

Article

Low-Cost Lettuce Height Measurement Based on Depth Vision and Lightweight Instance Segmentation Model

Yiqiu Zhao, Xiaodong Zhang, Jingjing Sun, Tingting Yu, Zongyao Cai, Zhi Zhang and Hanping Mao *

School of Agricultural Engineering, Jiangsu University, Zhenjiang 212013, China;
2112216012@stmail.ujs.edu.cn (Y.Z.); zxd700227@ujs.edu.cn (X.Z.); 2222016050@stmail.edu.cn (J.S.);
2222216048@stmail.ujs.edu.cn (T.Y.); 2222216051@stmail.ujs.edu.cn (Z.C.); zzhi@stmail.ujs.edu.cn (Z.Z.)
* Correspondence: maohp@ujs.edu.cn

Abstract: Plant height is a crucial indicator of crop growth. Rapid measurement of crop height facilitates the implementation and management of planting strategies, ensuring optimal crop production quality and yield. This paper presents a low-cost method for the rapid measurement of multiple lettuce heights, developed using an improved YOLOv8n-seg model and the stacking characteristics of planes in depth images. First, we designed a lightweight instance segmentation model based on YOLOv8n-seg by enhancing the model architecture and reconstructing the channel dimension distribution. This model was trained on a small-sample dataset augmented through random transformations. Secondly, we proposed a method to detect and segment the horizontal plane. This method leverages the stacking characteristics of the plane, as identified in the depth image histogram from an overhead perspective, allowing for the identification of planes parallel to the camera's imaging plane. Subsequently, we evaluated the distance between each plane and the centers of the lettuce contours to select the cultivation substrate plane as the reference for lettuce bottom height. Finally, the height of multiple lettuce plants was determined by calculating the height difference between the top and bottom of each plant. The experimental results demonstrated that the improved model achieved a 25.56% increase in processing speed, along with a 2.4% enhancement in mean average precision compared to the original YOLOv8n-seg model. The average accuracy of the plant height measurement algorithm reached 94.339% in hydroponics and 91.22% in pot cultivation scenarios, with absolute errors of 7.39 mm and 9.23 mm, similar to the sensor's depth direction error. With images downsampled by a factor of 1/8, the highest processing speed recorded was 6.99 frames per second (fps), enabling the system to process an average of 174 lettuce targets per second. The experimental results confirmed that the proposed method exhibits promising accuracy, efficiency, and robustness.

Keywords: RGB-D; plant height; machine vision; lightweight network; YOLOv8



Citation: Zhao, Y.; Zhang, X.; Sun, J.; Yu, T.; Cai, Z.; Zhang, Z.; Mao, H. Low-Cost Lettuce Height Measurement Based on Depth Vision and Lightweight Instance Segmentation Model. *Agriculture* **2024**, *14*, 1596. <https://doi.org/10.3390/agriculture14091596>

Academic Editor: Adrian Gracia-Romero

Received: 10 August 2024
Revised: 9 September 2024
Accepted: 12 September 2024
Published: 13 September 2024



Copyright: © 2024 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

Plant monitoring is essential for agronomists and breeders in assessing plant growth and effectively managing fields. Throughout the growth process, various morphological and physiological traits of plants are systematically monitored [1,2]. Notably, plant height is commonly used as a key indicator of field vitality, biomass estimation, and crop yield prediction [3]. The tender leaves of lettuce are vulnerable to injury in the process of mechanical harvesting. The size change of lettuce will affect the success rate of harvesting. The height information of lettuce also has important reference value for the intelligent harvesting of lettuce [4]. In contrast to traditional manual measurement techniques, non-contact visual measurement technology has gained popularity for crop characterization, thanks to its advantages of minimal damage, high efficiency, and exceptional precision [5].

In plant phenotype detection using machine vision, the crop recognition task is a critical component, as the accuracy of recognition directly impacts the reliability of phenotypic assessments. Due to the unstructured nature of scenes and objects in agricultural

settings [6], machine learning techniques, particularly those based on deep learning, frequently outperform traditional methods that rely solely on color, texture, or other features for identification [7]. Consequently, the application of deep learning technology has become standard [8] for various tasks, including crop identification and segmentation [9–13], weed detection [14], plant pest detection [15], yield prediction and estimation [16], classification and sequencing [17], soil analysis [18], and real-time decision-making [19]. YOLO series models, recognized as state-of-the-art (SOTA) models in this field, demonstrate excellent accuracy, performance, and adaptability to complex environments in agricultural settings [20–24]. In typical deep learning model design tasks, there are often dozens of classes to consider. However, agricultural scenarios usually require only one kind of crop and background. Consequently, optimizing deep learning models for agricultural applications primarily revolves around enhancing model efficiency. This optimization is achieved through methods such as utilizing lightweight backbone networks and modifying various model components [25–27]. Additionally, the construction of datasets consumes a significant portion of the time involved in model building. One research direction in deep learning technology focuses on developing usable models that can be trained on small sample sizes, ultimately reducing manual workload. In contrast to designing models specifically tailored for small sample sizes, the method of expanding datasets through sample augmentation can be applied across a broader range of models, further alleviating the burden of model design [28].

Three-dimensional crop data can be used to obtain the morphological structure and position coordinates of lettuce, such as those collected using depth cameras and lidar. These devices have the advantage of not being affected by perspective distortion [29–31]. The existing measurement methods of plant height can be divided into deep learning methods [32–34] and numerical calculation methods [35,36]. The numerical calculation method is mainly based on the difference method [37], and the plant height is obtained by identifying the canopy top and bottom of the plant and calculating the difference. Compared with deep learning methods, numerical methods have many advantages in simple tasks, such as their high operating efficiency and stability. The position of the plant's canopy is typically visible within the camera's field of view, whereas the position of the plant bottom may be obscured by the surrounding crops. Consequently, the height of the plant bottom must be estimated using indirect methods. One common approach to calculating crop plant height involves determining the height based on the sensor's installation position [38]. The height of plant bottom can be inferred from the relationship between the sensor's height and the position of the plant bottom. This method is commonly applied using fixed platforms; however, motion-related errors associated with moving platforms can diminish the accuracy of this approach. Multi-view stereo (MVS) [39] is a widely used offline method for measuring crop plant height, relying on a substantial amount of data to reconstruct the complete three-dimensional structure of plants. The declivity survey [40–42] is similar to MVS but offers a broader measurement range. However, its longer sight distance limits measurement accuracy, making it less suitable for applications requiring millimeter-level precision. Plane fitting [43] can be used to estimate the height of the plant bottoms when they are not visible; however, it is not well-suited for environments with significant height variations. Additionally, commonly employed multi-plane detection algorithms [44,45] substantially increase operational requirements. While the use of multiple sensors for imaging [46] can compensate for the limitations of a single sensor's field of view, it also raises both cost and spatial requirements.

In this paper, we propose a low-cost, rapid, multi-objective algorithm for measuring lettuce plant height that is suitable for application in multi-plane environments. Based on the YOLOv8n-seg model, we have built a lightweight instance segmentation model for lettuce through small sample augmentations, dataset expansion, integration of a lightweight backbone, and progressive redesign of the model's layer channel dimensions. Our algorithm identified the planes parallel to the camera's image plane by analyzing the accumulation of plane pixel values along the depth axis of the RGB-D depth data. We then determined the

cultivation substrate planes by measuring the distance between the candidate plane pixels and the center of the lettuce. Finally, we conducted experiments to evaluate the accuracy and performance of the lettuce plant height measurement algorithm in two scenarios, hydroponics and pot cultivation, using different sampling ratios for the sensor data.

2. Materials and Methods

2.1. Data

The data were collected in July 2024 at the glasshouse of Ganghuazijing Farmstead in Zhenjiang, China (latitude $32^{\circ}8'27''$ N, longitude $119^{\circ}8'21''$ E), as shown in Figure 1. The lettuce plants consisted of wrinkled leaf varieties in the hydroponics scenario and creaseless-leaf lettuce in the potting scenario. Data collection occurred between 8:00 AM and 6:00 PM on clear days. The plant spacing of the lettuce in the hydroponics scenario was 200 mm, and the plant spacing of the potted lettuce was 300 mm. The growth period of lettuce is 15–20 days after germination, and it grows well without disease. When collecting, it was ensured that there was no water shortage in the lettuce.



Figure 1. Lettuce growing environment.

Image data were collected using RGB-D cameras, and the camera information is shown in Table 1. Image data of the hydroponics scenario were acquired using Orbbec Gemini 336L (Orbbec, Shenzhen, China), including RGB and registered depth sensors, with a resolution of 1280×800 . During collection, the camera was positioned approximately 1000–1100 mm above the cultivation substrate from a top-down perspective. Image data of the potting scenario were acquired using Realsense D435 (Intel, Santa Clara, CA, the United States), including RGB and registered depth sensors, with a resolution of 1280×720 . During collection, the camera was positioned approximately 800–900 mm above the cultivation substrate from a top-down perspective.

Table 1. Camera information.

	Orbbec Gemini 336L	Realsense D435
Application scene	Hydroponics	Potting
Spatial error	<0.8% at 2 m	<2% at 2 m
Resolution	1280×800	1280×720
Field of view	$90^{\circ} \times 65^{\circ}$	$87^{\circ} \times 58^{\circ}$
Frame rate	30	30

The spatial error is the error of the depth distance calculated by the camera at a certain sight distance, which was provided by the manufacturer.

