Volume measurement of tulip bulbs

Roel Klein (roel.klein@wur.nl), Jos Ruizendaal (jos.ruizendaal@wur.nl)

Report WPR-1312

WAGENINGEN
UNIVERSITY & RESEARCH

Referaat
Als onderdeel van het Bollen 4.0-project onderzocht deze studie de mogelijkheid om het volume van tulpenbollen te meten met RGB-D-beelden. Bollen werden gesegmenteerd met behulp van een deep learning model. Met deze segmentaties werden verschillende methodes ontwikkeld en geëvalueerd om het volume te voorspellen. Verrassend genoeg maakte de beste methode geen gebruik van de diepte-informatie (R^2 =0.65). Dit zou een indicatie kunnen zijn dat er te veel ruis is om voldoende voorspellende waarde uit de (diepte)beelden te halen, waarbij onjuiste stitching, slecht gepelde bollen en onjuiste segmentatie door het deep learning model zijn geïdentificeerd als mogelijke oorzaken.

Abstract
As part of the Bollen 4.0 project, this study researched the possibility to measure volume of tulip bulbs using RGB-D images. Bulbs were segmented using a deep learning model. Given these segmentations, several volume estimation methods were developed and evaluated. Surprisingly, the method with the highest correlation to the ground truth volume did not make use of the depth information (R^2 =0.65). This could be an indication that there is too much noise to get sufficient predictive values from the (depth)-images, of which poor stitching, badly peeled bulbs and incorrect segmentation by the deep learning model have been identified as possible causes.

Reportinfo
Report WPR-1312
Project number: 3742275802-Bollenrevolutie 4.0 Verwerking
DOI: https://doi.org/10.18174/653183

This project is part of the PPS Bollenrevolutie 4.0 and had as goal to map the quality of the bulbs in the bulb cleaning line. Bollenrevolutie 4.0 is a four-year research program of the 'topsector Tuinbouw & Uitgangsmaterialen van het ministerie van LNV'. This Public-private partnership includes seven partners; KAVB, Anthos, Wageningen University & Research (WUR), Cremer Speciaalmachines B.V., Machinefabriek Steketee B.V., Agrisim B.V., BKD en TechnNature. Economic Board Greenport Duin & Bollenstreek and Rabobank Bollenstreek contribute to the financing from their Innovation funds. A brief explanaition of this PPP can be found on the website www.vitaleteelt.nl

Disclaimer

Address
Wageningen University & Research, BU Greenhouse Horticulture
Violierenweg 1, 2665 MV Bleiswijk
P.O. Box 20, 2665 ZG Bleiswijk
The Netherlands
+31 (0) 317 - 48 56 06
glastuinbouw@wur.nl
www.wur.eu/greenhousehorticulture

Table of contents

1 Introduction

The objective of this study was to research the possibilities to measure volume of tulip bulbs by means of 3D computer vision. Therefore, a method was developed and validated. The work presented is part of the Bollen 4.0 project.

2 Material and methods

2.1 Material

The volume measurements were done for 99 peeled bulbs from one category size sorted on the peeling line. From these bulbs 69 were used for the volume analysis using camera images. Some bulbs were not clearly visible in the image, the label was not clear enough or there were some image (depth or RGB) issues that made the image data inaccurate, so they were discarded from the dataset. Combined colour and distance images (RGB-D) of the bulbs were acquired using an Intel RealSense D415 camera (https://www.intelrealsense.com/depth-camera-d415/). When capturing an object with an RGB-D camera from above, the distance to the camera for every pixel is given.

There is one measured bulb per image, which is marked with a label to its right, as shown in Figure 1. In stitching the images there sometimes was a misalignment in two consecutive image lines, which resulted in 18 tulips that were deemed unusable for the volume measurements, such as the bulb in Figure 2. The stitching issue occurred randomly which made it hard to account for during the measurements on the running conveyor belts. There are several causes for the stitching issues possible, which would need further investigation. The RealSense camera framerate (fps) seems to be unstable, this has not to do with memory sizes, but real cause is not clear yet. For stitching the camera was set at is highest resolution, to be able to do proper abnormality detection. However the stitching works perfect on lower resolution, it did not on the higher resolution. From the recorded dataset, the remaining 49 well stitched bulbs were divided in two classes: 33 were considered correctly peeled and 16 poorly peeled.

