



Research article

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



B. Ambrus^a, G. Teschner^a, A.J. Kovács^a, M. Neményi^a, L. Helyes^b, Z. Pék^b, S. Takács^b, T. Alahmad^a, A. Nyéki^{a,*}

^a Széchenyi István University, Albert Kázmér Faculty of Mosonmagyaróvár, Department of Biosystems and Precision Technology, Vár 2., Mosonmagyaróvár 9200, Hungary

^b Hungarian University of Agriculture and Life Sciences, Institute of Horticultural, Páter Károly 1, Gödöllő, 2100, Hungary

ARTICLE INFO

Keywords:

Tomato yield estimation
Machine learning method
Convolution neural network
3D point cloud shaping
Approximation with sphere and 3D model
Image segmentation and calibration

ABSTRACT

The aim of this study was to estimate field-grown tomato yield (weight) and quantity of tomatoes using a self-developed robot and digital single lens reflex (DSLR) camera pictures. The authors suggest a new approach to predicting tomato yield that is based on images taken in the field, 3D scanning, and shape. Field pictures were used for tomato segmentation to determine the ripeness of the crop. A convolution neural network (CNN) model using TensorFlow library was devised for the segmentation of tomato berries along with a small robot, which had a 59.3 % F1 score. To enhance the accurate tomato crop model and to estimate the yield later, point cloud imaging was applied using a Ciclops 3D scanner. The best fitting sphere model was generated using the 3D model. The most optimal model was the 3D model, which gave the best representation and provided the weight of the tomatoes with a relative error of 21.90 % and a standard deviation of 17.9665 %. The results indicate a consistent object-based classification of the tomato crop above the plant/row level with an accuracy of 55.33 %, which is better than in-row sampling (images taken by the robot). By comparing the measured and estimated yield, the average difference for DSLR camera images was more favorable at 3.42 kg.

1. Introduction

Nowadays, there is increasing demand for new intelligent systems in precision agriculture that can enhance productivity and sustainable development, modern agrotechnological practices, and data processing methods. Yield prediction is one of the most important areas in crop production [1]. Representative data gathering, though time-consuming, is cost-effective, and replicable for larger fields. In precision agriculture, the application of on-line and on-the-go systems is essential in order to meet the measurement and data collection needs that also establish the basis for rapid intervention [2]. Therefore, scientific methods must be changed and big data's potential must be better exploited, which will require artificial intelligence (AI) [2]. The attributes deduced from the source objects are the input variables of the learning models because the source photos and videos cannot be provided directly for training procedures [3]. The ability to process in orders of magnitude more data quickly and in real time leads to a substantial increase in data

* Corresponding author.

E-mail address: nyeki.aniko@sze.hu (A. Nyéki).

volume, enhancing the economic viability and sustainability of agricultural practices. Moreover, the importance of cloud computing has to be emphasized [4].

The tomato is one of the produce items with the highest global economic values (*Lycopersicon esculentum* Mill.). Large amounts of premium raw materials must be harvested in order to meet customer demand for tomato products [5].

Various papers have been published on estimating tomato yield using artificial intelligence. The methods mainly focus on tomato size, color, shape, etc., to recognize and segregate (subdivide) the crops [3]. By altering the SegNet with VGG19, a deep-learning architecture model was produced. Recall, F1-score, and accuracy numbers for the recommended approach were 89.7 %, 72.55 %, and 80.22 %, respectively [3]. Counting and size estimation of fruits were used by Mu et al. [6] with transfer learning methods. However, this results in a longer training time; precision was 87.83 % for IoU (intersection over union), tomato count has an R2 coefficient of determination value of 0.87 and a threshold larger than 0.5. In the work of Rahnemoonfar and Sheppard [7] the testing and training accuracy was 90 % and 92 %, and a modified Inception-ResNet-A module for detection and counting was evaluated. Dynamic neural networks were used by Fitz-Rodrguez and Giacomelli [8] to create models utilizing data from greenhouse tomatoes, including fruit size, harvest rate (NN). With coefficients (R^2) higher than 70 for the corresponding harvest rate, fruit fresh weight, and developing time, the NN models accurately predicted the weekly and seasonal variability of the fruit-related characteristics during objective assessments. By combining two cutting-edge Deep Neural Network topologies used for temporal sequence processing, the Recurrent neural network and Temporal Convolutional Networks, Gong [9] et al. suggested a novel technique for predicting greenhouse tomato yield. Initially, a sequence of residual blocks in the Temporal Convolutional Networks module were treated after the representative features had been extracted during pre-processing using the Long short-term memory Recurrent neural network. A genetic algorithm based wavelet neural network was used to create a model to increase the accuracy of tomato production prediction in greenhouses [10]. For the purpose of estimating tomato yield, eight growth-related factors over a period of nine years straight were employed as input parameters. According to the experiment's findings, the backpropagation neural network model, the genetic algorithm based wavelet neural network model, and the wavelet neural network model all had mean relative errors (MREs) of 0.0067, 0.0104, and 0.0242, respectively.

A 90 % average accuracy Evolving Fuzzy Neural Network (EFuNN) was utilized to forecast greenhouse tomato output in weekly variations [11]. The characteristics of the greenhouse's internal environment, including temperature vapor pressure deficit, radiation, CO₂ and also previous yields between 2004 and 2008, made up the predictor's inputs. The neural network model was utilized for greenhouse tomato yield, which included automated monitoring station information [12]. Two UAV-derived photos were employed, according to Senthilnath [13] et al., to define and categorize the tomato fruits on the individual plants. However, tomato leaves and branches hid the fruits, therefore the plants were left out.

The object detection model was used by Liu [14] at first for detecting fruit objects and numbering images captured by cameras. Subsequently, a regression model was established for numbering tomato production on a single test plant and the real output of a single tomato. For the examination of 385 fruit photos taken at night, Xiong [15] improved an image segmentation technique; the best accuracy rate was greater than 90 %. The variations in value and location between the core pixel and adjacent pixels were used to build a pulse-coupled neural network model for segmentation.

For the 3D semantic segmentation of crops, Jayakumari [16] et al. developed a deep CNN model called CropPointNet. The point cloud was intended to be classified; however, in order to take overlap into consideration, the classification process also included the classifications of crops. The point cloud of tomato elements in the CropPointNet framework were divided with the highest rated reliability of 65 %. F1-score, recall and precision increased by 0.81, 0.85, and 0.79 points, respectively, as a result of our study. In order to construct the ML approach for identifying tomatoes, the color, shape, and size variables were extracted [17]. The training pictures' pixels were divided into four groups (fruits, leaves, stems, and background). For the purpose of identifying individual tomatoes, blob-based segmentation was employed. Throughout the testing period, the findings demonstrated accuracy and recall scores of 88 % and 80 %, respectively. Combining color and shape multi-features SVM was suggested and demonstrated to perform well in solving the issues of apple tree organs, natural apple orchard classification, and yield calculation [18].

Low-power microcontrollers can run miniature, highly optimized machine learning models like the robot described above. By analyzing sensor data immediately at the source, integrated systems may become intelligent, faster, and more energy-efficient as opposed to transmitting information to a server and awaiting an answer. Embedded machine learning or TinyML stands for this concept (URL¹).

3D scanners can be used to collect topological data about a real object or the surrounding area. A digital 3D model of the scanned components is produced after computation. Several low-cost, innovative methods have been applied recently to 3D scanning operations components [19]. The result is recorded by a single camera and the 3D information of the object is restored by software using different triangulation or projection geometries. This procedure can obtain very dense and accurate point clouds [20]. To produce 3D point clouds and R² values of more than 0.85 for transverse diameter, plant height, stem length and internode length, 3D phenotyping was suggested for tomato plants [21]. Moreover, several predicted tomato fruit weights using point cloud data were heavier than the real weights [22].

More than any other developing technology, computer vision technology will be advantageous to agriculture. In order to overcome existing issues, future agricultural production will primarily rely on computer visual intelligence based on large-scale datasets [18].

The current digital image processing techniques frequently run into issues because of lens distortion. The variation from the ideal projection taken into account by the camera model is known as distortion. This is an example of an optical aberration where the straight lines of the picture are not preserved in the electronic image [23]. Four types of lens distortion are distinguished: tangential (symmetric), radial (symmetric), asymmetric radial or tangential, and random [24]. Measurements of strain and displacement in the full-field have been successfully performed using computer vision and digital image correlation [25]. The picture's visual distortion is not distributed evenly. This demonstrates that extra movements brought on by picture distortion are visible in the detection findings and impair the actual measurements [26].

The study's objective was to calculate the precision of tomato production forecast in actual field circumstances. Various steps and data sources were involved in this research. This paper's main objectives were realized by employing a machine learning method on images sensed by a self-developed small robot and DSLR camera, along with 3D point cloud tomato shaping. The aim was to refine the best-fitting tomato crop model and estimate tomato crop production.

2. Materials and methods

2.1. Experimental site

The test site was at the Hungarian University of Agriculture and Life Sciences' horticulture experimental farm. (GPS: 47°34'51.6"N 19°22'39.0"E) in Gödöllő, Hungary (Fig. 1). The soil is characterized by loamy texture with 47.5 % silt, 41 % sand and 11.5 % clay.

The H1015 tomato hybrid (Pomodoro AGRO Kft., Mezőberény, Hungary) was used in the experiment. This is a determinate growth type hybrid with 2nd early maturity, blocky fruit shape and around 70 g fruit weight. The plants were spaced 20 cm and 140 cm apart, resulting in a plant density of 3.57 plants per square meter. The length of each row was 24 m. Transplantation of seedlings was conducted on the 10th of May. The fertilization was distributed uniformly in the treatments (excluding T0) with respect to the phenological phase of tomato plants in granulated form, meaning 159, 70, 137 kg/ha, nitrogen, phosphorus and potassium for the whole growing season, respectively. The water supply was provided with a drip irrigation system. The experiment has 5 treatments namely T0, T1, T2, T3, T4, in two repetitions. Overall, 10 rows in the field, with each row having 2 sample sites (T0-1, T0-2, T0-3, T0-4) were selected in the rows with 10 tomato plants. Altogether the data collection was exported from 20 sample sites, meaning 4 repetitions per treatment.

All treatments received 10 mm irrigation uniformly right after transplantation and once again the same amount after 6 days. In the T4 treatment, the plants received 100 % of crop evapotranspiration (ETc), and the T3 treatment was supplied with only 75 % the amount of water compared to T4 whilst the T2 received 50 % of T4. T1 was the non-irrigated control (it was irrigated only in the recovery period after transplantation, 20 mm in total). The different amounts of water were supplied with different types of drip tapes. The 10 cm emitter spacing provided 10.6 l h⁻¹ m⁻¹ (used in the T4 and T0 treatment). In the T3 treatment two lines of drip tapes were used with 20 cm emitter spacing (4 l h⁻¹ m⁻¹ × 2 = 8 l h⁻¹ m⁻¹), while the 15 cm emitter spacing provided 5.3 l h⁻¹ m⁻¹ (used in the T2 treatment) irrigation water. The positioning of the drip lines did not differ in the respective treatments; all lines were placed on the same side of each tomato row. The T0 treatment received the same amount of irrigation as T4 (100 % of crop ETc), but only the base fertilizer was applied (24, 45, 45 kg of NPK). Crop water demand was calculated by the Penman-Monteith method with the AquaCrop software [27], using meteorological data (daily minimum and maximum temperature, daily mean wind speed, global radiation, daily mean relative humidity), recorded by the meteorological station installed nearby. The ETc was calculated as the reference evapotranspiration was multiplied with the crop coefficient (Kc) [28]. The irrigation treatment started on the 14th of June and the last irrigation was carried out on the 8th of August. Precipitation was not significant in the 2022 growing season with only 119 mm in total. The cumulative water supply was 472, 472, 390, 306 and 139 mm in the T0, T1, T2, T3 and T4 treatments, respectively.



