Research Article

A comparative study of plant phenotyping workflows based on three-dimensional reconstruction from multi-view images

Someno Daiki[1], Koji Noshita[1*]

[1]Department of Biology, Kyushu University, Fukuoka, Fukuoka 819–0395, Japan

* **Corresponding author**: Koji Noshita

Affiliation: Kyushu University

Address: Fukuoka, Fukuoka 819–0395, Japan

Email: noshita@morphometrics.jp, Phone: +81-92-802-4316

Abstract

With the world facing escalating food demand, limited agricultural land, and rapid environmental change, there is a growing need for data-driven sustainable agricultural management approaches. The proliferation of next-generation sequencers and sensor networks has reduced the cost of acquiring genomic and environmental data, respectively. However, collecting phenotypic data, which is crucial for monitoring plant growth trajectories and detecting pests and diseases, continues to be a labor-intensive endeavor. Technological advances have enabled an efficient collection of three-dimensional (3D) data, yet this process currently involves a set of intricate steps. Therefore, the development of an effective phenotyping method is essential. In this study, we developed a phenotyping process based on 3D reconstruction, including mask image generation using deep neural network models, 3D reconstruction using the Structure from Motion/Multi-View Stereo (SfM/MVS) pipeline, and leaf surface reconstruction for leaf area estimation. Our investigation using soybean datasets into the optimal magnification for input images in mask generation models revealed that a 1/5.4× magnification was most effective for segmenting thin structures. Using mask images in the SfM/MVS pipeline limited the region of interest, and this could decrease the processing time and improve the quality of the point-cloud data. We assessed four scenarios regarding mask image usage and found that the scenario that set the mask images

of soybeans and stages before SfM and soybeans only after SfM yielded the highest-quality point-cloud data and was the second fastest in processing. Finally, we compared the Poisson reconstruction and B-spline surface fitting for leaf surface reconstruction from point clouds. B-spline fitting shows a greater correlation with destructive measurements. We proposed an optimal workflow for estimating leaf area based on these results. Additionally, to contribute to the development of plant phenotyping methods, we provided a web application for mask generation, a command-line tool for leaf surface reconstruction, and validation datasets.

Keywords: photogrammetry, multi-view image, phenomics, morphometrics, open science

45 # Introduction

46 The rise in global food demand [1–3], limited agricultural land [4–6], and climate change
47 [7–9] have heightened the necessity for the quantitative evaluation of crop phenotypic
48 information [10–12]. This approach would facilitate the development of efficient and robust
49 agricultural management approaches, which are essential for addressing the abovementioned
50 challenges in a data-driven manner [13,14]. Recent technological breakthroughs, particularly
51 in next-generation sequencing, have facilitated the accumulation of omics data, including
52 genomic and transcriptomic information [15–17]. In addition, advancements in Internet of
53 Things (IoT) sensor technologies have improved the acquisition of environmental data
54 [18,19]. However, despite the significance of phenotypic information in understanding plant
55 growth and physiological responses in various environments [20–22], the acquisition of
56 phenotypic data remains labor-intensive and yields limited data [23–25]. Consequently,
57 there has been increasing interest in plant phenotyping technologies among researchers and
58 agricultural producers [11,12,24–26]. Several studies have reported that the monitoring of
59 three-dimensional morphological properties of plants provides insights for improving
60 cultivation management, including optimizing the timing of pesticide and fertilizer
61 application [27–29].

62 High-resolution 3D morphological data can be acquired efficiently and cost-

effectively using 3D laser scanners [30–32], depth cameras [33–35], and photogrammetry [36–38]. A pipeline using structure from motion (SfM) [39,40] and multi-view stereo (MVS) [41,42], which reconstructs a 3D surface as point cloud data from a set of two-dimensional (2D) images captured from different angles, is widely available and implemented in several libraries and software products (e.g., [39,42–44]). Based on the resultant high-resolution point cloud data, a three-dimensional morphological model can be constructed to yield phenotypic data (e.g., leaf number, plant height, and leaf area index (LAI) [45–47]).

Optimizing the measurement process for three-dimensional morphological data directly facilitates the efficient acquisition of phenotypic data. Various approaches have been proposed for the acquisition of high-quality three-dimensional point cloud data using the SfM/MVS pipeline (e.g., [48–50]). Mask images offer several advantages in the SfM/MVS pipeline, such as reduced processing time by limiting the region of interest (ROI) and suppressing noise derived from background textures [51–53]. Moreover, there are several choices for creating morphological models for estimating phenotypic data from point cloud data acquired using SfM/MVS and 3D laser scanners, depending on the assumptions (e.g., target traits, precision requirements, and data acquisition conditions). Several methods have been proposed to provide robust quantification of leaf morphology, such as modeling the leaf surface as a triangular mesh [54], a set of small patches [55], and a piecewise polynomial

function [56]. Although Poisson surface reconstruction [57] is often used to reconstruct the target surface based on point cloud data in various fields [58], it is not ideal for leaf surface reconstruction [59,60]. This is because it assumes that the target surface is closed (i.e., topologically equivalent to a sphere). Although various techniques for improving the phenotyping processes have been proposed, as noted above, it is difficult to say that methods for optimizing each step in the process or the entire process for the target trait are well established.

In this study, we aimed to optimize the processes for acquiring phenotypic information based on 3D morphological data of plants. Through comparative analyses of various procedures on individual soybean plants, including mask image generation, SfM/MVS pipelines, and leaf surface reconstruction to estimate the leaf area (Fig. 1), we formulated an optimal phenotyping process. To generate mask images using a DNN-based semantic segmentation model, we proposed an ideal scale of input images for the two types of mask images. Within the SfM/MVS pipeline, we compared the processing time and reconstructed point-cloud data quality for the four mask image usage scenarios. Finally, we evaluated the leaf surface reconstruction methods to estimate the area of each leaf by comparing the two general methods. We propose an optimal 3D reconstruction-based plant phenotyping process and provide tools and data to accelerate these comparisons. Such

comprehensive comparisons for optimizing plant phenotyping processes can also be applied to other phenotyping procedures. Overall, these optimization processes can yield insights that contribute to the improvement of crop yield, quality, and the cultivation of various target crops.

Fig 1. A phenotyping workflow based on the 3D reconstruction. From multi-view images of a target plant individual, mask images are generated for the (A) SfM/MVS pipeline (B). The 3D morphological data is obtained as point cloud data using SfM/MVS pipeline. Several phenotypic values (e.g., leaf area) are estimated through morphological modeling (C).

Materials and Methods

Multi-view image datasets of individual soybeans

We used three multi-view image datasets of soybean (*Glycine max*) plants to train the DNN-based semantic segmentation model ($dataset_t$), evaluate different SfM/MVS scenarios based on the quality of the reconstructed point cloud data and their processing time ($dataset_m$), and evaluate the accuracy of leaf area estimation with two different morphological modeling methods against destructive measurements ($dataset_v$).

The $dataset_t$ consisted of four individuals from four different cultivars (Enrei, Zairai

117 51-2, Aoakimame, and Saga zairai), with each individual represented using 256 or 264
118 multiview images. All images were annotated in a pixel-wise manner across four classes
119 (soybean, stake, stage/pot, and background) using Labelbox (Labelbox, Inc., San Francisco,
120 CA, USA), a web application for generating training data.

121 The $dataset_m$ consisted of six individuals of three cultivars, namely, Seita, N2295,
122 and Keumdu, across different growth stages (4, 8, and 12 weeks after sowing, respectively).
123 The cultivars each included two replicates, with each individual represented using 256, 264,
124 or 272 multi-view images (Table 1).

125

126 **Table 1. Specimen list in the $dataset_m$.**

127

128 The $dataset_v$ consisted of 256 or 264 images per soybean individual obtained from
129 two cultivars (Hime daizu and Shiro mitsu mame) at two growth stages (6 and 12 weeks
130 after sowing) with two replicates. After acquiring multi-view images, all leaves were
131 sampled from individual plants, and their positions were recorded. The sampled leaves were
132 captured as two-dimensional images at 600 dots per inch (DPI) using a flatbed scanner
133 (CanoScan LiDE 400; Canon, Tokyo, Japan). The region of interest (ROI) for each leaf was
134 set semi-manually using the Selection Brush Tool in Affinity Photo 2 (Serif (Europe) Ltd.,

135 Nottingham, UK) to generate mask images for destructive leaf area measurements. These
136 mask images were processed using OpenCV via Python; specifically, this processing
137 involved binarization and the counting of pixels corresponding to each leaf for leaf area
138 measurements. The leaves of Shiro mitsu mame tended to have a smaller area
139 (approximately 1000 mm^2) than those of Hime daizu (1000–3000 mm^2).

140 All multiview images were obtained using a simple photogrammetry studio
141 consisting of a rotating table (MT320RL40, ComXim, Shenzhen, China), eight digital
142 single-lens reflex (DSLR) cameras (EOS Rebel SL1, Canon, Tokyo, Japan), coded markers,
143 and blue screen-like background sheets (S1 Fig). The focal length of each camera was set to
144 24 mm, and the images were captured approximately every 11° until the rotating table
145 reached 360 °(256, 264, or 272 images per soybean individual).

146

147 Mask image generation using U-Net

148 U-Net [61], a DNN model for semantic segmentation, was used to generate the mask images
149 in the SfM/MVS pipeline. The U-Net was trained for two classification tasks: three-
150 (soybean, stake, and background) and four-class classification (soybean, stake, stage, and
151 background) (Fig 2A). These different classes of mask images can be used across different
152 SfM/MVS scenarios (see **SfM/MVS pipelines** for details).

Fig 2. Mask image generation using U-Net model. Results of the semantic segmentation using the U-Net models for (A) three-class segmentation (soybean, stake, and background) and (B) four-class segmentation (soybean, stake, stage, and background) on an individual soybean (Saga zairai). Ground truth data was generated by manual annotation. Each color in the ground truth and the prediction corresponds to each class. (C) Each image in multi-view images was divided and resized into multiple 512×512 tiles with the six different magnifications (1/2×, 1/3×, 1/4×, 1/5.4×, 1/6× and 1/6.75×), for input images to the U-Net models. Lower magnification input image contains larger region (lower row). (D) Soybean mask images (foreground: soybean, background: stake and background) and soybean/stage mask images (foreground: soybean and stage, background: stake and background) were generated based on the predictions of three- and four-class segmentation models, respectively. The examples were mask images of a soybean individual (N2295 #1).