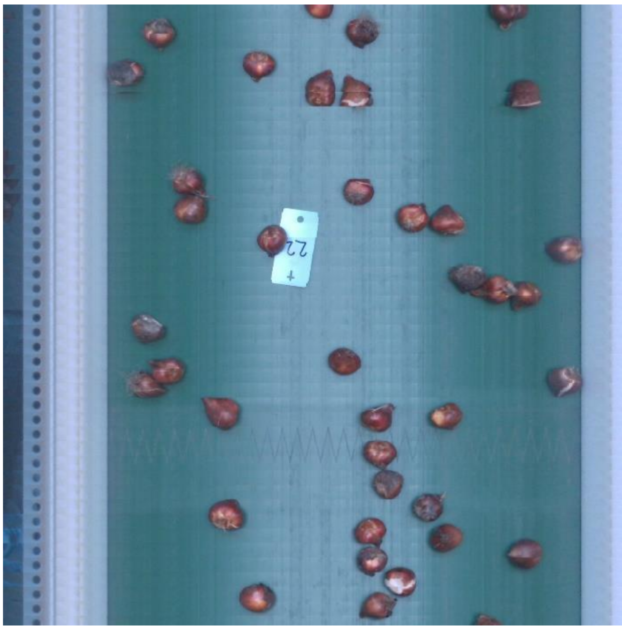


Figure 1 *Example image with bulb marked by tag. The bulb is placed to the upper-left of the tag.*

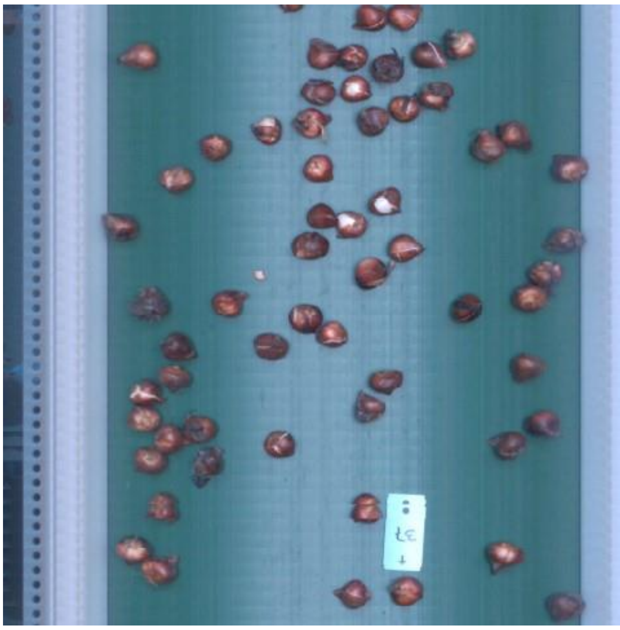


Figure 2 *Example Image with poor stitching. The volume of the bulb cannot easily be estimated, because it is not captured correctly in the image.*

2.2 Method

The manual volume measurements were done by weighing the water displacement of a bulb on a scale. A bulb was put in a small cage and submerged in a bucket of water on a scale. The weight of the water displacement is equal to the volume of the bulb, and with a density of 997 kg/m^3 this is equal to the weight of the bulb. From submerged bulbs an image was recorded with the camera box, adding a label next to the bulb as reference (see Figure 1).

For image analysis, the first step is to segment the bulbs. This is done by applying a trained deep learning model on the RGB image. This segmentation is assumed to be correct.

Next the volume is determined based on the depth map within that region. Here multiple settings have been experimented with to identify possible sources of errors. Specifically, we looked at:

1. How to deal with **zero-values** in the depth map? There are incidental zero-values in the depth map, which means that the real depth at those pixels is not known.
2. How to determine the **conveyor belt depth**? The conveyor belt depth is required to determine how high the bulbs are.
3. Which **volume calculation** to use given the conveyor belt and depth map?

