

Three-dimensional photogrammetry with deep learning instance segmentation to extract berry fruit harvestability traits

Xueping Ni ^{a,b}, Changying Li ^{b,d,*}, Huanyu Jiang ^{a,*}, Fumiomi Takeda ^c

^a College of Biosystems Engineering and Food Science, Zhejiang University, Hangzhou, China

^b College of Engineering, University of Georgia, Athens, GA, USA

^c Appalachian Fruit Research Station, USDA-ARS, Kearneysville, WV, USA

^d Phenomics and Plant Robotics Center, University of Georgia, Athens, GA, USA



ARTICLE INFO

Keywords:

3D reconstruction
Mask R-CNN
2D-3D projection
Blueberry traits

ABSTRACT

Fruit cluster characteristics such as compactness, maturity, berry number, and berry size, are important phenotypic traits associated with harvestability and yield of blueberry genotypes and can be used to monitor berry development and improve crop management. The goal of this study was to develop a complete framework of 3D segmentation for individual blueberries as they develop in clusters and to extract blueberry cluster traits. To achieve this goal, an image-capturing system was developed to capture blueberry images to facilitate 3D reconstruction and a 2D-3D projection-based photogrammetric pipeline was proposed to extract berry cluster traits. The reconstruction was performed for four southern highbush blueberry cultivars ('Emerald', 'Farthing', 'Meadowlark' and 'Star') with 10 cluster samples for each cultivar based on photogrammetry. A minimum bounding box was created to surround a 3D blueberry cluster to calculate compactness as the ratio of berry volume and minimum bounding box volume. Mask R-CNN was used to segment individual blueberries with the maturity property from 2D images and the instance masks were projected onto 3D point clouds to establish 2D-3D correspondences. The developed trait extraction algorithm was used to segment individual 3D blueberries to obtain berry number, individual berry volume, and berry maturity. Berry maturity was used to calculate cluster maturity as the ratio of the mature berry (blue colored fruit) number and the total berry (blue, reddish, and green colored fruit) number comprising the cluster. The accuracy of determining the fruit number in a cluster is 97.3%. The linear regression for cluster maturity has a R^2 of 0.908 with a RMSE of 0.068. The cluster berry volume has a RMSE of 2.92 cm³ compared with the ground truth, indicating that the individual berry volume has an error of less than 0.292 cm³ for clusters with a berry number greater than 10. The statistical analyses of the traits for the four cultivars reveals that, in the middle of April, 'Emerald' and 'Farthing' were more compact than 'Meadowlark' and 'Star', and the mature berry volume of 'Farthing' was greater than 'Emerald' and 'Meadowlark', while 'Star' had the smallest mature berry size. This study develops an effective method based on 3D photogrammetry and 2D instance segmentation that can determine blueberry cluster traits accurately from a large number of samples and can be used for fruit development monitoring, yield estimation, and harvest time prediction.

1. Introduction

Fruit traits derived from plant phenotyping offer useful data for monitoring growth stages and provide reference points for plant management (Kromdijk et al., 2014) and harvest strategies (Tran et al., 2017) for growers and plant breeders. Blueberries are an important economic fruit crop throughout the world (Brazelton and Strik, 2007). Blueberry fruit number and fruit size are the two key traits for yield estimation. The maturity of blueberries can predict the best time to

begin harvesting. Mechanical harvesting of blueberries (Takeda et al., 2013; Sargent et al., 2020) is necessary because the high cost of hand-harvesting and limited availability of workers have posed challenges for many blueberry growers producing fresh-pack fruit. Recently, the number of blueberry growers who harvest their fruit entirely with a machine or use a machine for second and third harvests has increased (DeVetter et al., 2019). However, with the cultivars that produce fruit on tight or compact clusters, the initial harvesting must be performed by hand picking. The reason is that mechanical harvesters are not capable

* Corresponding authors.

E-mail addresses: cyl@uga.edu (C. Li), hyjiang@zju.edu.cn (H. Jiang).

of selectively detaching mature berries if they are in a tight cluster along with immature (green and red) berries (Ballington et al., 1990). Thus, monitoring blueberry compactness can assist growers determining the best time to begin harvesting fruit with mechanical harvesters. The phenotyping of fruit traits (number counting, size, maturity and compactness) and integrating collected data for estimating yield and predicting harvest time are complicated, labor-intensive, and typically destructive (Tanksley, 2004), especially for specialty crops such as high-valued fresh market blueberries that have high costs associated with harvest operations. It has become important to develop an automated method to non-destructively measure fruit traits with optimal accuracy.

The recent development of convolutional neural networks (CNNs) (LeCun et al., 2015) has been successfully used in image classification, object detection, and semantic segmentation (LeCun et al., 1998, 2015). AlexNet is the first successful CNN architecture that was developed for image classification (Krizhevsky et al., 2012). VGG-16 increased the network depth based on AlexNet and a better performance was achieved. Subsequently, deeper networks VGG-19 (Simonyan and Zisserman, 2014) and GoogLeNet (Szegedy et al., 2015) were proposed for image classification, the latter of which reduced the computational time. To achieve object detection, an R-CNN architecture (Regions with CNN features) was proposed to generate and train 2000 regional proposals within an image (Girshick et al., 2014). Because of the large number of regional proposals, it took 84 h to train and 49 s to test each image. Thus, a Fast R-CNN model was proposed that used input images for training and then identified object regions from the feature map, which took 8.75 h to train and 2.3 s to test every image (Girshick, 2015). Based on the Fast R-CNN, a Faster R-CNN architecture was proposed to use two separated networks to identify object regions, and its performance for image testing was reduced to 0.2 s (Ren et al., 2015). In 2016, YOLO was proposed for high speed image streams capable of 45 frames per second (Redmon et al., 2016), although YOLO could not effectively detect small objects. In some applications, object detection has not been sufficient, while semantic segmentation is required to classify every pixel in an image. In contrast, fully convolutional networks (FCNs) provide one of the most popular semantic algorithms used for performing pixel-wise prediction and segmentation (Long et al., 2015). Furthermore, U-Net was developed based on FCN with a robust use of data augmentation. Based upon Faster R-CNN and FCN, Mask R-CNN was proposed for instance segmentation with reduced training time and better performance than FCN and it was largely considered the state-of-the-art CNN model for instance segmentation and has been widely used for benchmarking other models (He et al., 2017a). Unlike U-Net that segments all objects as a union mask (i.e., semantic segmentation), Mask R-CNN can predict objects as individual masks (i.e., instance segmentation) (Chen et al., 2017a), which is why Mask R-CNN was chosen in this study for blueberry detection and segmentation.

Deep learning methods have replaced traditional methods and have been used increasingly in fruit detection. In the past, traditional clustering methods (e.g., K-means or KNN clustering) have been used to detect fruits in 2D images based on color, shape, texture, and spectral properties (Wachs et al., 2010; Ji et al., 2012; Linker et al., 2012; Kurkulmus et al., 2014; Qiang et al., 2014). In recent years, deep learning methods have become more popular because they can achieve better performances than traditional methods and do not require hand-engineered features (LeCun et al., 2015). Sa et al. (2016) presented a sweet pepper detection method that used a Faster R-CNN model with RGB and NIR images and obtained a F1 score of 0.838. Six other fruits (rock melon, strawberry, apple, avocado, mango, and orange) also were trained and predicted with F1 scores of 0.848, 0.948, 0.938, 0.932, 0.942, and 0.915, respectively. Bargoti and Underwood (2017) used Faster R-CNN with a VGG-16 architecture to detect apples, mangoes, and almonds with F1 scores of 0.904, 0.908, and 0.775, respectively. In the same year, Chen et al. (2017b) used FCN to detect and segment oranges and green apples and achieved a mean Intersection over Union (mIOU) of 0.813 and 0.838, respectively. In 2018, more studies were

published detailing the application of CNNs to strawberries, citrus, mangoes, apples, and peaches, etc. (Habaragamuwa et al., 2018; Liang et al., 2018; Tao et al., 2018; Xiong et al., 2018; Zhang et al., 2020). Koirala et al. (2019) proposed ‘MangoYOLO’ to detect mangoes with higher speed and achieved a F1 score of 0.968 and an average precision of 0.983. Recent studies have used deep learning models for 2D object detection and tracking with video frames to solve challenging computer vision tasks of counting plant seedlings in the field (Jiang et al., 2019) and to quantify bruising by segmenting hyperspectral images of whole blueberry fruits (Zhang et al., 2020).

