81. An autonomous UAV system for precision greenhouse muskmelon growth monitoring

J.H. Lin and T.T. Lin*
*National Taiwan University, No. 1, Section 4, Roosevelt Road, Taipei, Taiwan, ROC, * m456@ntu.edu.tw*

Abstract

Greenhouse muskmelon cultivation requires precise monitoring to optimize management and yield, particularly during the vine growth stage. This study developed an image processing pipeline through a visual SLAM-based autonomous unmanned aerial vehicle system for continuous monitoring in greenhouse environment. The proposed system, equipped with a camera, followed a pre-defined flight path to capture visual data of the growing plants. Image processing and visual segmentation techniques were then employed to analyze the data. Specifically, YOLOv8 is used for plant detection, SORT for tracking, and DinoV2 for segmentation and measurement. Testing on 15 muskmelon plants over 21 days achieved a mean Average Precision (mAP) of 0.921 and Multiple Object Tracking Accuracy (MOTA) of 72.4%. Accurate height (R^2=0.979) and width (R^2=0.747) measurements highlight its potential for scalable greenhouse crop monitoring.

Keywords: autonomous navigation, object tracking, plant growth modeling, plant traits, visual SLAM

Introduction

The need to measure greenhouse plant growth parameters increases as technologies evolve. Especially for vertical breeding plants like musk melon (Kesh and Kaushik, 2021). Measuring plant height, canopy volume, diameter, and other morphological characteristics using depth imaging has proven effective. These traits can provide more information to monitor the yield and response from irrigation (Cabello *et al.*, 2009; Dong & Isler, 2018; Li *et al.*, 2011). Lidar and stereo cameras can generate 3D point clouds to monitor crop growth progress (Das *et al.*, 2015). However, when using a mobile platform like a drone, it is restricted by the payload; hence, a monocular camera is a suitable device to acquire depth information for measurement, and a visual transformation-based method can be used (Karpov and Makarov, 2022). This research aimed to apply an autonomous UAV system with multiple-object tracking and monocular depth-estimation techniques to monitor greenhouse muskmelon growth continuously.

Materials and methods

Autonomous UAV navigation platform design
The UAV system was designed specifically for indoor greenhouses with unavailable GPS signals. The platform measures 35 cm×30 cm and weighs 890 g, making it suitable for navigating confined spaces. The system employs a Pixhawk 6C flight controller for precise flight control and a Raspberry Pi 4 as the companion computer. The Raspberry Pi, running Ubuntu 22.04 with Robot Operating System 2 (ROS2, Humble), serves as the command middleware, collecting data and facilitating communication with the flight controller. The ROS2 architecture enables seamless integration among the UAV's components, including the companion computer, the remote localization processing unit, and the flight controller. For localization, ORB2-SLAM is executed on a laptop, enabling the system to follow a pre-scripted route to complete its mission (Mur-Artal & Tardos, 2017). The companion computer activates a 155° FOV, 12MP wide-view camera (Runcam Thumb Pro) mounted on the frame to capture a full view of the entire melon plant during flight. The whole system architecture, including data acquisition and the muskmelon monitoring pipeline, is illustrated in Figure 1.

Muskmelon growth monitoring pipeline

The image processing procedure begins with video clips captured by a UAV flying over the greenhouse. Each frame is corrected for lens distortions through an non-distortion process, ensuring precise spatial relationships within the image. The undistorted frames are then processed using YOLOv8 object detection algorithm to identify and localize muskmelon plants. Depth prediction is performed using the DinoV2 model (Oquab *et al.*, 2024), which is crucial for spatial analysis and filtering out plants in uninterested rows. The Simple Online and Real-time Tracking (SORT) algorithm is applied to the detection outputs, consistently identifying individual plants across multiple frames (Bewley *et al.*, 2016). Known reference points within the greenhouse establish a pixel-to-centimeter ratio for accurate scale calibration. For each tracked plant, measurement data is extracted when its bounding box traverses the central region of the camera view, as this positioning minimizes perspective distortion effects. Known reference points within the greenhouse establish a pixel-to-centimeter ratio for accurate scale calibration. A segmentation mask is then applied to isolate plant regions from the background. Plant height and width are calculated within these segmented regions, with depth information used to adjust measurements for perspective effects. The detected and tracked plants are registered in a spatial database, associating each with its location within the greenhouse. The illustration of the workflow is shown in Figure 2.