Fig. 1. The field-grown tomato experiment site.

2.2. Data sources

2.2.1. Data (picture) collection with RGB DSLR camera

The images were captured using a Canon EOS 1100D camera with Canon EF-S 18–55 mm f/3.5–5.6 lens, 4272 × 2848 pixels, which were taken above the rows of tomatoes in RGB (red, green and blue) color. The dataset of 10 RGB images per treatment was used for the data analysis. Using a tripod and an elevated position, every image was captured at a meter's height from the tomato plants.

2.2.2. Data (picture) collection and tomato fruit detection with the self-developed robot

An open-source small robot (Fig. 2) was modified as a data logger device. The equipment has a compact, modular structure, both in terms of hardware and software. In terms of hardware, it can be divided into three parts, such as the part containing the power supply, the control system, and the intervention devices. The robot itself is based on a metal frame structure. Thanks to the design of this structure, the height barrier of the robot can be easily adjusted to the specific plant culture, which helps with the positioning of the sensors and sampling equipment. Its walking structure is designed with a rubber belt for proper traction on the ground, and due to this design, the control of the robot can also be easily realized. It is powered by two direct current gear motors (DC). The heart of the system is a Raspberry Pi 4 B+ minicomputer, which is shielded by a custom PCB that also houses extra electronic parts and interfaces. A leaf sampling device on the robot's three-axis servo motor-based arm may be manipulated by a servomotor. Moreover, it has an RGB camera that has servomotors that allow it to be moved along two axes. The robot can move autonomously between the rows of the crop thanks to RPILIDAR A1 for location. Based on artificial intelligence, the robot can snap photographs and instantly recognize red tomatoes (with a neural network). The Python programming language was used to create the robot's control software. It enables total control and machine optimization for the user. Both cable (LAN) and wireless (Wifi, Bluetooth) connections can be used to control the device. The robot was employed in tomato cultivars using open-field technologies.

2.3. Calibration of the camera's pictures

Digital image correlation methods generally exhibited low detection accuracy because of lens distortion [29]. Calibration procedures were used to eliminate these distortions in order to achieve more accurate results.

2.3.1. Lens correction for elimination of lens distortion

A linear calibration was carried out by using a calibration pattern with Adobe Lens Profile Creator 1.0.4 (URL⁵) software. An accurate linear calibration was adapted using Eq. (1) as follows:

$$L \begin{bmatrix} x_i & y_i & 1 \end{bmatrix} = M \begin{bmatrix} X_V & Y_V & 1 \end{bmatrix} \quad (1)$$

where L is an scale factor, M a homography matrix, x_i and y_i coordinates [29].

2.3.2. Distance calibration to calculate tomato projection area on computer images

To execute a calibration using a reference object (distinguishing it from external adjustment), the projection area of tomatoes (objects) on the photos was established. The reference object must possess two crucial characteristics. First, there should be knowledge of the object's dimensions (width or height) in measurable units, such as millimeters or inches. Second, the reference object must be easily identifiable in the images. This can be achieved either by the object's placement (for example, it is always positioned in the top left corner of the image) or by its appearance (such as a distinctive color or shape that is unique and different from other objects visible in the image). The reference should be recognizable as a whole (URL⁶).

2.4. Tomato fruit detection with machine learning and digital image processing

Analysis of camera pictures can collect more information about cultivated crops. A shape recognition method based on machine learning was used for separating the red and green tomatoes. The automatic processing of this data was evaluated with Computer Vision based on machine learning. It was set up on a Raspberry Pi 4 microcomputer and the Tensorflow open source library developed by Google (URL¹) was chosen for the machine learning applications. Therefore, an Edge Impulse development platform was created for a shape recognition model. Edge Impulse with TensorFlow library (URL¹) was used to train, optimize, and deploy machine learning models on our self-developed robot.

The main goal is standardization, i.e., to bring different types of images into uniform format so that subsequent feature extraction is easier (Fig. 3) in the preprocessing method. A representative dataset must be collected to identify the objects, in this case the tomato

URL5

https://helpx.adobe.com/camera-raw/digital-negative.html#Adobe_Lens_Profile_Creator. Accessed: 18/09/2022.

URL6

<https://pyimagesearch.com/2016/03/28/measuring-size-of-objects-in-an-image-with-opencv/>.

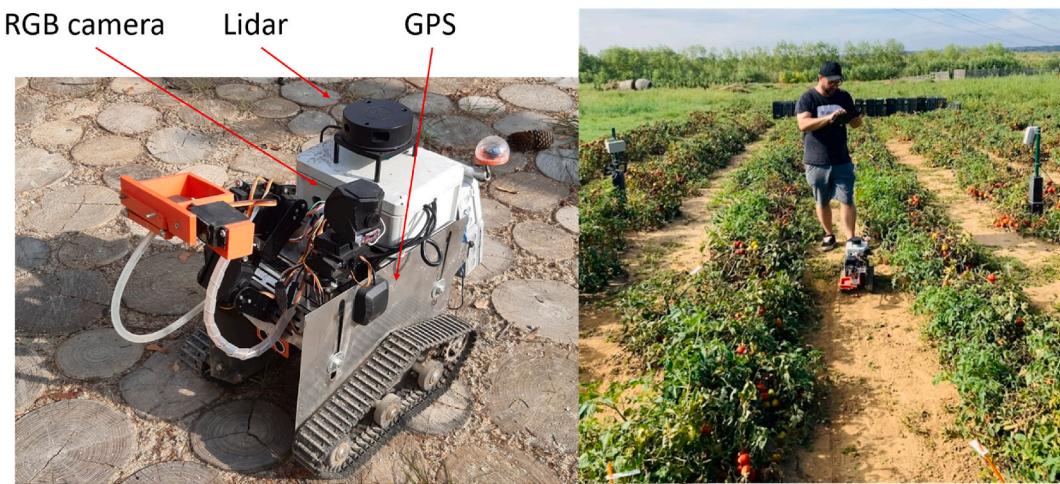


Fig. 2. The structure of the self-developed robot (left); the robot taking pictures in tomato rows.

fruits, with machine learning; a good baseline to start from is 1000 images per class. The dataset of this analysis was obtained during an earlier recording (on July 15, 2022) using 924 images per class with the small robot. The shots were cut to the size of 640×480 pixels. FOMO was employed as an object detection model based on MobileNetV2 throughout the filtering phase (URL²). After training, the model performance and confusion matrix was displayed. Edge Impulse will automatically generate the original model (float32) and tiny model (int8). 20 % of the images were utilized as the validation set. The evaluation of detection included a variety of performance indicators, including accuracy, recall, F1 score, etc. Accuracy gauges the proportion of real positives among positive observations, which in turn reveals the system's dependability. The capacity of a system to identify true positives is shown by recall.

Image classification was used during the autonomous movement of the robot in the rows. The pictures taken by the robot were classified into the predefined categories (green or red tomatoes in the image).

Low-level features such as image edges were identified for accurate analysis of images. The position, direction, and other properties of low-level elements were extracted using different algorithms. The low-level elements, in this study, were the mask layers featured in the images. To analyze the images, we applied OpenCV algorithms. As a first step, we converted the original images, consisting of three primary colors (RGB), into a form describable with HSV parameters for the precise selection of the color/colors to be segmented. From the images, we created a color-segmented mask, which was used to mark the contours of ripe tomatoes in each image. During the creation of the color segmentation mask, we took into account the HSV color values characteristic to ripe tomato berries, which were determined based on preliminary measurements. Since the contours delineated this way often did not separate well due to frequent overlap of fruits, we applied the “Watershed” algorithm for appropriate separation. We wouldn't be able to distinguish each tomato berry in the image using conventional algorithms for image processing like thresholding and contour recognition because of potential overlays. Using the resulting contour-bounded surface elements, we fitted circles around each individual tomato, using the smallest possible circle that can encompass each fruit. Great attention was paid to the precise adjustment of algorithm parameters during circle fitting, determining the minimum and maximum radius for fitting circles onto the surface elements, thereby avoiding inaccurate fittings. By utilizing the radii of the circles fitted to the tomato fruits in the image, the pixel-calculated area of each circle can be determined. Using the data of the distance between the camera and the object, along with the lens parameters, the camera-specific distance/pixel metric can be provided, allowing for the real area of each circle to be given in cm^2 dimensions, representing the projection of tomato berries onto the surface at the given distance. Utilizing this, the volume of the fruits can also be determined. For volume determination, we employed two methods: approximation with a sphere and approximation with the 3D model of the fruit.

2.5. Cloud point 3D scanning and model building for tomato fruits

In this research, a Ciclop 3D scanner was used for cloud point shaping and model building of tomato fruits. It was created by BQlabs in the Department of Innovation and Robotics, and developed in Python and released under the GPLv2 license. The scanner is equipped with an HD camera and two 650 nm red line lasers. This has a scanning region of 250 mm in diameter and 205 mm in height and can examine an object in 2–8 min. The selected item is positioned on a spinning platform. Two line lasers light the item from two opposed angles as it is spun, and a camera examines the object at each location. With the images taken, a 3D point cloud of the image is generated. (Fig. 4).

URL2

<https://www.edgeimpulse.com/blog/announcing-fomo-faster-objects-more-objects>. Accessed: 18/09/2022.

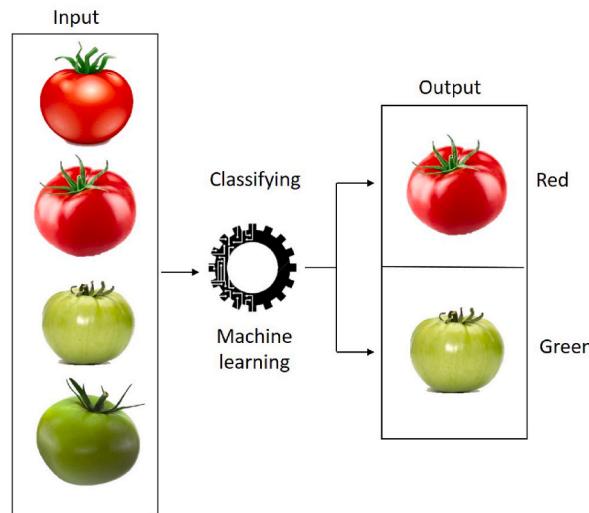


Fig. 3. Machine learning method for tomato classification (source: own).