2D image-based object detection and segmentation has been investigated but may not provide sufficient information for some fruit trait extraction because of the occlusion and ambiguity caused by perspective projection (Chaudhury et al., 2018). As a result, 3D fruit detection has emerged in recent years. Initially, studies used traditional clustering methods (like K-Means) or machine learning methods (like SVM) to detect and segment fruit based on color, shape, depth, and 3D information (Jiang et al., 2013; Cupec et al., 2014; Barnea et al., 2016; Tao and Zhou, 2017; Lin et al., 2020). Similar methods also have been used to detect and segment plant leaves and stems (Wahabzada et al., 2015; Xia et al., 2015; Sodhi et al., 2017). The related traits were then extracted based on the detected or segmented 3D fruits. For example, He et al. (2017a) segmented strawberries in 2D images by combining them with 3D reconstruction using the structure from motion (SfM) technique to segment 3D calyx, body, and achenes of strawberries based on color, after which size, calyx size, berry volume, and achene number were extracted. Gongal et al. (2018) estimated apple sizes in 2D and 3D, and by comparison demonstrated that apple sizes gleaned from 2D were more accurate than 3D, which illustrated that 3D trait extraction techniques have not advanced sufficiently. Deep learning methods also have been utilized in combination with 3D reconstruction techniques to detect and segment 3D fruit. In 2016, a Faster R-CNN model was used to detect mangoes in 2D images, and epipolar lines from multiple images were used to locate and count mangoes (Stein et al., 2016). Ge et al. (2019) used Mask R-CNN to segment strawberries in 2D images and then localized fruit with a depth camera. Liu et al. (2018) and Gené-Mola et al. (2020) both used SfM for 3D reconstruction, but the former used FCN to detect fruit in 2D images, while the latter used Mask R-CNN to detect fruit in 2D images. Both studies localized fruits to perform the counting.

Blueberry fruits develop as clusters on branches. Each fruit has a short pedicel that is connected to the cluster axis. The number of blueberry fruits on a cluster can vary from several to more than 10. The cluster axis may be short, long, or branched. Terms such as small, large, loose, and tight have been used by blueberry breeders to describe blueberry clusters. A few grape counting methods have been proposed, as grapes are similar to blueberry fruit since both types of fruit are produced as clusters (horticulturists used the terms “bunch” and “cluster” in describing the inflorescence that berries develop). Rist et al. (2018) used a 3D scanner to perform 3D reconstruction of grape bunches and used a clustering method to classify point clouds. Then, a RANSAC sphere model was used to fit individual berries further for berry counting. However, the 3D reconstruction in the study highly depended on an expensive 3D scanner with unknown integrated algorithms, so, therefore, the process and the result cannot be easily reproduced. The counting accuracy was not high enough because they counted berries indirectly based on the fitted sphere result instead of the original point clouds, a process that increased the counting error. Another study used stereo images for depth generation and Mask R-CNN to detect individual grape berries in 2D images in which 2D-3D mapping and sphere fitting were used to count the berries (Nellithimaru and Kantor, 2019). The generated 3D points in the study only reflects the information from the camera facing side, which means the counted berry number was not the real berry number for the whole grape clusters. One of advantages of this study is that the combination of 3D information and 2D segmentation have the potential to improve berry counting accuracy compares to

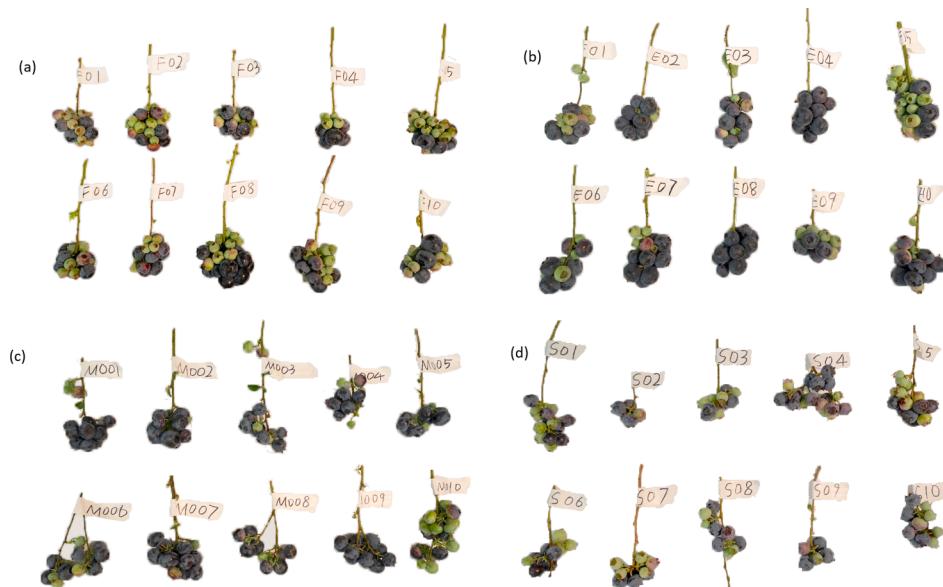


Fig. 1. Four cultivars of blueberry samples. (a) 'Emerald', (b) 'Farthing', (c) 'Meadowlark', (d) 'Star'. Total 40 samples with 10 samples for each cultivar.

earlier studies. For our study, we propose a method for counting berries based on the result of individual 3D berry segmentation. We used general digital cameras to collect image data for full 3D reconstruction and developed an algorithm to extract traits. The proposed method is expected to increase the counting accuracy compared to using sphere fitting to differentiate berries used in prior studies and enable quantitatively assessment of more fruit related traits (e.g., maturity,

compactness, and berry size) simultaneously.

The goal of this study was to develop a framework of 3D segmentation using photogrammetry and instance segmentation for individual blueberry fruits and to extract harvestability traits for fruit clusters. The specific objectives were to:

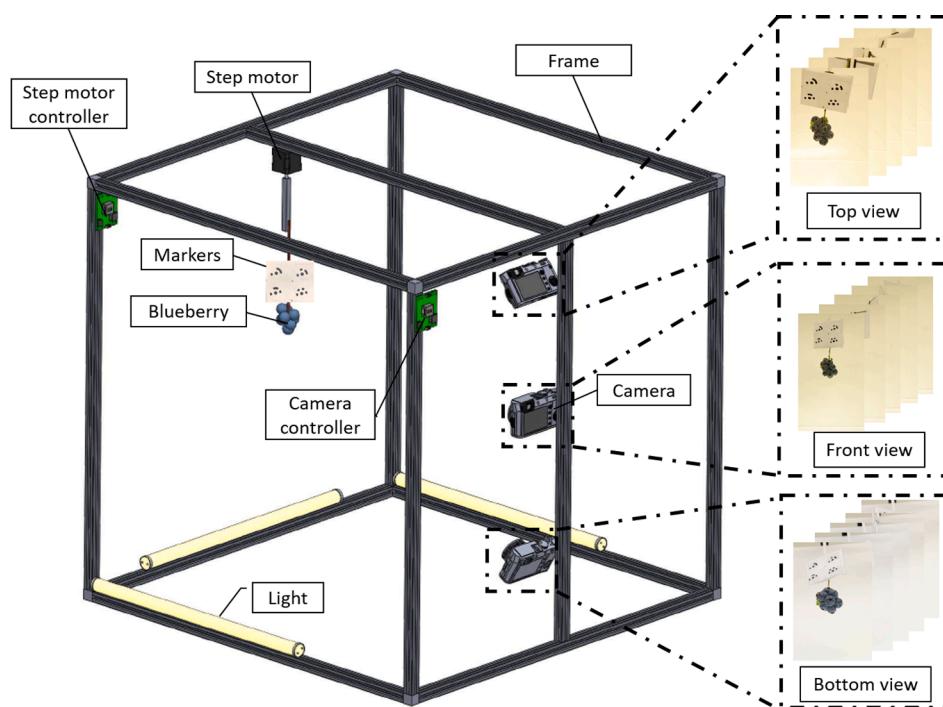


Fig. 2. Image capturing system to capture blueberry cluster information. 2D images captured by the three cameras are shown in top view, front view, and bottom view.

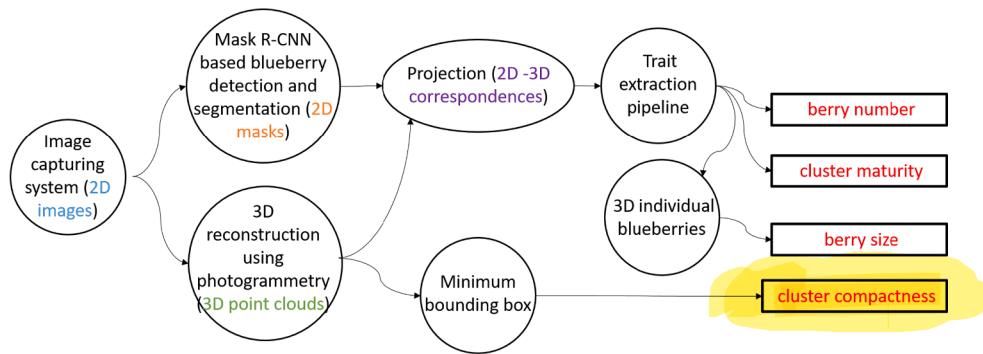


Fig. 3. Individual 3D blueberry segmentation and trait extraction pipeline.

- (1) Segment individual blueberries and classify their maturity levels in 2D images with a deep learning model.
- (2) Derive the 3D minimum bounding box to calculate fruit cluster compactness after 3D reconstruction.
- (3) Develop a trait extraction algorithm based on the 2D-3D projection to segment individual 3D blueberries, count berry number, calculate maturity, and estimate berry size.

