

Article

Tomato Growth Monitoring and Phenological Analysis Using Deep Learning-Based Instance Segmentation and 3D Point Cloud Reconstruction

Warut Timprae ¹, Tatsuki Sagawa ², Stefan Baar ¹, Satoshi Kondo ¹, Yoshifumi Okada ¹, Kazuhiko Sato ¹, Poltak Sandro Rumahorbo ¹, Yan Lyu ¹, Kyuki Shibuya ³, Yoshiki Gama ³, Yoshiki Hatanaka ³ and Shinya Watanabe ^{1,*}

¹ Graduate School of Engineering, Muroran Institute of Technology, 27-1 Mizumoto-cho, Muroran 050-8585, Japan; 23096501@muroran-it.ac.jp (W.T.); okada@muroran-it.ac.jp (Y.O.); sandro@muroran-it.ac.jp (P.S.R.)

² Hitachi Solutions, Co., Ltd., 4-12-7 Higashishinagawa, Shinagawa-ku, Tokyo 140-0002, Japan

³ Asai Nursery, Inc., Tsu 514-2221, Japan

* Correspondence: sin@muroran-it.ac.jp; Tel.: +81-143-46-5432

Abstract

Accurate and nondestructive monitoring of tomato growth is essential for large-scale greenhouse production; however, it remains challenging for small-fruited cultivars such as cherry tomatoes. Traditional 2D image analysis often fails to capture precise morphological traits, limiting its usefulness in growth modeling and yield estimation. This study proposes an automated phenotyping framework that integrates deep learning-based instance segmentation with high-resolution 3D point cloud reconstruction and ellipsoid fitting to estimate fruit size and ripeness from daily video recordings. These techniques enable accurate camera pose estimation and dense geometric reconstruction (via SfM and MVS), while Nerfacto enhances surface continuity and photorealistic fidelity, resulting in highly precise and visually consistent 3D representations. The reconstructed models are followed by CIELAB color analysis and logistic curve fitting to characterize the growth dynamics. When applied to real greenhouse conditions, the method achieved an average size estimation error of 8.01% compared to manual caliper measurements. During summer, the maximum growth rate ($g_{\max}$) of size and ripeness were 24.14%, and 95.24% higher than in winter, respectively. Seasonal analysis revealed that winter-grown tomatoes matured approximately 10 days later than summer-grown fruits, highlighting environmental influences on phenological development. By enabling precise, noninvasive tracking of size and ripeness progression, this approach is a novel tool for smart and sustainable agriculture.

Keywords: tomato growth estimation; instance segmentation; ellipsoid fitting; photogrammetry

Academic Editor: Marko Vinceković

Received: 2 October 2025

Revised: 7 November 2025

Accepted: 7 November 2025

Published: 12 November 2025

Citation: Timprae, W.; Sagawa, T.; Baar, S.; Kondo, S.; Okada, Y.; Sato, K.; Rumahorbo, P.S.; Lyu, Y.; Shibuya, K.; Gama, Y.; et al. Tomato Growth Monitoring and Phenological Analysis Using Deep Learning-Based Instance Segmentation and 3D Point Cloud Reconstruction. *Sustainability* **2025**, *17*, 10120. <https://doi.org/10.3390/su172210120>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In recent years, research on plant phenotyping using artificial intelligence has progressed significantly in large-scale greenhouse agriculture [1]. Among various crops, tomatoes are particularly important for large-scale greenhouse horticulture. Numerous studies have focused on facility-grown tomatoes, including automated detection of tomato fruits, leaf area estimation, and yield prediction [2–4]. In these studies, object detection and instance segmentation techniques played a vital role in automating the identification of tomato fruits, thereby contributing to labor cost reduction and improved operational efficiency [5–7].

However, accurately detecting tomato fruits and estimating their size and color remains a significant challenge in greenhouse environments, where a large number of plants are cultivated in dense, multi-row arrangements. This difficulty is exacerbated in the case of cherry tomatoes, which require high-precision size estimations. Although previous research has been successful in applying image-based methods for yield estimation. In situations where data collection is limited or object position angles are not suitable, the effectiveness of image-based assessment techniques decreases. Includes advances in tomato phenotypic analysis. However, this method has been underexplored.

To address these challenges, the proposed method detects individual cherry tomato fruits from daily video recordings captured in a greenhouse, extracts size and color information, and analyzes their growth dynamics over time. For a precise size estimation, this method reconstructs 3D point clouds for each detected fruit using instance segmentation with the YOLOv8x-seg model [8,9]. The 3D reconstruction pipeline integrates Structure-from-Motion (SfM), Multi-View Stereo (MVS), and the Nerfacto framework to generate detailed point clouds, which are then refined using ellipsoid-fitting techniques [10,11].

This research focuses on developing an integrated 3D growth tracking framework, which combines instance segmentation with point cloud visualization. The framework enables accurate fruit size estimation through a fitting shape approach. Facilitates the modeling of growth dynamics over time using logistic curve fitting, thus providing a comprehensive strategy for quantitative phenological analysis.

The proposed approach enables automated and quantitative growth assessment, contributing to sustainable greenhouse agriculture by reducing manual labor, increasing resource efficiency, and supporting data-driven crop management approaches that align with sustainable food production goals.

Growth estimation was performed based on morphological and color features. Temporal changes in fruit size and color from fruit set to harvest are modeled using logistic growth curves, enabling the objective estimation of current growth stages and future harvest timing [12].

The effectiveness of the proposed method is validated using real-world data collected from operational greenhouse farms.

The remainder of this paper is structured as follows. Section 2 reviews related studies to contextualize the background of this research. Section 3 details the proposed methodology for tomato phenotyping and the estimation of ripeness and size. Section 4 presents the experimental results obtained using the proposed approach, and Section 5 discusses these results in detail. Finally, Section 6 concludes the paper by summarizing the main findings.

2. Related Works

This section introduces related research on instance segmentation, the CIELAB color space, and 3D model reconstruction. These techniques form the basis of the methodology

proposed in this study. Additionally, comparisons of tomato growth in terms of size and color over time using a logistic function.

2.1. Image Processing and Instance Segmentation

In agricultural image processing, instance segmentation technique enables the detection and identification of individual fruits. Moreover, this method allows the extraction of key characteristics, such as approximate size, color, quantity, and spatial location. This information is crucial for the subsequent analysis and model development [13].

One of the most widely adopted and efficient approaches for object detection is the You Only Look Once (YOLO) framework [8,10], which offers real-time performance with high accuracy. YOLO has been used in various agricultural applications, including plant disease detection, pest management [9], and yield estimation of tomatoes in controlled environments, such as plant factories [14].

Traditional 2D image-based approaches have been widely used for ripeness estimation because of their simplicity and minimal hardware requirements. For example, we utilized RGB color analysis from 2D images to classify tomato maturity levels, demonstrating that ripeness stages can be effectively estimated using color features [3]. Similarly, Gómez et al. [15] showed that using the CIELAB color space improves the consistency of ripeness classification.

However, these methods are inherently limited in their ability to capture spatial and morphological characteristics, such as size, shape, and volume, which are essential for comprehensive phenotyping. The absence of three-dimensional information restricts the accuracy and robustness of ripeness evaluations [16]. To address this limitation, they introduced a 3D reconstruction pipeline (3DPhenoMVS), which significantly enhanced the ability to analyze fruit traits in three dimensions using multi-view imagery and point cloud generation.

2.2. 3D Model Reconstruction

Integrating computer vision with 3D reconstruction techniques is critical for advancing fruit phenotyping in agriculture. Recent studies have leveraged 3D phenotyping methods employing SfM and MVS to capture morphological traits more accurately [10,16,17].

However, SfM has certain limitations. When the number of 2D images is insufficient, the camera angles are too narrow, or the image quality is poor, the resulting 3D reconstruction may exhibit a high error percentage. For example, Lindenberger et al. [18] emphasized that the number of 2D observations significantly influenced the accuracy of object size estimation, with at least three views required to maintain an acceptable error margin. This finding highlights that 3D image-based measurements generally provide higher accuracy than traditional 2D image-based approaches, particularly for the assessment of object dimensions.

In 3D modeling, ellipsoid fitting is applied to shape approximations in various contexts. Maillard and Kunisky [19] presented a robust mathematical approach for ellipsoid fitting, which inspired the use of least-squares ellipsoid models in this study to estimate the fruit size and volume. In addition, accurate size estimation in outdoor environments often relies on metric calibration using reference objects, as emphasized by Gené-Mola et al. [20]. A spherical reference object addresses these challenges in a practical greenhouse setting.

Clustering algorithms such as DBSCAN are effective for segmenting point cloud data in agricultural applications [21,22]. These techniques help to isolate individual fruits within complex plant structures, particularly when combined with color cues. This study

adopted a similar strategy to segment tomatoes into dense canopies using both spatial and color-based clustering.

Recent developments in neural rendering, particularly in Neural Radiance Fields (NeRFs), have opened up new opportunities for high-fidelity 3D reconstruction from RGB images. Choi et al. [23] and Zheng et al. [24] proposed NeRF-based pipelines for crop morphology capture, aligning with the use of the Nerfacto framework in this study to improve the 3D quality of daily video recordings.

Although alternative technologies, such as LiDAR, have been explored for fruit phenotyping owing to their rapid depth-sensing capabilities [25], they often face limitations in greenhouse environments. For instance, typical LiDAR sensors may struggle with small fruits such as cherry tomatoes, producing insufficient point density and color information for ripeness assessment. Moreover, high-precision LiDAR systems remain costly compared with vision-based methods. In contrast, this study employed standard RGB videos from consumer-grade cameras to reconstruct high-resolution colored 3D point clouds, enabling size and ripeness analyses in a cost-effective manner. Therefore, LiDAR-based approaches were not adopted in this study because they are less suitable for small-scale fruits under greenhouse conditions and would introduce unnecessary costs and complexity compared to the proposed RGB video-based framework.

