors on Agri.Q platform. (b) Agri.Q while navigating the apple orchard.

Phase 2: Data processing and information extraction

The HLS was initially optimized using the Adaptive Data Replay tool provided by Stencil 2, which allows for the re-processing of the point cloud by adjusting parameters that control feature matching (Di Pietra et al., 2020) and the Loop Closure algorithm. The communication between the SLAM and GNSS systems enabled the time synchronization of the data acquisition, allowing for the association of SLAM and GNSS trajectories; the offsets between their centres of phase were considered when roto-translating the SLAM cloud in the global projected reference system (EPSG: 32632). The point cloud was then filtered and clipped to the area of interest.

The HLS point cloud was used to generate the 3D geometric model for data visualization and storage. We decided to use the single tree as the reference unit for our DT and associate measurements to each one. A local maxima individual tree detection (ITD) technique was used to identify individual trees. This approach was chosen because height and treetop-based models have proven effective so far, given the structure of the orchard, where all the plants have a distinct treetop and small heights that fall into Kaarta HLS range capabilities. Using the Cloth Simulation Filter (CSF) algorithm implemented in Cloud Compare software, the Digital Terrain Model (DTM) was extracted (Zhang et al., 2016), which was then used to calculate the normalized DTM (nDTM), as the difference between the Digital Surface Model (DSM) and the DTM. To reduce the computation time of the algorithm, the CHM was scaled to 25cm/pixel. The stems were detected using "local minima and maxima" algorithm in SAGA; outliers were removed successively according to maxima distribution. Finally, Voronoi polygons were computed upon the treetops to define the canopy geometry and then projected on the point cloud. The goodness of the ITD method was assessed by comparing the treetop detection with the ones manually identified and computing precision, recall, F1-score and accuracy metrics (Belcore et al., 2021).

The data gathered by the ZED 2 stereo camera was stored in StereoLabs' proprietary .svo format. Point clouds were extracted in .ply format using the ZEDFu software. 35,122 image pairs were captured by the stereo camera and used to automatically generate depth maps. 32,196 RGB image and depth map pairs were exported .png format with a resolution of 1920 x 1080. Only RGB images taken by the left camera were exported.

To examine the quality of the stereo camera point clouds and to test their correspondence with the HLS data, point cloud registration was performed on a test section of the orchard. This section was located at the end of the second row, contained 12 trees, and had a length of 33m and an area of 341m^2. This test section was selected based on the presence of red, highly visible apples on the trees. This allowed for easier visual assessment of apple representation in the stereo camera point cloud, using the corresponding RGB images as a reference.

Point cloud registration was carried out using Cloud Compare software. The HLS point cloud was used as the reference cloud. The HLS point cloud was clipped to the area of interest, and the stereo camera cloud was registered first manually, then using the ICP algorithm (Theoretical overlap = 70%) before being merged. The root of the mean square (RMS) error was calculated to assess the quality of the ICP registration.

Phase 3: Apple detection and segmentation

An Ultralytics YOLOv8 (Jocher et al., 2023) segmentation model was trained to detect and segment apples. A training dataset was prepared using the RGB images previously exported from the stereo camera. Images were subsampled and filtered to include only images of apple-bearing trees with minimal overlap, resulting in 65 images. To supplement the dataset, 670 images from the MinneApple dataset (Häni et al., 2020) were added. All images were sliced into smaller images of the optimal YOLOv8 input size (640 x 640), labelled, then split into training (1963), validation (379), and testing (405) datasets.

Training was conducted using a medium pre-weighted segmentation model (YOLOv8m-seg.pt) with default parameters (epochs = 100, batch size = 16, image size = 640, data augmentation = enabled) and validation was automatically conducted. At the conclusion of training, the best model was saved.

3. Results

Phase 1: Data acquisition

Agri.Q travelled a total of 580m in 30 minutes. All sensors remained securely fixed to the platform and gathered data for the duration of the journey.

Phase 2: Data processing and information extraction

The raw HLS point cloud consisted of 30,000,000 points and covered an area of approximately 6,000 m^2. The final cloud had 20,749,929 points and the RTK acquisition had an error of 4 centimetres.

