

Doctoral Thesis (博士論文)

**Studies on 3D-based plant phenotyping
by multi-scale data fusion**

(マルチスケールデータ融合による植物表現型の3次元計測に関する研究)

王 浩舟

Wang Haozhou

Abstract

Climate change affects agriculture and food production. It is essential to improve the efficiency of crop breeding to ensure a sustainable increase in crop productivity for global food security. Likewise, high throughput plant phenotyping that quantitatively evaluates crop growth and its interaction with the environment plays an important role in precision agriculture. Recently, many methodologies have been developed to improve the efficiency and accuracy of plant phenotyping using remote and proximal sensing and image processing techniques. These methods can now measure many phenotypic traits that are measurable in two-dimensional (2D), such as canopy coverage and organ number. However, it is still difficult to measure those traits that need to be measured in three-dimensional (3D), such as structure and volume, especially in field conditions. Recently, the photogrammetry technique has been widely used in 3D measurement in various fields. Photogrammetry can generate 3D model of an object from 2D images, it is normally combined with aerial (Long-range) or terrestrial (Close-range) sensing platforms. In agriculture, aerial sensing combined with photogrammetry can efficiently obtain the canopy 3D model of the entire field, but its quality is often not enough for 3D analysis. The quality is often affected by wind-caused blurring and canopy-occlusion-caused structure loss. Terrestrial sensing combined with photogrammetry can obtain a high-quality 3D model, but the efficiency and throughput of current approaches cannot satisfy the demand of thousands of plants in the field. Thus, outdoor 3D plant phenotyping with high quality and throughput remains a difficult challenge.

To address the challenges mentioned above, we presented a multi-scale data fusion method that combines the strengths of close-range (high quality) and aerial (high throughput) surveys. Taking field-grown broccoli as a representative of row-planted

crops having harvestable organs on the top of the canopy, the thesis consisted of three studies:

1. Develop the close-range 3D-based phenotyping pipeline, which can obtain high-quality 3D plant structural models and calculate morphological traits from 1D to 3D.
2. Develop the aerial 3D-based phenotyping pipeline, which can obtain the 3D canopy model of the entire field and improve the accuracy of 2D morphological traits by fusing with raw aerial images and time-series data.
3. Fuse the 3D structural models from the close-range photogrammetry and the aerial photogrammetry, which can calibrate the aerial-measured 2D morphological traits and provide better 3D visualization in the field.

1. Close-range 3D-based phenotyping pipeline

The objective was to develop and validate the close-range 3D phenotyping pipeline, which obtains high-quality 3D models of destructively sampled plants (a broccoli head as an example) and calculates morphological traits. Although lots of software (photogrammetry-based) could generate the 3D model from images, users still need to conduct the process from image acquisition to parameter tuning for trait measurement. Users often deal with one plant by one manually, which is not suited for large population sizes. This chapter first presented an almost-automatic workflow that captures and saves the close-range image of the target crop in multiple view angles (perspectives). Then, regions of interest (ROI) on the broccoli heads in the images were extracted by two pre-trained deep learning models. The preprocessed images were then fed into photogrammetry-based software (Agisoft Metashape) to generate 3D models using automatic processing scripts. Finally, the broccoli crown part was automatically segmented, the 3D model coordinates were corrected, and the phenotypic traits were calculated automatically. To evaluate the performance of the proposed pipeline, we compared some of the traits measured through the pipeline with manually measured ones. Statistically high correlations were observed between manual measurements and pipeline calculated

traits ($R^2 > 0.96$) with root mean square error (RMSE) $< 0.55\text{ cm}$. These findings indicated that the proposed pipeline has high feasibility and accuracy to achieve the final targets of the study.

2. Aerial 3D-based phenotyping pipeline

The objective was to develop and validate the large-scale aerial 3D phenotyping pipeline using drones equipped with commercial red-green-blue (RGB) cameras. To overcome the challenges of distinguishing small broccoli heads on low-quality 3D canopy models from photogrammetry, we proposed a novel data-fusion method. We first fused the time-series data. By using the broccoli seedling positions extracted from the early seedling stage, which is a very simple detection task, the processing regions of later broccoli head segmentation tasks were narrowed down dramatically. This method combined with active learning, significantly decreased the workload of data annotation for deep learning training and the data amount for deep learning processing. Then, we fused the data between pixel coordinates (raw images) and geographical coordinates (photogrammetry outputs). This method segmented the broccoli head region in the raw drone images, whose quality is better than the photogrammetry outputs but lacks actual scale in its pixel coordinates. To compensate for this scale missing, the projective transformation was used to convert the segmentation results to geographical coordinates in order to calculate the head size. Good correlations ($0.57 < R^2 < 0.74$) between the traits measured manually and those automatically measured through the pipeline were observed.

3. Data fusion for virtual visualization

The objective was to test the idea of data fusion on the pipelines built in the previous chapters. Though the projective transformation used in the previous chapter solved the scale missing problem, the transformed broccoli head positions were not so accurate due to the slight perspective differences. Hence, the piecewise affine transformation was used to better locate the plant positions in the geographical coordinates. Then, automatic machine learning was trained by the destructively sampled broccoli heads, which are

available in both close-range broccoli head database and the canopy 3D models from aerial photogrammetry. The trained auto machine learning (AutoML) model was then used to calibrate the 2D morphological traits from the aerial pipeline. Later, a modified normalized cross-correlation template matching method was used to find the closest template in the close-range 3D model database and transform it back to the aerial 3D canopy model for virtual visualization.

General discussion and conclusions

Overall, the results of three research chapters showed the feasibility of the proposed 3D-based phenotyping pipelines in the tested fields. The first close-range pipeline successfully obtained high-quality broccoli head models without heavy workloads. Based on this pipeline, we built the close-range broccoli head database. The second aerial pipeline successfully obtained better head segmentation results by fusing the raw unmanned aerial vehicle (UAV) images. Meanwhile, by using active learning and time-series data fusion, the workload for deep learning data annotation and processing was decreased. Lastly, the proposed AutoML calibration model successfully improved the accuracy of aerial measured morphological traits; and the template matching approach fixed the structure loss of broccoli heads in the aerial 3D models for virtual visualization. The results and statistical analysis concluded that the research objectives have been achieved, and we can conclude that the research has made a positive contribution to 3D-based plant phenotyping and precision agriculture. Lastly, there is currently no strong evidence that the proposed methods can achieve similar performance in different farmland or vegetables, which is suggested for further studies.

Contents

Abstract	i
List of Figures	vii
List of Tables	viii
List of Abbreviations	ix
1 General introduction	1
1.1 Challenges for agronomy and horticulture	1
1.2 Plant phenotyping techniques	2
1.2.1 Data collection	2
1.2.2 ROI extraction	7
1.2.3 Traits calculation and application	15
1.3 Challenges in plant phenotyping	17
1.3.1 Complex cultivation environments	17
1.3.2 Occlusion effects	18
1.3.3 Low quality from aerial	18
1.3.4 Low efficiency from close-range	19
1.3.5 Workloads in data annotation	19
1.3.6 Research reproducibility	21
1.4 Objectives of this study	21
1.5 Outline of this study	22
2 Close-range 3D-based phenotyping pipeline	25
2.1 Introduction	25
2.2 Methods and Materials	28
2.2.1 Plant materials	28
2.2.2 Plant 3D model acquisition	29
2.2.3 Morphological traits extraction	32
2.2.4 Validation	35
2.3 Results	35
2.3.1 Plant 3D model acquisition	36
2.3.2 Morphological traits extraction	37
2.4 Discussion	39
3 Aerial 3D-based phenotyping pipeline	41
3.1 Introduction	41
3.2 Materials and Methods	44

3.2.1	Plot conditions and field data collection	44
3.2.2	Data collection and preprocessing	46
3.2.3	Seedling position detection	48
3.2.4	Broccoli head segmentation	49
3.2.5	Phenotyping of head traits	51
3.2.6	Field validation for head traits	52
3.3	Results	52
3.3.1	Broccoli position detection	52
3.3.2	Broccoli head segmentation	53
3.3.3	Head diameter validation	54
3.4	Discussion	54
3.5	Appendix	57
3.5.1	Backward projection methodology	57
3.5.2	Supplementary video	62
3.5.3	Supplementary figure and tables	62
4	Data fusion for virtual visualization	65
4.1	Introduction	65
4.2	Methods and Materials	69
4.2.1	Data collection and preprocessing	69
4.2.2	Optimized positioning	69
4.2.3	Template matching and transformation	70
4.2.4	3D virtual visualization	71
4.3	Results	72
4.3.1	Optimized positioning	72
4.3.2	Template matching and transformation	72
4.3.3	3D virtual visualization	74
4.4	Discussion	75
5	General discussion	79
5.1	Reseach limitations	79
5.1.1	Lacking universality	79
5.1.2	Lacking real farmland practices	82
5.1.3	Lacking full participation	84
5.2	Future research prospects	85
5.2.1	Probabilistic of this study	85
5.2.2	Probabilistic of 3D reconstruction	87
5.2.3	Probabilistic of digital twin	90
5.2.4	Probabilistic of Foundation AI models	93
6	General conclusion	96
Acknowledgements		99
References		101
Publications		120

List of Figures

1.1	Example of platforms for plant phenotyping	3
1.2	Common approaches for 3D structure reconstruction	5
1.3	The compairson of 2D-based and 3D-based data format	14
2.1	Current photogrammetry (3D reconstruction) methods and challenges	26
2.2	The plant 3D model reconstruction workflow	30
2.3	The algorithms for plant 3D model analysis	33
2.4	Two deep learning segmentation results	36
2.5	The obtained 3D models	37
2.6	Examples of plant 3D model analysis at different growing stages	38
2.7	The comparison between the proposed close-range pipeline and manual measurement	39
3.1	Organ-level analysis challenges of UAV-based approach	43
3.2	Workflow and schematic of the UAV-based pipeline	44
3.3	Plot conditions and timelines for field operation and data collection	45
3.4	Workflow of broccoli seedling detection and head segmentation	48
3.5	Comparison of broccoli head diameter measured from UAV images and obtained using manual field measurements	55
S3.1	Backward projection illustration graph	59
S3.2	Examples of 2020 broccoli seedling position detection and head segmentation by interactive annotation	63
S3.3	Head size of all broccoli heads based on UAV imagery	64
4.1	Challenge to forward results from pixel coordinate to geographical coordinate	67
4.2	Control points between geographical coordinates on DOM and the pixel coordinate on raw image	70
4.3	Head segmentation forward location comparison between projective and piecewise affine transformation	73
4.4	Validation for the AutoML calibration model	74
4.5	The closest matched broccoli head template 3D model	75
4.6	One example of 3D virtual visualization for optimal harvest date	76
5.1	The differences between the experimental field and actual farmland	82

List of Tables

1.1	Softwares for 3D reconstruction using photogrammetry	8
2.1	Meteorological data during the experiment period (Nishida, 2023)	29
2.2	Destructively sampled broccoli head number during the flowering season	30
S3.1	Mid intersection of union (IoU) changes around each iteration in weakly supervised learning.	62
S3.2	Meteorological data during the broccoli growth period.	62
4.1	The components of AutoML calibration model	73

List of Abbreviations

AI	artificial intelligence.
ALS	aerial laser scanner.
ANN	artificial neural network.
API	application programming interface.
AR	augmented reality.
AutoML	auto machine learning.
CART	classification and regression tree.
CCD	charge-coupled device.
CG	computer graphics.
CMOS	complementary metal oxide semiconductor.
CNN	convolutional neural network.
CPU	central processing unit.
CUDA	compute unified device architecture.
DBSCAN	density-based spatial clustering of applications with noise.
DGCNN	convolutional neural network over large-scale labeled graphs.
DOM	digital orthomosaic map.
DSLR	digital single-lens reflex.
DSM	digital surface map.
ExG	excess green index.
GB	Gigabytes.
GBDT	gradient boosting decision tree.
GCN	graph neural network.
GCP	ground control point.
GIS	geographic information system.
GMM	Gaussian mixture model.
GNSS	Global Navigation Satellite System.
GPS	Global Positioning System.
GPU	graphics processing unit.
GUI	graphical user interface.
HLS	handheld laser scanner.
HSV	hue-saturation-value.
IoT	internet of things.
ISAS	Institute for Sustainable Agro-ecosystem Services.
KD-Tree	k-dimensional tree.

KNN	k-nearest neighbor.
LDA	linear discriminant analysis.
LiDAR	light detection and ranging.
LOWESS	locally weighted scatterplot smoothing.
ML/DL	machine learning / deep learning.
MLR	multiple linear regression.
MNVG	median normalizedvector growth.
MVCNN	multi-view region-based convolutional neural network.
MVS	multi-view stereo.
NARO	National Agriculture and Food Research Organization.
NDVI	normalized difference vegetation index.
NeRF	neural radiance fields.
NLP	natural language processing.
NMS	non-maximum suppression.
OBB	oriented bounding box.
ORCNN	occlusion-aware region-based convolutional neural network.
PCT	point cloud transformer.
RAM	random access memory.
RANSAC	random sample consensus.
RCNN	region-based convolutional neural network.
RF	random forest.
RGB	red-green-blue.
RGBD	red-green-blue-depth.
RGRI	red-green ratio index.
RMSE	root mean square error.
ROI	regions of interest.
RTK	real time kinematic.
SfM	structure-from-motion.
SVM	supported vector machine.
TLS	terrestrial laser scanner.
ToF	time of flights.
UAV	unmanned aerial vehicle.
UGV	unmanned ground vehicle.
UI	user interface.
URL	uniform resource locator.
VCNN	voxel-based convolutional neural network.
VDVI	visible-band difference vegetation index.
VR	virtual reality.

Chapter 1

General introduction

1.1 Challenges for agronomy and horticulture

The world population is expected to increase to 9.6 billion in 2050 and 10.9 billion in 2100 (Gerland et al., 2014). For sustainable food security, the current yield gains need to be doubled by 2050 (Ray et al., 2013; Tilman et al., 2011). There is a need for a significant increase in global agricultural productivity, particularly, maize, rice, wheat and soybean must increase to 2.4% of annual yield gains to meet the growing demand (Ray et al., 2013). Meanwhile, except for the food plants that can provide energy, fruits and vegetables are also important as sources of trace elements, vitamins, and minerals in human nutrition. However, for one thing, it is challenging for plant breeders to significantly increase crop productivity by developing new varieties, which often requires several years. For another, climate change also has negative impacts on productivity. For example, extreme weather like drought, flooding, frost, and hail can significantly decrease the yield (Lobell and Gourdji, 2012); while the global warming also hurts overwintering crops that need a cold accumulation period to produce harvestable organs (i.e. vernalization), like wheat. In addition, field management practices (e.g., tillage, density, and nutrients) also affect crop yield and quality (Jackson et al., 2004; Satodiwa et al., 2015). Hence, to better assist breeders to accelerate variety development and assist agronomists, extension workers and farmers to make proper management decisions to changing environments, it is important to monitor the plants more accurately

and more frequently during their growth stages.

Conventional agricultural management requires significant manual labor and is now facing several challenges. Firstly, it often requires professional knowledge and may be subject to human error. Secondly, the availability of human labor in agriculture is currently decreasing due to long-term global events, such as urbanization and aging populations, as well as short-term pandemics, such as economic recessions and coronavirus disease 2019 (Gallardo and Sauer, 2018; Larue, 2020). As a result, there is an unprecedented demand for labor-saving technologies. The development of technologies, such as automation, sensing, big data analysis, computer vision, and artificial intelligence, has made it possible to address such demands and thereby spark the digital revolution in the agricultural industry (Gallardo and Sauer, 2018).

1.2 Plant phenotyping techniques

Plant phenotyping, which includes labor-saving and digital revolution, has been rapidly developed and implemented in recent years (Araus and Cairns, 2014). It is defined as “*the application of methodologies and protocols to measure specific traits related to plant structure or function with traits ranging from cellular to whole-plant levels*” (Fiorani and Schurr, 2013; Ghanem et al., 2015). Although different workflows are required for different crops and applications, the general workflow of plant phenotyping is summarized as follows: 1) data collection, which gathers plant data using various sensors; 2) regions of interest (ROI) extraction, which detects or segments plant parts at expected levels (e.g., full canopy or single organ) from the collected data; and 3) crop trait calculation and guidance for practical applications (Zhao et al., 2019).

1.2.1 Data collection

For data collection, there are various sensors available for different purposes in plant phenotyping. These sensors can be categorized as environmental sensors (e.g. light intensity, temperature, and humidity) and plant sensors (e.g. organic compounds and plant images) (Garlando et al., 2020). Environmental sensors are typically fixed

in place as part of the internet of things (IoT) system to record critical environmental factors that affect plant phenotypes (Ghanem et al., 2015). On the other hand, plant sensors are usually mounted on platforms and moved around, as shown in Figure 1.1. These platforms enable data collection at different scales, from the organ level to the individual and canopy levels, as demonstrated in both indoor (Figures 1.1a-d) and outdoor (Figures 1.1e-k) settings.

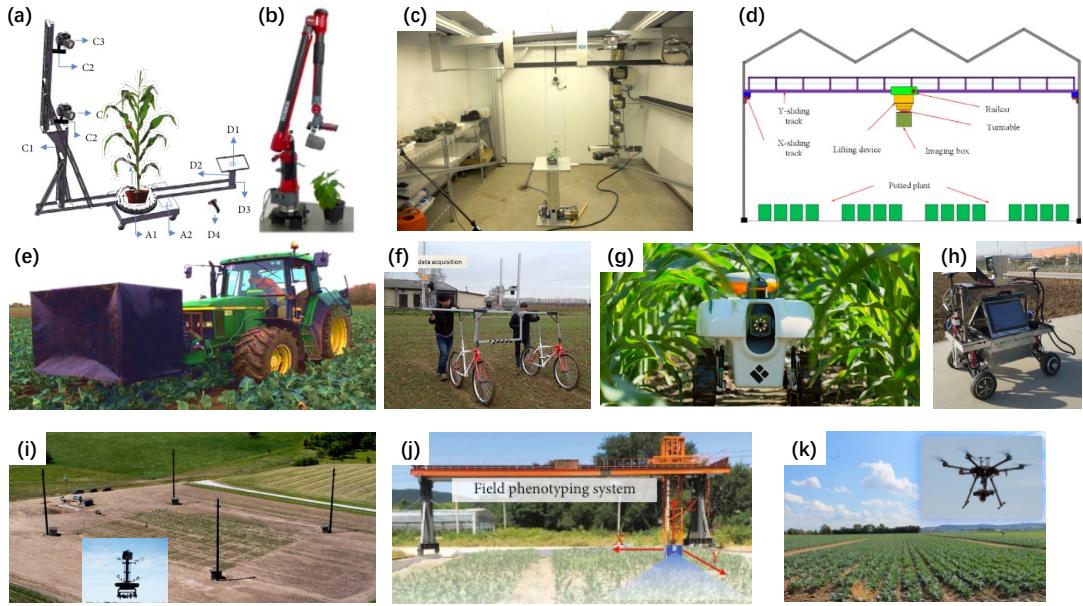


Figure 1.1: Example of platforms for plant phenotyping. (a-b) self-designed indoor devices (Schunck et al., 2021; Wu et al., 2020); (c-d) robotic arms in growth chamber or greenhouse (Chaudhury and Barron, 2018; Du et al., 2021); (e) tractor (Kusumam et al., 2017); (f) ground vehicle (Liu et al., 2017b); (g-h) unmanned ground vehicle (UGV) (McGuire et al., 2021; Qiu et al., 2019); (i-j) robotic outdoor instruments (Bai et al., 2019; Jin et al., 2021a); and (k) unmanned aerial vehicle (UAV) (Kierdorf et al., 2022)

An important part of plant sensors is imaging sensors, which can record the morphological information of crops and are therefore commonly used in many plant phenotyping studies (Feng et al., 2021; Paulus, 2019). For example, infrared, multi- and hyper-spectral imaging sensors were used to calculate several vegetation indices (Han et al., 2019), such as normalized difference vegetation index (NDVI), to assess biomass (Jimenez-Berni et al., 2018), yield, and water stress level (Herrero-Huerta et al., 2020b; Romano et al., 2011), and to evaluate postharvest senescence of broccoli (Guo et al., 2022). The common red-green-blue (RGB) camera was also used to count plant number

(Liu et al., 2022), plant density (Velumani et al., 2021), detect harvestable organs of cereal and horticultural crops (Blok et al., 2016), and assess their quality (Stansell et al., 2017). These studies demonstrated the feasibility of using imaging sensors in plant phenotyping due to their high-efficiency and non-destructive nature.

However, these imaging sensors are difficult to use in directly describing the 3D morphological structure of plants, due to occlusion and dimension loss when projecting onto the 2D plane of photosensitive elements. As a result, inaccuracies and uncertainties arise when describing the 3D structure. With the development of sensing techniques, several studies have reviewed the latest available approaches for 3D plant phenotyping (Kochi et al., 2021; Okura, 2022; Paulus, 2019). Figure 1.2 summarizes some of them, which are comprised of both active and passive scanners. A complete list of commercial 3D scanners is also provided by Bartol et al. (2021).

The active scanners use a light source to project on the object and analyze the reflection results to obtain the structure. It consists of two main parts: triangulation and ToF. The triangulation approach uses the shape changes on the projected straight laser line(s) to obtain the object structure. For example, Schunck et al. (2021, Figure 3) used a laser line (Fig. 1.2a) triangulation scanner (Perceptron Scan Works V5, Perceptron Inc., USA) to build a public 3D model dataset for maize and tomato. Structured light improves efficiency by projecting an array of lines which is often visible or infrared light, rather than laser (Fig. 1.2b). The Microsoft Kinect V1 is one of the famous red-green-blue-depth (RGBD) cameras based on this approach, which was released in November 2010. However, it was used only in a few plant phenotyping studies (Nguyen et al., 2015). Since the release of the second version (Kinect V2, in July 2014) and the third version (Azure Kinect, in March 2020) with a different ToF approach and better performance (Lachat et al., 2015; Tölgessy et al., 2021). The ToF measures the distance between the sensor and points on the object to obtain its structure. Depending on the type of light source used, it can be divided into two categories: infrared for ToF cameras (Fig. 1.2c) and laser for light detection and ranging (LiDAR) (Fig. 1.2d). Since the sun is a massive source of infrared, infrared-based ToF cameras typically do not

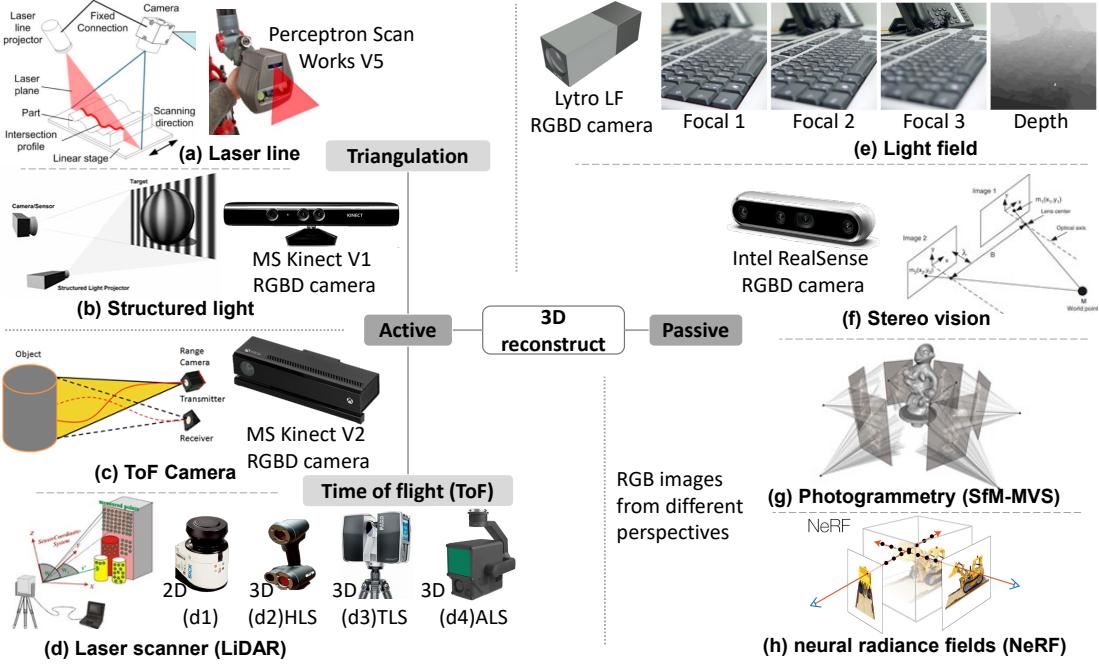


Figure 1.2: Common approaches for 3D structure reconstruction; (a-d) The active sensors rely on projecting lights on the object and analyzing the reflection results; (a) the theory of laser line (Zhou et al., 2020c, Fig. 1) and one commercial device used by Schunck et al. (2021, Fig. 2); (b) the theory of structured light (Paturkar et al., 2021, Fig. 4) and one commercial device (Duc et al., 2015, Fig. 1); (c) The theory of time of flights (ToF) camera (Jamtsho, 2010, Fig. 2.1) and one commercial device (<https://en.wikipedia.org/wiki/Kinect>); (d) the theory of light detection and ranging (LiDAR) (Khomsin et al., 2019, Fig. 4); (d1) 2D lidar scanner (<https://www.sick.com/de/en/lidar-sensors/2d-lidar-sensors/lms1xx/c/g91901>) (d2) handheld laser scanner (HLS) (Ameen et al., 2018, Fig. 2b); (d3) terrestrial laser scanner (TLS) (<https://downloads.faro.com/index.php/s/89t7CjYKcjGZd3R> Fig. 1-1); (d4) aerial laser scanner (ALS) (<https://enterprise.dji.com/zenmuse-l1>); (e-f) the passive sensors rely on analyzing the passively received image groups; (e) one commercial light field camera and its collected data (Schima et al., 2016, Fig. 1&6); (f) the theory of stereo vision (Kim et al., 2003, Fig. 1) and one commercial device (<https://www.intelrealsense.com/depth-camera-d455>) (g-h) the RGB-based 3D reconstruction algorithms from several perspective images, via (g) conventional geometric deduction (Shalma and Selvaraj, 2023, Fig. 6) and (h) deep learning (<https://www.matthewtancik.com/nerf>)

perform well in outdoor environments with intense sunlight (Tölgyessy et al., 2021). Therefore, they are commonly used for indoor plant phenotyping (Martinez-Guanter et al., 2019; Xu et al., 2023; Zhang et al., 2020a). In contrast, lasers are often more tolerant to sunlight due to their higher energy density, making them widely used in outdoor environments. There are different types of LiDAR sensors used in 3D plant phenotyping applications. These include 2D LiDAR scanners (Garrido et al., 2015),

3D handheld laser scanner (HLS) (Ma et al., 2019), 3D terrestrial laser scanner (TLS) (Qiu et al., 2019; Su et al., 2018; Wu et al., 2019), and 3D aerial laser scanner (ALS) (Nguyen et al., 2023; Ten Harkel et al., 2019).

The passive scanners rely on analyzing passively received image groups, mainly using RGB images. The light field camera takes a group of photos with different focal lengths, and the object structure is calculated according to different degrees of clarity caused by the different distances from the sensor (Fig. 1.2e). For example, Apelt et al. (2015) built a light field camera system for measuring morphological traits related to plant growth. A commercial low-cost light field camera (Lytro LF, Lytro Inc., USA) was used as an RGBD camera to monitor maize 3D morphological traits (Schima et al., 2016). Another commercial RGBD camera, Intel RealSense L515(Intel Corporation, USA), uses a stereo vision approach (Fig. 1.2f). It uses binocular vision like human eyes to obtain the object's structure. While for other RealSense models (D455, D435i, and D415) integrate the RGB camera with LiDAR (Bartol et al., 2021, Table 5). Blok et al. (2021b) used this RealSense RGBD camera to estimate the broccoli head size with different degrees of occlusion, which is a tough problem for common RGB cameras.

The photogrammetry approach is based on RGB images obtained from different perspectives (view angles) using common imaging sensors (Fig. 1.2g). It first uses the overlapped area among images to estimate the camera poses and object's rough 3D structure (tie points), called structure-from-motion (SfM). Then the multi-view stereo (MVS) is applied to densify the 3D point cloud of the tie points, and surface reconstruction and texture rendering are performed to obtain a 3D mesh model of the objects. For more details, please refer to Hartley and Zisserman (2003) and Snavely et al. (2010).

Unlike the aforementioned scanners (Figures 1.2a-f), which often require special, not-so-low-cost devices, the structure-from-motion (SfM) approach only requires an RGB camera and SfM software, with very flexible cost advantages. For the camera/sensor parts, even the cameras of smartphones at hand can be used to obtain photos for plant modeling (Li et al., 2020). If 3D plant models with ultra-high quality are

required, a digital single-lens reflex (DSLR) camera with better complementary metal oxide semiconductor (CMOS) (image sensor) and lens, can also be used to raise the 3D model quality to a higher level (Drofova et al., 2023; Nguyen et al., 2016), often exceeding the previously commercial RGBD cameras. Meanwhile, a large number of SfM open-source and commercial software packages are available for individual plants (close-range) and geo-referenced canopy (aerial) applications (Table 1.1). For this reason, many studies have used this approach to obtain 3D models of indoor individual plants (Wu et al., 2020; Zhou et al., 2019), in-field individual plants (Herrero-Huerta et al., 2023; Jay et al., 2015), and in-field canopy (Herrero-Huerta et al., 2020a; Kim et al., 2018).

Recently, a novel deep learning method by Mildenhall et al. (2022), called neural radiance fields (NeRF), was introduced to create 3D models from different perspectives. The proposed NeRF is based solely on the SfM step and can bypass the time-consuming MVS, 3D meshing, and 3D rendering steps required by photogrammetry. The official website (<https://www.matthewtancik.com/nerf>) demonstrates the feasibility of obtaining plant 3D structures, and Jignasu et al. (2023) obtained the maize 3D structure using this approach. However, the effectiveness of this approach relies on the quality of training data, and while there are many datasets available for industrial applications, there are not many in agricultural sciences. As a novel technology, its application in agriculture still requires further development and thus is not included in this study.

1.2.2 ROI extraction

After collecting the data, the next step is to extract the regions of interest (ROI) for further analysis. This section summarizes different methods and algorithms for image analysis and 3D data analysis, including data processing, 2D-based image analysis from imaging sensors, and 3D-based point cloud analysis from 3D photogrammetry or LiDAR scanner.

Table 1.1: Softwares for 3D reconstruction using photogrammetry; “Close-range” is mainly used for building 3D models of individual plants or organs at the local coordinate; “Aerial” is mainly used for building geo-referenced canopy 3D models using UAV imagery with geographical coordinates.

Types	Software name	close-range	aerial
Open source	AliceVision Meshrooms	✓	✓
	COLMAP	✓	✓
	Multi-View Environment (MVE)	✓	✗
	OpenDroneMap ¹	✗	✓
	OpenMVG ² → { MVE OpenMVS }	✓	✓
	VisualSFM ² → { MeshRecon OpenMVS PMVS }	✓	○ ³
	3DF Zephyr	✓	✓
	Agisoft Metashape	✓	✓
	Autodesk Recap Photo	✓	✓
	ContextCapture	✓	✓
Commercial	Correlator3D	✗	✓
	DJI Terra ⁴	✗	✓
	DroneDeploy	✗	✓
	Elcocision 10	✓	✓
	Pix4Dmapper	✗	✓
	Reality Capture	✓	✓

¹ charges for a complied installer for whom is difficult to install from source code;

² some software only provide SfM pipeline, need to integrate them with MVS software as a complete photogrammetry pipeline;

³ has video tutorials for processing UAV images, but no official documentation about making geo-referenced GeoTiff for common aerial products like digital orthomosaic map (DOM) and digital surface map (DSM);

⁴ is optimized for DJI (Shenzhen DJI Technology Co., Ltd. China) drones only.

1.2.2.1 Data preprocessing

For canopy-level or full-field phenotyping applications, the digital orthomosaic map (DOM) is often used as the field plot map (by UAV photogrammetry); the 3D point cloud is often used as 3D canopy models (by UAV photogrammetry or LiDAR scanner). However, processing the whole file image or 3D data directly is often impossible, since the file size of these whole canopy data is very large, often over 0.5 Gigabytes (GB) and can reach over 10 GB. One common data preprocessing step in plant phenotyping applications is to split the whole field into smaller parts. It can be split into equal-sized

grids. For example, split the full DOM image into several 250×250 pixel grids (Bauer et al., 2019) or by the boundary of each (micro) plot (Tresch et al., 2019). The latter option is more meaningful and often used.

To manually place plot boundaries and their labels on the DOM map, geographic information system (GIS) software such as ArcGIS Desktop (Esri, Redlands, USA) or QGIS (open source <https://qgis.org>) is often utilized. The (micro) plot results are typically saved in shapefile format (*.shp) for better compatibility with other software. However, manually editing and operating software graphical user interface (GUI) can be time-consuming for large fields. Consequently, several studies have attempted to solve automatic ROI detection and generation by using computer vision. Tresch et al. (2019) developed an open-source Python package called EasyMPE for the semi-automatic generation and cropping of ROI. Subsequently, this tool was re-built using C# by National Agriculture and Food Research Organization (NARO) and renamed PREPs (<http://cse.naro.affrc.go.jp/aitoh/PREPs>) for higher performance on the Windows platform. Chen and Zhang (2020) also developed a Python-based tool called GRID (<https://zzlab.net/GGRID>); Mortensen et al. (2019) used MATLAB to create such a tool, and Mardanisamani et al. (2021) extended these tools by introducing the capability to slightly adjust the ROI location for better canopy fitting. All of these tools also support cropping the entire DOM into smaller parts within the ROI for post-processing.

For the 3D canopy point cloud, point cloud processing software such as MeshLab (open source, <https://www.meshlab.net>) or CloudCompare (open source, <https://www.cloudcompare.org>) is often used for similar tasks of cropping ROI, but it is not very convenient. Instead, it is more common to use the ROI shapefiles created in the previous step and write batch scripts for processing. For example, Sun et al. (2018) used a Matlab script to crop the ROI on 3D point cloud, but they did not publish their code. Even after segmenting the point cloud into smaller parts, point downsampling and noise removal are often necessary to reduce data size and processing difficulties due to the disorderliness of the point cloud's data structure (Ma et al., 2019).

After preprocessing the entire field into smaller parts and data sizes, the following

2D- and 3D-based approaches can be applied to each part.

1.2.2.2 2D-based approaches

Most of the image analysis tasks for plant phenotyping can be summarized into three categories: classification, detection, and segmentation. The classification task involves determining the class of objects present in an image. For example, deciding whether a broccoli head is healthy or diseased (García-Manso et al., 2021). The detection task combines classification and localization to identify the kind of objects and their locations in an image. For instance, detecting the number of maize seedlings (Liu et al., 2022) or sorghum tassels (Ghosal et al., 2019) in an image. The segmentation task aims to separate an image into distinct regions with particular shapes and borders. It can be further divided into semantic segmentation (where a single label is assigned to all objects belonging to a particular class, e.g., all the regions of plants or harvestable organs in an image) and instance segmentation (where unique labels are assigned to each object, e.g., the regions of each harvestable organ in an image). These tasks (classification, detection, and segmentation) are not independent of one another, as semantic segmentation often involves classifying each pixel to obtain the segmentation result of the entire image (Guo et al., 2017), while instance segmentation usually requires combining the results of semantic segmentation and detection to identify each object's region (Lüling et al., 2021, see Fig. 2).

The most conventional approach is to use computer vision algorithms on common RGB images. In some simple cases, a manually defined color threshold can be used to separate plants from their backgrounds. For example, Das Choudhury et al. (2018) converted maize plant images from RGB to hue-saturation-value (HSV) and used $\{0.051 < H < 0.503; 0.102 < S < 0.804; 0.000 < V < 0.786\}$ to separate plant regions from the background. However, in most outdoor images with complex lighting and background conditions, manual thresholding does not perform well. Therefore, Meyer and Neto (2008) proposed color vegetation indices (NDI, ExG, and ExR) and combined them with Otsu (1979) thresholding algorithm for automated crop imaging applications.

To enhance compatibility with more complex scenarios and not be restricted to green plants, Guo et al. (2017) manually annotated the training data and trained machine learning classifiers based on extended color information. Additionally, Zou et al. (2019) and Blok et al. (2016) segmented broccoli buds and heads, respectively, using texture and color information and a trained supported vector machine (SVM) classifier.

In the same way, deep learning techniques show great potential for achieving better results in RGB image processing. The convolutional neural network (CNN) is a well-known deep learning framework for color imagery and has been used by many phenotyping studies. For instance, Ghosal et al. (2019) employed a pre-trained CNN model to detect and count sorghum heads while Liu et al. (2022) used the Fast region-based convolutional neural network (RCNN) to count maize seedlings. Additionally, Bender et al. (2020) used FastRCNN to detect and segment entire broccoli on RGB images and shared them as public datasets. García-Manso et al. (2021) applied FasterRCNN to differentiate between immature and diseased broccoli. Blok et al. (2021a) simplified the MaskRCNN with data augmentation to improve broccoli head segmentation on RGB images, but accuracy was affected by leaf occlusion. To address this issue, Blok et al. (2021b) used occlusion-aware region-based convolutional neural network (ORCNN) to recover the hidden parts of harvestable organs.

1.2.2.3 3D-based approaches

Although 2D-based machine learning and deep learning approaches show the feasibility of extracting ROI, for very complex scenes, it often requires labeling a considerable number of training sets to ensure better performance. The 3D-based approaches cannot only provide more morphological information but also greatly simplify complex tasks.

The depth information is often used as a common and simple 3D-based approach. The depth image is a common file format for this information, which uses pixels to record depth values. It can be obtained from either depth cameras (often for close-range; the depth value is the distance from the object to the camera) or photogrammetry (often for aerial; the depth value is the height or altitude). Such information has been widely used

for better plant segmentation for phenotyping studies. For close-range applications, Lüling et al. (2021) used the depth image generated by photogrammetry and color information to segment cabbage instances. For aerial photogrammetry applications, Guo et al. (2020) segmented the plant area inside each ROI using the color and depth information from DOM and DSM, respectively.

Another format for recording depth information is the 3D point cloud. Analyzing the 3D point cloud is also of great importance for the 3D-based approach. The general processing workflow for this analysis includes: 1) data preprocessing to decrease data size (see Subsection 1.2.2.1); 2) semantic segmentation for separating foreground (plants) and background (soils, etc.); 3) instance segmentation of each plant for individual-level studies; and 4) instance segmentation of each organ for organ-level studies. The conventional methods are introduced first, followed by stepping to 3D-based deep learning approaches.

To perform 3D-based semantic segmentation between plants and backgrounds, color thresholding can be applied to 3D point clouds with colors as well as in the 2D-based approach. In Xiao et al. (2020, Fig. 3), a manually defined RGB threshold ($R - G \geq 7$) was used to separate plant points from backgrounds like shadow and soil points. The geometry relationship among points can also be used for segmentation. Ge et al. (2019) applied Gaussian mixture model (GMM) clustering to recognize the points of broccoli buds and removed noises using the k-nearest neighbor (KNN) algorithm. Instead of extracting plant parts, Garrido et al. (2015) used random sample consensus (RANSAC) plane regression to identify flat ground points and removed them.