The point clouds generated by Ciclop are processed by the Horus software (URL³). Horus is an open source solution for 3D laser scanning. It has a graphical user interface designed to connect, configure, control, calibrate and scan with the Ciclop 3D scanner, which is also open source. The point cloud created during the 3D scanning process was post-processed and analyzed with the software called MeshLab, the Visual Computing Laboratory [30]. Before the scanning procedure, the equipment must be calibrated in order to obtain the correct result. It is necessary to determine whether the pattern, motor and lasers are configured properly. In the first step of post-processing with the “Compute normal for sets” command, the program calculates the normals of the vertices of the mesh triangular relationship. The next stage is to utilize a surface reconstruction filter to combine whole meshes into one new mesh, merging the components for seamless fitment. With the help of this 3D model of tomatoes, the volume of the berries can be determined. The whole process flowchart can be seen in Fig. 5.

To determine the mass of the tomato berries, we had to establish a connection between the 3D model and the sphere formed from the mapping surface; therefore the best-fitting sphere was attached to the generated 3D surface. The sphere that fit on point clouds was calculated using Least Squares Method with Eqs. (2)–(5).

$$A = 2 \left(\frac{1}{N} \sum_{k=1}^n v_k \cdot v_k^T - \bar{v}_k \cdot \bar{v}_k^T \right) \quad (2)$$

$$b = \bar{v}_k^3 - (\bar{v}_k^2) \bar{v}_k \quad (3)$$

$$m = A^{-1} \cdot b \quad (4)$$

$$r = \sqrt{\frac{1}{N} \sum_{k=1}^n (v_k - m)^2} \quad (5)$$

where v_k are the point vectors of the point clouds, r is the radius, and m is the coordinate of the center of the mass of the sphere (URL⁴).

2.6. Regression analysis

The link between the explanatory and response variables is examined and determined using regression analysis [31]. Linear Regression (LR) is used to establish the relationship between the independent variable (irrigation amount) and as an explanatory variable, the projected yield. The regression coefficient value demonstrates unequivocally that irrigation volume has a major impact on yield [32]. In order to examine a response variable Y that varies depending on the value of the treatment variable X, a method known as

URL3

<https://horus.readthedocs.io/en/release-0.2/#>. Accessed 18/09/2022.

URL4

<https://programming-surgeon.com/en/sphere-fit-python/#sphere-fit-script>. Accessed:18/09/2022.

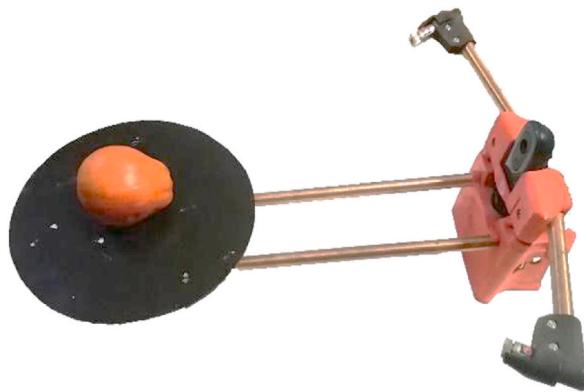


Fig. 4. Scanning a tomato with the Ciclops 3D scanner.

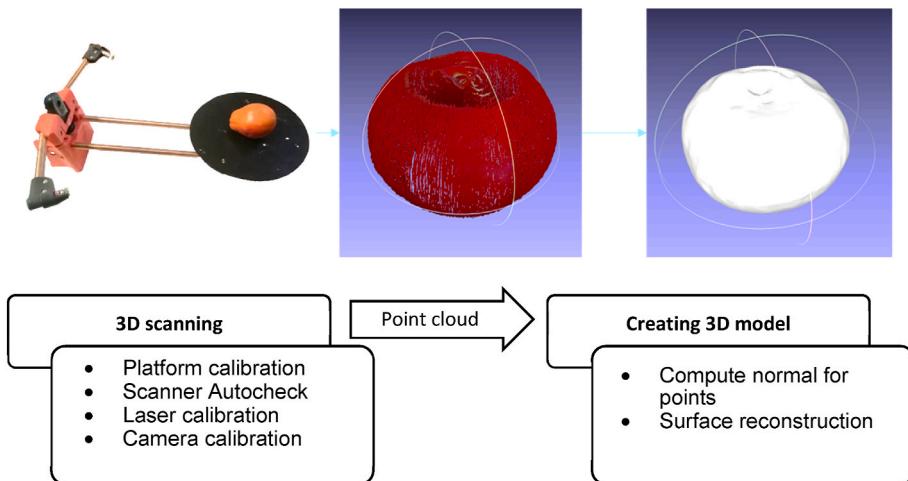


Fig. 5. Methodology flowchart of 3D scanning and model building for tomato fruit.

linear regression is presented. Prediction is also a strategy for estimating the value of a response variable from a known value of the explanatory variable. We used SPSS to calculate R^2 and adjusted R^2 , and the value of β (Unstandardized Beta and standardized Beta).

3. Results and discussion

In this work, a novel methodology was proposed for field-grown tomato yield estimation. To predict crop yield more accurately, it was taken from the field, from tomato fruits on two different dates by a self-developed robot and a DSLR camera. The suggested approach was built on the hierarchical integration of various picture segmentations centered on tomato redness. Furthermore, a 3D cloud point scanner was used for modelling the tomato fruit and estimating the best performance of volume/mass with a fit sphere model. The examined steps are shown in Fig. 6, namely: the robot and DSLR camera pictures detection with image classification and capturing (these classes represent the unique type of the object, in this case it is either a green or red tomato); the completed images are stored locally in the machine's memory; manual file transfer from robot to computer; computer images lens correction and images preprocessing and analysis; and 3D scanning.

3.1. Calibrated camera pictures

3.1.1. Lens correction for elimination of lens distortion

To eliminate the optical distortion errors a linear calibration was carried out using a calibration pattern. Information about the mapping of the optical system can be extracted from the grid mesh that is fitted to the calibration pattern by the program. During the process, the Adobe Lens Profile Creator 1.0.4 software created a calibration matrix based on images taken from different directions and distances of the calibration pattern. By utilizing Eq. (1) in this manner, a precise linear calibration model may be produced. In this way, it is possible to determine the projection surface of the tomato berries without distortion (Fig. 7).

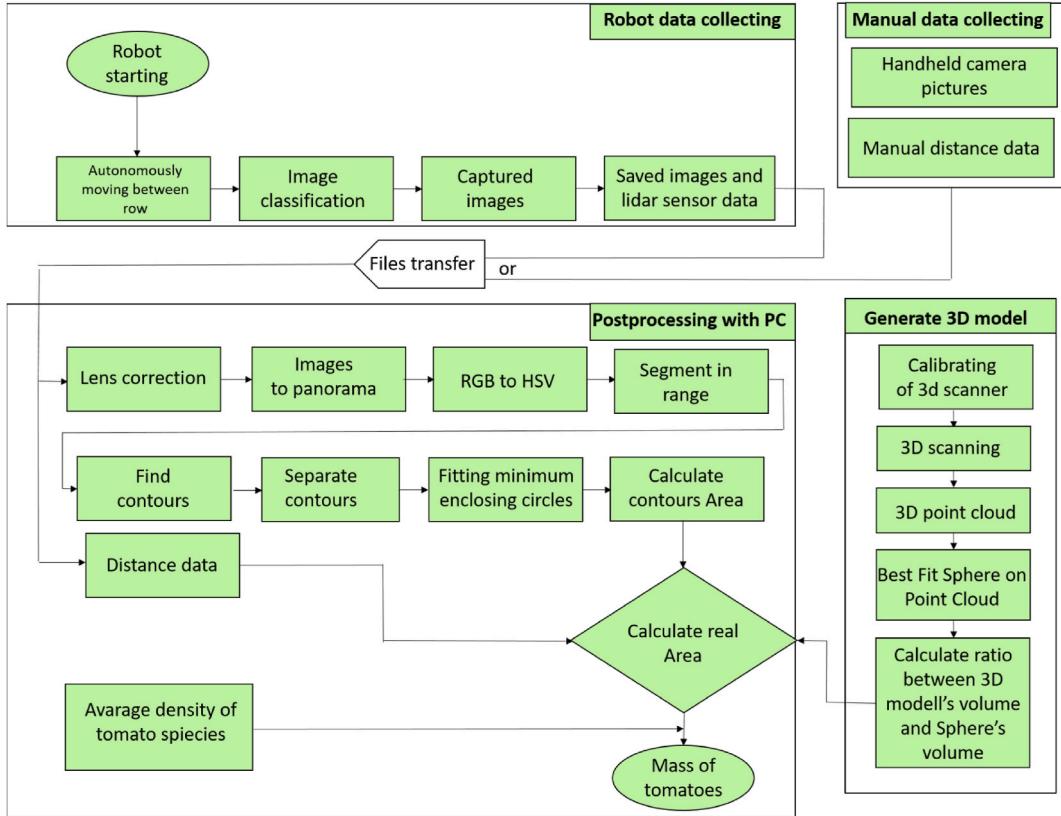


Fig. 6. Overall flow chart of the algorithm for the tomato yield estimation method.

3.1.2. Distance calibration to calculate tomatoes projection area

By knowing the shape of the images taken from different distances from 20 cm to 200 cm, the real surface of the reference object (red circle) (Fig. 8) and the surface of the mapped circle shown after color segmentation, a ratio between the surfaces can be determined using Eq. (6).

$$\varepsilon = \frac{A_{\text{projection}}}{A_{\text{real}}} \quad (6)$$

where A_{real} is the known area of the reference object (red circle), $A_{\text{projection}}$ is the projection area of the reference object on the image at a given distance. Knowing the ratio (ε) the real surface can be determined using the distance to the target with Eq. (7) from the correlation (Fig. 9).

$$\varepsilon = Ae^{bd} \quad (7)$$

where A and b are constants, d is the distance from the reference object and ε is the ratio from Eq. (6)

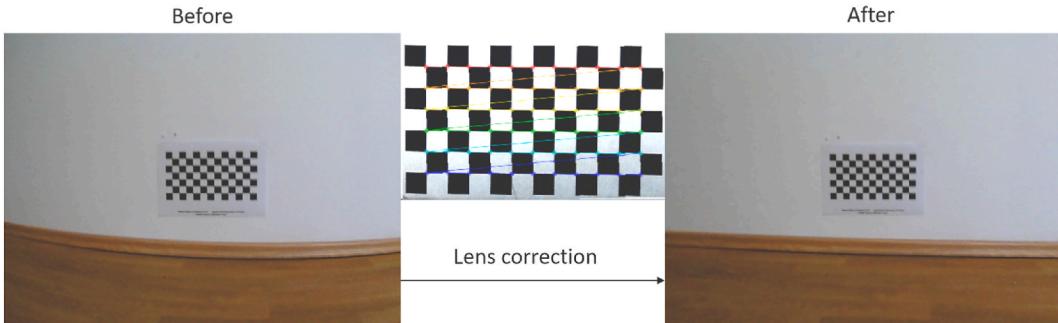


Fig. 7. Picture calibration with calibration pattern on the small robot's camera.