2. Methods and materials

2.1. Data collection

All blueberry clusters for this study were collected in Alma, Georgia, in the southeastern United States (Coordinates: 31°32'30"N, 82°28'0"W) on April 17th, 2019. Four southern highbush blueberry cultivars ('Emerald', 'Farthing', 'Meadowlark', and 'Star'), which typically bloom from late February to early March, were used for this study. Their fruits mature unevenly, so typically the mature fruit is picked selectively three to five times in five- to seven-day intervals from mid-April to May. To extract accurate traits from the blueberry fruits and clusters, branches with at least one cluster were collected before they were picked commercially to guarantee the complete clusters. Fig. 1 shows all 40

blueberry clusters harvested from (a) 'Emerald', (b) 'Farthing', (c) 'Meadowlark', and (d) 'Star', with 10 samples taken for each cultivar. The samples were removed from the bushes and were transported in an ice box to the Bio-sensing and Instrumentation Lab at the University of Georgia, Athens, Georgia, and kept at 4°C. All leaves were removed prior to imaging.

2.2. Image capturing system

The 3D image capturing system consists of a step motor, a step motor controller, three digital cameras (X-A10, Fujifilm, Japan), a camera controller (Raspberry Pi), and three LED lights. Twelve aluminum extrusions were used to construct a frame on which all components were affixed. The cameras were mounted at three positions (top, front, and bottom). The front camera was used to capture an overall view of the blueberries. The cameras at the top and bottom positions were angled to obtain blueberry information to supplement the information captured by the front camera, and a camera controller developed in-house automatically triggered all three cameras simultaneously. Each of the 40 blueberry clusters were fixed to a rod attached to the step motor that rotated the cluster. The step motor was controlled by a step motor controller using a Python program that coordinated the camera

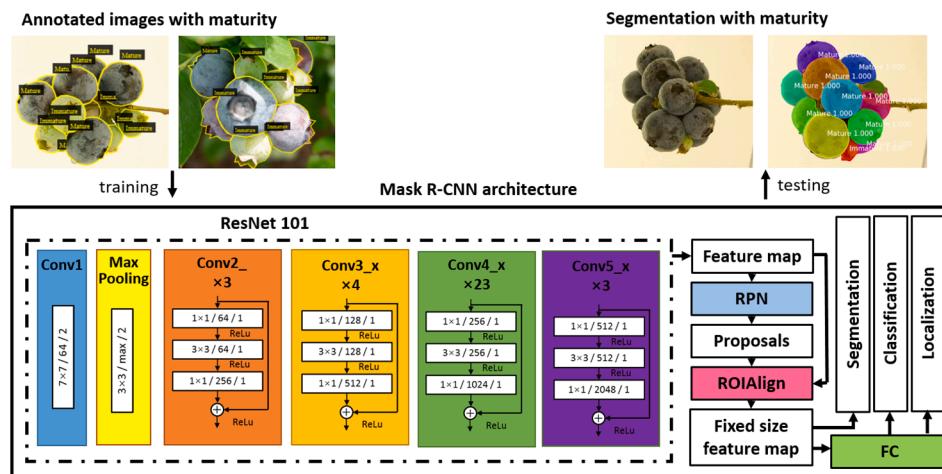


Fig. 4. Mask R-CNN based 2D blueberry instance segmentation. ResNet 101 was chosen as the backbone to extract features. A region proposal network (RPN) was used to generate regions of interest (ROIs) and ROIAlign was used to obtain the feature map with fixed size. Berries were segmented from the fixed size feature map and a fully connected (FC) layer was used for bounding box regression and instance classification (He et al., 2017b).

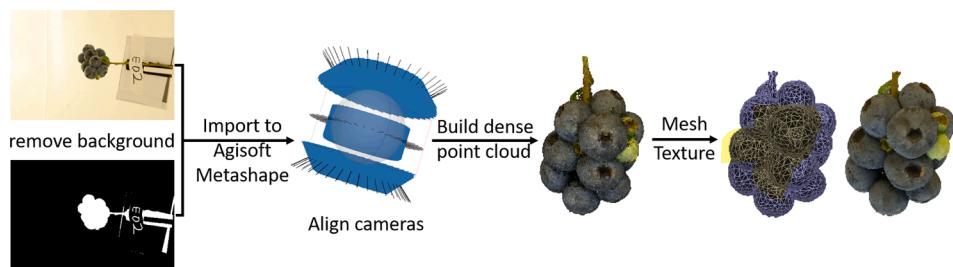


Fig. 5. 3D reconstruction using photogrammetry.

controller and step motor controller. The motor controller turned the step motor 9 degrees, sat stationary for 1.5 s, and then repeated the process 40 times while the cameras captured the cluster images at each stop. The step motor turned 360 degrees, so 120 images were captured from three cameras for each blueberry cluster. Three LED lights were fixed on the frame to provide uniform illumination to the blueberry cluster for better image quality. Four standard calibration markers were printed and fixed above the blueberry cluster for size references. The length of the two top markers was 40 mm, and the length of the two left markers was 30 mm. An example of the images of one cluster captured from the three different perspectives is shown in Fig. 2.

2.3. Individual 3D blueberry segmentation and trait extraction pipeline

The workflow for this study is shown in Fig. 3. The 2D images captured by the imaging system were used for both Mask R-CNN for individual berry detection and segmentation to output pixel-wise mask of the images and for 3D reconstruction based on photogrammetry. Because 3D points and 2D masks all originated from the same 2D images, a 2D mask-3D patches correspondence can therefore be established by means of image projection. According to the 2D-3D correspondence, the traits extracted from 2D were projected to 3D. Based on this principle, a trait extraction algorithm was developed to derive (1) berry number, (2) cluster maturity, (3) berry size, and (4) cluster compactness from the 3D blueberry point clouds.

2.4. 2D blueberry instance segmentation

Mask R-CNN (He et al., 2017b) was used to segment each individual blueberry in the cluster into 2D images (Ni et al., 2020), shown in Fig. 4. The training images were labelled as ‘Mature’ or ‘Immature’ to signify berry maturity based on skin color. The total number of the training images with both a natural and an artificial background was 724, of which 524, 145, and 55 were for training, validation, and testing, respectively. An iterative annotation method developed in an earlier study was used for image labelling (Ni et al., 2020). A Tesla V100 GPU card (Nvidia, Santa Clara, CA) with 16 GB of memory was used for training with 510 epochs. The performance of the training model was evaluated using a mean average precision (mAP) with a 0.5 IOU (Intersection over Union) threshold. 2D blueberry masks were segmented with maturity from all the images captured by the imaging system.

2.5. 3D blueberry reconstruction

2D images captured by the imaging system were used for 3D reconstruction based on photogrammetry that consists of four steps (Smith et al., 2016). The first step was to remove the background to improve computation efficiency because it is unnecessary to reconstruct the background. Otsu thresholding (Otsu, 1979) was used to remove the background of 2D images, generating the blueberry masks. In the second step, local feature points were detected using algorithms such as the

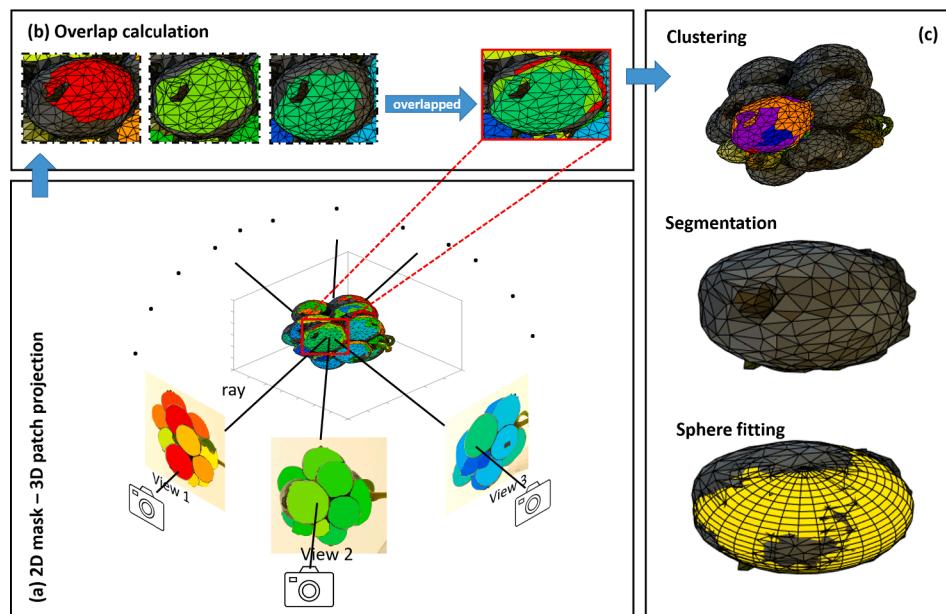


Fig. 6. Illustration of 2D masks projected onto 3D patches (a); overlap calculation (b); 3D berry clustering, segmentation, and sphere fitting (c).