We used U-Net with ResNet34 [62] pre-trained on ImageNet as the encoder block and the softmax function as the activation function on the last layer. To train and predict using high-resolution images, each input image was divided into a certain number of tiles, depending on magnification, and resized to 512 × 512 pixels; each tile was segmented into

171 three or four classes and merged into an entire image. An overlap-tile strategy that excluded
172 outer pixels that did not exceed a threshold during merging was adopted to avoid border
173 effects in each tile (Fig 2B). Class-weighted dice loss was used as a loss function, and we
174 adopted the weights calculated based on the occurrences of the pixels corresponding to each
175 class in the training data (Table 2). The number of epochs is set to 200. Adam, with a
176 learning rate of 0.0001, was initially used as the optimizer; this rate was then changed to
177 0.00001 after 150 epochs.
178
179 **Table 2. Cross-validations of the U-Net models for three- and four-class segmentation.**
180 For each fold, different values of class weights were used for calculating class-weighted
181 Dice loss.
182
183 Performance, i.e., intersection over union (IoU) and processing time, of U-Net, were
184 evaluated across six different magnifications of the input images (1/2×, 1/3×, 1/4×, 1/5.4×,
185 1/6× and 1/6.75×) (Fig 3); each image was divided into 30, 15, 6, 4, 4, and 2 tiles on the
186 magnifications 1/2×, 1/3×, 1/4×, 1/5.4×, 1/6× and 1/6.75×, respectively (Fig 2B). We
187 conducted a group four-fold cross-validation on a $dataset_t$; images belonging to a single
188 individual were excluded as test images, and other images were randomly divided into

training (80%) and validation (20%) images. In the training step, we used the Python package Albumentations [63] for data augmentation, including random resizing and cropping (i.e., while retaining the aspect ratio of the original images), geometrical transformation (i.e., vertical and horizontal flipping, and random rotation by 90° zero or more times), changing color, elastic transformation, and random noise (i.e., Gaussian noising, blurring, motion blurring, and color jittering). The abovementioned steps were conducted using a CPU (Intel(R) Core(TM) i7-6950X CPU @ 3.00GHz) and a GPU GeForce GTX 1080 Ti ×2.

Fig 3. The performance of U-Net models on different magnifications of input image. The IoU scores three-class segmentation for (A) all classes, and each class; (B) soybean, (C) stake, and (D) background. (E) The processing time of three-class segmentation for an individual plant. The IoU scores four-class segmentation for (F) all classes, and each class; (G) soybean, (H) stake, (I) stage, and (J) background. (K) The processing time of four-class segmentation for an individual soybean.

In subsequent analyses, we generated mask images using the U-Net model trained on the entire $dataset_t$.

Web application for generating mask images

We developed a web application for mask generation based on the trained U-Net models (Fig 4). This application allows users to easily manipulate several parameters (scale, overlap, and mask type) through a simple user interface using Streamlit, which is a Python framework for developing web applications (https://streamlit.io). The web application, operating without a GPU, is hosted on the Streamlit Community Cloud (https://mask-prediction-app.streamlit.app/), and its source code will be made publicly available on the GitHub repository.

Fig 4. User interface of the web application for mask image generation. (1) Images for generating masks are uploaded via file uploader widget. User sets (2) input image size and (3) overlapping ratio, and (4) selects an image for showing an example for dividing. Finally, (5) one of the mask types is selected, and (6) start prediction. Once processing is complete, the mask images can be downloaded as a zip file.

SfM/MVS pipelines

We adopted the Structure from Motion (SfM) and Multi-View Stereo (MVS) methods to

generate 3D point cloud data. SfM is a photogrammetric method for simultaneously estimating the camera parameters and depth of the corresponding points (i.e., sparse point clouds) from multiview images. The SfM was conducted as per the following steps: (1) key points (characteristic points to be candidates of tie points) were obtained from each image in multi-view images using local features (e.g., SIFT [64], AKAZE [65], ORB [66]); (2) the tie points, which are corresponding points between images (based on key point matching and outlier removal with epipolar constraints), were selected; (3) an essential matrix for each image pair based on tie points was calculated, and then the extrinsic camera parameters (position, orientation) and the depth of the tie points (sparse point cloud) were estimated; finally, (4) the bundle adjustment was conducted to optimize the extrinsic camera parameters and the depth of the tie points by minimizing the reprojection error. After SfM, the MVS generates dense point-cloud data based on the estimated camera parameters and tie points. Several algorithms have been proposed, e.g., SUrface REconstruction from imagery (SURE) [67] and patch-based multiview stereo (PMVS) [68]. We used Metashape v1.8.4 (Agisoft, St. Petersburg, Russia), to conduct the SfM/MVS pipeline. Metashape adopts a SURE-like algorithm to implement the MVS. SURE can offer a depth map, which is the distance from each pixel in a 2D image to the object or scene captured in each image using a semiglobal matching (SGM) -based algorithm [69], and convert the depth maps into 3D coordinates

243 using a fusion method based on geometric constraints [70].

244 We evaluated four scenarios based on SfM/MVS: (i) no mask, (ii) masks before
245 SfM, (iii) masks after SfM, and (iv) two masks (Fig 5). The basic procedure for the
246 SfM/MVS implementation for each scenario was as follows: (1) multi-view images of a
247 soybean individual were imported, (2) the coded markers were detected, (3) scale bars were
248 created from the pairs of detected coded markers, (4) sparse point clouds were generated
249 based on SfM, with the maximum number of key points set to 200,000 and the maximum
250 number of tie points set to 80,000 in this study, (5) a depth map was created, and (6) dense
251 point clouds were generated based on MVS. In scenarios (ii) and (iii), the soybean mask
252 images were imported before and after step (4), respectively. In scenario (iv), the
253 soybean/stage mask images were imported before step (4) and replaced with the soybean
254 mask images before step (6). These four scenarios were implemented using the Metashape
255 Python API.

256

257 **Fig 5. The 3D reconstruction based on SfM/MVS pipeline with the four different mask**
258 **usage scenarios.** Each column corresponds to (i) no masks, (ii) masks before SfM, (iii)
259 masks after SfM, and (iv) two masks. The reconstructed point clouds of a soybean individual
260 (Keumdu #1) at 12 WAS (the second row), 8 WAS (the third row), and 4 WAS (the fourth

261 row). The black bars represent 20 cm. (A-L) Leaflets corresponding to red circles in the first

262 left column. (M-P) Key points and tie points detected in the four different scenarios.

263

264

265 In addition, the abovementioned four scenarios were evaluated for each soybean

266 individual of $dataset_m$. In the implemented script, the processing time for each step was

267 measured using the time function in the time module of the standard Python library. The

268 Mann-Whitney *U* test with Bonferroni correction was subsequently performed to investigate

269 the differences in processing times among the four scenarios.

270

271 Leaf surface reconstruction

272 To estimate phenotypic values from the reconstructed point-cloud data, we employed two

273 prevalent methods for leaf surface reconstruction: (1) Poisson surface reconstruction and (2)

274 B-spline surface fitting.

275 Poisson surface reconstruction [57] is widely used to estimate surfaces from point-

276 cloud data using their normal vectors (i.e., oriented points). Assume that we will reconstruct

277 the boundary ∂A of a target structure A from the given oriented points. If a function (called

278 *indicator function* χ), which returns one if the point is located inside of the A, otherwise 0,

is given, the boundary is approximated by its isosurface. The best-fitted boundary will be estimated by minimizing the differences between the gradient of the indicator function $\nabla\chi$ and the vector field defined by the oriented points $\vec{P}$ (i.e., $\min_{\chi} ||\nabla\chi - \vec{P}||$). To solve the problem, this equation is transformed into the Poisson equation as follows: $\Delta\chi = \nabla \cdot \vec{P}$. Finally, the isosurface is extracted from the estimated χ. Additionally, the reconstructed surface is filtered based on the point density (i.e., only parts of the surface supported by a sufficient number of points remain) because the method potentially reconstructs a surface in which points are distributed with low density, including the outer region of the leaf surface (i.e., extrapolation). In this study, we used the create_from_point_cloud_poisson function implemented in Open3D [71] to estimate the leaf surface from the point-cloud data using Poisson surface reconstruction. Vertices of the reconstructed surface were filtered if less than 6.5 points supported them.

B-spline surface fitting is widely used for surface reconstruction from point-cloud data. In this study, we adopted the B-spline surface fitting and B-spline curve trimming approaches [72] to reconstruct the leaf surface. A B-spline basis function of degree d is defined by the de Boor-Cox recursion formula [73]:

$$b_{i,0}(l) = \begin{cases} 1 & \text{if } \tau_i \leq l < \tau_{i+1} \\ 0 & \text{otherwise} \end{cases},$$

$$b_{i,d}(l) = \frac{l - \tau_i}{\tau_{i+d} - \tau_i} b_{i,d-1}(l) + \frac{\tau_{i+d+1} + l}{\tau_{i+d+1} - \tau_{i+1}} b_{i+1,d-1}(l),$$

where $(\tau_1, \tau_2, \cdots, \tau_{n+d+1})$ is a knot vector that divides the parameter space into subintervals, and n is the number of basis functions used for the approximation. A continuous surface is approximated by a linear combination of the tensor products of the B-spline basis functions, as follows:

$$\mathbf{p}_s(u,v) = \sum_{i}^{n_u} \sum_{j}^{n_v} \mathbf{w}_{ij} b_{i,d}^{u}(u) b_{j,d}^{v}(v) \,,$$

where $\mathbf{w}_{ij}$ is a coefficient vector called a control point, and $b_{i,d}^{u}(u)$ and $b_{j,d}^{v}(v)$ are B-spline basis functions of degree d along the independent parameters (u, v) of the parameter space. In this study, we parameterized the leaf surface by applying a principal component analysis (PCA), and the 2D parameter space was defined using the 1st and the 2nd PC axes. The $\mathbf{w}_{ij}$ is estimated by minimizing the sum of the squared distances between the B-spline surface and point cloud data. Based on the reconstructed surface, the parameter space is trimmed using a closed B-spline curve to exclude the outer region of the leaf. The closed B-spline curve was fitted to the point cloud data projected into the parameter space using the asymmetric distance with a penalty for concavity and error-adaptive knot insertion (EAKI) [72]. In this study, we used the surface/on_nurbs submodule [74] in the Point Cloud Library (PCL; [75]) to reconstruct leaf surfaces using B-spline surface fitting and B-spline curve trimming. In the B-spline surface fitting, we estimated the 3-degree B-spline surface by incrementally adding knots three times using the following parameter values: the surface

314 smoothness, boundary smoothness, and boundary weight were 0.8, 0.5, and 0.2, respectively.
315 In the B-spline curve trimming, the 3-degree B-spline curve was estimated with a maximum
316 of 100 iterations using the following parameter values for the penalty for concavity (0.2) and
317 for EAKI: iterations without knot insertion and the maximum number of control points were
318 4 and 100, respectively.

319 Based on both methods, leaf area was estimated using $dataset_v$ through scenario iv.
320 Each leaflet was cropped from the point-cloud data using Cloud Compare v2.12.4
321 (CloudCompare), software for visualizing and manipulating point-cloud data. We then
322 applied color thresholding (i.e., excluding points with $0.3 < hue < 0.7$, saturation < 0.2, and
323 value > 0.8 in the HSV color space) and statistical outlier removal (i.e., excluding outlier
324 points for which the mean distance calculated from their 30 neighbor points is greater than
325 three standard deviations by using the statistical_outlier_removal function in Open3D).
326

327 Results

328 Optimizing image magnification for mask image generation

329 We identified the optimal magnification of input images for multiview images of soybean
330 individuals by evaluating the IoU and processing time of segmentation.

