The 11th Asian-Australasian Conference on Precision Agriculture (ACPA 11)
October 14-16, 2025, Chiayi, Taiwan

PLANTSAGA: INTEGRATING SEGMENT ANYTHING MODEL WITH GAUSSIAN SPLATTING FOR PLANT ORGAN-LEVEL 3D SEGMENTATION

Kuan-Chen Lin, Jieng-Heng Lin, Ta-Te Lin*

Department of Biomechatronics Engineering, National Taiwan University, Taipei, Taiwan, ROC

*Corresponding Author: m456@ntu.edu.tw

ABSTRACT

Organ-level 3D phenotyping is essential for crop breeding but remains limited by the high cost of manual annotations. To address this challenge, PlantSAGA (Plant Segment Anything Gaussian Splatting) is introduced as a reference-based framework that enables accurate organ segmentation with minimal annotation. Multi-view muskmelon plants were reconstructed using COLMAP for camera pose estimation and Gaussian Splatting for 3D modeling, while 1~10 reference masks guided organ-level discrimination. The pipeline proceeds in four stages: class-agnostic masks are generated by SAM, enriched with DINOv2 embeddings, aligned across views through a 3D-aware memory bank with contrastive loss, and refined by density–connectivity filtering to produce coherent 3D structures. In 10-shot settings, PlantSAGA achieved AP_{50} and AR_{50} above 0.80 and reconstructions with PSNR 28.62 dB and SSIM 0.83, establishing its potential as a scalable tool for 3D plant phenotyping.

Keywords: 3D Gaussian splatting, reference-based 3D instance segmentation, plant phenotyping, multi-view consistency, segment anything model.

INTRODUCTION

High-throughput organ-level phenotyping is critical for crop breeding but remains constrained by costly annotations and task-specific training. Recent advances have combined multi-view reconstruction with segmentation to extract organ traits in 3D. For example, PlantSegNeRF generates high-precision organ point clouds from multi-view images using neural radiance fields, yet its reliance on annotated masks limits scalability. In contrast, Leaf-Only SAM applies the Segment Anything Model with post-processing for leaf segmentation in 2D images, requiring no fine-tuning or training data, but offering no extension to 3D. To reduce annotation costs, reference-based methods have emerged. “No Time to Train” integrates SAM with DINOv2 to enable few-shot segmentation, though it does not ensure cross-view consistency (Espinosa et al., 2025). The Gaga (Group any Gaussians) framework addresses this limitation by linking inconsistent 2D masks across different views through a 3D-aware memory bank (Lyu et al., 2024). Building on these developments, this study introduces PLANTSAGA, a reference-based 3D instance segmentation framework that unites semantic filtering with multi-view consistency to achieve organ-level segmentation with minimal annotation.

MATERIALS AND METHODS

Multi-view images of muskmelon plants were collected in the NTU greenhouse and reconstructed with COLMAP for camera pose estimation, while Gaussian Splatting provided efficient radiance-field representation for 3D modeling. A small set of reference masks (1~10 per organ) was prepared to guide segmentation with minimal annotation. The PLANTSAGA pipeline integrates four stages: SAM generates class-agnostic masks; these are semantically enriched by matching DINOv2 embeddings with the reference set for organ-level discrimination (Espinosa *et al*., 2025); cross-view consistency is enforced through a 3D-aware memory bank (Lyu *et al*., 2024) enhanced with a contrastive loss; and the resulting organ-level Gaussians are refined with density and connectivity filtering to produce coherent 3D segmentations. Segmentation performance was evaluated using Average Precision (AP), Average Recall (AR), and Dice-Sørensen coefficient (DSC), following the metrics used in Leaf-Only. In addition, reconstruction fidelity was assessed using Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM).

RESULTS & DISCUSSION

PlantSAGA achieved accurate organ-level 3D segmentation with minimal annotation. With 10-shot settings, AP_{50} and AR_{50} exceeded 0.80, approaching the supervised performance of Mask R-CNN, while surpassing zero-shot approaches. Segmented leaves achieved an average PSNR of 28.62 dB and SSIM of 0.83, confirming high-fidelity reconstruction consistent with previous study (Chen *et al*., 2025). These results establish PLANTSAGA as a scalable alternative to training-based pipelines for efficient 3D organ phenotyping. It should be noted that the current evaluation was limited to a small number of plant types and tested only on leaves; performance also decreased on species with very small or highly occluded leaves.

CONCLUSIONS

This work presented PlantSAGA, a novel reference-based 3D instance segmentation framework that requires minimal annotation while achieving segmentation accuracy comparable to training-based approaches and preserving high-fidelity reconstruction. By integrating SAM, DINOv2, Gaussian Splatting, and a 3D-aware memory bank, the method advances beyond existing pipelines by uniting reference-based semantic filtering with multi-view consistency. Future work will extend PlantSAGA to additional plant organs and more diverse architectures to further demonstrate its broader applicability in 3D plant phenotyping.

REFERENCES

Chen, J., Jiao, Y., Jin, F., Qin, X., Ning, Y., Yang, M., & Zhan, Y. (2025). Plant Sam Gaussian Reconstruction (PSGR): A High-Precision and Accelerated Strategy for Plant 3D Reconstruction. *Electronics*, 14(11), 2291. https://doi.org/10.3390/electronics14112291

Espinosa, M., Yang, C., Ericsson, L., McDonagh, S., & Crowley, E. J. (2025). No time to train! Training-free reference-based instance segmentation. *arXiv preprint* arXiv:2507.02798. https://doi.org/10.48550/arXiv.2507.02798

Lyu, W., Li, X., Kundu, A., Tsai, Y.-H., & Yang, M.-H. (2024). Gaga: Group any Gaussians via 3D-aware memory bank. arXiv preprint arXiv:2404.07977. https://doi.org/10.48550/arXiv.2404.07977