Volume calculation algorithm:

1. Extract part of the depth image based on the bulb segmentation.
2. Handle **zero-values** in depth map, either:
   a. **Ignore**: remove zero-valued depth pixels.
   b. **Smooth**: replace zero-valued depth pixels with the result of a 5x5 median filter.
3. Set the **conveyor belt depth**, either:
   a. **MaxDepth**: set conveyor belt depth to the maximum depth within the bulb segmentation.
   b. **StableDepth**: Get a background segmentation by taking the inverse of all bulb segmentations in the image and applying a 3x3 erosion element. Set conveyor belt depth to the median of the background segmentation.
4. Subtract the conveyor belt depth from the depth image.
5. **Volume calculation** by either:
   a. **SumAll**: Set volume to the sum of depth image pixel values.
   b. **HalfMirror**: The bulb is round, so not everything between the top layer and the conveyor belt should be included in the volume. With this intuition, we try to find the volume of the top half and duplicate it to the bottom half. First, find the top bulb height by taking the 95 percentile. Remove the bottom half, by subtracting half of the top distance and removing negative values. Set the volume to two times the sum of remaining depth pixels.
   c. **SurfaceOnly**: Used as a check whether the depth map gives extra information. Simply take the number of pixels within the bulb segmentation, without looking at the depth values themselves.

Based on combinations of these options, six different methods were used, as seen in Table 1.

Table 1 *Different methods experimented with, and which options were used.*

	Zero-values	Conveyor belt depth	Volume Calculation
Method A	Ignore	MaxDepth	SumAll
Method B	Ignore	StableDepth	SumAll
Method C	Smooth	MaxDepth	SumAll
Method D	Smooth	StableDepth	SumAll
Method E	Ignore	StableDepth	HalfMirror
Method F	-	-	SurfaceOnly

3 Results

3.1 Quantitative analysis

The distribution of ground truth volumes, as measured by submerging the bulbs can be found in Figure 3 and Table 2.

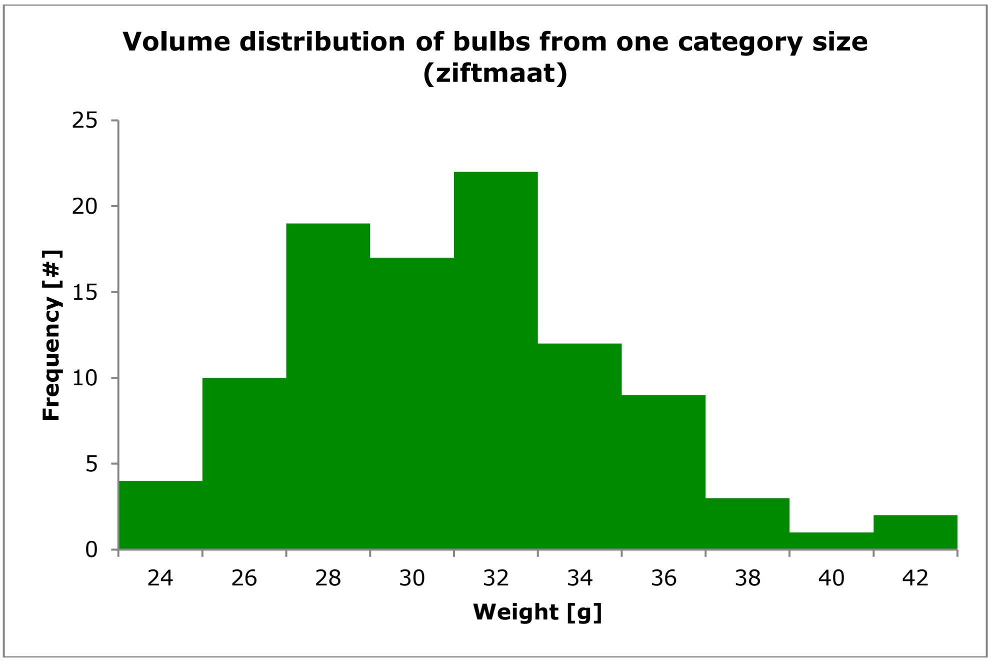


Figure 3 *Distribution of ground truth volume in cm3, measured by submerging bulbs in water.*

Table 2 *Weight distribution measured by submerging bulbs.*

Average weight [g]	30.1
Min	22.9
Max	40.5
stdev	3.7

The volume predictions from the image data do not make use of a conversion to cm^3, so at this point the units of the predicted volume are unknown. However, we only evaluate based on the correlation between predicted volumes and the ground truth volume in cm^3, which does not change under a constant scaling factor.

The metric we use is R-squared, which determines how much of the variation in the ground truth volume can be explained by the predicted volume. We report the R-squared on all of the bulbs combined, R-squared on the control bulbs, and R-squared on the badly peeled bulbs.