331 For both three- and four-class segmentations, the IoU increased with decreasing

332 magnification, peaking at approximately 1/4×, followed by a gradual decrease (Fig 3A, F).
333 The IoU by class showed the same pattern (Fig 3B-D, G-J); however, several classes peaked
334 at 1/5.4×: stake (three-class segmentation) 0.8670, background (four-class segmentation)
335 0.9939, and stage (four-class segmentation) 0.9903 (Fig 3C, H-J). There is a tradeoff: as the
336 magnification of the input images decreases, a single input tile to the U-Net covers larger
337 plant structures; conversely, as the magnification increases, the resolution decreases. In our
338 dataset, a magnification of 1/4× or 1/5.4× reached the optimal IoU.

339 The processing times for both the three- and four-class segmentations decreased
340 with decreasing magnification. This is because the total number of input tiles decreased (Fig
341 2C, 3E, and 3K). Although a marginal increase in processing time was shown to be linked
342 to reductions in the magnification (i.e., the expansion of the field of view), almost all the
343 processing times per image fall between 0.015–0.02 s (S2 Fig).

344 The U-Net models for the three- and four-class segmentations were trained using
345 the images of all four cultivars in the $dataset_t$ (see S3 Fig). These models generated high-
346 resolution mask images for SfM/MVS, in which the background was excluded (Fig 2D):
347 soybean mask images (soybean pixels were the foreground, whereas others were the
348 background) and soybean/stage mask images (soybean and stage pixels were the foreground,
349 whereas others were the background).

350 Moreover, we developed a web application for mask generation based on the trained
351 U-Net models. This application helps generate mask images with arbitrary scales and
352 overlapping thresholds via a simple user interface (Fig 4). The processing time was 126 s on
353 the GPU device for predicting 256 mask images of 5184 × 3456 pixels divided into two
354 columns and two rows of 2900 × 2900 pixel tiles (approximately 1/5.4×) with at least 10%
355 overlap using a four-class segmentation model.
356

357 Quality of reconstructed point cloud data through SfM/MVS with
358 four mask image usages

359 We reconstructed the dense point cloud data of six individuals of three cultivars at three
360 different growth stages of the $dataset_m$ based on SfM/MVS, following four possible
361 scenarios of mask image usage (Fig 5). Masking during SfM/MVS effectively suppressed
362 the noise derived from the background textures (Fig.5B-D, 5F-H, 5K, and 5L) compared to
363 scenario (i) without mask images (Fig. 5A, 5E, 5I), as noted in previous studies (e.g., [51–
364 53]).
365 The quality of the reconstructed point-cloud data varied depending on the mask type
366 and timing of the setting. Although in scenario (ii), setting soybean mask images before SfM
367 was able to suppress the noise derived from the background textures, the reconstructed points

368 along the leaf edges seemed inaccurate (Fig 5B, 5F) and often failed to reconstruct the leaf
369 surfaces for the small soybean plants (Fig 5J). These inaccurate reconstructions without
370 background texture-derived noise were a consequence of the low accuracy of camera
371 extrinsic parameter estimation in SfM, not using markers on the stage and corresponding
372 points on the pot, because the ROIs based on the mask images were limited to soybeans. In
373 scenario (ii), key points and tie points were located only in the region corresponding to
374 soybean (Fig 5N), and their detected numbers were lower than those in the other scenarios
375 (Table 3). There were few qualitative differences between scenarios (iii) and (iv), except for
376 the noise derived from the abaxial side of the leaves and the background textures (Fig 5C,
377 5D, 5G, 5H, 5K, and 5L). In both scenarios, leaf surfaces were well-reconstructed because
378 SfM precisely estimated extrinsic camera parameters using the markers on the stage and the
379 corresponding points on the pot.
380
381 **Table 3. The average numbers of key points and tie points.** The number of key points
382 and tie points detected in SfM/MVS pipeline were counted for $dataset_m$ across four scenarios.
383
384 **Evaluation of the throughput of SfM/MVS on four mask image**

385 **usages**

386 The processing times, including those for mask image generation, SfM, and MVS, were
387 measured for the four scenarios (Fig 6). Scenarios (ii), (iv), and (iii) achieved the first, second,
388 and third shortest processing times, respectively (Fig 6A). The decrease in the area of the
389 ROIs caused by setting mask images contributed to the processing time. The longest mean
390 processing time was 34.0 min for scenario (i), and the shortest one was 15.8 min for scenario
391 (ii).

392

393 **Fig 6. The processing times on the four different mask usage scenarios.** The processing
394 times of (A) total of mask image generation (MIG), SfM, and MVS, (B) MIG, (C) SfM, and
395 (D) MVS. Multiple comparison was conducted on the processing times among four scenarios
396 (Mann-Whitney *U* test, **: 0.001<p<=0.01, ***: 0.0001<p<=0.001, ****: p<=0.0001).

397

398 The processing time of mask image generation for scenario (iv) was 3.61 min,
399 approximately double that for 1.74 min of scenario (ii) and 1.74 min of scenario (iii) (Fig
400 6B), because two types of mask images (soybean and soybean/stage) were generated.

401 The processing times of SfM and MVS increased in the order of the ROIs, except
402 for scenarios (ii) and (iii) of MVS (Fig 6C and 6D). There was no significant difference

403 between the processing times of SfM in scenarios (i) and (iii), and the mean processing times
404 were ~5.4 min. The shortest mean processing time was 3.3 min for scenario (ii). The majority
405 of processing times came from the MVS; the longest mean processing time was 27.6 min in
406 scenario (i) and the shortest one was 9.6 min in scenario (ii). Although the ROIs for MVS
407 were the same between scenarios (ii) and (iii), scenario (ii) showed a shorter processing time
408 for MVS because fewer key points were detected in SfM in scenario (ii) (Table 3).

409 We adopted scenario (iv) for the following analysis because (1) it generated the
410 highest-quality point cloud data, and (2) its entire processing time was the second shortest.

411

412 Leaf surface reconstruction methods for evaluating phenotypic 413 values

414 We reconstructed the leaf surfaces of 190 leaflets from dataset_v, excluding dead leaves, using
415 both Poisson surface reconstruction and B-spline surface fitting, and then estimated their
416 leaf areas (S1 Table). The estimated leaf area had the following correlation coefficients with
417 the results of the destructive measurements: Poisson surface reconstruction ($\rho = 0.9123$; Fig
418 7A) and B-spline surface fitting ($\rho = 0.9943$; Fig 7B). The approach based on B-spline
419 surface fitting enabled accurate estimation of leaf area ($R^2 = 0.9871$; Fig 7B). Using the
420 Poisson surface reconstruction, the leaf area was overestimated, particularly for larger leaves.

421 Leaf surface reconstruction based on B-spline surface fitting achieved higher
422 accuracy than Poisson surface reconstruction for the leaf area estimation in almost all cases.
423 Both methods successfully reconstructed leaf surfaces and estimated areas on high-quality
424 point cloud data (Fig 7C); the Poisson surface reconstruction method enabled a more detailed
425 reconstruction of leaf surfaces, yet showed nearly equivalent performance in leaf area
426 estimation. Although the Poisson surface reconstruction occasionally produced the leaf
427 surface as dual layers for the point cloud data to reconstruct the abaxial side over the adaxial
428 side, the B-spline surface fitting consistently reconstructed a singular surface (Fig 7D).
429 However, when a single leaflet was reconstructed as multi-leaf-like point cloud data, neither
430 method could accurately reconstruct the surface, resulting in an overestimation of the area
431 (Fig 7E); we found four leaflets corresponding to this case (Table 4). Three of the four
432 leaflets were leaflets of the same individual plant, which implied blurring associated with
433 the motion of the target plants while capturing the multi-view images, which caused low
434 accuracy of the camera parameter estimation. This is different from the decrease in accuracy
435 in camera parameter estimation owing to the mask images observed for the small individuals
436 in scenario (ii) (Fig 5J); in fact, the quality of the reconstructed point cloud data is similar
437 across scenarios (ii) and (iv) (Fig 7F).
438

Fig 7. Leaf surface reconstruction and leaf area estimation. Scatter diagrams of leaf area estimated by (A) Poisson surface reconstruction and (B) B-spline surface fitting, against the destructive measurements. (C) Both methods estimated leaf area consistent with the destructive measurements (Hime daizu #7 S01_t). (D) Poisson surface reconstruction overestimated leaf area while B-spline surface fitting estimated accurately (Hime daizu #7 S07_t). (E) Both methods failed to reconstruct leaf surface well, then overestimated leaf area (Shiro mitsu mame #3 S06_r). (F) The reconstructed 3D point cloud data (Shiro mitsu mame #3 S07_t) based on scenario (i) and (iv).

Table 4. The four leaflets which were overestimated leaf area using B-spline surface fitting. The leaf areas of 190 leaflets were overestimated when B-spline surface fitting against the destructive measurements was used.

Discussion

Even though using mask images improves the quality of point cloud data reconstructed via the SfM/MVS pipeline and decreases the processing time, it remains crucial to understand which objects should be masked and in which operations these mask images should be applied. As noted in previous studies [51–53], the reconstructed point-cloud data based on

457 SfM/MVS were improved for large individuals by masking the target objects (Fig 5A-H).
458 However, the point cloud was not comprehensively reconstructed by setting the mask images
459 before SfM for small individuals (Fig 5J), although they were reconstructed properly without
460 mask images (Fig 5I). We found that appropriate point clouds were reconstructed either by
461 setting soybean mask images after SfM (scenario iii) or by setting soybean/stage mask
462 images before SfM and then soybean mask images after SfM (scenario iv) (Fig 5C, D, G, H,
463 K, and L). This implies the importance of setting mask images, which ensures a sufficient
464 number and accuracy of detected key points and tie points (Fig 5M-P, Table 3). There are
465 several considerations for generating the mask images used in this study, one of which is the
466 magnification rate of the input images. As a result of different scales of input images to the
467 U-Net, a magnification of 5.4× achieved the first or second optimal IoU for all classes (Fig
468 3). However, there was a trade-off between low magnification (i.e., long-range context is
469 available with low resolution) and high magnification (i.e., only local context is available
470 with high resolution) (Fig 2C). Setting a larger input image size could resolve this issue;
471 however, in practice, balancing them for the target objects is necessary because of various
472 constraints (e.g., available GPU memory).