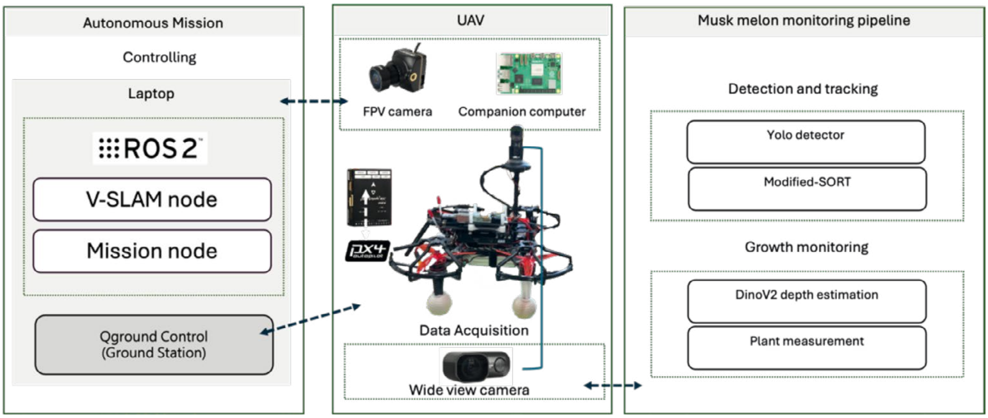

Figure 1. System architecture for greenhouse UAV-based muskmelon monitoring.

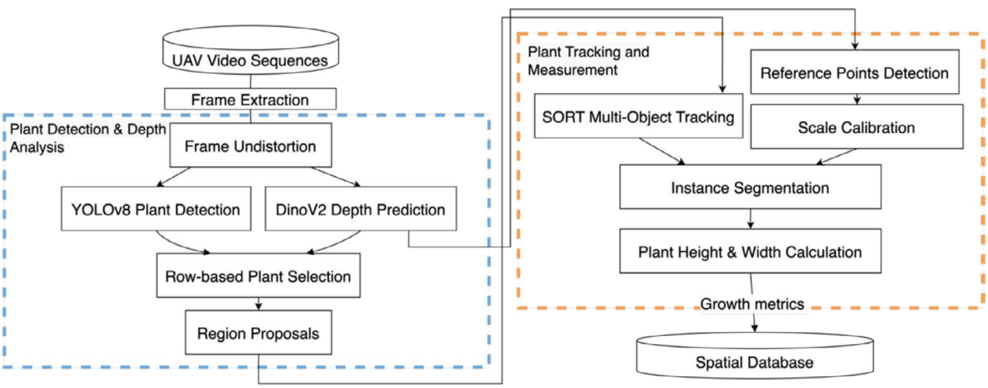

Figure 2. Muskmelon plant measurement workflow diagram.

The experiment also monitored muskmelon development from establishment to pre-pollination over one month. The detection dataset, comprising 431 annotated muskmelon plants, was collected from the experimental field. The experimental setup consisted of A, B, and C, three rows of muskmelon plants, with five plants per row. Data collection followed a pre-scripted UAV flight path along each aisle, maintaining a consistent height of 1.4 m.

Results and discussion

Muskmelon plant detection and tracking

The plant detection and tracking results are presented in the table below. A plant detection YOLOv8 model was trained using a resolution of 640×640 pixels, a learning rate of 0.01, and a momentum value of 0.937. The training process included 150 epochs with a batch size of 8. The detection performance yielded a precision of 0.833 and a recall of 0.89. The mean Average Precision at 50% intersection over union (mAP50) reached 0.921, while the mean Average Precision across 50–95% intersection over union thresholds (mAP50-95) was reduced to 0.639. While plant boundary differentiation was achieved, significant challenges emerged from occlusions caused by overlapping rows and extended foliage. This limitation affected the predicted bounding box's fit to the plant. An additional processing step was necessary to address these issues and achieve precise measurements during tracking. The tracking accuracy was evaluated with a Multiple Object Tracking Accuracy (MOTA) score of 0.724. The evaluation also recorded a false negative rate of 17.96% and a false positive rate of 8.6%. The detection process heavily influences the overall tracking performance, especially when dealing with challenges such as occlusions from overlapping rows, as illustrated in Figure 3a. From the camera's perspective, plants in the background move faster than those in the foreground due to parallax effects, leading to an increased false negative rate. However, the false positive rate remains relatively low, indicating that the algorithm is effective for plant identification. By integrating depth estimation, it was possible to filter and isolate the tracking of plants in the front row, as illustrated in Figure 3(b). This approach demonstrates the potential for improving tracking precision in complex environments.