The color and height distribution of the background of the hydroponics scene are relatively simple, but the sensor field of view contains a number of rice crops, which can be used to test the ability of the segmentation model to distinguish between different plant categories. The potting scene takes the soil as the background and contains only a small number of weeds. However, the potting scenes have more complex height distribution

relationships in the background, which can highlight the performance of the cultivation matrix plane detection algorithm in complex environments.

The ground truth value of the crop plant height was measured manually. The measurement tool, depicted in Figure 2, features a transparent board that slides vertically and a vertically mounted steel ruler. The precision of the steel ruler is 0.5 mm for measurements between 0 and 100 mm and 1 mm for measurements between 100 and 500 mm.

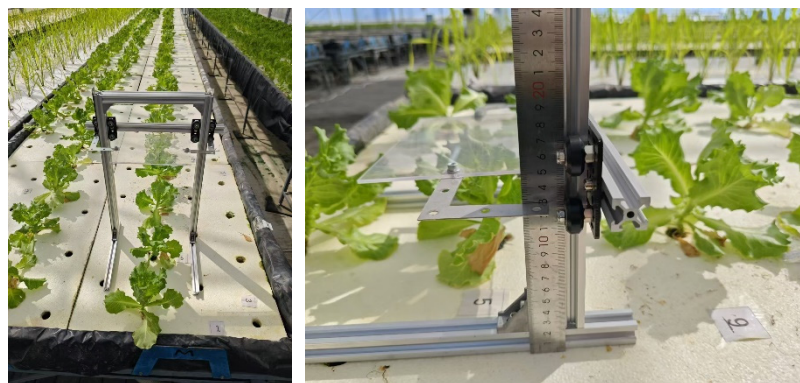


Figure 2. Plant height measurement tool.

Random transformation augmentation of datasets was employed to simulate variations in lighting conditions, fields of view, and viewing angles. The random transform augmentation included 11 methods: brightness transformation, contrast transformation, crop transformation, flip transformation, Gauss blur transformation, perspective transformation, resize transformation, rotation transformation, shadow transformation, spot transformation, and translate transformation. The methods of random transformations used in this paper are illustrated in Figure 3.

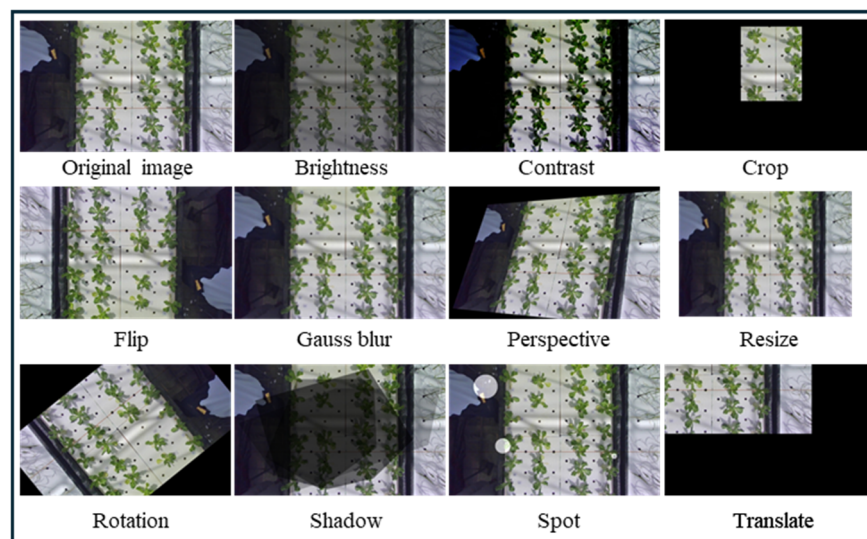


Figure 3. Examples of random transformations.

The random value ranges of the key parameters for the different methods are summarized in Table 2. Among these, five transformations were used to simulate changes in the light environment, including brightness, contrast, Gauss blur, shadow, and spot. The other six were utilized to simulate modifications in acquisition modes and sensor parameters, including crop, flip, perspective, resize, rotation, and translate.

Table 2. Parameters for random transformations.

Type	Parameters
Brightness	$I_t = \varphi_B \times I$, $\varphi \in [0.5-1]$
Contrast	$I_{Ct} = (255 \times \varphi_{Ct}) / (1 + e^{I-128})$, $\varphi_{Ct} \in [16, 32]$
Crop	$A_{cr} = \varphi_{Cr} \times A$, $\varphi_{Cr} \in [0.1, 1]$
Flip	One of left–right or up–down flipping, or both.
Gauss blur	Kernel size in [3, 9] pixels.
Perspective	X-axis and y-axis transform angles are in the range of $[-60, 60]$ degrees with a step size of 1.
Resize	X-axis and y-axis transform scales are in the range of [0.5, 1.5].
Rotation	Rotation angle of [0, 359] degrees with the step size of 1.
Shadow	Shadow contours number in [1, 3].
Spot	Spot number in [1, 10], spot size in [0, 160] pixels, transparency in [0.5, 0.7].
Translate	X-axis and y-axis transform scale range of $[-0.5, 0.5]$.

I is the image pixel value; I_t , I_{Ct} , are image pixel values after transformation; A is the image area size; A_{cr} is the image area size after transformation; φ_B , φ_{Ct} , φ_{Cr} are the transform ratios in brightness, contrast, and crop transformation, respectively.

Based on no more than 100 manually labeled images, three datasets were constructed, as shown in Table 3. The 100 manually labeled images were labeled by Labelme with lettuce contours as targets, including 50 hydroponic scenes and 50 potted scenes. The training set, validation set, and test set were divided by a random sampling algorithm. Dataset 1 consists of an untransformed training set of 60 images and a validation set of 20 images. Dataset 2 was created from the images in Dataset 1, applying a single type of transformation to each image, resulting in a total number of images equal to that of Dataset 1. Dataset 3 was generated from Dataset 1 through composite random transformations, with each original image yielding 11 new images after undergoing 11 different types of random transformations. Additionally, 20 images that were not included in the training or validation sets were used as an independent test dataset to compare the model’s training results.

Table 3. Dataset constructions.

Dataset	Total No. of Images	No. Images in Training Set	No. Images in Val Set	Train Targets	Val Targets	Random Transform
Dataset 1	80	60	20	1530	559	None
Dataset 2	80	60	20	1378	388	11 types
Dataset 3	880	660	220	15,572	5727	11 types

The instance segmentation model training device was equipped with an AMD Ryzen 7700 processor (AMD, Santa Clara, CA, USA), an RTX 2070 graphics card (Nvidia, Santa Clara, CA, USA), and 32 GB of memory (KingBank, Shenzhen, China). To verify the performance of the method using a low-cost device, the instance segmentation model inference and plant height measurement algorithm were tested using only the CPU.

2.2. Lightweight Instance Segmentation Model

The identification and segmentation of lettuce targets are prerequisites for multi-target plant height measurement. Vegetation index is a commonly used plant segmentation method [47]. It requires a small amount of calculation and usually has a good segmentation effect, but it is difficult to distinguish between plants and backgrounds in a complex light environment. In this paper, the YOLOv8n-seg instance segmentation model is employed to achieve multi-object recognition and segmentation of lettuce in a controlled environment. The model structure of YOLOv8n-seg utilized in this study is illustrated in Figure 4.

Compared to the original model based on the COCO dataset, which contains 80 classes, the requirement for the recognition task of identifying a single class is lower. This created an opportunity for optimizing the model’s efficiency. FasterNet is a lightweight network architecture that reduces the computational requirements of each layer in the network

by utilizing P-Conv [48]. In this paper, FasterNet is used to replace CSP DarkNet in the original YOLOv8n-seg model to enhance efficiency. Its structure is illustrated in Figure 5.

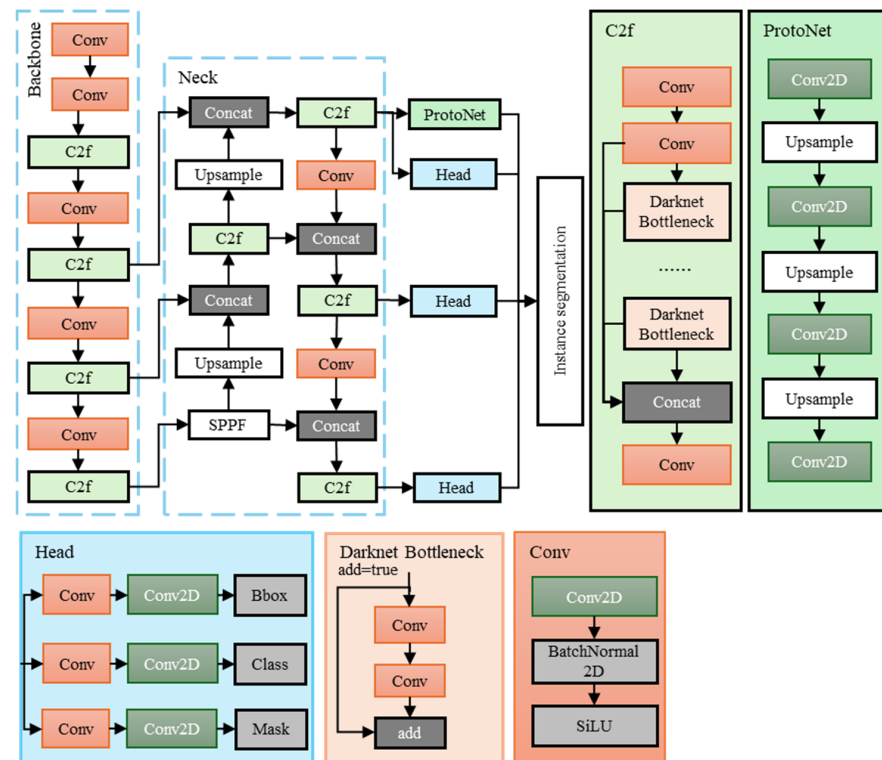


Figure 4. YOLOv8n-seg structure.