473 In plant phenotyping, it is advantageous to introduce known morphological
474 constraints to the target shape. In this study, we assumed that a leaf constitutes a single

475 bounded surface in a three-dimensional space during leaf surface reconstruction, which
476 enabled more robust measurements of leaf area (Fig 7). Thus, the surface was reconstructed
477 using Poisson surface reconstruction; this approach resulted in an overestimation of leaf area
478 (Fig 7A). It is worth noting that point cloud data derived from actual plants often contain
479 noise that is difficult to eliminate, even with denoising filters. In cases where B-spline
480 surface fitting is ineffective for reconstruction, methods exclusively derived for
481 characterizing leaf morphology could prove useful (e.g., [54–56,59,60]). The results of leaf
482 area estimation using B-spline surface fitting showed a high correlation with the results of
483 destructive measurements (Fig 7B), indicating its utility in the phenotyping process. In some
484 cases, the B-spline surface fitting did not reconstruct precisely (Fig 7E). However, this issue
485 often stems from the difficulty underlying camera parameter estimation in SfM (Fig 7F),
486 demonstrating that improvements in measurement devices can yield improvements in
487 accuracy. Concerning the qualitatively noted shortcomings in previous studies of the B-
488 spline surface fitting approach in reconstructing leaf surfaces [59,60], it is considered that
489 these issues do not significantly impact leaf area estimation as long as the point cloud is
490 reconstructed with high accuracy. Consequently, we propose an optimal workflow for
491 estimating leaf area of soybean plants, including mask image generation using U-Net with
492 1/5.4× magnification, 3D point cloud reconstruction based on scenario (iv) of SfM/MVS,

493 and leaf surface reconstruction by B-spline surface fitting to estimate leaf area. The optimal
494 values and methods depend on the target objects and constraints on the measurement
495 processes, and the optimal configurations are explored through comparative experiments, as
496 demonstrated in this study. The reconstructed leaf surfaces also allow for quantification of
497 leaf orientation in three-dimensional space; the resultant outputs can be used for functional
498 assessments, including light interception [76,77]. Utilizing this method in functional
499 structural plant models (FSPM) and other areas is a subject for future research on sustainable
500 food systems using data-driven approaches [11,12,78].

501 Moreover, we offer a tool and data that can be applied toward enhancing the
502 efficiency at which plant phenotyping can be conducted. Specifically, our web application
503 facilitates the generation of mask images through a trained U-Net model, and is designed to
504 allow the replacement of models, enabling its use in various subjects (Fig 4). We also
505 published the multiview soybean images used in model training as open data, including their
506 mask images and images from destructive measurements. These data will be available to
507 developers of future methodologies as benchmark datasets. We believe that disseminating
508 such tools and datasets will contribute to the field of phenomics, where data acquisition is
509 currently labor-intensive and slow.

510

Conclusion

Here, we investigated 3D reconstruction-based phenotyping processes including mask image generation using U-Net models, 3D reconstruction based on the SfM/MVS pipeline, and leaf surface reconstruction for leaf area estimation. Through comparative experiments using soybean data, we proposed an optimal workflow for estimating leaf area and generating mask images. The workflow entailed using U-Net models with 1/5.4× magnification, 3D point cloud reconstruction based on the SfM/MVS pipeline setting soybean/stage mask images before SfM and soybean mask images after SfM (scenario iv), and leaf surface reconstruction by B-spline surface fitting to estimate leaf area. However, it is important to note that in some cases, greater accuracy was achieved with 1/4 magnification in the U-Net model, and that high-quality point clouds can also be obtained by setting only the soybean mask images after SfM (scenario iii). The processes outlined in this study can be adapted as required. Moreover, we offer a web application for mask generation, a command-line tool for leaf surface reconstruction, and validation datasets to facilitate future plant phenotyping research and methodological development.

Acknowledgements

We thank Horiguchi, R., Inbe, N., Suzuki, M., and Kudo, Y. for their assistance in making multi-view image datasets.

530

References

1. Tilman D, Balzer C, Hill J, Befort BL. Global food demand and the sustainable intensification of agriculture. Proc Natl Acad Sci U S A. 2011;108: 20260–20264. doi:10.1073/pnas.1116437108
2. van Dijk M, Morley T, Rau ML, Saghai Y. A meta-analysis of projected global food demand and population at risk of hunger for the period 2010–2050. Nat Food. 2021;2: 494–501. doi:10.1038/s43016-021-00322-9
3. Godfray HCJ, Beddington JR, Crute IR, Haddad L, Lawrence D, Muir JF, et al. Food Security: The Challenge of Feeding 9 Billion People. Science (1979). 2010;327: 812–818. doi:10.1126/science.1185383
4. Phalan B, Bertzky M, Butchart SHM, Donald PF, Scharlemann JPW, Stattersfield AJ, et al. Crop Expansion and Conservation Priorities in Tropical Countries. PLoS One. 2013;8. doi:10.1371/journal.pone.0051759
5. Wu W, Yu Q, You L, Chen K, Tang H, Liu J. Global cropping intensity gaps: Increasing food production without cropland expansion. Land use policy. 2018;76: 515–525. doi:10.1016/j.landusepol.2018.02.032
6. Potapov P, Turubanova S, Hansen MC, Tyukavina A, Zalles V, Khan A, et al. Global maps of cropland extent and change show accelerated cropland expansion in the twenty-first century. Nat Food. 2022;3: 19–28. doi:10.1038/s43016-021-00429-z
7. Zhao C, Liu B, Piao S, Wang X, Lobell DB, Huang Y, et al. Temperature increase reduces global yields of major crops in four independent estimates. Proc Natl Acad Sci U S A. 2017;114: 9326–9331. doi:10.1073/pnas.1701762114
8. Adams RM, Hurd BH, Lenhart S, Leary N. Effects of global climate change on agriculture: An interpretative review. Clim Res. 1999;11: 19–30. doi:10.3354/cr011019
9. Wang X, Zhao C, Müller C, Wang C, Ciais P, Janssens I, et al. Emergent constraint on crop yield response to warmer temperature from field experiments. Nat Sustain. 2020;3: 908–916. doi:10.1038/s41893-020-0569-7
10. Costa C, Schurr U, Loreto F, Menesatti P, Carpentier S. Plant Phenotyping Research Trends, a Science Mapping Approach. Front Plant Sci. 2019;9. doi:10.3389/fpls.2018.01933
11. Herrero M, Thornton PK, Mason-D'Croz D, Palmer J, Benton TG, Bodirsky BL, et al. Innovation can accelerate the transition towards a sustainable food system. Nat Food. 2020;1: 266–272. doi:10.1038/s43016-020-0074-1
12. Fiorani F, Schurr U. Future Scenarios for Plant Phenotyping. Annu Rev Plant Biol. 2013;64: 267–291. doi:10.1146/annurev-arplant-050312-120137

13. van Etten J, de Sousa K, Cairns JE, Dell'Acqua M, Fadda C, Guereña D, et al. Data-driven approaches can harness crop diversity to address heterogeneous needs for breeding products. Proc Natl Acad Sci U S A. 2023;120. doi:10.1073/pnas.2205771120
14. Coppens F, Wuyts N, Inzé D, Dhondt S. Unlocking the potential of plant phenotyping data through integration and data-driven approaches. Curr Opin Syst Biol. 2017;4: 58–63. doi:10.1016/j.coisb.2017.07.002
15. Mochida K, Shinozaki K. Advances in Omics and Bioinformatics Tools for Systems Analyses of Plant Functions. Plant Cell Physiol. 2011;52: 2017–2038. doi:10.1093/pcp/pcr153
16. Goodwin S, McPherson JD, McCombie WR. Coming of age: Ten years of next-generation sequencing technologies. Nat Rev Genet. 2016;17: 333–351. doi:10.1038/nrg.2016.49
17. Stark R, Grzelak M, Hadfield J. RNA sequencing: the teenage years. Nat Rev Genet. 2019;20: 631–656. doi:10.1038/s41576-019-0150-2
18. Sathyamoorthy S, Okafor N. Advances and Challenges in IoT Sensors Data Handling and Processing in Environmental Monitoring Systems. TechRxiv. 2023; 1–8. doi:10.36227/techrxiv.24045951.v1
19. Ullo SL, Sinha GR. Advances in Smart Environment Monitoring Systems Using IoT and Sensors. Sensors. 2020;20: 3113. doi:10.3390/s20113113
20. Granier C, Tardieu F. Multi-scale phenotyping of leaf expansion in response to environmental changes: The whole is more than the sum of parts. Plant Cell Environ. 2009;32: 1175–1184. doi:10.1111/j.1365-3040.2009.01955.x
21. Langstroff A, Heuermann MC, Stahl A, Junker A. Opportunities and limits of controlled-environment plant phenotyping for climate response traits. Theoretical and Applied Genetics. Springer Science and Business Media Deutschland GmbH; 2022. doi:10.1007/s00122-021-03892-1
22. Mir RR, Zaman-Allah M, Sreenivasulu N, Trethowan R, Varshney RK. Integrated genomics, physiology and breeding approaches for improving drought tolerance in crops. Theoretical and Applied Genetics. Springer Verlag; 2012. pp. 625–645. doi:10.1007/s00122-012-1904-9
23. Furbank RT, Tester M. Phenomics - technologies to relieve the phenotyping bottleneck. Trends Plant Sci. 2011;16: 635–644. doi:10.1016/j.tplants.2011.09.005
24. Araus JL, Cairns JE. Field high-throughput phenotyping: The new crop breeding frontier. Trends Plant Sci. 2014;19: 52–61. doi:10.1016/j.tplants.2013.09.008
25. Yang W, Feng H, Zhang X, Zhang J, Doonan JH, Batchelor WD, et al. Crop Phenomics and High-Throughput Phenotyping: Past Decades, Current Challenges,

and Future Perspectives. Mol Plant. 2020;13: 187–214. doi:10.1016/j.molp.2020.01.008

26. Campbell ZC, Acosta-Gamboa LM, Nepal N, Lorence A. Engineering plants for tomorrow: how high-throughput phenotyping is contributing to the development of better crops. Phytochemistry Reviews. 2018;17: 1329–1343. doi:10.1007/s11101-018-9585-x

27. Dammer K-H, Wollny J, Giebel A. Estimation of the Leaf Area Index in cereal crops for variable rate fungicide spraying. European Journal of Agronomy. 2008;28: 351–360. doi:10.1016/j.eja.2007.11.001

28. Addai IK, Alimiyawo M. Graphical determination of leaf area index and its relationship with growth and yield parameters of Sorghum (Sorghum bicolor L. Moench) as affected by fertilizer application. Journal of Agronomy. 2015;14: 272–278. doi:10.3923/ja.2015.272.278

29. Zhu H, Dorner JW, Rowland DL, Derksen RC, Ozkan HE. Spray Penetration into Peanut Canopies with Hydraulic Nozzle Tips. Biosyst Eng. 2004;87: 275–283. doi:10.1016/j.biosystemseng.2003.11.012

30. Paulus S, Dupuis J, Mahlein A-K, Kuhlmann H. Surface feature based classification of plant organs from 3D laserscanned point clouds for plant phenotyping. 2013. Available: http://www.biomedcentral.com/1471-2105/14/238http://www.biomedcentral.com/1471-2105/14/238

31. Paulus S, Behmann J, Mahlein A-K, Plümer L, Kuhlmann H. Low-Cost 3D Systems: Suitable Tools for Plant Phenotyping. Sensors. 2014;14: 3001–3018. doi:10.3390/s140203001

32. Lin Y. LiDAR: An important tool for next-generation phenotyping technology of high potential for plant phenomics? Comput Electron Agric. 2015;119: 61–73. doi:10.1016/j.compag.2015.10.011

33. Chéné Y, Rousseau D, Lucidarme P, Bertheloot J, Caffier V, Morel P, et al. On the use of depth camera for 3D phenotyping of entire plants. Comput Electron Agric. 2012;82: 122–127. doi:10.1016/j.compag.2011.12.007

34. Andújar D, Dorado J, Fernández-Quintanilla C, Ribeiro A. An Approach to the Use of Depth Cameras for Weed Volume Estimation. Sensors. 2016;16: 972. doi:10.3390/s16070972

35. Milella A, Marani R, Petitti A, Reina G. In-field high throughput grapevine phenotyping with a consumer-grade depth camera. Comput Electron Agric. 2019;156: 293–306. doi:10.1016/j.compag.2018.11.026

36. Andújar D, Calle M, Fernández-Quintanilla C, Ribeiro Á, Dorado J. Three-Dimensional Modeling of Weed Plants Using Low-Cost Photogrammetry. Sensors.