The results of this can be seen in Table 3. Figure 4 also shows the scatter plots and fitted lines for the different methods. Surprisingly, method F, which does not use the depth information, has the highest overall correlation. Additionally, it performs best on the control bulbs, which could be an indication that the bulb height is not as important in this dataset. If we only look at the control bulbs, method B and D also outperform A and C, indicating that for this class StableDepth works better than MaxDepth. For the badly peeled bulbs, the methods that use MaxDepth work best instead.

From the comparisons of method A to C and method B to D we can determine that the way of dealing with zero-values has little influence, resulting in no more than 0.01 difference in R-squared. This is probably due to the small number of zero-values within the bulb segmentations.

Table 3 *R-squared scores of the different volume prediction methods to the measured volume. Additionally, scores on a subset with only control bulbs and only badly peeled bulbs are reported.*

	R-squared on all bulbs, n=49	**R-squared on control bulbs, n=33**	**R-squared on badly peeled bulbs, n=16**
Method A	0.60	0.56	**0.78**
Method B	0.54	0.60	0.37
Method C	0.60	0.56	**0.77**
Method D	0.54	0.60	0.38
Method E	0.49	0.53	0.42
Method F	**0.65**	**0.68**	0.54

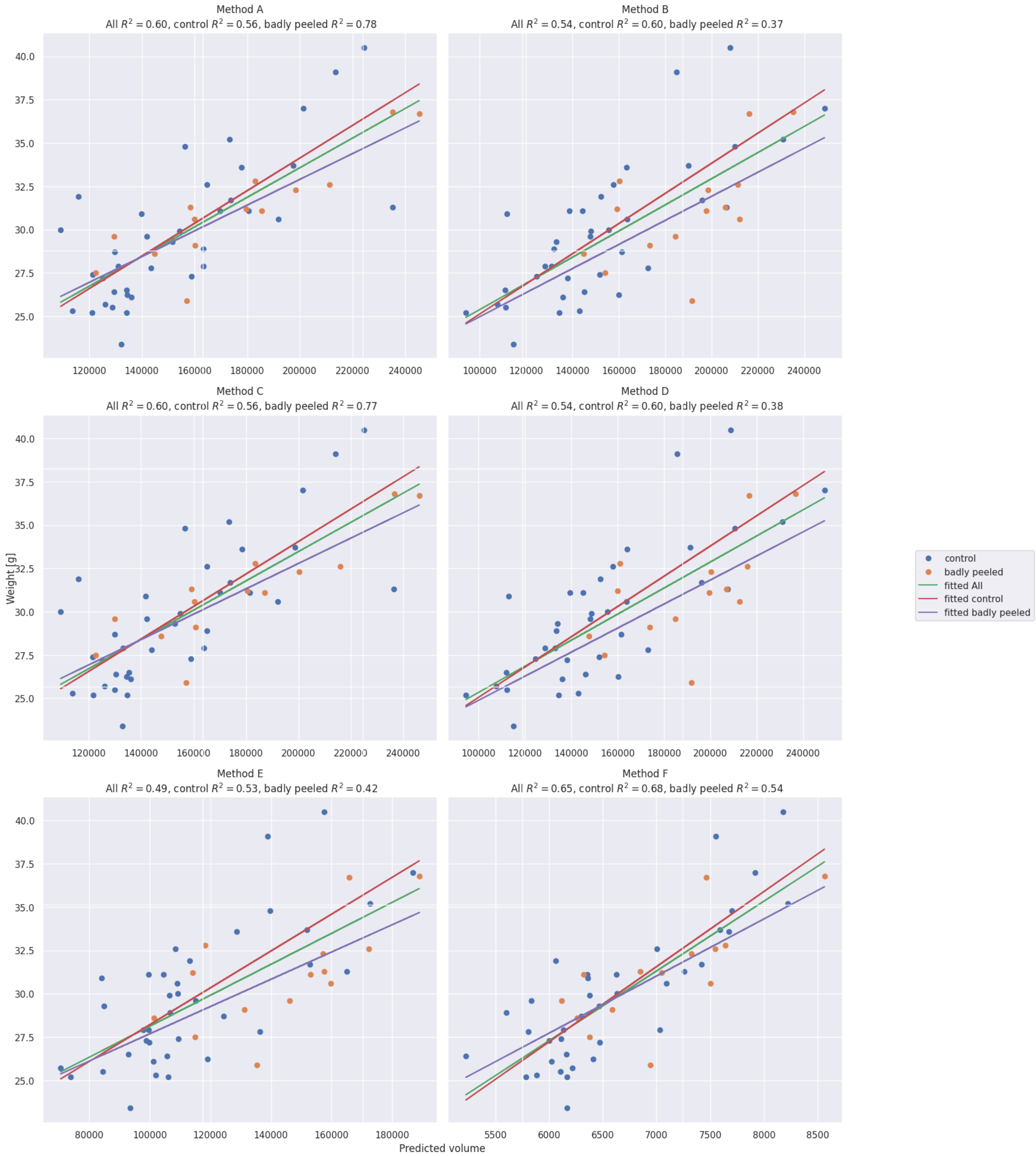


Figure 4 *Scatter plots of measured weight versus predicted volume, for each of the methods.*

3.2 Qualitative analysis

Overall, the results are not satisfactory for any of the methods. For the qualitative results, we inspect some of the highest-error bulbs of the different classes, both for the cases where the linear fit underpredicts and where it overpredicts.

Underprediction

In Figure 5 you can see the heaviest bulb, which gets underpredicted by all of the methods. One reason is that the bulb is not stitched correctly, which means that parts of the bulb are not in the image and will not contribute to the volume calculation. This can be seen by looking at the discontinuation of the white line on the bulb. Another reason is that the bulb segmentation is not completely correct: some edge pixels of the bulb are missing, which also means that these areas are not included in the volume calculations. A second example of this can be seen in Figure 6.

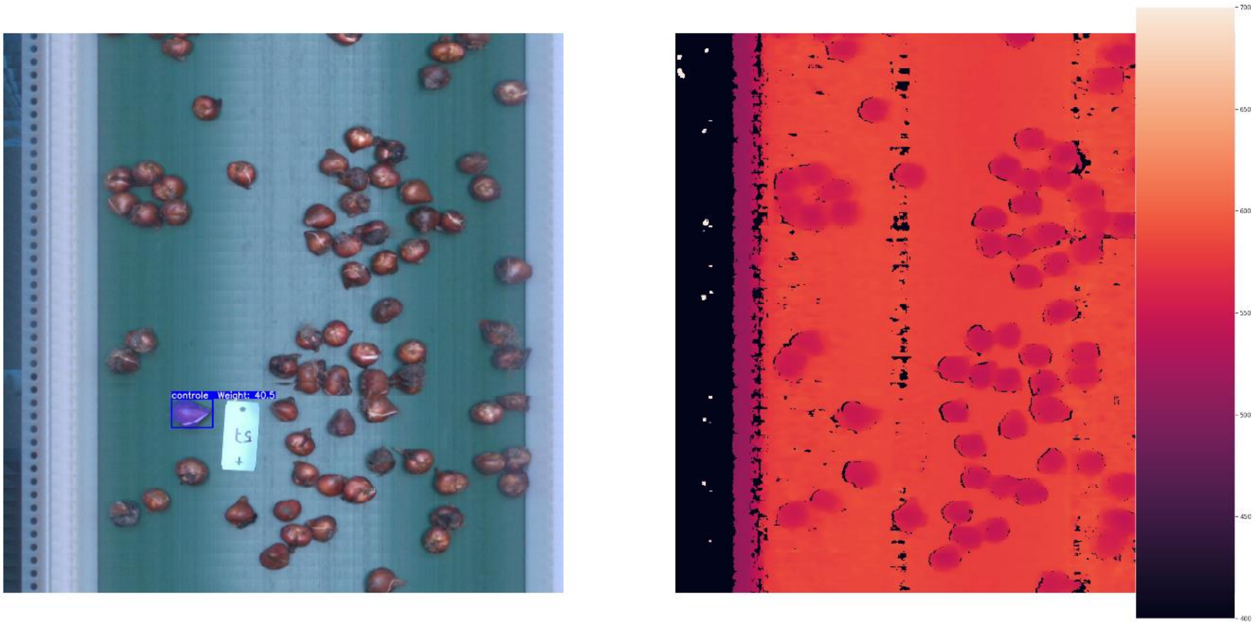


Figure 5 *Example of bulb which is heavier than our methods predict. Left: RGB image with predicted mask. Right: depth image. Measured weight: 40.5 gram.*
Method A to F errors: -4.85, -6.96, -4.92, -6.97, -7.24, -4.44.