The local maxima algorithm detected 243 treetops. Through the analysis if the heights distribution, the outlier threshold was selected (inflection point). 40 points were identified filtered out as outliers.

The goodness of the detection was then validated through comparison with the Voronoi polygons built on the treetop identified through visual interpretation. This resulted in 113 matched elements (the treetop falls within the reference polygon, true positive), 24 missed (no treetop within the reference polygon, false negative), 31 oversampled (more than one treetop within the reference polygon, false positive), and 35 (treetop outside the reference polygons, false positive).

The calculated precision, recall, F1 score, and accuracy metrics are listed in Table 3.

Metric	Precision	Recall	F1	Accuracy
Value	0.63	0.85	0.72	0.69

Table 3. Treetop detection accuracy metrics.

The stereo camera point cloud representing the registration test section consisted of 1,971,919 points while the HLS cloud clipped on the test rows consisted of 2,011,972 points. The RMS error of the ICP alignment was 0.064 m.

Phase 3: Apple detection and segmentation

The trained segmentation model was able to detect and segment apples in our images with high accuracy ($F_1 = 0.89$), detecting nearly all the apples (recall = 0.91) with a high degree of precision (0.87) (Fig. 5).

(a) (b)

Figure 5. (a) The original image and (b) the same image with YOLOv8-generated segmentation masks (red)

4. Discussion

Agri.Q proved to be a suitable UGV for orchard use. It easily navigated the terrain and the sensor-mounting platform allowed for the simple, secure installation of four sensors and auxiliary tools, including a battery and a laptop. While Agri.Q is currently controlled remotely by an onsite operator, automation would increase its utility, allowing farmers to concentrate on other tasks while Agri.Q performs data collection missions independently.

The HLS and GNSS receiver performed as expected and provided a high-quality, georeferenced point cloud. The ZED 2 stereo camera did not perform as well as expected, generating sparse, low-quality point clouds. Other studies using ZED and ZED 2 stereo cameras in orchards have encountered similar problems (Wang et al., 2017; Chen et al., 2021; Neupane et al., 2021). Still, stereo cameras have the advantage of (a) being significantly less expensive than LiDAR sensors and (b) integrating RGB data directly with depth data. Further testing of the ZED 2 under different lighting conditions and orientations, as well as comparisons with other available RGB-D sensors, is advisable. While images from the Mapir 3N were ultimately not included in our workflow, they could eventually be integrated to add information about photosynthetic activity and plant health.

The ITD method tends to overestimate the number of trees. This can be attributed to the tendency of lateral branches to grow upwards. The result is acceptable and consistent with the performance of automated treetop extraction algorithms from LiDAR (Ozdarici-ok and Ok, 2023). The advantage lies in the speed of the method, with the total processing time being 26 seconds to extract the maxima and 1 second to generate the Voronoi polygons, significantly faster than point cloud segmentation methods.

Upon visual inspection of the stereo camera point clouds, apples were identifiable by colour, however the sparseness of the point cloud resulted in somewhat undefined shapes and difficulty distinguishing individual apples where several were growing in close proximity. Tree branches and leaves were also often indistinct, while tree trunks were slightly better. Vertical wooden posts installed along the rows at roughly 9m intervals were often the most recognizable features. Due to their sparseness and lack of structural details, the stereo camera point clouds would not be suitable for the construction of a DT if used independently.

Therefore, it was necessary to register the stereo camera point cloud with the HLS point cloud. The accuracy of the alignment was poorer than expected, with an RMS error of 6.4cm and several persistent visual artifacts. Given that one of the potential uses of the DT is fruit-picking, which requires precise and accurate measurements, this is not acceptable.

Apple detection and segmentation was highly satisfactory, with excellent recall and precision. Visual analysis of results showed that the model was successful at detecting all or nearly all apples in clear, well-lit images, but struggled to detect all apples in blurry, poorly lit images, resulting in some false negatives. The abundance of high-quality, accurately labelled images from the MinneApple benchmark dataset (Häni et al., 2020) likely helped to compensate for the relatively small size of our own dataset. The variety of trees, apples, colours, and lighting conditions represented in the MinneApple dataset served to prevent overfitting and make the model more generalisable.