Muskmelon plant height and width measurement

Based on the identified results, frames for each plant were extracted and cropped. Depth estimation and segmentation masks were then applied, enabling precise delineation of plant contours throughout the sequence. The measurement results are depicted in the Figure 4. Figure 4a illustrates the raw measurement data from a plant tracker, with a representative mask selected from the sequence of masked frames. Pixel measurements were converted to physical dimensions using a reference point-based transformation. Figure 4b demonstrates how the application of segmentation masks improved precision compared to detection boxes, particularly in accurately capturing plant contours. Figure

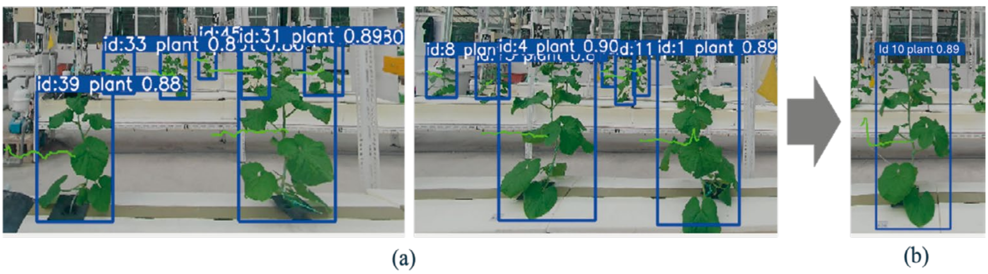

Figure 3 (a) Raw tracking results and (b) filtered tracking output, retaining only the plant in the front row.

4c depicts a muskmelon plant's growth measurements over time through its development phases, highlighting changes in plant dimensions over time.

Figure 5 presents the results of muskmelon growth monitoring over a 21-day period and the distribution of estimation errors. Height measurements generally exhibited a positive bias, with a mean overestimation of 4.34 cm and a standard deviation of 7.48 cm. In contrast, width measurements showed a slight negative bias, with a mean underestimation of -2.52 cm and a standard deviation of 5.56 cm. Kruskal-Wallis tests revealed statistically significant differences in measurement errors across dates for both height (χ^2=65.6768, p=9.101e-07) and width (χ^2=40.0952, p=0.004859), indicating temporal variations in measurement accuracy. This variation is particularly noticeable with larger plants or complex shapes, revealing the limitations of 2D-based methods for capturing certain growth details.

Modeling of muskmelon growth

Growth curves, modeled using the Gompertz function, exhibited sigmoidal trends for both plant height and width, demonstrating reasonable alignment with ground truth data. The analysis revealed that Row C exhibited the most significant growth, with the highest growth rate and maximum height, although this was accompanied by notable variability, as shown in Figure 6. The growth curves captured both the treatment-induced differences and within-row variations effectively. During the early and middle growth stages, the estimation method closely aligned with the actual growth curves, accurately identifying key growth trends. Relative Growth Rates (RGR) demonstrated the