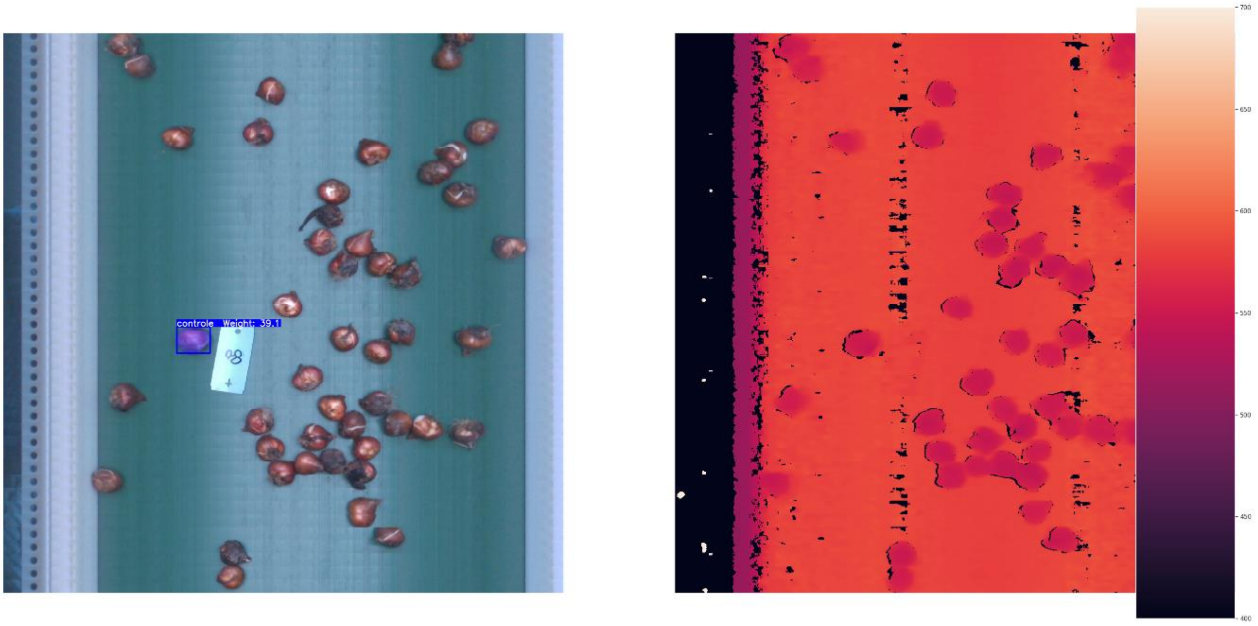


Figure 6 *Example of bulb which is heavier than our methods predict. Left: RGB image with predicted mask. Right: depth image. Measured weight: 39.1 gram.*
Method A to F errors: -4.37, -7.30, -4.43, -7.32, -7.50, -5.56.

Overprediction

An example bulb that gets overpredicted by all methods can be seen in Figure 7. This seems to be the case because the bulb is not completely peeled, and the attached skin still contributes to the volume calculations.