Comparing 2D image-based approaches and 3D model reconstruction methods for tomato growth estimation shows each approach has distinct strengths and limitations. For instance, 2D image-based analysis methods require fewer computing resources and are simpler to implement. However, their application scope is quite limited because the camera angle and position can affect the perceived size and shape of objects in the image, making the approach less flexible. On the other hand, 3D modeling is more robust to variations in camera angle and position. However, 3D model generation is relatively complex and requires a large amount of data and computational resources for training. In this study, the knowledge of each method was applied to create a model for estimating tomato growth rate, including a 2D image approach and 3D modeling, as well as quality improvement methods such as shape fitting, to achieve good efficiency.

3. Materials and Methods

This section presents the methods used in this study, including details of the data and study area, the tomato detection method, 3D reconstruction, feature extraction for estimating tomato size and ripeness, and the analysis used to estimate tomato growth.

3.1. Area of Study and Description of Data

This study focused on agricultural technology, specifically targeting cherry tomatoes, which are small tomato cultivars. The monitoring period spans approximately 30–40 days, from the initial fruiting stage to full maturity, depending on the seasonal conditions.

The data used in this study were obtained from video recordings collected at Asai Nursery, located in Tsu, Mie Prefecture, Japan (longitude 136°28'16.8" E, latitude 34°47'07.9" N). The tomatoes were cultivated in a greenhouse, and recordings were captured using a standard mobile phone camera (Lenovo Tab M9, 8-megapixel resolution). Videos were recorded in MP4 format at a resolution of 1920 × 1080 pixels. In this study, the camera was not fixed or calibrated for daily use, resulting in slight variations in camera position and angle across different days, with a capture distance of up to approximately 60 cm. To avoid disturbing plant growth, videos were taken from an approximate 180° angle, rather than a full 360° view. Consequently, this study required the application of additional techniques to compensate for incomplete viewpoints before performing accurate size estimation.

3.2. Approach Overview

The overall approach of this research involved estimating tomato growth patterns through the extraction of key features such as size and ripeness. This methodology combines 3D modeling for accurate measurements with mathematical modeling to represent the growth dynamics of tomato plants and fruits. By integrating computer vision and mathematical approaches, this study aims to provide a comprehensive framework for the nondestructive monitoring of tomato development in controlled-environment agriculture.

The process begins with Step 1, data preparation, which includes converting raw video data into individual frames, followed by annotation and preprocessing to ensure high-quality input for model training. The prepared data were then used to train an instance segmentation model in Step 2 for accurate tomato detection, enabling the identification and localization of individual fruits across different growth stages. Once tomatoes were detected, the process branched into two parallel paths to estimate their size and ripeness.

For size estimation, after detection, the workflow proceeds to Step 3, where a 3D reconstruction model is created to generate a point cloud representation of the tomatoes. This step ensured that the fine morphological details were preserved for reliable measurements. The reconstruction was followed by shape fitting to obtain a complete 3D tomato model, after which the actual size of the tomatoes, such as diameter and volume, was estimated with high precision.

Simultaneously, for ripeness estimation, the process continues to Step 4, which evaluates the ripeness of each detected tomato using the color features extracted from the Lab* color space. This approach enables the quantitative assessment of maturity levels, which are closely associated with fruit quality and harvest readiness.

Finally, once the size and ripeness estimations were complete, Step 5 involved the construction of a logistic growth model to describe the relationship between tomato size and ripeness over time using a mathematical formulation. This integration provides a dynamic representation of phenological changes, offering valuable insights into the growth trajectory and the maturation process.

An overview of the methodology is illustrated in Figure 1, and detailed descriptions of each step are provided in the following sections.

As illustrated in Figure 1, Step 1 corresponds to the input stage in which raw data are prepared for subsequent analysis. Step 2 involves tomato detection using an instance segmentation model; further details are provided in Section 3.3. Step 3 focuses on size estimation through 3D reconstruction and shape fitting, as described in Sections 3.4–3.6. In parallel, Step 4 addresses ripeness estimation based on color analysis, as discussed in Section 3.7. Finally, Step 5 presents the construction of a logistic growth model that integrates the outcomes of the size and ripeness estimations, as explained in Section 3.8. Together, these steps establish a unified framework that enables accurate phenotypic measurements and dynamic growth modeling of tomatoes, forming the basis for subsequent experimental evaluation.

3.3. Instance Segmentation and Color Overlay

In this study, an object detection model based on YOLOv8x-seg was trained to identify tomato plants in video frames. The tomatoes detected were identified using instance segmentation. YOLOv8x-seg is a deep learning object detection model in the YOLO series, released in 2023 by Ultralytics. YOLOv8 supports instance segmentation and includes five subversions (n, s, m, l, and x), each with different internal structures and performance. Yue et al. [9] compared the performance of various YOLOv8 versions, such as the segmentation of healthy and diseased tomato plants. The most efficient version, YOLOv8x-seg,

achieved a mean average precision (mAP; 0.5) of 90.7%. However, it also has the largest model size compared to the others. In this study, YOLOv8x-seg was selected to build the model because tracking performance was prioritized over the slightly larger model size. To improve model generalization, basic data augmentation strategies were applied during training, including vertical and horizontal flipping, 90° rotations clockwise and counterclockwise, random rotations within $\pm 45^\circ$, and shear transformations within $\pm 15^\circ$. Examples of these augmentations are shown in Figure 2.

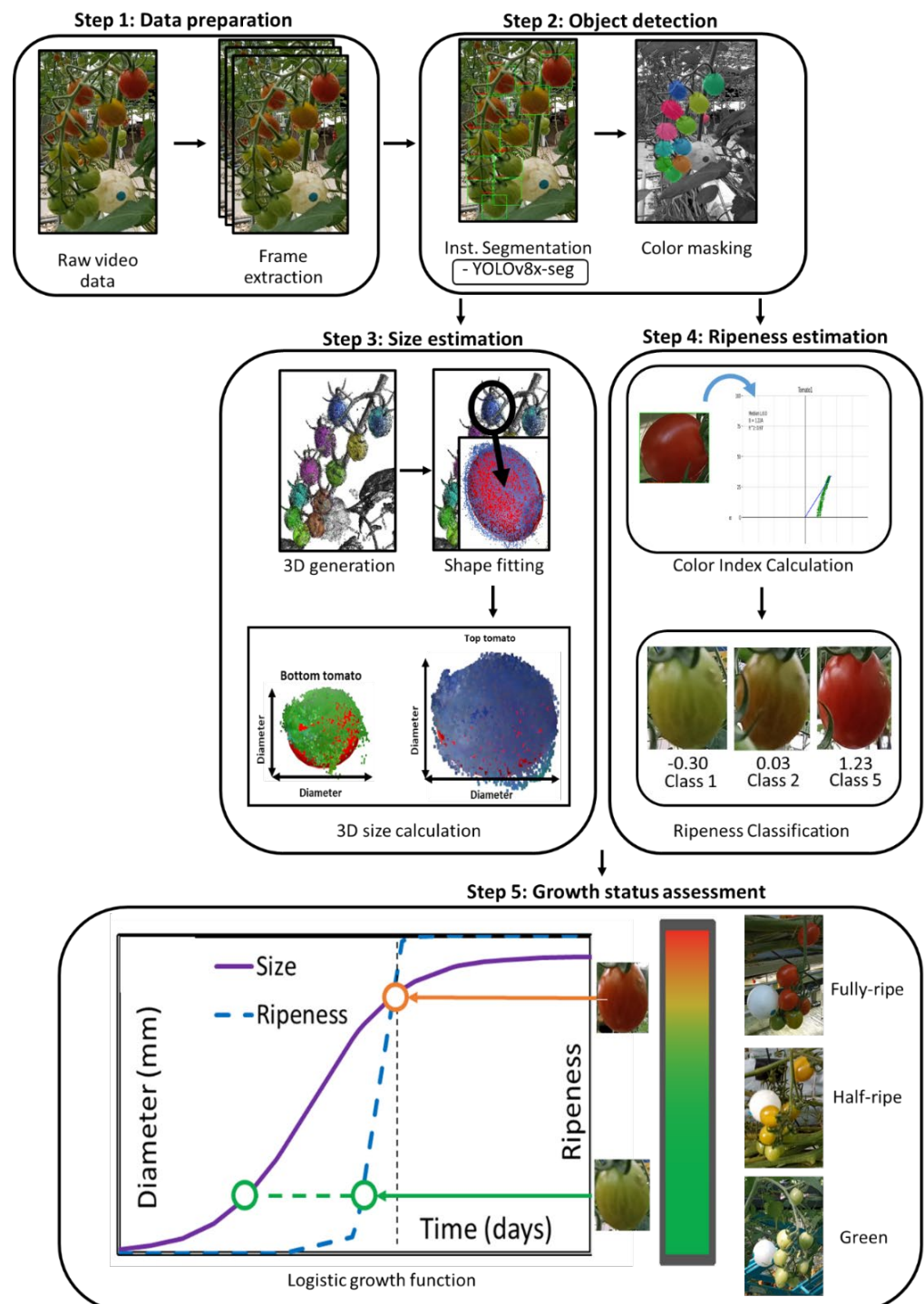


Figure 1. Flowchart diagram of the study methodology overview.