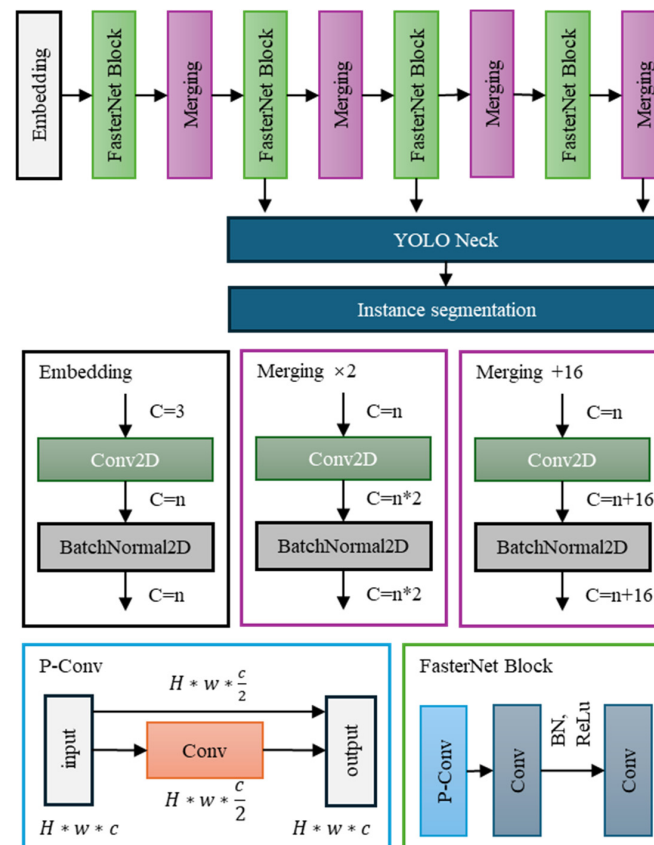


Figure 5. Structure of YOLOv8-seg with FasterNet as backbone.

To further optimize the model's efficiency, the inter-layer channel dimensions' growth relationship was reevaluated. The original exponential growth was replaced with the cumulative sum, effectively reducing the growth rate of inter-layer channel dimensions. This adjustment minimizes the number of parameters and enhances the model's efficiency [49].

2.3. Measurement of Lettuce Plant Bottom Height

Plant height is typically measured from the bottom to the top, with the position of the plant bottom generally indicated by the flat surface of the cultivation substrate. Consequently, in this paper, the measurement of the cultivation substrate plane is used as a substitute for measuring the plant bottom.

The distribution of depth image pixels along the depth axis, along with its histogram, is shown in Figure 6, which depicts a plane parallel to the sensor's imaging plane. The surface of the cultivation substrate is typically horizontal, and when the sensor is oriented vertically downward, the surface of the substrate will be parallel to the camera's image plane.

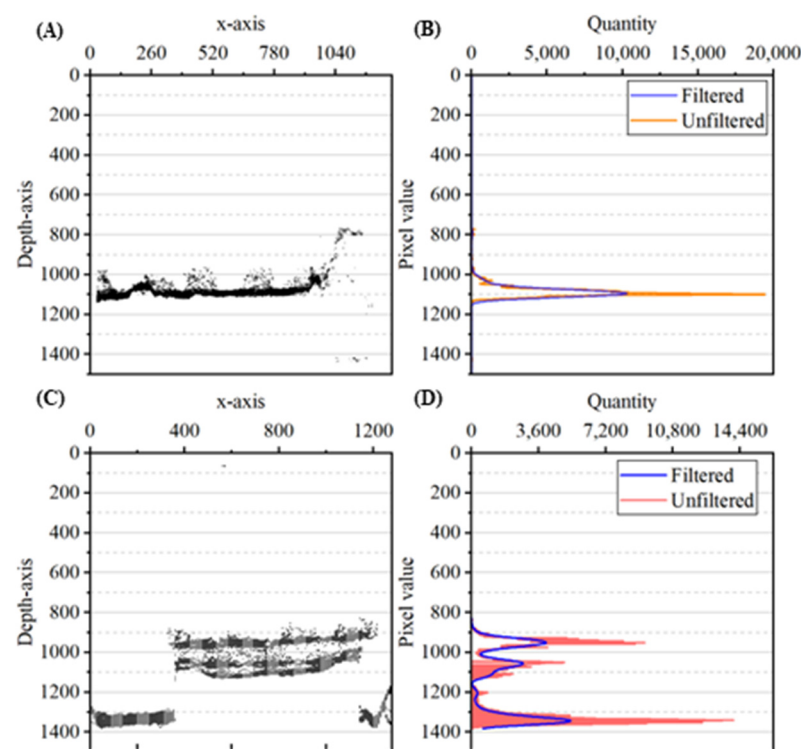


Figure 6. Hydroponics scenario: (A) distribution of depth image pixels along the depth axis, (B) histogram of depth image. Potting scenario: (C) distribution of depth image pixels along the depth axis, (D) histogram of depth image.

The detection of the cultivation substrate plane involves the following steps: First, the histogram is calculated from the depth image after removing the lettuce contour mask, as indicated by the red line in Figure 6B. A sliding window of length 30 is then applied for two rounds of mean filtering to eliminate the small amount of noise in the histogram, shown as an unfiltered line in Figure 6B,D, with the filtered signal represented by the blue line in Figure 6B,D. Finally, the wave peak point in the filtered histogram is identified as the plane depth, and the plane area is extracted within a depth value range of ± 50 . The results of this extraction are shown in Figure 7.

Each pixel was assigned to the nearest lettuce based on the distance between the pixel and the center of each lettuce. In each region, the distances between the pixels in the plane and the center of the corresponding plant were calculated. The distance (d) between a point

$P(x_p, y_p)$ in the plane and the center point $C(x_c, y_c)$ of the crop to which it belongs is given as follows:

$$d = \sqrt{(x_p - x_c)^2 + (y_p - y_c)^2} \quad (1)$$

The distance D between a plane containing n points and the center of the lettuce is given as follows:

$$D = \frac{\sum_{i=0}^n d_i}{n} \quad (2)$$

The average distance was calculated by merging the results from each plane sequentially, and the plane with the smallest average distance was identified as the cultivation substrate plane of the crop.

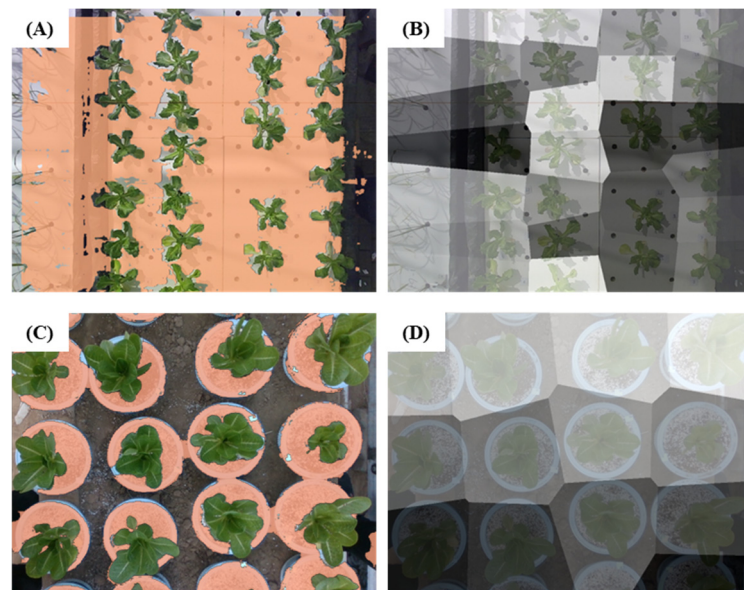


Figure 7. (A,C) Results of plane detection based on pixel stacking. (B,D) Image region division based on crop center.

2.4. Multiple Lettuce Height Measurements

A schematic diagram of the algorithm flow is shown in Figure 8. The height of each lettuce plant was calculated individually using the difference method. After removing zero values from the data, the 1% of the pixels with the lowest depth values in each crop region were discarded to eliminate discrete data in the image. Subsequently, the pixel with the lowest depth value within the crop contour was selected as the top point.

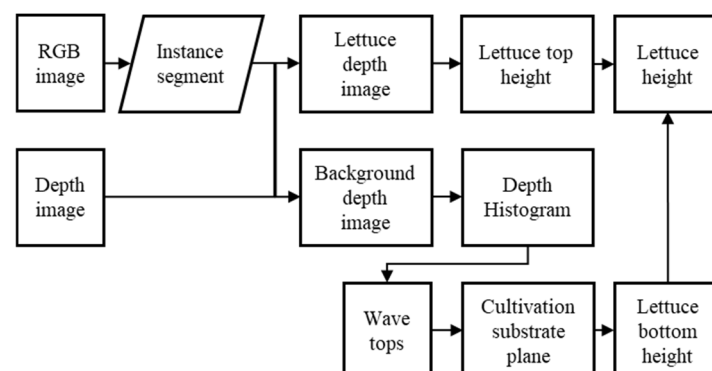


Figure 8. Algorithm flow diagram.