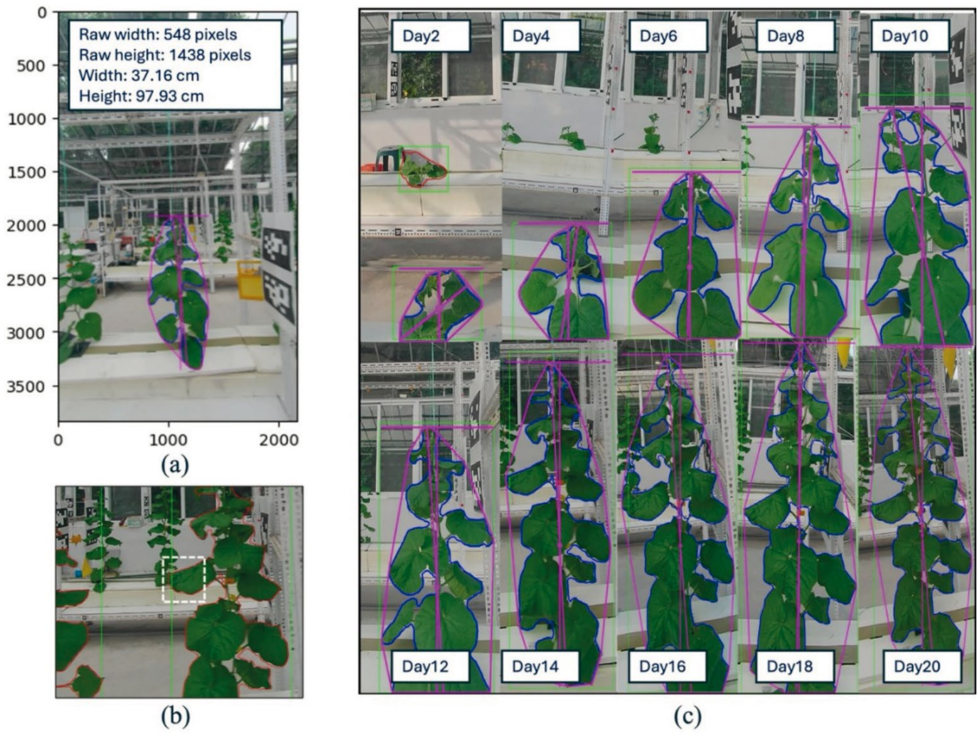

Figure 4 (a) Plant measurement output, (b) comparison of segmentation mask and detection bounding box, and (c) growth measurements over time of a muskmelon plant.

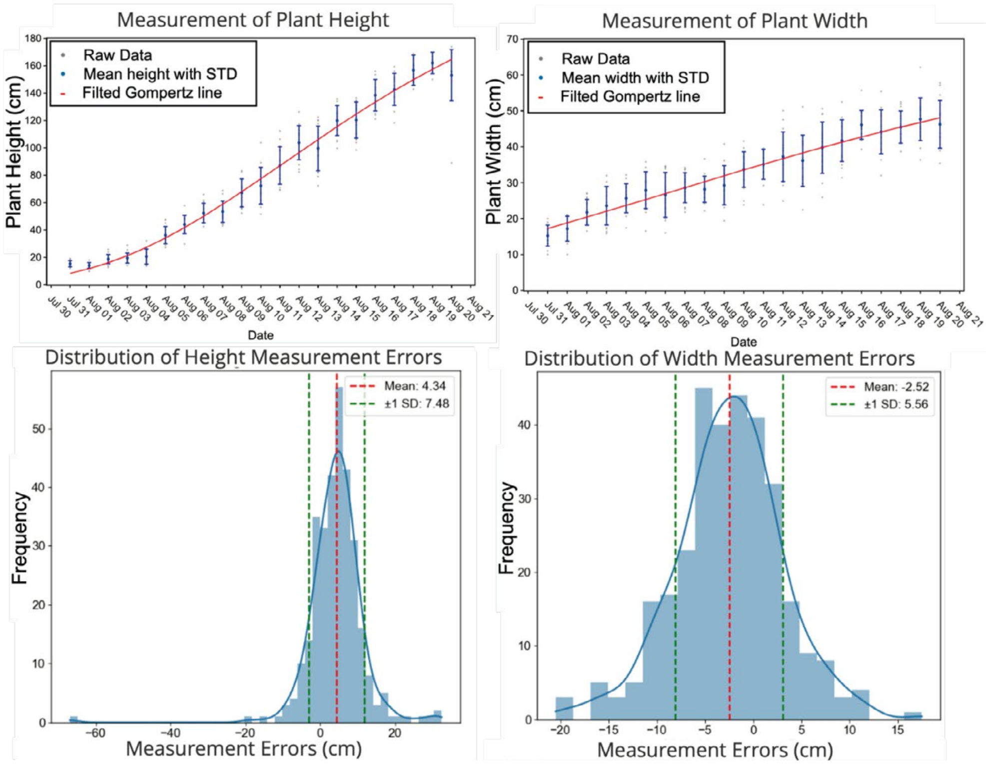

Figure 5. Comparison of 2D-based and ground truth measurements of muskmelon growth over 21 days.

highest changes during the initial stages but stabilized as the plants matured, reflecting the plants' early vertical growth phase. On August 6th, the plants were secured with ropes, an intervention that influenced their subsequent growth trajectories. Height measurements were consistently overestimated across all rows, whereas width measurements were generally underestimated, particularly in rows exhibiting substantial growth. Larger plants exhibited increased measurement uncertainty, as reflected by wider error bars in the later growth stages.

Conclusion

This research demonstrated the potential of a vision-based UAV system equipped with a wide-angle camera for monitoring plant growth in the greenhouse environment. The system's effectiveness was validated by repeatedly following predefined routes and collecting image data throughout the growth period. The visual analysis pipeline enabled the extraction of plant height and width, capturing differences between individual plants, thereby providing a valuable tool for agricultural research. However, as plants grew larger, measurement variability increased, particularly in later growth stages, due to challenges such as occlusions, plant positioning, and varying camera angles. These findings emphasize the importance of maintaining consistent camera placement and suggest that incorporating multi-angle image capture could reduce variability and improve accuracy, particularly during complex growth phases.