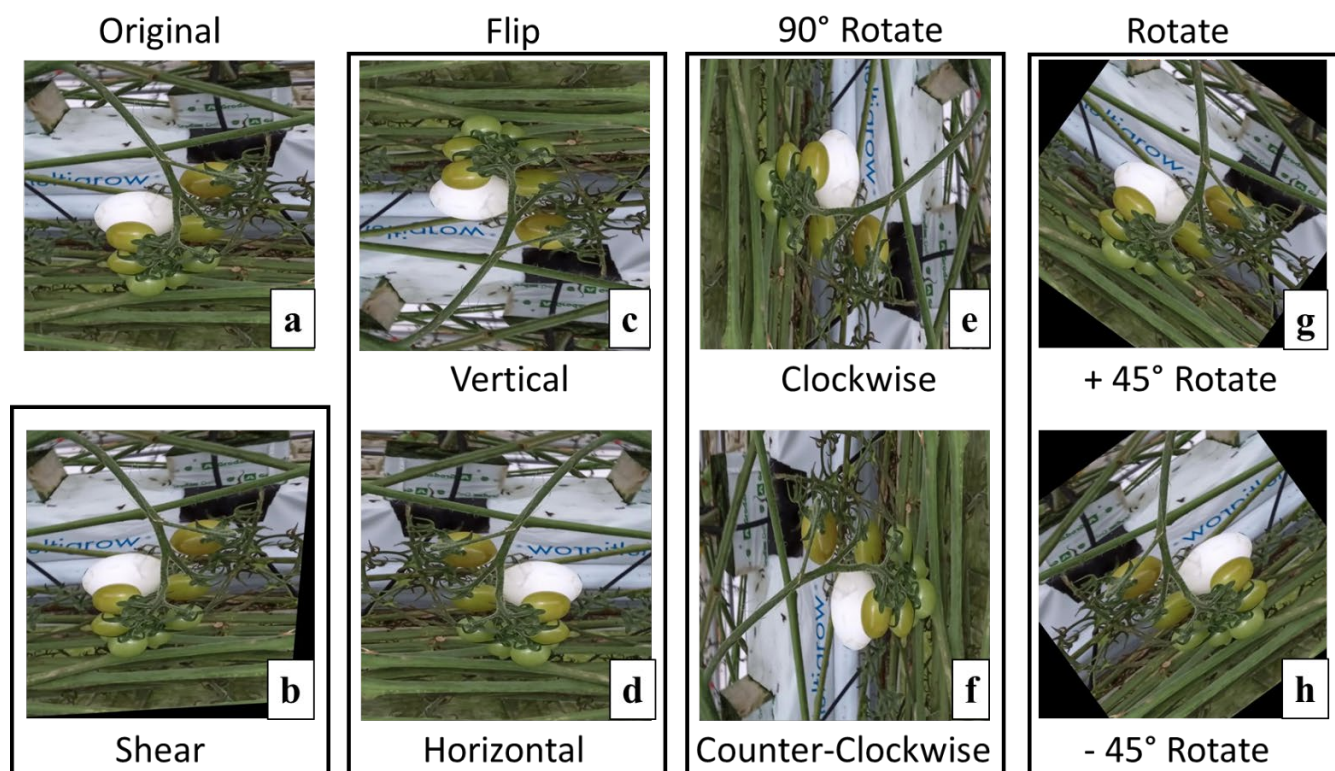


Figure 2. Examples of data augmentation applied to tomato images: (a) original image, (b) shear $\pm 15^\circ$ (c) vertical flip, (d) horizontal flip, (e) 90° rotation clockwise, (f) 90° rotation counterclockwise, (g,h) random rotation between $\pm 45^\circ$.

The YOLOv8x-seg model was trained under this environment using a batch size of 16, for 10 epochs, and a learning rate of 0.01. All input images were resized to 640×640 pixels. The hyperparameter settings used for training are summarized in Table 1.

Table 1. Hyperparameter settings.

Batch size	16
Number of epochs	10
Learning rate	0.01
Image size	640×640 pixels

Color masking refers to the process of highlighting important objects identified through object detection before generating 3D images. This step is crucial because various factors in the 3D image generation can distort the results. One significant factor is the color of the image. When an image is converted to 3D, many point clusters are created, each retaining its original color. This can cause background colors or unrelated objects to blend with the object of interest, thereby distorting the final 3D shape. For example, a green tomato surrounded by green leaves and branches, when converted into a cluster of 3D points without distinguishing colors, may result in overlapping green points, making accurate analysis impossible [26–28].

Color masking also simplifies point grouping. The process begins by converting an RGB image into grayscale. Subsequently, a unique color is assigned to prevent overlapping or closely positioned tomatoes from being confused. During the color masking process, the color values are converted from RGB to HSV, where H represents the hue, S represents the saturation, and V represents the value (or brightness). The advantage of using HSV is that the H value corresponds to a specific color, which can be easily adjusted.

Different colors can also be assigned, making it easier to distinguish between tomatoes. For example, in an 8-bit image with 255 possible H values, up to 255 different colors can be assigned to different tomatoes with S and V held constant.

3.4. 3D Point Clouds Generation

Cherry tomatoes have small dimensions, typically in millimeters (mm), making accurate size measurements highly dependent on image resolution. As discussed in Section 2, measurements based on 3D images generally provide a higher accuracy than those based on 2D images [18]. Furthermore, because the camera used in this study was not positioned at a fixed location, it was not possible to reliably calculate the focal length, making accurate size estimation from 2D images particularly challenging. Therefore, generation of accurate 3D point clouds played a critical role in this research [16].

In 3D image generation, the primary inputs are camera position information and point cloud coordinates. During the input acquisition process, the SfM and MVS techniques were employed. SfM calculates a 3D structure from multiple photographs captured from different camera positions or viewpoints. SfM detects specific features in each image such as corners or edges and matches these features across different images. Geometric calculations are then used to estimate the camera positions and generate 3D points.

This process effectively compensates for the lack of a fixed camera setup, since SfM automatically estimates the relative camera poses from overlapping images, allowing accurate 3D reconstruction even under slight variations in angle and position. The output of SfM is a point cloud that is not dense, and is primarily used to create the main structure of the image [18]. The fundamental principle SfM is shown in Figure 3.

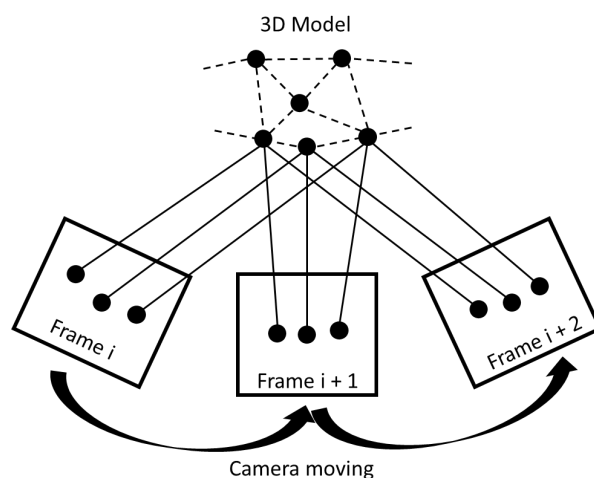


Figure 3. The fundamental principle of SfM, in which multiple overlapping images captured from different camera viewpoints are used to detect corresponding feature points and reconstruct their spatial positions as 3D coordinates. This process enables the generation of a 3D model from 2D observations.

Once the camera position information and main structure of the 3D image are obtained, the MVS technique is utilized to create a high-resolution 3D model based on multi-view photographs with known camera positions. This process results in a significantly higher-resolution image [29].

Subsequently, the data were used to train the Nerfacto model to produce high-quality 3D scenes. Nerfacto is a pipeline designed to generate NeRF models by integrating NeRF with instant-NGP, thereby improving the speed and efficiency of 3D scene reconstruction. The training process leverages neural networks to learn the radiance and point density values in 3D space, enabling the generation of photorealistic renderings [23,24].

This workflow utilizes COLMAP for SfM, MVS, and Nerfstudio to train Nerfacto, thus providing a user-friendly framework that enhances the efficiency of generating and rendering high-quality 3D models.

In summary, the 3D image generation process consists of the following steps. First, the SfM technique uses intrinsic and extrinsic parameters to generate a 3D image structure in the form of sparse point clusters. Then, the MVS technique increases the density of point clusters to make the image more complete. Finally, the Nerfacto framework utilizes the output from these SfM-MVS to enhance texture and render a more comprehensive and realistic 3D image.

3.5. 3D Point Cloud Identification and Shape Fitting

Individual tomatoes were identified after generating 3D point clouds from the processed images. This was achieved using a combination of clustering techniques and geometric analyses. Specifically, the point clouds were segmented into distinct groups representing individual tomatoes. Several steps were followed to achieve this goal.

1. Point Cloud Clustering: Using density-based spatial clustering (DBSCAN or another suitable clustering algorithm), groups of points corresponding to individual tomatoes were identified. Additionally, color values from the staining were used to assist with grouping alongside spatial clustering, allowing for finer differentiation between tomatoes. This helps to isolate tomatoes within a point cloud [21,22].

2. Ellipsoid fitting: The morphology of the cherry tomato closely resembles an ellipsoid rather than a perfect sphere, as its major and minor diameters differ slightly but consistently. Therefore, the ellipsoid model is suitable for fitting the fruit's natural geometry. The approach helped fit occluded or partially reconstructed tomatoes. Even when point clusters were incomplete due to leaf or fruit overlap during SfM-MVS, the least-squares fitting process could approximate missing regions by inferring the most likely fruit geometry. This provided a more reliable estimate of fruit size and shape, even under partial visibility conditions. Each cluster was fitted to a 3D ellipsoid model using the least-squares ellipsoid-fitting technique. The ellipsoid is defined by three semi-principal axes (a , b , and c , representing the principal dimensions of the fruit) and a center point. The goal is to minimize the distance between the points in the cluster and the surface of the ellipsoid, providing an estimate of the shape and size of the tomato in 3D space.

The ellipsoid equation is expressed as:

$$x^T S x = d \quad (1)$$

where S is a 3×3 symmetric matrix defined as:

$$S = \begin{bmatrix} S_{11} & S_{12} & S_{13} \\ S_{12} & S_{22} & S_{23} \\ S_{13} & S_{23} & S_{33} \end{bmatrix} \quad (2)$$

The parameter d represents the scale factor (set to $d = 3$ in this case for a 3D ellipsoid). Least-squares optimization was employed to minimize the residuals ($x^T S x = d$) for all points x in the cloud. The eigenvalues and eigenvectors of S were used to calculate the principal axes and volume of the ellipsoid. These results provide insights into the geometry and growth characteristics of tomatoes.

