Enhancing Strawberry Breeding with AI-Assisted Runner and Flower Identification

Xu "Kevin" Wang and Xue "Kara" Zhou

Summary

This study conducted large-scale strawberry breeding trial characterization using high-throughput phenotyping for two seasons and developed an AI-powered machine vision model for accurately quantifying strawberry runner, flower, and fruit traits in a cost-effective and timely manner. The machine vision model achieved 95%, 99%, and 95% accuracy for runner, flower, and fruit identification, respectively, demonstrating great potential for trait-based selection support. This study will be continued for another season to improve the accuracy of runner counting and empower the machine vision model with nutrient deficiency and disease identification functions for strawberry breeding.

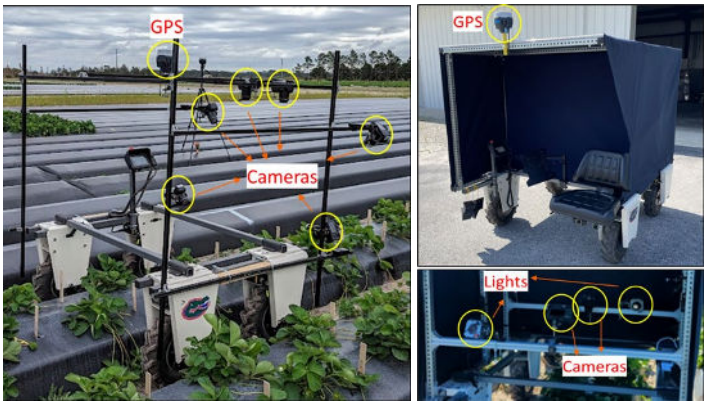
Background

Strawberry runners and flowers are critical in breeding, influencing vegetative propagation and fruit production, but manual data collection is labor-intensive and time-consuming. With advancements in AI technologies, we aim to utilize AI to enhance breeding by reducing manual data collection and accurately deriving trait values for variety selection. Additionally, precise flower and fruit counts will enable growers to estimate potential yields and make better crop management decisions.

Methods

Imaging system integration

Since the 2022-23 strawberry season, we began integrating a ground imaging system (Version 1) for capturing strawberry plant images in the open field. This system included an electric vehicle, a GNSS receiver, and six DSLR cameras positioned at various angles, capturing 4K photos at 1 fps. In the 2023-24 season, we refined the imaging system (Version 2) by adding a shade sail, reducing the number of cameras onboard while increasing image acquisition throughput. The shade sail provided stable lighting conditions, enhancing image quality. Two cameras, facing forward and backward, recorded 4K videos at 60 fps, significantly reducing data collection time.



A B

Figure 1. Ground imaging systems of the 1st version (A) and the 2nd version (B).

Data collection

In the 2022-2023 season, 4K images of 2,000 advanced-stage strawberry varieties were collected from Nov. 2022 to Jan. 2023 using the 1st version ground imaging system at GCREC. 4K video data of strawberry plants in the field was collected monthly using the 2nd version of the ground imaging system from November 2023 to February 2024 from the Clonal trial which included 2,210 selections. These images and videos were used for AI model training and testing for strawberry runner, flower, and fruit identification and quantification.

Machine vision model development

We trained an AI model (StrawberryV2023) based on YOLOv8 using 8,400 images to identify strawberry runners, flowers, and fruits, and we tested it with 2,100 images from the 2022-2023 season. After the 2023-2024 season, we added 1,600 new images to train the latest model (StrawberryV2024) to enhance the strawberry plant parts identification accuracy.

Results

Genotype ID associated with video frames

We developed a computational pipeline to link plant video frames with corresponding genotypes. In evaluating 442 plants across four experiments, the method achieved 100% accuracy in tagging frames with the correct genotypes.

Strawberry parts identification

The latest model (StrawberryV2024) significantly improved the accuracy in identifying strawberry plant parts (Table 1).

Table 1. Test accuracy for four strawberry parts identification.

	Runner	Flower	Immature Fruit	Mature Fruit
StrawberryV2023	71%	79%	80%	72%
StrawberryV2024	95%	99%	94%	96%
Improvement	24%	20%	14%	24%

Strawberry runner quantification

During the first funding period, the study emphasized quantifying strawberry runners to support breeding efforts. The latest model (StrawberryV2024) demonstrated enhanced accuracy (Table 2) in runner quantification during November 2023, a peak period for runner accumulation. The accuracy improved to 82.58% for runner quantification across the entire season, mainly due to the reduced number of runners produced in the late season.

Table 2. Strawberry runner quantification accuracy using two models during November 2023.

	Plants with no error in runner counting	Plants within one-runner error in runner counting
StrawberryV2023	51.36%	89.14%
StrawberryV2024	78.28%	95.25%
Improvement	26.92%	6.11%

Error analysis

After analysis of runner quantification errors, we summarized eight reasons that led to the incorrect counting of runners (Fig.2):

A. Missed runner identification (60.87%)
B. Runners from neighbor plants (17.39%)
C. Overlapped runners (11.96%)
D. Double counting (7.61%)
E. Leaf stems identified as runners (6.52%)
F. Background objects (light/tire print/dead stem) identified as runners (6.52%)
G. Calyx identified as runners (2.17%)
H. Runners out of frame (1.09%)

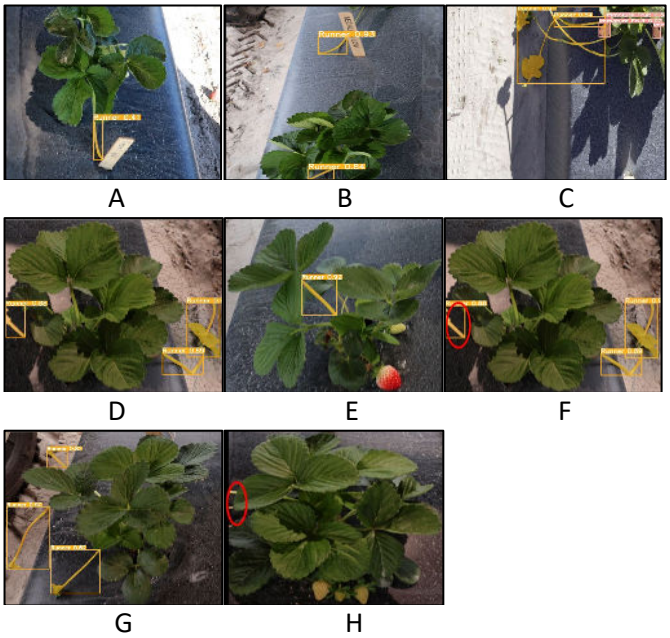


Figure 2. Example images showing causes of errors in runner identification and quantification using StrawberryV2024.

Takeaways

A machine vision model was developed to efficiently identify strawberry runners, flowers, and fruits, with strong potential for breeding acceleration. However, accuracy, especially in quantifying runners, needs to be improved in the next funding period, and quantification time must be reduced for future real-time applications like runner removal.

Contact

Dr. Xu "Kevin" Wang and Dr. Xue “Kara” Zhou
UF/IFAS Gulf Coast Research and Education Center
Phone: 813-419-6618
Email: xuwang1@ufl.edu, zhoux3@ufl.edu
Website: https://gcrec.ifas.ufl.edu/gcrec-facultystaff-directory/xu-kevin-wang/