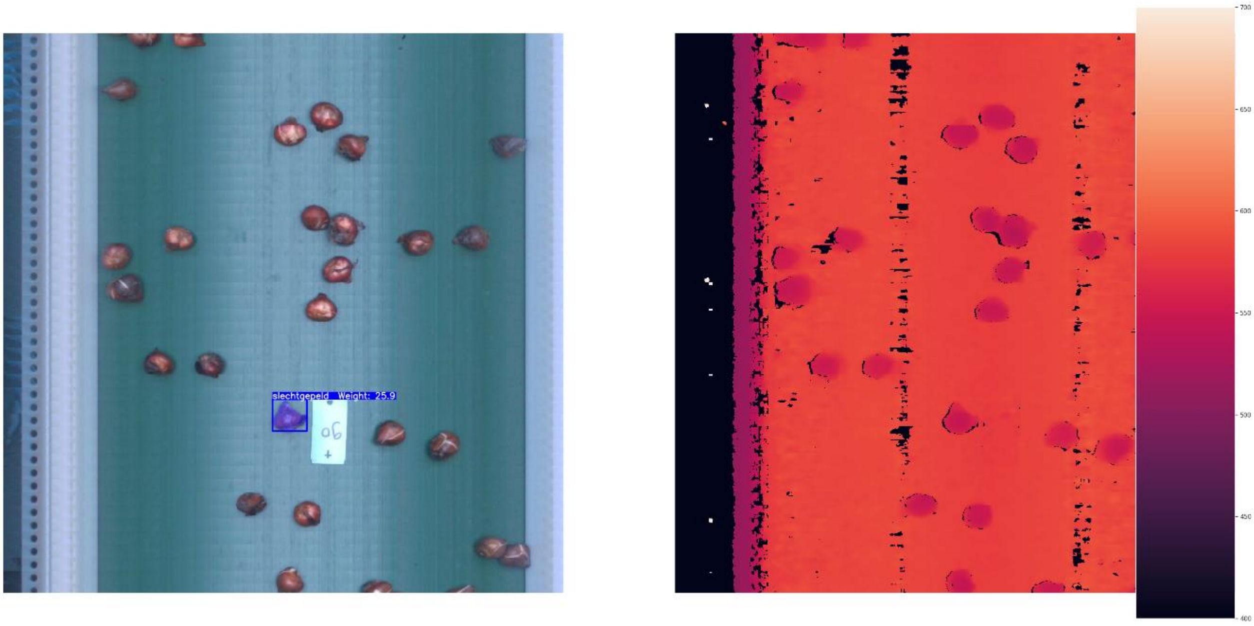


Figure 7 *Example of bulb which is heavier than our methods predict. Left: RGB image with predicted mask. Right: depth image. Measured weight: 25.9 gram.*
Method A to F errors: 3.99, 6.40, 3.95, 6.36, 5.38, 5.19.

An example of a bulb that is peeled correctly, but still get overpredicted by all methods can be found in Figure 8. No obvious mistakes can be seen in the images, which could be the cause of the overprediction. Since we fit one line through both the relatively good and poorly stitched images, the better stitched images will have too high volumes in comparison.

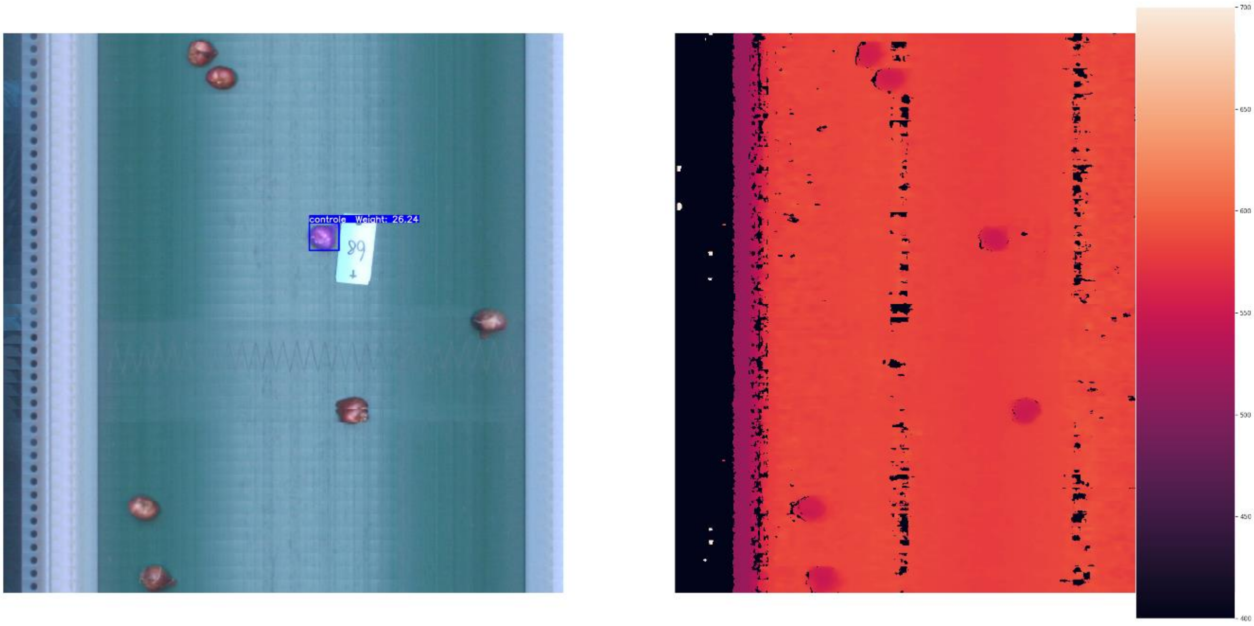


Figure 8 *Example of bulb which is heavier than our methods predict. Left: RGB image with predicted mask. Right: depth image. Measured weight: 26.24 gram.*
Method A to F errors: 1.72, 3.68, 1.69, 3.63, 3.59, 2.71.