The diameters of the ellipsoid along its axes were derived from the eigenvalues, and the volume of the ellipsoid was calculated using the equation:

$$V = \frac{4}{3} \pi abc \quad (3)$$

where a , b , and c are the lengths of the semi-principal axes. This method enables accurate measurements of tomato dimensions and volume, supporting detailed growth analysis [19].

3. Visibility estimation: For clusters that do not fit well to an ellipsoid owing to occlusions or incomplete data, interpolation methods are used to fill in the missing parts. In some cases, manual refinement may be necessary to correct significant occlusions caused by leaves or overlapping fruits.

Tomato growth rate estimation utilizes the morphological characteristics of tomato fruits [16], such as fruit volume, fruit surface area, vertical diameter, transverse diameter, fruit shape index, and fruit color, as follows:

Fruit volume: the volume of the convex hull of the generated 3D point cloud.

Fruit surface area: the total area of the surface meshes.

Vertical diameter: distance between the pedicel and opposite end of the tomato fruit.

Transverse diameter: diameter of the largest transverse section perpendicular to the vertical diameter.

Fruit shape index: the ratio between the vertical and transverse diameters.

Fruit color: the color range of tomatoes indicates the ripeness of the fruit.

However, before estimating the growth rate, the accuracy of 3D data must be evaluated. In this study, the transverse diameter obtained from 3D images was compared with the actual size measured using calipers.

3.6. Size Estimation

Once the 3D reconstruction and identification of each fruit were completed, the next crucial step was to calibrate the 3D model to ensure accurate size estimation. Metric calibration is essential for aligning the distances and dimensions in a point cloud using real-world measurements [20]. For this purpose, a reference object was introduced during the data collection process, which was approximately the size of a ping-pong ball (40 mm in diameter) that closely matches the average size of a real-world tomato. By calibrating the 3D model to a known reference object, the scales and sizes of the reconstructed tomatoes were ensured to be accurate. The fruit size was calculated using the following equation:

$$S_{\text{tomato}} = \frac{L_{\text{tomato3D}}}{L_{\text{ref3D}}} \times 40 \quad (4)$$

where S_{tomato} is the tomato size estimation, L_{tomato3D} is the length of the tomato obtained from the 3D image, L_{ref3D} is the length of the reference ball obtained from the 3D image, and 40 is the actual diameter of the reference ball. The ratio of the actual reference ball size to the 3D reference ball size represents the calibration process. When multiplied by the 3D tomato size, the size of the tomato is in standard units of millimeters.

3.7. Color-Based Ripeness Estimation

In this study, CIELAB standard color measurements were used to assess tomato ripeness, specifically the luminosity (L^*), red-green component (a^*), and yellow-blue component (b^*). These parameters are effective for distinguishing different ripening stages [15].

Following the segmentation of the tomatoes, the a^* and b^* values were extracted for each pixel. A scatter plot was generated using these values, and the slope of the linear interpolation was calculated to quantify the color-based ripeness characteristics.

The interpolation slope serves as a quantitative indicator of ripeness, reflecting the color transition during the maturation process. Once the color-based ripeness estimation was completed, the tomato ripeness grade was determined based on Figure 4 to classify tomatoes according to their ripeness stage.

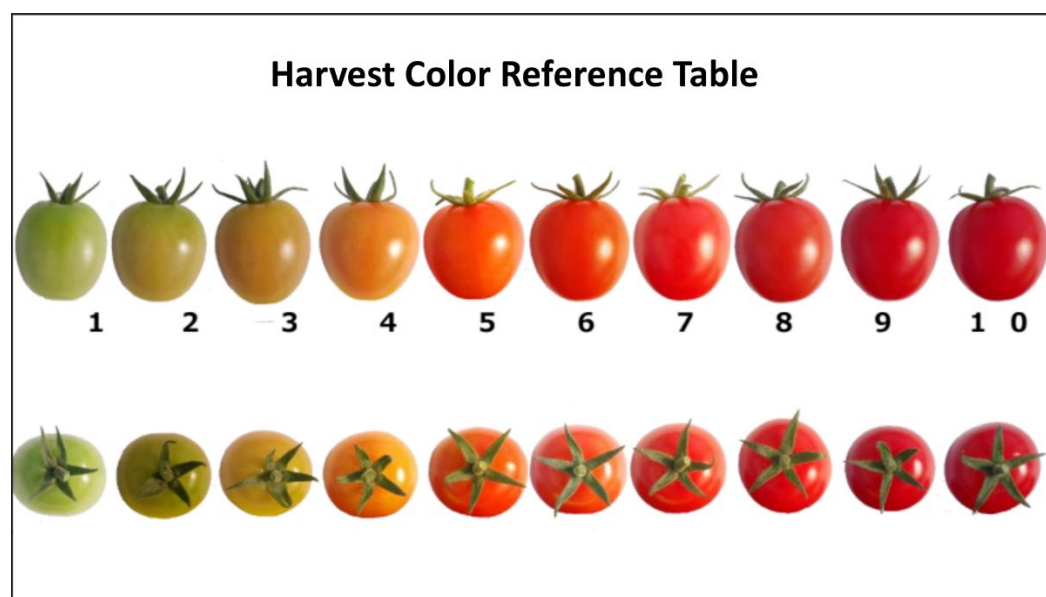


Figure 4. Tomato harvest color reference table provided by Asai Nursery, based on Poudel et al. [30].

3.8. Logistic Modeling of Tomato Size and Ripeness Transitions

Monitoring tomato fruit growth in large-scale greenhouse environments is essential for accurate yield estimation and for maximizing labor efficiency, which aligns with the objectives of the present study. To achieve this, mathematical modeling is commonly employed to characterize the growth patterns of plants and fruits and to understand their responses to environmental influences.

Baar et al. [12] implemented a logistic model for precise tomato fruit growth prediction based on fruit diameter measurements under greenhouse conditions and demonstrated the effectiveness of a simple logistic approach. Their findings highlighted that such models can accurately predict tomato maturation with minimal complexity, supporting their continued use in real-time greenhouse monitoring systems.

Among the various modeling approaches, the logistic function is one of the most suitable for describing plant and fruit growth dynamics. The general form of the logistic equation is as follows:

$$f(x) = \frac{A}{1 + e^{-d(x-x_0)}} \quad (5)$$

The equation was applied to this study, where A is the maximum size of tomato, x_0 is the position of the peak growth day, and d can describe the growth rate of the tomato. However, for comparisons with other datasets. If values of A and x_0 are the same, the values of d can be compared to determine which dataset has the faster growth rate. On the other hand, when the values of A and x_0 are different, the maximum growth rate ($g_{\max}$) value can be used to compare the results, which is calculated from the differential function as follows: $g_{\max} = df/dx(\max(df/dx)_x)$.

To obtain the optimal parameters, the minimum optimization approach was applied using the Generalized Reduced Gradient (GRG) nonlinear method, with the initial parameter limits assigned as shown in Table 2.

Table 2. Initial parameter settings.

	Description	Unit	Initial Value
A	Maximum tomato size	mm	50
x_0	Peak growth day	days	40
d	Growth rate	-	10

Various logistic-type functions, including the sigmoid, Richards, and Gompertz models, are commonly employed to describe biological growth processes. Fang et al. [31] provided a detailed analysis of these functions and proposed enhancements for several complex formulations.

Nevertheless, this study adopts the standard logistic model introduced earlier primarily to reduce computational complexity and simplify parameter estimation. Future work may explore the integration of more advanced models to enhance fitting accuracy and better capture biological variability.

4. Experimental Results

This section presents the experimental results obtained by applying the proposed methodology to video data collected in a greenhouse environment. The effectiveness of each component instance segmentation, 3D reconstruction, size and ripeness estimation, and growth modeling was evaluated. Quantitative assessments were performed to validate the accuracy of the detection and size estimation processes, whereas temporal analyses illustrated the growth trends and seasonal variations in tomato development. Visual examples and statistical comparisons are included to demonstrate the performance of the system under real-world conditions.

The experiments, which took approximately 40 min per instance segmentation model training and took approximately hour and half to generated a 3D model, were conducted on Windows 11 using a 13th Gen Intel® Core™ i7-13700KF CPU and an NVIDIA GeForce RTX 4070 Ti GPU. The deep learning framework was Python 3.10.0 with PyTorch 2.1.2 and CUDA 12.2, accelerated by cuDNN 9.1.0.2. The 3D reconstruction framework was nerfstudio 1.1.3 and colmap 3.10.

4.1. Accuracy of Instance Segmentation Model

The object detection model employed in this study was trained using the publicly available Laboro Tomato dataset hosted on GitHub [32]. This dataset included tomato images categorized into three ripeness stages: fully ripe, half-ripe, and green. It contains 804 images, each of which features multiple tomatoes of varying ripeness levels. The images were obtained at two resolutions: 3024×4032 and 3120×4160 pixels.

YOLOv8x-seg was selected based on the comparative subversion (n, s, m, l, and x) in Section 3.3, which demonstrated high segmentation performance and suitability for instance-level detection in tomato images. In addition, comparisons with other commonly used instance segmentation frameworks, including Mask R-CNN [8,9]. The performance comparison of each model is shown in Table 3.

Table 3. Comparison of segmentation results across models.