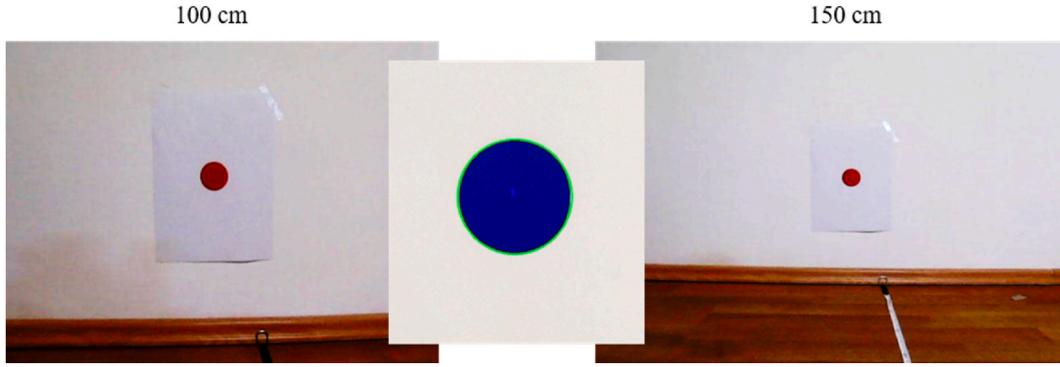


Fig. 8. Distance calibration with reference object.

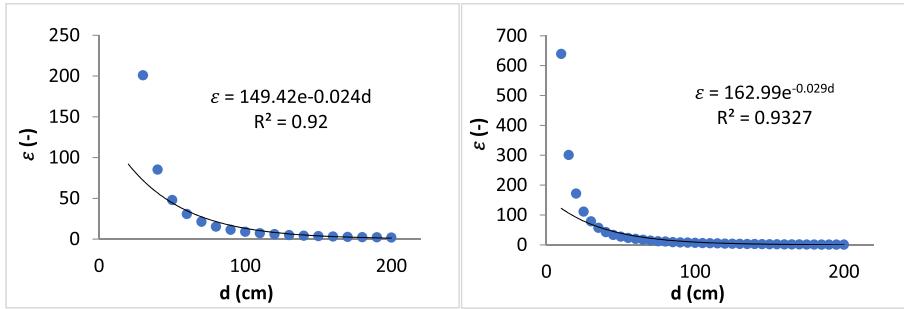


Fig. 9. DSLR and robot camera distance calibration.

3.1.3. Tomato fruit detection with machine learning

Several performance measures, including accuracy, recall, and F1-score, were computed to evaluate the detection. The last training performance for the validation set resulted in a 59.3 % F1-score. Tomato pictures on the experimental field were collected on July 15, 2022 and the number of images collected was 924. These pictures were tested and the results analyzed. The images of the test (validation) dataset provided the confusion matrix values; this matrix compares the actual target values (vertical columns) with those predicted (horizontal lines) using the machine learning model (Table 1). The low detection accuracy of green tomato is due to the same color of the background.

Evaluation occurred partially in the experimental field with the robot and partly in the processing of the photographs. The variables mentioned in the part of the study titled “Tomato fruit detection with machine learning” were used to train MobileNetV2 architecture on 80 % of the photos. The other 20 % of photos were verified after training.

The effectiveness of the suggested strategy was assessed. The metrics listed in the “Performance evaluation” section were used to assess the training and validation data. The Edge Impulse system did not work directly online with the robot, but instead ran a stand-alone application, omitting the wireless data connection used in the int8 model. Due to computing capacity, the frame rate was a maximum of 5 FPS, which proved to be sufficient with the low speed of the robot, which had a maximum of 0.2 m/s.

3.2. Cloud point 3D scanning and model building for tomato fruit

A point cloud was formed using the Ciclops 3D scanner to improve the morphological characteristics of the tomato crop and characterize the intrinsic area. The images of the multiview reconstruction were acquired manually by going around the tomato fruit (Fig. 10).

Post-processing was reached with the Meshlab software application, called Visual Computing Laboratory. To determine the mass of

Table 1
Confusion matrix (with validation set).

	Background	red tomato	green tomato
Background	100.0 %	0.0 %	0.0 %
red tomato	34.7 %	65.3 %	0 %
green tomato	83.3 %	0 %	16.7 %
F1 score	1.00	0.70	0.22

tomato fruit, a connection was established between the 3D model and the sphere formed from the mapping surface. In the first step of the processing process with the “Compute normal for sets” command, the program calculates the normals of the vertices of the mesh triangular relationship. With the help of this 3D model of tomatoes, the volume of the berries can be determined. To determine the volume of the tomato crop from the real projection surface, the relationship between the sphere and its projection was used in the simplest approximation with Eqs. 8–10.

$$V_{\text{sphere}} = \frac{4}{3}\pi r^3 \quad (8)$$

$$A_{\text{spherical projection}} = \pi r^2 \rightarrow r = \sqrt{\frac{A_{\text{spherical projection}}}{\pi}} \quad (9)$$

$$V_{\text{sphere}} = \frac{4}{3}\pi \left(\frac{A_{\text{spherical projection}}}{\pi}\right)^{\frac{3}{2}} = \frac{4A_{\text{spherical projection}}^{\frac{3}{2}}}{3\sqrt{\pi}} \quad (10)$$

In addition to approximating the tomato fruits with a sphere, a more accurate result can be achieved by using the 3D model. Therefore, the best-fitting sphere was fitted (chosen) to the generated 3D surface. The point clouds were calculated using the Least Squares Method (Fig. 11).

During this process, the center of the point cloud was determined, which will be the center of the fitted sphere, as well as the radius of the fitted sphere. Taking into account the ratio of the volume of the sphere to the volume of the tomato berries, as well as the average density of the tomato berry fruits, a more accurate estimate of the mass can be given.

3.3. Analysis and processing of the camera’s pictures with digital images processing

Low-level features, such as image edges, were essential for the accurate analysis of images. These features included the position, direction, and other properties of low-level elements, which, in this study, were the masks featured in the images. The images were segmented into color-based sections using the Python programming language’s implementation of the OpenCV package. The panoramic photos taken by the robot were then analyzed. Subsequent post-processing of the segmented images aimed to identify tomato fruits and estimate their yield. Specifically, the contours of the segmented images provided the pixels of the red tomato berries.

An algorithm, predefined to recognize a specific color range, was employed to select the red berries in the image. Another algorithm was implemented to distinguish individual berries from one another. The watershed method, a well-known technique, proved particularly useful for separating items that were touching or overlapping in the images. Following this, the program fitted circles that best matched the surface of the separated entities and calculated the area that predominated in the image. During field experiments with the robot, a radius range of 25–60 units was effective for outlining ripe tomato berries, provided the robot maintained a pre-defined distance with a maximum deviation of $\pm 3\text{--}4$ cm. However, significant deviations at certain sampling locations led to inaccurate circle fittings and, consequently, unreliable yield estimations. When capturing images manually, the fitted circle radius was effective within a range of 25–35 units.

The program-designated area, combined with the camera’s mapping, allowed for the determination of the actual surface area of the berries through calibration, which accounted for the distance between the camera and the target, measured using LiDAR. If the volumes were known, the weight of the tomato berries could be calculated using an average density to estimate the yield. More accurate weight calculations could be achieved by not approximating the berry surfaces as spheres but by using a model that also considered the berries’ morphological properties. To this end, a three-dimensional model of the berries was created using a 3D scanning process, facilitating the establishment of a relationship between their projection and volume.



Fig. 10. 3D scanning process.

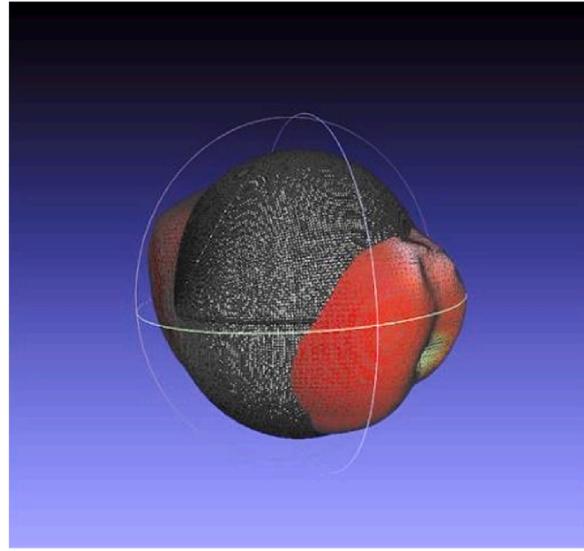


Fig. 11. Best fitting sphere of the 3D scanned tomatoes berry.

3.3.1. DSLR camera

The field-grown tomato experiment yielded a total of 100 photographs using a Canon EOS 1100D camera at harvesting time, on August 29, 2022. The image size was 4272×2848 pixels. During image data recording, the completed images were combined into a panoramic image after lens correction using Adobe Photoshop CC 2019. Fig. 12 shows the steps of the picture analysis methods, from the RGB, HSV (hue, saturation and value) segmentation to masking and contouring the images.

After the camera distance calibration and lens correction, we believe that the 3D model gives more accurate results for estimating tomato fruit number and weight.

Eqs. 11 and 12 were used, considering a linear connection between the actual and computed weights of the tomatoes.

$$m_{hs} = aM + b \quad (11)$$

$$m_{h3D} = cM + d \quad (12)$$

where m_{hs} is the calculated weight of tomatoes from DSLR camera pictures approximation with the sphere model, m_{h3D} is the calculated weight of tomatoes from DSLR camera pictures approximation with the 3D model, M is the measured weight of tomatoes, and a , b , c and d are constants.

Table 2 represents the correlation between the weight of tomato fruits using DSLR camera images and the measured weight

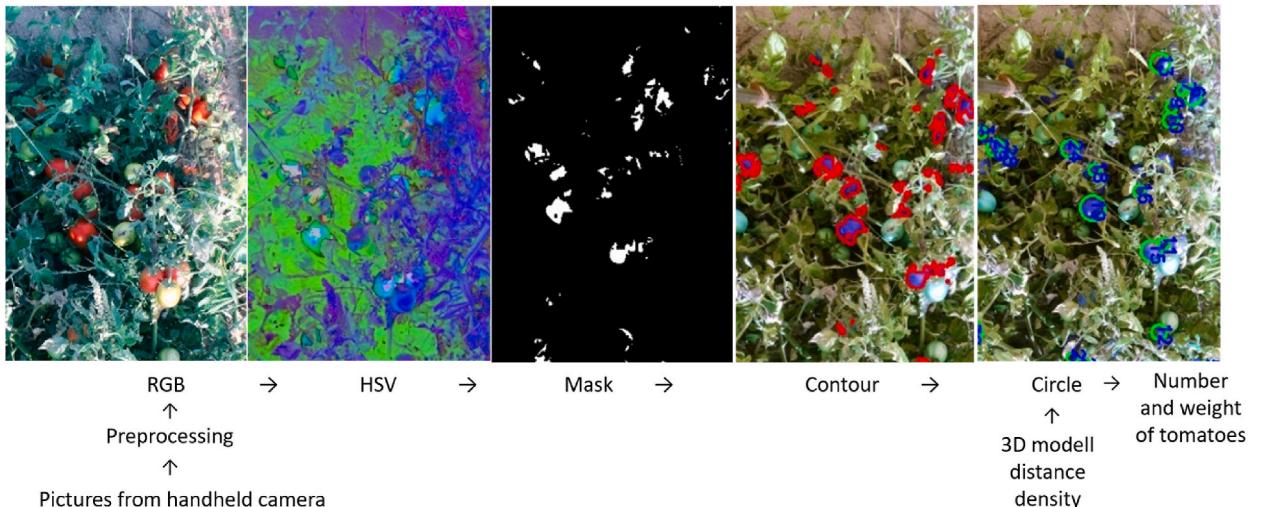


Fig. 12. General flowchart for the suggested analysis's algorithm (detail of full panoramic image).

approximation with the sphere and 3D models. The approximation using the 3D model shows a higher accuracy in every irrigation treatment (R^2 0.72–0.98).

