Cost-efficient broccoli head phenotyping using aerial imagery and SfM-based weakly supervised learning

Haozhou Wang[1†], Tang Li[1†], Erika Nishida[1], Yuya Fukano[1], Yoichiro Kato[1], and Wei Guo[1*]
[1]Graduate School of Agricultural and Life Sciences, The University of Tokyo, Tokyo, Japan
[*]Corresponding author. Email: guowei@g.ecc.u-tokyo.ac.jp
[†]Contributed equally to this work.

Background

Monitoring the broccoli head growth status is important for farmers to make prompt and reliable management decisions. For large-scale fields, aerial imagery and high-throughput phenotyping techniques are often used to decrease the investigation workloads. However, the broccoli head detection and segmentation tasks are still challenging due to the variate object size, texture, leaf occlusion, and various light conditions. Although deep learning shows great potential to tackle this task, it often requires annotating a tremendous number of labeled images and computation costs to train and apply a model.

Methods

In this study, we sued weakly supervised learning to decrease the labor cost of data annotation, the backward projection to original aerial images to increase the accuracy of broccoli head phenotyping. Firstly, we use the simple budding stage to obtain the reference position of broccoli on aerial images generated digital orthomosaic (DOM) by YOLO v5. Secondly, we divied the whole field into small grids. We chose one corresponding aerial image for each grid and added it to training datasets based on its distance to the image center. Thirdly, we used a well-trained BiSeNet model by weakly supervised learning to segment the broccoli head from the small neighbor region derived by the reference positions. Finally, the size-related traits for each broccoli head are calculated by the segmentation results.

Results & Conclusions

The results showed that the proposed method could produce annotation data with intersection over union (IoU) over 85% after just two iterations. The segmentation guided by reference position has better performance and faster training time than the Mask-RCNN segmentation framework. Moreover, the measured broccoli head-related traits by the proposed method have been evaluated by the field measurement. As for future steps, we would like to estimate the weight of each broccoli head and predict the best harvest time according to market price.

Keywords: time-series analysis, structure from motion (SfM), BiSeNet, YOLO, data annotation, unmanned aerial vehicle

Funding

The University of Tokyo Agricultural Science Emergence Fund, Japan Science and Technology Agency (JST) AIP Acceleration Research “Studies of CPS platform to raise big-data-driven AI agriculture”, SICORP Program JPMJSC16H2.