MOBILE-BASED AUTOMATED PHENOTYPING SYSTEM FOR ACCESSIBLE TOMATO BREEDING

Ssu-Chi Chen[1], Guan-Hua Chen[1], Ya-Ping Lin[2], Shih-Fang Chen[1]*

[1] Department of Biomechatronics Engineering, National Taiwan University, Taiwan.
[2] World Vegetable Center, Tainan, Taiwan
* Corresponding Author: sfchen@ntu.edu.tw

ABSTRACT

Tomato breeding programs require extensive phenotypic data collection including fruit development stages and critical timing parameters, yet manual monitoring is labor-intensive and limits breeding program scalability, particularly in resource-limited environments. This study presents a cost-effective automated phenotyping system that requires only smartphone video recording combined with pre-assigned plot numbers, eliminating the need for expensive mobile platforms and making advanced breeding technologies accessible to developing countries. The system employs a multi-stage architecture combining plant positioning technology with advanced trait detection. Individual plant extraction from smartphone-recorded multi-plant videos is achieved through an automated positioning system, followed by trait detection using DETR with Improved Denoising Anchor Boxes (DINO) architecture to identify four key developmental stages: flowers, green fruits, turning-pink fruits, and harvest-ready fruits. Comprehensive validation demonstrated robust system performance with 0.909 of accuracy in plant positioning, 0.862 mAP and 0.836 F1 score for trait detection. The system significantly reduces infrastructure barriers to precision breeding by leveraging widely available smartphone technology, enabling remote monitoring and data management without expensive hardware investments. This work democratizes access to advanced phenotyping tools, providing breeders worldwide with cost-effective solutions for accelerating variety development regardless of economic constraints.

Keywords: Tomato breeding, Automated phenotyping, Deep learning, Computer vision, Plant localization.

INTRODUCTION

Global tomato production reaches 192 million tons annually (Food and Agriculture Organization Statistical Database, 2023), establishing tomato (Solanum lycopersicum) as a critical vegetable crop. Climate change and biotic stresses increasingly threaten production yield and quality. Current breeding approaches rely on labor-intensive field trials with phenotypic evaluations that suffer from human subjectivity and time constraints. Agricultural labor shortages compound these limitations, restricting breeding efficiency at scale. This study proposes a cost-effective automated phenotyping system integrating deep learning algorithms with practical breeding expertise. Using only smartphone video recordings and pre-assigned plot numbers, the system replaces conventional manual evaluation. This approach enhances tomato phenotyping and precision while reducing labor demand and optimizing breeding research workflow.

MATERIALS AND METHODS

SYSTEM ARCHITECTURE

This integrated system combines two major tasks: plant positioning and trait detection. Spatial, temporal, and phenotypic data are consolidated into a comprehensive database (Fig. 1). Input sources include smartphone camera or automated robots. Outputs provide complete phenotype records and developmental stage classification.
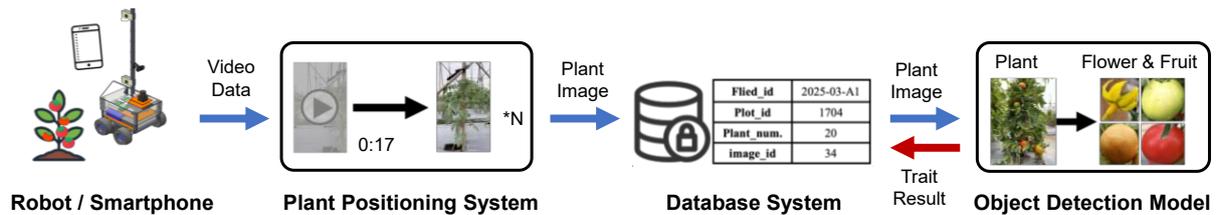

Fig. 1 Schematic diagram of the system framework.

DATA COLLECTION PROTOCOL

Field trials were conducted at an open-field site in Tainan, Taiwan (23.114N, 120.298E) spanning October 2023 through January 2025 across three growing seasons. Dynamic video capture utilized constant-speed camera movement for multi-plant recording, simulating handheld or robotic data acquisition. Static imagery focused on single-plant samples for model training (Fig. 2). Coverage included early and late growth stages under three daily lighting periods (Table 1).
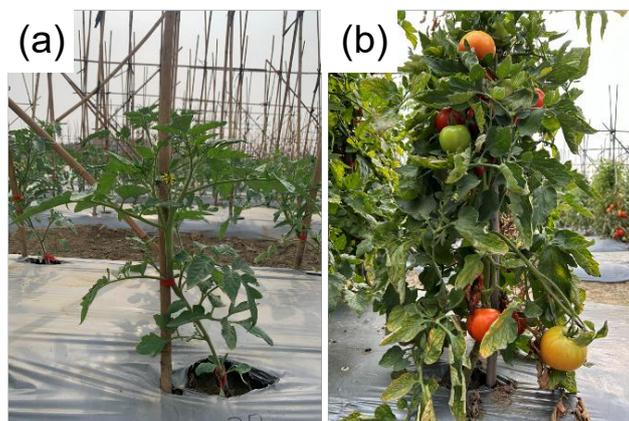

Fig. 2 Tomato plant sample images. (a) Early growth stage; (b) late growth stage.

Table 1 Distribution of images across collection periods and growth stages.