The modal depth value of the selected plane in each pixel region was taken as the depth of the lettuce bottom height. If the selected plane did not exist within the plant's

region, the modal depth value from the entire image was used as the depth value for the lettuce bottom height.

The height of each lettuce plant was then determined by subtracting the depth value of the bottom from the top point. Each pixel region was traversed sequentially to measure the heights of all lettuce plants in the image.

3. Experimental Details

3.1. Model Training

A comparison of the training models is given in Table 4. Model 1 is the YOLOv8n-seg model, trained using Dataset 1. Model 2 employs FasterNet as the backbone and was also trained on Dataset 1. Models 3, 4, and 5 utilized FasterNet as the backbone with a custom accumulation strategy to readjust the layer channel, trained on Dataset 1, Dataset 2, and Dataset 3, respectively. Using the default training parameters for YOLOv8, mosaic augmentation was turned off, and each model was trained for 200 epochs. In addition to these five models, two additional models using MobileNetv3 and ShuffleNetv2 as backbones, both trained on Dataset 1, were compared as Models 6 and 7.

Table 4. Models for comparative training.

Model	Dataset	Backbone	Channel Dimension Growth Relationships
Model 1	Dataset 1	CSP DarkNet	$+2^n$
Model 2	Dataset 1	FasterNet T0	$+2^n$
Model 3	Dataset 1	FasterNet T0	+16
Model 4	Dataset 2	FasterNet T0	+16
Model 5	Dataset 3	FasterNet T0	+16
Model 6	Dataset 1	MobileNetv3	Unmodified
Model 7	Dataset 1	ShuffleNetv2	Unmodified

The layer channel dimensions of each model are presented in Figure 9. After the number of channel dimensions of Model 2 increased to the sixth layer with a power multiple of 2 ($+2^n$, where n is the number of times the channel dimension increases), the growth of the sixth to eighth layers was the same as that of the fourth to sixth layers. Compared with Model 1, whose channel dimension growth of 0–8 layers all changed by a power multiple of 2, Model 2 had fewer channel dimensions as a whole. After redesigning the model in this paper, Model 3, with all channel dimensions changing with fixed growth in 16 ($+16$), had the fewest channel dimensions; that is, the smallest model size.

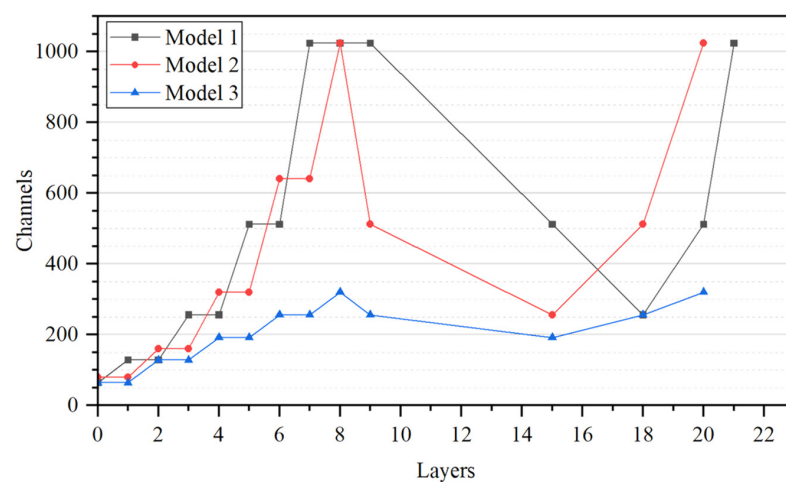


Figure 9. Model channel dimension comparisons (before multiplying the width coefficient of the model).

3.2. Lettuce Height Measurement Testing

The plant height measurement algorithm was tested in hydroponic and potting scenarios used 95 targets and 81 targets, respectively. The plant heights measured via manual measurements were compared. When measured manually, the plant heights in the hydroponic scene were recorded from the upper surface of the hydroponic board to the highest point of the plant, and the plant heights in the pot scene were recorded from the soil surface in the pot to the highest point of the plant.

Additionally, due to the numerous operations involved, the speed and accuracy of the algorithm were tested and compared under different downsampling ratios (based on the length of the side). The OpenCV resize function was used to downsample the images using the nearest neighbor interpolation method, ensuring that the pixel values remained unchanged. The processing speed of the algorithm was tested using hydroponic scene data. When testing, the downsampled data were obtained by multiplying the side length resolution of the image by the downsampling coefficient. The comparison size included images with downsampling coefficients of 1, 0.7, 0.5, and 0.354. These corresponded to about 100%, 50%, 25%, and 12.5% of the original size area, respectively.

3.3. Evaluation Indicators

The performance of the lettuce height measurement algorithm was evaluated based on the relative precision (P_h), correlation coefficient (R^2), and root mean square error (RMSE). The equation for P_h is as follows:

$$P_h = 1 - \frac{|H_m - H_g|}{H_g} \times 100\% \quad (3)$$

where H_m is the measurement results of the algorithm and H_g is the lettuce height ground truth based on manual measurement.

The performance of the lettuce instance segment model was evaluated based on precision (P), recall (R), average precision (AP), model inference time, model parameters size, and calculation amount. The equations of P , R , and AP are as follows:

$$P = \frac{TP}{TP + FN} \quad (4)$$

$$R = \frac{TP}{TP + FP} \quad (5)$$

$$AP = \frac{\sum_{k=1}^n P(k) \delta(\hat{y}_k = y_k)}{N_{gt}} \quad (6)$$

where True Positives (TP) reflect the samples accurately predicted as lettuce by the model, False Positives (FP) reflect the samples falsely predicted as lettuce by the model, and False Negatives (FN) describe the samples wrongly judged as background. The precision reflects the percentage of objects in the returned list that are correctly detected, and the recall evaluates the percentage of correctly detected objects in total. N_{gt} is the number of object instances in the ground truth, $\delta(\cdot)$ is an indicator function equaling 1 if the predicted label of the detected object is equal to the ground truth label, and $P(K)$ is the precision measured over the top- k results.

4. Results

4.1. Instance Segmentation Model Training

The changes in the mean Average Precision (mAP) throughout the model training process are depicted in Figure 10, showing that all of the models achieved convergence. Model 5, which was trained using Dataset 3, converged within 40 epochs due to its larger collective data volume, while the other models converged within approximately 110 epochs.

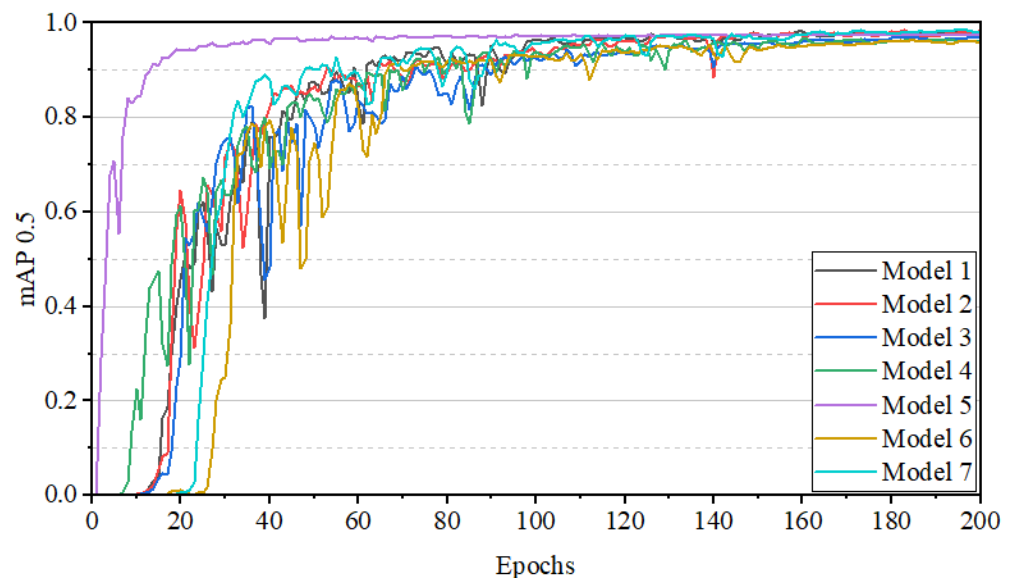


Figure 10. mAP changes of 7 models during model training.

The model training results are summarized in Table 5. Among the models trained with Dataset 1 (Model 1, Model 2, Model 6, and Model 7), Model 5, which employs FasterNet as its backbone, achieved the fastest inference speed of 0.040 s per frame. Though Model 3's mean Average Precision (mAP) is only 0.004 lower than that of Model 7, which has the highest mAP among the four models, Model 3 benefits from a 0.021 s reduction in inference time, making it approximately 34% faster than Model 7. Additionally, the mAP of Model 3 is just 0.01 lower than that of Model 2, demonstrating that reducing the model's layer channel dimensions offers promising performance improvements in both speed and accuracy.

Table 5. Model training results.