After performing 3D-based semantic segmentation of the plant part, the 3D-based instance segmentation to split each plant should proceed for individual-level studies. Hofle (2014) developed a region-growth method based on the local maxima in elevation to split each maize bud from semantic-segmented results. It shares a similar idea with the density-based spatial clustering of applications with noise (DBSCAN) clustering method, which several studies later used for maize (Lin et al., 2022) and cotton balls (Sun et al., 2020). Kusumam et al. (2017) segmented broccoli heads using the Euclidean

clustering method and the SVM classifier on the 3D features, and Montes et al. (2020) further improved its efficiency to almost real-time.

Finally, studies at the organ-level require segmentation of each organ, especially for crops with complex morphological structures (such as maize and tomato). Different studies proposed solutions depending on the morphological characteristics of the crops or cultivars. For maize leaf segmentation, Jin et al. (2019) proposed an median normalizedvector growth (MNVG) algorithm, Liu et al. (2021b) used skeletonization and region growth, Wang et al. (2023a) proposed a distance field-based segmentation pipeline, and Miao et al. (2021) used optimal transportation distance and published an interactive segmentation tool (Label3Dmaize, <https://github.com/syau-miao/Label3DMaize>). For tomato stem and leaf segmentation, Rossi et al. (2022) proposed a complex but automated workflow using Matlab, which includes radius detection, sphere climbing, and phyllotaxy retrieving. Dutagaci (2023) proposed a t-distributed stochastic neighbor embedding segmentation method and demonstrated its feasibility on five different plants (hibiscus, maple, tomato, tobacco, and rosebush). Additionally, Dutagaci et al. (2020) and Schunck et al. (2021) published annotated datasets for evaluating 3D plant organ segmentation methods.

One limitation of the conventional 3D-based analysis method, like the conventional 2D-based image analysis method mentioned above, is the necessity to develop specific algorithms based on the characteristics and objectives of different plants. Existing algorithms often have low compatibility with new crops and objectives. The 2D-based deep learning approach shows the possibility of changing the difficulty from specific algorithm development to training data annotation. But, unlike the well-structured 2D raster image data ($m \times n$ matrix, Fig. 1.3b), it is difficult to apply neural networks directly to the 3D point cloud data with disordered structures (Fig. 1.3d).

Currently, there are four solutions for applying deep learning to the 3D point cloud. The first solution is 3D convolution, which converts the disordered point cloud (Fig. 1.3d) into the ordered voxel model (Fig. 1.3e). One famous deep learning network based on this idea is VoxNet (Maturana and Scherer, 2015), but it is easily affected by

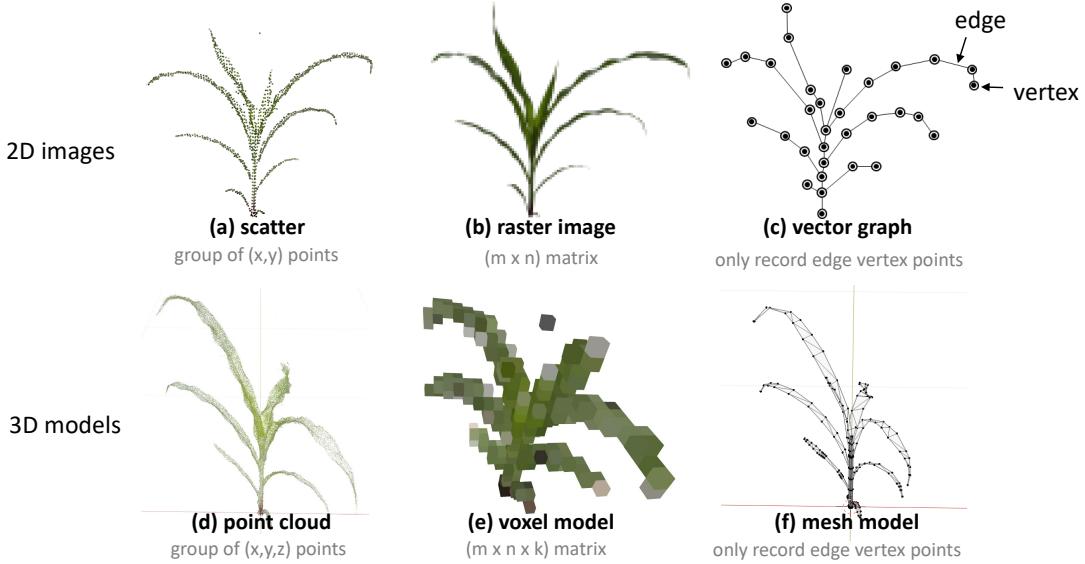


Figure 1.3: The comparison of 2D-based and 3D-based data format; (a) 2D point cloud also called scatter; (b) the common format for 2D images; (c) updated the scatter with relationship between points; (d) 3D point cloud, the most common format for 3D-based analysis; (e) the unit of 2D raster image is pixel, while the unit of 3D raster is voxel; (f) updated the point cloud with relationship, often used in computer graphics (CG) industry.

the resolution (sidelength of the voxel) selection and is limited with small data sizes (maximum of $32 \times 32 \times 32$ voxel numbers). Jin et al. (2020a) developed a similar voxel-based convolutional neural network (VCNN) for maize stem and leaf classification and segmentation.

The second solution is graph convolution, which converts the disordered point cloud into an ordered graph (Fig. 1.3c&f). The graph relationship already supports distinguishing different shape features like leaves (flat planes) and stems (long cylinders) according to self-defined rules (Mirande et al., 2022). However, it is also possible to apply techniques such as graph neural network (GCN) (Wu et al., 2021; Zhou et al., 2020b) or convolutional neural network over large-scale labeled graphs (DGCNN) to the graph data (Phan et al., 2018). For example, Du et al. (2023) tested the feasibility of DGCNN for instance segmentation on rapeseed leaves.

The third solution involves converting multi-view projections into 2D images for 2D-based detection or segmentation, and then re-projecting the results back to the 3D point cloud. One well-known network for this approach is the multi-view region-based

convolutional neural network (MVCNN), which was proposed by Su et al. (2015). Jin et al. (2018) applied this method to segment individual maize plants from a 3D point cloud canopy. Van de Zedde et al. (2019) re-used the geometric relationship between 3D models and the raw images from photogrammetry to project the 2D deep learning results onto the raw images and back onto the tomato 3D point cloud models.

The last solution is to apply deep learning directly on the point cloud. Qi et al. (2016) first trained a deep learning network called PointNet by randomly shuffling the order of points in point clouds to avoid the impact of their order. However, the algorithm's complexity is factorial N, which becomes unacceptable for cloud sizes over 1000 points. Therefore, the PointNet++ (Qi et al., 2017) was proposed to solve the lack of hierarchical feature aggregation in the previous version. Boogaard et al. (2021) tested its feasibility on cucumber segmentation, and PointNet++ also serves as a comparison standard for several newly proposed networks for plants (Du et al., 2023; Jin et al., 2020a; Li et al., 2022a; Zhou et al., 2022). Specifically, Zhou et al. (2022) and Du et al. (2023) proposed point cloud transformer (PCT)-based networks for broccoli head segmentation and rapeseed leaf segmentation, respectively. Li et al. (2022b) proposed a PlantNet and its updated version, PSegNet (Li et al., 2022a), for the organ segmentation of sorghum, tomato, and tobacco.

1.2.3 Traits calculation and application

After segmenting the ROI, several different kinds of trait calculations can be applied to obtain 2D-based and 3D-based results. Du et al. (2021, Table 1) summarized the six categories of 2D-based traits obtained from common RGB images, namely: 1) geometric (morphological) traits, such as projected area, projected perimeter, convex area, convex perimeter, circumcircle diameter, the angle of the main axis, and the length of the long and short axis; 2) structure traits, such as the length or proportion of intersection line segments in concentric circles; 3) color space traits, such as the mean and variance of Red in RGB; 4) color index traits, such as excess green index (ExG), visible-band difference vegetation index (VDVI), red-green ratio index (RGRI); 5) color component

traits, such as the first order color moment of B component; and 6) texture traits, such as contrast, dissimilarity, and homogeneity. Among these, geometric (morphological) traits, color index traits, and texture traits are commonly used in other studies. For example, Stansell et al. (2017) extracted texture traits like head color, head smoothness, head size, and head uniformity as broccoli head quality indices, while Bauer et al. (2019) applied the color index trait NDVI as an indicator of lettuce maturity.

When it comes to 3D-based approaches, besides the aforementioned 2D morphological traits (projected canopy area, convex and concave hull area), the canopy volume and canopy height are also available for cotton canopy (Jiang et al., 2018). For maize canopy applications, other canopy traits such as maize height (Hämmerle and Höfle, 2018; Qiu et al., 2019) and row spacing (Qiu et al., 2019) can also be calculated. Jin et al. (2020b, Table 3) calculated morphological traits of maize from canopy-level to leaf organ-level, including height, canopy cover, plant area index, projected leaf area, volume, stem diameter, leaf length, and width. Itakura and Hosoi (2018) obtained the leaf inclination angle, and Liu et al. (2021b) analyzed its effect on canopy light use efficiency. Additionally, 3D-based traits can also support advanced applications for drought stress (Sorrentino et al., 2020; Su et al., 2019), intercropping (Liu et al., 2021a) and light competition (Zhu et al., 2020) analyses.

Several advanced phenotyping applications can be implemented for breeding activities: 1) time-series analysis can be used to trace the growth changes (Li et al., 2013), to measure the flowering time in rice (Guo et al., 2015) and lettuce (Han et al., 2021); to analyze the seasonal and circadian rhythms in maize structural (Jin et al., 2021a); 2) The ray-tracing technologies can be applied to simulate lights within the canopy. For example, analysis of the contributions of foliar and nonfoliar tissues (Chang et al., 2022), leaf inclination angles (Liu et al., 2021b), and row spacing (He et al., 2021) to canopy photosynthesis efficiency of each crop; and 3) the lodging responses can be analyzed by either multi-spectral reflectance (Wang et al., 2021b,c) or mechanical analysis (Cook et al., 2019; Robertson et al., 2017) or mechanical simulation (Nakashima et al., 2023).

1.3 Challenges in plant phenotyping

Despite the advances in phenotyping approaches, it is still challenging to apply them directly to field-grown crops. The reasons are summarized as follows: 1) the difficulties of data processing caused by complex cultivation environments; 2) the target structure missing caused by canopy occlusion; 3) the insufficient outputs' quality of aerial photogrammetry for organ-level tasks; 4) the heavy labor and low efficiency of the close-range approach, though with better 3D model quality; 5) workloads in data annotation of deep learning; and 6) research reproducibility.

1.3.1 Complex cultivation environments

The complex weather conditions in the field often produce images with varying brightness and cloud shadows. Meanwhile, the in-field applications often have complex backgrounds with different soil textures and various weeds. These factors pose great challenges to ROI detection and segmentation tasks. For some 3D-based approaches using the RGBD cameras, the sun emits a massive amount of infrared radiation, and many infrared-based RGBD cameras cannot perform well in such outdoor environments (Tölgessy et al., 2021). For the photogrammetry method, which assumes the object is solid from different angles, wind-caused leaf movement conflicts with that assumption and often results in the ghosting effect (double mapping), excessive pixelation, and seamline distortions in the final products (Duan et al., 2017; Lin et al., 2021). Even for the high-cost LiDAR sensors using laser, dust particles and water droplets from rain and fog act as noise and affect the quality and accuracy of plant phenotyping applications. All of these environmental factors seriously impact the quality and accuracy of plant phenotyping applications.

Different crop management practices can also lead to variations. For instance, planting methods can differ across farmlands, resulting in varying row spacings and plant densities. In fields with clear between-row intervals, a line detection method can be used to locate the position of each ridge and plot regions (Tresch et al., 2019, Fig. 1). However, in fields lacking clear ridge intervals (Faye et al., 2016, Fig. 3), the

line detection method may not be effective, necessitating the development of alternative approaches. Additionally, different cultivars of the same vegetable may exhibit distinct growth patterns and textures. For example, the application of different fertilizers to the same broccoli cultivar can impact head size and branch patterns (Nishida, 2023). The combination of these factors makes it difficult to develop a universal approach that can handle all scenarios. As a result, the typical approach is to create targeted algorithms for different crops and site characteristics, which can be time-consuming and heavily reliant on the expertise of computer science professionals.

1.3.2 Occlusion effects

With the growth of plants, leaves can partially occlude vegetable fruits. For tomatoes (Yamamoto et al., 2014) and grapes (Liang et al., 2022), individual fruits can also occlude each other, resulting in plant structure loss in both 2D and 3D-based approaches. This can result in missed detection, and partial occlusion often leads to inaccuracies in area and other parameters. The issue of occlusion has become of great interest in the field of plant phenotyping (Blok et al., 2021b; Boogaard et al., 2020; Lehnert et al., 2019).

1.3.3 Low quality from aerial

The conventional analysis of the products (DOM and 3D point cloud) of the aerial survey is not sufficient for the organ level. For one thing, considering the efficiency of surveys and the impact of the wind produced by the drone's propellers, the flying height is often above 10 m. This resolution is sufficient for surveys at the canopy scale, but it is difficult to achieve organ-level details. For many organ-level applications, the images are often collected close to the ground (less than 1 m between a sensor and a plant) by a hand-held camera or an unmanned ground vehicle (UGV). For another, as mentioned in Subsection 1.3.1, the complex cultivation environment naturally conflicts with the solid object assumption of photogrammetry (object without shape changes from different perspectives). Although some of the drone cameras now have a resolution greater than 4K and some drones even support attaching DSLR cameras for photography, making the

organ-level details visible on the raw images, the quality of the generated 3D products is influenced and often decreases compared to the raw image. Hence, although the raw image has better resolution than the photogrammetry-produced DOM and point cloud, it does not have the pixel-level geographical coordinates. This makes it difficult to accurately match each plant from raw images (with the same pixel coordinate for each image) to corresponding locations (with geographic coordinates) in the field.

1.3.4 Low efficiency from close-range

When accurate organ-scale models cannot be obtained outdoors, a common method is to obtain 3D models through destructive sampling and close-range indoor 3D reconstruction. However, this approach has some problems: 1) destructive sampling cannot track the complete growth state of a plant, and can only obtain growth slices of a particular plant. However, when there is sufficient destructive sampling, it is possible to approximate the growth trend of the entire field; 2) currently, low-cost indoor reconstruction equipment still requires a lot of manual operation, such as manually replacing plants, moving cameras, and pressing the shutter to obtain images from different perspectives. After obtaining the images, a large amount of manual user interface (UI) operation is still required in the 3D reconstruction software to process individual plants, making it difficult to handle large numbers of plants; and 3) the most popular “fixing plant and moving camera approach” is limited by the viewing angle and the success rate of image stitching, resulting in incomplete modeling due to certain angles of the plant not being captured. Therefore, the current close-range plant 3D modeling method is still in a “low-speed and high-labor” stage, and it is difficult to obtain complete plant 3D models.

1.3.5 Workloads in data annotation

Efficient acquisition of large amounts of high-quality image training data and annotations has become an urgent need for obtaining robust deep learning models in agriculture (Yang and Xu, 2021). Many studies have attempted to solve these difficult-

ties and their solutions can be summarized as follows: 1) decreasing the workload of training data annotation using automated annotation devices (Beck et al., 2020b). 2) publicly sharing datasets. For the first point, reproducing such a system is challenging for researchers without a background in machinery, and it is also hard for outdoor applications. For the second, Bender (2019) shared a weekly scanned image dataset of cauliflower and broccoli, but without annotations. Beck et al. (2020a) shared the labeled weed seedling dataset collected by their automation system (Beck et al., 2020b). David et al. (2021) shared an extensive labeled wheat head dataset collected from 12 countries. However, current research is often distributed on various platforms with various label formats, making it cumbersome to use. The agricultural science community urgently needs a standardized data-sharing platform; 3) use transfer learning to reduce the amount of required training data and time for model fitting (Yang and Xu, 2021). Desai et al. (2019) used the ResNet-50 model pre-trained on the ImageNet dataset, while Blok et al. (2021a) used the Mask-RCNN model pre-trained on the COCO dataset. However, current models for transfer learning are trained on computer vision datasets, and Blok et al. (2021a) reported that broccoli in the CV datasets are dishes on plates, rather than broccoli grown in the field; 4) use data augmentation to generate a large number of annotations from small ones. Zhou et al. (2020a) used geometric transformations (random cropping and rotation), and Blok et al. (2021a) applied more geometric transformations and added photometric transformations to them. A more advanced data augmentation strategy involves generating fake images (Nesteruk et al., 2021). All these studies have shown that data augmentation can improve model performance; and 5) Use model-assisted labeling to significantly reduce the workload during data labeling. While some platforms provide model-assisted labeling, such as LabelBox (online, <https://docs.labelbox.com/docs/model-assisted-labeling>), v7labs (online, <https://www.v7labs.com>), and EISeg (offline, <https://github.com/PaddlePaddle/PaddleSeg>), they simply reduce the workload of training image annotation by using pre-trained models, rather than iteratively and interactively improving model performance for specific studies.

1.3.6 Research reproducibility

Reproducibility remains a problem in some plant phenotyping studies and tools. Many studies do not share their datasets or code publicly, making reproducing their experimental results difficult. This leads to a waste of resources in solving similar problems by reinventing the wheel. For studies that share their codes or tools publicly, the application effect on other cultivars is often suboptimal. As Lobet (2017) stated in the title “*Image Analysis in Plant Sciences: Publish Then Perish*”, many published phenotyping tools still suffer from severe reproducibility issues and a lack of ease of use. The main issues are: 1) code and example data are published on personal or laboratory websites and are often inaccessible due to slow download speeds caused by bandwidth costs and uniform resource locator (URL) changes and the inability to maintain them; 2) only code files and straightforward introductory files are provided, with detailed instructions for setting up the code environment, runnable examples, and detailed application programming interface (API) documentation often missing. This results in an inability to run and apply the code to their own data; and 3) a lack of follow-up maintenance and community support. Bugs may be solved through troubled email communications, but user experiences cannot be shared with other users.

1.4 Objectives of this study

The study targeted the improvement of phenotyping of field-grown broccoli (*Brassica oleracea* L.) as an example. Broccoli is an important crop for the global vegetable market that faces the aforementioned challenges. It grows in rows in the open field and has an uneven growth rate like other cereal and horticultural crops. Aerial survey is one of the most time-efficient approaches to obtain entire field information, but the broccoli head is often less than 20 cm, with different occlusions and positions, making the organ-level detection and segmentation of ROI challenging for current 3D plant phenotyping techniques. Additionally, due to the “mushroom” shape of its head crown, it is still difficult for the labor-intensive close-range approach to obtain a complete 3D model with

the bottom of the crown. Furthermore, except for providing accurate statistical value for broccoli head size estimation, 3D virtual visualization may be more straightforward for farmers to make decisions. This Ph.D. thesis aimed to improve the performance of 3D-based plant phenotyping on broccoli farmland using only low-cost RGB cameras.

To be more specific, the overall objectives were:

1. To develop a semi-automatic close-range 3D reconstruction pipeline that can obtain high-quality 3D models of broccoli heads as a template database.
2. To develop an unsupervised phenotyping data processing workflow that can automatically calculate several 1D to 3D morphological traits for close-range broccoli heads.
3. To develop an improved workflow for aerial 3D reconstruction of broccoli canopy that can obtain better 2D morphological traits of broccoli heads in complex outdoor conditions.
4. To automatically fuse the data between the close-range 3D database and the aerial canopy, which can calibrate the aerial-measured 2D morphological traits.
5. To provide 3D virtual visualization for the head in the actual field.
6. To publicly share all the codes used in this study (including the \LaTeX code of this thesis) on Github, the most famous community for programming communication and bug reports, making this contribution free to use for anyone who has similar needs.

1.5 Outline of this study

Chapter 1 provides an overview of the background information, related studies, and objectives of the study.

Chapter 2 presents the close-range 3D phenotyping pipeline, which obtains high-quality 3D models of destructively sampled broccoli heads. Firstly, an almost-automatic

workflow that captures and saves the image of the target crop in multiple view angles is developed. Then, ROI on the broccoli heads in the images are extracted by two pre-trained deep learning models. The preprocessed images are then fed into photogrammetry-based software (Agisoft Metashape) to generate 3D models using automatic processing scripts. Finally, the broccoli crown part is automatically segmented, the 3D model coordinates are corrected, and the phenotypic traits are calculated automatically. To evaluate the performance of the proposed pipeline, we compared some of the pipeline-measured traits with manually measured traits using the coefficient of determination and root mean square error (RMSE).

Chapter 3 presents the development of an aerial 3D phenotyping pipeline for UAV sensing on broccoli canopies. To overcome the challenges in distinguishing small broccoli heads on low-quality 3D canopy models from photogrammetry, we proposed a novel data-fusion method. We first fused the time-series data. By using the broccoli seedling positions extracted from the early seedling stage, which is a very simple detection task, the processing regions of later broccoli head segmentation tasks were narrowed down dramatically. This method combined with active learning, significantly decreased the workload of data annotation for deep learning training and the data amount for deep learning processing. Then, we fused the data between pixel coordinates (raw images) and geographical coordinates (photogrammetry outputs). This method allows for the segmentation of the broccoli head region on the original drone images (pixel coordinates without actual scale) and projective transformation of the results back to geographical coordinates to calculate actual size. Manual measurements in the field were carried out to validate the accuracy of this pipeline.

Chapter 4 tests the idea of data fusion on the pipeline built in Chapters 2 and 3. The piecewise affine transformation is used to locate the head segmentation regions more precisely in the geographical coordinates than in Chapter 3. Then, automatic machine learning is used to calibrate the 2D morphological traits for each broccoli head from the aerial pipeline (Chapter 3) using the destructively sampled broccoli heads. Later, a modified normalized cross-correlation template matching method is used to find the

closest broccoli head in the template database (Chapter 2) and transform it back to the aerial 3D canopy model for virtual visualization.

Chapter 5 discusses study limitations and potential future research prospects and Chapter 6 summarizes the study's general conclusions.

Chapter 2

Close-range 3D-based phenotyping pipeline

2.1 Introduction

Broccoli (*Brassica oleracea* L.) is an important crop for the global vegetable market. However, a certain amount of non-standardly harvested heads are wasted owing to uneven sizes and the failure to meet cosmetic standards such as shape and aesthetics. Traditional manual judgments for the size and quality of broccoli heads are time-consuming, laborious, and often inaccurate. Hence, there is a need to develop novel methods to quantify specific traits and accelerate the evaluation process.

Several authors have developed 2D-based phenotyping methods using images for the above purpose, which are efficient, non-destructive, and have high-throughput (Guo et al., 2017; Yang et al., 2015; Zou et al., 2019). However, these approaches are unable to describe the plant 3D structure due to occlusion and dimension loss when projecting onto the 2D plane. As a result, it produces inaccuracies and uncertainties for advanced phenotyping applications.

To overcome the drawbacks of 2D-based phenotyping, several studies have paid attention to 3D approaches. Paulus (2019) and Kochi et al. (2021) have summarized the current approaches to obtain 3D plant models; and a large number of studies have chosen 3D reconstruction by photogrammetry using common red-green-blue (RGB) cameras due to the low device cost (Xiao et al., 2021; Zermas et al., 2020; Zhang et al., 2016). The 3D model of an object can be calculated from images from different perspectives

(view angles) with enough overlap (Fig. 2.1). The full process includes structure-from-motion (SfM) to calculate relative positions between images and the rough 3D point cloud of the object; multi-view stereo (MVS) to densify the 3D point cloud of the object; and surface reconstruction to obtain object 3D mesh model. For details, please refer to Hartley and Zisserman (2003) and Snavely et al. (2010). Although the proposed 3D-based phenotyping was proved feasible for many phenotyping applications, when apply to the broccoli head measurements, the current 3D method is facing several challenges.

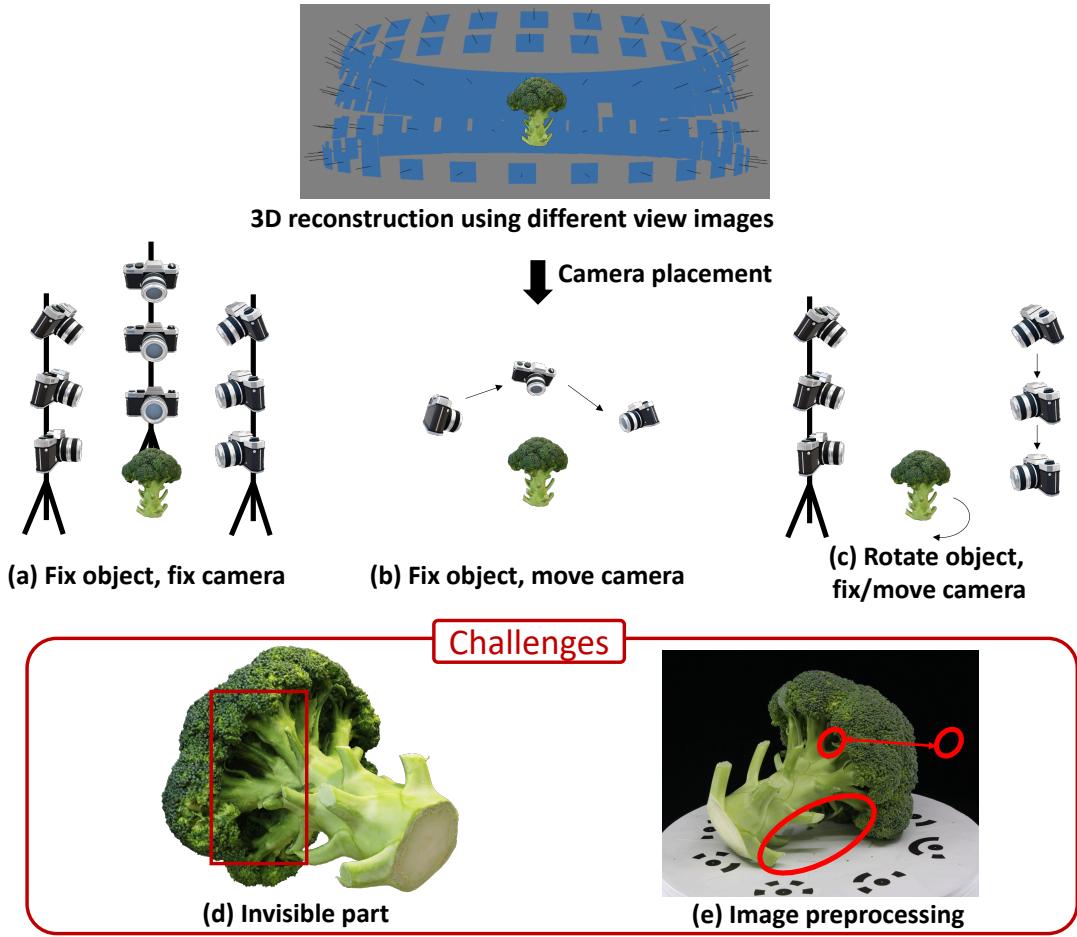


Figure 2.1: The current photogrammetry (3D reconstruction) methods and challenges; (a-c) the current image acquisition approaches (a) fixing the object and taking images using multiple fixed cameras at the same time, also called forward intersection; (b) fixing the object but taking images by using a moved camera, also called backward resection; and (c) rotating the object and taking images using fewer multiple fixed cameras, or a camera fixed at different locations for each rotation. The challenges of current approaches: (d) the limited view angles of current image occlusion approaches has visual dead area, which will cause incomplete plant 3D models; and (e) the difficulties to segment foreground (plant) area in the image preprocessing.

The first challenge is image acquisition for cost-effective reconstruction. Several authors used the following three approaches to acquire images (Fig. 2.1): (a) fix the object and fix cameras (Nguyen et al., 2015); (b) fix the object and move a camera manually (Xiao et al., 2020) or using robotic arms (Cao et al., 2019; Nguyen et al., 2016); and (c) rotate the object and fix or move the camera(s) (Gao et al., 2021; Kochi et al., 2018). The main drawback of these imaging approaches is that it is difficult to provide fully complete perspectives of plants within acceptable cost. Most methods can only capture images of the sunlit side (such as the adaxial surface of leaves and the top of broccoli crowns); while capturing images of the shaded side (such as the abaxial surface of leaves and the downside of broccoli crowns) is more difficult and requires extreme upward viewing angles (Fig. 2.1d). Even if such images can be acquired, they often fail to align with the 3D model due to insufficient feature matching points. Hence, a solution is required to improve the completion of the reconstructed plant 3D model, especially for solid, closed harvestable organs like broccoli heads.

The second challenge is image preprocessing for the acquired images. Compared to the approach of building the full scene, which includes both plants and backgrounds, following the segmenting of the plant parts (Ge et al., 2019) and denoising (Wu et al., 2020); some studies segmented and masked the foreground (plant) area before doing the 3D reconstruction (Kochi et al., 2018; Nguyen et al., 2016), to improve the model quality and eliminate the effect of background noise. Although the controlled environment can provide pure background and stable lighting, it is still challenging to develop a robust algorithm to segment the broccoli area perfectly, especially to handle the shadows caused by the irregular broccoli head shape at the different growing stages. Some outdoor studies showed the power of deep learning for plant area detection and segmentation (Blok et al., 2021a; García-Manso et al., 2021; Zhou et al., 2020a), but due to limitation of the graphics processing unit (GPU) memory, the original images are often resized to smaller sizes for deep learning. For example, Zhou et al. (2020a) resized to 1440×1080 , Blok et al. (2021a) resized to 1024×1024 , and García-Manso et al. (2021) resized to 640×480 . The plant mask produced in such a resolution cannot fit well into the original

images in detail (in our case, is 5184×3456), and hence also need a method to produce detailed masks with the original resolution.

In this chapter, several strategies were applied to overcome the aforementioned challenges and provide a labor-saving pipeline for obtaining high-quality 3D models for broccoli heads. The objectives of this chapter were to (1) develop the dual-rotation object strategy with a fixed camera to collect images with abundant perspectives for complete 3D reconstruction; (2) implement the two-step deep learning workflow to acquire the detailed plant masks on the original images; (3) calculate the 3D morphological traits of broccoli head and crown; and (4) validate the estimated head size using the manual measurements.

2.2 Methods and Materials

The pipeline proposed in this chapter has two main parts, the workflow to acquire high-quality broccoli head 3D models and the workflow to measure the morphological traits of broccoli heads. For the first workflow, we developed an automatic imaging device using commercial photographic equipment, which extended the design of Kochi et al. (2018) to dual-rotation of objects; then applied marker detection (software built-in function) and deep learning segmentation of plant masks to the collected images, as image preprocessing; afterward, we extended the batch scripts that Feldman et al. (2021) developed to automatically operate the 3D reconstruction on a large number of broccoli heads to get 3D model results. For the second workflow, similar to the work by Feldman et al. (2021), we first developed the unsupervised algorithm to split broccoli crown and stem parts; then corrected the broccoli head top direction of the 3D models (as the z-axis positive direction); lastly, we calculated several 2D and 3D morphological traits of the broccoli head and segmented crown.

2.2.1 Plant materials

Field trials were conducted at the experimental farm of the Institute for Sustainable Agro-ecosystem Services (ISAS), Nishi-tokyo, Tokyo, Japan ($35^{\circ}43'N$, $139^{\circ}32'E$) from

October 2021 to April 2022. The field size was approximately 3000 m^2 ; The meteorological data during the growth period were collected by a local weather station and are shown in Table 2.1, which comes from Nishida (2023). The soil was tilled and leveled using Half-Soiler and a Rotavator. Initial fertilizer was mechanically spread and mixed thoroughly on October 11, 2021. From October 14 to 16 approximately 11,000 broccoli plants (cultivar: Suzuka) were transplanted using a broccoli transplanter with an in-row spacing of 35 cm and between-row spacing of 70 cm . Additional fertilizer was applied by hand on October 28. Herbicides, insecticides, and fungicides were applied as needed. Detailed crop management practices were described in the graduation thesis of Nishida (2023).

Table 2.1: Meteorological data during the experiment period (Nishida, 2023)

	Mean temperature ($^{\circ}\text{C}$)	Total precipitation (mm)	Mean daily solar radiation ($\text{MJ} \cdot \text{m}^{-2}$)
2021.10	17.6	156.5	11.2
2021.11	12.2	85.0	11.0
2022.12	6.4	123.0	9.6
2022.01	3.6	16.5	10.6
2022.02	3.8	60.0	13.3
2022.03	10.3	94.5	15.3
2022.04	15.0	213.0	15.9

Three fertilizing treatments (inorganic, organic, and mixed) with three replicates (total 9 plots) were applied to produce different sizes and shapes of broccoli heads (Nishida, 2023). For each plot, several neighbored broccoli heads were destructively sampled during the flowering seasons (152 days after transplanting, Table 2.2) for indoor 3D reconstruction, the fieldwork was contributed by ISAS staff and Nishida.

2.2.2 Plant 3D model acquisition

The general workflow to acquire a broccoli head 3D model is shown in Figure 2.2, including imaging devices, image pre-processing, and batch reconstruction.

The imaging device is assembled by a marker board, an automatic horizontal rotating

Table 2.2: Destructively sampled broccoli head number during the flowering season. A total of 189 broccoli heads were destructively sampled. The “rep” means one replicate with 3 plots, while “all” means all 9 plots

Date	Plot id	head count	Total
Mar 17	3; 6; 9 (rep 3)	5 per plot	15
Mar 21	1; 4; 7 (rep 1)	5 per plot	15
Mar 24	2; 5; 8 (rep 2)	8 per plot	24
Mar 29	1-9 (all)	6 per plot	54
Mar 31	1-9 (all)	6 per plot	54
Apr 5	1-9 (all)	3 per plot	27

platform (Foidio360, <https://orangemonkie.com/ja/products/foldio360-turntable>), a small photography studio (with lighting and pure color background), a Canon EOS 60D digital single-lens reflex (DSLR) camera, and a tripod (Fig. 2.2a). The rotating platform can rotate at a given angle (15° in this study) and then stop for a short interval for image

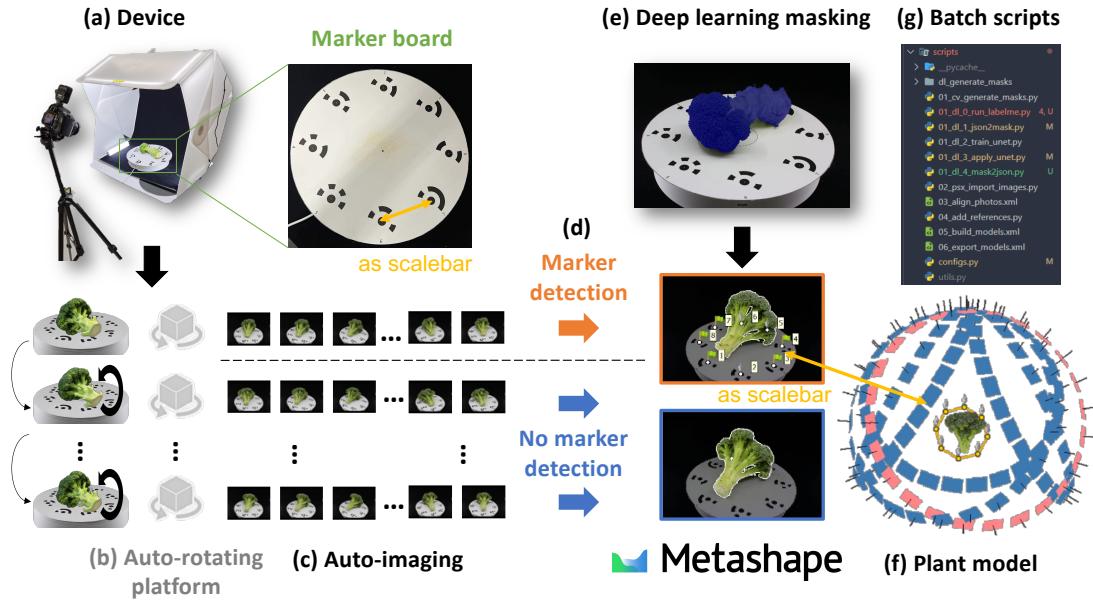


Figure 2.2: The Plant 3D model reconstruction workflow; (a) devices for taking plant images, with a marker plane used as scalebar references; (b) an automatic horizontal rotating platform, which can rotate at a given angle and then stop for a short interval for image taking. The vertical rotation of plants (flip) requires manual operation; (c) different photographic perspectives images taken by the infrared signal auto-imaging system; (d) partial marker detection to avoid misleading image alignment, where only markers on one rotating image group (orange) are detected and not on the others (blue); (e) plant part segmentation using deep learning; (f) the final 3D plant model produced by Metashape; and (g) the scripts for batch processing a large number of plants.

taking. It was programmed to emit infrared signals to command the DSLR camera to take photos and then images were directly transferred to a computer through a data cable using Canon EOS Utility 2 software (<https://cweb.canon.jp/drv-upd/dc/euw21401.html>, Fig. 2.2c). Manually rotating the broccoli head vertically (black rotation arrow in Fig. 2.2b) to capture three to five photograph perspectives was required for complete 3D reconstruction. The commercial software, Agisoft Metashape (Agisoft LLC, St. Petersburg, Russia, <https://www.agisoft.com>), was used for 3D reconstruction.

The 3D reconstruction and the later image and data processing were conducted on a computer with an AMD Ryzen 9 5950x with 16-Core processors, and DDR4 128GB random access memory (RAM). The computer had one Nvidia GeForce RTX 3090 GPU. The operating system is Windows 11 Professional. The codes are implemented in Python 3.8, compute unified device architecture (CUDA) 11.1, Pytorch 1.8.2, and torchversion 0.9.2.

The first step of image preprocessing was partial marker detection (Fig. 2.2d). The marker board not only provides reference points for camera alignment but also provides the size calibration as a scalebar. The Agisoft Metashape 3D reconstruction software provides the built-in function to detect markers on the given images. However, only markers on the first rotating image group (orange color) were detected and not on the others (blue color). It was done since we vertically rotated the plants, thus changing the relationship between the plant and markers and potentially misleading the expected image alignment from other rotation groups. In other words, we aimed to “cheat” the software by inverting the relationship between static and rotating object. Thus, the software treated the broccoli as static and the camera as rotating (Fig. 2.2f). The marker plane from the first image sequence was used for size calibration. Meanwhile, that marker plane was used as the X-Y plane for defining the 3D model coordinate.