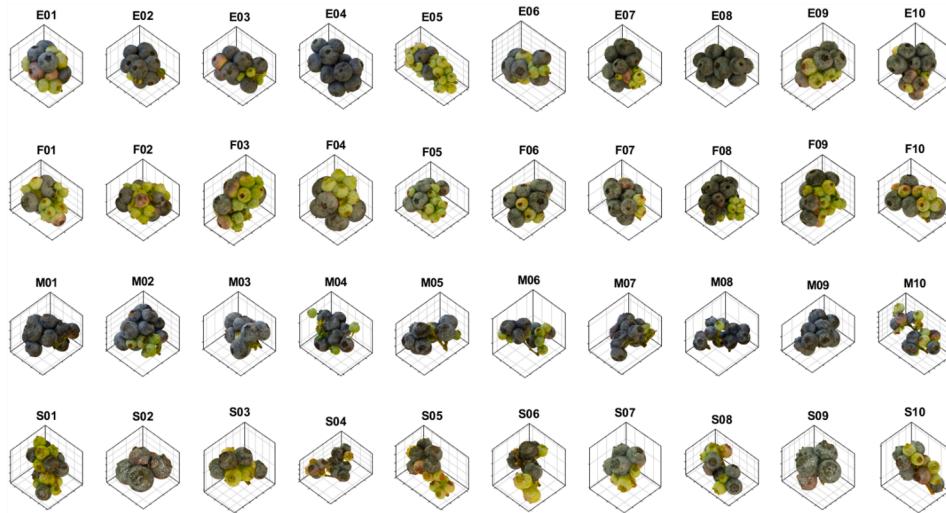


Fig. 7. 3D reconstruction result with texture for four blueberry cultivars, 10 samples for each cultivar. ‘E’, ‘F’, ‘M’ and ‘S’ represent ‘Emerald’, ‘Farthing’, ‘Meadowlark’, and ‘Star’, respectively.

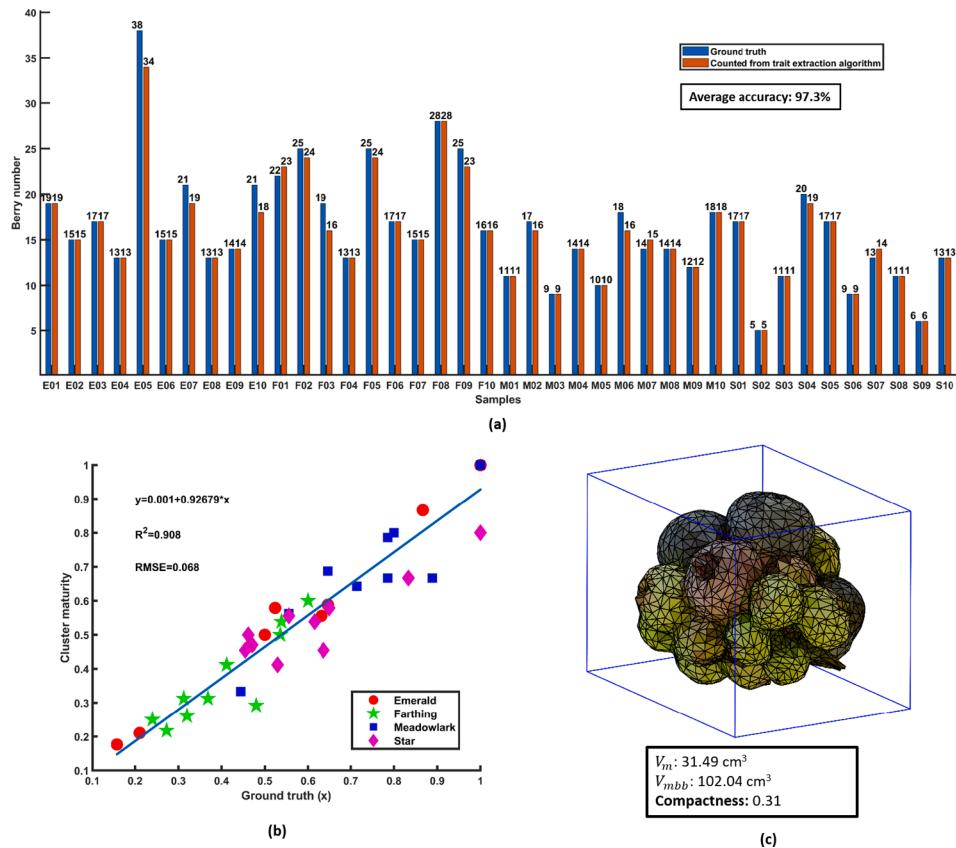


Fig. 8. Three traits results. (a) Counting result for all 40 samples compared with the ground truth, (b) Linear regression for calculated maturity based on counting result with the ground truth, (c) 3D blueberry sample (E01) surrounded by the minimum bounding box, and the compactness was calculated by V_m/V_{mbb} .

scale-invariant feature transform (SIFT) (Lowe, 1999) and the detected feature points (or key points) between image pairs were matched. Based on the matched point pairs (x_1, x_2), the third step was to compute both the camera intrinsic and extrinsic (pose) parameters as well as the 3D positions of the matched key points using triangulation, which is known as bundle adjustment (Szeliski, 2010). More specifically, the fundamental matrix F was computed by $x_1^T F x_2 = 0$ utilizing the matched

points, from which the intrinsic parameters (e.g., focal length, principal point, and axis skew) as well as the extrinsic parameters (rotation and translation) can be recovered (Hartley and Zisserman, 2003). According to the camera intrinsic parameters and pose (including rotation and translation), positions of the matched key points (also known as sparse 3D point cloud) were generated by triangulation (i.e., epipolar

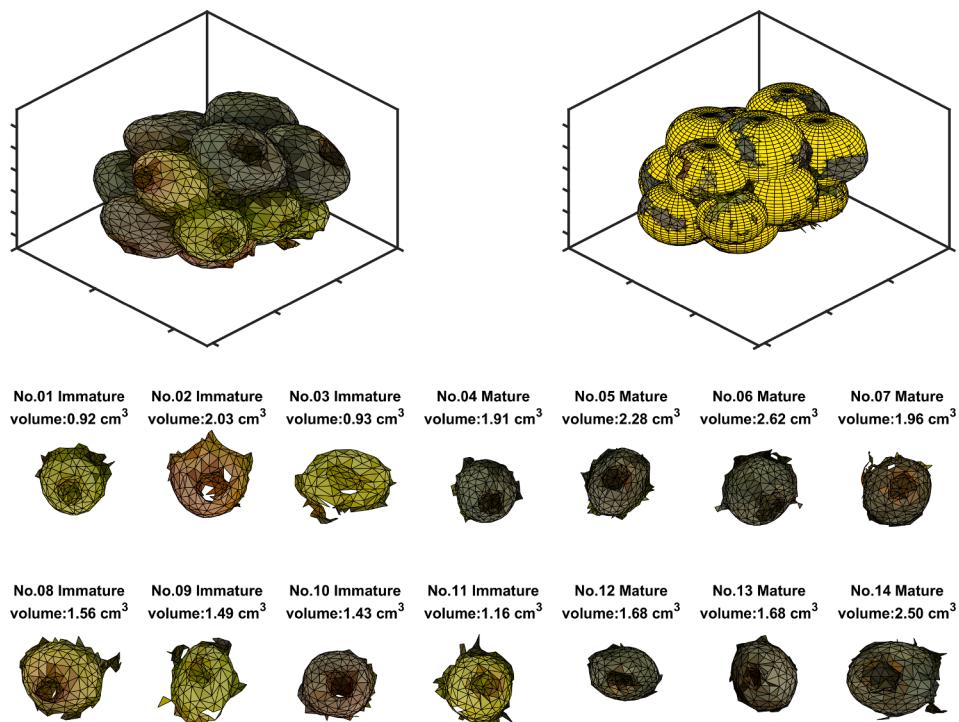


Fig. 9. 3D individual berry segmentation for sample E09. 3D point cloud with 5000 faces mesh (a), sphere fitting for all individual berries (b), and individual berries with their maturity classification and berry volume (c).

geometry). The last step was to compute the dense point cloud according to the camera parameters and the sparse points (or tie-points) computed in step three by using multi-view stereo matching (Seitz et al., 2006). The above 3D reconstruction process was implemented in Agisoft Metashape (Agisoft LLC, Saint Petersburg, Russia). The background noises were manually removed. Mesh and texture were conducted for better visualization (Fig. 5). All 40 samples were 3D reconstructed by this photogrammetric procedure. At the same time, the sample ID and size reference markers on 2D images were also reconstructed to identify the cluster and the actual berry size.

2.6. Trait extraction algorithm

After obtaining 2D masks and 3D point clouds, individual blueberries were segmented and their traits (berry number, cluster maturity, cluster compactness, and berry volume) were extracted in 3 steps with the trait extraction algorithm (Algorithm 1):

Step 1: 2D mask to 3D patch projection

Since the 3D point cloud was reconstructed from 2D images, every pixel on the 2D images corresponds to a 3D point. Thus, every 2D mask on the 2D images projected to a unique 3D patch on the 3D point cloud (Fig. 6(a)). However, in the 2D-3D projection, the same point in the 2D image would project to multiple 3D points that had the same x and y pixel coordinates but had different depths (z). To eliminate this error, a distance (between the 2D points in the 2D images to the 3D points) threshold was set in the 2D-3D projection. The 3D point was taken if the distance was less than the threshold even if there were multiple 3D points corresponding to the same point in the 2D image. If this distance threshold is too large, some erroneous points would manifest; if the distance threshold is too small, no points would manifest. In this study, the distance threshold was set to 30 cm according to the distance between the camera to the blueberry cluster shown in Fig. 2, which was measured manually. From all 2D images, each 2D mask projected onto the 3D point clouds and corresponding patches were identified (Fig. 6 (a)). Every projected 3D patch had the same maturity with the 2D mask.