Model	P	R	mAP0.5	Inference Time (s)	Parameters (M)	GFLOPs
Model 1	0.951	0.963	0.977	0.054	3.258	11.973
Model 2	0.935	0.961	0.982	0.054	3.158	11.568
Model 3	0.967	0.922	0.981	0.040	1.369	5.960
Model 4	0.962	0.957	0.989	0.039	1.369	5.960
Model 5	0.998	<u>0.985</u>	<u>0.995</u>	<u>0.040</u>	1.369	5.960
Model 6	0.916	0.937	0.973	0.058	3.749	14.679
Model 7	0.969	0.947	0.986	0.061	2.502	10.383

The accuracy of Models 4 and 5, trained on Dataset 2 and Dataset 3, respectively, significantly increased, particularly in recall, which improved by 0.063 compared to Model 3. The precision increased by 0.031, and mAP rose by 0.013. Compared to Model 1, Model 5 exhibited a 25.56% improvement in speed, alongside a 0.024 increase in mAP. A segmentation performance comparison of the different models is illustrated in Figure 11, demonstrating that Model 5 achieved the highest average confidence for detecting the lettuce targets. The results of the model comparison confirm the superior performance of Model 5.

The heat maps of images processed by the last layer of the different backbone networks are shown in Figure 12. It shows the relative attention of the model to different positions of the image. Compared to other backbones, the model utilizing FasterNet demonstrates a greater focus on target areas, effectively enhancing the model's detection performance for the lettuce canopy.

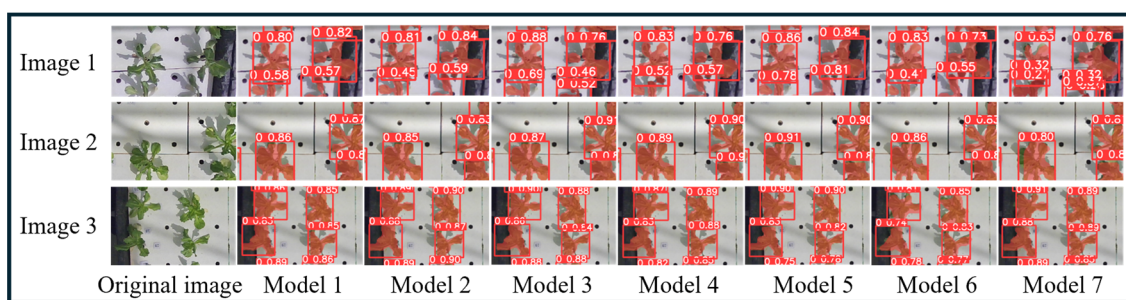


Figure 11. Segmentation performance comparison of 7 models with target confidence scores.

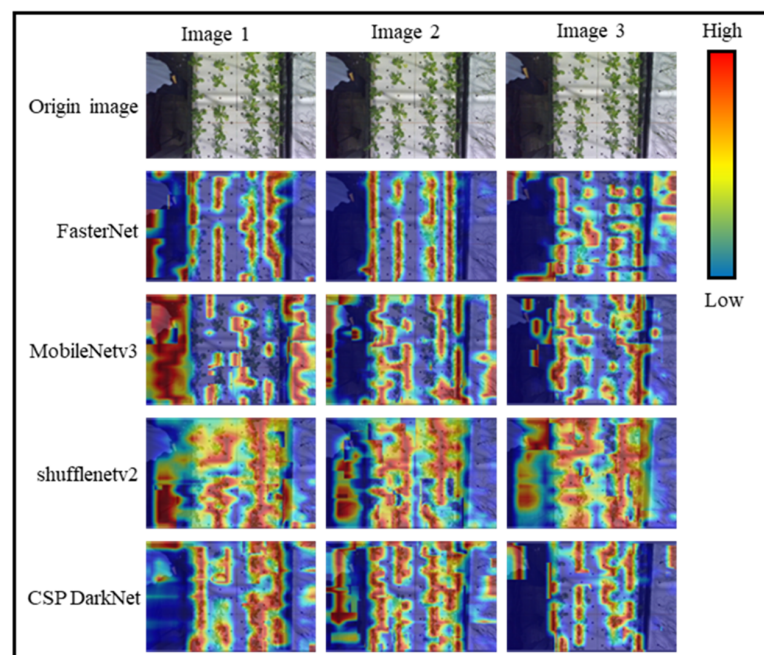


Figure 12. Heat maps of the last layer of different backbones.

4.2. Plant Height Measurement Results

Based on the segmentation results from Model 5, the accuracy of the plant height measurement algorithm was evaluated. As illustrated in Figure 13, the measurement results were printed on the image. Figure 14 shows that the measurement accuracy for 95 test samples was 93.666%, with an R^2 value of 0.882 and an 11.346 mm RMSE. This proved that the algorithm had good measurement accuracy in the plant height measurement tasks.

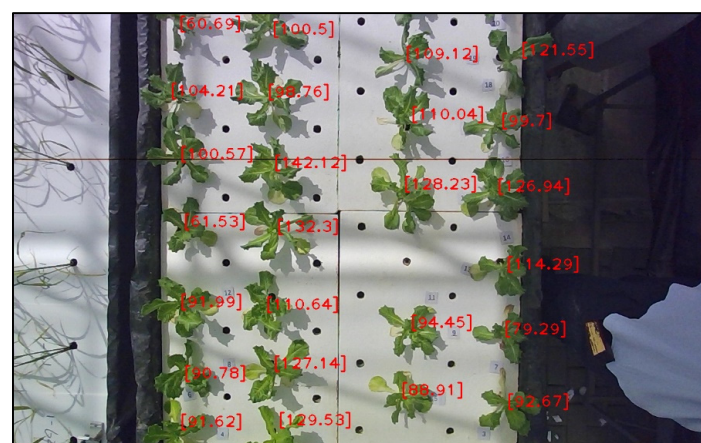


Figure 13. Lettuce height measurement outputs (mm).

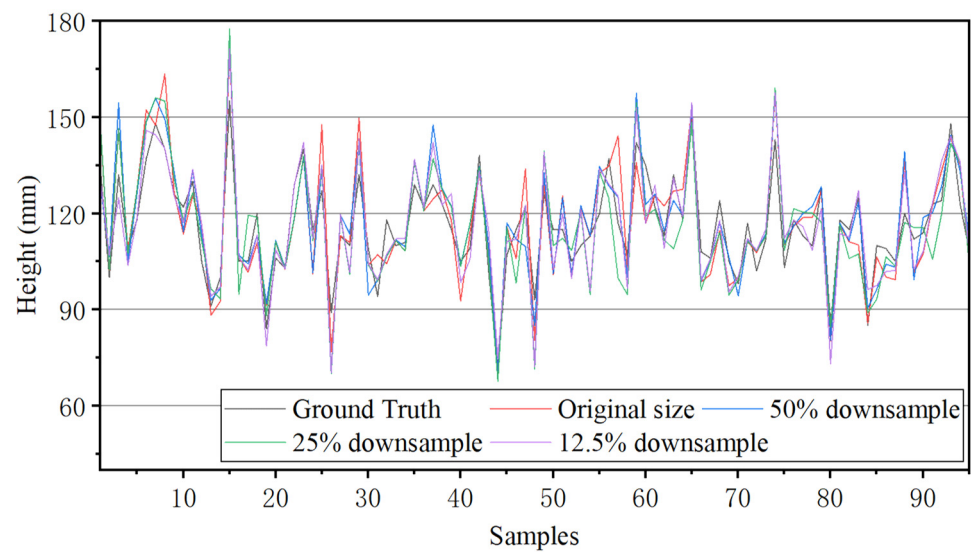


Figure 14. Plant height measurement results of hydroponics scenario.

In the potting scene, the plant height measurement algorithm achieved an accuracy of 91.22% across 81 samples, with an R^2 value of 0.895 and a 16.411 mm RSME. The mean absolute error was 9.23 mm. This proved that the surface extraction method of cultivation substrate proposed in this paper also provided good detection in multi-plane environments (Figure 15).

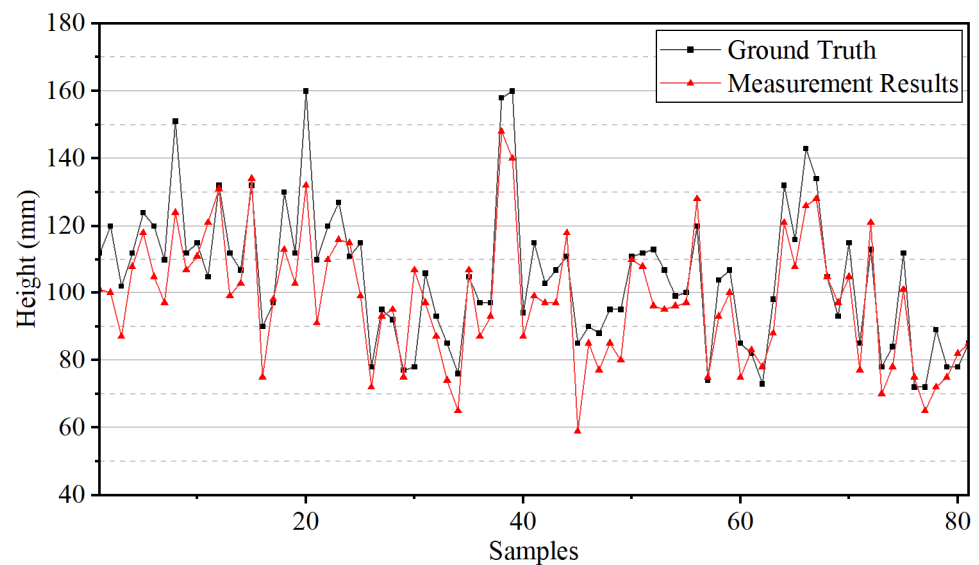


Figure 15. Plant height measurement results of potted lettuce.