Models	Precision (%)	Recall (%)	F1 Score (%)	Segment mAP@0.5 (%)	Model Size (MB)
YOLOv8n-Seg	0.815	0.755	0.784	0.842	6.8
YOLOv8s-Seg	0.824	0.800	0.812	0.87	23.8
YOLOv8m-Seg	0.808	0.812	0.810	0.875	54.8
YOLOv8l-Seg	0.809	0.817	0.813	0.873	92.3

YOLOv8x-Seg	0.806	0.818	0.812	0.881	143.9
Mask RCNN	0.421	0.455	0.437	0.523	334.84

The instance masks produced by the model achieved a mAP of 0.881 at an Intersection over Union (IoU) threshold of 0.8, demonstrating a robust performance in segmentation tasks. The overall precision and recall across all classes were 0.806 and 0.818, respectively, indicating effective identification of tomato fruits at various ripening stages. The model demonstrated high performance for fully ripened tomatoes, achieving a mAP of 0.886, followed by 0.919 for green tomatoes, and 0.839 for half-ripened tomatoes. These results highlight the effectiveness of the model in accurately distinguishing different tomato ripeness levels, underscoring its potential for practical applications in agricultural monitoring and management.

4.2. 3D Tomato Phenotyping and Metric Calibration

Evaluating the accuracy of a reconstructed 3D point cloud is challenging. A common method involves comparison with 3D images acquired using precise instruments, such as LiDAR and structured light scanners [10,16,25]; however, the high cost of such devices can be a limitation. In this study, a metric calibration technique was used to evaluate accuracy by employing a reference object to compare the measurements obtained from the 3D model with the corresponding real-world values. The variables selected for comparison were the transverse diameter of the tomato and the diameter of the reference ball [20].

After instance segmentation to detect tomatoes, 3D reconstruction was performed, beginning with staining to highlight key features before generating the 3D model. The process involved segmenting the point cloud data, followed by selecting tomatoes of interest using the DBSCAN technique and color value clustering. The resulting 3D point groups were then analyzed to isolate the target tomatoes, allowing for a more focused assessment of their growth characteristics and other relevant features (Figure 5).

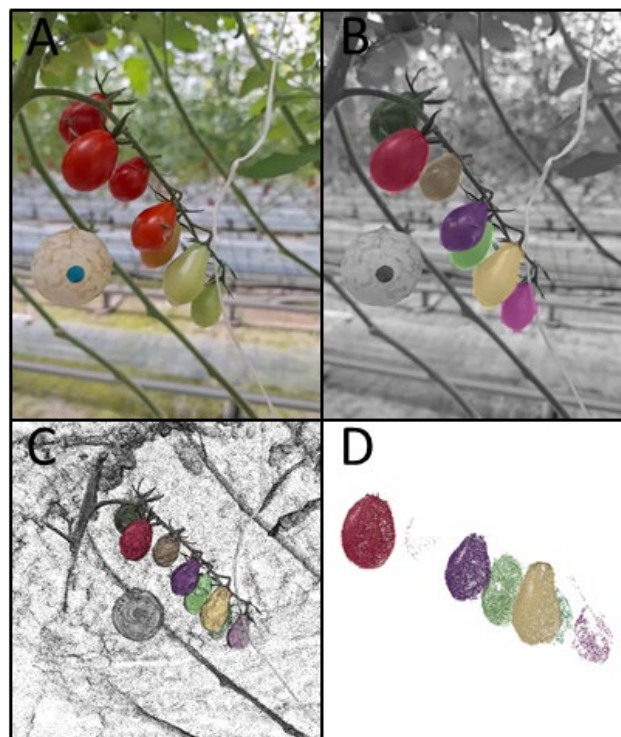


Figure 5. Results from detection to 3D segmentation: (A) Original image. (B) Recolored image. (C) 3D point cloud representation. (D) 3D point cloud with background clipping and DBSCAN applied to color values for tomato clustering.

4.3. Shape Fitting Techniques and Size Estimation Analysis for Tomatoes

The point clouds were separated into reference spheres and tomato fruits (Figure 6). Each point cloud was then shaped based on the principles outlined in Section 3.5; the reference sphere was adjusted to form a circular shape, whereas the ellipsoid was fitted to the tomato, as this tomato variety exhibited an ellipsoid shape. After completing these processes, the diameter of the tomato was measured, and the variables specified in Section 3.6 were calculated to estimate the growth rate.

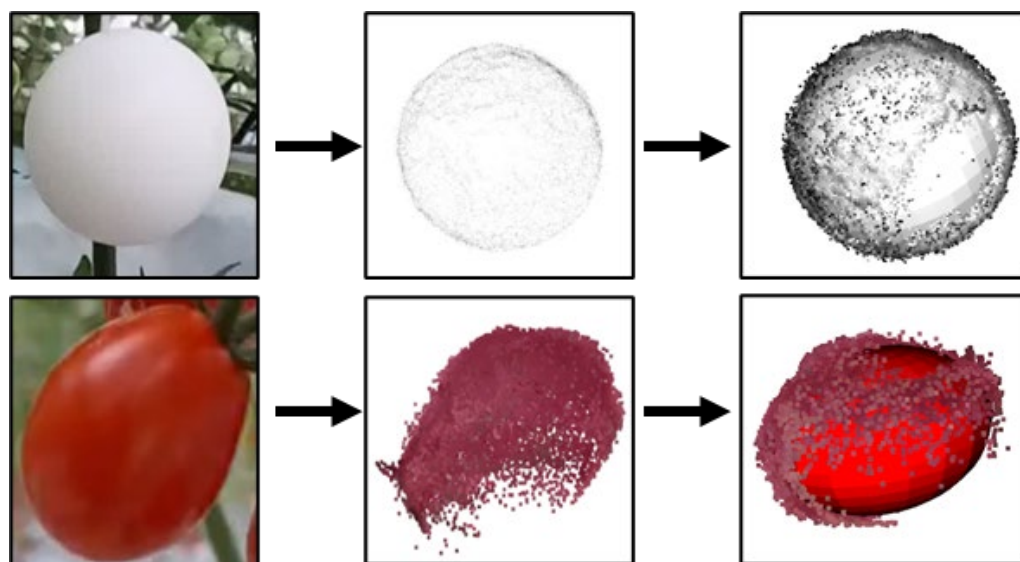


Figure 6. Shape fitting results: The upper section shows sphere fitting for the spherical reference ball, and the bottom section shows ellipsoid fitting for the tomato.

After measuring tomato size from 3D point cloud, the accuracy of the size estimation was evaluated by comparison with the actual size measured using calipers. The results indicate that the growth trend characteristics were similar, with an average percentage error of 8.01% (Figure 7). However, the size estimation slightly overestimated growth, although it remained within the acceptable criteria for agricultural use.

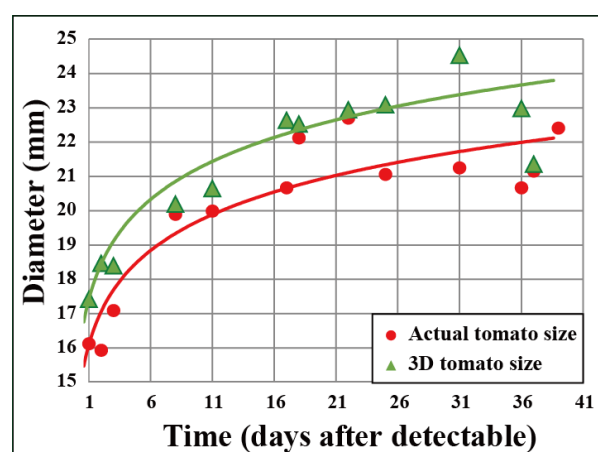


Figure 7. Evolution of fruit diameter over time: Graph comparing the growth of tomato size derived from 3D images with the actual size.

From Figure 7 The time (days) begins from when the tomato's size was first detectable by instance segmentation and continues until the fruit reached harvestable size. The growth curves in the graph were fitted using logarithmic approximation equations for comparison [12].

After evaluating the accuracy of the tomato size estimation, a logistic function was used for curve fitting to construct a transition graph of tomato growth for analysis based on growth principles, as presented in Section 4.5.

4.4. Assessment of Tomato Ripeness Status Based on Color Estimation

The results of the color-based ripeness estimation calculation are shown in Figure 8, following the method described in Section 3.7. After calculating the pixel-by-pixel color values, the data form a cluster of points distributed on the graph, where the x -axis represents the a^* values and the y -axis represents the b^* values. The slope was then determined using linear interpolation. For example, on 1 July, 2 July, 4 July, and 12 July, the calculated slope values were -0.3 , 0.03 , 0.83 , and 1.31 , respectively, with ripeness levels classified as classes 1, 2, 5, and 8, respectively.

After estimating the ripeness of tomatoes based on color values and classifying the ripeness levels, a logistic function was used for curve fitting to generate a tomato ripeness progression graph for the growth pattern analysis, as presented in Section 4.5.

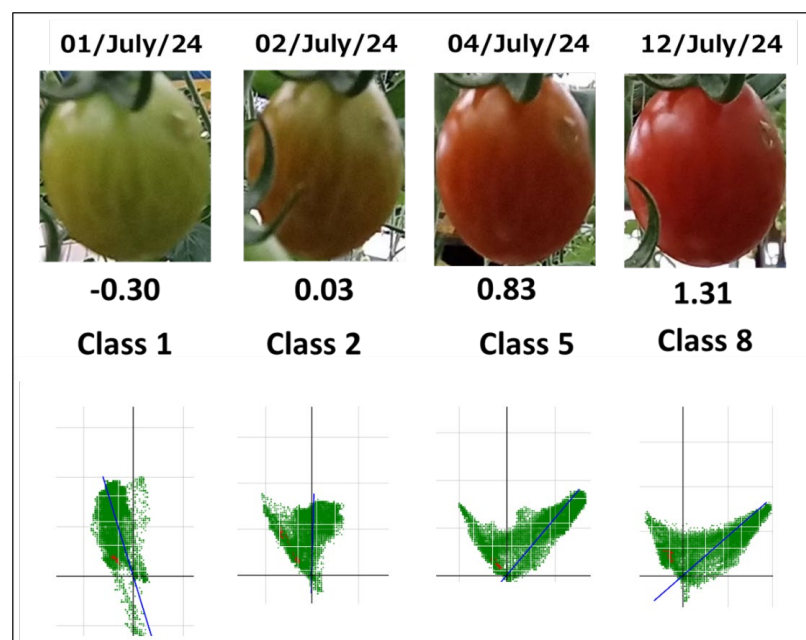


Figure 8. Results of daily color value calculations for tomatoes on a pixel-by-pixel basis, followed by slope calculation using linear approximation and classification into respective categories.

