Integrated Approach for Precise Strawberry Yield Forecasting

Daeun "Dana" Choi

Summary

Accurate yield forecasting is crucial for optimizing labor management, allocating resources efficiently, and minimizing waste due to overproduction or underproduction. This project is designed to improve the profitability of strawberry growers in Florida by providing tools that will enable them to accurately predict harvest throughout the growing season.

Methods

Objectives were (1) designing machine vision hardware modified to raised bed system for Florida's strawberry production, (2) collecting data to train machine learning algorithm for strawberry flowering stage classification with an estimated time frame to harvest, and (3) developing a machine learning algorithm to measure estimated volume, weight and maturity of the fruit. This year, we successfully developed a computer vision-based and comprehensive strawberry crop load prediction system for yield forecasting. This system combined machine learning, sensors, and data analysis for seasonal harvest predictions. The project laid the foundation for an integrated approach to strawberry yield forecasting, addressing several critical aspects of the process, including flower and fruit stage classification, fruit volume estimation.

Prototype Development

Our first objective revolved around designing machine vision hardware tailored specifically to the raised bed system employed for Florida's strawberry production. We have achieved significant progress in this area, creating a unique vision system that accounts for the intricacies of the raised bed method (Figure 1). We used cameras mounted on the vehicle for enhanced bed coverage and fruit visibility. These cameras were optimized for close-range fruit detection, ensuring a compact system without sacrificing image quality (Figure 2).

Figure 1. The ground vehicle with stereo vision cameras for capturing images of strawberry plants and LiDAR.

Figure 2. Sample images of strawberry bed.

Machine Vision Models

Pursuing our second and third objectives, we gathered extensive data from physical fields and a virtual farm for our machine learning algorithms (Figure 3).

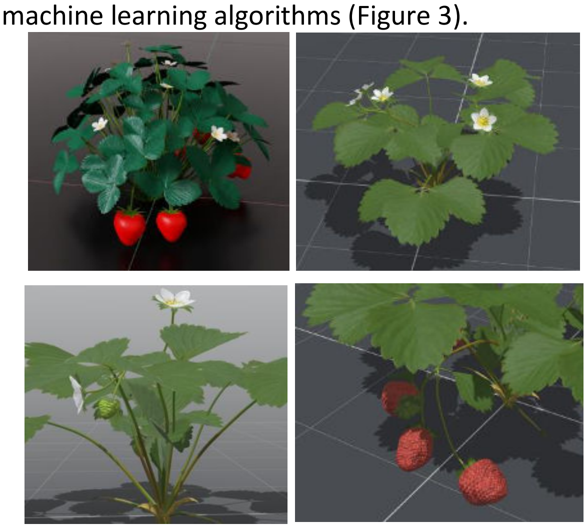

Figure 3. Sample screenshots of the simulated strawberry plants farm with different growth. Source: Mirbod and Choi, 2023.

Two 3D strawberry models were used; M1 had five plant variations, while M2, with higher detail, represented strawberries at varied growth stages (Figure 3). These algorithms detect strawberry growth stages, aiding harvest timing predictions. We're advanced in creating an algorithm that gauges fruit volume, weight, and maturity.

Results

The project utilized TensorFlow 2's Object Detection API, specifically, employing the "Faster R-CNN Inception ResNet V2 1024x1024" model (Huang et al., 2017) for image training. A total of four distinct fruit detection models were trained, each emphasizing different aspects, including synthetic images (both M1 and M1+M2 variations), real images, and combinations of synthetic and real images (Table 1). The subsequent fruit detection accuracies for each model are summarized in Table 2, where the percentage accuracy is derived from the ratio of the model's predicted count to the actual count.

Table 1. Image datasets used to train Faster R-CNN models for strawberry fruit detection. Source: Mirbod and Choi, 2023.

	# of Training Images	# of Fruit Training Labels	# of Validation Images	# of Testing Images
Synthetic Images (M1)	75	189	10	25
Synthetic Images (M1+M2)	75	201	10	25
Real Images	65	278	10	25
Synthetic (M1) + Real	65	306	10	25

Table 2. Fruit detection performance on test image dataset using models trained on synthetic, real, and mixed image data. Source: Mirbod and Choi, 2023.

	TP	FP	FN	Precision	Recall	F1-Score	% Accuracy
Synthetic (M1)	56	2	86	0.96	0.39	0.56	39.4
Synthetic (M1+M2)	96	3	44	0.97	0.69	0.80	68.6
Real	129	2	12	0.98	0.91	0.95	91.4
Mixed (Real+M1)	127	6	14	0.95	0.90	0.93	90.0

The simulated strawberry plant images from the M1 model, despite their simplicity in texture and uniformity in color, managed to identify over one-third of the fruit in the test images. The synthetic plants' lack of structural variability compared to real plants did not hinder detection. Both M1 and M2 models' detection seemed indifferent to variations in size, color, or orientation of the fruit, though inconsistencies in detection were noted. Both synthetic data models struggled to detect immature fruit with visible calyx features. While the M1 entirely missed this feature, the M2 captured one-third. Using a simulated farm for yield prediction can minimize field trials and adapt faster to new crop management changes, given improving fruit detection accuracy (Figure 4).

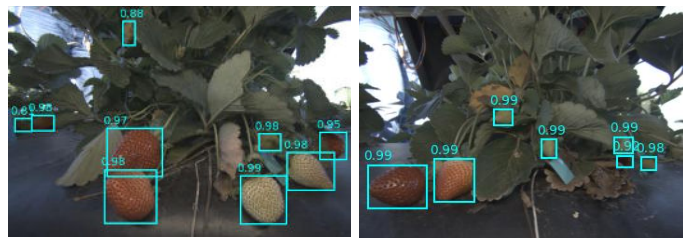

Figure 4. Sample images of strawberry fruits detected in the field using a training model developed in the simulated strawberry farm.

Takeaways

1. The insights gained from these measurements have the potential in determining optimal harvesting periods and strategizing other decisions related to the distribution and marketing of the harvested fruit.
2. The various accuracy with synthetic, real, and mixed image data training emphasize the value of diversifying data sources. By investing in mixed-data models that utilize both real and synthetic images, growers can achieve enhanced detection accuracy of various fruit stages at a faster pace with potentially lower cost.
3. Next step: there will be an emphasis on refining the existing models to enhance their ability to detect features like immature fruit and calyx attributes. Collaboration with growers and field experts will allow for the integration of real-world feedback, ensuring that the models are tailored to practical farming needs.

References

Huang, J., Rathod, V., Birodkar, V., Myers, A., Lu, Z., Votel, R., Chen, Y., & Chow, D. (2017). TensorFlow 2 Detection Model Zoo. https://github.com/tensorflow/models/blob/master/research/object_detection/g3doc/tf2_detection_zoo.md. (Accessed 26 April 2023).

Mirbod, O., & Choi, D. (2023). Synthetic Data-Driven AI Using Mixture of Rendered and Real Imaging Data for Strawberry Yield Estimation. In 2023 ASABE Annual International Meeting (p. 1). American Society of Agricultural and Biological Engineers.

Contact

Dr. Daeun “Dana” Choi
UF/IFAS Gulf Coast Research and Education Center
P: 813-419-6591
E: dana.choi@ufl.edu
smartaglab.com