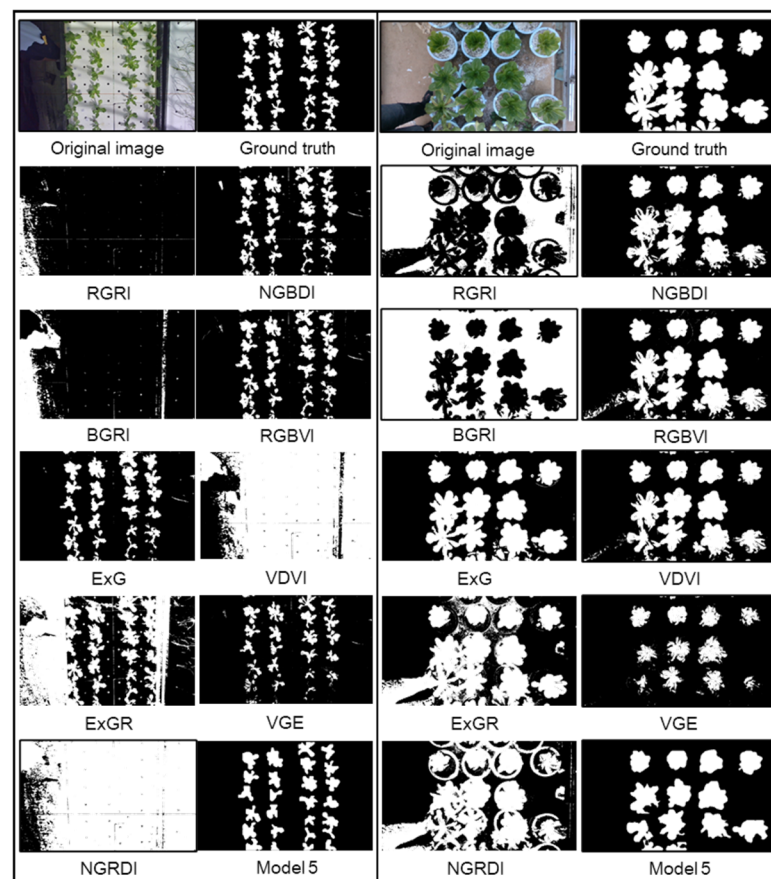
The image pixels were compressed to approximately 100%, 50%, 25%, and 12.5% of the original image area size, and the test results are presented in Table 6. The maximum fluctuation in accuracy across the different downsampling ratios was 0.875%. Notably, the measurement time for the 12.5% downsampled images was only 5% of the time required for the original size. The results indicated that the algorithm was robust across different downsampling ratios.

Table 6. Comparison of plant height measurements under different down sampling ratios.

Downsampling Ratios (Length of Side)	Pixels in Image	Instance Segmentation Time (s)	Height Measurement Time (s)	Time for All (s)	Average Accuracy	R ²
1	921,600	0.071	0.748	0.818	93.666%	0.882
0.7	451,584	0.075	0.307	0.381	94.115%	0.902
0.5	230,400	0.067	0.172	0.239	93.464%	0.876
0.354	115,491	0.070	0.073	0.143	94.339%	0.907

5. Discussion

The task of recognition and segmentation of images is a complex problem. The color-based crop segmentation method is one of the most used methods for vegetation segmentation. However, it is highly susceptible to changes in the light environment, especially the uniformity of light. Due to the structure of the facility, the shadow changes in the facility are complex, and the light environment in many cases cannot be made more uniform than the light in an open field. Therefore, the general method based on color segmentation is difficult to adapt. As shown in Figure 16, we compared the segmentation effects of instance-based segmentation methods and multiple color segmentation methods in two scenarios.

**Figure 16.** Segmentation comparison between vegetation index method and Model 5.

Nine visible light-based vegetation indices [50] used as comparison methods include the red–green ratio index (RGRI), blue–green ratio index (BGRI), excess green index (ExG), excess green minus red index (ExGR), normalized green–red difference index (NGRDI), normalized green–blue difference index (NGBDI), red–green–blue vegetation index (RGBVI), visible-band difference vegetation index (VDAI), and vegetative index (VGE). The

binary segmentation method was the Otsu algorithm, which selects the best threshold based on the maximum inter-class variance criterion [51]. It has been proven to be one of the best threshold techniques due to its excellent uniformity and shape measure [52]. It is not difficult to find that there are many noise points in the segmentation results of most vegetation indices compared with the instance segmentation method. NGBDI, RGBVI, VEG, and ExG, which have better segmentation effects in both scenarios, cannot identify weeds in the images. In addition to the choice of color index, the post-processing method after segmentation also requires elaborate design to adapt to different situations. This requires significant analysis and design work while requiring data inclusiveness, which is not simpler than the method based on deep learning. Therefore, this paper used statistical and numerical methods to achieve a relatively simple part, such as plane picking and height difference calculation. For complex recognition problems, a more adaptive deep learning method is used. The comparison results in Figure 16 prove the correctness of this design. This paper also greatly reduced the additional workload and computational requirements of deep learning methods by improving the model structure and dataset construction methods. These improvements have achieved promising results.

Compared with the measurement method that relies on the structural size of the support mechanism [53], the method designed in this paper reduces the systematic error of the measurement process. At the same time, it can also avoid the random error caused by the motion fluctuation of the support structure. As shown in the Figure 17, the plane fitting method cannot adapt to the multi-plane environment in the greenhouse, and the results based on neighborhood interception do not align well with the target plane [43,54]. In this paper, a pure numerical operation method is used to obtain the target plane in the image. The result has a higher stability than the fitting method and may also better evaluate the computing resources required by the algorithm.

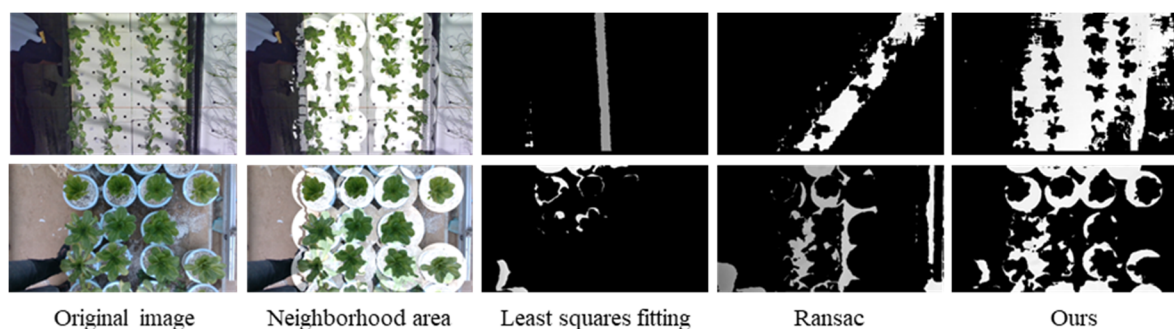


Figure 17. Comparison of different plane detection algorithms.

The proposed method depends on the visibility of the crop cultivation substrate plane within the image view, which imposes a limitation: the substrate plane must be observable and perpendicular to the z-axis of the depth image. Additionally, the method of using the modal value to evaluate depth requires that the cultivation substrate plane maintains a certain level of flatness. The algorithm employs the difference method, which directly captures sensor data after identifying the contour of the lettuce to calculate plant height. Consequently, the accuracy of the plant height data is heavily dependent on the sensor's precision. It can be seen from the performance parameters given by the manufacturer that the nominal depth error of the Orbbec Gemini 336L and Realsense D435 are 0.8% and 2%, respectively, and the distance from the camera to the crop cultivation substrate is approximately 800–1100 mm, resulting in a theoretical sensor error of about 6.4 to 16 mm from the sensor itself. The absolute mean errors of the plant height test results in the two scenarios were 7.39 mm and 9.23 mm, which were close to the sensor's error. That shows that the system error caused by the algorithm is very small and the plant height measurement error primarily stems from sensor inaccuracies.

Further research should be conducted in the following areas:

- (a) Data acquisition improvement: Utilize a mobile acquisition platform to reduce the sensor's sight distance, thereby enhancing the stability and consistency of the data.
- (b) Data optimization: Optimize the original data by increasing the calibration, filtering, and cleaning processes to improve the sensor's data accuracy.
- (c) Engineering optimization: Redesign the engineering construction and deployment processes of deep learning models to further enhance code stability and operational efficiency.

6. Conclusions

This paper proposed an algorithm for measuring the height of multi-target lettuce plants. The low-cost Orbbec Gemini 336L camera and Realsense D435 were utilized to analyze test image data collected in a controlled environment. This study demonstrates that the efficiency of a single-class instance segmentation model can be improved while maintaining accuracy by reducing the channel dimension growth rate of the model layers. Additionally, the horizontal distribution in the image can be quickly detected based on the stacking characteristics of the overhead RGB-D depth image histogram, and the method exhibits high robustness across various application scenarios.

In this paper, we first compared several improved models based on the YOLOv8n-seg instance segmentation model. Three datasets were utilized for training the models, while one test dataset was employed for comparison. Model 5, which features FasterNet as the backbone and a reconstructed channel dimensions distribution, demonstrated a 25.56% improvement in inference speed compared to the original YOLOv8n-seg model, while its accuracy increased by 4.7%.