Step 2: 3D individual berry clustering and segmentation

The same berry was captured in multiple images with different masks and every mask had a projected 3D patch. The 3D patches overlapped if they were from the same berry. Overlaps were calculated by the intersection over union of every two different 3D patches (Fig. 6(b)). When classifying 3D berries, an overlap threshold α was set to remove small overlaps to eliminate false positives. For instance, two 3D berries would be segmented as one because some 3D patches from different berries had a small overlap near the boundary. In this study, $\alpha = 0.3$ had the highest berry segmentation accuracy after testing multiple different thresholds. All the overlapped 3D patches exceeding the threshold were assigned the same berry ID membership and therefore clustered (segmented) as one. After this initial segmentation, there were a few potential errors. One error occurred when some parts of the stem were segmented as one berry because they were segmented as the mask in a few of the 2D images. Normally, one berry would be captured in 60 images (3D patches), so this berry should contain about 60 masks. Therefore, the 3D berries were considered as noise and removed from the result if they only consisted of fewer than 25 3D patches. Another error occurred because some segmented 3D berries only had a small number of faces while most of the faces were noise. To correct this error, a 3D berry was considered as noise and removed if it had fewer than 20 faces. After optimization, all 3D individual blueberries were segmented based on the clustering membership of the 3D patches.

Step 3: Extraction of traits

Berry number was determined by the number of segmented 3D berries. Every 3D berry had multiple corresponding 2D masks from multiple view images and every 2D mask had either a mature or immature property based on classification results from the trained Mask R-CNN model. Generally, each blueberry matures uniformly on the whole surface. Thus, if a berry's number of mature 3D patches was greater than the number of immature 3D patches, this 3D berry was considered as mature; otherwise, this 3D berry was classified as immature. Regarding the maturity of the whole cluster, it was defined as the ratio of mature berry number to the total berry number in the cluster.

A minimum bounding box (MBB) that surrounds the whole blueberry cluster was generated using the method described by Barequet and Har-

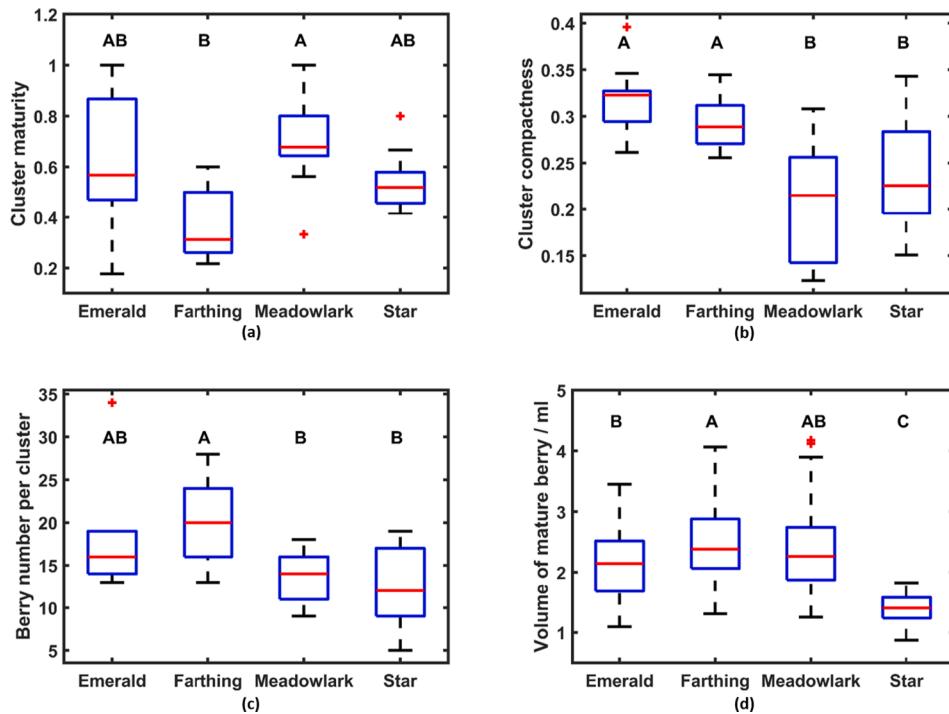


Fig. 10. Trait differences in cultivars. (a) Cluster Maturity, (b) Cluster Compactness, (c) Berry Number per cluster, and (d) Volume of mature berries.

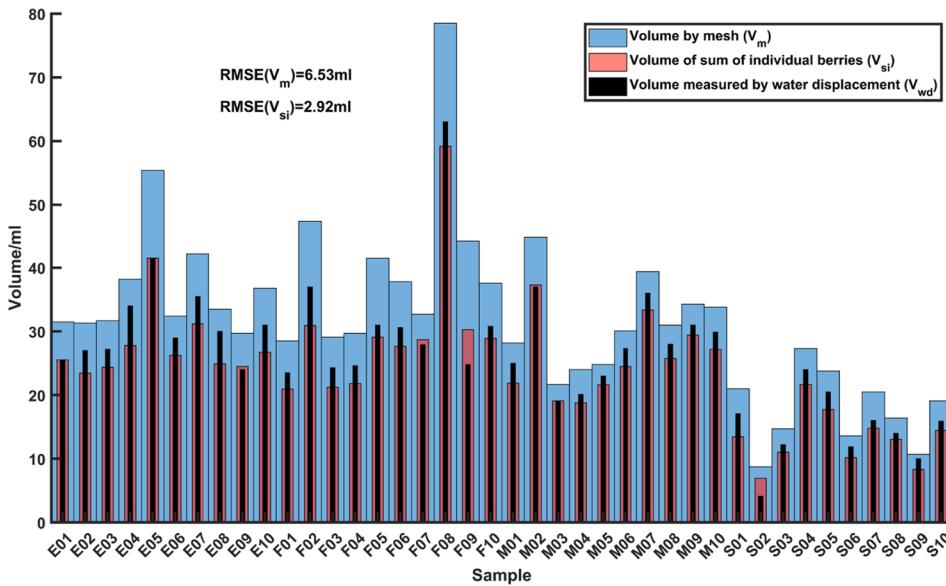


Fig. 11. The calculated volume for the whole blueberry cluster compared with the ground truth. The black (narrowest) bar is the volume measured by water displacement (V_{wd}) as ground truth. The blue (widest) bar is the meshed 3D model (V_m). The red bar is the sum volume of individual berries (V_{si}) added with assumed stem volume within the whole cluster calculated by the fitting sphere after sphere intersections removing. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

[Peled \(2001\)](#). Different from the regular bounding box, the minimum bounding box is only based on the outer shape regardless of the orientation or position of the cluster. According to the high correlation between the cluster compactness and the ratio of berry volume of the whole cluster over the sum of cluster length and width ([Rist et al., 2018](#)), we defined the cluster compactness in our study as the ratio of the whole cluster volume over the MBB volume (shown in [Fig. 8\(c\)](#)), which not only considered the cluster length and width but also the cluster height. Since the meshed 3D point cloud surface was composed of many triangles, the 3D solid could be considered as the combination of multiple tetrahedrons. The berry cluster volume V_m was defined as the sum of every single tetrahedron volume. The volume of the generated minimum bounding box V_{mbb} was calculated by the width, length, and height.

As every individual berry was segmented, sphere fitting was used to fit each berry ([Fig. 6\(c\)](#)). However, the fitted sphere (sv) would overestimate the real volume of a berry because there were considerable intersections between fitted spheres. To achieve higher berry volume accuracy, the sphere intersection volumes (isv) were subtracted from the sum of the individual sphere volumes in a cluster, and a correction factor for the volume was defined as the ratio of $(\sum sv - \sum isv)$ over $(\sum sv)$. This correction factor was multiplied with every individual sphere volume as the final single berry volume.

Algorithm 1. 3D blueberry trait extraction

```

input : 2D masks, 3D point cloud, overlap threshold ( $\alpha$ )
output: berry number, cluster maturity, cluster compactness, berry volume

// Step 1:2D mask - 3D patch projection
foreach 2D image do
    foreach mask do
        3D patch  $\leftarrow$  2D mask; // the maturity of the 3D patch is the
        same with the maturity of the 2D mask
    end
end

// Step 2:3D individual berry clustering and segmentation
for all 3D patches do
    calculate overlap for every two different 3D patches ;
    if overlap  $>$   $\alpha$  then
        | cluster the patches as the same 3D berry ;
    end
end

foreach 3D berry do
    if the number of 3D patch  $<$  25  $\vee$  the number of face  $<$  20 then
        | remove this 3D berry ; // result optimization
    end
end

// Step 3:Trait extraction
berry number  $\leftarrow$  the number of segmented 3D berry ;
foreach 3D berry do
    // set berry maturity ;
    if the number of mature patch  $>$  the number of immature patch then
        | 3D berry  $\leftarrow$  mature ;
    else
        | 3D berry  $\leftarrow$  immature ;
    end
end

cluster maturity  $\leftarrow$  mature berry number / berry number ;
cluster compactness  $\leftarrow$   $V_m / V_{mbb}$  ;
foreach 3D berry do
    sphere fitting;
    berry volume  $\leftarrow$  the volume of fitted sphere (sv);
end

correction factor =  $\frac{\sum sv - \sum isv}{\sum sv}$ ; // berry volume optimization
berry volume  $\leftarrow$  berry volume  $\times$  correction factor

```

3. Results

3.1. 3D reconstruction results

High quality 3D reconstruction results for all 40 samples are shown in Fig. 7. Hundreds of thousands of points were generated for each sample meshing with actual textures. Visually, ‘Emerald’ and ‘Farthing’ are more compact and have more blueberries per cluster than the other two cultivars and ‘Meadowlark’ is more mature than the other three cultivars. However, berry size is difficult to estimate by visually observing these images, as the scales are not uniform for all samples.