4.5. Growth Rate Estimation and Comparative Analysis Across Seasons

This section demonstrates a technique for estimating the tomato growth rate by analyzing the size and ripeness, and generating a graph using a logistic function, as shown in Figure 9.

Figure 9 shows the comparative growth patterns of the 11 tomatoes from the same bunch, with Tomatoes 1 and 11 representing the lowest and topmost tomatoes, respectively.

The graph shows a clear consistency between the data and the actual images: the top tomato grows faster than the bottom tomato. For example, the sample image shows that Tomato 11 is almost fully grown, whereas Tomato 1 remains green and immature. This aligns with the graph in which the top tomato exhibited a faster increase in size and ripeness than the bottom tomato.

Moreover, the graph shows that tomatoes grew rapidly in size during the first 40 days and remained relatively constant after day 45. In terms of ripeness, the tomatoes began to change color around day 20, progressing sequentially from the top to the bottom.

Various environmental factors significantly influenced the growth rate of tomatoes. The main factors in this study include daily temperature, humidity deficit, and solar radiation [33]. These factors varied seasonally, even under greenhouse conditions, particularly during winter and summer. When environmental data from these two periods were compared, the differences were clearly visible, as shown in Figure 10.

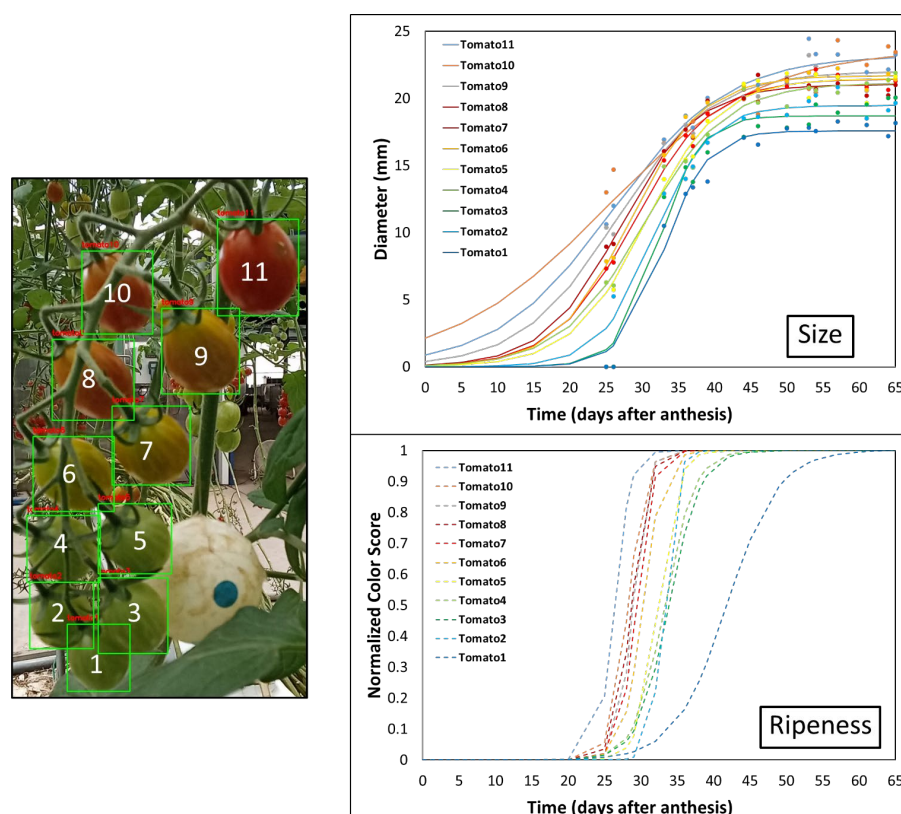


Figure 9. The estimated transition of individual tomato sizes and ripeness over time, using a logistic function for curve fitting.

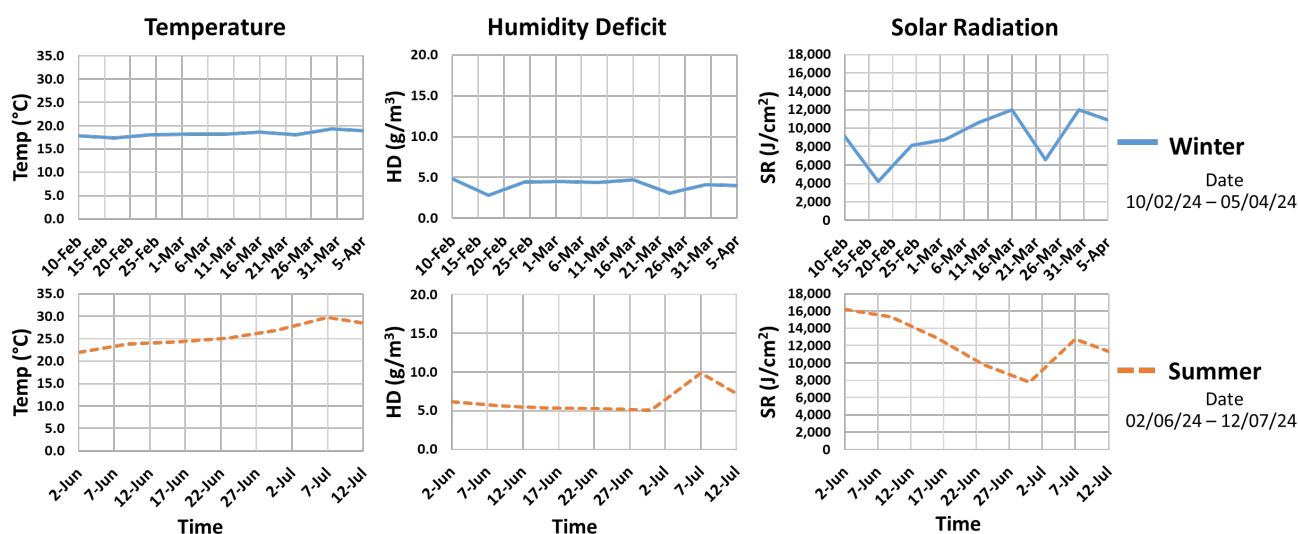


Figure 10. The measured data of four environmental factors inside the greenhouse were collected during winter and summer.

Figure 10, data collected in 2024, compares the measured environmental factors between winter and summer. The graph shows that these factors differ across seasons. Winter temperatures do not exceed 20 °C, while summer temperatures average 25.7 °C and reach a maximum of 30 °C. Temperatures decrease at night. Furthermore, humidity deficits are little higher in summer than in winter, and solar radiation levels are also higher in summer, with an average of 12,226 J/cm², compared to 9125 J/cm² in winter.

Different environmental factors in each season affect the growth period of tomatoes, both size and ripening period. By using the estimated transition method with the logistic function to analyze the growth rate of tomatoes in summer and winter, the differences can be compared more clearly. The relationship between growth size and ripeness, as well as the differences in tomato cultivation between summer and winter, are shown in Figure 11.

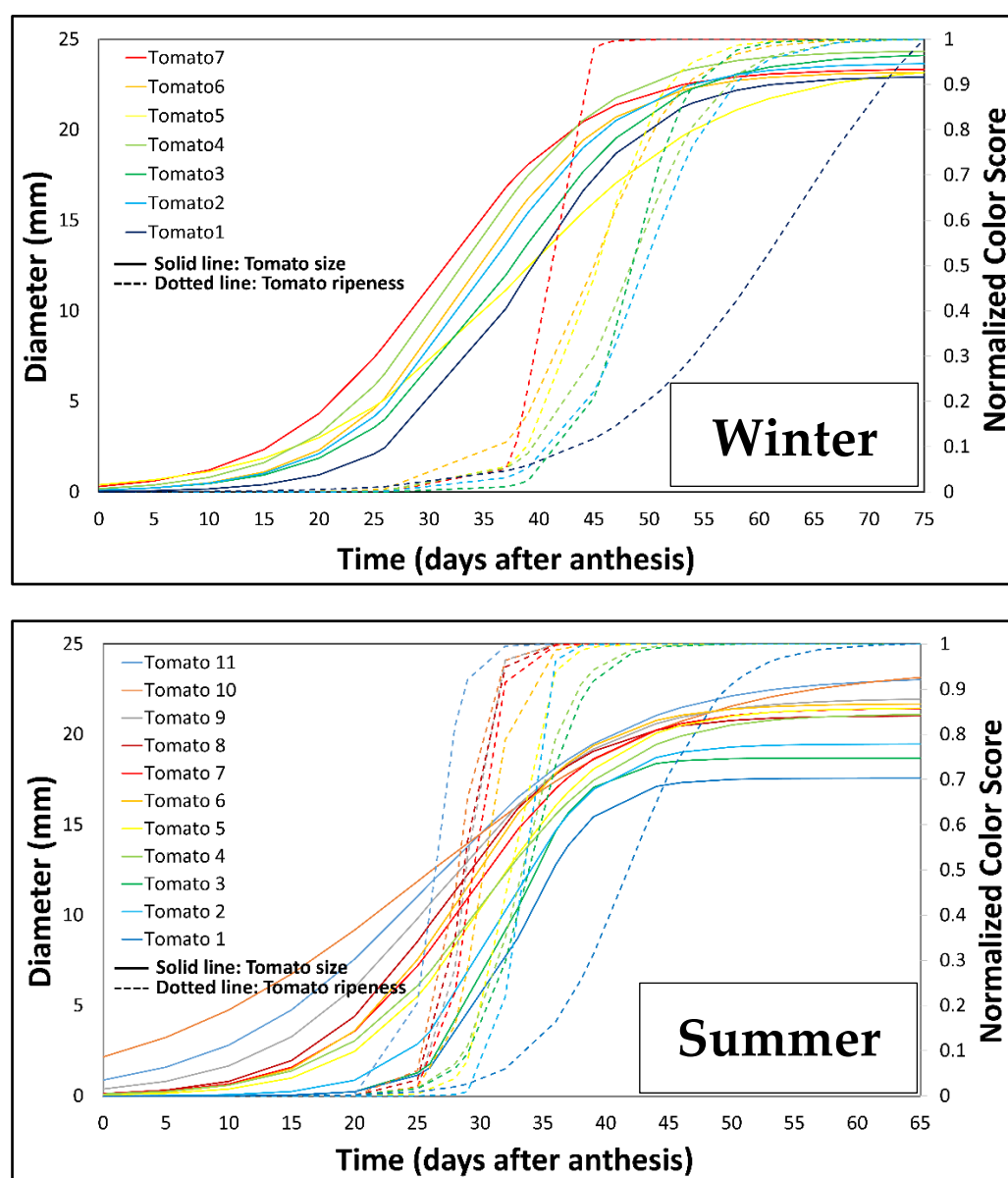


Figure 11. Graph comparing the relationship between tomato size and ripeness value, along with an analysis of growth differences between winter and summer.