This was a negative correlation of the higher irrigation amount treatment (T4), as the yield estimate from DSLR camera pictures was underestimated. The delineation values were because tomato branches and leaves eclipsed the fruits related to Senthilnath [13] et al.

Figs. 13 and 14 show the correlation between the measured and calculated tomato weights for all treatments using DSLR camera pictures approximation with the sphere and 3D models. The R^2 was 0.867, which indicates a clear connection and proves the effectiveness of the method. In the case of both approximations, the outlier of the T4 treatment is noticeable.

Fig. 15 summarizes the measured and calculated weights based on the treatment units, in order of increasing yields. In the T4 treatment, which received the most irrigation, the biggest difference is due to the obscuration caused by the larger amount of foliage.

$$n_h = eN + f \quad (13)$$

where n_h is the calculated number of tomatoes from DSLR camera pictures, N is the calculated number of tomatoes, and e and f are constants.

Table 3 show the correlation between the calculated number of tomato fruits from DSLR camera images and the counted number. The R^2 value was between 0.88 and 0.9965 and the relative error was 28.57 %.

Fig. 16 represent the correlation between the counted and calculated tomato numbers for all treatments from DSLR camera pictures. The R^2 was 0.744.

Fig. 17 summarizes the counted and calculated number of tomato berries based on the treatment units, depending on the increase in tomato quantities. The biggest difference can be observed in the number of tomatoes counted in the T4 treatment that received the most irrigation.

3.3.2. Analysis of the pictures by the small robot

The robot saved the number, position, and detection accuracy of red tomatoes in tabular form. In addition, when a red tomato was detected, it also recorded a picture with a 640×480 pixel resolution for later processing. During harvest time on August 29, 2022, 453 photographs were taken from the field-grown tomato experiment utilizing the self-developed robot. The recorded images were dependent on the treatments in the tomato experimental field and on the number of detectably red tomatoes, so they differed at each sampling location. During image data recording, the completed images were combined into a panoramic image after lens correction using Adobe Photoshop CC 2019. The tomato fruit region was converted from RGB to the HSV color model. Then it used a mask and contour model, which was followed by circle approximations (Fig. 12).

Eqs. 14 and 15 were used in a linear relationship between the measured and computed weights of tomatoes.

$$m_{rs} = gM + h \quad (14)$$

$$m_{r3D} = iM + j \quad (15)$$

where m_{rs} is the calculated weight of tomatoes using the small robot camera pictures approximation with the sphere model, m_{r3D} is the calculated weight of tomatoes using the small robot camera pictures approximation with the 3D model, M is the measured weight of tomatoes, and g , h , i and j are constants.

Results show that the approximation with the 3D model and the sphere model have no differences in correlation (R^2 0.77–0.95) relating to the irrigation treatments (Table 4). In this case, the calculated weight of tomato also resulted in negative correlation with measured yield in the T4 treatment.

Figs. 18 and 19 represent the correlation between the measured and calculated tomato weights for all treatments using the small robot camera images approximation with the sphere and 3D models. The R^2 was 0.907, which is a value that is higher than the images of the DSRL camera. In this case too, the outlier of the T4 treatment can be noticed for both approximations.

Fig. 20 also summarizes the measured and calculated weights based on the treatment units in order of increasing yields. In the case

Table 2

Correlation of measured and calculated tomato weight approximation with the sphere and 3D models for each treatment using DSLR camera images.

TREATMENTS	EQUATION CONSTANTS (APPROXIMATION WITH SPHERE)		EQUATION CONSTANTS (APPROXIMATION WITH 3D MODEL)		R^2	SIGN
	a	b	c	d		
T0	0.6004	3.9148	0.4962	3.2354	0.9139	▲
T1	1.965	-1.3574	1.6239	-1.1219	0.9676	■
T2	0.5407	2.1386	0.4469	1.7675	0.7346	◆
T3	0.7523	1.7713	0.6217	1.4639	0.9863	●
T4	-1.7826x	38.452	-1.4732x	31.778	0.7289	×

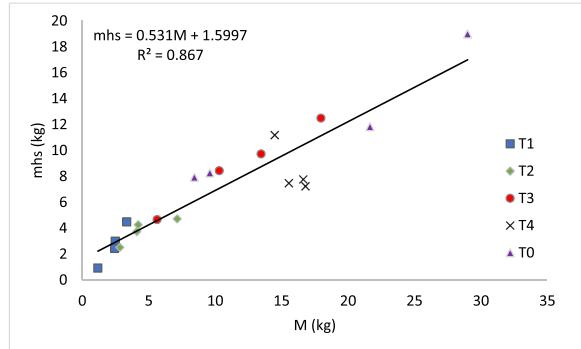


Fig. 13. Correlation of measured and calculated tomato weight approximation with the sphere model for all treatments from DSLR camera images (Appendix 2).

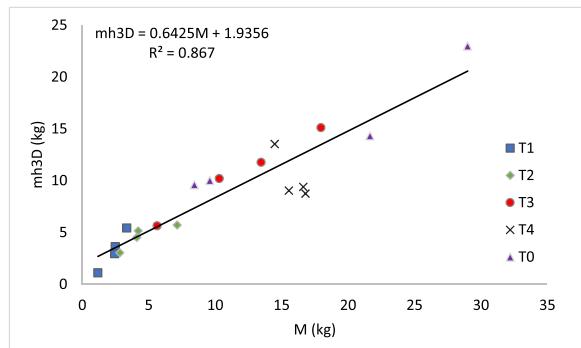


Fig. 14. Correlation of measured and calculated tomato weight approximation with the 3D model for all treatments from DSLR camera images (Appendix 2).

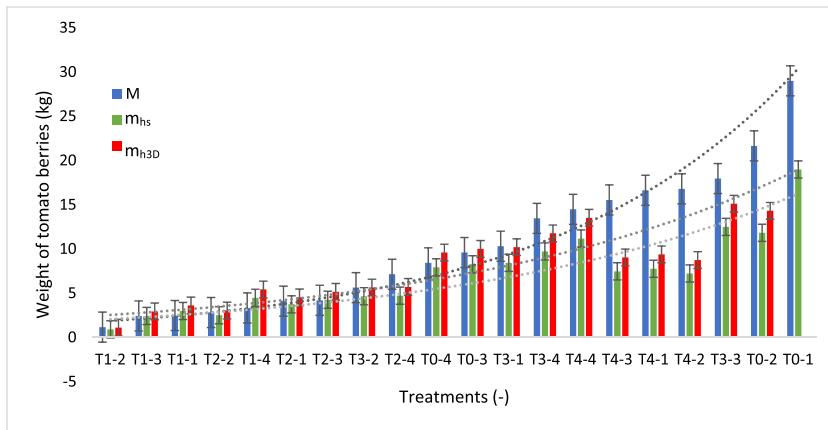


Fig. 15. Comparison of tomato berry weight based on growing yield of each treatment using different methods from DSLR camera images (Appendix 2). Assuming a linear relationship between the counted and calculated number of tomatoes, Eq (13) was applied.

of the T4 treatment, which received the most irrigation, the difference is also the largest, as in the case of DSLR images. The high value of the relative error of 76.07525 % was caused by the image recording method, according to which only one side of the plant stock was recorded, so no information is available about the other details (tomatoes).

Based on a linear relationship between the counted and calculated number of tomatoes, we used Eq. (16) as follows:

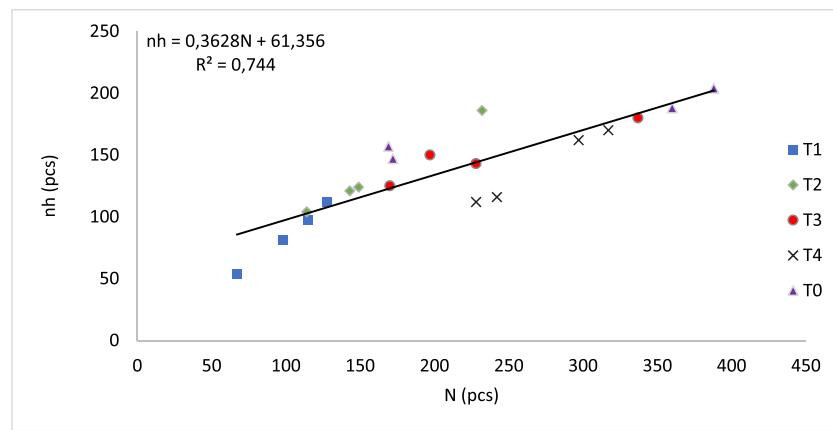
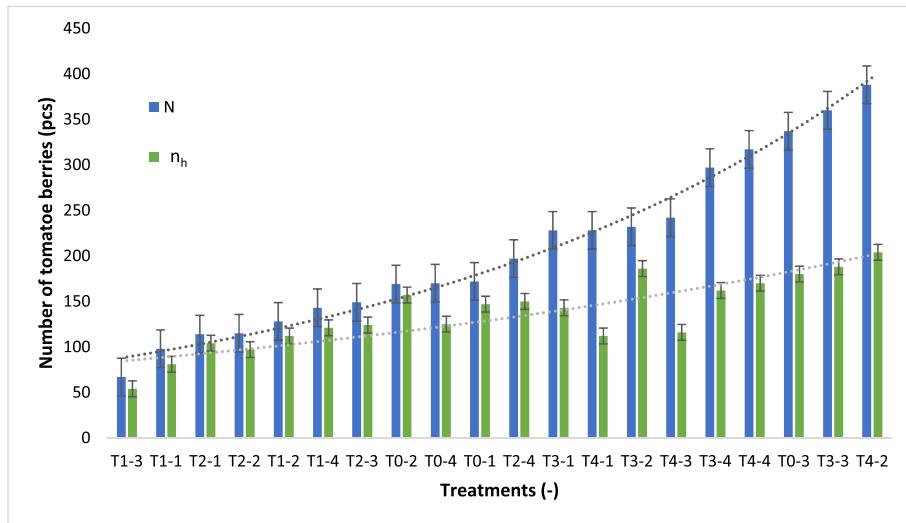
$$n_r = kN + l \quad (16)$$

where n_r is the calculated number of tomatoes from the small robot camera pictures, N is measured weight of tomatoes, and k and l are

Table 3

Correlation of measured and calculated number of tomato berries from DSLR camera images.

TREATMENTS	EQUATION CONSTANTS		R^2	SIGN
	e	f		
T0	0.2192	114.33	0.9500	▲
T1	0.9391	-9.79	0.9966	■
T2	0.7075	20.896	0.9965	◆
T3	0.2937	81.073	0.8831	●
T4	0.7027	-50.422	0.9864	×

**Fig. 16.** Correlation of counted and calculated number of tomato berries from DSLR camera images (Appendix 3).**Fig. 17.** Comparison of number of tomato berries based on growing yield of each treatment using different methods from DSLR camera images (Appendix 3).

constants.

