Refining AI-Assisted Runner Identification in Strawberry Breeding (AGR00031844)

Xue "Kara" Zhou and Xu "Kevin" Wang

Summary

Over the past two years, our work has focused on developing a pipeline that leverages AI-powered machine vision to quantify strawberry traits, specifically runners, flowers, and fruits, at 0.5 m above ground level (AGL), achieving satisfactory accuracy. In the current reporting period, we expanded this effort by capturing both ground-level (0.5 m) images and aerial imagery at altitudes of 5 m AGL and 10 m AGL to train a more robust AI model for strawberry runner identification across multiple domains. The updated model achieved a mean average precision (mAP) of 0.83 for detection and 0.74 for segmentation of runners. Additionally, we collected aerial video data to segment individual strawberry plants and quantify the number of flowers and fruits per plant.

Background

Building on our previous work with ground-based imaging, this year we expanded our research to aerial platforms to improve scalability and operational efficiency. Ground vehicles, while effective at close-range imaging, are limited by slower speeds and the challenges of large, heavy equipment in the field. In contrast, aerial platforms can rapidly collect high-resolution data over large-scale trials. By integrating novel cameras capable of capturing high-resolution imagery and video at altitudes as low as 5 m, we were able to obtain data of comparable quality to ground-based systems while significantly increasing data collection efficiency. This transition allows us to support large-scale phenotyping with greater speed and flexibility.

Methods

Imaging platforms

To support robust runner detection, we deployed three imaging platforms at different heights: **Ground Imaging at 0.5 m AGL** (GI, Figure 1-a, 1-b): A 24 MP camera mounted on a vehicle with LED lighting and a shade cover captured 4K videos at 60 fps; **Aerial Imaging at 10 m AGL** (AI10, Figure 1-c): A UAV (Autel EVO II) captured top-down 20 MP images; and **Aerial Imaging at 5 m AGL** (AI5, Figure 1-d) **and video recording at 4 m AGL**: A UAV (DJI Mavic Pro 3) captured 20 MP images and recorded 4K videos.

Figure 1. Ground imaging and aerial imaging platforms.

Data collection

To support deep learning model training for strawberry runner detection, images containing visible runners were annotated for the three platforms: 461 (GI), 445 (AI10), and 453 (AI5). Annotated images were split into training and validation sets, with data augmentation (horizontal flips and 90° rotations) applied to the training sets to increase diversity. The dataset is publicly available on Dryad (DOI: 10.5061/dryad.bzkh189nw).

Machine vision model development

In the 2024–2025 season, we trained strawberry runner detection and segmentation models using YOLOv8x-seg and YOLOv11x-seg architectures. These models were trained on ground and aerial imagery to enable robust identification of runners across multiple imaging platforms and conditions.

Depth-based plant segmentation and parts counting

We used the Video Depth Anything model to generate depth maps from UAV videos, enabling background removal and segmentation of individual strawberry plants. The YOLOv11x-seg-based model was trained to detect and count parts such as runners, flowers, and fruits for each plant, allowing accurate per-plant quantification across large fields.

Results

Multi-altitude deep learning model performance

YOLOv11x-seg outperformed YOLOv8x-seg in both detection and segmentation tasks, offering higher accuracy and faster inference overall. Detection models achieved better F1 scores and AP50 values than segmentation models. Models trained on individual datasets performed best on their own validation sets but showed limited cross-domain generalization.

In contrast, models trained on the combined dataset (GI + AI5 + AI10) achieved consistent performance across all platforms, with YOLOv11x-seg showing the best overall results. Among imaging methods, aerial imaging at 5 m AGL offered the best trade-off between resolution and coverage for large-scale runner detection. Table 1 presents the cross-dataset detection performance of YOLOv11x-seg.

Table 1. Cross-dataset runner detection performance of YOLOv11x-seg models trained on different datasets.