Growth stage \ Collection Period	Number of images		
	8 am to 10 am	11 am to 1 pm	2 pm to 4 pm
Early	536	127	617
Late	160	199	1053

PLANT POSITIONING SYSTEM

Monocular depth estimation using Marigold model processes continuous field videos to extract individual plant frames. Generated depth maps capture structural relationships between plants and surrounding environments. This approach outperforms traditional depth sensors by reducing occlusion artifacts across diverse plant sizes (Fig. 3). Transformer-based temporal modeling identifies centered plant frames using 64-dimensional *x*-axis depth features. Sequential relationships are encoded through temporal context integration. Frame-wise probability computation utilizes fully connected networks with ReLU and sigmoid activations. Peak detection algorithms retrieve target plant images.
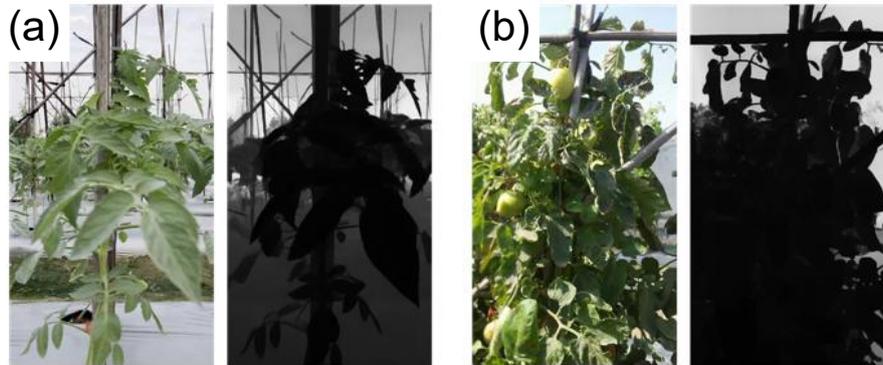

Fig. 3 Original plants image and depth map generated by Marigold. (a) Early growth stage plant; (b) late growth stage plant.

TRAIT DETECTION MODEL

Object annotation employed Labelme software across 23,475 instances spanning four maturity categories (Table 2): flowers, green fruits, turning-pink fruits, and harvest-ready fruits (Fig. 4). DETR with Improved deNoising anchOr boxes (DINO) architecture performed multi-class detection. While DINO demands greater computational resources, its architecture demonstrates significant advantages for agricultural applications where accurate detection of small, partially occluded objects in complex environments is essential (Zhang et al., 2022).

Table 2 Distribution of the annotated objects across four maturity categories.

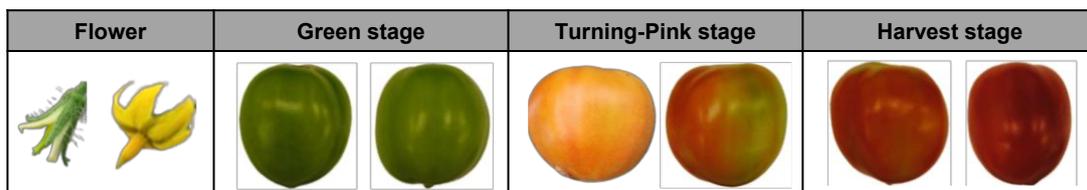| No. of Images	No. of Objects			
	Flower	Green stage	Turning-Pink stage	Harvest stage
2916	5155	13073	1983	3264

Fig.4 Examples of the annotated objects across four phenotypic categories.

RESULTS & DISCUSSION

Plant positioning achieved 0.909 accuracy in frame extraction, successfully identifying centered plant locations. Early-stage performance exceeded late-stage results due to reduced vegetation density and morphological complexity. Increased plant density in later stages reduced depth contrast, elevating positioning errors. DINO detection demonstrated robust multi-class performance (Table 3). Turning-pink stage recognition showed lower precision (0.762) and recall (0.686) compared to other categories. Color variation during this developmental phase, combined with outdoor lighting fluctuations, increased classification difficulty. Temporal analysis of detection sequences predicted three key breeding dates with 4.3-day mean error across 44 genotypes and 880 plants. The system achieved 0.862 mean average precision and 0.836 F1 score for phenotypic recognition.

Table 3 DINO detection model performance of tomato growth stages.

Class	Precision	Recall	F1 score
Flower	0.770	0.779	0.774
Green stage	0.882	0.884	0.883
Turning-Pink stage	0.762	0.686	0.722
Harvest stage	0.808	0.702	0.751
Overall	**0.826**	**0.817**	**0.836**

CONCLUSIONS

This study establishes an automated tomato phenotyping platform integrating monocular depth estimation with deep learning-based trait detection. Plant localization achieved 90.9% accuracy while phenotypic classification reached 86.2% mAP performance. Temporal breeding prediction demonstrated practical utility with sub-week accuracy. Future development will expand seasonal data coverage and phenotypic trait diversity. User interface development will enable real-time genotype tracking and phenological monitoring. This research provides accessible, smartphone-based solutions for global breeding programs, reducing infrastructure barriers while maintaining scientific rigor.

ACKNOWLEDGMENTS

REFERENCES

Xiang, L., Gai, J., Bao, Y., Yu, J., Schnable, P. S., & Tang, L. (2023). Field-based robotic leaf angle detection and characterization of maize plants using stereo vision and deep convolutional neural networks. *Journal of Field Robotics, 40*, 1034-1053.

Zhang, H., Li, F., Liu, S., Zhang, L., Su, H., Zhu, J., Lionel M Ni., & Shum, H. Y. (2022). Dino: Detr with improved denoising anchor boxes for end-to-end object detection. *The Eleventh International Conference on Learning Representations.*

Lee, S.-Y., Wang, J.-C., Guo, M.-C., Jiang, J.-A., Hsieh, M.-H., Peng, J.-C., & Chen, S.-F. (2022). Developing a Self-Guided Field Robot for Greenhouse Asparagus Monitoring. *2022 ASABE Annual International Meeting,* (p. 1).

Food and Agriculture Organization Statistical Database. (2023). Crops and livestock products (QCL). FAOSTAT Database. Retrieved May 11, 2025, from https://www.fao.org/faostat/en/#data/QCL/visualize