3.2. Counting result

The counting accuracy for blueberry cultivars is high with only a few exceptions. Fig. 8(a) shows the counting results compared with the ground truth, i.e., the manual count in each cluster. The average counting accuracy for the 40 samples is 97.3%. The fruit clusters with a low fruit number generally have a higher accuracy, resulting in almost 100% accuracy in many clusters. Because ‘Star’ has the lowest berry number, it has only 2 samples with a 1 berry counting error. The errors occurred primarily in clusters with a larger number of fruits, such as E05, E07, and E10. The average counting accuracies of ‘Emerald’ and ‘Farthing’ are 96.6% and 96.4%, respectively, because of the higher overall fruit numbers, whereas ‘Meadowlark’ and ‘Star’ (97.6% and

98.7%, respectively) have lower overall fruit numbers. There are several explanations for this error. The first is that some samples have very small, young fruit, and these berries could have been considered as part of the other fruits after meshing, so this situation could have resulted in underestimation. The second possibility is that if two berries were segmented as one fruit, even in only one 2D image among the 120 images captured, this resulted in overlaps of two 3D individual berries, which could also have caused underestimation. Also, an overestimation occurred when one berry was erroneously segmented as two because the overlap between two 3D berry patches did not exceed the threshold (α).

3.3. Fruit maturity and cluster compactness

Linear regression was conducted on extracted cluster maturity values regressed with the ground truth (manually classified as mature or immature berries). The results show that the coefficient of determination (R^2) is 0.908 with a root mean square error (RMSE) of 0.068. Fig. 8 (b) shows that the samples of each cultivar are distributed evenly on both sides of the line, which means that the linear regression is suitable for all four cultivars. However, there are still some reasons for errors in calculating maturity. First, counting errors likely have a significant effect on calculating maturity. Second, there may have been blueberries that were just turning blue in a portion of the fruit surface and the remaining part of the fruit was still red or green. This could have led to a mismatch between the manual classifications and the method used in the trait extraction process. For example, a berry determined to have more immature 3D patches than mature 3D patches would be considered as immature, but manual classification considered it a mature fruit.

A 3D blueberry cluster (sample ID: E01) with a minimum bounding box is shown in Fig. 8(c) with compactness result. In this example, the meshed berry volume is 31.49 cm^3 . The minimum bounding box (blue box) was obtained based on the point cloud and the volume is 102.04 cm^3 based on the scale value of 0.0865 according to the marker reference for the E01 model. Thus, the compactness for this sample was calculated as 0.31.

3.4. Individual berry segmentation and size estimation

Individual 3D blueberries were segmented with maturity using the trait extraction algorithm (Fig. 9). For each individual blueberry in the cluster, sphere fitting was performed and the results are shown in Fig. 9 (b). The segmentation results are overall promising but include some minor unconnected faces. In the entire cluster sample, all blueberries are connected when information of those unseen parts behind the connection could not be obtained. Thus, all the segmented blueberries were not complete. However, the berry volume was still estimated based on the major parts of the berry. The volume of mature berries (e.g. No.05 and No.07) is generally larger than that those of immature berries (e.g. No.01 and No.03). In all 40 cluster samples with over 600 berries, there are several berries with abnormally large volume in sample E05, F08, F10, M02, and M07. This is because some berries only show a small part in a very compact cluster. The small segmented part with small curvature resulted in large sphere fitting. Average single berry volume is lower than 3 cm^3 according to the volume measured by the water displacement method, so the berries were removed when the fitted sphere volumes exceeded 6 cm^3 which is two folds of 3 cm^3 in the later volume statistical analysis.

3.5. Statistical analyses for all traits

Blueberry traits of cluster maturity, cluster compactness, berry number, and mature berry volume were compared among the four cultivars (Fig. 10). The results show that ‘Meadowlark’ fruit was more mature than that of ‘Farthing’ in the middle of April. There is no significant difference in cluster maturity at the time of sampling in the field

between ‘Emerald’ and ‘Star’. However, the maturity of ‘Emerald’ varies more than for ‘Star’ which means that ‘Star’ fruits within a cluster ripened more evenly than ‘Emerald’. In terms of cluster compactness, ‘Emerald’ and ‘Farthing’ were more compact than ‘Meadowlark’ and ‘Star’ in the middle of April before they would have been picked for the first time. The cluster compactness of ‘Emerald’ and ‘Farthing’ varies less among clusters examined in this study than ‘Meadowlark’ and ‘Star’, which means that clusters of ‘Emerald’ and ‘Farthing’ were uniformly compact while the compactness values ranged from loose to compact in ‘Meadowlark’ and ‘Star’. The clusters of ‘Farthing’ have a higher fruit number than ‘Meadowlark’ and ‘Star’. Only mature berry volume was estimated because there is no need to analyze the immature berry volume since the immature berries varies considerably from a very young stage to the near mature stage. The mature berry volume of ‘Farthing’ is larger than that of ‘Emerald’. ‘Star’ has the smallest mature berries among the four cultivars. The median mature berry volume of ‘Farthing’ is 2.385 cm^3 , which is about 70% larger than that of ‘Star’ whose median value is 1.405 cm^3 . The mature berry volume of ‘Meadowlark’ is not significantly different from that of ‘Farthing’ and ‘Emerald’.

4. Discussion

4.1. Instance segmentation models

Although Mask R-CNN has been largely considered as the state-of-the-art instance segmentation model and used for benchmarking other technique, new convolutional neural network models have been proposed recently, such as YOLACT, SOLO, PolarMask, BlendMask, and SOLOv2 (Bolya et al., 2019; Wang et al., 2019; Xie et al., 2020; Chen et al., 2020; Wang et al., 2020). Some of these new models may achieve better performance in the 2D berry segmentation. We tested SOLOv2 with the same data used in Mask R-CNN and results show that SOLOv2 achieved average precision (AP) of 0.892 and 0.877 for validation and testing dataset, respectively, while Mask R-CNN achieved 0.750 and 0.720, respectively. However, the better performance of SOLOv2 in 2D berry segmentation may have little effect in 3D berry segmentation and trait extraction because the model presented in this study already achieved a high accuracy based on Mask R-CNN with over 97% counting accuracy. In future studies, newer deep learning models could be investigated.

4.2. Cluster volume comparison

The berry cluster volume was calculated with three different methods (Fig. 11). Berry cluster volume determined by the mesh (V_m) give the largest volume because the empty space inside the berries was meshed and considered as part of the whole volume. Furthermore, stems also contributed to the volume as they were included in the volume determined by the water displacement method (V_{wd}). The berry cluster volume was also calculated as the sum of all individual berry volumes (V_{si}) within the same clusters, which is less than the volume measured by water displacement. This is because the stem’s volume was not included in the sum of all individual berry volumes, which makes this method (i.e., sum of V_{si}) more accurate. The method of taking the sum of individual fitted sphere volumes did not take into account the considerable intersection among those spheres (Fig. 9(b)), and therefore overestimated the total cluster volume. To counter this error, the intersection volumes of the fitted spheres were calculated and subtracted from the total volume to obtain the final volume of the cluster. Thus, it is reasonable that the cluster volume obtained by summation of individual berry volumes is much less than the volume calculated from the entire cluster 3D meshes because the stem volume and empty space volume were included, and less than the volume measured by water displacement because the stem volume was included. The ratio between the sum volumes after and before sphere intersection removing was used to calculate the final

individual berry volume and more accurate individual berry volumes were obtained. The volume of stems for every cluster was assumed to 0.2 cm³ and was added to V_{si} . As mentioned previously, berries were removed when the fitted sphere volume exceeded 6 cm³. Therefore, the berry volumes that exceeded 6 cm³ were added according to the medium value of mature berry volume for each of the cultivars (Fig. 10(d)). After these two steps, both RMSEs of V_m and V_{si} were calculated and compared with the ground truth (V_{wd}). The RMSE of V_{si} is 2.92 cm³, which means that the average error for each berry volume should have been less than 0.292 cm³ when the berry number for a cluster is more than 10. The average accuracy of V_{si} is 89.68%, which is larger than V_m , indicating that the volume calculated by our trait extraction pipeline is more accurate than the mesh volume.