2018;18: 1077. doi:10.3390/s18041077

37. An N, Palmer CM, Baker RL, Markelz RJC, Ta J, Covington MF, et al. Plant high-throughput phenotyping using photogrammetry and imaging techniques to measure leaf length and rosette area. Comput Electron Agric. 2016;127: 376–394. doi:10.1016/j.compag.2016.04.002

38. Santos TT, Rodrigues GC. Flexible three-dimensional modeling of plants using low-resolution cameras and visual odometry. Mach Vis Appl. 2016;27: 695–707. doi:10.1007/s00138-015-0729-3

39. Schonberger JL, Frahm J-M. Structure-from-Motion Revisited. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE; 2016. pp. 4104–4113. doi:10.1109/CVPR.2016.445

40. Snavely N, Seitz SM, Szeliski R. Modeling the World from Internet Photo Collections. Int J Comput Vis. 2008;80: 189–210. doi:10.1007/s11263-007-0107-3

41. Seitz SM, Curless B, Diebel J, Scharstein D, Szeliski R. A Comparison and Evaluation of Multi-View Stereo Reconstruction Algorithms. 2006. Available: http://vision.middlebury.edu/mview.

42. Schönberger JL, Zheng E, Frahm JM, Pollefeys M. Pixelwise view selection for unstructured multi-view stereo. In: Leibe B, Matas J, Sebe N, Welling M, editors. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). Cham: Springer International Publishing; 2016. pp. 501–518. doi:10.1007/978-3-319-46487-9_31

43. Cernea D. OpenMVS: Multi-View Stereo Reconstruction Library. 2020.

44. Moulon P, Monasse P, Perrot R, Marlet R. OpenMVG: Open multiple view geometry. In: Kerautret B, Colom M, Monasse P, editors. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). Cham: Springer International Publishing; 2017. pp. 60–74. doi:10.1007/978-3-319-56414-2_5

45. Zhang Y, Teng P, Aono M, Shimizu Y, Hosoi F, Omasa K. 3D monitoring for plant growth parameters in field with a single camerby multi-view approach. Journal of Agricultural Meteorology. 2018;74: 129–139. doi:10.2480/agrmet.D-18-00013

46. Lu X, Ono E, Lu S, Zhang Y, Teng P, Aono M, et al. Reconstruction method and optimum range of camera-shooting angle for 3D plant modeling using a multi-camera photography system. Plant Methods. 2020;16: 118. doi:10.1186/s13007-020-00658-6

47. Zermas D, Morellas V, Mulla D, Papanikolopoulos N. Estimating the Leaf Area Index of crops through the evaluation of 3D models. IEEE International Conference on Intelligent Robots and Systems. IEEE; 2017. pp. 6155–6162.

doi:10.1109/IROS.2017.8206517

48. Koutsoudis A, Ioannakis G, Vidmar B, Arnaoutoglou F, Chamzas C. Using noise function-based patterns to enhance photogrammetric 3D reconstruction performance of featureless surfaces. J Cult Herit. 2015;16: 664–670. doi:10.1016/j.culher.2015.01.008

49. Zhao J, Monno Y, Okutomi M. Polarimetric Multi-View Inverse Rendering. IEEE Trans Pattern Anal Mach Intell. 2023;45: 8798–8812. doi:10.1109/TPAMI.2022.3232211

50. Pierrot-Deseilligny M, De Luca L, Remondino F. Automated Image-Based Procedures for Accurate Artifacts 3D Modeling and Orthoimage Generation. Geoinformatics FCE CTU. 2011;6: 291–299. doi:10.14311/gi.6.36

51. Nooruddin M, Rahman M. Improved 3D reconstruction for images having moving object using semantic image segmentation and binary masking. 4th International Conference on Electrical Engineering and Information and Communication Technology, iCEEiCT 2018. IEEE; 2018. pp. 32–37. doi:10.1109/CEEICT.2018.8628064

52. Tanabata T, Isobe S, Hayashi A, Kochi N. Development of a semi-automatic 3D modeling system for phenotyping morphological traits in plants. Proceedings: IECON 2018 - 44th Annual Conference of the IEEE Industrial Electronics Society. IEEE; 2018. pp. 5439–5444. doi:10.1109/IECON.2018.8592904

53. Kochi N, Tanabata T, Hayashi A, Isobe S. A 3D Shape-Measuring System for Assessing Strawberry Fruits. International Journal of Automation Technology. 2018;12: 395–404. doi:10.20965/ijat.2018.p0395

54. Dornbusch T, Wernecke P, Diepenbrock W. A method to extract morphological traits of plant organs from 3D point clouds as a database for an architectural plant model. Ecol Modell. 2007;200: 119–129. doi:10.1016/j.ecolmodel.2006.07.028

55. Pound MP, French AP, Murchie EH, Pridmore TP. Automated recovery of three-dimensional models of plant shoots from multiple color images. Plant Physiol. 2014;166: 1688–1698. doi:10.1104/pp.114.248971

56. Kempthorne DM, Turner IW, Belward JA, McCue SW, Barry M, Young J, et al. Surface reconstruction of wheat leaf morphology from three-dimensional scanned data. Functional Plant Biology. 2015;42: 444–451. doi:10.1071/FP14058

57. Kazhdan M, Bolitho M, Hoppe H. Poisson Surface Reconstruction. Eurographics Symposium on Geometry Processing. 2006.

58. Yang M Der, Chao CF, Huang KS, Lu LY, Chen YP. Image-based 3D scene reconstruction and exploration in augmented reality. Autom Constr. 2013;33: 48–60. doi:10.1016/j.autcon.2012.09.017

59. Zhu F, Thapa S, Gao T, Ge Y, Walia H, Yu H. 3D Reconstruction of Plant Leaves for High-Throughput Phenotyping. Proceedings - 2018 IEEE International Conference on Big Data, Big Data 2018. IEEE; 2018. pp. 4285–4293. doi:10.1109/BigData.2018.8622428
60. Ando R, Ozasa Y, Guo W. Robust Surface Reconstruction of Plant Leaves from 3D Point Clouds. Plant Phenomics. 2021;2021. doi:10.34133/2021/3184185
61. Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). 2015. pp. 234–241. doi:10.1007/978-3-319-24574-4_28
62. He K, Zhang X, Ren S, Sun J. Deep Residual Learning for Image Recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE; 2016. pp. 770–778. doi:10.1109/CVPR.2016.90
63. Buslaev A, Iglovikov VI, Khvedchenya E, Parinov A, Druzhinin M, Kalinin AA. Albumentations: Fast and flexible image augmentations. Information (Switzerland). 2020;11. doi:10.3390/info11020125
64. Lowe DG. Distinctive image features from scale-invariant keypoints. Int J Comput Vis. 2004;60: 91–110. doi:10.1023/B:VISI.0000029664.99615.94
65. Alcantarilla PF, Nuevo J, Bartoli A. Fast explicit diffusion for accelerated features in nonlinear scale spaces. BMVC 2013 - Electronic Proceedings of the British Machine Vision Conference 2013. British Machine Vision Association; 2013. pp. 13.1-13.11. doi:10.5244/C.27.13
66. Rublee E, Rabaud V, Konolige K, Bradski G. ORB: An efficient alternative to SIFT or SURF. Proceedings of the IEEE International Conference on Computer Vision. IEEE; 2011. pp. 2564–2571. doi:10.1109/ICCV.2011.6126544
67. Rothermel M, Wenzel K, Fritsch D, Haala N. SURE - Photogrammetric Surface Reconstruction from Imagery. Proceedings LC3D Workshop. Berlin; 2012.
68. Furukawa Y, Curless B, Seitz SM, Szeliski R. Towards internet-scale multi-view stereo. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE; 2010. pp. 1434–1441. doi:10.1109/CVPR.2010.5539802
69. Hirschmuller H. Stereo Processing by Semiglobal Matching and Mutual Information. IEEE Trans Pattern Anal Mach Intell. 2008;30: 328–341. doi:10.1109/TPAMI.2007.1166
70. Remondino F, Spera MG, Nocerino E, Menna F, Nex F. State of the art in high density image matching. The Photogrammetric Record. 2014;29: 144–166. doi:10.1111/phor.12063

71. Zhou Q-Y, Park J, Koltun V. Open3D: A Modern Library for 3D Data Processing. 2018. Available: http://arxiv.org/abs/1801.09847
72. Mörwald T, Balzer J, Vincze M. Modeling connected regions in arbitrary planar point clouds by robust B-spline approximation. Rob Auton Syst. 2016;76: 141–151. doi:10.1016/j.robot.2015.11.006
73. de Boor C. On calculating with B-splines. J Approx Theory. 1972;6: 50–62. doi:10.1016/0021-9045(72)90080-9
74. Mörwald T. Object modelling for cognitive robotics. Technische Universität Wien. 2013.
75. Rusu RB, Cousins S. 3D is here: Point Cloud Library (PCL). Proc IEEE Int Conf Robot Autom. 2011; 1–4. doi:10.1109/ICRA.2011.5980567
76. Pearcy RW, Muraoka H, Valladares F. Crown architecture in sun and shade environments: assessing function and trade-offs with a three-dimensional simulation model. New Phytologist. 2005;166: 791–800. doi:10.1111/j.1469-8137.2005.01328.x
77. Niinemets Ü. Photosynthesis and resource distribution through plant canopies. Plant Cell Environ. 2007;30: 1052–1071. doi:10.1111/j.1365-3040.2007.01683.x
78. Evers JB, Letort V, Renton M, Kang M. Computational botany: Advancing plant science through functional-structural plant modelling. Ann Bot. 2018;121: 767–772. doi:10.1093/aob/mcy050
79. Noshita K. A simple photogrammetry studio. figshare; 2023. doi:10.6084/m9.figshare.24920118.v2