Two pre-trained deep-learning networks were used for the second image processing step of plant area segmentation (masking). As mentioned in the introduction, deep learning can handle complicated image analyzing tasks, but often not with the original image resolution. Here we first use a pre-trained UNet with default model setting ([https:](https://)

[//github.com/qubvel/segmentation_models.pytorch](https://github.com/qubvel/segmentation_models.pytorch)) and do the transfer training with only 19 broccoli images. The training data was annotated using the open-source software, Labelme (<https://github.com/wkentaro/labelme>). The UNet model produces rough masks (256×256 pixel resolution) with good performance and acceptable processing speed on almost all broccoli head images during the flowering season. Afterward, we directly used the pre-trained CascadePSP (Cheng et al., 2020, <https://github.com/hkchengrex/CascadePSP>) to refine the rough masks to the original image resolution without any image annotation (Fig. 2.2e). Afterward, to further address the potential flaws of the previous segmentation results, we assumed the plant area should be one continuous region. Hence, only the largest region as the plant area was kept and the small holes inside the plant area was filled, as the final result.

Finally, the batch scripts were developed for the previous image preprocessing and the 3D reconstruction control in the Metashape (Fig. 2.2g). All the scripts can be found at https://github.com/UTokyo-FieldPhenomics-Lab/Foldio360_3D_Reconstruct_Platform.

2.2.3 Morphological traits extraction

The workflow of Morphological traits extraction is shown in Figure 2.3, which includes broccoli crown segmentation, direction correction, and trait calculation.

The broccoli crown segmentation step aimed to split the crown and the stem. Since they have significant color differences, the color information can be used for splitting. A method employed a simple color threshold like Otsu (1979) threshold did not work in a pilot study, because the color also varies among individual heads during the flowering season. Thus, a more robust method was developed. To begin with, broccoli head 3D mesh models were transferred to a point cloud with RGB colors (Fig. 2.3a1). An unsupervised Kmeans clustering algorithm split all point clouds into two classes (Fig. 2.3a2). Later, the mean RGB color values of these two classes were calculated and transformed into hue-saturation-value (HSV) color space. The V represents the “darkness” of color, and the class with lower V values (darker color) was picked as the crown part. And

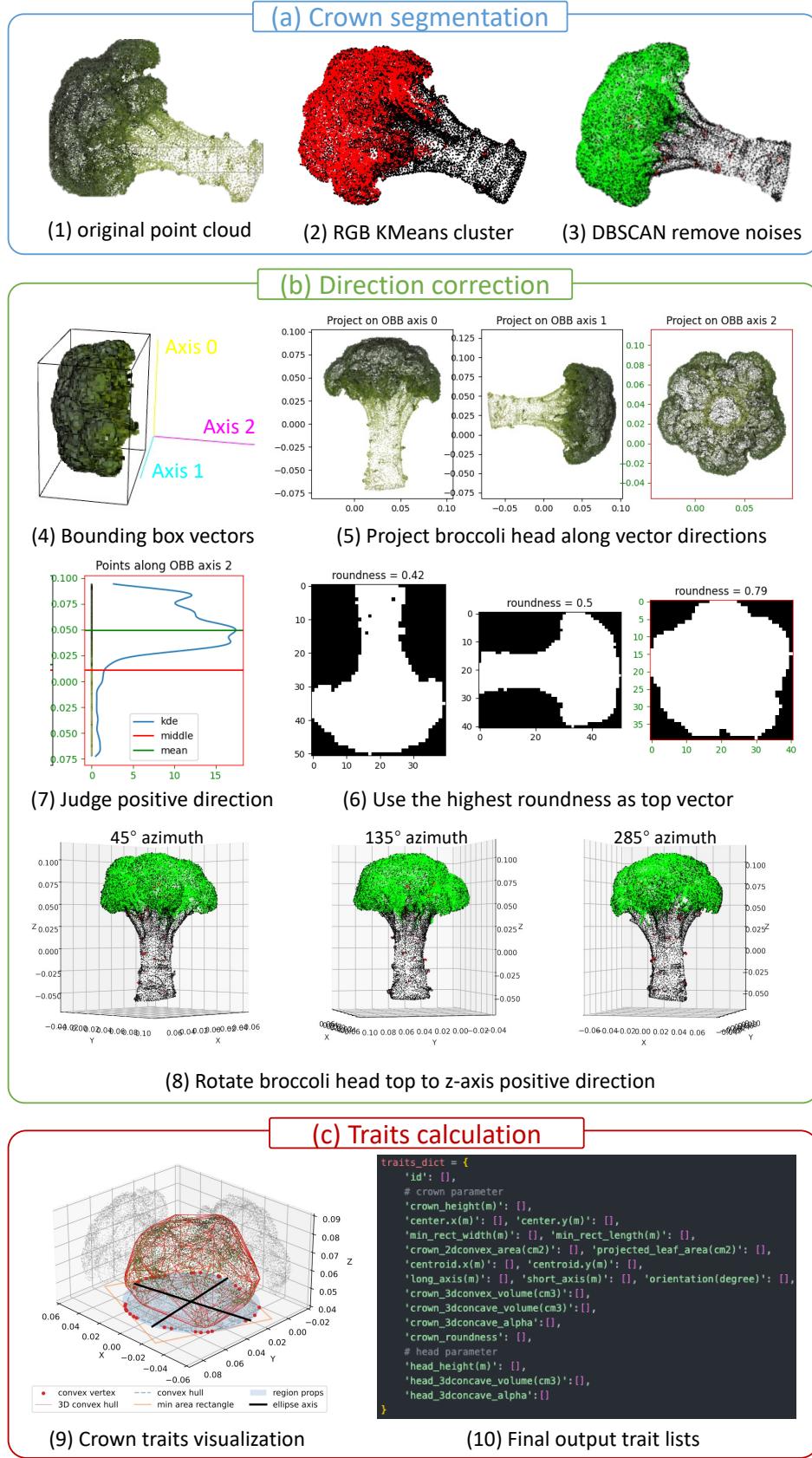


Figure 2.3: The algorithms for plant 3D model analysis; (a) crown segmentation, to split the broccoli crown part and the stem part; (b) direction correction, to rotate the broccoli vertically to the z-axis positive direction; and (c) traits calculation.

finally, the density-based spatial clustering of applications with noise (DBSCAN) algorithm was used to cluster the crown points into several groups according to neighbor point density. These groups are clustered by 2-class Kmeans again according to their size, and the clusters with smaller means sizes were discarded as noise (red dots in Fig. 2.3a3).

Due to the broccoli heads always laying on the marker board (the X-Y plane in the coordinates), the broccoli head top did not correspond to the z-axis positive direction for the reconstructed 3D model. Broccoli head top direction was corrected to better find the projecting plane for the crown projected area and crown thickness calculations. To begin with, the minimum volume bounding box, also called oriented bounding box (OBB), and the vectors of its three perpendicular edges were calculated (Fig. 2.3b4). Since the broccoli head has variable shapes during the flowering season, the shortest perpendicular edge was not always the top direction. Hence, the full broccoli head was projected into the corresponding planes perpendicular to OBB vectors (Fig. 2.3b5); the roundness of the projected shape was also calculated. The vector with the highest roundness was used as the top direction vector (red boundary, Fig. 2.3b6). Later on, to find the positive direction (Axis 2 in Fig. 2.3b4 was a reversed direction), full broccoli along the vector (left greenish vertical line in Fig. 2.3b7) was projected; the middle value and the mean value of these points are also calculated. The positive direction should be the middle value greater than the mean value. Figure 2.3b8 shows the corrected direction along the previous vector, with azimuth angle views at 45° , 165° , and 285° .

In the end, we calculated the traits for the broccoli crown (Fig. 2.3c9) and the whole broccoli according to Feldman et al. (2021). For the 1D traits, we calculate the crown height (thickness) and whole broccoli height; For the 2D traits, we project the broccoli crown on the X-Y plane and calculate its center, centroid, roundness, minimum area rectangle width and length, projected area, convex hull area, and the short and long axis and azimuth angle of the fitted ellipse. For the 3D traits, we calculated the 3D convex hull volume for both the crown and whole broccoli and the 3D concave hull volume for the crown only (Fig. 2.3c10). All the codes can be found at

<https://github.com/UTokyo-FieldPhenomics-Lab/BroccoliHead3D.ipynb>

2.2.4 Validation

To further validate the effectiveness of our workflow, we compared the size of broccoli heads measured by manual [Nishida (2023)’s contribution] and pipeline (my contribution). Given the challenges associated with measuring 2D and 3D traits (e.g. area and volume) in reality, we focused solely on the 1D length (size) of the broccoli heads for this comparison. In agriculture practices, the head lengths are manually measured in the 0° , 45° , 90° , and 135° directions (the “union jack” shape) for validation. And then we picked up the longest and shortest ones as the final size. All the manual measurements were contributed by Nishida (2023). The workflow calculated trait values, such as the width and length of the minimum area rectangle, and the long and short axis of the fitted ellipse. The agreement of workflow measured and manually measured was estimated using the coefficient of determination (R^2 , calculated as the squared Pearson’s correlation coefficient) and the root mean square error (RMSE):

$$RMSE = \sqrt{\frac{1}{n} \cdot \sum_i^n (W_i - M_i)^2} \quad (2.1)$$

where, n is the total destructive sampled head number, W_i is the Workflow calculated length of broccoli head i , M_i is the Manually measured length of broccoli head i .

2.3 Results

Following the workflow outlined above, a total of 189 destructively sampled broccoli heads were successfully 3D reconstructed during the flowering season with high model quality and completion. The intermediate results also show the feasibility of the proposed algorithms for variate broccoli head shapes. The calculated trait values also had high correlations with those by manual measurements.

2.3.1 Plant 3D model acquisition

Figure 2.4 showed some intermediate results for the plant area segmentation with variable broccoli head sizes and shapes. Using pre-trained UNet and transfer training with just 19 broccoli head images, rough masks (red parts in Fig. 2.4b) were successfully generated for different broccoli shapes. UNet image mask quality, especially on object boundaries, was limited by resolution and a low amount of training data. However, masks were substantially improved by CasadePSP and the hole-filling postprocessing (blue parts in Fig. 2.4c).

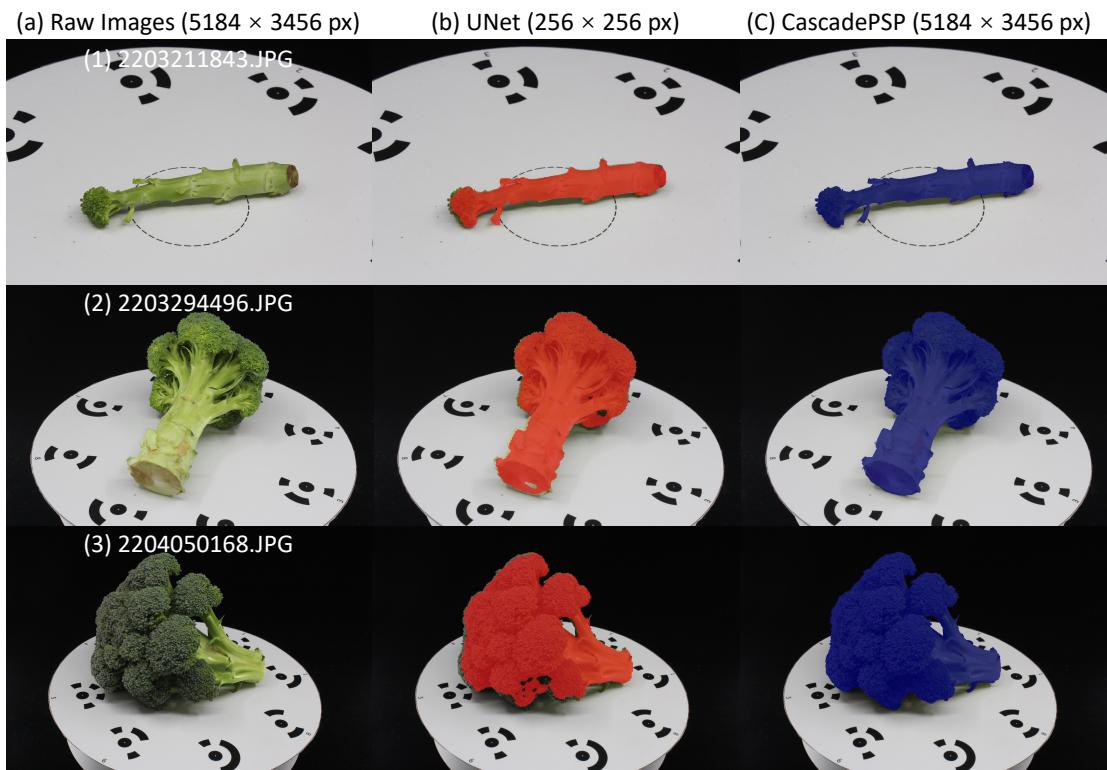


Figure 2.4: Examples of two deep learning segmentation results; (a) is the original image of the broccoli head, the resolution is 5184×3456 pixel; (b) is the output of the UNet segmentation, the resolution is 256×256 pixel, which is not enough as final plant masks; and (c) the output of the refined high-resolution mask using CasadePSP based on the UNet output. The mask has the same resolution as the original image.

Some examples of reconstructed 3D head models were shown in Figure 2.5. Here, we generate a computer graphics (CG) image using the 3D head models (Fig. 2.5b) and compare it visually to the real-world photo (Fig. 2.5a). Due to limitations in shooting angles and manual editing, the two images are not the same in detail, but the

overall details and sizes of the broccoli are quite similar. Figure 2.5c showed different perspectives of the same broccoli head. Unlike other studies, in which the underside of the model (Kochi et al., 2018) or leaf backside (Cao et al., 2019) are not available, the proposed pipeline does include the entire broccoli head, with bottoms of the crown and stem available. In general, our proposed dual-rotation workflow produces completely enclosed high-quality 3D models, it will not only contribute to better analysis accuracy but also build a base for future digital farm applications.

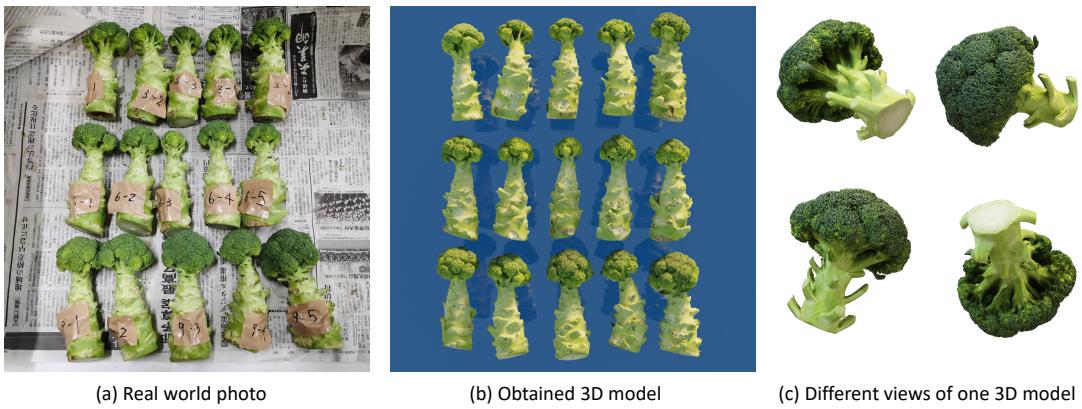


Figure 2.5: Some examples of obtained 3D models; (a) the image of 15 broccoli heads in the real world; (b) the obtained 3D models of corresponding broccoli heads; and (c) the different views of one broccoli head 3D model.

2.3.2 Morphological traits extraction

As mentioned above (see Subsection 2.2.3), the crown and stem parts of the broccoli head should be segmented and rotated upwards for easier trait calculation. Figure 2.6 shows the feasibility of our proposed unsupervised algorithm on broccoli heads with different shapes. Figure 2.6a is a simple case of a small broccoli head with a straight stem; Figure 2.6b is a challenging case with a curved stem; Figure 2.6c and d are challenging cases for head and stem segmentation. Instead of concluding that our algorithm rotates them to the exact status in the field correctly, it rotates them to a proper horizontal angle with the largest projection area on the X-Y plane and the shortest thickness on the Z-axis. This provides an initialized starting point for later trait calculation, especially very important for the 1D (on the Z-axis) traits and the 2D traits (on the X-Y plane).

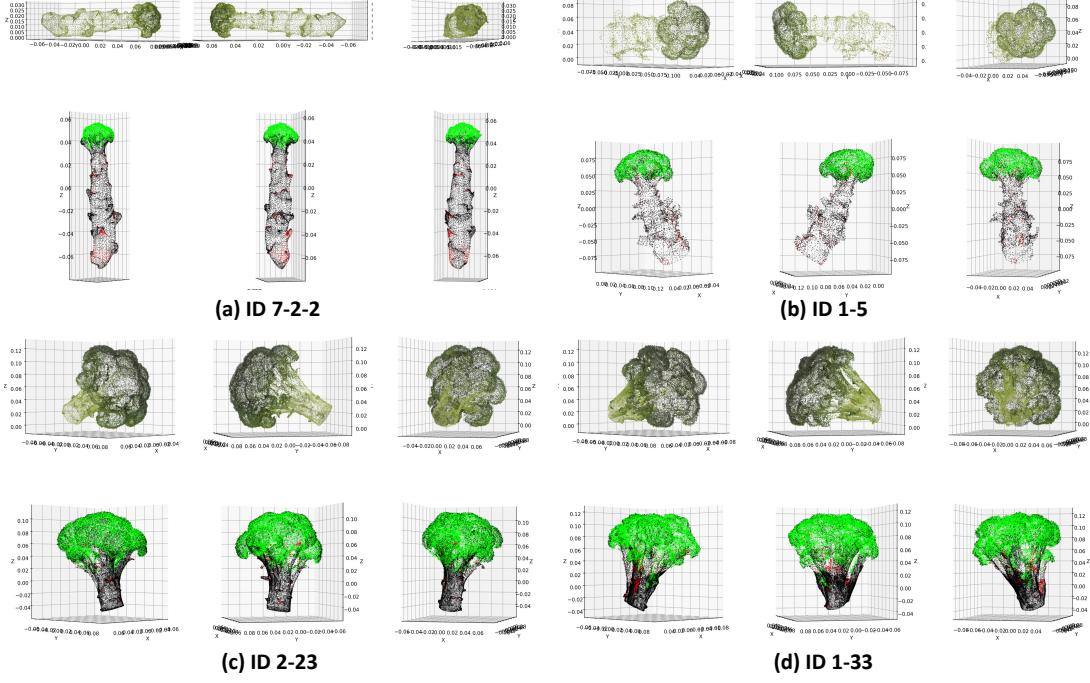


Figure 2.6: Examples of plant 3D model analysis at different growing stages. The upper parts are the original coordinates of the obtained 3D models, while the lower parts are the segmented and direction-corrected results, the red parts are removed noise. The three columns show the corresponding azimuth angle views at 45° , 165° , and 285° for the same 3D model.

To validate the accuracy of the proposed workflow calculated traits, they were compared with manual measurement using standard agricultural practices. Figure 2.7 shows the head size comparison results. The broccoli head size can be represented by the minimum area bounding rectangle and fitted ellipse axis lengths. Both the rectangle and the ellipse methods showed high correlations with the manually measured results ($R^2 \geq 0.97$), but the rectangle has a higher correlation than the ellipse. This is because the rectangle width and length are more like the minimum and maximum of the manually measured “union jack” lengths, respectively. Assuming the complicated broccoli head is an ellipse inherently involves a certain degree of error, but the RMSEs of the ellipse assumption (0.553 cm and 0.448 cm) were still better than that of the circular assumption (Blok et al., 2021b, Table 5, RMSE=0.97 cm)

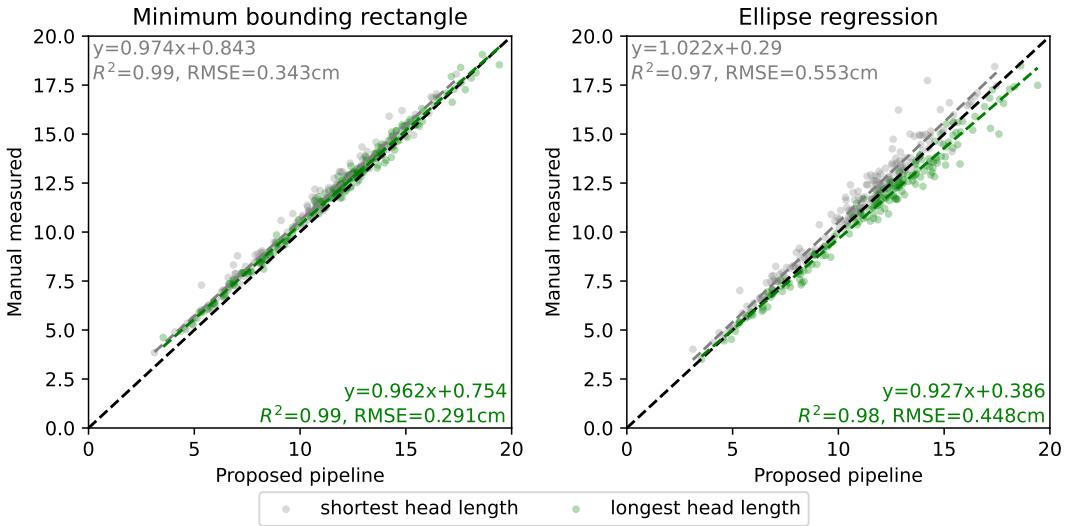


Figure 2.7: The comparison between the proposed close-range pipeline (my contribution) and manual measurement [Nishida (2023)'s contribution]. The shortest length and longest length of the broccoli head are compared. For the proposed pipeline, it uses two methods to estimate those lengths. One is using the length and width of the minimum bounding rectangle, the other is using the major and minor axes of the fitted ellipse.

2.4 Discussion

In this chapter, we implemented a 3D plant phenotyping pipeline to obtain high-quality 3D broccoli head models and calculate the 3D morphological traits of head shape and size. Our 3D model reconstruction method proved superior to image acquisition methods used by many 3D-based phenotyping studies (Fig. 2.1a-c), because they often failed to capture complete plant structure due to limited perspective. Here, we extended the object rotation method (Fig. 2.1c) to the dual-rotation (Fig. 2.2d) using the commercial auto-rotation platform, which ensures enough perspectives without excessively increasing workload. Meanwhile, generating masks for the plant area on the images to exclude the impact of background noise has proved beneficial for 3D model quality. Rather than developing customized computer vision algorithms to extract plant area (Kochi et al., 2022, 2018; Nguyen et al., 2016), which is not robust for different environments and plant objects, we used the pre-trained deep learning networks to decrease the workload of algorithm development and data annotation. The results showed good feasibility of this pipeline on different shapes of broccoli heads during

the flowering seasons. The open-source unsupervised algorithm to segment broccoli crown, correct the head direction and calculate several agronomical traits is available at <https://github.com/UTokyo-FieldPhenomics-Lab/BroccoliHead3D.ipynb>.

Some limitations remained to be unsolved. First, the dual-rotation method is only suitable for solid objects, such as broccoli or cauliflower head or fruits. It cannot be applied to soft objects, such as flowers or plant leaves, whose shape or structure can change after vertical flipping on the marker plane. The object shape changing from different perspectives conflicts with the assumption of 3D reconstruction and results in poor model quality or even alignment failure. Second, the dual-rotation is also not fully automated, requiring manually flipping the object after each sequence of horizontally rotated images. Although flipping and replacing the broccoli only takes a few seconds each time, a person needs to be constantly waiting for each automation cycle to complete. The waiting time accounts for at least 80% of the total imaging time. And lastly, limited by the developing time, we only published the open-source code with simple instructions for this pipeline, there is no easy-to-use software with a user-friendly graphical user interface (GUI) for researchers without any programming background.

For future work, it is suggested to apply and examine the pipeline on more agricultural applications for crop organs, such as sweet potato quality judgment and potato tuber bud detection for agronomic, breeding and post-harvest quality objectives. The proposed pipeline also provides an ultra-high-quality broccoli head template database, which can integrate with the unmanned aerial vehicle (UAV)-based outdoor survey for the entire field to substantially improve model quality (In Chapter 3). Also, with the fast development of deep learning, more advanced models, like “The Segment Anything” (Kirillov et al., 2023), could be implemented to further reduce the workload in the image preprocessing process.

Chapter 3

Aerial 3D-based phenotyping pipeline

3.1 Introduction

Smart farming, which involves new technologies such as remote sensing, high-throughput phenotyping, and artificial intelligence in agricultural production, has received considerable attention from researchers, farmers, and governments. The unmanned aerial vehicle (UAV)-based aerial pipeline provides a cost-efficient method to capture images for crop canopies. The 3D canopy model can also be reconstructed using photogrammetry-based software. Several studies extended this approach to estimate canopy architectural traits (Herrero-Huerta et al., 2020a; Shu et al., 2021; Wang et al., 2021c) and even traits for individual lettuce plants (Bauer et al., 2019). But such trait phenotyping at the canopy level or individual level cannot meet the accuracy demand for harvestable organ size judgment; for example, broccoli head grades vary by a few centimeters. Therefore, extracting the organ-level traits can significantly improve the accuracy of whole-field estimation.

For many organ-level applications, the images are often collected close to the ground (less than 1 *m* between a sensor and a plant) by a hand-held camera or an unmanned ground vehicle (UGV). Although several studies successfully proved the accuracy and feasibility of this approach on small-area experimental fields (Blok et al., 2021a; García-Manso et al., 2021; Lüling et al., 2021), its efficiency is not always applicable to the large-area field with thousands of individuals. To apply the organ-level analysis on the

UAV-based approach, there are two main challenges to be solved.

One challenge is the poor quality of the reconstructed canopy model (2D field map and 3D point cloud). Due to the plant structure movement in different UAV images caused by wind, the canopy model often has the effects of double mapping (ghost effect) and seamline distortion (Lin et al., 2021) (Figure 3.1b). Many studies tried to fix the low quality using machine learning algorithms (Hu et al., 2021, 2019; Velumani et al., 2021) or multi-spectral sensors (Guo et al., 2021a; Lu et al., 2022), these can be time-consuming, costly, and are often not robust on different crops. Little attention has been paid to the original UAV images which often have better quality. Without pixel-level Global Positioning System (GPS) coordinates on original images, it is not possible to accurately match plants from images to corresponding locations in the field. The solution is to reuse the intermediate parameters during the 3D reconstruction. In the 3D reconstruction procedure, the relative rotation angle and position between the UAV images and the object (field) have been calculated and calibrated by the ground control point (GCP) (Fig. 3.1c). Therefore, the transformation matrix from image pixel coordinates to real-world 3D coordinates is available from the intermediate parameters. Duan et al. (2017) and Guo et al. (2018) tested this idea to find the real-world regions of interest (ROI) on corresponding original UAV images, while Lin et al. (2021) developed their algorithms to fix the field map. But they did not publish any source code or tools. It is difficult and time-consuming to reproduce and use these research achievements directly, especially for those who are not familiar with professional photogrammetry programming. So the previous research outputs remain far from the development of user-friendly tools.

The other challenge is the complexity of crop images. As Figure 3.1a shows, there are huge differences in time, sunlight, soil condition, growing stage, and cultivar. It is quite difficult to design a conventional computer vision algorithm to handle all types of variation at the same time. Many studies have applied deep learning algorithms on the broccoli head image segmentation tasks (Blok et al., 2021b; García-Manso et al., 2021; Zhou et al., 2020a), but most of them took images under controlled conditions (indoor

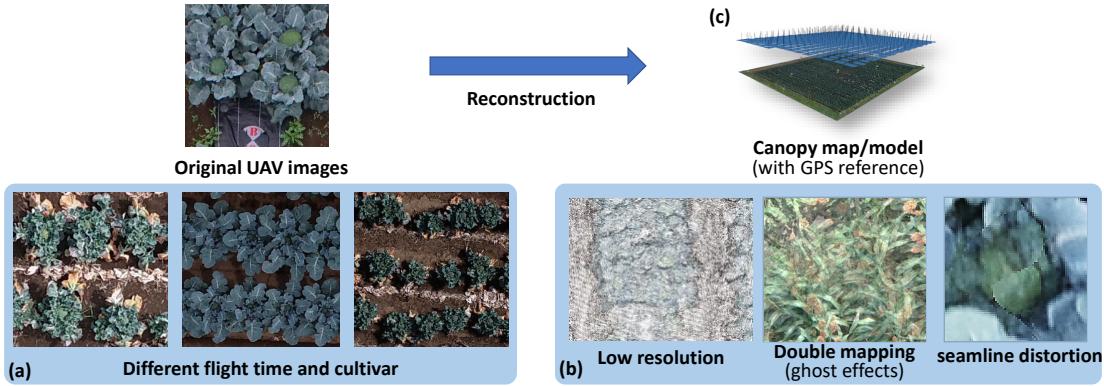


Figure 3.1: The main challenges for organ-level analysis on the UAV-based approach; (a) shows the produced canopy 2D map and 3D model often have poor quality, and suffer from double mapping (ghost effect) and seamline distortion; (b) shows the huge differences in time, sunlight, soil condition, growing stage, and cultivars, which makes the image analysis quite difficult; and (c) is a figure to show the relationship between raw UAV images and the reconstructed canopy model.

or inside a black box). To fulfill the more complex outdoor tasks, a large amount of training data need to be manually labeled. Though there is a public training dataset for cauliflower available (Kierdorf et al., 2022), it cannot be used directly on the broccoli head or another crop. Besides, labeling 14,000 individual plants manually as in the cauliflower dataset is not feasible for all kinds of crops. To decrease the workload of this deep learning solution, it is preferable to simplify the image segmentation, minimize the number of computation tasks, and increase the efficiency of training data acquisition as much as possible.

In this chapter, we aimed to develop several techniques (backward projection, pre-position-guided head segmentation, and interactive annotation) to overcome the aforementioned challenges and provide a highly accurate and labor-saving pipeline for broccoli head size estimation. The objectives were to: (1) develop a general workflow of the broccoli head size estimation pipeline; (2) implement the pipeline for field trials of broccoli growth in 2020 and 2021; (3) validate the estimated head size by comparison with field measurements; and (4) open all the source code of this pipeline to the public domain.

3.2 Materials and Methods

The general workflow of the broccoli head size measurement pipeline is shown in Fig. 3.2 and the Supplementary Video Subsection 3.5.2. The time-series data of all broccoli were collected and visualized using a UAV-based pipeline. The pipeline included the following steps: 1) aerial data collection; 2) data preprocessing and 3D reconstruction; 3) broccoli position detection using YOLO v5 at the seedling stage; 4) broccoli head segmentation using BiSeNet v2, and; 5) geometry trait calculation.

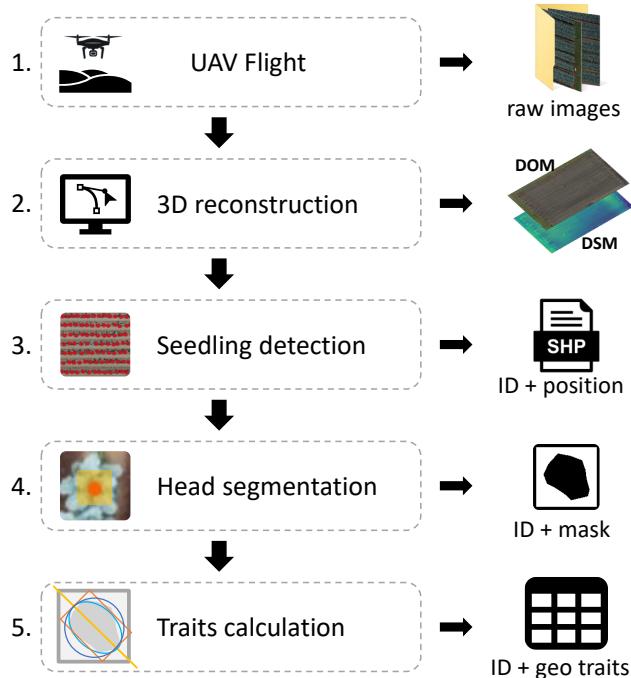


Figure 3.2: Workflow and schematic of the UAV-based pipeline. The time-series head size information (geometry traits) of all broccoli heads was obtained during the growing season. The outputs of each step are the inputs of the next step.

3.2.1 Plot conditions and field data collection

Field trials were conducted at the experimental farm of the Institute for Sustainable Agro-ecosystem Services (ISAS), Nishi-tokyo, Tokyo, Japan ($35^{\circ}43'N$, $139^{\circ}32'E$) in 2020 and 2021 (Fig. 3.3). Detailed meteorological data during the growth period were collected by a local weather station and are shown in Supplementary Table S3.2. The plot sizes were approximately 0.2 and 0.1 ha for 2020 and 2021, respectively. During

the 2-year experiment, the same broccoli cultivar (Jet dome) was planted under the same field management. Machine planting of seedlings at 35 cm intervals in rows 70 cm apart is consistent with local commercial broccoli cultivation regimes (Nishida, 2023).

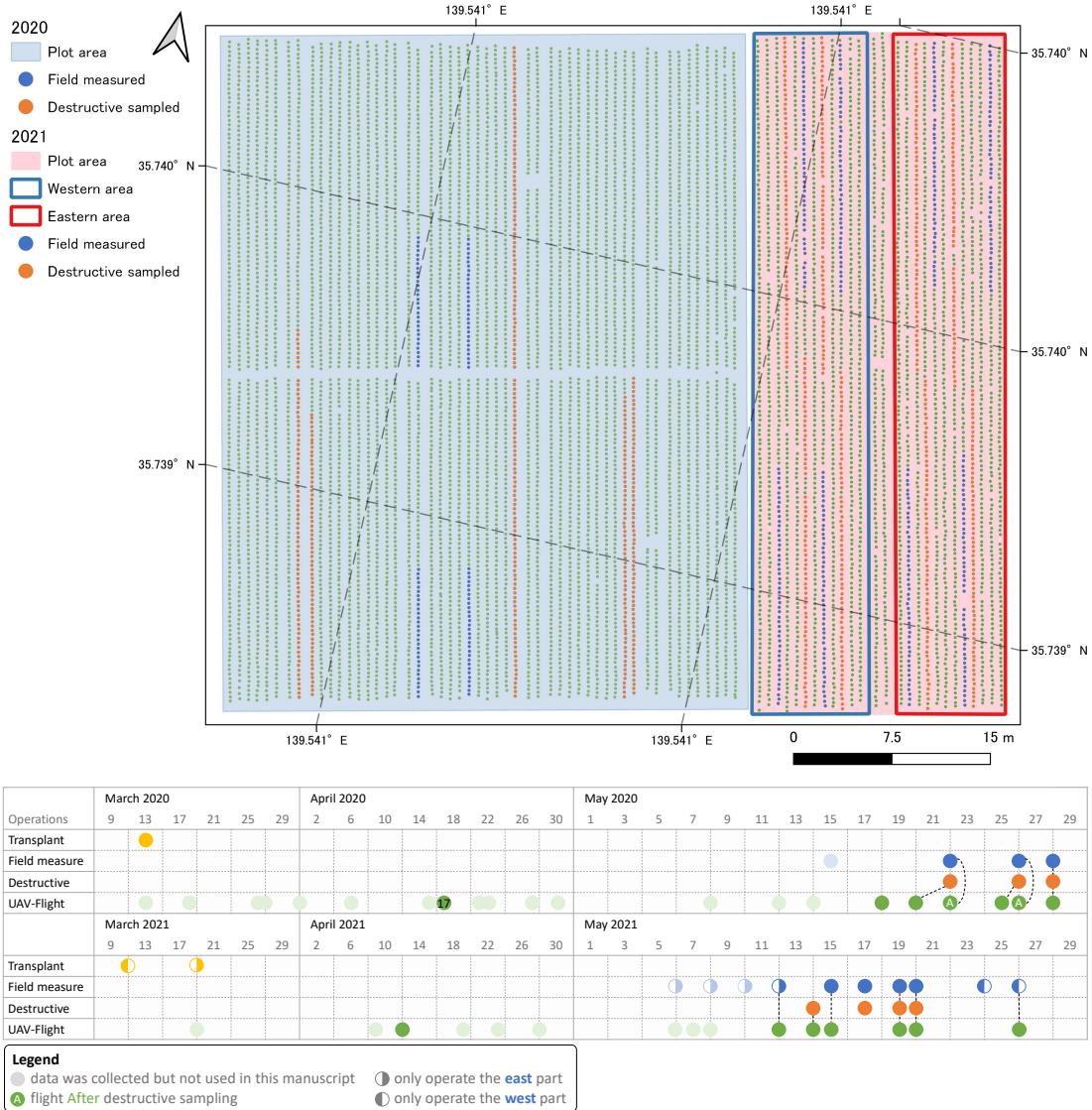


Figure 3.3: Plot conditions and timelines for field operation and data collection. The plots are connected but do not have overlap for both years. Field operations include transplanting, broccoli head non-destructive measurement in the field (field measure), broccoli head destructive measurement, and UAV field investigation. In the field map, the field measurements were conducted on the fixed blue positions on every occasion, whereas the orange positions were the final of all destructively sampled positions. In the timeline, the broken dash line demonstrates how these data were paired.

The field data of the broccoli head size were measured by Nishida (2023) as ground truth for validation. This was conducted manually every 2–3 d using both destructive and non-destructive measurement methods. Non-destructive measures were conducted

directly in the field and destructive measures were conducted indoors. In 2020, the maximum broccoli head length (as head diameter) was measured by the visual judgment of the longest axis. A total of 360 (120×3 times) non-destructive and 434 destructive measurements of 7438 individual broccoli were recorded.

In 2020, the potential of the proposed algorithm was identified. To further validate this algorithm, we improved the measurement method and increased the number of field samples in 2021. The maximum head length (as head diameter) was measured using the maximum value of the length in the 0° , 45° , 90° , and 135° directions. To increase the variation in the broccoli head for each survey, there was an 8-day interval between seeding in the western and eastern parts (timeline in Fig. 3.3, yellow half circle). A total of 2,000 (400×5 times) non-destructive and 557 destructive measurements of 3276 individual broccoli were Nishida (2023). To reduce the workload, only half of the area (east or west) was measured on a certain day (timeline in Fig. 3.3, blue half circle). Head diameter measurements ranged from 2–25 cm.

3.2.2 Data collection and preprocessing

After transplanting, several GCP boards (8 for 2020 and 9 for 2021) were set at the four corners and within the field for UAV data collection. This was an important resource for later 3D reconstruction and time-series alignment. All GCPs were measured using hemisphere real time kinematic (RTK) differential Global Navigation Satellite System (GNSS) devices to obtain geographical coordinates. For developing regions without RTK services, the GCPs' coordinates can be replaced by measuring distances (as scalebar corrector) among each GCP and building a referencing map at the very beginning. The relative coordinates of those GCPs on the referencing map can function the same as the actual geographic coordinates for time-series alignment.

Aerial images were collected by ISAS technical staff using a DJI (Shenzhen DJI Technology Co., Ltd. China) Phantom 4 v2 (camera model FC6310s), a DJI Mavic 2 Pro (camera model Hasselblad) in 2020, and a DJI Phantom 4 RTK (camera model FC6310R) in 2021. The image resolution was the same at 5472×3648 pixels. The

flight height in 2020 was initially 15 m and then decreased to 10 m when the broccoli head turned up. The flight height in 2021 was constantly 15 m. Most of the flights were conducted before the field operation, except on May 22 and 26, 2020. On both these days, the destructively sampled broccoli did not exist in the UAV image; hence, the destructive data were linked to the previous flight (the black broken lines in the timeline in Fig. 3.3). For all other sampling events, data collected on the same day were paired together.