4.3. Traits of different blueberry cultivars

The blueberry cluster maturity and compactness calculated from our study are consistent with the reference descriptions for ‘Emerald’, ‘Farthing’, ‘Meadowlark’, and ‘Star’ (Lyrene, 1998, 2001, 2010, 2017). The harvest period for southern highbush cultivars varies by as much as 3 weeks from early ripening to late ripening cultivars, which allows the harvesting to be stretched over a 6 to 8-week period at a given location. Within an individual cultivar, the harvest may last 3 weeks since blueberry fruit on a bush do not mature uniformly and would require selective picking of mature berries 4 to 7 days apart. The first harvest date for ‘Meadowlark’ was April 10, which is 10 days and 12 days earlier, respectively, than for ‘Farthing’ and ‘Emerald’ in Gainesville, FL, which is 250 km south of Alma, GA (Coordinates: 29°65'20" N; 82°32'50" W). All samples in this study were collected on the same date—April 17 in Alma, GA, when growers had already started harvesting ‘Meadowlark’ and ‘Star’ but not ‘Emerald’ and ‘Farthing’. The mean mid-harvest date for ‘Meadowlark’ was April 24, which was 10 days earlier than for ‘Farthing’. The results of this study indicate that ‘Meadowlark’ was significantly more mature than ‘Farthing’ on April 17. There are no significant maturity differences between ‘Emerald’ and ‘Star’, and the mid-harvest date of ‘Emerald’ was 7 days after ‘Meadowlark’. The order of the median maturity dates as calculated by our analysis (Fig. 10(a)) is consistent with the mean mid-harvest date order (‘Meadowlark’ > ‘Emerald’ = ‘Star’ > ‘Farthing’) as mentioned by Lyrene. Lyrene also described the cluster tightness for ‘Meadowlark’, ‘Farthing’, ‘Emerald’, and ‘Star’ as ‘loose’, ‘medium’, ‘medium’ and ‘loose’, respectively, which is consistent with the results obtained in this study (Fig. 10(b)).

The berry volume for mature berry estimation obtained in this study reveals a different pattern from the data presented by Lyrene (1998, 2001, 2010, 2017). The reported height × width (cm × cm) for a single berry fruit from ‘Meadowlark’, ‘Farthing’, ‘Emerald’ and ‘Star’ are 1.5 × 1.9, 1.27 × 1.71, 1.3 × 1.8 and 1.4 × 1.8, respectively. If berries were considered as ellipsoid, the berry volumes for the four cultivars would be 2.83 cm³, 2.08 cm³, 2.21 cm³ and 2.37 cm³, respectively. ‘Meadowlark’ had the largest berry volume, ‘Farthing’ had the smallest volume, and ‘Emerald’ and ‘Star’ ranked between ‘Meadowlark’ and Farthing. However, the actual order of mature berry volume from the data obtained in this study is ‘Farthing’ (‘Meadowlark’) > ‘Emerald’ (‘Meadowlark’) > Star. Some horticultural factors could have contributed to these inconsistent findings (e.g. difference in environment conditions, soil, fruit load, pruning practices, pollinator activity during bloom, and fertilization practice). NeSmith demonstrated that ‘Emerald’ has a much larger five-year average (2005–2010) mature berry size than ‘Star’ in Ware County (Alma) in Georgia (NeSmith, 2014). A later study by NeSmith reported that ‘Farthing’ and ‘Star’ have similar mature berry sizes for 3 years (2014–2016) in Alapaha, GA, located about 80 km west of Alma (NeSmith, 2017).

4.4. Advantages and limitations

The most significant advantage of this study is the ability to achieve a complete individual blueberry segmentation from a 3D point cloud, whereas some earlier studies only used sphere fitting to represent the individual berry to count berry number (Mack et al., 2017; Nellithimaru and Kantor, 2019). The berry counting accuracy for blueberries in a cluster is higher than similar studies on grapes (Rist et al., 2018). This is also the first study to describe extraction methods for phenotypic traits for blueberry clusters (e.g. cluster maturity, cluster compactness, berry number and berry size) synchronously in 3D with the aid of deep learning instance segmentation from 2D images. The limitation of this study is that it was performed indoors under an illumination-controlled environment and was limited to one cluster at a time. Image capturing systems and procedures can become more robust for capturing images under ambient, field conditions, as some recent studies have demonstrated (Nellithimaru and Kantor, 2019; Gené-Mola et al., 2020). The extension from indoor conditions to outdoor conditions will be conducted in a future study extra efforts in: (1) transferring our sensor system to a platform that could traverse the field, and (2) making a cover for the whole system to block the wind and uncontrollable illumination and give controllable artificial illumination.

5. Conclusions

This study provides an individual 3D blueberry segmentation pipeline for extracting blueberry cluster maturity, cluster compactness, berry number, and berry size. The trait extraction pipeline combined with masks segmented from 2D images and mesh faces from 3D point clouds demonstrates promising functionality for the segmentation of individual berries in 3D. The average counting accuracy is 97.3% for all cultivars, and the RMSE for maturity is 0.068 for all cultivars. The average error for individual berry volume is less than 0.295 cm³ if berry number per cluster is larger than 10. The phenotypic traits derived from our methods for four blueberry cultivars are generally in agreement with observations and prior reports from plant breeders. The method developed in this study provides a reliable means for performing blueberry fruit phenotypic trait extraction from a large sample size. It can also offer blueberry breeders, growers, and packers a new technique for objectively phenotyping blueberry cluster/fruit traits, as well as potentially monitor/track cluster and berry development throughout the season.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

This study is supported by the USDA National Institute of Food and Agriculture Specialty Crop Research Initiative (Award No: 2019-51181-30015). The authors thank China Scholarship Council (CSC) for the financial support to the author (X. N) to conduct her doctoral research in the Bio-Sensing and Instrumentation Lab at the University of Georgia. The author also gratefully thank Rui Xu and Renee Allen for the help on data collection in Alma, GA, in USA.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at <https://doi.org/10.1016/j.isprsjprs.2020.11.010>.