The final goal of this workflow is to generate a 3D model of the orchard structure using the LiDAR system so that the information collected by the YOLO algorithm can be associated with each tree and georeferenced. By knowing the depth camera pose coordinates and their relation to the depth camera point cloud, it is possible to transform them into the global reference system. The YOLO model can be uploaded to the depth camera and run during the acquisition, enabling instantaneous detection.

In our case study, we did not load the detection model into the camera because it was developed after the acquisitions. The depth camera's point cloud was aligned with that of the HLS, also thanks to the markers on the poles. This procedure can be further automated if the markers are encoded.

The major limitations and difficulties encountered in the work are related to the data acquisition methods. While the terrestrial laser scanner, being an active sensor, did not encounter particular problems, the depth camera was mainly affected by light exposure and camera orientation. The rows oriented north were in the shade, and the images taken were underexposed. We believe this limiting factor can be overcome with better scheduling of the surveys avoiding hours when the sun is low on the horizon.

The second limiting aspect was the difficulty of reconstructing the depth camera's point cloud. The repetitive geometries and the serpentine trajectory caused inaccuracies in the visual odometry algorithm, resulting in the data only being exportable and processable in different segments. To reduce these errors, we intend to orient the camera not perpendicular to the row but rotated by about 10° so that the entire row is visible and more points are available for the visual odometry algorithm. We are currently researching these aspects to improve and automate the prototype further. Having successfully prepared the 3D and 2D components of a DT, the next steps will focus on their unification via the projection of the 2D colour data and apple detections onto the segmented 3D model. Then, fruit-count data and other metrics can be assigned to each tree, and the full DT will be stored in a queryable webGIS.

5. Conclusions

Here we present a comprehensive workflow for the creation of a DT of an apple orchard. We demonstrate the suitability of an agricultural UGV, Agri.Q, as a mobile sensor carrier in an apple orchard. We also obtained a detailed 3D model and successfully performed ITD, allowing for the eventual association of crop measurements to individual trees. While we did not gather measurements such as fruit count and size, these can be derived from the results of apple detection and segmentation. Using a refined AI model, we performed apple detection and segmentation on RGB images with excellent results. The next phase of this work will focus on the automatization of the data collection and the projection of the 2D detections and RGB data onto the 3D model in order to realize the final georeferenced, segmented, colourised DT.

As the Agriculture 4.0 revolution ushers in the age of smart farming, agricultural DTs will move rapidly from the conceptual to the real world. To facilitate this transition and ensure the readiness of the technology, the development and testing of DT generation methods must begin now. Our proposed methodology represents one of the first start-to-finish workflows for the creation of a DT of an apple orchard and provides a foundation on which future DTs of complex, living agricultural systems can be built.

Acknowledgements

The authors acknowledge the Agrion research center for providing their experimental orchards for these tests.

This study was carried out within the Agritech National Research Center and received funding from the European Union Next-GenerationEU (PIANO NAZIONALE DI RIPRESA E RESILIENZA (PNRR) – MISSIONE 4 COMPONENTE 2, INVESTIMENTO 1.4 – D.D. 1032 17/06/2022, CN00000022). This manuscript reflects only the authors’ views and opinions, neither the European Union nor the European Commission can be considered responsible for them.

References

Alves, R. G., R. F. Maia, and F. Lima. 2023. Development of a Digital Twin for smart farming: Irrigation management system for water saving. Journal of Cleaner Production. 388, 135920. https://doi.org/10.1016/j.jclepro.2023.135920.

Apeinans, I., L. Litavniece, S. Kodors, I. Zarembo, G. Lacis, and J. Deksne. 2023. Smart Fruit Growing Through Digital Twin Paradigm: Systematic Review and Technology Gap Analysis. Engineering Management in Production and Services. 15 (4), 128–143.