The configuration of the computer for 3D reconstruction was as follows: Intel i9-7980XE central processing unit (CPU) 2.6GHz, 64GB random access memory (RAM), two NVIDIA GeForce GTX 1080Ti graphics processing unit (GPU)s, and Windows 10 Pro 64-bit operating system. Pix4DMapper Pro (Pix4D, S.A., Prilly, Switzerland) was used to perform flight investigations. The default software parameters were used and the GCPs were marked manually on the closest 3–5 images. The reconstruction parameters, digital orthomosaic map (DOM), and digital surface map (DSM) were produced for later use.

The open-source software package QGIS (www.qgis.org) was used to prepare and modify GIS shapefiles, such as field boundaries and field grids on the field map (DOM). The field boundary (plot area) was a rectangular region that tightly wrapped around the broccoli (Fig. 3.3). It could be used to filter the noise outside the broccoli plot. The long boundary edges were parallel to the ridge direction (Figs. 3.3 and 3.4b). Grid plots (yellow squares in Fig. 3.4b) shared the same direction as that of the broccoli ridge and overlapped the boundary. For each grid, the edge length was 2.5 m (approximately 1000–1500 pixels) and contained approximately 50–100 broccolis.

The LabelMe annotation tool (<https://github.com/wkentaro/labelme>) was used to label the training data for the deep learning models. EasyIDP (<https://github.com/UTokyo-FieldPhenomics-Lab/EasyIDP>) was developed and used to locate and crop the same field region imagery on the original UAV images when the DOM was not sufficiently clear for head segmentation (also known as backward projection or reverse calculation, see the Supplementary material 3.5.1 for details about the methodology).

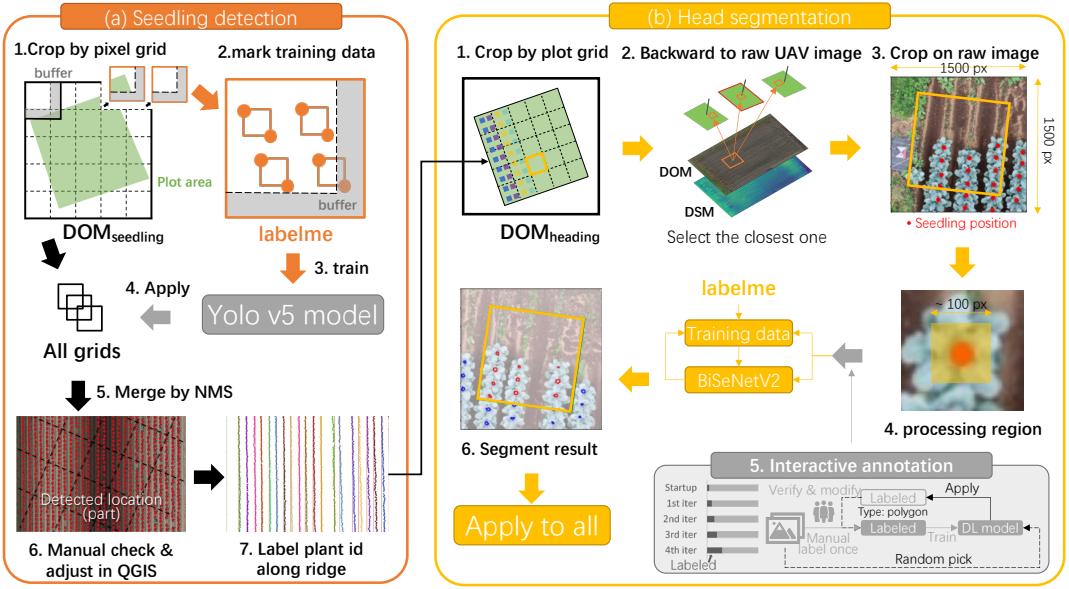


Figure 3.4: Workflow of broccoli (a) seedling detection and (b) head segmentation.

3.2.3 Seedling position detection

The difficulty in detection varied throughout the broccoli growth season. Detection during the seedling stage was simpler than that during the flowering stage, and the latter had complex leaf occlusion. The seedling position of the broccoli was almost the same as that of the broccoli head. Aligned by the GCPs, the positions were consistent through the time-series DOM and DSM of the full growing season. Information of the positions detected during the early stage was directly used for the image analysis during the flowering stage (pre-position-guided), in order to reduce the difficulty in data analysis of each broccoli head.

The flight at approximately 1 month after transplanting was used to detect seedling positions (Fig. 3.3, dark green circle in April). During this period, most broccoli leaves were sufficiently large to be clearly observed from the DOM, but the leaves did not overlap. The uniform light conditions and differentiation between leaves and backgrounds were also taken into consideration when selecting the most suitable detection time.

The broccoli detection procedure is illustrated in Fig 3.4a. The full DOM was first split into several small pixel grids (named ‘sectors’, Fig. 3.4a) along the DOM pixel matrix as the input of YOLO v5 (<https://github.com/ultralytics/yolov5>) by EasyIDP.

The edge length of each sector was 1300 pixels and was buffered with 200 pixels (gray L-shape in Fig. 3.4a) on the lower right corner (1300×1300 to 1500×1500) to avoid individual division on sector edges. Training data with only two sectors were labeled using LabelMe (Fig. 3.4a).

Subsequently, the YOLO v5 detection model with default settings was trained and applied to all sectors. Outliers and noise outside the broccoli field were filtered by the plot area boundary. The duplicated buffer zone detection results were merged using the non-maximum suppression (NMS) algorithm. The center point of the bounding box was then viewed as the broccoli position. Finally, we manually inspected and adjusted the results in QGIS, ensuring no missing or duplicate detections (Fig. 3.4a). The broccoli ID was given from the north to the south of each ridge and ridges from the west to the east by the ridge detection algorithm (Fig. 3.4a; please refer to the links in the data availability section for further details).

3.2.4 Broccoli head segmentation

Leaf movement and occlusion often cause DOM with double mapping, excessive pixelation, and seamline distortions (Lin et al., 2021). It is difficult to obtain a DOM that meets the consistent quality of the entire field for head size estimation. To manage this situation, the same region on the raw UAV images (rather than that on the DOM) was used for analysis by backward projection (Supplementary material 3.5.1). To reduce the workload of segmentation model training, only the area around the broccoli seedling positions was used. Meanwhile, interactive annotation (Ghosal et al., 2019) was applied to decrease the workload of label annotation.

The general workflow for this section is illustrated in Fig. 3.4b. For image data preparation, only flights with visible broccoli heads were chosen for all 12 flight investigations over both years (Fig. 3.3, dark green circles in Mays). The plot area was divided into grids fig:bro6b1). For each grid, the grid boundary and broccoli positions were backward-projected onto the closest raw aerial image (Fig. 3.4b2). Each raw aerial image was cropped into small sectors (1500×1500 px) which were located in the center

and contained broccoli seedling positions (Fig. 3.4b3). Only the square area (approximately 100 px side length) buffered from the seedling positions was used for broccoli head segmentation (Fig. 3.4b4) to eliminate the effects of soil and weeds.

Considering the efficiency of the interactive annotation, the deep learning model can be trained and applied in just a few minutes. Therefore, BiSeNet v2 (Yu et al., 2020) was used as the segmentation model. BiSeNet v2 is a network structure that employs multiple branches to process inputs of different sizes to strike a balance between efficiency and computational cost (Yu et al., 2020, Fig. 1). The BiSeNetV2 network used in the present study was based on the <https://github.com/CoinCheung/BiSeNet> GitHub project. We used both geometric (G) and photometric (P) transformations in data augmentation, as per Blok et al. (2021a). The G strategy consisted of *ShiftScaleRotate* (shift limit = 0.5, scale limit = 0.2, rotate limit = 90) and *VerticalFlip*, which were used to solve the problem caused by our point-based or position-guided segmentation workflow. Ideally, the input images had only one broccoli in the middle of the input image, but in actual practice, the broccoli head position appeared randomly caused by bud-head growth shifting, or even with the probability of multiple or no broccoli heads in the input images. The P strategy simulated “cloudy, sunny” and “day, night” transitions using red-green-blue (RGB) shift (r shift limit = 25, g shift limit = 25, b shift limit = 25) and *RandomBrightnessContrast* (brightness limit = 0.3, contrast limit = 0.3) to address the effects of different weather and light conditions.

Interactive annotation was used to decrease the workload of image labeling (Fig. 3.4b5). Initially, a small number of startup training data (approximately five to ten broccoli heads per flight) were manually marked using LabelMe; then, the segmentation model was trained using the startup data. Next, images were randomly selected and applied to the segmentation results. These results were transformed into LabelMe JSON formats using Python scripts. Subsequently, manual adjustment produced annotations as new training data in LabelMe. The previous steps were iteratively repeated until no significant adjustment was required for the newly applied results.

The verification dataset for the model performance evaluation was also prepared

using the previous interactive annotation. The modified intersection of union (IoU) was used as the evaluation metric. In this case, only the segmentation results inside the grid region (Fig. 3.4b6, red polygon inside the yellow square) were chosen as the final results. The segmentation results attached to the grid bottom and right edge were also removed because of duplication with the neighboring grids. Here, we renamed the modified IoU inside the grid region as “mid-IoU” (middle IoU).

The segmentation model was first trained using only the 2020 dataset for several iterations until a good performance was achieved (no manual modification was required for the model segmentation results). The model was then applied to the 2021 dataset over several iterations. When the model performed well in both years, it was applied to all dataset images, and the segmentation results after the grid boundary filter were saved for the next procedure.

3.2.5 Phenotyping of head traits

The unit of the segmentation polygon in raw aerial images was a pixel, which could not represent the actual head size. However, the actual head size could be calculated using the pixel scale bar from the geo-referenced DOM. The pixel scalebar on the raw images could be derived from the ratio between the grid size in pixels on the raw image and that on the DOM. This concept was then implemented using projective transformation in scikit-image (<https://scikit-image.org>). On some locations, one broccoli may have had duplicated polygons; only the polygon with the largest area was retained, which was accelerated by a k-dimensional tree (KD-Tree) in SciPy (<https://scipy.org>).

For each polygon of broccoli head, the following geometry traits were calculated: 1) max length and min length, from the side lengths of the minimum area bounding box; 2) perimeter, circularity and area of polygon; 3) area of polygon convex hull; 4) major length, minor length and eccentricity of ellipse has the same second-moments as the polygon. 5) equivalent diameter of circle has the same area as the polygon.

The source code used in this manuscript can be accessed at <https://github.com/UTokyo-FieldPhenomics-Lab/UAVbroccoli>.

3.2.6 Field validation for head traits

To test the correlation between field-measured length [dependent variable contributed by Nishida (2023)] and aerial pipeline measured length (independent variable contributed by myself), the coefficient of determination (R^2) using simple linear regression was used as the evaluator. To assess the trends in differences in broccoli size in detail, locally weighted scatterplot smoothing (LOWESS) regression and distribution comparison were also used.

3.3 Results

3.3.1 Broccoli position detection

To provide a general understanding of this procedure, Supplementary Figs. S3.2a to f shows some intermediate results during broccoli position detection. The DOM of the entire broccoli plot was first divided into several small sectors with a buffer zone. All broccoli heads were labeled using bounding box (rectangle) annotation in LabelMe (Supplementary Fig. S3.2a, example sector). The detection model performed as expected in the other sectors (Supplementary Fig. S3.2b: one example sector). The green mask in Supplementary Fig. S3.2b shows the buffer zone that overlapped with neighboring sectors, with the aim of avoiding incomplete broccoli detection on the sector boundary. When merging the detection results for all sectors, duplicate detections were removed by the NMS algorithm (black rectangles in Supplementary Fig. S3.2c). When adjusting the zoom to the full map view, the removed black duplicates were clearly distributed on the sector boundaries (grid lines). The center of the detected bounding box was viewed as the broccoli position; however, it required manual inspections to correct false detections and missing broccoli. In Supplementary Fig. S3.2d, the green dots show the manually corrected (removed) YOLO detection results, and the red dots indicate the final positions used by manual adjustment. Supplementary Figure S3.2e shows the results of ridge-line detection, and Supplementary Fig. S3.2f shows some results of the automatic broccoli ID assignment. In general, broccoli position detection

and semi-automatic labeling functioned as expected.

3.3.2 Broccoli head segmentation

To clearly demonstrate the interactive annotation procedure, Supplementary Figs. S3.2g to i and Supplementary Table S3.1 provide example images and a simple summary from the first to the last iteration. As startup training data, one image was randomly selected from each flight investigation in 2020, and only a few representative broccoli heads were annotated as simply as possible (Supplementary Fig. S3.2g). The BiSeNet model (v0) was then trained using this annotation. For each flight in 2020, one image was randomly selected and applied to the v0 model (Supplementary Fig. S3.2h, red mask). The results were manually adjusted and saved as new training data (Supplementary Fig. S3.2h, blue boundary). The previous steps were iteratively repeated until the model achieved good segmentation results, and this version (in this case, v2) was used to produce 30 validation data with low labor costs. The model performance was evaluated using validation data.

The detailed model performance for each iteration version is presented in Supplementary Table S3.1. With the proposed guidance of the broccoli position and the buffered mask, even a startup with only a few annotations could achieve a mid-IoU of 78.15%. The model performance improved considerably after four iterations and achieved a mid-IoU of 88.33% after the 4th iteration (Supplementary Fig. S3.2i). Then, when the v4 model trained by 2020 data was applied to the 2021 data, the performance decreased to 79.16%, as expected. However, it significantly improved to 91.70% after one additional iteration with six training data points added from 2021. Meanwhile, broccoli head segmentation was more challenging at an early stage (May 18, 2020, and May 12, 2021, with the lowest mid IoU) when the head size was small. This suggested that weakly supervised annotation improved the model performance and decreased the workload in data annotation with iteration.

3.3.3 Head diameter validation

The full map of all calculated HDs from the UAV is shown in Supplementary Fig. S3.3. We found a high correlation between the maximum length of the broccoli head measured (as head diameter) by the UAV image and that measured manually in the field (Fig. 3.5). The UAV method tended to underestimate broccoli growth (trend line above the reference line). However, the overall distribution of the estimated head size was almost the same between the two groups (UAV vs. manual, Figs. 3.5c and d).

3.4 Discussion

The main aim of this chapter was to test the use of UAV-mounted digital measurements of broccoli head size and their use in monitoring for different sellable size classes. Here, we could successfully create a broccoli head size estimating pipeline based on UAV imagery, machine learning / deep learning (ML/DL) for the entire broccoli field. We demonstrated that the head sizes estimated by UAV imagery were highly correlated with the field measurements. Although the current chapter focused on broccoli as a model crop, this framework could be readily applied to other crops such as cauliflower, artichoke, and cabbage. Thus, our case study shows that smart farming techniques have great potential to contribute to the sustainable development of vegetable production.

In addition to estimating temporal variations in head size distribution, our pipeline could visualize spatial variations in individual head size (Supplementary Fig. S3.3). However, it was difficult to visually observe the differences on the ground. Such a spatial variation can often occur on field-grown crops, especially in large-scale agriculture (Quine and Zhang, 2002).

Although the pipeline we developed highlights the benefits and importance of UAV imagery powered by ML/DL for sustainable agricultural development, there are some limitations to its use. First, our system is neither fully automated nor app-based; therefore, farmers without computer science backgrounds cannot use this system directly in their own fields. However, because the source code is open to the public ([https:](https://)

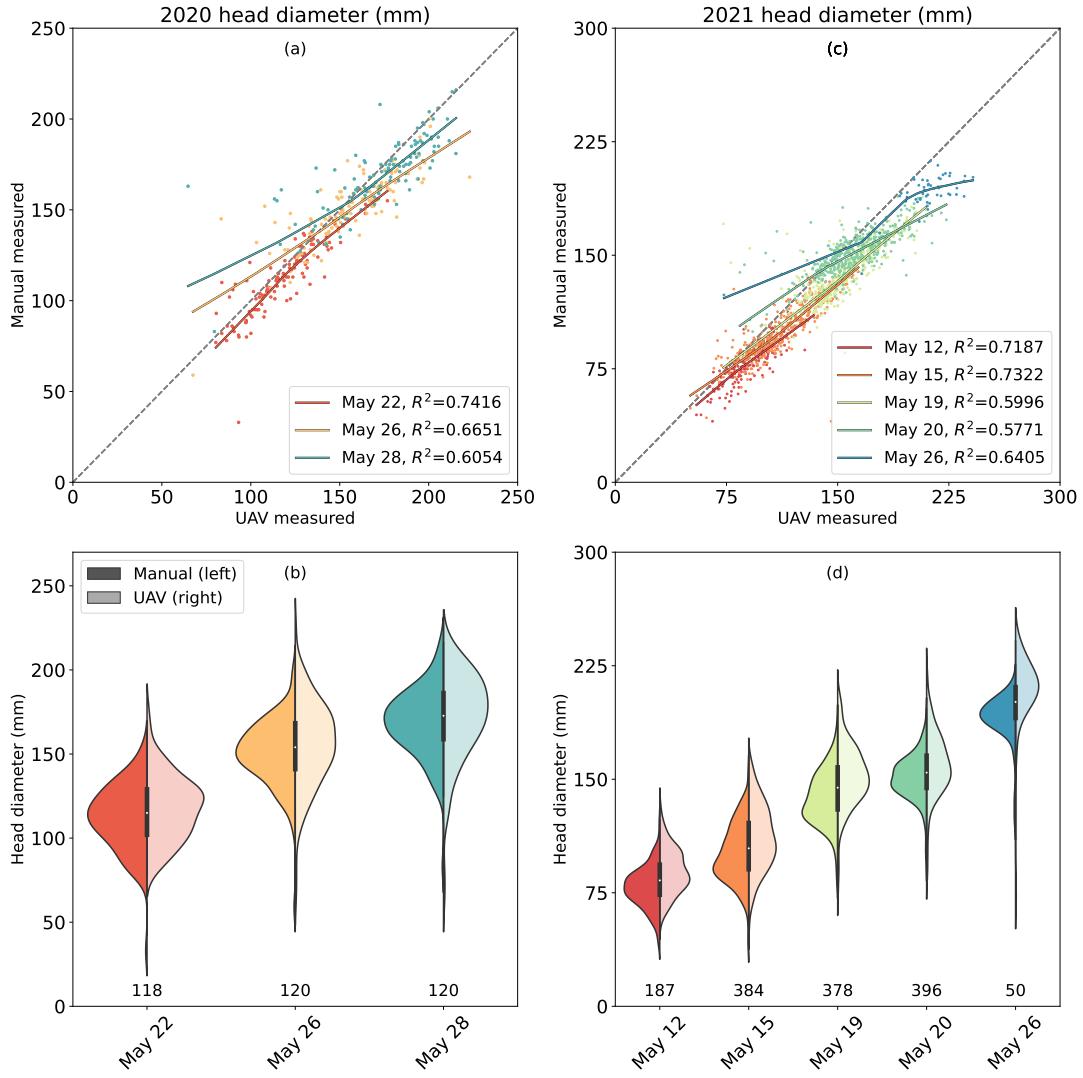


Figure 3.5: Comparison of broccoli head diameter measured from UAV images and obtained using manual field measurements. Different colors represent different sampling dates. In (a) and (c), the curved solid lines represent the trends by locally weighted scatterplot smoothing (LOWESS) regression. In (b) and (d), the violin plots show comparisons of the value distribution; darker colors (left part) show manual field measurements, brighter colors (right part) show UAV measurements, and the values below show the broccoli count.

//github.com/UTokyo-FieldPhenomics-Lab/UAVbroccoli), local agricultural institutes and agricultural companies are able to modify and use the system according to their target. This study is definitely not a one-stop solution, but is a pioneer in real agriculture applications. Second, unlike traditional manual methods with limited throughput, the proposed method can sample every plant in the field at much higher frequency, leading to higher temporal and spatial resolution. The large amount of data generated by

this method is therefore appropriate for data-driven modeling, which could lead to breakthroughs in smart farming. Third, manual inspection of the seedling position detection is required; this step cannot be omitted because this result is the basis for subsequent broccoli segmentation. Detection omissions, duplications, and drifts need to be checked manually on a case-by-case basis. Although it saves considerable effort compared to adding them manually one by one, this inspection still requires several hours to complete in large-scale fields. Additionally, if a broccoli plant dies before the flowering stage, there is a risk of wrong head segmentation generating incorrect results (but in some cases that we observed, it gave an empty segmentation result and was easily discarded). Fourth, the problem of leaf occlusion has not been solved, which remains a challenging problem for plant phenotyping (Zhang et al., 2020b). As broccoli heads are essentially round, one approach is to restore the roundness of the stubs. The circularity and eccentricity of the broccoli head may be used to describe the severity of occlusion. The least squares for round fitting can be used, or the deep learning framework occlusion-aware region-based convolutional neural network (ORCNN) can be applied to obtain improved recovery results (Blok et al., 2021b). However, it requires a depth camera and image pairs before and after occlusion as training data is collected on the ground, which is inconvenient for current UAV applications but warrants further study. For example, multi-spectral and even light detection and ranging (LiDAR) sensors are becoming increasingly cheap, combined with the rapid development of artificial intelligence (AI) algorithms, suggesting this problem could be resolved without unaffordable cost increases. Also, integration of the method with other common management practices such as mulching films with bioplastics could assist in the identification of plants and broccoli heads, particularly when used in conjunction with a multi-spectral sensor. Finally, hardware and software instrument costs should not be omitted. The UAV with RTK (\$6,500), 3D reconstruction software (\$3,499), and a high-performance computer (\$6,000) for computation employed in this chapter would limit the pipeline's widespread use. However, even for a small farm (0.2 ha) with 7,000 broccoli plants, only 2-day difference from the optimal harvest date can result in

an income loss of almost \$2,000. The feasibility of our pipeline on larger farms is also worth being tested in the future. For companies that provide this type of agricultural consulting service, this one-off investment can be offset by the increased profit of many producers. For those economically and socially disadvantaged rural regions, the RTK or the expense of a base-station should not be mandatory. It can be replaced by setting more GCP boards and measuring distances among them as scalebar correctors, to get similar results with relative geographical coordinates. It was suggested that cooperating with local broccoli farmers to test the proposed system without RTK dependences and keeping improving the algorithm performance on the occlusion area will be needed in further work. The head quality and transport costs were also suggested to be integrated into the system to refine its applicability.

3.5 Appendix

3.5.1 Backward projection methodology

The function of this backward projection is projecting ROI from world coordinate to relative UAV raw images. In this section, using “zipfile” and “xml” Python packages to load the external and internal camera parameters generated from structure-from-motion (SfM)-multi-view stereo (MVS) reconstruction projects. Then apply the backward calculation using the pinhole camera model. Finally camera distortion calibration is conducted.

3.5.1.1 Camera parameters loading

The relationship between field and UAV raw images is built after running SfM-MVS software. It has two main parts, external and internal parameters. The external parameters are different for each raw image, including the camera position (x , y , z) in the real-world coordinate (O_{world} , Fig. S3.1.a) and the camera rotation (yaw, pitch, roll). The internal parameters describe the characteristics of the sensor and are the same as each raw image, such as focal length, camera charge-coupled device (CCD) size, and

lens distortion calibration parameters (<https://support.pix4d.com/hc/en-us/articles/202559089-How-are-the-Internal-and-External-Camera-Parameters-defined>). These parameters are available under the Agisoft Metashape and Pix4D project intermediate files.

For Pix4D projects, all these parameters are located in the “1_initial/params” folder, the “calibrated internal camera parameters.cam”, “calibrated camera paramters.txt”, “pmatrix.txt”, and “offset.xyz” are loaded as text directly and parsed in the EasyIDP package without any external packages. For more details about those files, please refer to the Pix4D official documentation (<https://support.pix4d.com/hc/en-us/articles/202559089-How-are-the-Internal-and-External-Camera-Parameters-defined>).

For Agisoft Metashape projects, all these parameters can be obtained either by calling APIs (Professional license required) or by reading zipped xml files in the project file “project.files/0/chunks.zip/doc.xml”. The EasyIDP package chose the zipped xml way without a professional license. The “zipfile” and “xml” packages were used to un-zip and parse these parameters in xml files.

3.5.1.2 Backward projection formulas

The geometry from the real-world coordinate (O_{world}) to image pixel coordinate (O_{pix}) is shown in Fig. S3.1.a-c. There are four coordinate systems, first is the O_{world} whose unit is often meter (Fig. S3.1.a). The second one is the camera coordinate (O_{cam} , Fig. S3.1.b), which makes the camera position to the origin (0,0,0) of coordinates, and the camera optical axis is used as the z-axis (commonly, the point O_{img} is not the center point of plane). The third is the camera CCD coordinate (O_{img} , Fig. S3.1.c) whose unit is often mm. The last one is pixel coordinate (O_{pix}) whose origin is the top left corner in the O_{img} and the unit is pixel.

Let us assume a point $P_{world}(x_w, y_w, z_w)$ in O_{world} , to transform that point into $P_{cam}(x_c, y_c, z_c)$ in O_{cam} (Fig. S3.1.a), the 3×4 transform matrix T could be derived from camera position (t , translational transformation) and camera rotation (R , rotational transformation):

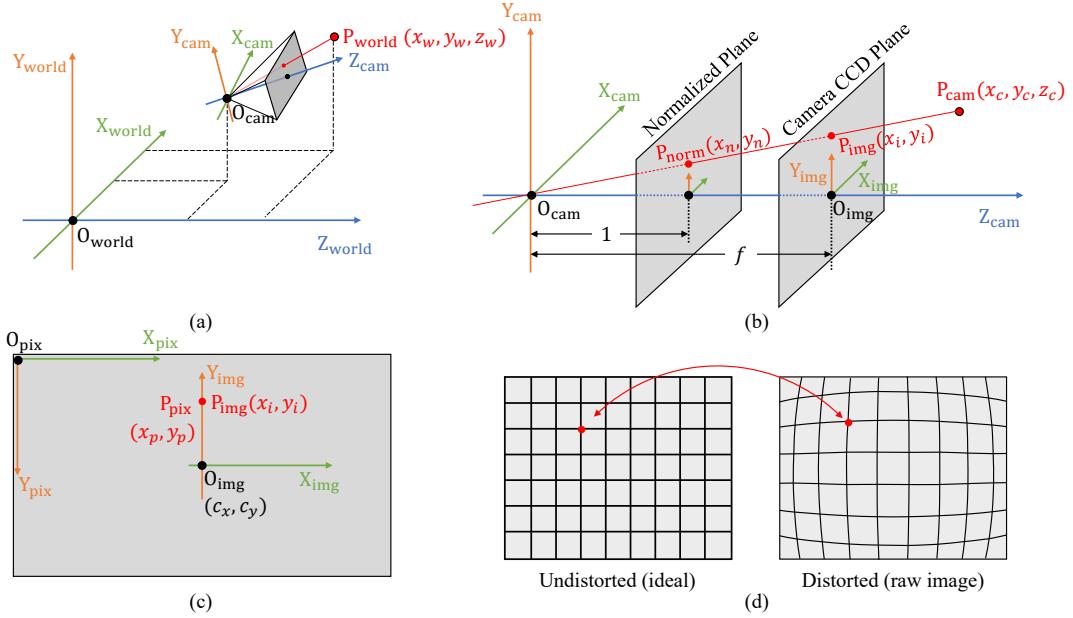


Figure S3.1: Backward projecting one point from the 3D world coordinates to the 2D pixel coordinates on raw UAV images by pinhole camera model. (a) The relationship between world coordinate (O_{world}) and camera coordinate (O_{cam}), linked by camera external parameters (position and rotation). (b) the relationship between camera coordinate (O_{cam}) and image coordinate (O_{img}). (c) The relationship between image coordinate (O_{img}) and pixel coordinate (O_{pix}). (d) The camera distortion calibration between undistorted images and distorted images caused by the lens.

$$\begin{aligned}
 P_{cam} &= T \cdot P_{world} \\
 \begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} &= \begin{bmatrix} R_{11} & R_{12} & R_{13} & t_1 \\ R_{21} & R_{22} & R_{23} & t_2 \\ R_{31} & R_{32} & R_{33} & t_3 \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix}
 \end{aligned} \tag{3.1}$$

Where, t is the 3×1 position matrix, and R is the 3×3 rotation matrix derived by $(\omega, \varphi, \kappa)$ from camera rotation parameters (yaw, pitch, roll) (<https://support.pix4d.com/hc/en-us/articles/202558969-Yaw-Pitch-Roll-and-Omega-Phi-Kappa-angles>):

The distance from the normalized plane to the origin O_{cam} is 1 mm while the distance from the camera CCD plane to the origin O_{cam} is focal length f (Fig. S3.1.b) in mm. The transformation from $P_{cam}(x_c, y_c, z_c)$ to normalized plane $P_{norm}(x_n, y_n)$ and camera CCD plane $P_{img}(x_i, y_i)$ can be derived by triangle similarity:

$$\begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} = f \begin{bmatrix} x_n \\ y_n \\ 1 \end{bmatrix} = f \begin{bmatrix} \frac{x_c}{z_c} \\ \frac{y_c}{z_c} \\ 1 \end{bmatrix} = \frac{f}{z_c} \begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} \quad (3.2)$$

To transform $P_{img}(x_i, y_i)$ in mm to the image pixel coordinate position $P_{pix}(x_p, y_p)$ in pixel (Fig. S3.1.c), the following set of equations should be applied:

$$\begin{cases} x_p = \alpha \cdot x_i + c_x = f \cdot \alpha \cdot x_n + c_x = f_\alpha \cdot x_n + c_x \\ y_p = \beta \cdot y_i + c_y = f \cdot \beta \cdot y_n + c_y = f_\beta \cdot y_n + c_y \end{cases} \quad (3.3)$$

Where, α and β are the pixel resolution whose unit is pixel/mm, and often are the same in pinhole camera models. f_α and f_β is the focal length in pixel. Notably, Pix4D (c_x, c_y) can be obtained directly, while for Agisoft Metashape (https://www.agisoft.com/pdf/metashape-pro_1_7_en.pdf, p. 176), (c_x, c_y) in the xml file is not what is defined here, it is the offset to image center, which actually equals to $(0.5w + c_x, 0.5h + c_y)$, where w and h are the pixel width and pixel height, respectively.

Equations (3.3) can be expressed in the following homogeneous coordinate form:

$$\begin{bmatrix} x_p \\ y_p \\ 1 \end{bmatrix} = \begin{bmatrix} f_\alpha & 0 & c_x \\ 0 & f_\beta & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_n \\ y_n \\ 1 \end{bmatrix} = K \begin{bmatrix} x_n \\ y_n \\ 1 \end{bmatrix} \quad (3.4)$$

Sum up equations (3.1) to (3.4); to transform $P_w(x_w, y_w, z_w)$ directly to $P_{pix}(x_p, y_p)$:

$$\begin{bmatrix} x_p \\ y_p \\ 1 \end{bmatrix} = K \begin{bmatrix} x_n \\ y_n \\ 1 \end{bmatrix} = \frac{1}{z_c} K \begin{bmatrix} x_c \\ y_c \\ 1 \end{bmatrix} = \frac{1}{z_c} \begin{bmatrix} K & 1 \end{bmatrix} T \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} = P_{mat} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} \quad (3.5)$$

where the 3×4 matrix P_{mat} is often called the projection matrix, which can directly transform points in 3D world coordinates to 2D pixel coordinates.

3.5.1.3 Camera distortion calibration

The equation (3.4) transformation is idealized, and the distortion caused by the camera lens is neglected (Fig. S3.1.d). Several camera calibration parameters are used to correct this distortion, including three or four radial distortion coefficients (K_i in MetaShape and R_i in Pix4D) and two tangential distortion coefficients (P_i in MetaShape and T_i in Pix4D). Metashape sometimes provides affinity (B_1) and non-orthogonality (B_2) coefficients in pixels. The correction equation to distorted pixel position (x'_p, y'_p) are as follows:

$$\text{Pix4D} \quad = \quad \begin{cases} x'_p = c_x + x'f \\ y'_p = c_y + y'f \end{cases} \quad (3.6)$$

$$\text{Metashape} \quad = \quad \begin{cases} x'_p = c_x + x'f + x'B_1 + y'B_2 \\ y'_p = c_y + y'f \end{cases}$$

where:

$$\text{Pix4D} \quad = \quad \begin{cases} x' = k_0x_n + 2T_2x_ny_n + T_1(r^2 + 2x_n^2) \\ y' = k_0y_n + 2T_1x_ny_n + T_2(r^2 + 2y_n^2) \end{cases} \quad (3.7)$$

$$\text{Metashape} \quad = \quad \begin{cases} x' = k_0x_n + 2P_2x_ny_n + P_1(r^2 + 2x_n^2) \\ y' = k_0y_n + 2P_1x_ny_n + P_2(r^2 + 2y_n^2) \end{cases}$$

and:

$$r \quad = \quad \sqrt{x_n^2 + y_n^2}$$

$$k_0 \quad = \quad \begin{cases} 1 + R_1r^2 + R_2r^4 + R_3r^6 & (\text{Pix4D}) \\ 1 + K_1r^2 + K_2r^4 + K_3r^6 + K_4r^8 & (\text{Metashape}) \end{cases} \quad (3.8)$$

3.5.2 Supplementary video

An illustration video about the general pipeline of this study can be accessed here <https://youtu.be/SYuOCVqgtrU>. The background music used in this video is copyright-free music from freepd.com.

3.5.3 Supplementary figure and tables

Table S3.1: Mid intersection of union (IoU) changes around each iteration in weakly supervised learning. Firstly, four iterations were conducted on the 2020 dataset only, and then applied to the 2021 dataset.

Training data file number	Model version	Training time (s)	Mid IoU in 2020 (%)							Mid IoU in 2021 (%)						
			May 18	May 20	May 22	May 25	May 26	May 28	Mean	May 12	May 14	May 15	May 19	May 20	May 26	Mean
Startup 2020x6	v0	100.8	69.82	82.08	84.49	77.82	68.09	86.60	78.15	-	-	-	-	-	-	-
v0 add 2020x6	v1	166.1	76.28	87.60	88.60	84.16	81.75	89.02	84.57	-	-	-	-	-	-	-
v1 add 2020x9	v2	420.4	84.44	89.35	89.07	85.94	88.17	89.76	87.79	-	-	-	-	-	-	-
v2 add 2020x14	v3	649.2	84.91	89.77	89.30	86.56	90.12	90.03	88.45	-	-	-	-	-	-	-
v3 add 2020x12	v4	1085.2	83.37	90.04	89.57	86.49	90.15	90.35	88.33	47.09	79.98	80.29	91.77	91.26	84.56	79.16
v4 add 2021x6	v5	1267.5	85.04	90.20	90.35	86.13	90.63	90.49	88.81	85.07	88.84	91.24	94.59	94.55	95.88	91.70

Table S3.2: Meteorological data during the broccoli growth period.

Date	Mean temperature (°C)	Mean precipitation (mm)	Sunshine duration (h)	Wind speed ($m \cdot s^{-1}$)
2020.03	10.5	103.5	183.4	2.0
2020.04	12.8	228.5	218.1	1.9
2020.05	19.6	103.0	174.9	1.3
2021.03	12.5	143.0	186.5	2.0
2021.04	15.0	104.5	218.4	1.7
2021.05	19.7	72.0	145.0	1.3

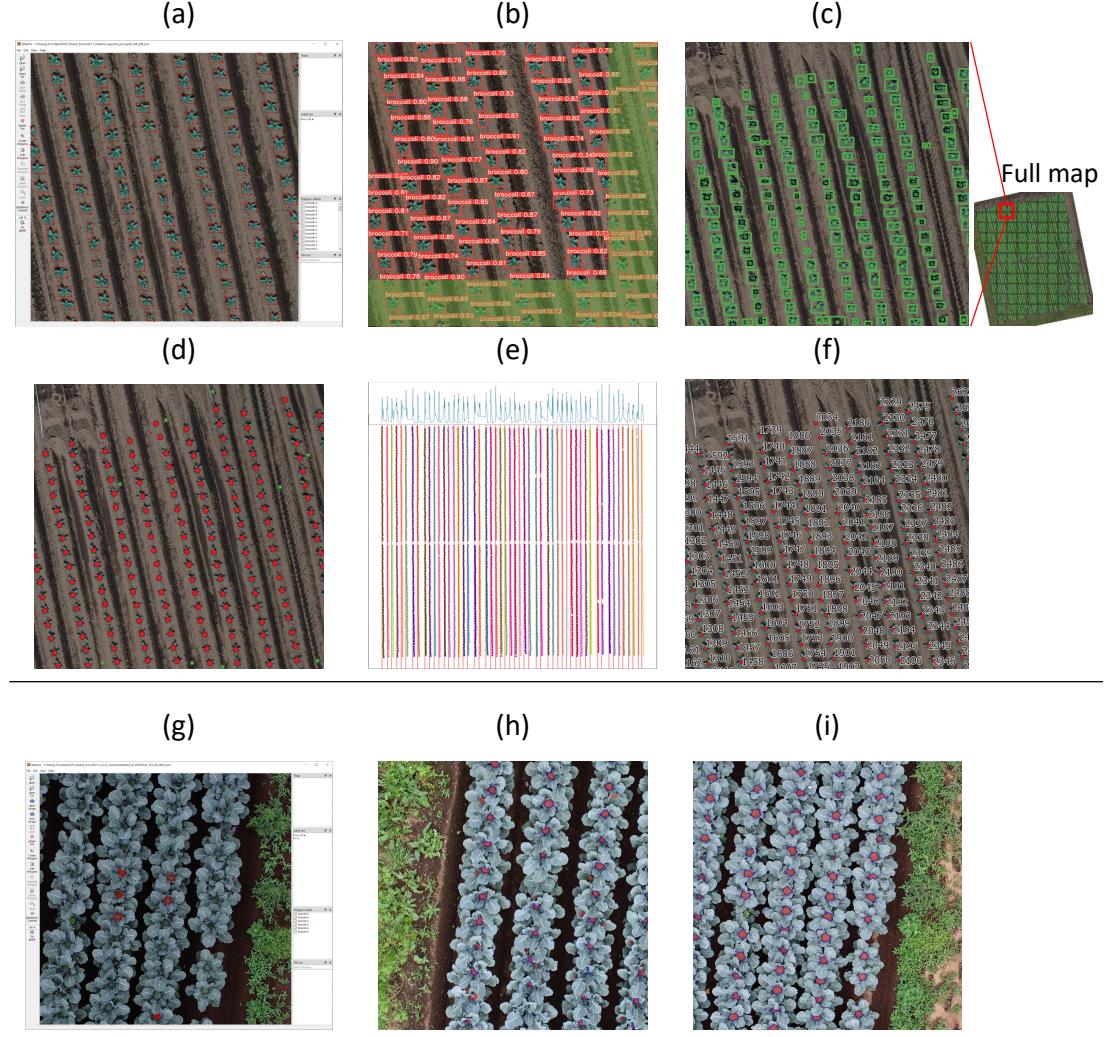


Figure S3.2: Examples of 2020 broccoli seedling position detection (a-f) and head segmentation by interactive annotation (g-i). (a) One annotated training data by LabelMe. (b) YOLO v5 detected results; the green part is the buffer zone to avoid the broken broccoli plants at the edge. (c) Duplicate detection in buffer zone removed by the non-maximum suppression (NMS) algorithm. Black shows removed duplicate detection, green shows those that were retained, and green dots are the center points as the broccoli position. (d) Red dots show the manually adjusted positions by QGIS. (e) Ridge detection by identifying the peak of points distribution. (f) Automatic placing of plant ID along the ridge. (g) Startup training data annotation made by LabelMe; only a few annotations were required. (h) After the first iteration. The red polygons are the segmentation results (as auxiliary annotations) trained by the startup data and the blue polygons are manually adjusted according to the previous results. (i) After the fourth iteration, almost no manual adjustment was required in this case.