References

- Ballington, J., et al., 1990. Germplasm resources available to meet future needs for blueberry cultivar improvement. *Fruit Varieties J.* 44, 54–62.
- Barequet, G., Har-Peled, S., 2001. Efficiently approximating the minimum-volume bounding box of a point set in three dimensions. *J. Algorithms* 38, 91–109.
- Bargoti, S., Underwood, J., 2017. Deep fruit detection in orchards. In: 2017 IEEE International Conference on Robotics and Automation (ICRA), IEEE, pp. 3626–3633.
- Barnea, E., Mairon, R., Ben-Shahar, O., 2016. Colour-agnostic shape-based 3d fruit detection for crop harvesting robots. *Biosyst. Eng.* 146, 57–70.
- Bolya, D., Zhou, C., Xiao, F., Lee, Y.J., 2019. Yolact: Real-time instance segmentation. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 9157–9166.
- Brazelton, D., Strik, B.C., 2007. Perspective on the us and global blueberry industry. *J. Am. Pomological Soc.* 61, 144.
- Chaudhury, A., Ward, C., Talasaz, A., Ivanov, A.G., Brophy, M., Grodzinski, B., Hüner, N.P., Patel, R.V., Barron, J.L., 2018. Machine vision system for 3d plant phenotyping. *IEEE/ACM Trans. Comput. Biol. Bioinf.* 16, 2009–2022.
- Chen, H., Sun, K., Tian, Z., Shen, C., Huang, Y., Yan, Y., 2020. Blendmask: Top-down meets bottom-up for instance segmentation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 8573–8581.
- Chen, L.C., Papandreou, G., Kokkinos, I., Murphy, K., Yuille, A.L., 2017a. DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE Trans. Pattern Anal. Mach. Intell.* 40, 834–848.
- Chen, S.W., Shivakumar, S.S., Deunha, S., Das, J., Okon, E., Qu, C., Taylor, C.J., Kumar, V., 2017b. Counting apples and oranges with deep learning: a data-driven approach. *IEEE Robot. Automat. Lett.* 2, 781–788.
- Cupec, R., Filko, D., Vidović, I., Nyarko, E.K., Hocenski, Ž., 2014. Point cloud segmentation to approximately convex surfaces for fruit recognition. In: Proceedings of the Croatian Computer Vision Workshop, Zagreb, Croatia, pp. 56–61.
- DeVetter, L.W., Yang, W.Q., Takeda, F., Korthuis, S., Li, C., 2019. Modified over-the-row machine harvesters to improve northern highbush blueberry fresh fruit quality. *Agriculture* 9, 13.
- Ge, Y., Xiong, Y., Tenorio, G.L., From, P.J., 2019. Fruit localization and environment perception for strawberry harvesting robots. *IEEE Access* 7, 147642–147652.
- Gené-Mola, J., Sanz-Cortiella, R., Rosell-Polo, J.R., Morros, J.R., Ruiz-Hidalgo, J., Vilaplana, V., Gregorio, E., 2020. Fruit detection and 3d location using instance segmentation neural networks and structure-from-motion photogrammetry. *Comput. Electron. Agric.* 169, 105165.
- Girshick, R., 2015. Fast R-CNN. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 1440–1448.
- Girshick, R., Donahue, J., Darrell, T., Malik, J., 2014. Rich feature hierarchies for accurate object detection and semantic segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 580–587.
- Gongal, A., Karkee, M., Amatya, S., 2018. Apple fruit size estimation using a 3d machine vision system. *Int. Process. Agric.* 5, 498–503.
- Habaragamuwa, H., Ogawa, Y., Suzuki, T., Shiigi, T., Ono, M., Kondo, N., 2018. Detecting greenhouse strawberries (mature and immature) using deep convolutional neural network. *Eng. Agric. Environ. Food* 11, 127–138.
- Hartley, R., Zisserman, A., 2003. Multiple View Geometry in Computer Vision. Cambridge University Press.
- He, J.Q., Harrison, R.J., Li, B., 2017a. A novel 3d imaging system for strawberry phenotyping. *Plant Methods* 13, 93.
- He, K., Gkioxari, G., Dollár, P., Girshick, R., 2017b. Mask r-cnn. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 2961–2969.
- Ji, W., Zhao, D., Cheng, F., Xu, B., Zhang, Y., Wang, J., 2012. Automatic recognition vision system guided for apple harvesting robot. *Comput. Electrical Eng.* 38, 1186–1195.
- Jiang, L., Koch, A., Scherer, S.A., Zell, A., 2013. Multi-class fruit classification using rgbd data for indoor robots. In: 2013 IEEE International Conference on Robotics and Biomimetics (ROBIO), IEEE, pp. 587–592.
- Jiang, Y., Li, C., Paterson, A.H., Robertson, J.S., 2019. DeepSeedling: Deep convolutional network and kalman filter for plant seedling detection and counting in the field. *Plant Methods* 15, 141.
- Koirala, A., Walsh, K., Wang, Z., McCarthy, C., 2019. Deep learning for real-time fruit detection and orchard fruit load estimation: benchmarking of ‘mangoyolo’. *Precision Agric.* 20, 1107–1135.
- Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. Imagenet classification with deep convolutional neural networks. *Adv. Neural Inf. Process. Syst.* 1097–1105.
- Kromdijk, J., Bertin, N., Heuvelink, E., Molenaar, J., de Visser, P., Marcelis, L., Struijk, P., 2014. Crop management impacts the efficiency of quantitative trait loci (qtl) detection and use: case study of fruit load×qtl interactions. *J. Exp. Bot.* 65, 11–22.
- Kurtulmus, F., Lee, W.S., Vardar, A., 2014. Immature peach detection in colour images acquired in natural illumination conditions using statistical classifiers and neural network. *Precision Agric.* 15, 57–79.
- LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. *Nature* 521, 436–444.
- LeCun, Y., Bottou, L., Bengio, Y., Haffner, P., 1998. Gradient-based learning applied to document recognition. *Proc. IEEE* 86, 2278–2324.
- Liang, Q., Zhu, W., Long, J., Wang, Y., Sun, W., Wu, W., 2018. A real-time detection framework for on-tree mango based on ssd network. International Conference on Intelligent Robotics and Applications, Springer, 423–436.
- Lin, G., Tang, Y., Zou, X., Xiong, J., Fang, Y., 2020. Color-, depth-, and shape-based 3d fruit detection. *Precision Agric.* 21, 1–17.
- Linker, R., Cohen, O., Naor, A., 2012. Determination of the number of green apples in rgbd images recorded in orchards. *Comput. Electron. Agric.* 81, 45–57.
- Liu, X., Chen, S.W., Aditya, S., Sivakumar, N., Deunha, S., Qu, C., Taylor, C.J., Das, J., Kumar, V., 2018. Robust fruit counting: combining deep learning, tracking, and structure from motion. In: 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), IEEE, pp. 1045–1052.
- Long, J., Shelhamer, E., Darrell, T., 2015. Fully convolutional networks for semantic segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3431–3440.
- Lowe, D.G., 1999. Object recognition from local scale-invariant features. In: Proceedings of the Seventh IEEE International Conference on Computer Vision, IEEE, pp. 1150–1157.
- Lyrene, P., 1998. Low-chill highbush blueberry ‘Star’. US Patent App. 08/523,357.
- Lyrene, P.M., 2001. Blueberry plant called ‘Emerald’. US Patent App. 09/392,389.
- Lyrene, P.M., 2010. Southern highbush blueberry plant named ‘FL01-173’. US Patent App. 12/587,285.
- Lyrene, P.M., 2017. Blueberry plant named ‘FL03-228’. US Patent App. 14/544,722.
- Mack, Jennifer, Lenz, C., Teutrine, J., Steinlage, V., 2017. High-precision 3D detection and reconstruction of grapes from laser range data for efficient phenotyping based on supervised learning. *Comput. Electron. Agric.* 135, 300–311.
- Nellithimaru, A.K., Kantor, G.A., 2019. Rols: Robust object-level slam for grape counting. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops.
- NeSmith, D.S., 2014. ‘TH-819’ southern highbush blueberry Georgia Dawn. HortScience 49, 674–675.
- NeSmith, D.S., 2017. ‘TH-921’ southern highbush blueberry Miss Alice Mae. HortScience 52, 196–197.
- Ni, X., Li, C., Jiang, H., Takeda, F., 2020. Deep learning image segmentation and extraction of blueberry fruit traits associated with harvestability and yield. *Hortic. Res.* 7, 1–14.
- Otsu, N., 1979. A threshold selection method from gray-level histograms. *IEEE Trans. Syst. Man Cybernet.* 9, 62–66.
- Qiang, L., Jianrong, C., Bin, L., Lie, D., Yajing, Z., 2014. Identification of fruit and branch in natural scenes for citrus harvesting robot using machine vision and support vector machine. *Int. J. Agric. Biol. Eng.* 7, 115–121.
- Redmon, J., Divvala, S., Girshick, R., Farhadi, A., 2016. You only look once: unified, real-time object detection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 779–788.
- Ren, S., He, K., Girshick, R., Sun, J., 2015. Faster R-CNN: Towards real-time object detection with region proposal networks. *Adv. Neural Inf. Process. Syst.* 91–99.
- Rist, F., Herzog, K., Mack, J., Richter, R., Steinlage, V., Töpfer, R., 2018. High-precision phenotyping of grape bunch architecture using fast 3d sensor and automation. *Sensors* 18, 763.
- Sa, I., Ge, Z., Dayoub, F., Upcroft, B., Perez, T., McCool, C., 2016. Deepfruits: A fruit detection system using deep neural networks. *Sensors* 16, 1222.
- Sargent, S.A., Takeda, F., Williamson, J.G., Berry, A.D., 2020. Harvest of southern highbush blueberry with a modified, over-the-row mechanical harvester: use of handheld shakers and soft catch surfaces. *Agriculture* 10, 4.
- Seitz, S.M., Curless, B., Diebel, J., Scharstein, D., Szeliski, R., 2006. A comparison and evaluation of multi-view stereo reconstruction algorithms. In: 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’06), IEEE, pp. 519–528.
- Simonyan, K., Zisserman, A., 2014. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.
- Smith, M., Carrivick, J., Quincey, D., 2016. Structure from motion photogrammetry in physical geography. *Prog. Phys. Geogr.* 40, 247–275.
- Sodhi, P., Vijayarangan, S., Wettergreen, D., 2017. In-field segmentation and identification of plant structures using 3d imaging. In: 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), IEEE, pp. 5180–5187.
- Stein, M., Bargoti, S., Underwood, J., 2016. Image based mango fruit detection, localisation and yield estimation using multiple view geometry. *Sensors* 16, 1915.
- Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A., 2015. Going deeper with convolutions. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1–9.
- Szeliski, R., 2010. Computer Vision: Algorithms and Applications. Springer Science & Business Media.
- Takeda, F., Kremer, G., Li, C., MacLean, D., Olmstead, J.W., 2013. Techniques for increasing machine harvest efficiency in highbush blueberry. *HortTechnology* 23, 430–436.
- Tanksley, S.D., 2004. The genetic, developmental, and molecular bases of fruit size and shape variation in tomato. *Plant Cell* 16, S181–S189.
- Tao, Y., Zhou, J., 2017. Automatic apple recognition based on the fusion of color and 3d feature for robotic fruit picking. *Comput. Electron. Agric.* 142, 388–396.
- Tao, Y., Zhou, J., Wang, K., Shen, W., 2018. Rapid detection of fruits in orchard scene based on deep neural network. In: Proc. ASABE Annu. Int. Meeting, p. 1.
- Tran, D.T., Hertog, M.L., Tran, T.L., Quyen, N.T., Van de Poel, B., Mata, C.I., Nicolaï, B.M., 2017. Population modeling approach to optimize crop harvest strategy: the case of field tomato. *Front. Plant Sci.* 8, 608.
- Wachs, J.P., Stern, H.I., Burks, T., Alchanatis, V., 2010. Low and high-level visual feature-based apple detection from multi-modal images. *Precision Agric.* 11, 717–735.
- Wahabzada, M., Paulus, S., Kersting, K., Mahlein, A.K., 2015. Automated interpretation of 3d laserscanned point clouds for plant organ segmentation. *BMC Bioinformatics* 16, 248.
- Wang, X., Kong, T., Shen, C., Jiang, Y., Li, L., 2019. Solo: Segmenting objects by locations. *arXiv preprint arXiv:1912.04488*.

- Wang, X., Zhang, R., Kong, T., Li, L., Shen, C., 2020. Solov2: Dynamic, faster and stronger. arXiv preprint arXiv:2003.10152.
- Xia, C., Wang, L., Chung, B.K., Lee, J.M., 2015. In situ 3D segmentation of individual plant leaves using a rgb-d camera for agricultural automation. Sensors 15, 20463–20479.
- Xie, E., Sun, P., Song, X., Wang, W., Liu, X., Liang, D., Shen, C., Luo, P., 2020. Polarmask: Single shot instance segmentation with polar representation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 12193–12202.
- Xiong, J., Liu, Z., Tang, L., Lin, R., Bu, R., Peng, H., 2018. Visual detection technology of green citrus under natural environment. Trans. Chinese Soc. Agric. Mach. 49, 45–52.
- Zhang, M., Jiang, Y., Li, C., Yang, F., 2020. Fully convolutional networks for blueberry bruising and calyx segmentation using hyperspectral transmittance imaging. Biosyst. Eng. 192, 159–175.