Figure 11 illustrates the relationship between the tomato growth rate, size, and ripeness at each growth stage and shows that these features are correlated. For example,

tomatoes exhibit a rapid growth rate when they remain green until the point of color change. After the tomatoes were fully ripe, their size remained relatively constant.

Based on the calculation of $g_{\max}$ values using the method in Section 3.8, the $g_{\max}$ values of tomato size and tomato ripeness in summer are 1.08 and 0.41, respectively, while those in winter are 0.87 and 0.21, respectively. The comparison of the results from both seasons reveal that the growth rate of tomato size in summer was approximately 24.14% higher than in winter. In addition, tomatoes in summer ripened approximately 95.24% faster than in winter.

A comparison of tomato growth between winter and summer showed that tomatoes grew faster and changed color earlier in summer than in winter. Moreover, the growth rate in winter was approximately 10 days slower than that in summer.

5. Discussion

5.1. Instance Segmentation Framework

This study presents a novel instance segmentation framework for tomato tracking and 3D tomato phenotyping based on daily video images. The YOLOv8x-seg model demonstrated high segmentation performance, achieving a mAP of 0.881 at an IoU threshold of 0.8, confirming its effectiveness in distinguishing tomatoes at different ripening stages.

5.2. 3D Reconstruction and Morphological Analysis

A combination of SfM, MVS, and the Nerfacto framework was used to generate high-quality, high-resolution 3D point clouds of tomatoes. This reconstruction process enabled a detailed and precise analysis of fruit morphology. To separate the individual tomatoes from the point cloud data, the DBSCAN clustering algorithm was applied, which effectively grouped the points corresponding to each fruit for further analysis. A key contribution of this study is the implementation of ellipsoid fitting for 3D shape analysis. This noninvasive approach allows the estimation of important morphological traits, such as volume, surface area, and vertical diameter, which serve as reliable indicators of tomato growth. Compared with the image-based approach, the proposed 3D framework provides the natural shape and spatial structure of the fruit more effectively. However, these techniques can be applied to increase efficiency in smart and sustainable agriculture.

5.3. Tomato Feature Extraction: Size Estimation and Ripeness Assessment

To accurately assess tomato growth, metric calibration was performed using a spherical reference object, enabling reliable size estimation with an average error of 8.01% compared with manual caliper measurements over a 40-day cultivation period. This calibration improved the accuracy of the morphological traits derived from the ellipsoid fitting, such as volume and diameter, and provided precise size information.

The observed errors could be due to slight inaccuracies in manual caliper measurements. For example, when the data collector applies too much force to the instrument, soft-skinned tomatoes may collapse, causing size discrepancies. Additionally, the error may be caused by limitations in the 3D point cloud generation using daily video. To further reduce these errors, future plans will focus on validating and verifying the manual measurement process, and possibly considering the use of LiDAR technology for performance comparisons and potential application in 3D image generation.

For ripeness assessment, the CIELAB color space was utilized, focusing on the luminosity (L^*), red-green (a^*), and yellow-blue (b^*) components. After segmenting the tomatoes from the images, pixel-wise a^* and b^* values were extracted to generate scatter plots representing the color distribution. The slope of the linear interpolation through these

points quantified the color transition during ripening, serving as a robust numerical indicator of ripeness. This slope-based metric aligned well with the tomato color standard (Figure 4) and allowed classification into ripeness stages consistent with visual observations.

5.4. Growth Rate Estimation and Seasonal Comparison

Growth rate estimation using logistic function curve fitting effectively captured the dynamic development of individual tomatoes over time. The modeled growth patterns revealed a clear trend in which tomatoes positioned higher in the bunch exhibited faster increases in size and ripeness than those at the bottom (Figure 9). This observation aligned well with the actual images, validating the ability of the framework to reflect real-world growth variability within a single cluster.

The growth trajectory demonstrated a rapid increase in size during the initial 40 days, followed by a stabilization phase after day 45, consistent with typical fruit maturation stages. Similarly, ripeness progression, indicated by color changes starting around day 20, followed a top-to-bottom sequence within the bunch, reflecting biological developmental patterns.

The analysis across seasons reveals that in summer, the tomato growth rate was higher than in winter, with tomato size growing and ripeness faster by 24.14% and 95.24%, respectively.

Moreover, the seasonal comparison highlighted the significant environmental influences on tomato growth. Tomatoes cultivated in summer displayed faster growth rates and earlier color changes than those grown in winter, where maturation was delayed by approximately 10 days. This finding emphasizes the importance of considering seasonal effects in phenotyping studies, and supports the utility of the proposed system for adaptive management practices to optimize yield and quality across different growing conditions.

6. Conclusions

This study successfully developed a framework for 3D tomato phenotyping and growth monitoring using daily video recordings under real greenhouse conditions. Integrating instance segmentation (YOLOv8x-seg) and 3D reconstruction techniques (SfM, MVS, and Nerfacto) enabled the accurate morphological and ripeness analyses of individual tomato fruits. Ellipsoid fitting and metric calibration supported reliable size estimation, whereas ripeness classification was achieved through color analysis based on the CIELAB color space.

Logistic modeling of tomato growth revealed consistent developmental trends and highlighted seasonal differences in size and color progression, with summer-grown tomatoes maturing faster than those in winter. These findings demonstrate the potential of the proposed framework for supporting nondestructive, data-driven crop monitoring in precision agriculture.

Although the system performed well, some limitations remained, particularly regarding occlusions, lighting variability, and the need for manual corrections. Future enhancements will aim to improve automation, reduce the sensitivity to environmental variations, and expand the applicability of the system to other crops and growing conditions. These developments will be the key to realizing a fully practical and scalable phenotyping solution for smart farming. Despite the proposed system demonstrating high accuracy in instance segmentation and 3D reconstruction, some limitations remain. One major challenge is the occlusion caused by leaves and overlapping fruits, which can interfere with accurate detection and shape fitting. Although clustering and coloring techniques are used

to mitigate this problem, reconstruction errors can still occur, particularly when tomatoes are partially hidden or only partially visible in video frames.

In addition, variations in camera position and lighting owing to handheld recordings may introduce inconsistencies in 3D modeling and ripeness estimation. These variations may have affected the stability of the results on different days.

Furthermore, although many processes in the pipeline are automated, some steps such as object selection, shape verification, and tracking still require manual intervention. This reduces the efficiency of the system when applied over long periods of monitoring.

Future work will focus on improving the robustness of the system under actual greenhouse conditions, minimizing manual adjustments, and exploring methods for adapting the method to other tomato varieties and environmental conditions. These improvements will enhance the practicality and scalability of the proposed approach for broader agricultural applications.

Author Contributions: Conceptualization, W.T.; methodology, W.T., S.B. and S.W.; software, W.T.; validation, W.T., S.B. and S.W.; formal analysis, W.T. and S.W.; investigation, W.T.; resources, K.S. (Kyuki Shibuya), Y.G. and Y.H.; data curation, K.S. (Kyuki Shibuya), Y.G. and Y.H.; writing—original draft preparation, W.T.; writing—review and editing, W.T. and S.W.; visualization, W.T. and T.S.; supervision, S.K., Y.O., K.S. (Kazuhiko Sato), P.S.R., Y.L. and S.W.; project administration, S.W.; All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Council for Science, Technology and Innovation (CSTI), Cross-ministerial Strategic Innovation Promotion Program (SIP), the 3rd period of SIP “Creation of new ways of learning and working in a post-COVID-19 era,” Grant Number JPJ012347 (Funding agency: JST) and JSPS KAKENHI (grant number 22H02463).

Data Availability Statement: The data presented in this study are available upon request from corresponding author. The data are not publicly available because of privacy and ethical restrictions.