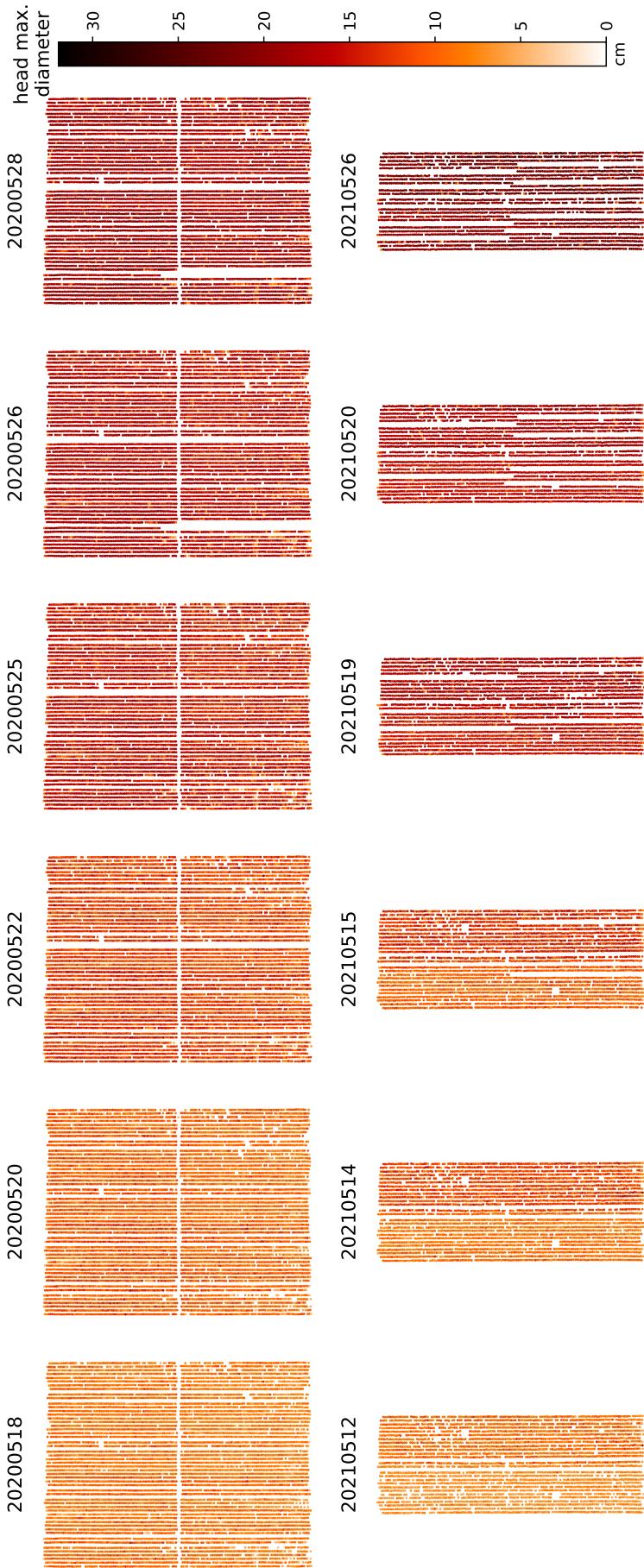


Figure S3.3: Head size of all broccoli heads based on UAV imagery; color represents the maximum head diameter.

Chapter 4

Data fusion for virtual visualization

4.1 Introduction

The previous two chapters developed two approaches that can obtain broccoli head parts using close-range and aerial approaches. The close-range approach (Chapter 2) can obtain high-quality 3D models of the broccoli heads, while the aerial approach (Chapter 3) can obtain low-quality broccoli canopy models of the entire field. As the next step, we aim to transform and place the high-quality head 3D models into the low-quality full-field canopy models, to enable 3D virtual visualization and explanation for a more intuitive assessment of the growth status as compared to numerical statistical values. This technique is fundamental and critical for virtual farmland and digital twin technologies (Pylianidis et al., 2021; Slob et al., 2023). To the best of our knowledge, no published research has yet addressed this problem in the actual vegetable crop fields. However, several challenges need to be addressed before this goal can be achieved.

The first challenge is to determine a more accurate geographical location. In Chapter 3, we segmented the broccoli head on the raw images without geo-coordinates due to the low quality of the aerial model. However, converting 3D geographical coordinates (3D canopy model) to the 2D pixel coordinate (2D raw image), named backward projection, one dimension is lost and cannot be reversed. Therefore, when convecting the broccoli head segmentation results in 2D pixel coordinates back to the 3D geographical coordinate, only a ray (Z_{cam}) connecting the camera position (O_{cam}) and that point on

the camera charge-coupled device (CCD) plane can be produced for each vertex (point) of broccoli head polygons (Supplementary Fig. S3.1a), rather than obtaining accurate 3D geographical points.

To solve the accurate 3D position, there are currently two solutions: 1) intersecting two or more rays from different perspectives to determine the 3D intersection point, which is almost the same as the structure-from-motion (SfM) process in photogrammetry; and 2) computing the intersection points between the ray and the 3D surface of the broccoli canopy model. For the first solution, matching the segmentation polygons from different perspectives for the same broccoli is not difficult, but matching each vertex is challenging. Since the shapes in different perspectives have slight differences, the vertex count may not match, let alone the vertex with "random" orders. For the second solution, it is necessary to check whether that ray intersects with all the triangular faces and calculate the intersection point if it does. On one hand, this approach requires immense computation, as a 3D broccoli canopy model typically contains millions of triangular faces, and each broccoli head polygon has dozens of vertices while there may be several thousand of them. On the other hand, a ray may intersect with multiple triangular faces, such as both the leaf and head triangular faces, so it is necessary to determine which is the actual triangular face of the broccoli head. Shao et al. (2020) optimized this approach by projecting the 3D triangular faces onto the raw image to generate a depth image in order to resolve the dimension loss, which is still computationally intensive for rendering each pixel, and code is not available.

In Chapter 3, we proposed a simplified solution. Since only the actual size is required, rather than accurate positions for that chapter, we assumed that the relative sizes of the broccoli polygon and the grid boundary remain unchanged in the two perspectives (Fig. 4.1a-b). In practice, we applied projective transformation, which directly stretches the image according to the regions of interest (ROI) vertices (blue broken lines in Fig. 4.1). Therefore, although the positions may not fit perfectly due to slight angle variations in perspectives between the raw image (Fig. 4.1a, with a slight slant) and the digital orthomosaic map (DOM) (Fig. 4.1b, the vertical view), we can

obtain the actual geographical sizes of broccoli heads, which is enough for the objective of Chapter 3. However, the position accuracy of this method may not be enough for the 3D visualization objective of this chapter and requires further optimization.

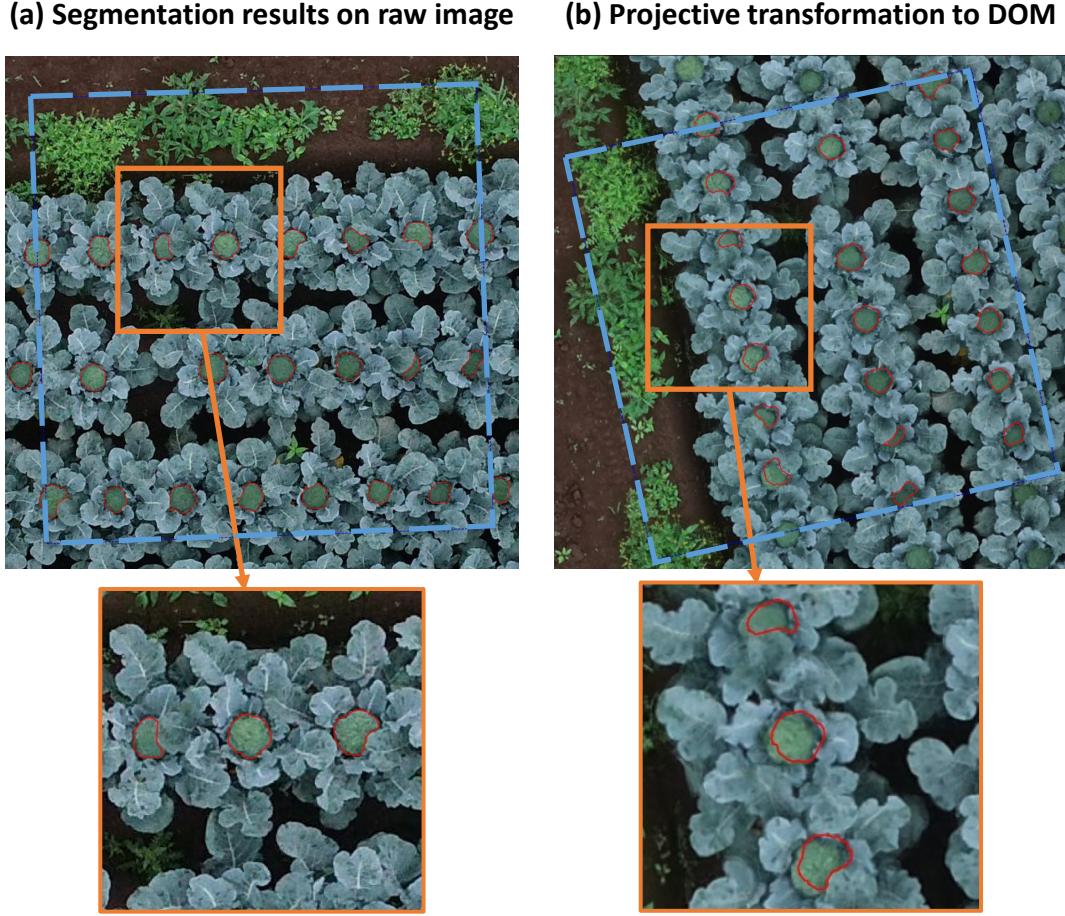


Figure 4.1: Challenge to forward results from (a) pixel coordinate on raw image to (b) geographical coordinate on DOM. The red polygons are broccoli head segmentation results while the blue broken lines are the boundary of the current plot grid. (b) shows the projective transformation on the grid vertices has slight deviations.

The second challenge is how to properly match and adjust the close-range broccoli models to the aerial results. The first step is calibration, which makes the morphological traits of broccoli heads measured from the aerial closer to the “ground truth”. For the model calibration of complex cases, compared to linear regression, the machine learning models such as supported vector machine (SVM) and random forest (RF) were often used by several studies (Lu et al., 2022; Nguyen et al., 2023). However, the selection and tuning of the machine learning model also pose a challenge. Wang et al. (2019) reported

that classification and regression tree (CART) outperformed SVM, RF, and gradient boosting decision tree (GBDT) on plant classification tasks. However, Han et al. (2021) found that SVM performed best on flower classification tasks compared to RF, CART, linear discriminant analysis (LDA), and k-nearest neighbor (KNN). Furthermore, Han et al. (2019) reported that RF yielded better results on biomass prediction than multiple linear regression (MLR), SVM, and artificial neural network (ANN). These studies indicate that the performance of different machine learning algorithms varies depending on the task at hand. From our experience, model selection is an empirical process that may require extra time to compare mainstream algorithms. To be effective in practice, the demand for an optimized machine learning system that can automatically choose the algorithm and set its hyperparameters for non-experts is increasing (Feurer et al., 2015). The next step after calibration is converting the broccoli heads in the database and fitting them to the calibrated morphological traits. Since we only destructively sampled around 200 broccoli heads to obtain high-quality models, we still cannot cover all shapes of broccoli heads grown in the entire field. Therefore, template matching and transformation methods should also be developed.

In this chapter, we aimed to develop a 3D virtual visualization technique that can fuse high-quality 3D head models, obtained using close-range reconstruction (Chapter 2), with full-field canopy models, obtained through aerial photogrammetry (Chapter 3). Our objectives were to: 1) develop an optimized method to obtain more accurate geographical locations of segmented broccoli heads from raw images; 2) use the latest auto machine learning (AutoML) technique to generate a calibration model for aerial measurements of head traits; 3) develop a template matching method that can quickly find the broccoli model in the database with the smallest difference to the calibrated traits; 4) apply geometric transformations to the matched head 3D model template to reduce differences to the calibrated traits; and 5) develop a visualization script to display the results in 3D.

4.2 Methods and Materials

The general workflow for 3D visualization can be summarized into three main parts: 1) obtain the accurate geographical position of each broccoli head from the aerial approach (optimize Chapter 3); 2) calibrate the morphological traits of the broccoli head from aerial measurements, then locate and transform the nearest template from the high-quality broccoli head database (produced by Chapter 2); and 3) fuse the close-range and aerial results and visualize them in 3D.

4.2.1 Data collection and preprocessing

The broccoli data used in this chapter was obtained in the spring of 2022 (Chapter 2) and the detailed plot conditions were introduced in Subsection 2.2.1. The high-quality 3D model of 189 destructively sampled broccoli heads was obtained from the workflow in Subsection 2.2.2. Additionally, the field 3D model of the aerial photogrammetry from this year was obtained from the workflow introduced in Subsection 3.2.2. However, we changed the aerial photogrammetry software from Pix4DMapper Pro (Pix4D, S.A., Prilly, Switzerland) to Agisoft Metashape (Agisoft LLC, St. Petersburg, Russia). Then, the same workflows for broccoli detection (Subsection 3.2.3) and segmentation (Subsection 3.2.4) were applied to obtain the broccoli head polygons from the raw images. The morphological traits of broccoli heads from the close-range database and the aerial survey were calculated by Subsection 2.2.3 and Subsection 3.2.5, respectively.

4.2.2 Optimized positioning

Inspired by the depth image rendering proposed by Shao et al. (2020) and the projective transformation used in Chapter 3, a control point array covering the ROI and the broccoli segmentation results (Fig. 4.2a) was generated in geographical coordinates on the DOM. These control points were then backward projected onto the raw images (Fig. 4.2b) and the piecewise affine transformation was used to revert coordinates according to these control points in both coordinates. Pitiot et al. (2006) previously applied this technique to register biological images for volume reconstruction. This transfor-

mation was applied to the broccoli head segmentation results to obtain more accurate geographical positions.

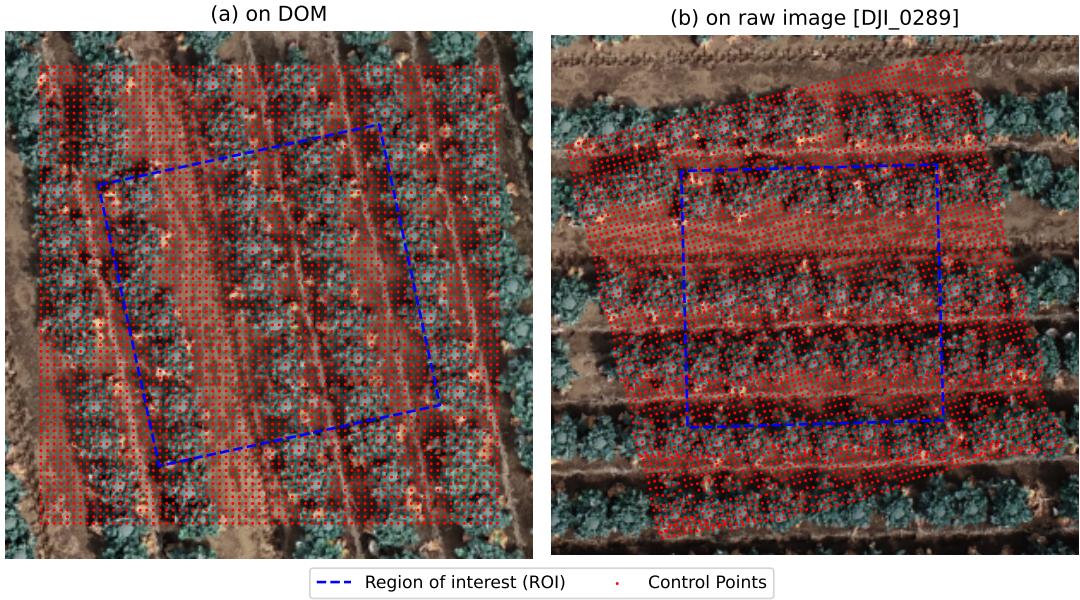


Figure 4.2: Control points between geographical coordinates on DOM and the pixel coordinate on raw image. The blue broken lines are the ROI boundary of the current grid, the red dots are the control points.

After converting the broccoli head segmentation results to the graphical coordinates, the center of the head was calculated in order to position the high-quality 3D head models from close-range database. The circle fitting procedure described in Blok et al. (2021b) was applied to the convex hull of the broccoli head to determine its center, which partially solves the issue of leaf occlusion to some degree. To obtain the height (z) values of that center point, the mean value of all points in the head region ≥ 95 th percentile height was calculated.

4.2.3 Template matching and transformation

Before performing template matching and transformation, a calibration model is required to improve the performance of the system. This model should be trained by relating the aerial morphological traits to the close-range morphological traits. Since we sampled the close-range broccoli heads destructively after the aerial survey, we were able to obtain such data pairs. These data pairs were divided into training data

and validation data with a 4:1 ratio. The training data was used to train an AutoML multi-output regression model (Feurer et al., 2020) to calibrate the morphological traits from aerial surveys, while the validation data was used to assess the calibration model’s performance.

The next step is to find the template broccoli head from the close-range database. We modified the normalized cross-correlation of image template matching (Yoo and Han, 2009), which helped us find the closest 3D head template with the smallest differences:

$$\bar{j} = \underset{j}{\operatorname{argmin}} \sum_i^n \sqrt{\sum_{i,j}^m |T_i - D_{i,j}|^2} \quad (4.1)$$

where, T_i is the i-th of calibrated morphological traits of the current aerial broccoli head with a total number of n (in our data, $n = 6$, including length of major axis, length of minor axis, width of minimum area rectangle, length of minimum area rectangle, area, and convex area); j is the j-th of broccoli head in the close-range template database with a total number of m (in our data, $m = 189$); $D_{i,j}$ is the i-th of morphological traits of j-th of broccoli head in the template database; \bar{j} is the j-th of broccoli head that finally matched.

After obtaining the closest template, the ratios between the calibrated values and template values were used as the zoom ratios for template transformation. To be more specific, the width and length of the minimum area rectangle were used as the horizontal and vertical zoom ratios along the minor and major axes, respectively.

4.2.4 3D virtual visualization

Since the triangular face count of each close-range high-quality broccoli head model is often over 50,000, it is not feasible to directly render the entire field with thousands of broccoli. Instead, we render one ROI grid at a time to decrease the demand for computational resources. The EasyIDP package (<https://github.com/UTokyo-FieldPhenomics-Lab/EasyIDP>, Supplementary material 3.5.1) was used to crop the ROI broccoli canopy 3D point cloud (by aerial photogrammetry), and the transformed broccoli heads were placed at the 3D geographic head center points. The Open3D Python package (Zhou

et al., 2018, <https://github.com/isl-org/Open3D>) was used to process and visualize the 3D models. The code and script for data processing and interactive visualization can be found at <https://github.com/UTokyo-FieldPhenomics-Lab/BroccoliHead3D.ipynb>.

4.3 Results

The aforementioned workflow demonstrated the feasibility of most ROI grids, particularly those with dates close to the optimal harvest date with larger broccoli head sizes. In this section, some representative examples were used to illustrate.

4.3.1 Optimized positioning

Figure 4.3a shows the segmentation results of broccoli on the raw image, while Figure 4.3b displays a comparison between the optimized positioning method (blue polygons) and the one used in Chapter 3 (red polygons). The broccoli heads in the bottom row (Fig. 4.3a) correspond to the left column in Figure 4.3b, which has an almost vertical perspective and thus conforms to the position in geographic coordinates. On the other hand, in the top row of Figure 4.3a (corresponding to the right column in Fig. 4.3b), the perspective is not perfectly vertical. As mentioned in the introduction of this chapter, the method used in Chapter 3 (red polygons) shows clear deviation, while the proposed optimized method (blue polygons) still accurately displays on the geographical positions of the broccoli heads.

4.3.2 Template matching and transformation

The full pairwise data was split into training and testing sets using a 4:1 ratio, and the calibration AutoML model was trained on the training set only. The components of the calibration AutoML model are shown in Table 4.1. Three RF models (IDs 72, 73, and 69) with larger ensemble weights (≥ 0.20) were selected, as well as two KNN models (IDs 67 and 20). These models were combined as a mixture model, and the results output by each model were weighted according to their respective weights to obtain the final result.

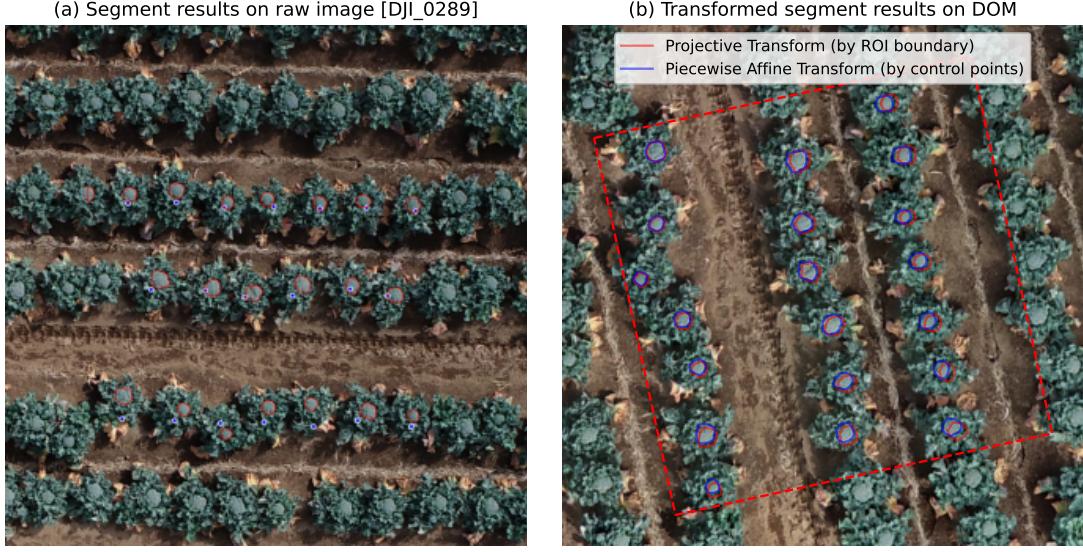


Figure 4.3: Head segmentation forward location comparison between projective (in Chapter 3) and piecewise affine transformation (this chapter). (a) is the head segmentation results on the raw image. They are forward transformed to corresponding positions on the DOM using two different transformations.

Table 4.1: The components of AutoML calibration model

model id	rank	ensemble weight	type
72	1	0.22	random forest
55	2	0.08	random forest
2	3	0.02	random forest
73	4	0.28	random forest
69	5	0.24	random forest
67	6	0.06	k nearest neighbors
20	7	0.10	k nearest neighbors

The validation was conducted on the remaining testing dataset. According to the results (Fig. 4.4), the calibration model obtained good correlation ($R^2 \geq 0.84$) between calibrated values and validation values for the six morphological traits. Compared to the morphological traits without calibration (red), the calibrated traits (blue) have a higher correlation and lower root mean square error (RMSE). The calibrated traits (solid regression line) are closer to the standard line, which means that the AutoML calibration model corrects the underestimated trend (mainly due to occlusion for small broccoli heads) of aerial measurements.

The calibration also contributed to better template-matching results. Take one plot

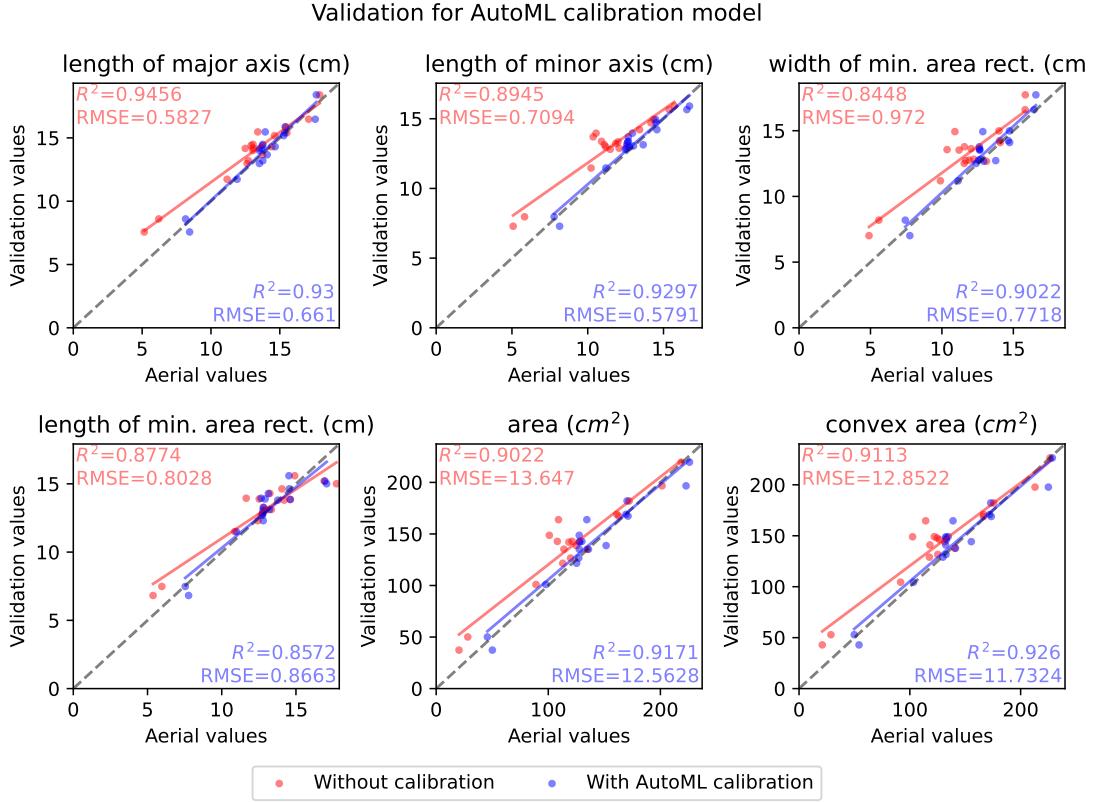


Figure 4.4: Validation for the AutoML calibration model.

with variable broccoli head sizes (ranging from 10cm to 22cm) as an example (Fig. 4.5). The calibrated results showed higher correlation and lower RMSE compared to those without calibration. The distribution was also closer to the standard line, which could result in a smaller transformation for the templates.

4.3.3 3D virtual visualization

Figure 4.6 shows a representative plot grid for 3D virtual visualization with varying broccoli head sizes before the optimal harvest date. Figure 4.6a shows the broccoli head segmentation results on the raw image, while Figure 4.6b shows the aerial broccoli canopy 3D point cloud models. It was almost impossible to observe the broccoli head directly on the point cloud model (Fig. 4.6b). By using the workflow proposed in this chapter, the high-quality 3D models were transformed and accurately placed into the point cloud model with low resolution (Fig. 4.6c). Figure 4.6d shows another perspective of the 3D visualization, and Figure 4.6e shows the model details after zooming closer.

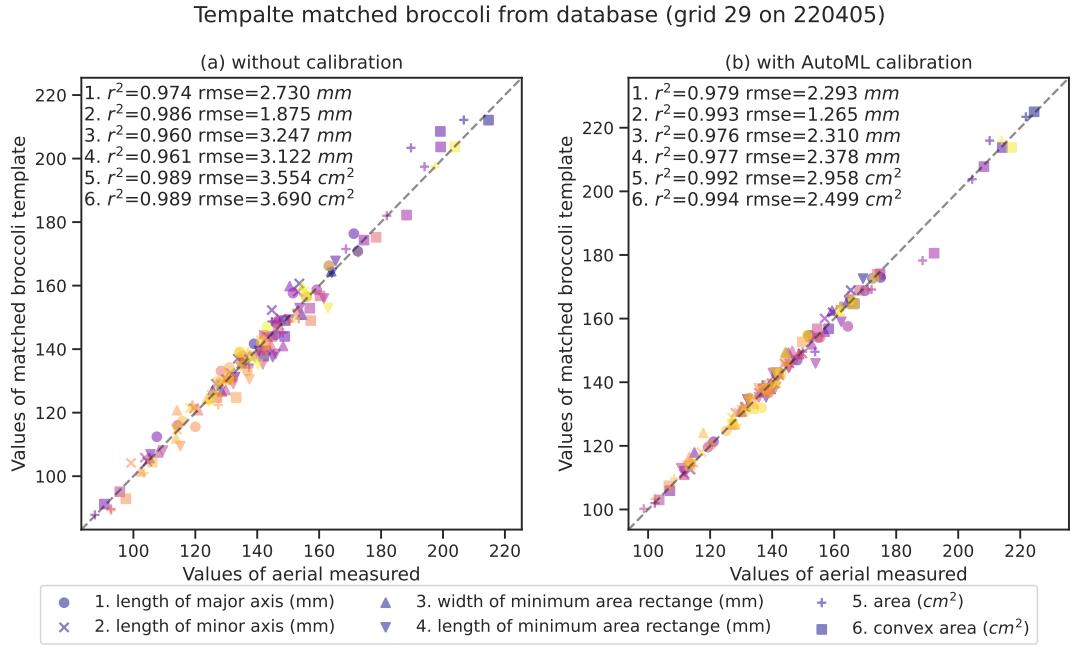


Figure 4.5: The closest matched broccoli head template 3D model to each aerial segmentation results. Different colors represent different broccoli heads.

Some of the broccoli heads touched the plot boundary (two on the top right and one on the bottom left, Fig. 4.6c) and did not show up; because we defined the lower and right boundaries as belonging to the neighbor grids. These broccolis showed up in that grid.

4.4 Discussion

The 3D virtual visualization of broccoli heads introduced in this chapter provides, for the first time, a data fusion method that combines the advantages of high-quality close-range reconstruction (Chapter 2) and high throughput aerial survey (Chapter 3). We observed that the proposed method, including optimized positioning (Fig. 4.3), automated machine learning calibration (Fig. 4.4), and template matching and transformation (Fig. 4.5), had clear improvements compared to previous or controlled methods in both correlation and RMSE. The final 3D virtual visualization (Fig. 4.6) shows the feasibility of this technique as a fundamental and critical component for virtual farmland and digital twin for future studies.

Obtaining paired training data (the same broccoli plant appears in both close-range

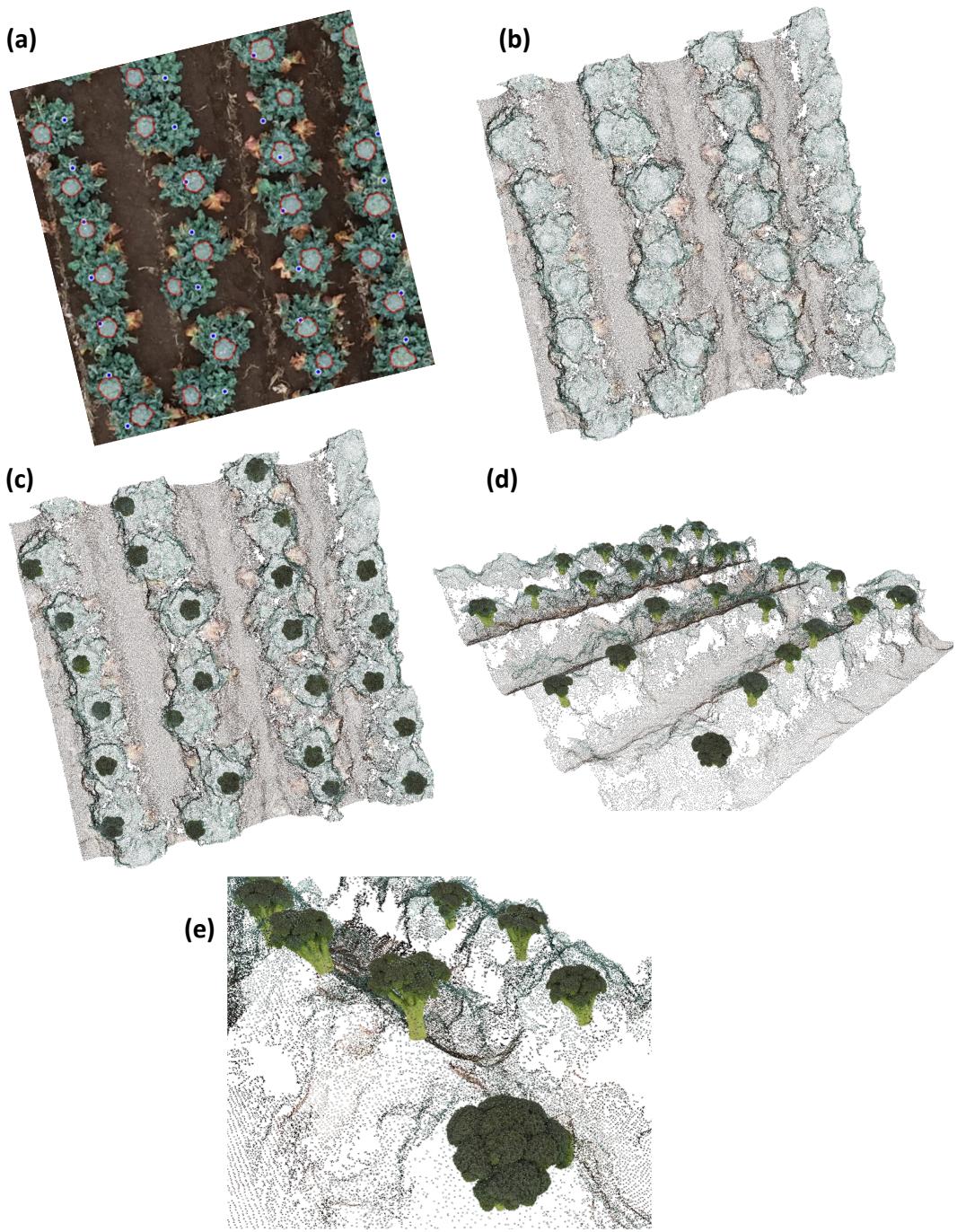


Figure 4.6: One example of 3D virtual visualization for optimal harvest date; (a) is the broccoli head segmentation results on raw aerial image; (b) is the top view of original aerial 3D point cloud models for broccoli canopy; (c) is the top view of canopy model with matched high-quality 3D broccoli heads; (d) is the side view; and (e) is the zoom in detailed view.

and aerial view) for model calibration is a time-consuming and challenging process. It involves careful verification of the field and destructive sampling ids to ensure matching accuracy, as well as data cleaning to remove outliers. For example, in this experiment,

due to the complexity of field conditions, the segmentation results for 3 broccoli heads on the original image were not perfect, resulting in great differences in morphological traits between aerial and close-range views. Adding them to our pre-experiment would decrease the model's correlation to only around 0.5. After removing these outliers, the calibration model correlation (R^2) increased to over 0.84 (Fig. 4.4). However, this does not mean that the process is mandatory when applied to new broccoli fields. Due to the similarities in the growth and occlusion of broccoli, the model we trained can theoretically be transferred to new broccoli fields for calibration. In addition, based on our comparison results (Fig. 4.4), even without calibration, the RMSE is $\leq 0.97\text{ cm}$ for length and $\leq 13.6\text{ cm}^2$ for the area. We can still find corresponding templates in the template database for transformation with an acceptable level of error (RMSE $\leq 3.1\text{ mm}$ for length and $\leq 2.76\text{ cm}^2$ for head area) without calibration. But we need to consider that smaller broccoli may be underestimated. Since the optimal harvesting date usually consists of larger broccoli, the impact of this underestimation is still acceptable. It was suggested for further experiments to evaluate the transferability of the AutoML calibration model and the impact of skipping the calibration in new broccoli fields.

As a pioneer study in its early stage, we have to admit that the proposed method in this chapter has some limitations. Firstly, the current visualization is still rough and limited by performance, which requires breaking down the entire field into smaller parts for 3D virtual visualization. It was suggested to use game engines to develop a user-friendly graphical user interface (GUI), or even a virtual reality (VR) or augmented reality (AR) interface. Also, integrating all the methods in the thesis into a software product for users to conveniently apply in their field should be considered. Secondly, the occlusion problem has not been completely solved yet. The current method of addressing occlusion is through AutoML model calibration, but the statistic-based calibration is difficult to surpass the shape-based occlusion repair. Blok et al. (2021b) proposed an approach based on occlusion-aware region-based convolutional neural network (ORCNN) to repair missing parts. It requires a camera with its position fixed over a broccoli head. Then a fully exposed broccoli head image was collected as the ground truth and leaves were

used to create training data with varying degrees of occlusion. Further adjustments are needed for the application of drone imagery in the future. Thirdly, there is only one cultivar in the close-range 3D broccoli head model database, which is not sufficient to reflect the differences between different broccoli cultivars. Further cooperation with farms is needed to acquire more high-quality 3D models of different broccoli cultivars for further expansion of the database. Fourthly, the current broccoli placement strategy simply performs geometric transformations and places the broccoli directly back to the center point, which cannot guarantee a perfect fit of the broccoli crowns to that of 3D point cloud of broccoli canopy from the aerial photogrammetry. Additionally, the base of the model is suspended in the air rather than connected to the soil. For some inclined growing broccoli, it is noticeable that the stem does not bend towards the soil mound, which does not truly reflect the actual growth situation. In the future, matching and more delicate transformation should be considered with the broccoli crown surface of the aerial model, as well as the relative growth relationship between the stem and the soil, to achieve a more realistic display effect. Fifthly, simply expanding the database of destructive sampling is not sufficient to achieve the above goals, as it involves many complex structural transformations and combinations, thus a virtual broccoli model controlled by several parameters needs to be developed. Similar to the virtual maize model implemented in Cieslak et al. (2021), which can use parameters to change the details of plant structure. It was suggested to develop such a procedural model for broccoli heads.

Chapter 5

General discussion

The general discussion begins by summarizing the limitations of this doctoral study (Section 5.1). Following that, discussing the potential solutions for future research opportunities are outlined (Section 5.2).

5.1 Reseach limitations

Although the three-year experiment conducted in the broccoli field yielded promising outcomes, the study was also subject to certain limitations. This section provides a summary of three specific limitations encountered: the overreliance on broccoli-specific algorithms (Subsection 5.1.1), the lack of real farmland testing (Subsection 5.1.2), and inadequate contributions in fieldwork and data collection (Subsection 5.1.3).