Validation Data Sets	Models Trained on: GI				AI5				AI10				GI+AI5+AI10			
	F1	P	R	AP50	F1	P	R	AP50	F1	P	R	AP50	F1	P	R	AP50
GI	0.79	0.83	0.74	0.83	0.66	0.74	0.60	0.69	0.58	0.74	0.48	0.56	0.79	0.83	0.75	0.84
AI5	0.61	0.68	0.55	0.60	0.81	0.83	0.80	0.84	0.64	0.67	0.62	0.65	0.81	0.85	0.77	0.82
AI10	0.10	0.36	0.06	0.22	0.47	0.59	0.39	0.48	0.81	0.81	0.81	0.84	0.82	0.87	0.77	0.86
Average	0.50	0.62	0.45	0.55	0.65	0.72	0.60	0.67	0.68	0.74	0.64	0.68	0.80	0.85	0.76	0.84

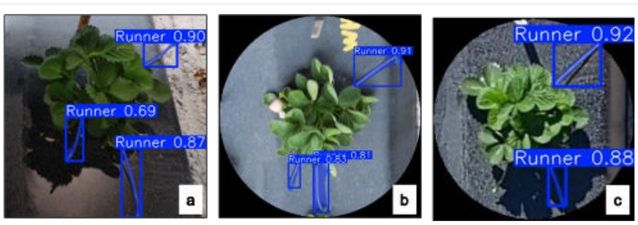

Figure 2. Runner successfully identified within images of (a) GI, (b) AI5, and (c) AI10 by the YOLOv11-seg model retrained on the integrated GI+AI5+AI10 dataset.

Plant segmentation performance

We developed a robust video processing pipeline (Figure 3) to segment individual strawberry plants with 100% accuracy. The original UAV RGB frames (Figure 3-a) are first processed using a monocular depth estimation model to generate grayscale depth maps (Figure 3-b), which guide background removal and foreground extraction (Figure 3-c). Small, non-plant regions are then filtered out, retaining only valid strawberry plants (Figure 3d). A YOLO-based model detects each plant, and circular regions-of-interest (ROIs) are applied (Figure 3-e) to ensure precise segmentation and enable accurate counting of plant parts such as flowers, fruits, and runners.

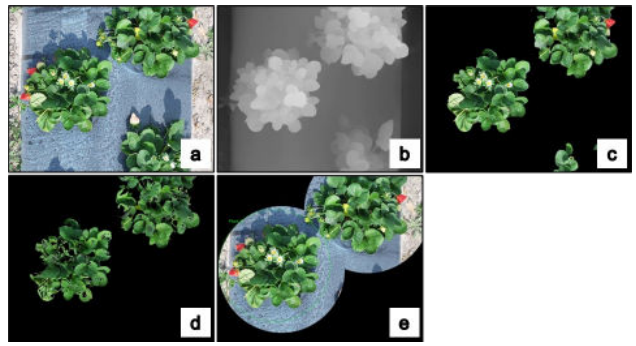

Figure 3. Video processing workflow for individual plant segmentation and strawberry part quantification.

Parts quantification performance

By combining depth-based segmentation with YOLO detection, we were able to quantify flowers and fruits per plant from UAV videos with promising accuracy. As shown in Table 2, predicted counts closely followed visual counts, indicating the potential of this approach. With continued algorithm improvements, this pipeline can provide a reliable solution for automated, large-scale plant trait analysis.

Table 2. Comparison of Visual and Predicted Counts of Strawberry Plant Parts

Classes	Flower	Green Fruit	White Fruit	Pink Fruit	Mature Fruit
Correct Predictions (%)	97.25	85.71	92.59	100.00	100.00
Total Samples	109	119	27	6	13

Takeaways

We developed an automated pipeline combining UAV imaging, depth-based segmentation, and YOLO models for per-plant trait quantification. Multi-altitude platforms enabled robust runner detection, with 5 m AGL offering the best balance of resolution and coverage. We also developed a pipeline that showed strong potential for accurate flower and fruit counting to support high-throughput phenotyping.

Contact

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