Table 5 shows the correlation between the calculated number of tomato fruits from the small robot camera images and the counted number. The R^2 value was between 0.618 and 0.870, and 79.15 % of relative error was caused by the image capture method.

Table 4

Correlation of measured and calculated tomato weight approximation with the sphere and 3D models for each treatment from the small robot camera images.

TREATMENTS	EQUATION CONSTANTS (APPROXIMATION WITH SPHERE)		EQUATION CONSTANTS (APPROXIMATION WITH 3D MODEL)		R^2	SIGN
	<i>g</i>	<i>h</i>	<i>i</i>	<i>j</i>		
T0	0.2518	0.858	0.2081	0.7091	0.8553	◆
T1	0.2622	-0.1765	0.2167	-0.1459	0.8793	■
T2	0.4107	-0.9199	0.3394	-0.7602	0.8938	×
T3	0.2557	-0.296	0.2113	-0.2447	0.9582	●
T4	-0.6924	15.245	-0.5722	12.599	0.7764	▲

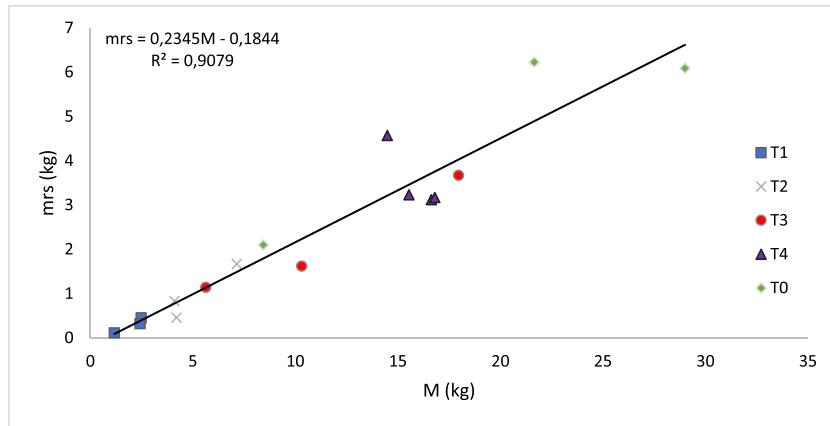


Fig. 18. Correlation of measured and calculated tomato weight approximation with the sphere model for all treatments from the small robot camera images (Appendix 4).

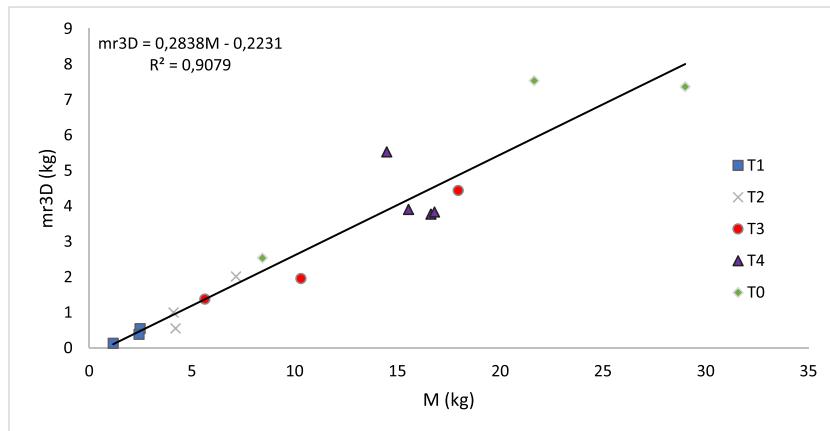


Fig. 19. Correlation of measured and calculated tomato weight approximation with the 3D model for all treatments from the small robot camera images (Appendix 4).

Fig. 21 represents the correlation between the counted and calculated tomato numbers for all treatments from the small robot camera pictures. The R^2 was 0.806.

Fig. 22 summarizes the counted and calculated number of tomato berries based on the treatment units, depending on the increase in tomato quantities. The biggest difference in proportions can be observed in the number of tomatoes counted in the T4 treatment that received the most irrigation.

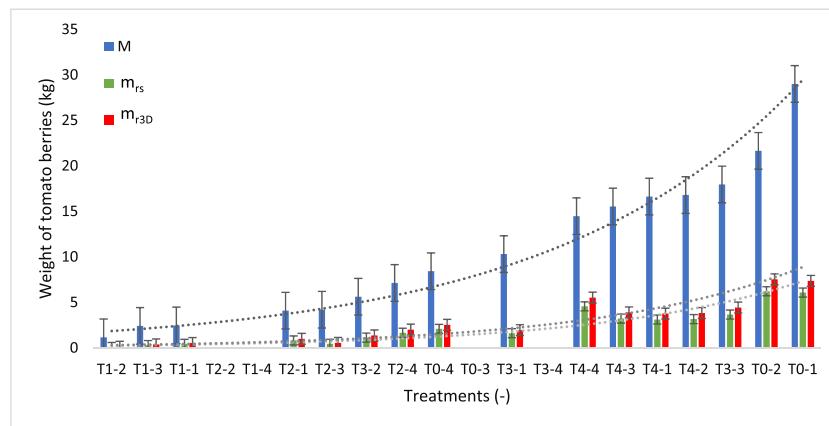


Fig. 20. Comparison of tomato berry weight based on different methods from the small robot camera images (Appendix 4).

3.3.3. Regression analysis

The findings of the regression analysis are presented in Table 6. When the computed R² value exceeds 0.5, there is a strong relationship between the response variable and the explanatory variable. The R² results unequivocally show that irrigation volume has a significant impact on tomato crop output (IrrA).

In this study, the irrigation amount has a positive value for the regression coefficient in all traits; the highest value (measured red tomato weight) was in R², and almost 58 % of variance in this trait can be explained by the difference in irrigation amount, with the B value indicating that any increase in irrigation amount by one unit corresponds to an increase of 0.045 in this trait (Table 7).

For the DSLR camera, irrigation amount has also a positive value for the regression coefficient in all traits. The highest value (measured red tomato weight and calculated tomato fruit Nr from Image) was in R². Almost 57 % of variance in this traits can be explained by the difference in irrigation amount, and the B value indicates that any increase in irrigation amount by one unit corresponds to an increase of 0.045 and 0.222 in these traits, respectively.

4. Conclusions

In this study, tomato fruit detection, maturity level, and yield estimation were examined and implemented. According to an examination of the data, red tomato berries could be distinguished with an average accuracy of 65.3 %., which is lower than the one used by Peng [33] et al., which was 99.31 %, and also lower than that of Xiong [15], which was 91.67 %. The 924 images used to train the CNN are roughly the same size as the 2000 images used in Jun's [34] study. Firstly, the tomato images were obtained using a DSLR camera and a RGB camera on the self-developed robot using a machine learning model. Next, the captured images were segmented after the development of camera lens correction and distance calibration. To enhance the accurate tomato crop model and to estimate the crop yield, a point cloud forming was used with a Ciclops 3D scanner. The best-fitting sphere model was generated using a 3D model. The feature color extraction was identified based on the accurate color separation of tomato fruits with watershed algorithms. For picture masking and contouring, the RGB and HSV feature color values were also retrieved. Finally, the tomato numbers and weight were calculated from the DSLR and robot camera's photos.

The weight and quantity of the berries were underestimated by the approximation curves used to evaluate the link between measured and predicted yield and count of tomato fruits. This was because it was difficult for the robot to take a picture behind the leaves of the tomato plants in rows. The field-grown tomato canopy structure resulted in missing points (values, fruits) in tomato detection. The R² of the calculated tomato fruit counts and weight was smaller than that of the measured tomato (Figs. 13–14, 18–19):

Table 5
Correlation of counted and calculated number of tomatoes from the small robot camera images.

TREATMENTS	EQUATION CONSTANTS		R^2	SIGN
	k	l		
T0	0.136	16.622	0.8333	◆
T1	0.277	-6.7166	0.6974	■
T2	0.2784	-7.959	0.8703	×
T3	0.0962	+27.756	0.6182	●
T4	0.2979	-28.727	0.6813	▲

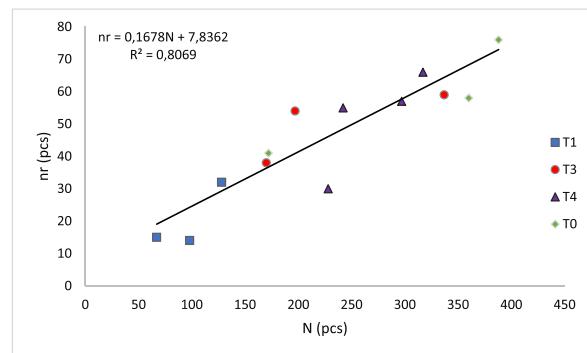


Fig. 21. Correlation of counted and calculated number of tomatoes from the small robot camera images (Appendix 5).

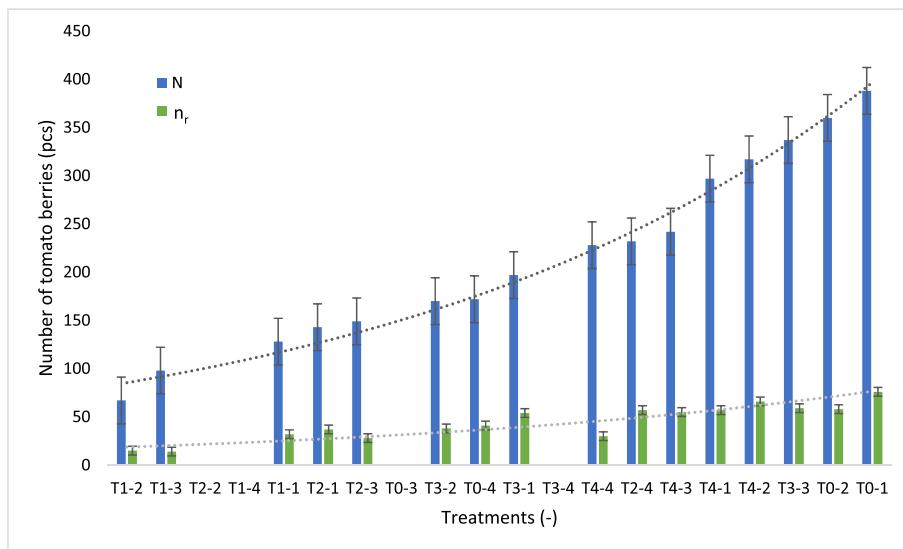


Fig. 22. Comparison of number of tomato berries based on different methods from the small robot camera images (Appendix 5).

Table 6
Linear regression between irrigation amount and all studied traits (robot).

IRRIGATION/TRAITS	R SQUARE	ADJUSTED R SQUARE	UNSTANDARDIZED COEFFICIENTS BETA	STANDARDIZED COEFFICIENTS BETA
M (kg)	0.578	0.554	0.045	0.760
M_{RS} (kg)	0.412	0.379	0.009	0.642
M_{R3D} (kg)	0.479	0.450	0.047	0.692
N (Nr)	0.534	0.509	0.529	0.731
N_R (Nr)	0.532	0.506	0.096	0.729

Table 7
Linear regression between irrigation amount and all studied traits (DSLR camera).