5.1.1 Lacking universality

This study relied on certain assumptions about the characteristics of broccoli heads when developing the methods and algorithms. However, at present, there is insufficient evidence to support the claim that the proposed methods can achieve comparable results across various farmland or plant species.

In Chapter 2, an important assumption for the high-quality 3D reconstruction pipeline is that the broccoli head remains rigid and does not change when vertically rotated on the rotation platform. This assumption allows for the alignment and production of good quality 3D models through structure-from-motion (SfM)-multi-view stereo (MVS)

algorithms using images captured from different view angles. While this assumption holds for cauliflowers, sweet potatoes, and fruits like apples and oranges, whose bodies are almost rigid, it does not hold for plants with soft stems, leaves, and even roots. In such cases, the study proposed pipeline is not suitable, and a better solution is to utilize fixed-object methods (Figures 2.1a and 2.1b) to achieve improved results.

The postprocessing algorithms for broccoli crown segmentation and rotation in this study were also developed based on the characteristics of the broccoli head. Specifically, the algorithm combines the distinct color differences between the crown (dark green) and stem (light green) with a 2-class clustering algorithm to separate the two parts. Additionally, the flattened, mushroom-shaped form of the broccoli crown is utilized to determine the normal vector in the upward direction. These algorithms are specifically designed for broccoli head cases and may not apply to other vegetables and fruits.

In Chapter 3, the assumption of minimal positional variation in broccoli from seeding to heading is crucial. In our study, we employed the results of seeding position detection to refine the image processing region during the challenging heading stage. The concept of simplifying complex detection or segmentation tasks by incorporating data collected from an earlier period has also proven effective in other application scenarios. For instance, Mu et al. (2018) successfully obtained the convex hulls of peach tree crowns during the winter season, which were then utilized to guide the demanding crown segmentation task in summer. Similarly, Li et al. (2023) accurately located maize positions during the seeding stage, and these positions served as a guide for the challenging segmentation tasks when the maize canopy has severe leaf overlapping. However, this assumption does not apply to targets that frequently change due to wind, such as wheat and sorghum tassels. It is necessary and valuable to develop different processing algorithms to address such cases in the future.

We utilized EasyIDP as another valuable tool in this chapter, but it is reported to have a potential performance bottleneck. EasyIDP enables the linking of the same broccoli from low-quality geo-referenced map images to high-quality non-geo-referenced raw images, thereby greatly enhancing the accuracy of head segmentation. In our experiment

field, which spanned just one to 0.2 hectares, no significant performance bottlenecks were observed during the data processing. However, Dr. Wei Guo reported that when processing a large forest area of over 58 hectares with file sizes exceeding 23GB using EasyIDP, the process took over 3 hours. This performance bottleneck needs to be solved to facilitate the widespread application of this approach in large farmlands.

In Chapter 4, to fuse the high-quality head 3D model with the low-quality aerial canopy model, we assume that all broccoli heads share similar structures for the same cultivar in the same field condition. This allows us to create a representative template database through proper sampling. Each head can then be approximated after a simple template transformation, which involves scaling along the short and long axes and rotating to overlap. Regarding the specific data fusion operation of putting the transformed template back into the field canopy, we make further simplifying assumptions: 1) The shape of the broccoli head is approximately circular. Therefore, even if some heads are occluded by leaves in the aerial images, we can obtain the location of the center by fitting the exposed portion. This allows us to determine the X and Y coordinates for placing the transformed template back. 2) The broccoli heads tend to grow upright. As a result, the center obtained in assumption (1) represents the highest point of the broccoli crown in the vertical direction. This allows us to determine the Z coordinate for placing the template back. While our results demonstrate significant quality improvement for the broccoli heads after this data fusion process (refer to Figure 4.6), it is important to note that this is merely an approximation of size and better visualization of the calibrated morphological traits. It does not capture the details present in the actual field. Firstly, the shape of a broccoli head is a more complex fractal shape. Additionally, when the planting density is too high, some heads may be "squeezed" and grow at a skewed angle, which contradicts the simplifying assumptions mentioned earlier. Furthermore, this approach is only applicable to visible broccoli heads in the aerial images and does not work for heavily occluded or completely invisible heads. Especially, this method is not suitable for other challenging crops, such as potato and sweet potato, where the harvestable parts are underground. Exploring how to deal with such complex and invis-

ible conditions, like maize canopies and potato canopies, is a valuable avenue for future research.

5.1.2 Lacking real farmland practices

One unavoidable limitation of this study is that it has not been tested in actual farmlands, which are much larger and more complex than the current experimental field. One noticeable difference is the variation in canopy patterns caused by different plant densities. For instance, our experimental field strictly maintained a spacing of 35 cm in rows and 70 cm apart, with a smaller deviation of less than 3 cm (Fig. 5.1.a). In contrast, in the broccoli farmland we collaborated with in another project in Fukushima, the plants were spaced approximately 35 cm apart in rows, with a considerable variation greater than 7 cm (Fig. 5.1.b). This difference in density and deviation not only leads to distinct canopy patterns but also affects the direction and straightness of the ridges. The challenges posed by this difference in our proposed pipeline can be summarized in the following three points.

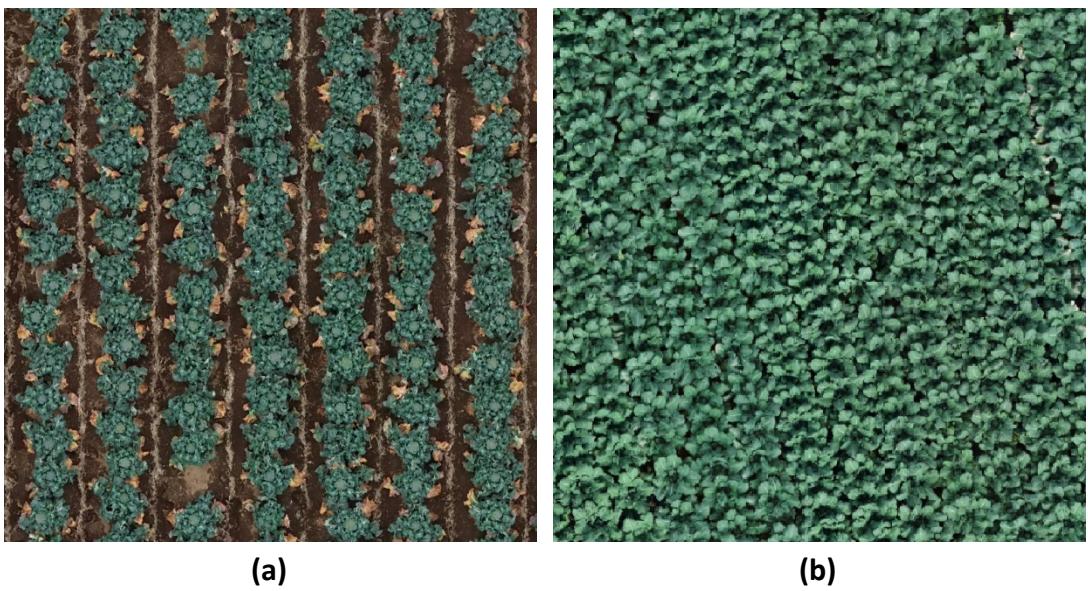


Figure 5.1: The differences between the experimental field and actual farmland. (a) is our experimental field in Tanashi of this study. (b) is a broccoli farmland in Fukushima from another project.

Firstly, in Fukushima, the broccoli tends to have smaller leaf angles compared to

the experimental fields. This leads to the broccoli heads being heavily covered by the shadows of leaves. These shadows create dark areas, making it challenging to visually identify the flower heads or leaves in the aerial images and also complicating the segmentation process. To effectively address this issue, the utilization of advanced sensors such as multispectral or light detection and ranging (LiDAR) is warranted, but at a much higher cost as compared to the conventional RGB camera. Therefore, there is a need to develop a more cost-effective LiDAR solution to cater to the future demands of farmland applications.

Secondly, the closer row spacing and higher deviation make it infeasible to use our previous linear-based ridge detection algorithm. Without it, assigning regular ID numbers to individual broccoli based on the order of the ridge becomes impractical. One potential approach is to abandon the use of sequential numbers for identification purposes along the ridges. Instead, numbering can be determined based on the sorted XY coordinate positions. This alternative not only reduces the computational burden for ridge detection but also guarantees the uniqueness of the identification numbers. However, this solution suffers from a disadvantage: it becomes challenging to conveniently associate the numbers with the plants in the field. Given that the actual measurements typically occur along the ridges, during recording, it becomes necessary to repeatedly verify the ID instead of simply incrementing it by 1 as in the previous solution. However, it is important to note that in the Fukushima farmland, conducting field measurements along ridges is not feasible. This issue will be discussed further in the upcoming point.

Thirdly, the high planting density poses difficulties in sampling and field measurement. In Figure 5.1.b, it is evident that the density of the actual farmland makes it difficult for people to enter without causing damage to the broccoli and soil. However, the farmland owner does not allow such damage, leaving only the broccolis along the plot boundaries available for measuring and sampling. Unfortunately, these boundary broccolis do not represent the same growing conditions as those in the plot center. This poses a critical problem for the current pipeline since it leads to a lack of accurate ground truth data for model calibration. To address this issue in the future, a closer collaboration

with more farmland owners is necessary. For one thing, a carefully designed sampling strategy needs to be implemented before transplanting the broccoli. For another, we can enlarge our broccoli database through such collaborations. We can potentially provide prediction services for farmlands where ground truth data is unavailable. This can be achieved by identifying similar environmental and cultivar patterns already present in our database.

5.1.3 Lacking full participation

This is a personal limitation rather than the study itself. One thing that needs to be pointed out is I did not contribute entirely to the full process of the broccoli project. The main focus of my doctoral research lies in the development of algorithms and engineering aspects. Many in-field contributions were made by several collaborators. For example, the plant materials used in this study were provided by another student, Nishida (2023), in the laboratory. To be more specific, the experimental design of obtaining variate broccoli sizes by different fertilizations and the measurement of field data were primarily carried out by Nishida. Meanwhile, the transplanting and daily field management of broccoli were conducted by technicians from ISAS. While the operation of the unmanned aerial vehicle (UAV) and data collection were mainly completed by Dr. Wei Guo. Additionally, I implemented some parts of the deep learning codes with the assistance of Tang Li from another laboratory. With his assistance, I was able to apply deep learning methods to process certain data and provide technical reserves for subsequent engineering applications. Thanks to their help, I was able to focus my efforts on the engineering aspects of indoor reconstruction systems (Chapter 2), the batch processing workflow of aerial data, and the accuracy improvement of spatiotemporal scale fusion (Chapter 3), as well as the calibration, transformation, and matching of the broccoli model across different scales (Chapter 4). Such research experience has provided me with the opportunity to engage in interdisciplinary collaboration and have successfully achieved breakthroughs in my area of study.

However, this may not be beneficial to my overall academic development in some

content. Due to not being involved throughout the experimental process and data collection, it limits my understanding of the physiological and structural aspects of broccoli in the study. It has made me aware of my shortcomings in research and motivated me to address them. For example, during the period of broccoli transplanting, I visited the site several times to observe and inquire about the relevant details of plot settings. When Nishida conducted field measurements and destructive sampling, I assisted with some measurements and destructive sampling work to gain insight into the details of field measurements. Furthermore, when Dr. Wei Guo carried out drone data collection, I took the opportunity to learn the basic procedures of operating a drone and understood the effects of setting overlap, shutter speed, and exposure on the quality of photographs. After this collaboration experience of my doctoral study, I have come to recognize the importance of actively participating in experiments and data collection for future research.

5.2 Future research prospects

In this thesis, several new approaches or pipelines were proposed to enhance the performance of 3D-based plant phenotyping. While our experimental field exhibits promising results, as discussed previously, this study still faces unresolved limitations at its current stage. In this section, I will explore the future research prospects that not only benefit this study but also inspire further advancements in smart agriculture.

5.2.1 Probabilistic of this study

Although our current research has some shortcomings and limitations, a set of tools and pipelines we developed in this study has been helping many subsequent research projects, not only limited to the agricultural field.

Chapter 2 presents the establishment of high-quality close-range crop modeling pipelines. We obtained 189 high-resolution 3D models of broccoli plants through it. As mentioned previously, this workflow can be readily extended to modeling more similar crops, including potatoes, sweet potatoes, and various other fruits. These high-quality

3D plant organ models can subsequently undergo phenotypic analysis. Moreover, our 3D database holds potential for application in the computer graphics (CG) or gaming industry, as it offers high-resolution assets for use in AAA games.

As a dependency tool to Chapter 3, we first developed an Intermediate Data Processing (IDP) tool, named EasyIDP (Wang et al., 2021a), to address the issue of low quality of drone-based 3D reconstruction outputs. This tool not only found good application in our third chapter but has also been applied and highly praised in other industries that require the use of drone measurement. For example, Prof. Benjamin Weinstein, from the University of Florida, used our tool for avian ecology surveys; Prof. Derek Young, from the University of California Davis, used our tool for optimizing forest species research; and Engineer K. Senda, from the Japanese company Yumeni, used our tool for improving urban greening monitoring. As of September 2023, our GitHub project community (<https://github.com/UTokyo-FieldPhenomics-Lab/EasyIDP>) has received 34 stars and 22 discussion threads. We have established a positive open-source community feedback mechanism, which is a level of achievement that other similar researchers have not reached. It is important to further improve and promote these types of open-source software and communities in order to ensure the reproducibility and contribution of our work.

In Chapter 3, an improved aerial survey pipeline for broccoli size was established, which obtained the specific size of each broccoli plant in the time series. This is difficult to achieve with traditional methods in acceptable labor costs. Subsequently, other members of the laboratory based on this technology, combined with meteorological data and growing degree models, successfully developed an individual-based broccoli growth prediction method (Wang et al., 2023b). It was also found that the deviation of only 2 days from the optimal date, The profit difference in 0.2 hectares of the experimental field could reach \$2000, indicating significant implications for reducing food loss and improving farmers' income (Wang et al., 2023b). Although it currently cannot be directly applied to densely planted farmlands, with the subsequent integration of other sensors to solve the broccoli segmentation problem, this process can be smoothly

applied.

Chapter 4 uses data fusion to achieve a better visualization effect, transforming numerical phenotype data (such as broccoli size) into a 3D intuitive model. This also serves as the technical basis for subsequent digital twins and smart agriculture, which is further discussed in Section 5.2.3. If augmented reality (AR) and virtual reality (VR) technologies can be combined in the future, there is hope to transform the laborious field patrol work into indoor activities.

5.2.2 Probabilistic of 3D reconstruction

The study primarily utilizes photo-based 3D reconstruction (SfM-MVS) to acquire 3D data for both close-range and aerial plants. In comparison to other reconstruction methods (Fig. 1.2), SfM-MVS offers the key advantage of relatively low sensor costs. Only common RGB cameras are necessary to yield satisfactory results for close-range or aerial reconstructions. However, the workflow of photo-based 3D reconstruction also presents several challenges for data management and processing algorithms that deserve to be considered for future research.

The first point is about data storage and management. Typically, modeling a scene requires a large number of images, which can occupy a minimum of 1GB of storage space. In our 3-year study, for both close-range broccoli head modeling (Chapter 2) and aerial survey (Chapter 3), the raw image data reached over 350GB (135GB in 2020, 30.3GB in 2021, and 101.7GB in 2022 for aerial survey; 87GB for close-range modeling). Furthermore, the data size from data preprocessing, reconstruction process, and the resulting model is essentially the same as the raw image data. This amount of data is only applicable to a 0.2-hectare experimental field, and the volume of data increases proportionally with the area. While it is theoretically possible to delete the original data and intermediate processes after obtaining the final modeling results to free up space, this is not feasible for our pipeline. We need to merge the final results with the raw images to achieve more accurate segmentation results. Additionally, these raw data hold significant value as digital resources for smart agriculture research. When properly

preserved, they can serve multiple research groups in answering complementary research questions over the years (Guo et al., 2021b). Therefore, compared to using local hard drives for data storage and archiving, the use of centralized cloud storage not only ensures better data security but also enables fast data sharing across institutions. This is a crucial aspect that smart agriculture must consider for future data storage and management (Guo et al., 2021b). Simultaneously, it is imperative to collaborate with various research institutions to establish a unified data archiving and metadata storage standard, similar to that of meteorological data, for subsequent cross-regional and long-term research tracking.

The second point concerns the processing algorithm of 3D reconstruction. It should be noted that this algorithm assumes the object to be rigid and consistent throughout all points of view (photos). This assumption is commonly valid for urban buildings, mine shafts, and cultural relics, rather than capturing dynamic objects (such as those affected by wind-induced movement), repeatable structures (such as similar plants in a field), and reflective objects (for instance, paddy fields). These cases commonly arise in agricultural applications, and additionally along with complex combinations of lighting, humidity (cloud and fog), presence of weeds, and occlusions. Several studies have been conducted to address the challenges posed by complex situations in algorithm adaptation. For instance, Zheng et al. (2015) focused on enhancing the reconstruction effect of dynamic components through algorithmic improvements. However, it is not feasible to manually develop special algorithms for all different kinds of natural conditions, a more efficient and robust approach should be found.

The potential application of deep learning technology to partially or entirely replace the SfM-MVS algorithm has also raised many research interests (Li et al., 2022c). This includes the substitution of camera pose estimation, also known as bundle adjustment (Wei et al., 2020), and point cloud densification (Yao et al., 2018). Furthermore, the neural radiance fields (NeRF) builds upon the outcomes of the SfM algorithm and employs deep learning networks to supplant subsequent MVS and 3D object generation processes (Mildenhall et al., 2022). For example, Jignasu et al. (2023) successfully

achieved excellent reconstruction outcomes for maize, underscoring the effectiveness of this approach.

There is now an application of NeRF based on the latest iPad’s LiDAR sensor for camera position estimation. In our preliminary experiments in early May 2023, especially for landscape plants with complex structures, better 3D models can be obtained than with photogrammetry. However, currently, the iPad needs to be manually moved to match the path guidance in the software. If this process can be interactively realized with a robotic arm, it also has broad potential applications.

In addition to enhancing the SfM-MVS workflow, there have been studies exploring the utilization of deep learning for direct 3D structure prediction. For instance, Vicini et al. (2022) employ direct estimation to determine the 3D shape, while Magistri et al. (2021) utilize a single image in conjunction with deep learning to directly render the 3D plant model. These studies, based on deep learning techniques, aim to overcome the technical limitations associated with conventional 3D reconstruction algorithms. It is important to verify and implement these approaches in the acquisition of 3D plant structures in the future.

The third point is about multi-sensor integration. The aforementioned attempts to reconstruct 3D structures from photos and optimize various algorithms are compromised solutions due to equipment costs. However, if we neglect equipment costs, the use of LiDAR is a more accurate method of obtaining spatial 3D structures. LiDAR has been extensively applied and tested in both close-range and outdoor studies (Jin et al., 2021b; Lin et al., 2022; Thapa et al., 2018). With the continuous development of autonomous driving and sensor technology, the price of LiDAR is continuously decreasing (Ackerman, 2016). It is expected that within the next 10 years, the price of LiDAR will approach that of advanced digital single-lens reflex (DSLR) cameras, potentially revolutionizing the acquisition of 3D structural data. In terms of applications in smart agriculture in the future, the fusion of LiDAR with other sensors further enhances the performance of high-throughput models (Li et al., 2023; Nguyen et al., 2023).

The final point is regarding the automation technologies utilized in aerial data collec-

tion. Presently, the quality of aerial image data still greatly depends on the expertise of device operators. Due to varying weather conditions during different survey times, there are currently no universal settings that are suitable for all conditions. Consequently, operators must manually adjust camera settings, such as exposures and shutter speed, according to the prevailing weather conditions. However, these adjustments heavily rely on experience, which encompasses not only drone operation but also a profound understanding of post-data processing. Often, the data collection work is outsourced to technicians or personnel from other institutions working remotely, and its quality frequently proves challenging to meet the high-precision analysis requirements. The implementation of this personnel training approach is not feasible. Instead, it is necessary to establish collaboration with drone companies to develop a camera mode, which can automatically adjust based on the surrounding environment for effective agricultural data collection. This approach will diminish the requirement for personnel expertise. This has already been demonstrated in previous technological advancements. Initially, manual control the drone flight and camera photography was essential during aerial surveys around the 2010s. However, the acquired images often failed to meet the required level of overlap, resulting in reconstruction failure. Consequently, when I attended the internship at the Chinese Academy of Forestry, we developed the software for flight route control with required overlaps (Wang et al., 2017, www.uav-hirap.org). These features allow for automatic flight route planning on a computer and transmitting to the drone. This technology has since been completely superseded by the built-in flight planning function of drones within three years. Hence, it can be predicted that this issue of camera parameter settings will also be integrated into drones shortly to ensure the stability of data quality obtained by different operators.

5.2.3 Probabilistic of digital twin

The advent of digital twin technology points to a very different future for smart agriculture and plant phenotyping, the researchers can operate directly on the virtual plants and preview its impact on the plant immediately (Verdouw et al., 2021). The

fundamental in implementing such technology is to present the 3D crop model (virtual plant) accurately in the computer first, and then implement architecture fine-tuning and phenotyping applications on them. This study is the first to explore the possibility of presenting the virtual plants from actual fields accurately.

In Chapter 4, the data fusion between close-range and aerial surveys showed a good result in the visualization hidden part of the broccoli head, although based on statistical regression for visualization rather than actual status. As discussed in Chapter 4, for different cultivars of broccoli, more broccoli heads should be collected to provide a sufficient template database to meet more detailed transformation needs. However, the proposed method is facing challenges when dealing with crops with complex structures, such as maize or soybean canopies. On one hand, the severity of occlusion makes it very difficult to distinguish individual maize or soybean; on the other hand, due to the multilateral structure of crops, it is difficult to collect enough samples for template transformation. Thus, there is a demand to change the virtual plant architecture with lower cost

Currently, there are three main methods of acquiring virtual plants that can be architecture fine-tuned. 1) By template stitching. Chang et al. (2022) obtained the database of wheat organ 3D models at different growth stages by destructive sampling and manual measurement of parameters. Then combine those organs to one wheat by random picking and placing with random position angles within the range obtained by real measurements. Wen et al. (2021) applied the similar idea for maize. 2) By static deformation. The 3D static model of whole maize (in point cloud format) was obtained by 3D reconstruction. The maize model was then automatically segmented into organs, and these segmented parts were transformed by applying geometric transformations to implement changes in architecture Liu et al. (2021b). 3) By parametric approximation. The L-system was used to construct a parameter-adjustable virtual crop model and then adjusted the parameters to approximate the real crop by three-view photos Cieslak et al. (2021).

In this condition, the advent of procedural modeling points to very different solutions

that use several parameters to control the structure of 3D models. For example, The L-system was used to construct virtual crop models that adjusted the parameters to approximate the real crop by three-view photos (Cieslak et al., 2021). Although the model obtained using the above method is only structural, it is not realistic enough in terms of detail and appearance. Mikami et al. (2022) used photo-quality texture mapping on the plant 3D models with relatively simple geometries and finally obtained very realistic 3D models in texture. In our preliminary experiment, a similar corn model using parameter controls has already been achieved in Blender, but due to the complexity of the research, the ability to repair the corn canopy has not yet been realized. In subsequent studies, we hope to develop a parameter-controlled broccoli model and complete the repair of the corn canopy.

Furthermore, several advanced techniques can be applied to the abovementioned virtual 3D plant models. 1) The ray-tracing technologies can be applied to simulate lights within the canopy. For example, analyzing the contribution of foliar and non-foliar tissues to canopy photosynthesis under different canopy structures (Chang et al., 2022), or the effects of different spacing (He et al., 2021) and leaf angles (Liu et al., 2021b) on the light use efficiency. 2) The point cloud simulation technology can be used to simulate point clouds from the virtual canopy. For the canopy traits which are easy to get from virtual plants but hard from the real field, the corresponding models between features from the simulated point cloud and traits from the virtual canopy can be trained and then applied to the field-scanned point cloud data to inverse the actual data (Liu et al., 2017a). 3) The CG rendering techniques can decrease the workload of data annotating in the phenotyping data processing by deep learning. The rendering engine can yield CG photos that are very close to the real world. At the same time, by adjusting the model texture to solid pure color, the corresponding labeled information for both CG photos (Mikami et al., 2022) and point clouds (Chaudhury et al., 2020) can be generated in batch from random virtual canopies.

5.2.4 Probabilistic of Foundation AI models

This study incorporated machine learning (AutoML) and deep learning (Yolo, BiSeNet) techniques to effectively reduce labor costs associated with occlusion calibration and broccoli head segmentation. However, it is important to note that the robustness and universality of this study are confined to broccoli heads within the experimental field, as discussed in Section 5.1. To make further advancements, the emergence of Foundation (large) AI models presents an inspiring opportunity for future research in the field of smart agriculture.

In the field of natural language processing (NLP) tasks, such as translation, researchers initially focused on developing algorithms that followed grammar rules for translating between languages (McDonald, 1994). This approach bears resemblance to the manual rule-writing in smart agriculture a few years ago, where image-processing rules (e.g., color space threshold) were written to detect and segment regions of interest (ROI) from images (Kapur et al., 1985; Nobis and Hunziker, 2005; Wang, 2019; ?). However, this method is time-consuming and only applicable under specific and limited conditions, with limited ability to be readily applied to other conditions.

Then, machine learning algorithms shifted their focus from developing rules to collecting as much data as possible. By inputting paired translations of the same book or article, The machine itself could identify patterns between different languages. This technique quickly led to a revolution in NLP, significantly improving translation quality and making it practical for everyday use (e.g., Google Translate, Bing Translate, and more recently, DeepL). This is similar to what deep learning has brought to smart agriculture, where the emphasis has shifted from developing image-processing algorithms to maximizing the amount of annotated training data (Bauer et al., 2019; Blok et al., 2021b; Zhou et al., 2020a). Chapter 3 contributed to integrating the characteristics of agriculture to reduce the workload of data annotation. This is now at the cutting edge of research in smart agriculture, but not necessarily in NLP.

The continuous development of deep learning, particularly the introduction of trans-

former architecture in 2017, has led to a significant increase in the parameter count of deep learning models. For instance, the BERT model (Devlin et al., 2018) consists of over 30 billion parameters, whereas the GPT-3 model (Brown et al., 2020) exceeds 175 billion parameters. The combination of massive training data and massive model parameters has led to a qualitative leap in the performance of models. In the past, algorithms for various subtasks in NLP differed greatly, but after the emergence of foundation models, machine translation, intelligent question answering, and other fields have almost been dominated by foundation models (Bommasani et al., 2021). With the emergence of commercial products like ChatGPT, this technology has been rapidly progressing from the academic field to the public domain, thereby driving the recent AI fever all around the world. In addition to its remarkable achievements in the field of NLP, large models have also led to technological innovations in computer vision and weather forecasting. For instance, Facebook’s Segment Anything (Kirillov et al., 2023) has demonstrated the possibility of training a universal segmentation model by providing a huge amount of segmentation annotation data. Moreover, Huawei’s Pangu-Weather foundation model has achieved a speed improvement of over 10,000 times compared to traditional forecast methods, enabling a single graphics card to complete a 24-hour global weather forecast within a few minutes, with greater accuracy in predicting typhoons (Bi et al., 2023). The outstanding performance of these large models in these fields has paved the way for further research in smart agriculture.

Agriculture is also a highly complex system, and it may be possible to obtain a general model similar to Pangu-Weather by collecting a sufficient amount of data and conducting training. However, there is still a long way to go for large-scale agricultural models. The most critical issue is the lack of high-quality agricultural datasets, compared to NLP, computer vision, and meteorology studies. On the one hand, there are not many agricultural datasets available; On the other hand, the majority of datasets consist of image data often lacking the additional information required in agriculture such as genotype, soil type, and management methods. Furthermore, different datasets vary greatly in terms of recorded phenotypic data types and formats, making it very difficult

to merge and analyze the data. Therefore, for the future development of large-scale agricultural models, as emphasized in Section 5.2.1, strengthening cooperation between institutions and improving dataset management are urgent issues to be addressed.

Chapter 6

General conclusion

This is the first study that combined the advances of aerial (high throughput) and close-range (high quality) surveys in plant phenotyping applications to field-grown crops. This Ph.D. thesis aimed to improve the performance of 3D-based plant phenotyping on field-grown broccoli as a representative of row-planted crops having harvestable organs on the top of canopy, using only low-cost red-green-blue (RGB) cameras. Using unmanned aerial vehicle (UAV) for aerial surveys and photogrammetry allows for the efficient acquisition of 2D field maps and 3D models of the crop canopy for the entire farmland. However, due to limitations in survey efficiency and wind blurring caused by propellers, the UAV cannot fly too close to plants, resulting in inadequate resolution and quality for directly analyzing the broccoli heads at the organ level. This thesis attempted to fuse the close-range and aerial 3D phenotype data, as well as some latest machine learning and deep learning techniques, to accurately and efficiently obtain the position and size of broccoli heads in the field. Furthermore, this thesis also provided a better 3D virtual visualization for those broccoli heads in the field and builds a foundation for digital twins and virtual farmland for smart agriculture.

In the second chapter, we have implemented an almost fully automated plant phenotyping pipeline based on the 3D reconstruction of broccoli heads. The described workflow has been designed to minimize the effects of limited perspective on the 3D model quality and completion. By using the dual-rotation of the object and the dual deep learning segmentation, our proposed workflow performed well on over 180 broccoli

heads during the flowering season. The accuracy of pipeline calculated head size traits was validated by simple manual measurements using standard agricultural practices. The results suggest that our pipeline offers a great opportunity for high-throughput 3D phenotyping applications on the solid and enclosed plant organs (e.g. oranges, potatoes, cauliflowers, and sweet potatoes), in which size is directly related to harvest timing and profitability.

The third chapter aimed to adapt the UAV-based aerial sensing technology to the monitoring of spacial growth variation of field-grown crops. By using aerial photogrammetry and machine learning / deep learning (ML/DL), we developed a system for estimating and predicting the head size of whole broccoli with high accuracy. This UAV-based prediction system is based on several technical improvements and requires minimal labor and computational costs. Therefore, it could be applied to support broccoli farming, and with modifications, to a variety of similar vegetables (i.e., cabbage, cauliflower, artichoke, and lettuce). Because our developed pipeline uses a simple sensor, not a complex integration of multiple sensors, it would be more applicable and offers user-friendliness for economically and socially disadvantaged rural regions, and it has the potential to be widely adopted by vegetable farmers worldwide.

The work in the fourth chapter aimed to develop a 3D virtual visualization technique that can fuse high-quality head 3D models (from Chapter 2) into low-quality full-field canopy models (from Chapter 3). By using piecewise affine transformation, auto machine learning (AutoML) calibration model, and template matching and transformation, a 3D virtual visualization system was developed to visualize the broccoli head sizes at their field positions. The statistical analysis supports the improvements of the proposed method in geographical locations and broccoli morphological traits. The proposed 3D visualization system offers a great opportunity for virtual farmland and digital twin technology, which can provide a more intuitive feeling of growth status compared to numeric statistical values.

Overall, the results of these three research chapters showed that the proposed pipelines including active learning, deep learning, backward projection, auto machine

learning, template matching, and data fusion, led to improved performance in the tested broccoli fields from 2020 to 2022. The results and statistical analysis concluded that the research objectives have been achieved and all the source codes are published on GitHub for replication and any usage purposes, we can conclude that the research has made a positive contribution to the 3D-based plant phenotyping and precision agriculture for broccoli farmlands.

Acknowledgements

I would like to express my sincere gratitude to my thesis supervisor, Prof. Yoichiro Kato of the Graduate School of Agricultural and Life Sciences, for his invaluable advice, support, guidance, and encouragement throughout my doctoral research journey. He has constantly challenged and guided me to push the boundaries of my research and finally bring me to the point of accomplishing my degree.

I would also like to extend my thanks to the members of my thesis committee for their valuable feedback and insights. Their contributions have significantly improved the quality and relevance of my research.

Furthermore, I would like to thank Dr. Wei Guo. He not only served almost like a role of an advisor who greatly helped with my doctoral studies but also as a close friend who gave me many valuable suggestions for both academic and daily life in Japan. Also thank Dr. Yuji Yamasaki and Dr. Jimmy Burridge from The University of Tokyo, Dr. Yuya Fukano from Chiba University, Dr. Wenli Zhang from Beijing University of Technology, Dr. Yun Shi and Dr. Yulin Duan from Chinese Academy of Agricultural, and Dr. Feng Wang from Chinese Academy of Forestry, for providing me with advanced resources and excellent academic support.

I would also like to thank Naomi Morinaga, Kozue Wada, Kunihiro Kodama, Masanori Ishii, and all the technical support staff of Institute for Sustainable Agro-ecosystem Services (ISAS) for providing me with a conducive research environment and excellent support for fieldwork and document-related work. It enabled me to freely explore and develop my research interests.

I am also extremely grateful to my lab mates. Tang Li is a good friend who helped me a lot with deep learning and life in Japan, and Nishida Erika contributed greatly to

the broccoli fieldwork. I also thank my tutors, Thum Chun Hau and Jaesuk IM, for their kind help, and all my other friends for being the pillars of my life.

I would also like to express my gratitude to my family for their constant support, encouragement, and love, especially my girlfriend who has never appeared in my actual life during my doctoral studies and has allowed me to fully focus on my academic work.

Thank you all for being an integral part of my doctor journey and for helping me achieve this milestone.

References

- Ackerman, E., 2016. Lidar that will make self-driving cars affordable [News]. IEEE Spectrum 53, 14–14. doi:[10.1109/MSPEC.2016.7572525](https://doi.org/10.1109/MSPEC.2016.7572525).
- Ameen, W., Al-Ahmari, A., Hammad Mian, S., 2018. Evaluation of Handheld Scanners for Automotive Applications. Applied Sciences 8, 217. doi:[10.3390/app8020217](https://doi.org/10.3390/app8020217).
- Apelt, F., Breuer, D., Nikoloski, Z., Stitt, M., Kragler, F., 2015. Phytotyping ^{4D}: A light-field imaging system for non-invasive and accurate monitoring of spatio-temporal plant growth. The Plant Journal 82, 693–706. doi:[10.1111/tpj.12833](https://doi.org/10.1111/tpj.12833).
- Araus, J.L., Cairns, J.E., 2014. Field high-throughput phenotyping: The new crop breeding frontier. Trends in Plant Science 19, 52–61. doi:[10.1016/j.tplants.2013.09.008](https://doi.org/10.1016/j.tplants.2013.09.008).
- Bai, G., Ge, Y., Scoby, D., Leavitt, B., Stoerger, V., Kirchgessner, N., Irmak, S., Graef, G., Schnable, J., Awada, T., 2019. NU-Spidercam: A large-scale, cable-driven, integrated sensing and robotic system for advanced phenotyping, remote sensing, and agronomic research. Computers and Electronics in Agriculture 160, 71–81. doi:[10.1016/j.compag.2019.03.009](https://doi.org/10.1016/j.compag.2019.03.009).
- Bartol, K., Bojanic, D., Petkovic, T., Pribanic, T., 2021. A Review of Body Measurement Using 3D Scanning. IEEE Access 9, 67281–67301. doi:[10.1109/ACCESS.2021.3076595](https://doi.org/10.1109/ACCESS.2021.3076595).
- Bauer, A., Bostrom, A.G., Ball, J., Applegate, C., Cheng, T., Laycock, S., Rojas, S.M., Kirwan, J., Zhou, J., 2019. Combining computer vision and deep learning to enable ultra-scale aerial phenotyping and precision agriculture: A case study of lettuce production. Horticulture Research 6, 70. doi:[10.1038/s41438-019-0151-5](https://doi.org/10.1038/s41438-019-0151-5).
- Beck, M., Liu, C.Y., Bidinosti, C., Henry, C., Godee, C., Ajmani, M., 2020a. Weed seedling images of species common to Manitoba, Canada. doi:[10.5061/DRYAD.GTHT76HHZ](https://doi.org/10.5061/DRYAD.GTHT76HHZ).
- Beck, M.A., Liu, C.Y., Bidinosti, C.P., Henry, C.J., Godee, C.M., Ajmani, M., 2020b. An embedded system for the automated generation of labeled plant images to enable machine learning applications in agriculture. PLOS ONE 15, e0243923. doi:[10.1371/journal.pone.0243923](https://doi.org/10.1371/journal.pone.0243923).
- Bender, A., 2019. Ladybird Cobbitty 2017 Brassica Dataset. doi:[10.25910/5C941D0C8BCCB](https://doi.org/10.25910/5C941D0C8BCCB).