Conflicts of Interest: Author Tatsuki Sagawa was employed by the company Hitachi Solutions, Co., Ltd. Authors Kyuki Shibuya, Yoshiki Gama and Yoshiki Hatanaka were employed by the company Asai Nursery, Inc. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Maraveas, C. Image Analysis Artificial Intelligence Technologies for Plant Phenotyping: Current State of the Art. *AgriEngineering* **2024**, *6*, 3375–3407. <https://doi.org/10.3390/agriengineering6030193>.
2. Zhu, T.; Ma, X.; Guan, H.; Wu, X.; Wang, F.; Yang, C.; Jiang, Q. A calculation method of phenotypic traits based on three-dimensional reconstruction of tomato canopy. *Comput. Electron. Agric.* **2023**, *204*, 107515. <https://doi.org/10.1016/j.compag.2022.107515>.
3. Baar, S.; Kobayashi, Y.; Horie, T.; Sato, K.; Watanabe, S. Tomato fruit maturity estimation from RGB images. In Proceedings of the 2022 IEEE 11th Global Conference on Consumer Electronics (GCCE), Osaka, Japan, 18–21 October 2022; pp. 615–616.
4. Rossi, R.; Costafreda-Aumedes, S.; Leolini, L.; Leolini, C.; Bindi, M.; Moriondo, M. Implementation of an algorithm for automated phenotyping through plant 3D-modeling: A practical application on the early detection of water stress. *Comput. Electron. Agric.* **2022**, *197*, 106937. <https://doi.org/10.1016/j.compag.2022.106937>.
5. Verma, U.; Rossant, F.; Bloch, I. Segmentation and size estimation of tomatoes from sequences of paired images. *EURASIP J. Image Video Process.* **2015**, *2015*, 33. <https://doi.org/10.1186/s13640-015-0087-0>.
6. Mbouembe, P.L.T.; Liu, G.; Sikati, J.; Kim, S.C.; Kim, J.H. An efficient tomato-detection method based on improved YOLOv4-tiny model in complex environment. *Front. Plant Sci.* **2023**, *14*, 1150958. <https://doi.org/10.3389/fpls.2023.1150958>.
7. Zheng, S.; Liu, Y.; Weng, W.; Jia, X.; Yu, S.; Wu, Z. Tomato Recognition and Localization Method Based on Improved YOLOv5n-seg Model and Binocular Stereo Vision. *Agronomy* **2023**, *13*, 2339. <https://doi.org/10.3390/agronomy13092339>.
8. Lyu, Z.; Lu, A.; Ma, Y. Improved YOLOv8-Seg Based on Multiscale Feature Fusion and Deformable Convolution for Weed Precision Segmentation. *Appl. Sci.* **2024**, *14*, 5002. <https://doi.org/10.3390/app14125002>.

9. Yue, X.; Qi, K.; Na, X.; Zhang, Y.; Liu, Y.; Liu, C. Improved YOLOv8-Seg network for instance segmentation of healthy and diseased tomato plants in the growth stage. *Agriculture* **2023**, *13*, 1643. <https://doi.org/10.3390/agriculture13081643>.
10. Rose, J.C.; Paulus, S.; Kuhlmann, H. Accuracy analysis of a multi-view stereo approach for phenotyping of tomato plants at the organ level. *Sensors* **2015**, *15*, 9651–9665. <https://doi.org/10.3390/s150509651>.
11. Yu, Z.; Poching, T.; Aono, M.; Shimizu, Y.; Hosoi, F.; Omasa, K. 3D monitoring for plant growth parameters in field with a single camera by multi-view approach. *Agric. Meteorol.* **2018**, *74*, 129–139. <https://doi.org/10.2480/agrmet.D-18-00013>.
12. Baar, S.; Kobayashi, Y.; Horie, T.; Sato, K.; Kondo, S.; Watanabe, S. A logistic model for precise tomato fruit-growth prediction based on diameter-time evolution. *Comput. Electron. Agric.* **2024**, *227*, 109500. <https://doi.org/10.1016/j.compag.2024.109500>.
13. Miranda, J.C.; Gené-Mola, J.; Zude-Sasse, M.; Tsoulas, N.; Escolà, A.; Arnó, J.; Rosell-Polo, J.R.; Sanz-Cortiella, R.; Martínez-Casasnovas, J.A.; Gregorio, E. Fruit sizing using AI: A review of methods and challenges. *Postharvest Biol. Technol.* **2023**, *206*, 112587. <https://doi.org/10.1016/j.postharvbio.2023.112587>.
14. Wang, X.; Vladislav, Z.; Viktor, O.; Wu, Z.; Zhao, M. Online recognition and yield estimation of tomato in plant factory based on YOLOv3. *Sci. Rep.* **2022**, *12*, 8686. <https://doi.org/10.1038/s41598-022-12732-1>.
15. Gómez, R.; Varón, R.; Amo, M.; Tardaguila, J.; Pardo, J.E. Differences in the rate of coloration in tomato fruit. *J. Food Qual.* **1998**, *21*, 329–339. <https://doi.org/10.1111/j.1745-4557.1998.tb00526.x>.
16. Wang, Y.; Hu, S.; Ren, H.; Yang, W.; Zhai, R. 3DPhenoMVS: A low-cost 3D tomato phenotyping pipeline using 3D reconstruction point cloud based on multiview images. *Agronomy* **2022**, *12*, 1865. <https://doi.org/10.3390/agronomy12081865>.
17. Roshan, T.R.; Jafari, M.; Golami, M.; Kazemi, M. Evaluating geometric measurement accuracy based on 3D model reconstruction of nursery tomato plants by Agisoft photscan software. *Comput. Electron. Agric.* **2024**, *221*, 109000. <https://doi.org/10.1016/j.compag.2024.109000>.
18. Lindenberger, P.; Sarlin, P.-E.; Larsson, V.; Pollefeys, M. Pixel-perfect structure-from-motion with featuremetric refinement. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, BC, Canada, 11–17 October 2021; pp. 5987–5997.
19. Maillard, A.; Kunisky, D. Fitting an Ellipsoid to Random Points: Predictions Using the Replica Method. *IEEE Trans. Inf. Theory* **2024**, *70*, 7273–7296. <https://doi.org/10.1109/TIT.2024.3415654>.
20. Gené-Mola, J.; Sanz-Cortiella, R.; Rosell-Polo, J.R.; Escolà, A.; Gregorio, E. In-field apple size estimation using photogrammetry-derived 3D point clouds: Comparison of 4 different methods considering fruit occlusions. *Comput. Electron. Agric.* **2021**, *188*, 106343. <https://doi.org/10.1016/j.compag.2021.106343>.
21. Sun, S.; Li, C.; Chee, P.W.; Paterson, A.H.; Jiang, Y.; Xu, R.; Robertson, J.S.; Adhikari, J.; Shehzad, T. Three-dimensional photogrammetric mapping of cotton bolls in situ based on point cloud segmentation and clustering. *ISPRS J. Photogramm. Remote Sens.* **2020**, *160*, 195–207. <https://doi.org/10.1016/j.isprsjprs.2019.12.011>.
22. Vanbrabant, Y.; Delalieux, S.; Tits, L.; Pauly, K.; Vandermaesen, J.; Somers, B. Pear Flower Cluster Quantification Using RGB Drone Imagery. *Agronomy* **2020**, *10*, 407. <https://doi.org/10.3390/agronomy10030407>.
23. Choi, H.B.; Park, J.K.; Park, S.H.; Lee, T.S. NeRF-based 3D reconstruction pipeline for acquisition and analysis of tomato crop morphology. *Front. Plant Sci.* **2024**, *15*, 1439086. <https://doi.org/10.3389/fpls.2024.1439086>.
24. Zheng, X.; Xinyi, A.; Qin, H.; Rong, J.; Zhang, Z.; Yang, Y.; Yuan, T.; Li, W. Tomato-nerf: Advancing tomato model reconstruction with improved neural radiance fields. *IEEE Access* **2024**, *12*, 184206–184215. <https://doi.org/10.1109/ACCESS.2024.3424908>.
25. Ambrus, B.; Teschner, G.; Kovács, A.J.; Neményi, M.; Helyes, L.; Pék, Z.; Takács, S.; Alahmad, T.; Nyéki, A. Field-grown tomato yield estimation using point cloud segmentation with 3D shaping and RGB pictures from a field robot and digital single lens reflex cameras. *Heliyon* **2024**, *10*, e37997. <https://doi.org/10.1016/j.heliyon.2024.e37997>.
26. Kellner, M.; Stahl, B.; Reiterer, A. Fused Projection-Based Point Cloud Segmentation. *Sensors* **2022**, *22*, 1139. <https://doi.org/10.3390/s22031139>.
27. Yang, J.; Lee, C.; Ahn, P.; Lee, H.; Yi, E.; Kim, J. PBP-Net: Point Projection and Back-Projection Network for 3D Point Cloud Segmentation. In Proceedings of the 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Las Vegas, NV, USA, 24 October 2020–24 January 2021; pp. 8469–8475.
28. Kneip, L.; Scaramuzza, D.; Siegwart, R. A novel parametrization of the perspective-three-point problem for a direct computation of absolute camera position and orientation. In Proceedings of the CVPR 2011, Colorado Springs, CO, USA, 20–25 June 2011; pp. 2969–2976.
29. Guo, J.; Xu, L.; Reinoso, O. Automatic Segmentation for Plant Leaves via Multiview Stereo Reconstruction. *Math. Probl. Eng.* **2017**, *2017*, 9845815. <https://doi.org/10.1155/2017/9845815>.

30. Poudel, S.; Aryal, P.; Basnet, M. Effect of different packaging materials on shelf life and postharvest quality of tomato (*Lycopersicum esculentum* var. Srijana). *Adv. Hortic. Sci.* **2022**, *36*, 127–134. <https://doi.org/10.36253/ahsc-12801>.
31. Fang, S.-L.; Kuo, Y.-H.; Kang, L.; Chen, C.-C.; Hsieh, C.-Y.; Yao, M.-H.; Kuo, B.-J. Using Sigmoid Growth Models to Simulate Greenhouse Tomato Growth and Development. *Horticulturae* **2022**, *8*, 1021. <https://doi.org/10.3390/horticulturae8111021>.
32. Trigubenko, R., Xu, D.; Fujihara, H. Laboro Tomato Dataset. Available online: <https://github.com/laboroai/LaboroTomato> (accessed on 2 October 2025).
33. Doan, C.C.; Tanaka, M. Relationships between tomato cluster growth indices and cumulative environmental factors during greenhouse cultivation. *Sci. Hortic.* **2022**, *295*, 110803. <https://doi.org/10.1016/j.scienta.2021.110803>.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.