In the second part, we designed a method to identify and segment the horizontal plane in the overhead RGB-D image to calculate the height of the lettuce bottom. This method leverages the stacking characteristics of the histogram of the horizontal plane in the RGB-D image, enabling the swift identification of multiple horizontal planes. By evaluating the distance between each lettuce center and the partitioned planes, the plane corresponding to the lettuce bottom height was identified.

Finally, using the lightweight instance segmentation model alongside the plane detection algorithm, we developed a multi-objective real-time measurement algorithm for lettuce plant heights. The segmentation model was employed to identify and segment the crop center and contour, and combined with the identification of the wave crest point in the depth image histogram, the correct plane at the base of the plant was selected from the image. The height of the lettuce plants was measured using the difference method. The accuracy for 95 lettuce plants, with an average height of 116.2 mm, was 94.339%, and the average absolute error was 7.39 mm. The processing time for a single frame image containing 25 lettuce plants was 0.818 s, which can potentially be improved to 0.143 s per frame through subsampling techniques.

In conclusion, this paper proposed a lightweight method for measuring lettuce height in a controlled environment. By utilizing a single, low-cost sensor, the method enabled real-time evaluation of lettuce growth height. This approach provides valuable support for adjusting planting decisions and ensures the quality and yield of lettuce production.

Author Contributions: Conceptualization: Y.Z.; data curation: Y.Z., J.S., T.Y. and Z.C.; formal analysis: Y.Z.; funding acquisition: H.M.; investigation: Y.Z.; methodology: Y.Z.; project administration: Y.Z., X.Z. and H.M.; resources: Y.Z.; software: Y.Z.; supervision: Y.Z., X.Z. and H.M.; validation: Y.Z.; visualization: Y.Z. and Z.Z.; writing—original draft: Y.Z. and H.M.; writing—review and editing: Y.Z., Z.Z., J.S. and H.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Key R&D Program (2022YFD2002302); a project funded by the Priority Academic Program Development of Jiangsu Higher Education Institutions (No.PAPD-2023-87); General Program of Basic Science (Natural Science) Research in Higher Education Institutions of Jiangsu Province (23KJB210004).

Institutional Review Board Statement: Not applicable.

Data Availability Statement: Data are contained within the article.

Acknowledgments: The authors express their gratitude to the School of Agricultural Engineering, Jiangsu University, for providing the essential instruments without which this work would not have been possible. The authors also thank the reviewers for their important feedback.

Conflicts of Interest: The authors declare no conflicts of interest.

References

- Petropoulou, A.S.; van Marrewijk, B.; de Zwart, F.; Elings, A.; Bijlaard, M.; van Daalen, T.; Jansen, G.; Hemming, S. Lettuce Production in Intelligent Greenhouses—3D Imaging and Computer Vision for Plant Spacing Decisions. *Sensors* **2023**, *23*, 2929. [\[CrossRef\]](#) [\[PubMed\]](#)
- Li, H.; Wang, Y.; Fan, K.; Mao, Y.; Shen, Y.; Ding, Z. Evaluation of Important Phenotypic Parameters of Tea Plantations Using Multi-Source Remote Sensing Data. *Front. Plant Sci.* **2022**, *13*, 898962. [\[CrossRef\]](#) [\[PubMed\]](#)
- Muñoz, M.; Guzmán, J.L.; Sánchez-Molina, J.A.; Rodríguez, F.; Torres, M.; Berenguel, M. A New IoT-Based Platform for Greenhouse Crop Production. *IEEE Internet Things J.* **2022**, *9*, 6325–6334. [\[CrossRef\]](#)
- Carli, D.; Brunelli, D.; Benini, L.; Ruggeri, M. An Effective Multi-Source Energy Harvester for Low Power Applications. In Proceedings of the 2011 Design, Automation & Test In Europe (Date), Dresden, Germany, 14–18 March 2011; IEEE: Grenoble, France, 2011; pp. 836–841.
- Neupane, C.; Pereira, M.; Koirala, A.; Walsh, K.B. Fruit Sizing in Orchard: A Review from Caliper to Machine Vision with Deep Learning. *Sensors* **2023**, *23*, 3868. [\[CrossRef\]](#)
- Yang, B.; Yang, S.; Wang, P.; Wang, H.; Jiang, J.; Ni, R.; Yang, C. FRPNet: An Improved Faster-ResNet with PASPP for Real-Time Semantic Segmentation in the Unstructured Field Scene. *Comput. Electron. Agric.* **2024**, *217*, 108623. [\[CrossRef\]](#)
- Rehman, T.U.; Mahmud, M.S.; Chang, Y.K.; Jin, J.; Shin, J. Current and Future Applications of Statistical Machine Learning Algorithms for Agricultural Machine Vision Systems. *Comput. Electron. Agric.* **2019**, *156*, 585–605. [\[CrossRef\]](#)
- Thakur, A.; Venu, S.; Gurusamy, M. An Extensive Review on Agricultural Robots with a Focus on Their Perception Systems. *Comput. Electron. Agric.* **2023**, *212*, 108146. [\[CrossRef\]](#)
- Gai, J.; Tang, L.; Brian, S. Plant Localization and Discrimination Using 2D+3D Computer Vision for Robotic Intra-Row Weed Control. In Proceedings of the 2016 ASABE International Meeting; American Society of Agricultural and Biological Engineers, Orlando, FL, USA, 17 July 2016.
- Wang, L.; Zheng, L.; Wang, M. 3D Point Cloud Instance Segmentation of Lettuce Based on PartNet. In Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), New Orleans, LA, USA, 19–20 June 2022; pp. 1646–1654.
- Ji, W.; Pan, Y.; Xu, B.; Wang, J. A Real-Time Apple Targets Detection Method for Picking Robot Based on ShufflenetV2-YOLOX. *Agriculture* **2022**, *12*, 856. [\[CrossRef\]](#)
- Xu, B.; Cui, X.; Ji, W.; Yuan, H.; Wang, J. Apple Grading Method Design and Implementation for Automatic Grader Based on Improved YOLOv5. *Agriculture* **2023**, *13*, 124. [\[CrossRef\]](#)
- Hu, T.; Wang, W.; Gu, J.; Xia, Z.; Zhang, J.; Wang, B. Research on Apple Object Detection and Localization Method Based on Improved YOLOX and RGB-D Images. *Agronomy* **2023**, *13*, 1816. [\[CrossRef\]](#)
- Xu, B.; Fan, J.; Chao, J.; Arsenijevic, N.; Werle, R.; Zhang, Z. Instance Segmentation Method for Weed Detection Using UAV Imagery in Soybean Fields. *Comput. Electron. Agric.* **2023**, *211*, 107994. [\[CrossRef\]](#)
- Zhang, X.; Li, F.; Zheng, H.; Mu, W. UPFormer: U-Sharped Perception Lightweight Transformer for Segmentation of Field Grape Leaf Diseases. *EXPERT Syst. Appl.* **2024**, *249*, 123546. [\[CrossRef\]](#)
- Wang, Y.; Yang, L.; Chen, H.; Hussain, A.; Ma, C.; Al-gabri, M. Mushroom-YOLO: A Deep Learning Algorithm for Mushroom Growth Recognition Based on Improved YOLOv5 in Agriculture 4.0. In Proceedings of the 2022 IEEE 20th International Conference on Industrial Informatics (INDIN), Perth, Australia, 25–28 July 2022; pp. 239–244.
- Cuong, N.H.H.; Trinh, T.H.; Meesad, P.; Nguyen, T.T. Improved YOLO Object Detection Algorithm to Detect Ripe Pineapple Phase. *J. Intell. Fuzzy Syst.* **2022**, *43*, 1365–1381. [\[CrossRef\]](#)
- Kose, E.; Cicek, A.; Aksu, S.; Tokatli, C.; Emiroglu, O. Spatio-Temporal Sediment Quality Risk Assessment by Using Ecological and Statistical Indicators: A Review of the Upper Sakarya River, Türkiye. *Bull. Environ. Contam. Toxicol.* **2023**, *111*, 38. [\[CrossRef\]](#)
- Liu, C.; Gu, B.; Sun, C.; Li, D. Effects of Aquaponic System on Fish Locomotion by Image-Based YOLO v4 Deep Learning Algorithm. *Comput. Electron. Agric.* **2022**, *194*, 106785. [\[CrossRef\]](#)
- Wang, A.; Qian, W.; Li, A.; Xu, Y.; Hu, J.; Xie, Y.; Zhang, L. NVW-YOLOv8s: An Improved YOLOv8s Network for Real-Time Detection and Segmentation of Tomato Fruits at Different Ripeness Stages. *Comput. Electron. Agric.* **2024**, *219*, 108833. [\[CrossRef\]](#)
- Wang, C.; Wang, Y.; Liu, S.; Lin, G.; He, P.; Zhang, Z.; Zhou, Y. Study on Pear Flowers Detection Performance of YOLO-PEFL Model Trained With Synthetic Target Images. *Front. Plant Sci.* **2022**, *13*, 911473. [\[CrossRef\]](#)
- Chen, C.; Zheng, Z.; Xu, T.; Guo, S.; Feng, S.; Yao, W.; Lan, Y. YOLO-Based UAV Technology: A Review of the Research and Its Applications. *Drones* **2023**, *7*, 190. [\[CrossRef\]](#)
- Dai, G.; Hu, L.; Fan, J. DA-ActNN-YOLOV5: Hybrid YOLO v5 Model with Data Augmentation and Activation of Compression Mechanism for Potato Disease Identification. *Comput. Intell. Neurosci.* **2022**, *2022*, e6114061. [\[CrossRef\]](#)