IRRIGATION/TRAITS	R SQUARE	ADJUSTED R SQUARE	UNSTANDARDIZED COEFFICIENTS BETA	STANDARDIZED COEFFICIENTS BETA
M (kg)	0.578	0.554	0.045	0.760
M_{HS} (kg)	0.505	0.478	0.024	0.711
M_{R3D} (kg)	0.505	0.478	0.029	0.711
N (Nr)	0.534	0.509	0.529	0.731
N_H (Nr)	0.570	0.546	0.222	0.755

its value was less than in the article by Bini [35] et al., which was 0.98). This was especially true in T4 with the highest irrigation amount because of higher leaf area (biomass). The red tomato fruit were acquired with more difficulty inside the canopy in these treatments.

Additionally, the irrigation amount has a positive regression coefficient across all traits; the highest value (measured red tomato weight) was in R², which means that almost 58 % of variance in this trait can be explained by the difference in the irrigation amount in the case of the robot pictures. In all traits, the irrigation amount also had a positive regression coefficient for the DSLR camera. The R² value indicates that almost 57 % of the variance can be explained by the irrigation amount difference.

Besides, other research focuses mainly on greenhouse-grown tomatoes [36], because it is more important and relevant than the industrial open-field grown. Open-field research emphasizes the remote sensing approach, primarily the use of drones and vegetation indices [37,38].

In summary, the robot images taken from just one side of the tomato plant rows showed a significant difference. In contrast, the hand-held camera provided more accurate values because it captured a larger range, photographing from over the rows.

The aim should be to have a separation of overlapping tomatoes for a more accurate prediction and more accurate distance measurements. This would be possible if each image were taken at the same distance as that of the robot LIDAR; however, as it wasn't the case, this is the basis for the poor results.

Data availability statement

Data are available from the authors upon request.

CRediT authorship contribution statement

B. Ambrus: Writing – original draft, Software. **G. Teschner:** Software. **A.J. Kovács:** Investigation, Formal analysis. **M. Neményi:** Supervision, Project administration. **L. Helyes:** Supervision, Conceptualization. **Z. Pék:** Data curation. **S. Takács:** Validation, Data curation. **T. Alahmad:** Software. **A. Nyéki:** Writing – original draft, Data curation, Conceptualization.

Declaration of competing interest

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

Acknowledgements

The research was carried out by the “Precision Bioengineering Research Group”, supported by the “Széchenyi István University Foundation”.

Appendices.

Appendix. 1. – The Python program of the image processing

Appendix 2Measured and calculated tomatoes weights from DSLR camera images

TREATMENTS	ITERATION	WATER SUPPLIES (MM)	MEASURED WEIGHT (KG)	CALCULATED WEIGHT (APPROXIMATION WITH SPHERE)	ABSOLUTE ERROR (APPROXIMATION WITH SPHERE) (KG)	RELATIVE ERROR (APPROXIMATION WITH SPHERE) (%)	CALCULATED WEIGHT (APPROXIMATION WITH 3D MODEL) (KG)	ABSOLUTE ERROR (APPROXIMATION WITH 3D MODEL) (KG)	RELATIVE ERROR (APPROXIMATION WITH 3D MODEL) (%)
T1	T1-1	109.8	2.48	2.99	0.51	20.56	3.6179	1.1379	45.88
	T1-2	109.8	1.17	0.91	0.26	22.22	1.1011	0.0689	5.89
	T1-3	109.8	2.43	2.43	0	0.00	2.9403	0.5103	21.00
	T1-4	109.8	3.34	4.48	1.14	34.13	5.4208	2.0808	62.30
T2	T2-1	276.1	4.1	3.74	0.36	8.78	4.5254	0.4254	10.38
	T2-2	276.1	2.82	2.51	0.31	10.99	3.0371	0.2171	7.70
	T2-3	276.1	4.2	4.26	0.06	1.43	5.1546	0.9546	22.73
	T2-4	276.1	7.14	4.72	2.42	33.89	5.7112	1.4288	20.01
T3	T3-1	360.8	10.31	8.43	1.88	18.23	10.2003	0.1097	1.06
	T3-2	360.8	5.63	4.66	0.97	17.23	5.6386	0.0086	0.15
	T3-3	360.8	17.96	12.49	5.47	30.46	15.1129	2.8471	15.85
	T3-4	360.8	13.46	9.72	3.74	27.79	11.7612	1.6988	12.62
T4	T4-1	442.4	16.63	7.76	8.87	53.34	9.3896	7.2404	43.54

(continued on next page)

(continued)

TREATMENTS	ITERATION	WATER SUPPLIES (MM)	MEASURED WEIGHT (KG)	CALCULATED WEIGHT (APPROXIMATION WITH SPHERE)	ABSOLUTE ERROR (APPROXIMATION WITH SPHERE)	RELATIVE ERROR (%)	CALCULATED WEIGHT (APPROXIMATION WITH 3D MODEL)	ABSOLUTE ERROR (APPROXIMATION WITH 3D MODEL)	RELATIVE ERROR (%)
T0	T4-2	442.4	16.8	7.23	9.57	56.96	8.7483	8.0517	47.93
	T4-3	442.4	15.54	7.47	8.07	51.93	9.0387	6.5013	41.84
	T4-4	442.4	14.48	11.18	3.3	22.79	13.5278	0.9522	6.58
	T0-1	442.4	29	18.99	10.01	34.52	22.9779	6.0221	20.77
	T0-2	442.4	21.65	11.83	9.82	45.36	14.3143	7.3357	33.88
	T0-3	442.4	9.6	8.27	1.33	13.85	10.0067	0.4067	4.24
	T0-4	442.4	8.43	7.93	0.5	5.93	9.5953	1.1653	13.82
	AVERAGE	326.3	10.3585	7.1	3.4295	25.52	8.591	2.45817	21.9085
STANDARD DEVIATION		127.7670	7.5971	4.3325	3.7402	17.03	5.2423	2.82144889	17.9665

Appendix 3

Measured and calculated number of tomatoes from DSLR camera images

TREATMENTS	ITERATION	WATER SUPPLIES (MM)	MEASURED NUMBER (PCS)	CALCULATED NUMBER (PCS)	ABSOLUTE ERROR (PCS)	RELATIVE ERROR (%)
T1	T1-1	109.8	128	112	16.00	12.50
	T1-2	109.8	67	54	13.00	19.40
	T1-3	109.8	98	81	17.00	17.35
	T1-4	109.8	115	97	18.00	15.65
T2	T2-1	276.1	143	121	22.00	15.38
	T2-2	276.1	114	104	10.00	8.77
	T2-3	276.1	149	124	25.00	16.78
	T2-4	276.1	232	186	46.00	19.83
T3	T3-1	360.8	197	150	47.00	23.86
	T3-2	360.8	170	125	45.00	26.47
	T3-3	360.8	337	180	157.00	46.59
	T3-4	360.8	228	143	85.00	37.28
T4	T4-1	442.4	297	162	135.00	45.45
	T4-2	442.4	317	170	147.00	46.37
	T4-3	442.4	242	116	126.00	52.07
	T4-4	442.4	228	112	116.00	50.88
T0	T0-1	442.4	388	204	184.00	47.42
	T0-2	442.4	360	188	172.00	47.78
	T0-3	442.4	169	157	12.00	7.10
	T0-4	442.4	172	147	25.00	14.53
AVERAGE		326.3	207.55	136.65	70.90	28.57
STANDARD DEVIATION		127.7	92.4602	38.88616991	62.11610521	16.02

Appendix 4

Measured and calculated tomatoes weights from the robot camera images

TREATMENTS	ITERATION	WATER SUPPLIES (MM)	MEASURED WEIGHT (KG)	CALCULATED WEIGHT (APPROXIMATION WITH SPHERE)	ABSOLUTE ERROR (APPROXIMATION WITH SPHERE)	RELATIVE ERROR (%)	CALCULATED WEIGHT (APPROXIMATION WITH 3D MODEL)	ABSOLUTE ERROR (APPROXIMATION WITH 3D MODEL)	RELATIVE ERROR (%)
T1	T1-1	109.8	2.48	0.45	2.03	81.85 %	0.5445	1.1379	45.88
	T1-2	109.8	1.17	0.11	1.06	90.60 %	0.1331	0.0689	5.89
	T1-3	109.8	2.43	0.32	2.11	86.83 %	0.3872	0.5103	21.00
	T1-4	109.8	3.34	NA	NA	NA	NA	NA	NA
T2	T2-1	276.1	4.1	0.83	3.27	79.76 %	1.0043	0.4254	10.38
	T2-2	276.1	2.82	NA	NA	NA	NA	NA	NA
	T2-3	276.1	4.2	0.46	3.74	89.05 %	0.5566	0.9546	22.73
	T2-4	276.1	7.14	1.67	5.47	76.61 %	2.0207	1.4288	20.01
T3	T3-1	360.8	10.31	1.62	8.69	84.29 %	1.9602	0.1097	1.06
	T3-2	360.8	5.63	1.14	4.49	79.75 %	1.3794	0.0086	0.15
	T3-3	360.8	17.96	3.67	14.29	79.57 %	4.4407	2.8471	15.85
	T3-4	360.8	13.46	NA	NA	NA	NA	NA	NA
T4	T4-1	442.4	16.63	3.12	13.51	81.24 %	3.7752	7.2404	43.54
	T4-2	442.4	16.8	3.17	13.63	81.13 %	3.8357	8.0517	47.93
	T4-3	442.4	15.54	3.23	12.31	79.21	3.9083	6.5013	41.84
	T4-4	442.4	14.48	4.57	9.91	68.44	5.5297	0.9522	6.58
T0	T0-1	442.4	29	6.09	22.91	79.00	7.3689	6.0221	20.77
	T0-2	442.4	21.65	6.23	15.42	71.22	7.5383	7.3357	33.88
	T0-3	442.4	9.6	NA	NA	NA	NA	NA	NA
	T0-4	442.4	8.43	2.1	6.33	75.09	2.541	1.1653	13.82

(continued on next page)

Appendix 4 (continued)

TREATMENTS	ITERATION	WATER SUPPLIES (MM)	MEASURED WEIGHT (KG)	CALCULATED WEIGHT (APPROXIMATION WITH SPHERE)	ABSOLUTE ERROR (APPROXIMATION WITH SPHERE)	RELATIVE ERROR (%)	CALCULATED WEIGHT (APPROXIMATION WITH 3D MODEL)	ABSOLUTE ERROR (APPROXIMATION WITH 3D MODEL)	RELATIVE ERROR (%)
AVERAGE		326.3	10.3585	2.4237	8.6981	80.2274	2.9327	8.1891375	76.07525
STANDARD DEVIATION		127.7670	7.5971	1.9802	6.1882	5.8472	2.3960	5.808465838	7.075181