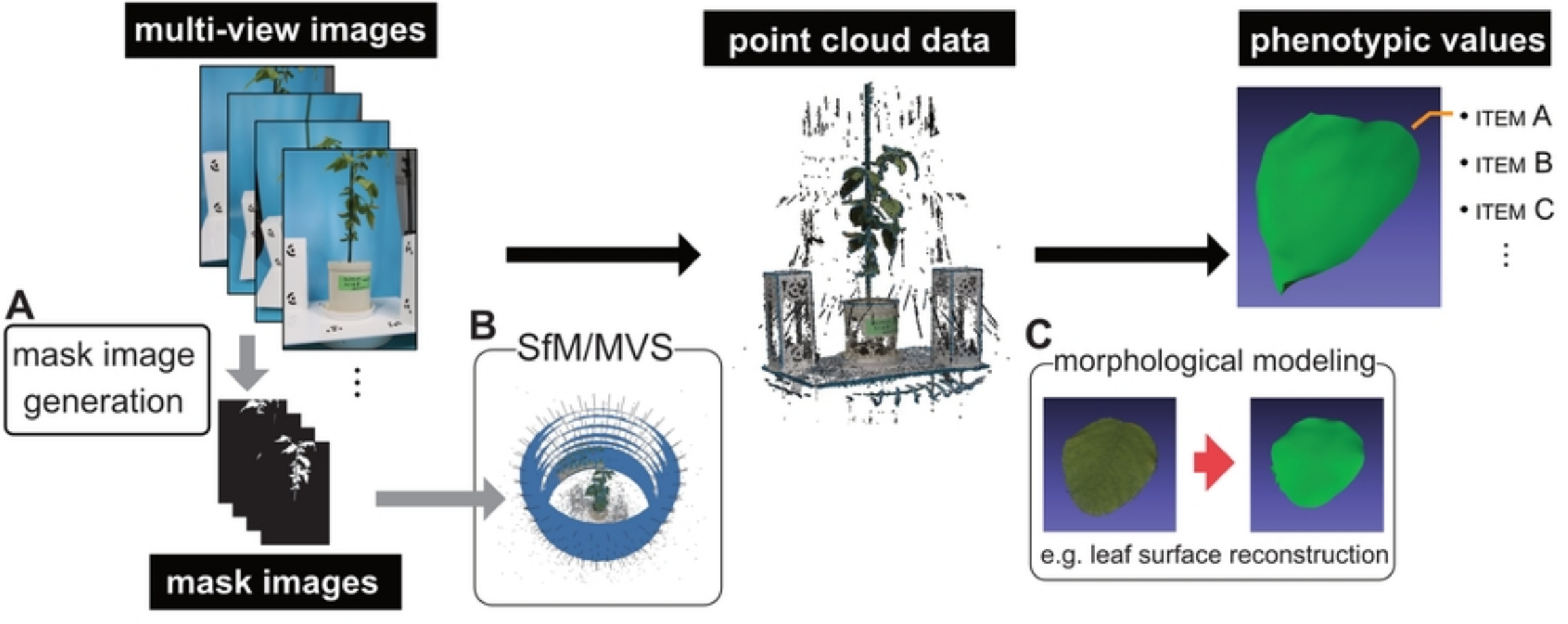
Supplementary Information

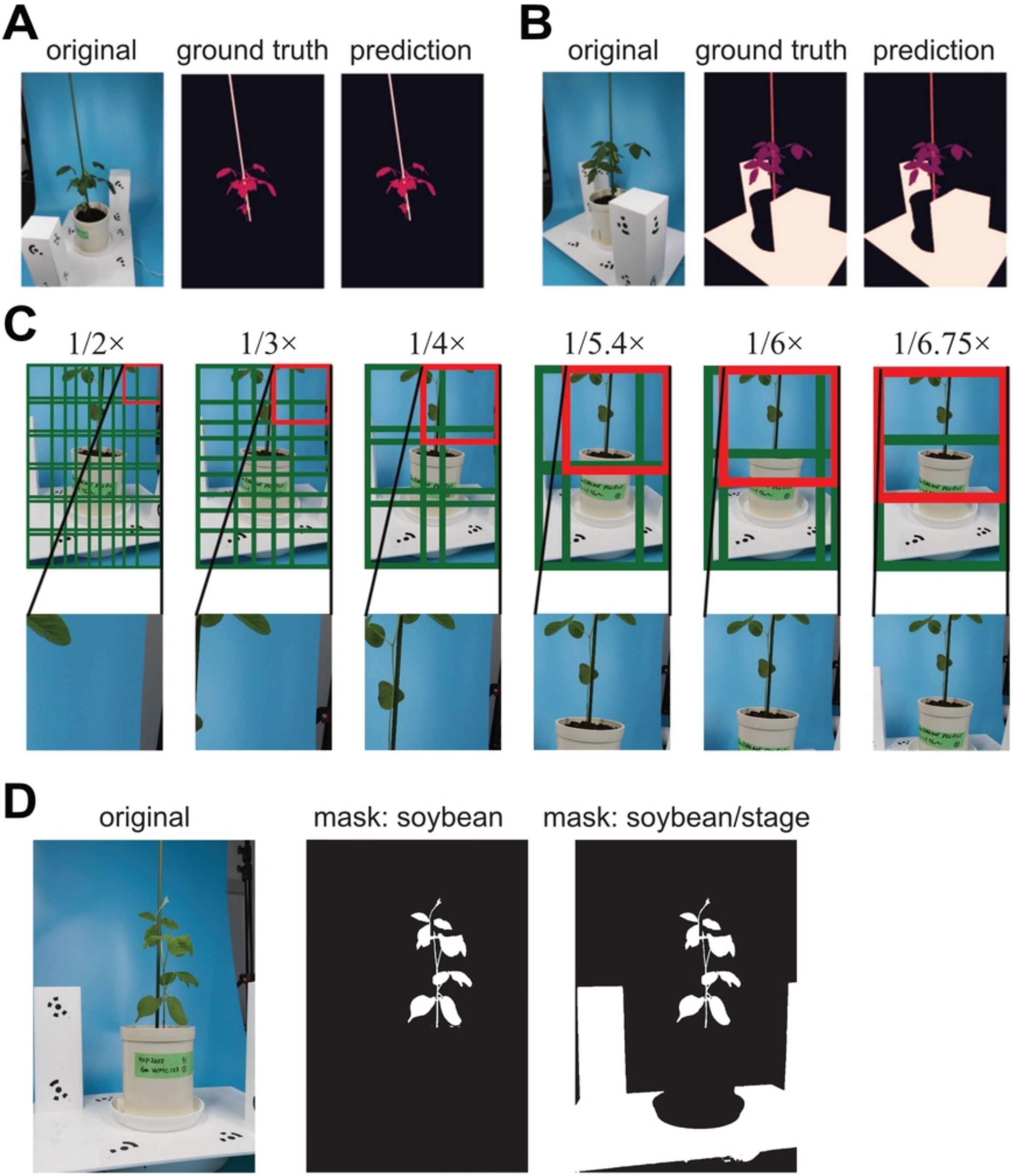
S1 Figure. Photogrammetry system. This figure was created based on [79] (Licensed under CC BY 4.0).

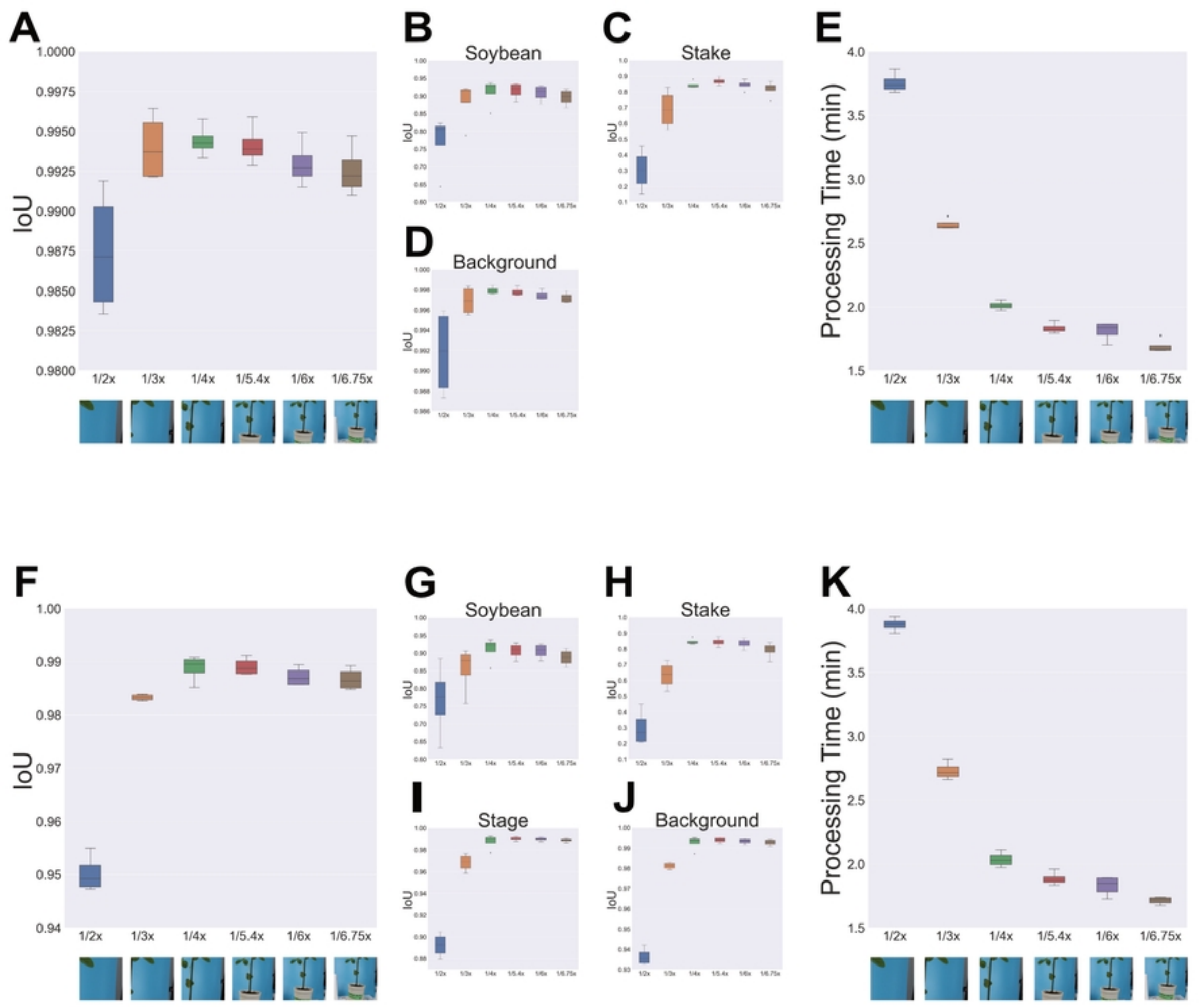
S2 Figure. Processing times of mask image generation per single 512×512 tile. Box plots of processing times among six magnifications of (a) three-class and (b) four-class segmentation.