- Bender, A., Whelan, B., Sukkarieh, S., 2020. A high-resolution, multimodal data set for agricultural robotics: A *Ladybird* 's-eye view of Brassica. *Journal of Field Robotics* 37, 73–96. doi:[10.1002/rob.21877](https://doi.org/10.1002/rob.21877).
- Bi, K., Xie, L., Zhang, H., Chen, X., Gu, X., Tian, Q., 2023. Accurate medium-range global weather forecasting with 3D neural networks. *Nature* 619, 533–538. doi:[10.1038/s41586-023-06185-3](https://doi.org/10.1038/s41586-023-06185-3).
- Blok, P.M., Barth, R., van den Berg, W., 2016. Machine vision for a selective broccoli harvesting robot. *IFAC-PapersOnLine* 49, 66–71. doi:[10.1016/j.ifacol.2016.10.013](https://doi.org/10.1016/j.ifacol.2016.10.013).
- Blok, P.M., Evert, F.K., Tielen, A.P.M., Henten, E.J., Kootstra, G., 2021a. The effect of data augmentation and network simplification on the image-based detection of broccoli heads with Mask R-CNN. *Journal of Field Robotics* 38, 85–104. doi:[10.1002/rob.21975](https://doi.org/10.1002/rob.21975).
- Blok, P.M., van Henten, E.J., van Evert, F.K., Kootstra, G., 2021b. Image-based size estimation of broccoli heads under varying degrees of occlusion. *Biosystems Engineering* 208, 213–233. doi:[10.1016/j.biosystemseng.2021.06.001](https://doi.org/10.1016/j.biosystemseng.2021.06.001).
- Bommasani, R., Hudson, D.A., Adeli, E., Altman, R., Arora, S., von Arx, S., Bernstein, M.S., Bohg, J., Bosselut, A., Brunskill, E., Brynjolfsson, E., Buch, S., Card, D., Castellon, R., Chatterji, N., Chen, A., Creel, K., Davis, J.Q., Demszky, D., Donahue, C., Doumbouya, M., Durmus, E., Ermon, S., Etchemendy, J., Ethayarajh, K., Fei-Fei, L., Finn, C., Gale, T., Gillespie, L., Goel, K., Goodman, N., Grossman, S., Guha, N., Hashimoto, T., Henderson, P., Hewitt, J., Ho, D.E., Hong, J., Hsu, K., Huang, J., Icard, T., Jain, S., Jurafsky, D., Kalluri, P., Karamcheti, S., Keeling, G., Khani, F., Khattab, O., Koh, P.W., Krass, M., Krishna, R., Kuditipudi, R., Kumar, A., Ladhak, F., Lee, M., Lee, T., Leskovec, J., Levent, I., Li, X.L., Li, X., Ma, T., Malik, A., Manning, C.D., Mirchandani, S., Mitchell, E., Munyikwa, Z., Nair, S., Narayan, A., Narayanan, D., Newman, B., Nie, A., Niebles, J.C., Nilforoshan, H., Nyarko, J., Ogut, G., Orr, L., Papadimitriou, I., Park, J.S., Piech, C., Portelance, E., Potts, C., Raghunathan, A., Reich, R., Ren, H., Rong, F., Roohani, Y., Ruiz, C., Ryan, J., Ré, C., Sadigh, D., Sagawa, S., Santhanam, K., Shih, A., Srinivasan, K., Tamkin, A., Taori, R., Thomas, A.W., Tramèr, F., Wang, R.E., Wang, W., Wu, B., Wu, J., Wu, Y., Xie, S.M., Yasunaga, M., You, J., Zaharia, M., Zhang, M., Zhang, T., Zhang, X., Zhang, Y., Zheng, L., Zhou, K., Liang, P., 2021. On the Opportunities and Risks of Foundation Models. arXiv doi:[10.48550/ARXIV.2108.07258](https://doi.org/10.48550/ARXIV.2108.07258).
- Boogaard, F.P., Rongen, K.S.A.H., Kootstra, G.W., 2020. Robust node detection and tracking in fruit-vegetable crops using deep learning and multi-view imaging. *Biosystems Engineering* 192, 117–132. doi:[10.1016/j.biosystemseng.2020.01.023](https://doi.org/10.1016/j.biosystemseng.2020.01.023).
- Boogaard, F.P., van Henten, E.J., Kootstra, G., 2021. Boosting plant-part segmentation of cucumber plants by enriching incomplete 3D point clouds with spectral data. *Biosystems Engineering* 211, 167–182. doi:[10.1016/j.biosystemseng.2021.09.004](https://doi.org/10.1016/j.biosystemseng.2021.09.004).
- Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J.D., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G.,

- Henighan, T., Child, R., Ramesh, A., Ziegler, D., Wu, J., Winter, C., Hesse, C., Chen, M., Sigler, E., Litwin, M., Gray, S., Chess, B., Clark, J., Berner, C., McCandlish, S., Radford, A., Sutskever, I., Amodei, D., 2020. Language models are few-shot learners, in: Larochelle, H., Ranzato, M., Hadsell, R., Balcan, M., Lin, H. (Eds.), *Advances in Neural Information Processing Systems*, Curran Associates, Inc.. pp. 1877–1901.
- Cao, W., Zhou, J., Yuan, Y., Ye, H., Nguyen, H.T., Chen, J., Zhou, J., 2019. Quantifying Variation in Soybean Due to Flood Using a Low-Cost 3D Imaging System. *Sensors* 19, 2682. doi:[10.3390/s19122682](https://doi.org/10.3390/s19122682).
- Chang, T.G., Shi, Z., Zhao, H., Song, Q., He, Z., Van Rie, J., Den Boer, B., Galle, A., Zhu, X.G., 2022. 3dCAP-Wheat: An Open-Source Comprehensive Computational Framework Precisely Quantifies Wheat Foliar, Nonfoliar, and Canopy Photosynthesis. *Plant Phenomics* 2022, 1–19. doi:[10.34133/2022/9758148](https://doi.org/10.34133/2022/9758148).
- Chaudhury, A., Barron, J.L., 2018. Machine Vision System for 3D Plant Phenotyping. *IEEE/ACM Transactions on Computational Biology and Bioinformatics* 16, 2009–2022. doi:[10.1109/TCBB.2018.2824814](https://doi.org/10.1109/TCBB.2018.2824814).
- Chaudhury, A., Boudon, F., Godin, C., 2020. 3D plant phenotyping: All you need is labelled point cloud data, in: CVPPP-ECCV 2020 - Workshop on Computer Vision Problems in Plant Phenotyping, Glasgow, United Kingdom. pp. 1–17.
- Chen, C.J., Zhang, Z., 2020. GRID: A Python Package for Field Plot Phenotyping Using Aerial Images. *Remote Sensing* 12, 1697. doi:[10.3390/rs12111697](https://doi.org/10.3390/rs12111697).
- Cheng, H.K., Chung, J., Tai, Y.W., Tang, C.K., 2020. CascadePSP: Toward class-agnostic and very high-resolution segmentation via global and local refinement, in: 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA. pp. 8887–8896. doi:[10.1109/CVPR42600.2020.00891](https://doi.org/10.1109/CVPR42600.2020.00891).
- Cieslak, M., Khan, N., Ferraro, P., Soolanayakanahally, R., Robinson, S.J., Parkin, I., McQuillan, I., Prusinkiewicz, P., 2021. L-system models for image-based phenomics: Case studies of maize and canola. *in silico Plants* 4, diab039. doi:[10.1093/insili copants/diab039](https://doi.org/10.1093/insili/copants/diab039).
- Cook, D.D., de la Chapelle, W., Lin, T.C., Lee, S.Y., Sun, W., Robertson, D.J., 2019. DARLING: A device for assessing resistance to lodging in grain crops. *Plant Methods* 15, 1–8. doi:[10.1186/s13007-019-0488-7](https://doi.org/10.1186/s13007-019-0488-7).
- Das Choudhury, S., Bashyam, S., Qiu, Y., Samal, A., Awada, T., 2018. Holistic and component plant phenotyping using temporal image sequence. *Plant Methods* 14, 35. doi:[10.1186/s13007-018-0303-x](https://doi.org/10.1186/s13007-018-0303-x).
- David, E., Serouart, M., Smith, D., Madec, S., Velumani, K., Liu, S., Wang, X., Espinosa, F.P., Shafiee, S., Tahir, I.S.A., Tsujimoto, H., Nasuda, S., Zheng, B., Kichgessner, N., Aasen, H., Hund, A., Sadeghi-Tehran, P., Nagasawa, K., Ishikawa, G., Dandrifosse, S., Carlier, A., Mercatoris, B., Kuroki, K., Wang, H., Ishii, M., Badhon, M.A., Pozniak, C., LeBauer, D.S., Lilimo, M., Poland, J., Chapman, S., de Solan, B., Baret, F., Stavness, I., Guo, W., 2021. Global Wheat Head Dataset 2021: More diversity to improve the benchmarking of wheat head localization methods. arXiv [arXiv:2105.07660](https://arxiv.org/abs/2105.07660).

- Desai, S.V., Balasubramanian, V.N., Fukatsu, T., Ninomiya, S., Guo, W., 2019. Automatic estimation of heading date of paddy rice using deep learning. *Plant Methods* 15, 76. doi:[10.1186/s13007-019-0457-1](https://doi.org/10.1186/s13007-019-0457-1).
- Devlin, J., Chang, M.W., Lee, K., Toutanova, K., 2018. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. arXiv doi:[10.48550/ARXIV.1810.04805](https://doi.org/10.48550/ARXIV.1810.04805).
- Drofova, I., Guo, W., Wang, H., Adamek, M., 2023. Use of scanning devices for object 3D reconstruction by photogrammetry and visualization in virtual reality. *Bulletin of Electrical Engineering and Informatics* 12, 868–881. doi:[10.11591/eei.v12i2.4584](https://doi.org/10.11591/eei.v12i2.4584).
- Du, J., Fan, J., Wang, C., Lu, X., Zhang, Y., Wen, W., Liao, S., Yang, X., Guo, X., Zhao, C., 2021. Greenhouse-based vegetable high-throughput phenotyping platform and trait evaluation for large-scale lettuces. *Computers and Electronics in Agriculture* 186, 106193. doi:[10.1016/j.compag.2021.106193](https://doi.org/10.1016/j.compag.2021.106193).
- Du, R., Ma, Z., Xie, P., He, Y., Cen, H., 2023. PST: Plant segmentation transformer for 3D point clouds of rapeseed plants at the podding stage. *ISPRS Journal of Photogrammetry and Remote Sensing* 195, 380–392. doi:[10.1016/j.isprsjprs.2022.11.022](https://doi.org/10.1016/j.isprsjprs.2022.11.022).
- Duan, T., Zheng, B., Guo, W., Ninomiya, S., Guo, Y., Chapman, S.C., 2017. Comparison of ground cover estimates from experiment plots in cotton, sorghum and sugarcane based on images and ortho-mosaics captured by UAV. *Functional Plant Biology* 44, 169. doi:[10.1071/FP16123](https://doi.org/10.1071/FP16123).
- Duc, N., Vu, N., Cong, T., Thi, T., 2015. Combined Plane and Point Registration of Sequential Depth Images for Indoor Localization, in: Third International Conference on Advances in Computing, Electronics and Electrical Technology - CEET 2015, Institute of Research Engineers and Doctors. pp. 136–140. doi:[10.15224/978-1-63248-056-9-85](https://doi.org/10.15224/978-1-63248-056-9-85).
- Dutagaci, H., 2023. Using t-distributed stochastic neighbor embedding for visualization and segmentation of 3D point clouds of plants. arXiv [arXiv:2302.03442](https://arxiv.org/abs/2302.03442).
- Dutagaci, H., Rasti, P., Galopin, G., Rousseau, D., 2020. ROSE-X: An annotated data set for evaluation of 3D plant organ segmentation methods. *Plant Methods* 16, 28. doi:[10.1186/s13007-020-00573-w](https://doi.org/10.1186/s13007-020-00573-w).
- Faye, E., Rebaudo, F., Yáñez-Cajo, D., Cauvy-Fraunié, S., Dangles, O., 2016. A toolbox for studying thermal heterogeneity across spatial scales: From unmanned aerial vehicle imagery to landscape metrics. *Methods in Ecology and Evolution* 7, 437–446. doi:[10.1111/2041-210X.12488](https://doi.org/10.1111/2041-210X.12488).
- Feldman, A., Wang, H., Fukano, Y., Kato, Y., Ninomiya, S., Guo, W., 2021. EasyDCP: An affordable, high-throughput tool to measure plant phenotypic traits in 3D. *Methods in Ecology and Evolution* 12, 1679–1686. doi:[10.1111/2041-210X.13645](https://doi.org/10.1111/2041-210X.13645).

- Feng, L., Chen, S., Zhang, C., Zhang, Y., He, Y., 2021. A comprehensive review on recent applications of unmanned aerial vehicle remote sensing with various sensors for high-throughput plant phenotyping. *Computers and Electronics in Agriculture* 182, 106033. doi:[10.1016/j.compag.2021.106033](https://doi.org/10.1016/j.compag.2021.106033).
- Feurer, M., Eggensperger, K., Falkner, S., Lindauer, M., Hutter, F., 2020. Auto-sklearn 2.0: Hands-free AutoML via meta-learning. arXiv [arXiv:2007.04074](https://arxiv.org/abs/2007.04074).
- Feurer, M., Klein, A., Eggensperger, K., Springenberg, J., Blum, M., Hutter, F., 2015. Efficient and robust automated machine learning, in: *Advances in Neural Information Processing Systems* 28 (2015), pp. 2962–2970.
- Fiorani, F., Schurr, U., 2013. Future Scenarios for Plant Phenotyping. *Annual Review of Plant Biology* 64, 267–291. doi:[10.1146/annurev-arplant-050312-120137](https://doi.org/10.1146/annurev-arplant-050312-120137).
- Gallardo, R.K., Sauer, J., 2018. Adoption of Labor-Saving Technologies in Agriculture. *Annual Review of Resource Economics* 10, 185–206. doi:[10.1146/annurev-resource-100517-023018](https://doi.org/10.1146/annurev-resource-100517-023018).
- Gao, T., Zhu, F., Paul, P., Sandhu, J., Doku, H.A., Sun, J., Pan, Y., Staswick, P., Walia, H., Yu, H., 2021. Novel 3D Imaging Systems for High-Throughput Phenotyping of Plants. *Remote Sensing* 13, 2113. doi:[10.3390/rs13112113](https://doi.org/10.3390/rs13112113).
- García-Manso, A., Gallardo-Caballero, R., García-Orellana, C.J., González-Velasco, H.M., Macías-Macías, M., 2021. Towards selective and automatic harvesting of broccoli for agri-food industry. *Computers and Electronics in Agriculture* 188, 106263. doi:[10.1016/j.compag.2021.106263](https://doi.org/10.1016/j.compag.2021.106263).
- Garlando, U., Bar-On, L., Avni, A., Shacham-Diamond, Y., Demarchi, D., 2020. Plants and Environmental Sensors for Smart Agriculture, an Overview, in: *2020 IEEE SENSORS*, IEEE, Rotterdam, Netherlands. pp. 1–4. doi:[10.1109/SENSORS47125.2020.9278748](https://doi.org/10.1109/SENSORS47125.2020.9278748).
- Garrido, M., Paraforos, D., Reiser, D., Vázquez Arellano, M., Griepentrog, H., Valero, C., 2015. 3D Maize Plant Reconstruction Based on Georeferenced Overlapping LiDAR Point Clouds. *Remote Sensing* 7, 17077–17096. doi:[10.3390/rs71215870](https://doi.org/10.3390/rs71215870).
- Ge, L., Yang, Z., Sun, Z., Zhang, G., Zhang, M., Zhang, K., Zhang, C., Tan, Y., Li, W., 2019. A Method for Broccoli Seedling Recognition in Natural Environment Based on Binocular Stereo Vision and Gaussian Mixture Model. *Sensors* 19, 1132. doi:[10.3390/s19051132](https://doi.org/10.3390/s19051132).
- Gerland, P., Raftery, A.E., Ševčíková, H., Li, N., Gu, D., Spoorenberg, T., Alkema, L., Fosdick, B.K., Chunn, J., Lalic, N., Bay, G., Buettner, T., Heilig, G.K., Wilmoth, J., 2014. World population stabilization unlikely this century. *Science* 346, 234–237. doi:[10.1126/science.1257469](https://doi.org/10.1126/science.1257469).
- Ghanem, M.E., Marrou, H., Sinclair, T.R., 2015. Physiological phenotyping of plants for crop improvement. *Trends in Plant Science* 20, 139–144. doi:[10.1016/j.tplants.2014.11.006](https://doi.org/10.1016/j.tplants.2014.11.006).

- Ghosal, S., Zheng, B., Chapman, S.C., Potgieter, A.B., Jordan, D.R., Wang, X., Singh, A.K., Singh, A., Hirafuji, M., Ninomiya, S., Ganapathysubramanian, B., Sarkar, S., Guo, W., 2019. A weakly supervised deep learning framework for sorghum head detection and counting. *Plant Phenomics* 2019, 1–14. doi:[10.34133/2019/1525874](https://doi.org/10.34133/2019/1525874).
- Guo, A., Huang, W., Dong, Y., Ye, H., Ma, H., Liu, B., Wu, W., Ren, Y., Ruan, C., Geng, Y., 2021a. Wheat Yellow Rust Detection Using UAV-Based Hyperspectral Technology. *Remote Sensing* 13, 123. doi:[10.3390/rs13010123](https://doi.org/10.3390/rs13010123).
- Guo, W., Carroll, M.E., Singh, A., Swetnam, T.L., Merchant, N., Sarkar, S., Singh, A.K., Ganapathysubramanian, B., 2021b. UAS-Based Plant Phenotyping for Research and Breeding Applications. *Plant Phenomics* 2021, 1–21. doi:[10.34133/2021/9840192](https://doi.org/10.34133/2021/9840192).
- Guo, W., Fukano, Y., Noshita, K., Ninomiya, S., 2020. Field-based individual plant phenotyping of herbaceous species by unmanned aerial vehicle. *Ecology and Evolution* 10, 12318–12326. doi:[10.1002/ece3.6861](https://doi.org/10.1002/ece3.6861).
- Guo, W., Fukatsu, T., Ninomiya, S., 2015. Automated characterization of flowering dynamics in rice using field-acquired time-series RGB images. *Plant Methods* 11, 7. doi:[10.1186/s13007-015-0047-9](https://doi.org/10.1186/s13007-015-0047-9).
- Guo, W., Zheng, B., Duan, T., Fukatsu, T., Chapman, S., Ninomiya, S., 2017. EasyPCC: Benchmark Datasets and Tools for High-Throughput Measurement of the Plant Canopy Coverage Ratio under Field Conditions. *Sensors* 17, 798. doi:[10.3390/s17040798](https://doi.org/10.3390/s17040798).
- Guo, W., Zheng, B., Potgieter, A.B., Diot, J., Watanabe, K., Noshita, K., Jordan, D.R., Wang, X., Watson, J., Ninomiya, S., Chapman, S.C., 2018. Aerial Imagery Analysis – Quantifying Appearance and Number of Sorghum Heads for Applications in Breeding and Agronomy. *Frontiers in Plant Science* 9, 1544. doi:[10.3389/fpls.2018.01544](https://doi.org/10.3389/fpls.2018.01544).
- Guo, X., Ahlawat, Y.K., Liu, T., Zare, A., 2022. Evaluation of Postharvest Senescence of Broccoli via Hyperspectral Imaging. *Plant Phenomics* 2022, 1–12. doi:[10.34133/2022/9761095](https://doi.org/10.34133/2022/9761095).
- Hämmerle, M., Höfle, B., 2018. Mobile low-cost 3D camera maize crop height measurements under field conditions. *Precision Agriculture* 19, 630–647. doi:[10.1007/s11119-017-9544-3](https://doi.org/10.1007/s11119-017-9544-3).
- Han, L., Yang, G., Dai, H., Xu, B., Yang, H., Feng, H., Li, Z., Yang, X., 2019. Modeling maize above-ground biomass based on machine learning approaches using UAV remote-sensing data. *plant methods* 15, 1–19. doi:[10.1186/s13007-019-0394-z](https://doi.org/10.1186/s13007-019-0394-z).
- Han, R., Wong, A.J.Y., Tang, Z., Truco, M.J., Lavelle, D.O., Kozik, A., Jin, Y., Michelmore, R.W., 2021. Drone phenotyping and machine learning enable discovery of loci regulating daily floral opening in lettuce. *Journal of Experimental Botany* 72, 2979–2994. doi:[10.1093/jxb/erab081](https://doi.org/10.1093/jxb/erab081).
- Hartley, R., Zisserman, A., 2003. Multiple View Geometry in Computer Vision. Cambridge university press.

- He, L., Sun, W., Chen, X., Han, L., Li, J., Ma, Y., Song, Y., 2021. Modeling Maize Canopy Morphology in Response to Increased Plant Density. *Frontiers in Plant Science* 11, 533514. doi:[10.3389/fpls.2020.533514](https://doi.org/10.3389/fpls.2020.533514).
- Herrero-Huerta, M., Bucksch, A., Puttonen, E., Rainey, K.M., 2020a. Canopy Roughness: A New Phenotypic Trait to Estimate Aboveground Biomass from Unmanned Aerial System. *Plant Phenomics* 2020, 1–10. doi:[10.34133/2020/6735967](https://doi.org/10.34133/2020/6735967).
- Herrero-Huerta, M., Gonzalez-Aguilera, D., Yang, Y., 2023. Structural Component Phenotypic Traits from Individual Maize Skeletonization by UAS-Based Structure-from-Motion Photogrammetry. *Drones* 7, 108. doi:[10.3390/drones7020108](https://doi.org/10.3390/drones7020108).
- Herrero-Huerta, M., Rodriguez-Gonzalvez, P., Rainey, K.M., 2020b. Yield prediction by machine learning from UAS-based mulit-sensor data fusion in soybean. *Plant Methods* 16, 78. doi:[10.1186/s13007-020-00620-6](https://doi.org/10.1186/s13007-020-00620-6).
- Hofle, B., 2014. Radiometric Correction of Terrestrial LiDAR Point Cloud Data for Individual Maize Plant Detection. *IEEE Geoscience and Remote Sensing Letters* 11, 94–98. doi:[10.1109/LGRS.2013.2247022](https://doi.org/10.1109/LGRS.2013.2247022).
- Hu, P., Chapman, S.C., Zheng, B., 2021. Coupling of machine learning methods to improve estimation of ground coverage from unmanned aerial vehicle (UAV) imagery for high-throughput phenotyping of crops. *Functional Plant Biology* doi:[10.1071/FP20309](https://doi.org/10.1071/FP20309).
- Hu, P., Guo, W., Chapman, S.C., Guo, Y., Zheng, B., 2019. Pixel size of aerial imagery constrains the applications of unmanned aerial vehicle in crop breeding. *ISPRS Journal of Photogrammetry and Remote Sensing* 154, 1–9. doi:[10.1016/j.isprsjprs.2019.05.008](https://doi.org/10.1016/j.isprsjprs.2019.05.008).
- Itakura, K., Hosoi, F., 2018. Automatic Leaf Segmentation for Estimating Leaf Area and Leaf Inclination Angle in 3D Plant Images. *Sensors* 18, 3576. doi:[10.3390/s18103576](https://doi.org/10.3390/s18103576).
- Jackson, L., Ramirez, I., Yokota, R., Fennimore, S., Koike, S., Henderson, D., Chaney, W., Calderón, F., Klonsky, K., 2004. On-farm assessment of organic matter and tillage management on vegetable yield, soil, weeds, pests, and economics in California. *Agriculture, Ecosystems & Environment* 103, 443–463. doi:[10.1016/j.agee.2003.11.013](https://doi.org/10.1016/j.agee.2003.11.013).
- Jamtsho, S., 2010. Geometric Modelling of 3D Range Cameras and Their Application for Structural Deformation Measurement. Master of Science. University of Calgary. Calgary, Alberta, Canada.
- Jay, S., Rabatel, G., Hadoux, X., Moura, D., Gorretta, N., 2015. In-field crop row phenotyping from 3D modeling performed using Structure from Motion. *Computers and Electronics in Agriculture* 110, 70–77. doi:[10.1016/j.compag.2014.09.021](https://doi.org/10.1016/j.compag.2014.09.021).
- Jiang, Y., Li, C., Paterson, A.H., Sun, S., Xu, R., Robertson, J., 2018. Quantitative Analysis of Cotton Canopy Size in Field Conditions Using a Consumer-Grade RGB-D Camera. *Frontiers in Plant Science* 8, 2233. doi:[10.3389/fpls.2017.02233](https://doi.org/10.3389/fpls.2017.02233).

Jignasu, A., Herron, E., Jubery, T.Z., Afful, J., Balu, A., Ganapathysubramanian, B., Sarkar, S., Krishnamurthy, A., 2023. Plant geometry reconstruction from field data using neural radiance fields, in: 2nd AAAI Workshop on AI for Agriculture and Food Systems, pp. 1–6.

Jimenez-Berni, J.A., Deery, D.M., Rozas-Larraondo, P., Condon, A.T.G., Rebetzke, G.J., James, R.A., Bovill, W.D., Furbank, R.T., Sirault, X.R.R., 2018. High Throughput Determination of Plant Height, Ground Cover, and Above-Ground Biomass in Wheat with LiDAR. *Frontiers in Plant Science* 9, 237. doi:[10.3389/fpls.2018.00237](https://doi.org/10.3389/fpls.2018.00237).

Jin, S., Su, Y., Gao, S., Wu, F., Hu, T., Liu, J., Li, W., Wang, D., Chen, S., Jiang, Y., Pang, S., Guo, Q., 2018. Deep Learning: Individual Maize Segmentation From Terrestrial Lidar Data Using Faster R-CNN and Regional Growth Algorithms. *Frontiers in Plant Science* 9, 866. doi:[10.3389/fpls.2018.00866](https://doi.org/10.3389/fpls.2018.00866).

Jin, S., Su, Y., Gao, S., Wu, F., Ma, Q., Xu, K., Ma, Q., Hu, T., Liu, J., Pang, S., Guan, H., Zhang, J., Guo, Q., 2020a. Separating the Structural Components of Maize for Field Phenotyping Using Terrestrial LiDAR Data and Deep Convolutional Neural Networks. *IEEE Transactions on Geoscience and Remote Sensing* 58, 2644–2658. doi:[10.1109/TGRS.2019.2953092](https://doi.org/10.1109/TGRS.2019.2953092).

Jin, S., Su, Y., Song, S., Xu, K., Hu, T., Yang, Q., Wu, F., Xu, G., Ma, Q., Guan, H., Pang, S., Li, Y., Guo, Q., 2020b. Non-destructive estimation of field maize biomass using terrestrial lidar: An evaluation from plot level to individual leaf level. *Plant Methods* 16, 69. doi:[10.1186/s13007-020-00613-5](https://doi.org/10.1186/s13007-020-00613-5).

Jin, S., Su, Y., Wu, F., Pang, S., Gao, S., Hu, T., Liu, J., Guo, Q., 2019. Stem-Leaf Segmentation and Phenotypic Trait Extraction of Individual Maize Using Terrestrial LiDAR Data. *IEEE Transactions on Geoscience and Remote Sensing* 57, 1336–1346. doi:[10.1109/TGRS.2018.2866056](https://doi.org/10.1109/TGRS.2018.2866056).

Jin, S., Su, Y., Zhang, Y., Song, S., Li, Q., Liu, Z., Ma, Q., Ge, Y., Liu, L., Ding, Y., Baret, F., Guo, Q., 2021a. Exploring Seasonal and Circadian Rhythms in Structural Traits of Field Maize from LiDAR Time Series. *Plant Phenomics* 2021, 1–15. doi:[10.34133/2021/9895241](https://doi.org/10.34133/2021/9895241).

Jin, S., Sun, X., Wu, F., Su, Y., Li, Y., Song, S., Xu, K., Ma, Q., Baret, F., Jiang, D., Ding, Y., Guo, Q., 2021b. Lidar sheds new light on plant phenomics for plant breeding and management: Recent advances and future prospects. *ISPRS Journal of Photogrammetry and Remote Sensing* 171, 202–223. doi:[10.1016/j.isprsjprs.2020.11.006](https://doi.org/10.1016/j.isprsjprs.2020.11.006).

Kapur, J.N., Sahoo, P.K., Wong, A.K.C., 1985. A new method for gray-level picture thresholding using the entropy of the histogram. *Computer Vision, Graphics, and Image Processing* 29, 273–285. doi:[10.1016/0734-189X\(85\)90125-2](https://doi.org/10.1016/0734-189X(85)90125-2).

Khomsin, Pratomo, D.G., Anjasmara, I.M., Ahmad, F., 2019. Analysis of The Volume Comparation of 3'S (TS, GNSS and TLS). *E3S Web of Conferences* 94, 01014. doi:[10.1051/e3sconf/20199401014](https://doi.org/10.1051/e3sconf/20199401014).

- Kierdorf, J., Junker-Frohn, L.V., Delaney, M., Olave, M.D., Burkart, A., Jaenicke, H., Muller, O., Rascher, U., Roscher, R., 2022. GrowliFlower: An image time-series dataset for GROWth analysis of cauLIFLOWER. *Journal of Field Robotics* 40, 173–192. doi:[10.1002/rob.22122](https://doi.org/10.1002/rob.22122).
- Kim, D.W., Yun, H.S., Jeong, S.J., Kwon, Y.S., Kim, S.G., Lee, W.S., Kim, H.J., 2018. Modeling and Testing of Growth Status for Chinese Cabbage and White Radish with UAV-Based RGB Imagery. *Remote Sensing* 10, 563. doi:[10.3390/rs10040563](https://doi.org/10.3390/rs10040563).
- Kim, H., Seung-jun Yang, Kwanghoon Sohn, 2003. 3D reconstruction of stereo images for interaction between real and virtual worlds, in: The Second IEEE and ACM International Symposium on Mixed and Augmented Reality, 2003. Proceedings., IEEE Comput. Soc, Tokyo, Japan. pp. 169–176. doi:[10.1109/ISMAR.2003.1240700](https://doi.org/10.1109/ISMAR.2003.1240700).
- Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A.C., Lo, W.Y., Dollár, P., Girshick, R., 2023. Segment anything. arXiv [arXiv:2304.02643](https://arxiv.org/abs/2304.02643).
- Kochi, N., Hayashi, A., Shinohara, Y., Tanabata, T., Kodama, K., Isobe, S., 2022. All-around 3D plant modeling system using multiple images and its composition. *Breeding Science* 72, 75–84. doi:[10.1270/jsbbs.21068](https://doi.org/10.1270/jsbbs.21068).
- Kochi, N., Isobe, S., Hayashi, A., Kodama, K., Tanabata, T., 2021. Introduction of All-Around 3D Modeling Methods for Investigation of Plants. *International Journal of Automation Technology* 15, 301–312. doi:[10.20965/ijat.2021.p0301](https://doi.org/10.20965/ijat.2021.p0301).
- Kochi, N., Tanabata, T., Hayashi, A., Isobe, S., R&D Initiative, Chuo University 1-13-27 Kasuga, Bunkyo-ku, Tokyo 112-8551, Japan, Kazusa DNA Research Institute, Kisarazu, Japan, 2018. A 3D Shape-Measuring System for Assessing Strawberry Fruits. *International Journal of Automation Technology* 12, 395–404. doi:[10.20965/ijat.2018.p0395](https://doi.org/10.20965/ijat.2018.p0395).
- Kusumam, K., Krajník, T., Pearson, S., Duckett, T., Cielniak, G., 2017. 3D-vision based detection, localization, and sizing of broccoli heads in the field. *Journal of Field Robotics* 34, 1505–1518. doi:[10.1002/rob.21726](https://doi.org/10.1002/rob.21726).
- Lachat, E., Macher, H., Landes, T., Grussenmeyer, P., 2015. Assessment and Calibration of a RGB-D Camera (Kinect v2 Sensor) Towards a Potential Use for Close-Range 3D Modeling. *Remote Sensing* 7, 13070–13097. doi:[10.3390/rs71013070](https://doi.org/10.3390/rs71013070).
- Larue, B., 2020. Labor issues and COVID-19. *Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie* 68, 231–237. doi:[10.1111/cjag.12233](https://doi.org/10.1111/cjag.12233).
- Lehnert, C., Tsai, D., Eriksson, A., McCool, C., 2019. 3D Move to See: Multi-perspective visual servoing towards the next best view within unstructured and occluded environments, in: 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), IEEE, Macau, China. pp. 3890–3897. doi:[10.1109/IROS40897.2019.8967918](https://doi.org/10.1109/IROS40897.2019.8967918).

- Li, C., Adhikari, R., Yao, Y., Miller, A.G., Kalbaugh, K., Li, D., Nemali, K., 2020. Measuring plant growth characteristics using smartphone based image analysis technique in controlled environment agriculture. *Computers and Electronics in Agriculture* 168, 105123. doi:[10.1016/j.compag.2019.105123](https://doi.org/10.1016/j.compag.2019.105123).
- Li, D., Li, J., Xiang, S., Pan, A., 2022a. PSegNet: Simultaneous Semantic and Instance Segmentation for Point Clouds of Plants. *Plant Phenomics* 2022. doi:[10.34133/2022/9787643](https://doi.org/10.34133/2022/9787643).
- Li, D., Shi, G., Li, J., Chen, Y., Zhang, S., Xiang, S., Jin, S., 2022b. PlantNet: A dual-function point cloud segmentation network for multiple plant species. *ISPRS Journal of Photogrammetry and Remote Sensing* 184, 243–263. doi:[10.1016/j.isprsjprs.2022.01.007](https://doi.org/10.1016/j.isprsjprs.2022.01.007).
- Li, Y., Fan, X., Mitra, N.J., Chamovitz, D., Cohen-Or, D., Chen, B., 2013. Analyzing growing plants from 4D point cloud data. *ACM Transactions on Graphics* 32, 1–10. doi:[10.1145/2508363.2508368](https://doi.org/10.1145/2508363.2508368).
- Li, Y., Wang, F., Hu, X., 2022c. Deep-Learning-Based 3D Reconstruction: A Review and Applications. *Applied Bionics and Biomechanics* 2022, 1–6. doi:[10.1155/2022/3458717](https://doi.org/10.1155/2022/3458717).
- Li, Y., Wen, W., Fan, J., Gou, W., Gu, S., Lu, X., Yu, Z., Wang, X., Guo, X., 2023. Multi-Source Data Fusion Improves Time-Series Phenotype Accuracy in Maize under a Field High-Throughput Phenotyping Platform. *Plant Phenomics* 5, 0043. doi:[10.34133/plantphenomics.0043](https://doi.org/10.34133/plantphenomics.0043).
- Liang, C.M., Li, Y.W., Liu, Y.H., Wen, P.F., Yang, H., 2022. Segmentation and weight prediction of grape ear based on SFNet-ResNet18. *Systems Science & Control Engineering* 10, 722–732. doi:[10.1080/21642583.2022.2110541](https://doi.org/10.1080/21642583.2022.2110541).
- Lin, C., Hu, F., Peng, J., Wang, J., Zhai, R., 2022. Segmentation and Stratification Methods of Field Maize Terrestrial LiDAR Point Cloud. *Agriculture* 12, 1450. doi:[10.3390/agriculture12091450](https://doi.org/10.3390/agriculture12091450).
- Lin, Y.C., Zhou, T., Wang, T., Crawford, M., Habib, A., 2021. New orthophoto generation strategies from UAV and ground remote sensing platforms for high-throughput phenotyping. *Remote Sensing* 13, 860. doi:[10.3390/rs13050860](https://doi.org/10.3390/rs13050860).
- Liu, F., Hu, P., Zheng, B., Duan, T., Zhu, B., Guo, Y., 2021a. A field-based high-throughput method for acquiring canopy architecture using unmanned aerial vehicle images. *Agricultural and Forest Meteorology* 296, 108231. doi:[10.1016/j.agrformet.2020.108231](https://doi.org/10.1016/j.agrformet.2020.108231).
- Liu, F., Song, Q., Zhao, J., Mao, L., Bu, H., Hu, Y., Zhu, X.G., 2021b. Canopy occupation volume as an indicator of canopy photosynthetic capacity. *New Phytologist* 232, 941–956. doi:[10.1111/nph.17611](https://doi.org/10.1111/nph.17611).
- Liu, S., Baret, F., Abichou, M., Boudon, F., Thomas, S., Zhao, K., Fournier, C., Andrieu, B., Irfan, K., Hemmerlé, M., de Solan, B., 2017a. Estimating wheat green area index from ground-based LiDAR measurement using a 3D canopy structure model. *Agricultural and Forest Meteorology* 247, 12–20. doi:[10.1016/j.agrformet.2017.07.007](https://doi.org/10.1016/j.agrformet.2017.07.007).

- Liu, S., Baret, F., Andrieu, B., Burger, P., Hemmerlé, M., 2017b. Estimation of Wheat Plant Density at Early Stages Using High Resolution Imagery. *Frontiers in Plant Science* 8, 739. doi:[10.3389/fpls.2017.00739](https://doi.org/10.3389/fpls.2017.00739).
- Liu, S., Yin, D., Feng, H., Li, Z., Xu, X., Shi, L., Jin, X., 2022. Estimating maize seedling number with UAV RGB images and advanced image processing methods. *Precision Agriculture* doi:[10.1007/s11119-022-09899-y](https://doi.org/10.1007/s11119-022-09899-y).
- Lobell, D.B., Gourdji, S.M., 2012. The Influence of Climate Change on Global Crop Productivity. *Plant Physiology* 160, 1686–1697. doi:[10.1104/pp.112.208298](https://doi.org/10.1104/pp.112.208298).
- Lobet, G., 2017. Image Analysis in Plant Sciences: Publish Then Perish. *Trends in Plant Science* 22, 559–566. doi:[10.1016/j.tplants.2017.05.002](https://doi.org/10.1016/j.tplants.2017.05.002).
- Lu, N., Wu, Y., Zheng, H., Yao, X., Zhu, Y., Cao, W., Cheng, T., 2022. An assessment of multi-view spectral information from UAV-based color-infrared images for improved estimation of nitrogen nutrition status in winter wheat. *Precision Agriculture* 23, 1653–1674. doi:[10.1007/s11119-022-09901-7](https://doi.org/10.1007/s11119-022-09901-7).
- Lüling, N., Reiser, D., Stana, A., Griepentrog, H.W., 2021. Using depth information and colour space variations for improving outdoor robustness for instance segmentation of cabbage. arXiv [arXiv:2103.16923](https://arxiv.org/abs/2103.16923).
- Ma, X., Zhu, K., Guan, H., Feng, J., Yu, S., Liu, G., 2019. Calculation Method for Phenotypic Traits Based on the 3D Reconstruction of Maize Canopies. *Sensors* 19, 1201. doi:[10.3390/s19051201](https://doi.org/10.3390/s19051201).
- Magistri, F., Chebrolu, N., Behley, J., Stachniss, C., 2021. Towards in-field phenotyping exploiting differentiable rendering with self-consistency loss, in: IEEE International Conference on Robotics and Automation (ICRA).
- Mardanisamani, S., Ayalew, T.W., Badhon, M.A., Khan, N.A., Hasnat, G., Duddu, H., Shirtliffe, S., Vail, S., Stavness, I., Eramian, M., 2021. Automatic Microplot Localization Using UAV Images and a Hierarchical Image-Based Optimization Method. *Plant Phenomics* 2021, 9764514. doi:[10.34133/2021/9764514](https://doi.org/10.34133/2021/9764514).
- Martinez-Guanter, J., Ribeiro, Á., Peteinatos, G.G., Pérez-Ruiz, M., Gerhards, R., Bengochea-Guevara, J.M., Machleb, J., Andújar, D., 2019. Low-Cost Three-Dimensional Modeling of Crop Plants. *Sensors* 19, 2883. doi:[10.3390/s19132883](https://doi.org/10.3390/s19132883).
- Maturana, D., Scherer, S., 2015. VoxNet: A 3D convolutional neural network for real-time object recognition, in: IEEE/RSJ International Conference on Intelligent Robots and Systems, Hamburg, Germany. pp. 922–928. doi:[10.1109/IROS.2015.7353481](https://doi.org/10.1109/IROS.2015.7353481).
- McDonald, D.D., 1994. Reversible NLP by Linking the Grammar to the Knowledge Base, in: Strzalkowski, T. (Ed.), *Reversible Grammar in Natural Language Processing*. Springer US, Boston, MA, pp. 257–291. doi:[10.1007/978-1-4615-2722-0_11](https://doi.org/10.1007/978-1-4615-2722-0_11).
- McGuire, M., Soman, C., Diers, B., Chowdhary, G., 2021. High Throughput Soybean Pod-Counting with In-Field Robotic Data Collection and Machine-Vision Based Data Analysis. arXiv [arXiv:2105.10568](https://arxiv.org/abs/2105.10568).