Appendix 5

Measured and calculated number of tomatoes from the robot camera images

TREATMENTS	ITERATION	WATER SUPPLIES (MM)	MEASURED NUMBER (PCS)	CALCULATED NUMBER (PCS)	ABSOLUTE ERROR (PCS)	RELATIVE ERROR (%)
T1	T1-1	109.8	128	32	96.00	75.00
	T1-2	109.8	67	15	52.00	77.61
	T1-3	109.8	98	14	84.00	85.71
	T1-4	109.8	115	NA	NA	NA
T2	T2-1	276.1	143	37	106.00	74.13
	T2-2	276.1	114	NA	NA	NA
	T2-3	276.1	149	28	121.00	81.21
	T2-4	276.1	232	57	175.00	75.43
T3	T3-1	360.8	197	54	143.00	72.59
	T3-2	360.8	170	38	132.00	77.65
	T3-3	360.8	337	59	278.00	82.49
	T3-4	360.8	228	NA	NA	NA
T4	T4-1	442.4	297	57	240.00	80.81
	T4-2	442.4	317	66	251.00	79.18
	T4-3	442.4	242	55	187.00	77.27
	T4-4	442.4	228	30	198.00	86.84
T0	T0-1	442.4	388	76	312.00	80.41
	T0-2	442.4	360	58	302.00	83.89
	T0-3	442.4	169	NA	NA	NA
	T0-4	442.4	172	41	131.00	76.16
AVERAGE		326.3	207.55	44.8125	175.5	79.15
STANDARD DEVIATION		127.7	92.4602	18.0894	80.9584	4.17

References

- [1] A. Nyéki, M. Neményi, Crop yield prediction in precision agriculture, *Agronomy* 12 (2022) 2460, <https://doi.org/10.3390/agronomy12102460>.
- [2] A. Nyéki, C. Kerepesi, B. Daróczy, A. Benczúr, G. Milics, J. Nagy, et al., Application of spatio-temporal data in site-specific maize yield prediction with machine learning methods, *Precis. Agric.* 22 (2021) 1397–1415, <https://doi.org/10.1007/s11119-021-09833-8>.
- [3] P. Maheswari, P. Raja, V.T. Hoang, Intelligent yield estimation for tomato crop using SegNet with VGG19 architecture, *Sci. Rep.* 12 (2022), <https://doi.org/10.1038/s41598-022-17840-6>.
- [4] B. Ambrus, Application possibilities of robot technique in arable plant protection, *Acta Agronomica Óvariensis* 62 (1) (2021) 67–97.
- [5] E. Nemeskéri, A. Neményi, A. Bócs, Z. Pék, L. Helyes, Physiological factors and their relationship with the productivity of processing tomato under different water supplies, *Water* 11 (2019) 586, <https://doi.org/10.3390/w11030586>.
- [6] Y. Mu, T.-S. Chen, S. Ninomiya, W. Guo, Intact detection of highly occluded immature tomatoes on plants using deep learning techniques, *Sensors* 20 (2020) 2984, <https://doi.org/10.3390/s20102984>.
- [7] M. Rahnamoonfar, C. Sheppard, Deep count: fruit counting based on deep simulated learning, *Sensors* 17 (2017) 905, <https://doi.org/10.3390/s17040905>.
- [8] E. Fitz-Rodríguez, G.A. Giacomelli, Yield prediction and growth mode characterization of greenhouse tomatoes with neural networks and Fuzzy logic, *Transactions of the ASABE* 52 (2009) 2115–2128, <https://doi.org/10.13031/2013.29200>.
- [9] L. Gong, M. Yu, S. Jiang, V. Cutsuridis, S. Pearson, Deep learning based prediction on greenhouse crop yield combined TCN and RNN, *Sensors* 21 (2021) 4537, <https://doi.org/10.3390/s21134537>.
- [10] Y. Wang, R. Xiao, Y. Yin, T. Liu, Prediction of tomato yield in Chinese-style solar greenhouses based on wavelet neural networks and genetic algorithms, *Information* 12 (2021) 336, <https://doi.org/10.3390/info12080336>.
- [11] K. Qaddoum, E.L. Hines, D.D. Iliescu, Yield prediction for tomato greenhouse using EFuNN, *ISRN Artificial Intelligence* 2013 (2013) 1–9, <https://doi.org/10.1155/2013/430986>.
- [12] D.L. Ehret, B.D. Hill, T. Helmer, D.R. Edwards, Neural network modeling of greenhouse tomato yield, growth and water use from automated crop monitoring data, *Comput. Electron. Agric.* 79 (2011) 82–89, <https://doi.org/10.1016/j.compag.2011.07.013>.
- [13] J. Senthilnath, A. Dokania, M. Kandukuri, N.R. K, G. Anand, S.N. Omkar, Detection of tomatoes using spectral-spatial methods in remotely sensed RGB images captured by UAV, *Biosyst. Eng.* 146 (2016) 16–32, <https://doi.org/10.1016/j.biosystemseng.2015.12.003>.
- [14] J. Liu, Tomato yield estimation based on object detection, *J. Adv. Comput. Intell. Inf. Sci.* 22 (2018) 1120–1125, <https://doi.org/10.20965/jaci.ii.2018.p1120>.
- [15] R. Xiang, Image segmentation for whole tomato plant recognition at night, *Comput. Electron. Agric.* 154 (2018) 434–442, <https://doi.org/10.1016/j.compag.2018.09.034>.
- [16] R. Jayakumari, R.R. Nidamanuri, A.M. Ramiya, Object-level classification of vegetable crops in 3D LiDAR point cloud using deep learning convolutional neural networks, *Precis. Agric.* 22 (2021) 1617–1633, <https://doi.org/10.1007/s11119-021-09803-0>.
- [17] K. Yamamoto, W. Guo, Y. Yoshioka, S. Ninomiya, On plant detection of intact tomato fruits using image analysis and machine learning methods, *Sensors* 14 (2014) 12191–12206, <https://doi.org/10.3390/s140712191>.
- [18] L. Ge, K. Zou, H. Zhou, X. Yu, Y. Tan, C. Zhang, et al., Three dimensional apple tree organs classification and yield estimation algorithm based on multi-features fusion and support vector machine, *Information Processing in Agriculture* 9 (2022) 431–442, <https://doi.org/10.1016/j.inpa.2021.04.011>.

- [19] D. Allegra, G. Gallo, L. Inzerillo, M. Lombardo, F.L.M. Milotta, C. Santagati, F. Stanco, Low Cost DSLR 3D Scanning for Architectural Elements Acquisition. Smart Tools and Apps in Computer Graphics, Giovanni Pintore and Filippo Stanco, 2016.
- [20] M. Javaid, A. Haleem, R. Pratap Singh, R. Suman, Industrial perspectives of 3D scanning: features, roles and it's analytical applications, Sensors International 2 (2021) 100114, <https://doi.org/10.1016/j.sintl.2021.100114>.
- [21] Y. Wang, S. Hu, H. Ren, W. Yang, R. Zhai, 3DPhenoMVS: a low-cost 3D tomato phenotyping pipeline using 3D reconstruction point cloud based on multiview images, Agronomy 12 (2022) 1865, <https://doi.org/10.3390/agronomy12081865>.
- [22] Y. Ohashi, Y. Ishigami, E. Goto, Monitoring the growth and yield of fruit vegetables in a greenhouse using a three-dimensional scanner, Sensors 20 (2020) 5270, <https://doi.org/10.3390/s20185270>.
- [23] O. Stankiewicz, G. Lafruit, M. Domański, Multiview video: acquisition, processing, compression, and virtual view rendering, Academic Press Library in Signal Processing ume 6 (2018) 3–74. Elsevier.
- [24] S.G. Tzafestas, Mobile Robot Control IV. Introduction to Mobile Robot Control, Elsevier, 2014, pp. 269–317.
- [25] J. Zhao, P. Zeng, B. Pan, L. Lei, H. Du, W. He, et al., Improved Hermite finite element smoothing method for full-field strain measurement over arbitrary region of interest in digital image correlation, Opt Laser. Eng. 50 (2012) 1662–1671, <https://doi.org/10.1016/j.optlaseng.2012.04.008>.
- [26] B. Pan, L. Yu, D. Wu, High-accuracy 2D digital image correlation measurements with bilateral telecentric lenses: error analysis and experimental verification, Exp. Mech. 53 (2013) 1719–1733, <https://doi.org/10.1007/s11340-013-9774-x>.
- [27] S. Takács, E. Csengeri, Z. Pék, T. Bíró, P. Szuvandzsiev, G. Palotás, et al., Performance evaluation of AquaCrop model in processing tomato biomass, fruit yield and water stress indicator modelling, Water 13 (2021) 3587, <https://doi.org/10.3390/w13243587>.
- [28] S. Takács, Z. Pék, D. Csányi, H.G. Daood, P. Szuvandzsiev, G. Palotás, et al., Influence of water stress levels on the yield and lycopene content of tomato, Water 12 (2020) 2165, <https://doi.org/10.3390/w12082165>.
- [29] Q. Sun, Y. Hou, J. Chen, Lens distortion correction for improving measurement accuracy of digital image correlation, Optik 126 (2015) 3153–3157, <https://doi.org/10.1016/j.ijleo.2015.07.068>.
- [30] P. Cignoni, M. Callieri, M. Corsini, M. Dellepiane, F. Ganovelli, G. Ranzuglia, MeshLab: an open-source mesh processing, in: V. Scarano, R. De Chiara, U. Erra (Eds.), Tool Eurographics Italian Chapter Conference, 2008.
- [31] R. Sarathy, K. Muralidhar, Perturbation Methods for Protecting Numerical Data: Evolution and Evaluation, Elsevier, 2012, December 31.
- [32] E. López-Mata, J.M. Tarjuelo, J.A. de Juan, A. Domínguez, Effect of Irrigation Uniformity on the Profitability of Crops, Elsevier Masson, 2010, December 1.
- [33] P. Wan, A. Toudeskki, H. Tan, R. Ehsani, A methodology for fresh tomato maturity detection using computer vision, Comput. Electron. Agric. 146 (2018) 43–50, <https://doi.org/10.1016/j.compag.2018.01.011>.
- [34] J. Liu, Tomato yield estimation based on object detection, J. Adv. Comput. Intell. Inf. 22 (2018) 1120–1125, <https://doi.org/10.20965/jaciii.2018.p1120>.
- [35] D. Bini, D. Pamela, T.B. Mary, D. Shama, S. Prince, Intelligent agrobots for crop yield estimation using computer vision, Comput. Assist. Mech. Eng. Sci. 29 (1–2) (2022) 161–175.
- [36] J. Kim, H. Pyo, I. Jang, J. Kang, B. Ju, K. Ko, Tomato harvesting robotic system based on Deep-ToMaToS: deep learning network using transformation loss for 6D pose estimation of maturity classified tomatoes with side-stem, Comput. Electron. Agric. 201 (2022) 107300, <https://doi.org/10.1016/j.compag.2022.107300>.
- [37] K. Tatsumi, N. Igarashi, X. Mengxue, Prediction of plant-level tomato biomass and yield using machine learning with unmanned aerial vehicle imagery, Plant Methods 17 (2021) 77. <https://doi.org/10.1186/s13007-021-00761-2>.
- [38] M. Lillo-Saavedra, A. Espinoza-Salgado, A. García-Pedrero, C. Souto, E. Holzapfel, C. Gonzalo-Martín, et al., Early estimation of tomato yield by decision tree ensembles, Agriculture 12 (2022) 1655, <https://doi.org/10.3390/agriculture12101655>.