S3 Figure. Time series data of class-weighted Dice loss and IoU of the U-Net models during training on the entire dataset$_t$. Line plots of class-weighted Dice losses of (a) three-class and (b) four-class segmentation models. Line plots of IoU scores of (c) three-class and (d) four-class segmentation.

S1 Table. Leaf area. The leaf area data of 190 leaflets (Hime daizu 87 leaves, Shiro mitsu mame 103 leaves) in dataset$_v$ were obtained by destructive measurement, Poisson surface reconstruction, and B-spline surface fitting. NA indicates missing values.


multi-view images
point cloud data
phenotypic values
A
mask image generation
mask images
B
SfM/MVS
C
morphological modeling
e.g. leaf surface reconstruction
• ITEM A
• ITEM B
• ITEM C


A
original
ground truth
prediction
B
original
ground truth
prediction
C
1/2×
1/3×
1/4×
1/5.4×
1/6×
1/6.75×
D
original
mask: soybean
mask: soybean/stage


A
IoU
1.0000
0.9975
0.9950
0.9925
0.9900
0.9875
0.9850
0.9825
0.9800
1/2x
1/3x
1/4x
1/5.4x
1/6x
1/6.75x
B
Soybean
C
Stake
D
Background
E
Processing Time (min)
4.0
3.5
3.0
2.5
2.0
1.5
F
1.00
0.99
0.98
0.97
0.96
0.95
0.94
G
Soybean
H
Stake
I
Stage
J
Background
K

Soybean Mask Prediction

Please upload your image files

Drag and drop files here
Limit 200MB per file • PNG, JPG, JPEG

① Browse files

SSP_032_00256_01_71074004861.JPG 6.1MB ×

SSP_032_00255_05_31074008483.JPG 5.9MB ×

SSP_032_00254_04_71074003363.JPG 6.1MB ×

Showing page 1 of 86

Uploading images 256

Input an integer number from 512 to 3456

② 512

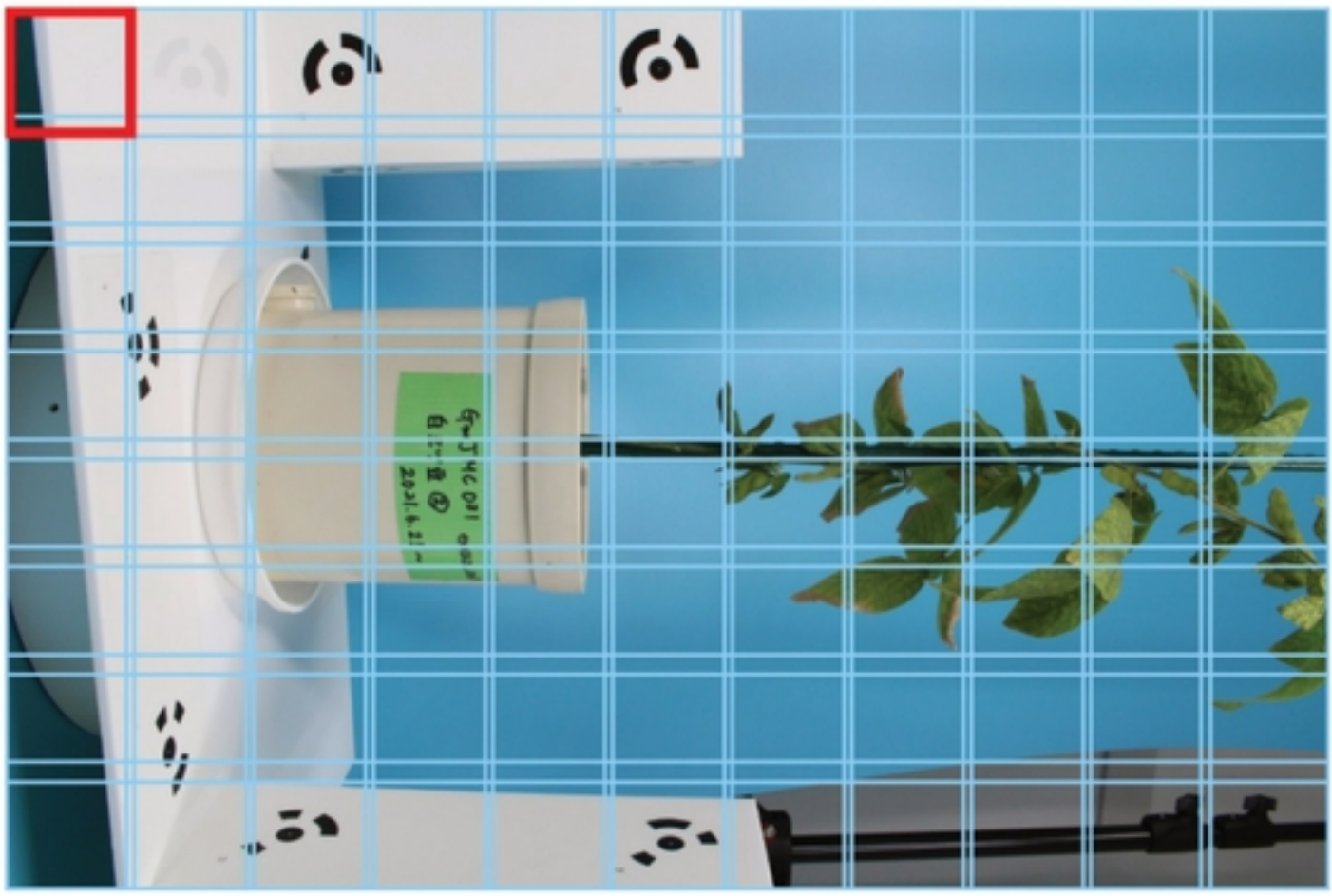
The input number is 512

Input an overlap ratio from 0.05 to 0.95

③ 0.05

Select an image of all you uploaded

④ 0



Each cropped image will be resized into 512x512

Image Name : SSP_001_00001_08_71074003452.JPG

Image Size : (width, height) = (5184 px, 3456 px)

The number of cropped images with blue color : n_column = 11, n_row = 8

The size of the cropped image with red color : (width, height) = (512 px, 512 px)

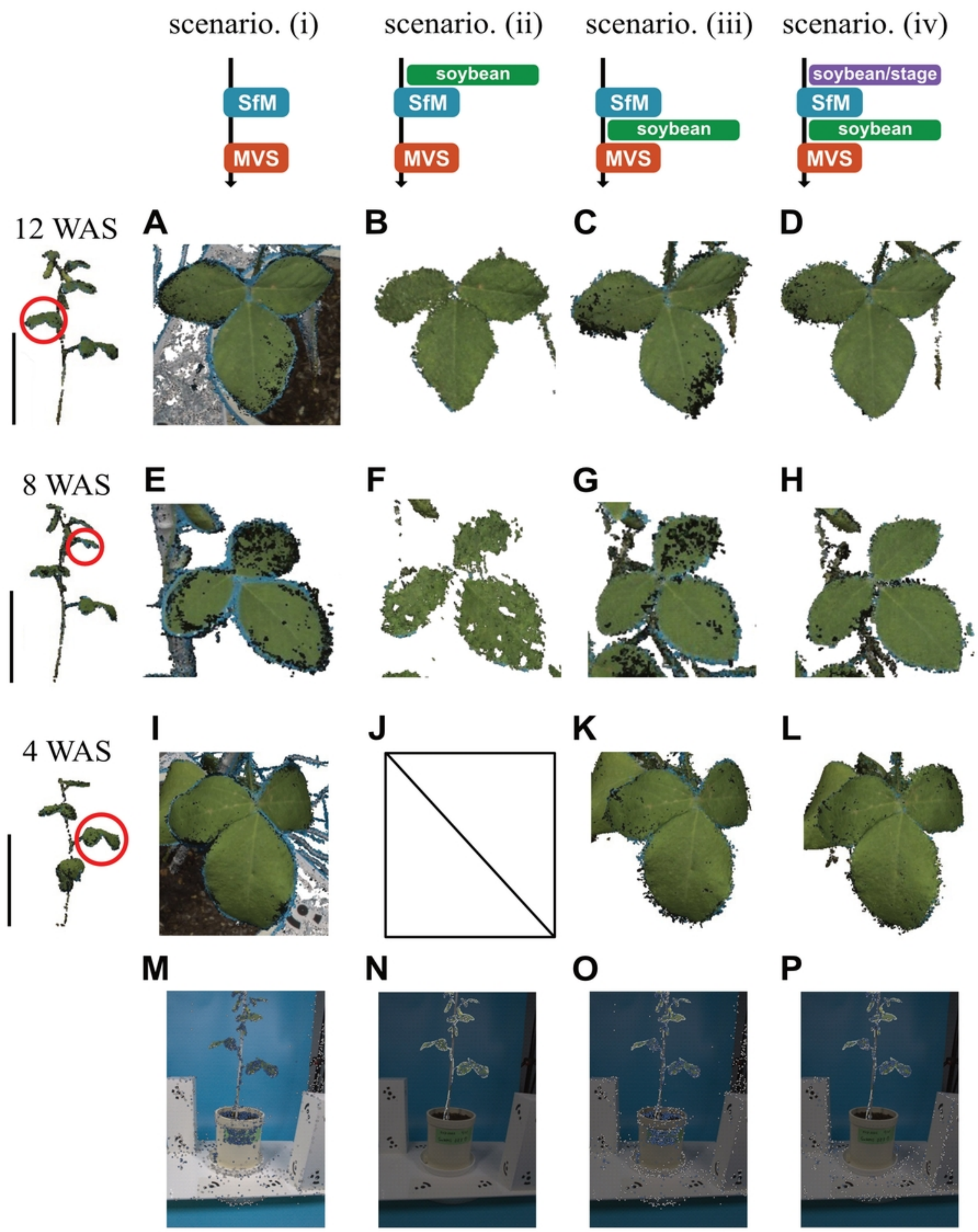
The overlapping value : horizontal direction = 44 px, vertical direction = 91 px