4 Discussion

The variation in measured bulb volume of the submerged bulbs from one size category (ziftmaat) shows that although it's a robust method, there is a variation of volume up to 30% compared to the average volume. As the volume of the bulb has influence on the expected plant and growth, this will have influence on the uniformity of the crop.

The number of analysed bulbs with measured ground truth volume (49) is relatively low, so no strong statements can be made about the effectiveness for different imaging methods. Added to this, the stitching introduces another kind of measurement noise. This makes it hard to make a good prediction, but also makes it hard confidently assert which method performs best. For a proper evaluation, stitching would need to be improved.

The methods assumed that the depth pixels were the distance from the camera along the viewing axis. This is not necessarily the case for bulbs that are not in the center of the image, so accounting for this can improve results. However, the impact of this is likely low, since the depth values of the conveyor belt look homogeneous across the images. The predicted volume is also not converted to a known volume unit at this point, which is generally desired and necessary when you combine data from cameras at different heights. For analysis of performance of a vision system for volume measurements, a correlation is sufficient. The correlations as seen in this study are lower than the correlations that were seen in earlier measurements. This is probably caused by the higher speed of the conveyor belt in the commercial peeling line, compared to the belt speed setting in the earlier lab measurements. With the lower speeds the stitching quality of the image was better, thereby giving more reliable volume measurements.

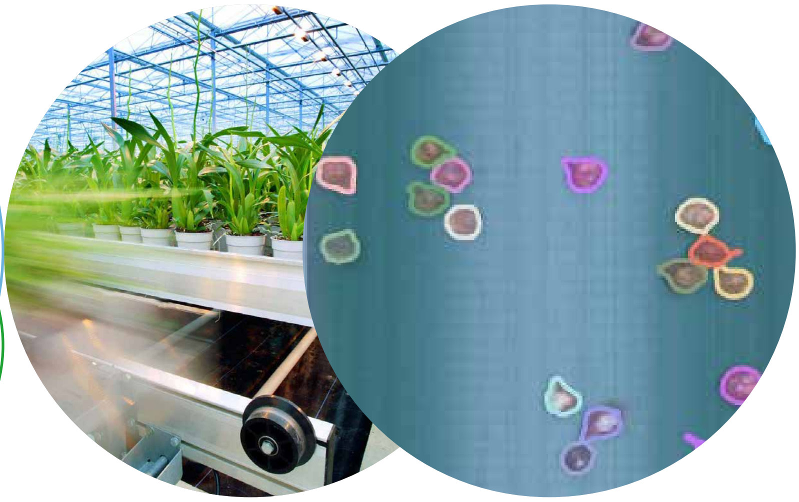
5 Conclusion

Methods to predict volume of tulip bulbs using an RGB-D camera have been developed and evaluated, using a deep learning model to segment the bulbs. Of the developed methods, surprisingly the one with the highest correlation to the measured volume did not make use of the depth information (R^2 =0.65). This could be an indication that there is too much noise to get sufficient predictive value from the (depth)-images.

Possible causes of this have been identified. Poor image stitching in the line scan likely had a negative impact on the prediction quality. In some cases, incorrect segmentation by the deep learning model and badly peeled bulbs also could have had a negative impact.

There are several possible improvement to be made, but with the aforementioned drawbacks, some of these could be hard to evaluate. Removing the image stitching errors would make it more clear where other errors originate from.

To explore the potential of nature to improve the quality of life



Wageningen University & Research,
BU Greenhouse Horticulture
P.O. Box 20
2665 ZG Bleiswijk
Violierenweg 1
2665 MV Bleiswijk
The Netherlands
T +31 (0)317 48 56 06
www.wur.nl/glastuinbouw

Report WPR-1312

The mission of Wageningen University & Research is "To explore the potential of nature to improve the quality of life". Under the banner Wageningen University & Research, Wageningen University and the specialised research institutes of the Wageningen Research Foundation have joined forces in contributing to finding solutions to important questions in the domain of healthy food and living environment. With its roughly 30 branches, 7,600 employees (6,700 fte) and 13,100 students and over 150,000 participants to WUR's Life Long Learning, Wageningen University & Research is one of the leading organisations in its domain. The unique Wageningen approach lies in its integrated approach to issues and the collaboration between different disciplines.