- Meyer, G.E., Neto, J.C., 2008. Verification of color vegetation indices for automated crop imaging applications. Computers and Electronics in Agriculture 63, 282–293. doi:[10.1016/j.compag.2008.03.009](https://doi.org/10.1016/j.compag.2008.03.009).
- Miao, T., Wen, W., Li, Y., Wu, S., Zhu, C., Guo, X., 2021. Label3DMaize: Toolkit for 3D point cloud data annotation of maize shoots. GigaScience 10, giab031. doi:[10.1093/gigascience/giab031](https://doi.org/10.1093/gigascience/giab031).
- Mikami, T., Ishikawa, H., Hara, N., Natsusako, K., Kuga, T., Ohwada, K., Kitaura, H., Ueda, R., 2022. Hidden Main Stem Detection of Tree in Dense Fruit Vegetable Field: —Image-to-Image Translation by Deep Convolutional Neural Networks Learned with Realistic Computer Graphics—(In Japanese). Journal of the Robotics Society of Japan 40, 143–153. doi:[10.7210/jrsj.40.143](https://doi.org/10.7210/jrsj.40.143).
- Mildenhall, B., Srinivasan, P.P., Tancik, M., Barron, J.T., Ramamoorthi, R., Ng, R., 2022. NeRF: Representing scenes as neural radiance fields for view synthesis. Communications of the ACM 65, 99–106. doi:[10.1145/3503250](https://doi.org/10.1145/3503250).
- Mirande, K., Godin, C., Tisserand, M., Charlaix, J., Besnard, F., Hétroy-Wheeler, F., 2022. A graph-based approach for simultaneous semantic and instance segmentation of plant 3D point clouds. Frontiers in Plant Science 13, 1012669. doi:[10.3389/fpls.2022.1012669](https://doi.org/10.3389/fpls.2022.1012669).
- Montes, H.A., Le Louedec, J., Cielniak, G., Duckett, T., 2020. Real-time detection of broccoli crops in 3D point clouds for autonomous robotic harvesting, in: 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 10483–10488. doi:[10.1109/IROS45743.2020.9341381](https://doi.org/10.1109/IROS45743.2020.9341381).
- Mortensen, A., Laursen, M., Jørgensen, R., Gislum, R., 2019. Drone dataflow - a MATLAB toolbox for extracting plots from images captured by a UAV, in: Precision agriculture '19, Wageningen Academic Publishers, Montpellier, France. pp. 959–965. doi:[10.3920/978-90-8686-888-9_118](https://doi.org/10.3920/978-90-8686-888-9_118).
- Mu, Y., Fujii, Y., Takata, D., Zheng, B., Noshita, K., Honda, K., Ninomiya, S., Guo, W., 2018. Characterization of peach tree crown by using high-resolution images from an unmanned aerial vehicle. Horticulture Research 5, 74. doi:[10.1038/s41438-018-0097-z](https://doi.org/10.1038/s41438-018-0097-z).
- Nakashima, T., Tomobe, H., Morigaki, T., Yang, M., Yamaguchi, H., Kato, Y., Guo, W., Sharma, V., Kimura, H., Morikawa, H., 2023. Non-destructive high-throughput measurement of elastic-viscous properties of maize using a novel ultra-micro sensor array and numerical validation. Scientific Reports 13, 4914. doi:[10.1038/s41598-023-32130-5](https://doi.org/10.1038/s41598-023-32130-5).
- Nesteruk, S., Shadrin, D., Pukalchik, M., 2021. Image Augmentation for Multitask Few-Shot Learning: Agricultural Domain Use-Case. arXiv [arXiv:2102.12295](https://arxiv.org/abs/2102.12295).
- Nguyen, C., Sagan, V., Bhadra, S., Moose, S., 2023. UAV Multisensory Data Fusion and Multi-Task Deep Learning for High-Throughput Maize Phenotyping. Sensors 23, 1827. doi:[10.3390/s23041827](https://doi.org/10.3390/s23041827).

- Nguyen, C.V., Fripp, J., Lovell, D.R., Furbank, R., Kuffner, P., Daily, H., Sirault, X., 2016. 3D Scanning System for Automatic High-Resolution Plant Phenotyping, in: 2016 International Conference on Digital Image Computing: Techniques and Applications (DICTA), Gold Coast, QLD, Australia. pp. 1–8. doi:[10.1109/DICTA.2016.7796984](https://doi.org/10.1109/DICTA.2016.7796984).
- Nguyen, T., Slaughter, D., Max, N., Maloof, J., Sinha, N., 2015. Structured Light-Based 3D Reconstruction System for Plants. *Sensors* 15, 18587–18612. doi:[10.3390/s150818587](https://doi.org/10.3390/s150818587).
- Nishida, E., 2023. Estimation of Optimal Harvesting Date for Broccoli Using a Model for Whole-Individual Measurement and Growth Prediction with the Use of UAV (In Japanese, UAV を用いた全個体計測と生育予測モデルによる ブロッコリーの最適収穫日推定). Master's thesis. The University of Tokyo.
- Nobis, M., Hunziker, U., 2005. Automatic thresholding for hemispherical canopy-photographs based on edge detection. *Agricultural and Forest Meteorology* 128, 243–250. doi:[10.1016/j.agrformet.2004.10.002](https://doi.org/10.1016/j.agrformet.2004.10.002).
- Okura, F., 2022. 3D modeling and reconstruction of plants and trees: A cross-cutting review across computer graphics, vision, and plant phenotyping. *Breeding Science* 72, 31–47. doi:[10.1270/jsbbs.21074](https://doi.org/10.1270/jsbbs.21074).
- Otsu, N., 1979. A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems, Man, and Cybernetics* 9, 62–66. doi:[10.1109/TSMC.1979.4310076](https://doi.org/10.1109/TSMC.1979.4310076).
- Paturkar, A., Sen Gupta, G., Bailey, D., 2021. Making Use of 3D Models for Plant Physiognomic Analysis: A Review. *Remote Sensing* 13, 2232. doi:[10.3390/rs13112232](https://doi.org/10.3390/rs13112232).
- Paulus, S., 2019. Measuring crops in 3D: Using geometry for plant phenotyping. *Plant Methods* 15, 103. doi:[10.1186/s13007-019-0490-0](https://doi.org/10.1186/s13007-019-0490-0).
- Phan, A.V., Nguyen, M.L., Nguyen, Y.L.H., Bui, L.T., 2018. DGCNN: A convolutional neural network over large-scale labeled graphs. *Neural Networks* 108, 533–543. doi:[10.1016/j.neunet.2018.09.001](https://doi.org/10.1016/j.neunet.2018.09.001).
- Pitiot, A., Bardinet, E., Thompson, P.M., Malandain, G., 2006. Piecewise affine registration of biological images for volume reconstruction. *Medical Image Analysis* 10, 465–483. doi:[10.1016/j.media.2005.03.008](https://doi.org/10.1016/j.media.2005.03.008).
- Pylianidis, C., Osinga, S., Athanasiadis, I.N., 2021. Introducing digital twins to agriculture. *Computers and Electronics in Agriculture* 184, 105942. doi:[10.1016/j.compag.2020.105942](https://doi.org/10.1016/j.compag.2020.105942).
- Qi, C.R., Su, H., Mo, K., Guibas, L.J., 2016. PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation. arXiv [arXiv:1612.00593](https://arxiv.org/abs/1612.00593).
- Qi, C.R., Yi, L., Su, H., Guibas, L.J., 2017. PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space. arXiv [arXiv:1706.02413](https://arxiv.org/abs/1706.02413).

- Qiu, Q., Sun, N., Bai, H., Wang, N., Fan, Z., Wang, Y., Meng, Z., Li, B., Cong, Y., 2019. Field-Based High-Throughput Phenotyping for Maize Plant Using 3D LiDAR Point Cloud Generated With a “Phenomobile”. *Frontiers in Plant Science* 10. doi:[10.3389/fpls.2019.00554](https://doi.org/10.3389/fpls.2019.00554).
- Quine, T.A., Zhang, Y., 2002. An investigation of spatial variation in soil erosion, soil properties, and crop production within an agricultural field in Devon, United Kingdom. *Journal of Soil and Water Conservation* 57, 55–65.
- Ray, D.K., Mueller, N.D., West, P.C., Foley, J.A., 2013. Yield Trends Are Insufficient to Double Global Crop Production by 2050. *PLoS ONE* 8, e66428. doi:[10.1371/journal.pone.0066428](https://doi.org/10.1371/journal.pone.0066428).
- Robertson, D.J., Julias, M., Lee, S.Y., Cook, D.D., 2017. Maize Stalk Lodging: Morphological Determinants of Stalk Strength. *Crop Science* 57, 926–934. doi:[10.2135/cropsci2016.07.0569](https://doi.org/10.2135/cropsci2016.07.0569).
- Romano, G., Zia, S., Spreer, W., Sanchez, C., Cairns, J., Araus, J.L., Müller, J., 2011. Use of thermography for high throughput phenotyping of tropical maize adaptation in water stress. *Computers and Electronics in Agriculture* 79, 67–74. doi:[10.1016/j.compag.2011.08.011](https://doi.org/10.1016/j.compag.2011.08.011).
- Rossi, R., Costafreda-Aumedes, S., Leolini, L., Leolini, C., Bindi, M., Moriondo, M., 2022. Implementation of an algorithm for automated phenotyping through plant 3D-modeling: A practical application on the early detection of water stress. *Computers and Electronics in Agriculture* 197, 106937. doi:[10.1016/j.compag.2022.106937](https://doi.org/10.1016/j.compag.2022.106937).
- Satodiya, B., Patel, H., Soni, N., 2015. Effect of planting density and integrated nutrient management on flowering, growth and yield of vegetable cowpea [*Vigna unguiculata* (L) Walp]. *The Asian Journal of Horticulture* 10, 232–236. doi:[10.15740/HAS/TAJH/10.2/232-236](https://doi.org/10.15740/HAS/TAJH/10.2/232-236).
- Schima, R., Mollenhauer, H., Grenzdörffer, G., Merbach, I., Lausch, A., Dietrich, P., Bumberger, J., 2016. Imagine All the Plants: Evaluation of a Light-Field Camera for On-Site Crop Growth Monitoring. *Remote Sensing* 8, 823. doi:[10.3390/rs8100823](https://doi.org/10.3390/rs8100823).
- Schunck, D., Magistri, F., Rosu, R.A., Cornelissen, A., Chebrolu, N., Paulus, S., Léon, J., Behnke, S., Stachniss, C., Kuhlmann, H., Klingbeil, L., 2021. Pheno4D: A spatio-temporal dataset of maize and tomato plant point clouds for phenotyping and advanced plant analysis. *PLoS ONE* 16, e0256340. doi:[10/gnzfdm](https://doi.org/10/gnzfdm).
- Shalma, H., Selvaraj, P., 2023. A review on 3D image reconstruction on specific and generic objects. *Materials Today: Proceedings* 80, 2400–2405. doi:[10.1016/j.mtpr.2021.06.371](https://doi.org/10.1016/j.mtpr.2021.06.371).
- Shao, W., Kawakami, R., Yoshihashi, R., You, S., Kawase, H., Naemura, T., 2020. Cattle detection and counting in UAV images based on convolutional neural networks. *International Journal of Remote Sensing* 41, 31–52. doi:[10.1080/01431161.2019.1624858](https://doi.org/10.1080/01431161.2019.1624858).

- Shu, M., Shen, M., Zuo, J., Yin, P., Wang, M., Xie, Z., Tang, J., Wang, R., Li, B., Yang, X., Ma, Y., 2021. The Application of UAV-Based Hyperspectral Imaging to Estimate Crop Traits in Maize Inbred Lines. *Plant Phenomics* 2021, 9890745. doi:[10.34133/2021/9890745](https://doi.org/10.34133/2021/9890745).
- Slob, N., Hurst, W., van de Zedde, R., Tekinerdogan, B., 2023. Virtual reality-based digital twins for greenhouses: A focus on human interaction. *Computers and Electronics in Agriculture* 208, 107815. doi:[10.1016/j.compag.2023.107815](https://doi.org/10.1016/j.compag.2023.107815).
- Snavely, N., Simon, I., Goesele, M., Szeliski, R., Seitz, S.M., 2010. Scene Reconstruction and Visualization From Community Photo Collections. *Proceedings of the IEEE* 98, 1370–1390. doi:[10.1109/JPROC.2010.2049330](https://doi.org/10.1109/JPROC.2010.2049330).
- Sorrentino, M., Colla, G., Rouphael, Y., Panzarová, K., Trtílek, M., 2020. Lettuce reaction to drought stress: Automated high-throughput phenotyping of plant growth and photosynthetic performance. *Acta Horticulturae* 1268, 133–142. doi:[10.17660/ActaHortic.2020.1268.17](https://doi.org/10.17660/ActaHortic.2020.1268.17).
- Stansell, Z., Björkman, T., Branham, S., Couillard, D., Farnham, M.W., 2017. Use of a Quality Trait Index to Increase the Reliability of Phenotypic Evaluations in Broccoli. *HortScience* 52, 1490–1495. doi:[10.21273/HORTSCI12202-17](https://doi.org/10.21273/HORTSCI12202-17).
- Su, H., Maji, S., Kalogerakis, E., Learned-Miller, E., 2015. Multi-view Convolutional Neural Networks for 3D Shape Recognition. arXiv [arXiv:1505.00880](https://arxiv.org/abs/1505.00880).
- Su, W., Zhu, D., Huang, J., Guo, H., 2018. Estimation of the vertical leaf area profile of corn (*Zea mays*) plants using terrestrial laser scanning (TLS). *Computers and Electronics in Agriculture* 150, 5–13. doi:[10.1016/j.compag.2018.03.037](https://doi.org/10.1016/j.compag.2018.03.037).
- Su, Y., Wu, F., Ao, Z., Jin, S., Qin, F., Liu, B., Pang, S., Liu, L., Guo, Q., 2019. Evaluating maize phenotype dynamics under drought stress using terrestrial lidar. *Plant Methods* 15, 11. doi:[10.1186/s13007-019-0396-x](https://doi.org/10.1186/s13007-019-0396-x).
- Sun, S., Li, C., Chee, P.W., Paterson, A.H., Jiang, Y., Xu, R., Robertson, J.S., Adhikari, J., Shehzad, T., 2020. Three-dimensional photogrammetric mapping of cotton bolls in situ based on point cloud segmentation and clustering. *ISPRS Journal of Photogrammetry and Remote Sensing* 160, 195–207. doi:[10.1016/j.isprsjprs.2019.12.011](https://doi.org/10.1016/j.isprsjprs.2019.12.011).
- Sun, S., Li, C., Paterson, A.H., Jiang, Y., Xu, R., Robertson, J.S., Snider, J.L., Chee, P.W., 2018. In-field High Throughput Phenotyping and Cotton Plant Growth Analysis Using LiDAR. *Frontiers in Plant Science* 9, 16. doi:[10.3389/fpls.2018.00016](https://doi.org/10.3389/fpls.2018.00016).
- Ten Harkel, J., Bartholomeus, H., Kooistra, L., 2019. Biomass and Crop Height Estimation of Different Crops Using UAV-Based Lidar. *Remote Sensing* 12, 17. doi:[10.3390/rs12010017](https://doi.org/10.3390/rs12010017).
- Thapa, S., Zhu, F., Walia, H., Yu, H., Ge, Y., 2018. A Novel LiDAR-Based Instrument for High-Throughput, 3D Measurement of Morphological Traits in Maize and Sorghum. *Sensors* 18, 1187. doi:[10.3390/s18041187](https://doi.org/10.3390/s18041187).

- Tilman, D., Balzer, C., Hill, J., Befort, B.L., 2011. Global food demand and the sustainable intensification of agriculture. *Proceedings of the National Academy of Sciences* 108, 20260–20264. doi:[10.1073/pnas.1116437108](https://doi.org/10.1073/pnas.1116437108).
- Tölgessy, M., Dekan, M., Chovanec, Ľ., Hubinský, P., 2021. Evaluation of the Azure Kinect and Its Comparison to Kinect V1 and Kinect V2. *Sensors* 21, 413. doi:[10.3390/s21020413](https://doi.org/10.3390/s21020413).
- Tresch, L., Mu, Y., Itoh, A., Kaga, A., Taguchi, K., Hirafuji, M., Ninomiya, S., Guo, W., 2019. Easy MPE: Extraction of Quality Microplot Images for UAV-Based High-Throughput Field Phenotyping. *Plant Phenomics* 2019, 1–9. doi:[10.34133/2019/2591849](https://doi.org/10.34133/2019/2591849).
- Van de Zedde, R., Kootstra, G., Shi, W., Jiang, H., 2019. Plant-part segmentation using deep learning and multi-view vision. *Biosystems Engineering* 187, 81–95. doi:[10.1016/j.biosystemseng.2019.08.014](https://doi.org/10.1016/j.biosystemseng.2019.08.014).
- Velumani, K., Lopez-Lozano, R., Madec, S., Guo, W., Gillet, J., Comar, A., Baret, F., 2021. Estimates of maize plant density from UAV RGB images using Faster-RCNN detection model: Impact of the spatial resolution. *Plant Phenomics* 2021, 9824843. doi:[10.34133/2021/9824843](https://doi.org/10.34133/2021/9824843).
- Verdouw, C., Tekinerdogan, B., Beulens, A., Wolfert, S., 2021. Digital twins in smart farming. *Agricultural Systems* 189, 103046. doi:[10.1016/j.agsy.2020.103046](https://doi.org/10.1016/j.agsy.2020.103046).
- Vicini, D., Speierer, S., Jakob, W., 2022. Differentiable signed distance function rendering. *ACM Transactions on Graphics* 41, 1–18. doi:[10.1145/3528223.3530139](https://doi.org/10.1145/3528223.3530139).
- Wang, D., Song, Z., Miao, T., Chao, Z., Yang, X., Yang, T., Zhou, Y., Den, H., Xu, T., 2023a. DFSP: A fast and automatic distance field-based stem-leaf segmentation pipeline for point cloud of maize shoot. *Frontiers in Plant Science* 14, 1109314. doi:[10.3389/fpls.2023.1109314](https://doi.org/10.3389/fpls.2023.1109314).
- Wang, H., 2019. Estimating Forest Attributes from Spherical Images. MSc Forestry thesis. The University of New Brunswick.
- Wang, H., Duan, Y., Shi, Y., Kato, Y., Ninomiya, S., Guo, W., 2021a. EasyIDP: A python package for intermediate data processing in UAV-based plant phenotyping. *Remote Sensing* 13, 2622. doi:[10.3390/rs13132622](https://doi.org/10.3390/rs13132622).
- Wang, H., Han, D., Mu, Y., Jiang, L., Yao, X., Bai, Y., Lu, Q., Wang, F., 2019. Landscape-level vegetation classification and fractional woody and herbaceous vegetation cover estimation over the dryland ecosystems by unmanned aerial vehicle platform. *Agricultural and Forest Meteorology* 278, 107665. doi:[10.1016/j.agrformet.2019.107665](https://doi.org/10.1016/j.agrformet.2019.107665).
- Wang, H., Li, T., Nishida, E., Kato, Y., Fukano, Y., Guo, W., 2023b. Drone-Based Harvest Data Prediction Can Reduce On-Farm Food Loss and Improve Farmer Income. *Plant Phenomics* 5, 0086. doi:[10.34133/plantphenomics.0086](https://doi.org/10.34133/plantphenomics.0086).

- Wang, H., Wang, F., Yao, X., Mu, Y., Bai, Y., Lu, Q., 2017. UAV-HiRAP: A Novel Method to Improve Landscape-Level Vegetation Classification and Coverage Fraction Estimation with Unmanned Aerial Vehicle Platform. Oral. The 12th International Congress of Ecological (INTECOL). Beijing, China.
- Wang, J., Wu, B., Kohnen, M.V., Lin, D., Yang, C., Wang, X., Qiang, A., Liu, W., Kang, J., Li, H., Shen, J., Yao, T., Su, J., Li, B., Gu, L., 2021b. Classification of Rice Yield Using UAV-Based Hyperspectral Imagery and Lodging Feature. *Plant Phenomics* 2021, 1–14. doi:[10.34133/2021/9765952](https://doi.org/10.34133/2021/9765952).
- Wang, Z., Nie, C., Wang, H., Ao, Y., Jin, X., Yu, X., Bai, Y., Liu, Y., Shao, M., Cheng, M., Liu, S., Wang, S., Tuohuti, N., 2021c. Detection and Analysis of Degree of Maize Lodging Using UAV-RGB Image Multi-Feature Factors and Various Classification Methods. *ISPRS International Journal of Geo-Information* 10, 309. doi:[10.3390/ijgi10050309](https://doi.org/10.3390/ijgi10050309).
- Wei, X., Zhang, Y., Li, Z., Fu, Y., Xue, X., 2020. DeepSFM: Structure from Motion via Deep Bundle Adjustment, in: Vedaldi, A., Bischof, H., Brox, T., Frahm, J.M. (Eds.), *Computer Vision – ECCV 2020*, Springer International Publishing, Cham. pp. 230–247. doi:[10.1007/978-3-030-58452-8_14](https://doi.org/10.1007/978-3-030-58452-8_14).
- Wen, W., Wang, Y., Wu, S., Liu, K., Gu, S., Guo, X., 2021. 3D phytomer-based geometric modelling method for plants—the case of maize. *AoB PLANTS* 13, plab055. doi:[10.1093/aobpla/plab055](https://doi.org/10.1093/aobpla/plab055).
- Wu, S., Wen, W., Wang, Y., Fan, J., Wang, C., Gou, W., Guo, X., 2020. MVS-Pheno: A Portable and Low-Cost Phenotyping Platform for Maize Shoots Using Multiview Stereo 3D Reconstruction. *Plant Phenomics* 2020, 1–17. doi:[10.34133/2020/1848437](https://doi.org/10.34133/2020/1848437).
- Wu, S., Wen, W., Xiao, B., Guo, X., Du, J., Wang, C., Wang, Y., 2019. An Accurate Skeleton Extraction Approach From 3D Point Clouds of Maize Plants. *Frontiers in Plant Science* 10, 248. doi:[10.3389/fpls.2019.00248](https://doi.org/10.3389/fpls.2019.00248).
- Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C., Yu, P.S., 2021. A Comprehensive Survey on Graph Neural Networks. *IEEE Transactions on Neural Networks and Learning Systems* 32, 4–24. doi:[10.1109/TNNLS.2020.2978386](https://doi.org/10.1109/TNNLS.2020.2978386).
- Xiao, S., Chai, H., Shao, K., Shen, M., Wang, Q., Wang, R., Sui, Y., Ma, Y., 2020. Image-Based Dynamic Quantification of Aboveground Structure of Sugar Beet in Field. *Remote Sensing* 12, 269. doi:[10.3390/rs12020269](https://doi.org/10.3390/rs12020269).
- Xiao, S., Chai, H., Wang, Q., Shao, K., Meng, L., Wang, R., Li, B., Ma, Y., 2021. Estimating economic benefit of sugar beet based on three-dimensional computer vision: A case study in Inner Mongolia, China. *European Journal of Agronomy* 130, 126378. doi:[10.1016/j.eja.2021.126378](https://doi.org/10.1016/j.eja.2021.126378).
- Xu, N., Sun, G., Bai, Y., Zhou, X., Cai, J., Huang, Y., 2023. Global Reconstruction Method of Maize Population at Seedling Stage Based on Kinect Sensor. *Agriculture* 13, 348. doi:[10.3390/agriculture13020348](https://doi.org/10.3390/agriculture13020348).

- Yamamoto, K., Guo, W., Yoshioka, Y., Ninomiya, S., 2014. On Plant Detection of Intact Tomato Fruits Using Image Analysis and Machine Learning Methods. *Sensors* 14, 12191–12206. doi:[10.3390/s140712191](https://doi.org/10.3390/s140712191).
- Yang, B., Xu, Y., 2021. Applications of deep-learning approaches in horticultural research: A review. *Horticulture Research* 8, 1–31. doi:[10.1038/s41438-021-00560-9](https://doi.org/10.1038/s41438-021-00560-9).
- Yang, W., Wang, S., Zhao, X., Zhang, J., Feng, J., 2015. Greenness identification based on HSV decision tree. *Information Processing in Agriculture* 2, 149–160. doi:[10.1016/j.inpa.2015.07.003](https://doi.org/10.1016/j.inpa.2015.07.003).
- Yao, Y., Luo, Z., Li, S., Fang, T., Quan, L., 2018. MVSNet: Depth Inference for Unstructured Multi-view Stereo, in: Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y. (Eds.), *Computer Vision – ECCV 2018*, Springer International Publishing, Cham. pp. 785–801. doi:[10.1007/978-3-030-01237-3_47](https://doi.org/10.1007/978-3-030-01237-3_47).
- Yoo, J.C., Han, T.H., 2009. Fast Normalized Cross-Correlation. *Circuits, Systems and Signal Processing* 28, 819–843. doi:[10.1007/s00034-009-9130-7](https://doi.org/10.1007/s00034-009-9130-7).
- Yu, C., Gao, C., Wang, J., Yu, G., Shen, C., Sang, N., 2020. BiSeNet V2: Bilateral Network with Guided Aggregation for Real-time Semantic Segmentation. *International Journal of Computer Vision* 129, 3051–3068. doi:[10.1007/s11263-021-01515-2](https://doi.org/10.1007/s11263-021-01515-2), [arXiv:2004.02147](https://arxiv.org/abs/2004.02147).
- Zermas, D., Morellas, V., Mulla, D., Papanikolopoulos, N., 2020. 3D model processing for high throughput phenotype extraction – the case of corn. *Computers and Electronics in Agriculture* 172, 105047. doi:[10.1016/j.compag.2019.105047](https://doi.org/10.1016/j.compag.2019.105047).
- Zhang, L., Jin, J., Wang, L., Huang, P., Ma, D., 2020a. A 3D white referencing method for soybean leaves based on fusion of hyperspectral images and 3D point clouds. *Precision Agriculture* doi:[10.1007/s11119-020-09713-7](https://doi.org/10.1007/s11119-020-09713-7).
- Zhang, Q., Liu, Y., Gong, C., Chen, Y., Yu, H., 2020b. Applications of deep learning for dense scenes analysis in agriculture: A review. *Sensors* 20, 1520. doi:[10.3390/s20051520](https://doi.org/10.3390/s20051520).
- Zhang, Y., Teng, P., Shimizu, Y., Hosoi, F., Omasa, K., 2016. Estimating 3D Leaf and Stem Shape of Nursery Paprika Plants by a Novel Multi-Camera Photography System. *Sensors* 16, 874. doi:[10.3390/s16060874](https://doi.org/10.3390/s16060874).
- Zhao, C., Zhang, Y., Du, J., Guo, X., Wen, W., Gu, S., Wang, J., Fan, J., 2019. Crop Phenomics: Current Status and Perspectives. *Frontiers in Plant Science* 10, 714. doi:[10.3389/fpls.2019.00714](https://doi.org/10.3389/fpls.2019.00714).
- Zheng, E., Ji, D., Dunn, E., Frahm, J.M., 2015. Sparse Dynamic 3D Reconstruction from Unsynchronized Videos, in: 2015 IEEE International Conference on Computer Vision (ICCV), IEEE, Santiago, Chile. pp. 4435–4443. doi:[10.1109/ICCV.2015.504](https://doi.org/10.1109/ICCV.2015.504).
- Zhou, C., Hu, J., Xu, Z., Yue, J., Ye, H., Yang, G., 2020a. A Monitoring System for the Segmentation and Grading of Broccoli Head Based on Deep Learning and Neural Networks. *Frontiers in Plant Science* 11. doi:[10.3389/fpls.2020.00402](https://doi.org/10.3389/fpls.2020.00402).

- Zhou, C., Ye, H., Sun, D., Yue, J., Yang, G., Hu, J., 2022. An automated, high-performance approach for detecting and characterizing broccoli based on UAV remote-sensing and transformers: A case study from Haining, China. International Journal of Applied Earth Observation and Geoinformation 114, 103055. doi:[10.1016/j.jag.2022.103055](https://doi.org/10.1016/j.jag.2022.103055).
- Zhou, J., Cui, G., Hu, S., Zhang, Z., Yang, C., Liu, Z., Wang, L., Li, C., Sun, M., 2020b. Graph neural networks: A review of methods and applications. AI Open 1, 57–81. doi:[10.1016/j.aiopen.2021.01.001](https://doi.org/10.1016/j.aiopen.2021.01.001).
- Zhou, J., Fu, X., Zhou, S., Zhou, J., Ye, H., Nguyen, H.T., 2019. Automated segmentation of soybean plants from 3D point cloud using machine learning. Computers and Electronics in Agriculture 162, 143–153. doi:[10.1016/j.compag.2019.04.014](https://doi.org/10.1016/j.compag.2019.04.014).
- Zhou, J., Pan, L., Li, Y., Liu, P., Liu, L., 2020c. Real-Time Stripe Width Computation Using Back Propagation Neural Network for Adaptive Control of Line Structured Light Sensors. Sensors 20, 2618. doi:[10.3390/s20092618](https://doi.org/10.3390/s20092618).
- Zhou, Q.Y., Park, J., Koltun, V., 2018. Open3D: A Modern Library for 3D Data Processing. arXiv [arXiv:1801.09847](https://arxiv.org/abs/1801.09847).
- Zhu, B., Liu, F., Xie, Z., Guo, Y., Li, B., Ma, Y., 2020. Quantification of light interception within image-based 3D reconstruction of sole and intercropped canopies over the entire growth season. Annals of Botany 126, 701–712. doi:[10.1093/aob/mcaa046](https://doi.org/10.1093/aob/mcaa046).
- Zou, K., Ge, L., Zhang, C., Yuan, T., Li, W., 2019. Broccoli Seedling Segmentation Based on Support Vector Machine Combined With Color Texture Features. IEEE Access 7, 168565–168574. doi:[10.1109/ACCESS.2019.2954587](https://doi.org/10.1109/ACCESS.2019.2954587).

Publications

Journals

Related to the doctor studies

1. **Wang, H.**, Duan, Y., Shi, Y., Kato, Y., Ninomiya, S., Guo, W., 2021. EasyIDP: A python package for intermediate data processing in UAV-based plant phenotyping. *Remote Sensing* 13, 2622. <https://doi.org/10.3390/rs13132622>
2. **Wang, H.**, Tang, L., Nishida, E., Fukano, Y., Kato, Y., Guo, W., 2023. Drone-based harvest data prediction can reduce on-farm food loss and improve farmer income. *Plant Phenomics* 2023:5, 0086. <https://doi.org/10.34133/plantphenomics.0086>
3. **Wang, H.**, Tang, L., Nishida, E., Fukano, Y., Kato, Y., Guo, W., Virtual broccoli farmland by drone-based phenotyping and cross-scale assimilation. (In preparation)

Not related to the doctor studies

1. Zhang, W., Peng, X., Cui, G., **Wang, H.**, Takata, D., Guo, W., 2023. Tree Branch Skeleton Extraction from Drone-Based Photogrammetric Point Cloud. *Drones* 7, 65. <https://doi.org/10.3390/drones7020065>
2. Drofova, I., Guo, W., **Wang, H.**, Adamek, M., 2023. Use of scanning devices for object 3D reconstruction by photogrammetry and visualization in virtual reality. *Bulletin of Electrical Engineering and Informatics* 12, 868–881. <https://doi.org/10.11591/eei.v12i2.4584>
3. Zhao, L., Guo, W., Wang, J., **Wang, H.**, Duan, Y., Wang, C., Wu, W., Shi, Y., 2021. An Efficient Method for Estimating Wheat Heading Dates Using UAV Images. *Remote Sensing* 13, 3067. <https://doi.org/10.3390/rs13163067>
4. Feldman, A., **Wang, H.**, Fukano, Y., Kato, Y., Ninomiya, S., Guo, W., 2021. EasyDCP: An affordable, high-throughput tool to measure plant phenotypic traits in 3D. *Methods in Ecology and Evolution* 12, 1679–1686. <https://doi.org/10.1111/2041-210X.13645>
5. **Wang, H.**, Yang, T.-R., Waldy, J., Kershaw, J.A., 2021b. Estimating Individual Tree Heights DBHs from Vertically Displaced Spherical Image Pairs. *Mathematical and Computational Forestry & Natural-Resource Sciences* 13, 1–14. <https://mcfns.net/index.php/Journal/article/view/13.1>

6. Hsu, Y.-H., Kershaw, J.A., Ducey, M.J., Yang, T.-R., **Wang, H.**, 2021. Sampling with probability proportional to prediction (3P sampling) using covariates derived from spherical images. *Canadian Journal of Forest Research* 51, 1140–1147. <https://doi.org/10.1139/cjfr-2020-0498>
7. Dai, X., Ducey, M.J., Kershaw, J.A., **Wang, H.**, 2021a. Sector subsampling for basal area ratio estimation: an alternative to big BAF sampling. *Canadian Journal of Forest Research* 1–9. <https://doi.org/10.1139/cjfr-2020-0496>
8. Dai, X., Ducey, M.J., **Wang, H.**, Yang, T.-R., Hsu, Y.-H., Ogilvie, J., Kershaw, J.A., Jr, 2021b. Biomass estimates derived from sector subsampling of 360° spherical images. *Forestry: An International Journal of Forest Research* 94, 565–575. <https://doi.org/10.1093/forestry/cpab023>
9. **Wang, H.**, Kershaw, J.A., Yang, T.-R., Hsu, Y.-H., Ma, X., Chen, Y., 2020. An Integrated System for Estimating Forest Basal Area from Spherical Images. *Mathematical and Computational Forestry & Natural-Resource Sciences* 12, 0–14. <http://mcfns.net/index.php/Journal/article/view/12.1>
10. **Wang, H.**, Han, D., Mu, Y., Jiang, L., Yao, X., Bai, Y., Lu, Q., Wang, F., 2019. Landscape-level vegetation classification and fractional woody and herbaceous vegetation cover estimation over the dryland ecosystems by unmanned aerial vehicle platform. *Agricultural and Forest Meteorology* 278, 107665. <https://doi.org/10.1016/j.agrformet.2019.107665>

Conferences

Related to the doctor studies

1. **Wang, H.**, Tang, L., Nishida, E., Fukano, Y., Kato, Y., Guo, W. Sept 27-30, 2022. Estimate Optimal Harvest Time by Cross-scale Assimilated Digital Broccoli Farmland (**poster**), *7th International Plant Phenotyping Symposium: "Plant Phenotyping for a Sustainable Future"*, Wageningen, Netherlands.
2. **Wang, H.**, Tang, L., Nishida, E., Fukano, Y., Kato, Y., Guo, W. July 20-22, 2021. Cost-efficient broccoli head phenotyping using aerial imagery and structure-from-motion (SfM)-based weakly supervised learning (**poster**), *The 8th International Horticulture Research Conference*, Nanjing, Jiangsu, China.
3. **Wang, H.**, Kato, Y., Guo, W., June 3-4, 2021. EasyIDP: A python package for intermediate data processing in UAV based plant phenotyping (**poster**), 超分野植物科学研究会の第1回研究集会 2021, Zoom online, Tokyo, Japan.
4. **Wang, H.**, Kato, Y., Guo, W., May 22, 2021. EasyIDP: A python package for intermediate data processing in UAV based plant phenotyping (**poster**), 農業情報学会 JSAL 2021 年次大会, Zoom online, Tokyo, Japan.

Not related to the doctor studies

1. **Wang, H.**, Kato, Y., Guo, W. May 21-22, 2022. Procedural Geometric Modeling for Plant Phenomics by Blender: Case Study of Maize (**oral**), 農業情報学会 JSAL 2022 年次大会, Kyoto, Japan.

2. Feldman, A., **Wang, H.**, Fukano, Y., Kato, Y., Ninomiya, S., Guo, W., February 24-27, 2020. Affordable high-throughput processing of handheld camera images of container plants to phenotypic data (**poster**), *Phenome 2020*, Tucson Convention Center, Tucson, Arizona, U.S.
3. Feldman, A., **Wang, H.**, Fukano, Y., Guo, W., October 22-25, 2019. Affordable high-throughput processing of multi-scale images to phenotypic data (**poster**). *The 6th International Plant Phenotyping Symposium*, Nanjing, Jiangsu, China.
4. **Wang, H.**, Kershaw, J.A., June 23-25, 2019. Estimating Forest Attributes from Spherical Images (**oral**, **poster**). *The Western Mensurationists 2019 Annual Meeting*, Kamloops Hotel and Conf. Center, Kamloops, British Columbia, Canada.
5. **Wang, H.**, Kershaw, J.A., March 23, 2018. Measuring Plant Area Index (PAI) from panorama photo images (**oral**). *The 25th Annual UNB Graduate Research Conference (GRC)*, Wu Conference Center, Fredericton, New Brunswick, Canada.
6. **Wang, H.**, Kershaw, J.A., November 5-7, 2017. Extracting DBH Measurements from RGB Photo Images (**oral**). *The Northeastern Mensurationists 2017 Annual Meeting*, The Inn at Saratoga. Saratoga Springs, New York, U.S.
7. **Wang, H.**, Wang, F., Yao, X., Mu, Y., Bai, Y., Lu, Q., August 20-25, 2017. UAV-HiRAP: A novel method to improve landscape-level vegetation classification and coverage fraction estimation with unmanned aerial vehicle platform (**oral**). *The 12th International Congress of Ecological (INTECOL)*, China National Convention Center, Beijing, China.

Awards

1. **May 22, 2021:** Young researcher innovation award (若手研究者イノベーション賞), on Conference 農業情報学会 *JSAI 2021* 年次大会
2. **October 2020 to March 2021:** Honors scholarship from *Japan Student Service Organization*

Source Code

Easy-Series Software packages (some of them are not publicly published at the current stage)

1. EasyIDP - A handy tool for dealing with regions of interest (ROI) on the image reconstruction (Metashape & Pix4D) outputs, mainly in agriculture applications. <https://github.com/UTokyo-FieldPhenomics-Lab/EasyIDP>
2. EastAMS - A GUI plugin tool for Agisoft Metashape with extended functions for smart agriculture. <https://github.com/UTokyo-FieldPhenomics-Lab/EasyAMS>
3. EastBPY - The wrapper for blender python application programming interface (API) without pip install extra packages. <https://github.com/UTokyo-FieldPhenomics-Lab/EasyBPY>

4. EasyDCP - Easy Dense Cloud Phenotyping. <https://github.com/UTokyo-FieldPhenomics-Lab/EasyDCP>

Projects

1. UAVbroccoli - Protocol for aerial broccoli data analysis by Metashape. <https://github.com/UTokyo-FieldPhenomics-Lab/UAVbroccoli>
2. Foldio360_3D_Reconstruct_Platform - a protocol for close-range 3D reconstruction for small objects. https://github.com/UTokyo-FieldPhenomics-Lab/Foldio360_3D_Reconstruct_Platform
3. BroccoliHead3D.ipynb - the scripts for close-range 3D broccoli head phenotyping. <https://github.com/UTokyo-FieldPhenomics-Lab/BroccoliHead3D.ipynb>

Website maintainer

1. www.global-wheat.com - The website for global wheat dataset. <https://github.com/UTokyo-FieldPhenomics-Lab/global-wheat.github.io>
2. 2023.mlcas.site - The website for Fifth International Workshop on Machine Learning for Cyber-Agricultural Systems (MLCAS2023). <https://github.com/UTokyo-FieldPhenomics-Lab/mlcas2023.github.io>
3. lab.fieldphenomics.com - The official website for Laboratory of Field Phenomics at UTokyo. <https://github.com/UTokyo-FieldPhenomics-Lab/utokyo-fieldphenomics-lab.github.io>
4. mlcas2021.github.io - The website for Third International Workshop on Machine Learning for Cyber-Agricultural Systems (MLCAS2021). <https://github.com/mlcas2021/mlcas2021.github.io>