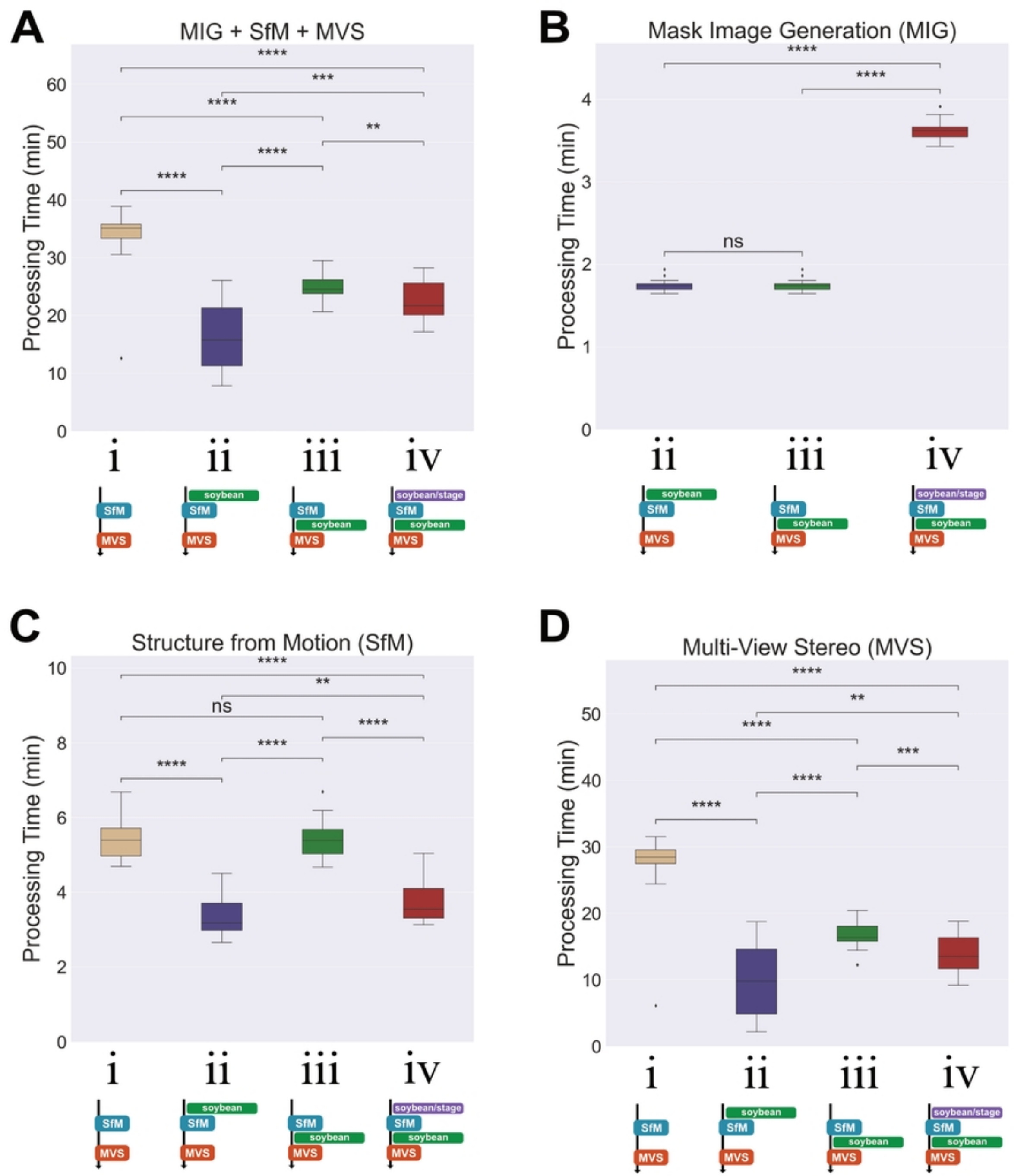
Please select which mask type you want to predict

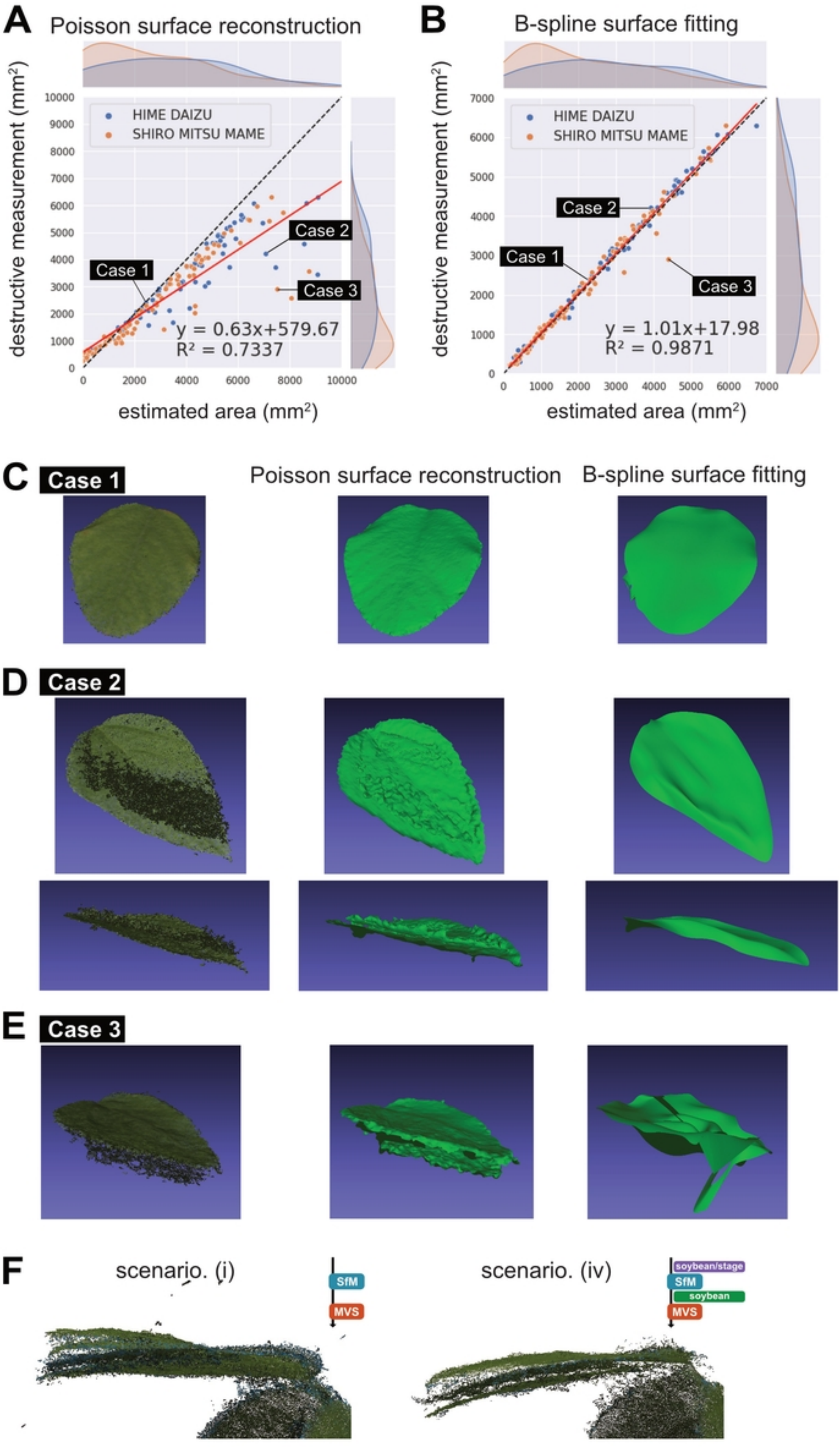
⑤ ◉ soybean
○ soybean & stage

You selected "soybean".

⑥ Start prediction


scenario. (i)
scenario. (ii)
scenario. (iii)
scenario. (iv)
SfM
MVS
soybean
SfM
MVS
SfM
soybean
MVS
soybean/stage
SfM
soybean
MVS
12 WAS
8 WAS
4 WAS
A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P


A
MIG + SfM + MVS
Processing Time (min)
0
10
20
30
40
50
60
****
***
****
**
****
****
i
ii
iii
iv
SfM
MVS
soybean
soybean/stage
B
Mask Image Generation (MIG)
Processing Time (min)
0
1
2
3
4
****
****
ns
ii
iii
iv
C
Structure from Motion (SfM)
Processing Time (min)
0
2
4
6
8
10
****
**
ns
****
****
****
i
ii
iii
iv
D
Multi-View Stereo (MVS)
Processing Time (min)
0
10
20
30
40
50
****
**
****
***
****
****
i
ii
iii
iv


A
Poisson surface reconstruction
HIME DAIZU
SHIRO MITSU MAME
Case 1
Case 2
Case 3
y = 0.63x+579.67
R² = 0.7337
destructive measurement (mm²)
estimated area (mm²)
B
B-spline surface fitting
HIME DAIZU
SHIRO MITSU MAME
Case 1
Case 2
Case 3
y = 1.01x+17.98
R² = 0.9871
destructive measurement (mm²)
estimated area (mm²)
C
Case 1
Poisson surface reconstruction
B-spline surface fitting
D
Case 2
E
Case 3
F
scenario. (i)
SfM
MVS
scenario. (iv)
soybean/stage
SfM
soybean
MVS

			Observation date		
Cultivar	**Individual num.**	**Sowing date**	**4 WAS**	**8 WAS**	**12 WAS**
Seita	1	5/24/2022	6/21/2022	7/19/2022	8/16/2022
Seita	2	5/24/2022	6/21/2022	7/19/2022	8/16/2022
N2295	1	6/1/2022	6/29/2022	7/27/2022	8/24/2022
N2295	2	6/1/2022	6/29/2022	7/27/2022	8/24/2022
Keumdu	1	6/29/2022	7/27/2022	8/24/2022	9/21/2022
Keumdu	2	6/29/2022	7/27/2022	8/24/2022	9/21/2022

Model	Fold num.	Train/Validation	Test	Weights			
				background	soybean	stake	stage
three-class segmentation	1	Zairai 51-2, Aoakimame, Saga zairai	Enrei	0.02	0.484	0.496	-
	2	Enrei, Aoakimame, Saga zairai	Zairai 51-2	0.019	0.486	0.496	-
	3	Enrei, Zairai 51-2, Saga zairai	Aoakimame	0.016	0.488	0.496	-
	4	Enrei, Zairai 51-2, Aoakimame	Saga zairai	0.019	0.486	0.495	-
four-class segmentation	1	Zairai 51-2, Aoakimame, Saga zairai	Enrei	0.098	0.323	0.33	0.249
	2	Enrei, Aoakimame, Saga zairai	Zairai 51-2	0.095	0.324	0.33	0.25
	3	Enrei, Zairai 51-2, Saga zairai	Aoakimame	0.094	0.326	0.33	0.25
	4	Enrei, Zairai 51-2, Aoakimame	Saga zairai	0.094	0.324	0.33	0.252

Scinario	Average number of points	
	Key points	Tie points
(i) No Masks	169,246	84,034
(ii) Masks before SfM	59,978	29,683
(iii) Masks after SfM	169,246	84,040
(iv) Two Masks	107,532	48,759

Cultivar Name	Individual Num.	Leaf ID	Leaf area (mm^2)	
			Destructive measurements	B-spline surface fitting
Hime daizu	7	S05_t	6287.910363	6741.43262
Shiro mitsu mame	3	S07_t	2568.184828	3214.35689
Shiro mitsu mame	3	S06_t	3557.466003	4076.095297
Shiro mitsu mame	3	S06_r	2898.617859	4396.995716