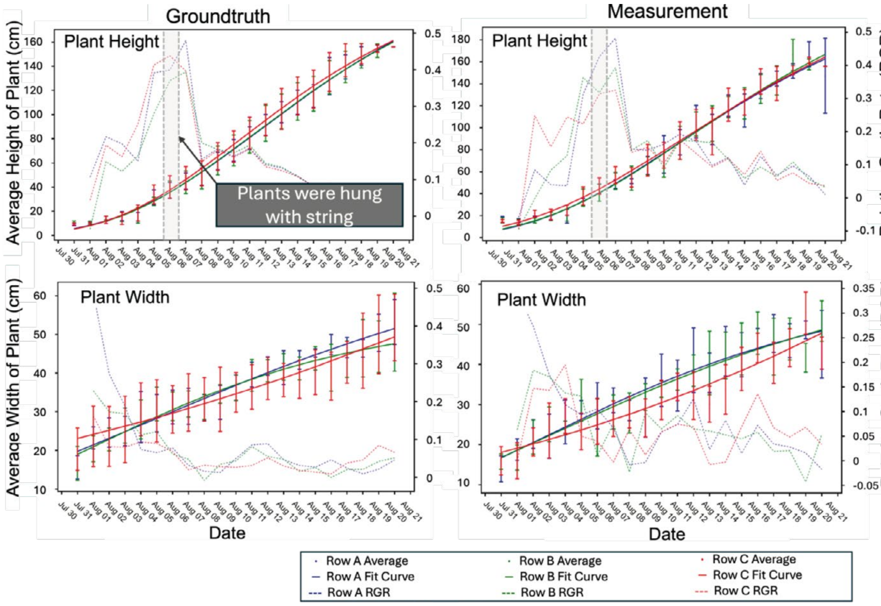

Figure 6. Comparison of growth curves and relative growth rates across different rows of muskmelon.

References

Bewley, A., Ge, Z., Ott, L., Ramos, F., & Upcroft, B. (2016). Simple online and realtime tracking. 2016 IEEE International Conference on Image Processing (ICIP), pp 3464–3468. https://doi.org/10.1109/ICIP.2016.7533003

Cabello, M. J., Castellanos, M.T., Romojaro, F., Martínez-Madrid, C., & Ribas, F. (2009). Yield and quality of melon grown under different irrigation and nitrogen rates. Agricultural Water Management, 96(5), 866–874. https://doi.org/10.1016/j.agwat.2008.11.006

Das, J., Cross, G., Qu, C., Makineni, A., Tokekar, P., Mulgaonkar, Y., & Kumar, V. (2015). Devices, systems, and methods for automated monitoring enabling precision agriculture. 2015 IEEE International Conference on Automation Science and Engineering (CASE), pp. 462–469. https://doi.org/10.1109/CoASE.2015.7294123

Dong, W., & Isler, V. (2018). Tree morphology for phenotyping from semantics-based mapping in orchard environments Non-peer reviewed preprint at arXiv: 1804.05905. https://doi.org/10.48550/arXiv.1804.05905

Karpov, A., & Makarov, I. (2022). Exploring efficiency of vision transformers for self-supervised monocular depth estimation. 2022 IEEE International Symposium on Mixed and Augmented Reality (ISMAR), pp 711–719. https://doi.org/10.1109/ISMAR55827.2022.00089

Kesh, H., & Kaushik, P. (2021). Advances in melon (*Cucumis melo* L.) breeding: An update. Scientia Horticulturae, 282, 110045. https://doi.org/10.1016/j.scienta.2021.110045

Li, Y. J., Yuan, B. Z., & Bie, Z. L. (2011). Response of muskmelon to drip irrigation water inside a plastic greenhouse. In 2011 International Conference on New Technology of Agricultural. IEEE, Piscataway, NJ, pp. 333–337. https://doi.org/10.1109/ICAE.2011.5943814

Mur-Artal, R., & Tardos, J. D. (2017). ORB-SLAM2: An open-source SLAM system for monocular, stereo, and RGB-D cameras. IEEE Transactions on Robotics, 33(5), 1255–1262. https://doi.org/10.1109/TRO.2017.2705103

Oquab, M., Darcet, T., Moutakanni, T., Vo, H., Szafraniec, M., Khalidov, V., …, & Bojanowski, P. (2024). DINOv2: Learning robust visual features without supervision Non-peer reviewed preprint at arXiv: 2304.07193. https://doi.org/10.48550/arXiv.2304.07193