24. Bai, B.; Wang, J.; Li, J.; Yu, L.; Wen, J.; Han, Y. T-YOLO: A Lightweight and Efficient Detection Model for Nutrient Buds in Complex Tea-plantation Environments. *J. Sci. Food Agric.* **2024**, *104*, 5698–5711. [\[CrossRef\]](#)
25. Shi, H.; Shi, D.; Wang, S.; Li, W.; Wen, H.; Deng, H. Crop Plant Automatic Detecting Based on In-Field Images by Lightweight DFU-Net Model. *Comput. Electron. Agric.* **2024**, *217*, 108649. [\[CrossRef\]](#)
26. Zhang, Z.; Lu, Y.; Zhao, Y.; Pan, Q.; Jin, K.; Xu, G.; Hu, Y. TS-YOLO: An All-Day and Lightweight Tea Canopy Shoots Detection Model. *Agronomy* **2023**, *13*, 1411. [\[CrossRef\]](#)
27. Jiao, Z.; Huang, K.; Jia, G.; Lei, H.; Cai, Y.; Zhong, Z. An Effective Litchi Detection Method Based on Edge Devices in a Complex Scene. *Biosyst. Eng.* **2022**, *222*, 15–28. [\[CrossRef\]](#)
28. Zhu, H.; Lu, Z.; Zhang, C.; Yang, Y.; Zhu, G.; Zhang, Y.; Liu, H. Remote Sensing Classification of Offshore Seaweed Aquaculture Farms on Sample Dataset Amplification and Semantic Segmentation Model. *REMOTE Sens.* **2023**, *15*, 4423. [\[CrossRef\]](#)
29. Xiang, L.; Wang, D. A Review of Three-Dimensional Vision Techniques in Food and Agriculture Applications. *Smart Agric. Technol.* **2023**, *5*, 100259. [\[CrossRef\]](#)
30. Liu, Y.; Yuan, H.; Zhao, X.; Fan, C.; Cheng, M. Fast Reconstruction Method of Three-Dimension Model Based on Dual RGB-D Cameras for Peanut Plant. *Plant Methods* **2023**, *19*, 17. [\[CrossRef\]](#) [\[PubMed\]](#)
31. Stilla, U.; Xu, Y. Change Detection of Urban Objects Using 3D Point Clouds: A Review. *ISPRS J. Photogramm. Remote Sens.* **2023**, *197*, 228–255. [\[CrossRef\]](#)
32. Zhang, Y.; Wu, M.; Li, J.; Yang, S.; Zheng, L.; Liu, X.; Wang, M. Automatic Non-Destructive Multiple Lettuce Traits Prediction Based on DeepLabV3+. *J. Food Meas. Charact.* **2023**, *17*, 636–652. [\[CrossRef\]](#)
33. Ye, Z.; Tan, X.; Dai, M.; Lin, Y.; Chen, X.; Nie, P.; Ruan, Y.; Kong, D. Estimation of Rice Seedling Growth Traits with an End-to-End Multi-Objective Deep Learning Framework. *Front. Plant Sci.* **2023**, *14*, 1165552. [\[CrossRef\]](#)
34. Zhang, Q.; Zhang, X.; Wu, Y.; Li, X. TMSNet: A Three-Stage Multi-Branch Self-Correcting Trait Estimation Network for RGB and Depth Images of Lettuce. *Front. Plant Sci.* **2022**, *13*, 982562. [\[CrossRef\]](#)
35. Ma, Y.; Zhang, Y.; Jin, X.; Li, X.; Wang, H.; Qi, C. A Visual Method of Hydroponic Lettuce Height and Leaves Expansion Size Measurement for Intelligent Harvesting. *Agronomy* **2023**, *13*, 1996. [\[CrossRef\]](#)
36. Song, P.; Li, Z.; Yang, M.; Shao, Y.; Pu, Z.; Yang, W.; Zhai, R. Dynamic Detection of Three-Dimensional Crop Phenotypes Based on a Consumer-Grade RGB-D Camera. *Front. Plant Sci.* **2023**, *14*, 1097725. [\[CrossRef\]](#) [\[PubMed\]](#)
37. Grenzdörffer, G.J. Crop Height Determination with UAS Point Clouds. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2014**, *XL-1*, 135–140. [\[CrossRef\]](#)
38. Zhang, Y.; Li, M.; Li, G.; Li, J.; Zheng, L.; Zhang, M.; Wang, M. Multi-Phenotypic Parameters Extraction and Biomass Estimation for Lettuce Based on Point Clouds. *Measurement* **2022**, *204*, 112094. [\[CrossRef\]](#)
39. Hu, Y.; Wang, L.; Xiang, L.; Wu, Q.; Jiang, H. Automatic Non-Destructive Growth Measurement of Leafy Vegetables Based on Kinect. *Sensors* **2018**, *18*, 806. [\[CrossRef\]](#) [\[PubMed\]](#)
40. Malambo, L.; Popescu, S.C.; Murray, S.C.; Putman, E.; Pugh, N.A.; Horne, D.W.; Richardson, G.; Sheridan, R.; Rooney, W.L.; Avant, R.; et al. Multitemporal Field-Based Plant Height Estimation Using 3D Point Clouds Generated from Small Unmanned Aerial Systems High-Resolution Imagery. *Int. J. Appl. Earth Obs. Geoinf.* **2018**, *64*, 31–42. [\[CrossRef\]](#)
41. Hämmerle, M.; Höfle, B. Direct Derivation of Maize Plant and Crop Height from Low-Cost Time-of-Flight Camera Measurements. *Plant Methods* **2016**, *12*, 50. [\[CrossRef\]](#) [\[PubMed\]](#)
42. Song, Y.; Wang, J. Winter Wheat Canopy Height Extraction from UAV-Based Point Cloud Data with a Moving Cuboid Filter. *Remote Sens.* **2019**, *11*, 1239. [\[CrossRef\]](#)
43. Qiu, R.; Zhang, M.; He, Y. Field Estimation of Maize Plant Height at Jointing Stage Using an RGB-D Camera. *Crop J.* **2022**, *10*, 1274–1283. [\[CrossRef\]](#)
44. Xia, S.; Chen, D.; Wang, R.; Li, J.; Zhang, X. Geometric Primitives in LiDAR Point Clouds: A Review. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2020**, *13*, 685–707. [\[CrossRef\]](#)
45. Jin, Z.; Tillo, T.; Zou, W.; Zhao, Y.; Li, X. Robust Plane Detection Using Depth Information From a Consumer Depth Camera. *IEEE Trans. Circuits Syst. Video Technol.* **2019**, *29*, 447–460. [\[CrossRef\]](#)
46. Gupta, C.; Tewari, V.K.; Machavaram, R.; Shrivastava, P. An Image Processing Approach for Measurement of Chili Plant Height and Width under Field Conditions. *J. Saudi Soc. Agric. Sci.* **2022**, *21*, 171–179. [\[CrossRef\]](#)
47. Guo, X.; Guo, Q.; Feng, Z. Detecting the Vegetation Change Related to the Creep of 2018 Baige Landslide in Jinsha River, SE Tibet Using SPOT Data. *Front. Earth Sci.* **2021**, *9*. [\[CrossRef\]](#)
48. Chen, J.; Kao, S.; He, H.; Zhuo, W.; Wen, S.; Lee, C.-H.; Chan, S.-H.G. Run, Don't Walk: Chasing Higher FLOPS for Faster Neural Networks. In Proceedings of the 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Vancouver, BC, Canada, 18–22 June 2023; pp. 12021–12031.
49. Han, D.; Yun, S.; Heo, B.; Yoo, Y. Rethinking Channel Dimensions for Efficient Model Design. In Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Nashville, TN, USA, 20–25 June 2021; pp. 732–741.
50. Chen, R.; Han, L.; Zhao, Y.; Zhao, Z.; Liu, Z.; Li, R.; Xia, L.; Zhai, Y. Extraction and Monitoring of Vegetation Coverage Based on Uncrewed Aerial Vehicle Visible Image in a Post Gold Mining Area. *Front. Ecol. Evol.* **2023**, *11*, 1171358. [\[CrossRef\]](#)
51. Otsu, N. A Threshold Selection Method from Gray-Level Histograms. *IEEE Trans. Syst. Man Cybern.* **1979**, *9*, 62–66. [\[CrossRef\]](#)
52. Sezgin, M.; Sankur, B. Survey over Image Thresholding Techniques and Quantitative Performance Evaluation. *J. Electron. Imaging* **2004**, *13*, 146–165. [\[CrossRef\]](#)

53. Liu, W.; Li, Y.; Liu, J.; Jiang, J. Estimation of Plant Height and Aboveground Biomass of *Toona Sinensis* under Drought Stress Using RGB-D Imaging. *Forests* **2021**, *12*, 1747. [[CrossRef](#)]
54. Bahman, L. Height Measurement of Basil Crops for Smart Irrigation Applications in Greenhouses Using Commercial Sensors. Master's Thesis, The University of Western Ontario, London, ON, Canada, 2019.

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.