



UNIVERSITY OF TRENTO
DEPARTMENT OF INFORMATION ENGINEERING AND
COMPUTER SCIENCE

KIWI-RECON: A Computer Vision Approach

Signal, Image and Video

Ettore Saggiorato

February 2024

Abstract

Accurate size estimation of kiwi fruits in orchard collection bins is crucial for agricultural analysis and research. This project employs a novel approach, utilizing image processing and machine learning techniques, to estimate the size and weight distribution of kiwi fruits. The workflow involves camera calibration, bin detection, and perspective transformation for accurate image processing. Instance segmentation using Mask-RCNN, fine-tuned on a manually labeled dataset, enables precise identification of individual kiwi fruits. Further steps involve obstruction removal, stem detection, and ellipsoid fitting to enhance measurement accuracy. The project's results, presented through visualizations and plots, demonstrate promising outcomes. Future work includes optimizing execution times, expanding the dataset for improved model precision, and exploring mobile device implementation for field use. The project marks a significant advancement in non-invasive, efficient, and reliable fruit size estimation, contributing to the enhancement of orchard productivity.

1 Introduction

Estimating the size of kiwi fruits on collection bins is crucial for farmers and researchers, providing the means to compare production between orchards or sections. This process yields precise data contributing to the continuous enhancement of production volume and quality over time. The significance of this estimation becomes particularly evident in light of climate change, which has a substantial impact on agriculture [1, 2].

This approach addresses a crucial gap in the current fruit measurement methods. Industrial measurements involve machines that, while efficient, cannot be practically employed in orchards. On the other hand, manual measurements on fields are time-consuming and resource-intensive.

The conventional industrial measurement method involves a conveyor belt, where individual fruits are rotated to visualize their shape and imperfections, measure their size, and subsequently undergo individual weighing and classification. This process, however, is impractical for field use.

2 Project Objectives

This project aims to assess the quality distribution of kiwi fruits within orchards, utilizing 2D images of filled collection bins. The underlying assumption is that kiwi-fruit mass, similar to mango mass, can be estimated from its geometric dimensions through optical measurements [3]. The calculation of kiwi-fruit mass is expressed as $M = kV$, where the bulk density $k = 544.73 - 572.17 \text{ kg/m}^3$ [4] and $V [\text{cm}^3]$ denotes the volume.

Kiwi fruit exhibits variations in size and shape among different varieties. In this project, *Actinidia deliciosa* ‘Hayward’ was specifically studied, revealing sizes at harvest maturity of approximately 5-8 cm in length and 4.5-5.5 cm in diameter. The fruit typically exhibits an oval or elliptical shape, characterized by two different measurements at the diameter, as illustrated in Figure 1.



Figure 1: Definitions of fruit length, maximum diameter of the equatorial section (MaDES), and minimum diameter of the equatorial section (MiDES). (a) Top view of a kiwifruit; (b) Side view of the kiwifruit.

Source: [5].

3 Kiwi-Fruit Categories in Italy

In Italy, kiwi-fruit categories are classified based on weight, size, defects, and shape. Fruits with a non-oval or elliptical shape or smaller defects are categorized for the industry. Those with slight defects fall into the II category and are further classified by size and weight. Kiwis with no defects and an oval or elliptical shape are designated as I category and are reclassified based on size and weight.

4 Datasets

4.1 Fruit-Images-Dataset

The Fruit-Images-Dataset [6, 7, 8] is a dataset with 131 calsses of fruits and vegetables. Images size is 100x100 pixels, with pure white background for easy masking. In this project only kiwi-fruit related images were considered. These images comprehend two different fruits, rotated by 360 degrees on the x and y axis. An example is shown in Figure 2.



Figure 2: Example images from Fruit-Images-Dataset.

4.2 Orchard-Captured Images

A dataset comprising 43 images of kiwifruit bins was collected directly in the orchard, providing essential input data for the project. All images were captured from a top-down perspective, ensuring a comprehensive view of filled collection bins while minimizing side views. Therse boundaries allow to have a uniform physical dimension of the collection bins, so that they can be used as a scale, and that the top layer of the fruit is close to the top of the bin so that when calculating its size the distance between each fruit and the camera doesn't influence the measurements, as seen by the thin lens theorem.

A dataset consisting of 43 images of kiwifruit bins was directly collected in the orchard, serving as input data for the project. All images were captured from a top-down perspective, ensuring a comprehensive view of filled collection bins while minimizing side views. These constraints are crucial to maintaining a uniform physical dimension of the collection bins, allowing them to be used as a scale. Moreover, positioning the top layer of the fruit flat and close to the top of the bin ensures that, when calculating its size, the distance between each fruit and the camera doesn't influence the measurements, as dictated by the thin lens theorem. An example is shown in Figure 3.



Figure 3: Example of images captured in the orchard.

4.3 Reference Data

The reference dataset comprises measurements and weights recorded by the co-operative during the fruit assessment process. Notably, the dataset's errors and approximations remain unknown. Its data are shown in Table 1.

| Category 1 | | Category 2 | | Industry | |
|------------|------------|------------|------------|----------|------------|
| Grade | Percentage | Grade | Percentage | Grade | Percentage |
| 150+ | 3.35% | 120+ | 1.99% | Industry | 6.74% |
| 125/150 | 13.78% | 100/120 | 1.44% | | |
| 115/125 | 10.39% | 80/100 | 1.17% | | |
| 105/115 | 11.99% | 70/80 | 0.34% | | |
| 95/105 | 13.27% | | | | |
| 85/95 | 12.96% | | | | |
| 75/85 | 11.64% | | | | |
| 70/75 | 4.72% | | | | |
| 65/70 | 3.17% | | | | |
| 60/65 | 3.05% | | | | |

Table 1: Reference dataset

5 Methodology

The project follows a linear workflow, which is elucidated in this section. Each image from the Orchard-Captured Images undergoes the following processing steps. The entire workflow is summarized in Figure 7.

5.1 Camera Calibration

Camera calibration parameters are computed based on an 11x9 grid, as per the OpenCV documentation [9]. This process is crucial for correcting image distortions and enhancing measurement accuracy.

5.2 Bin Detection and Perspective Transform

This step is dedicated to the precise delineation of the region containing the fruit bin. Utilizing a series of image processing operations, the visibility of the bin is enhanced. To counteract the impact of ambient conditions on bin detection, color thresholding and morphological operations are systematically applied. Subsequently, the edges are detected using the Canny edge detection algorithm. The subsequent phase involves amalgamating both color and edge masks, creating a comprehensive mask that ensures a more accurate representation of the fruit bin region. Contours are then extracted from this combined mask, and the contour with the maximum area is presumed to correspond to the fruit bin.

Finally, a perspective transformation is implemented to crop and correct the original image based on the identified bin region. Acknowledging the typical square shape of the bin, the side of the resulting cropped image is defined as the mean length of the polygon. This strategic decision is crucial in preserving optimal pixel density—a critical factor for maintaining measurement precision. Notably, stretching the image to higher resolutions may compromise accuracy, emphasizing the fundamental role of this step in the overall image processing pipeline.

A more precise method that searched for straight lines was developed but deemed too unreliable, and hence, it wasn't employed in the final implementation. Both methods are visually demonstrated in Figure 4.

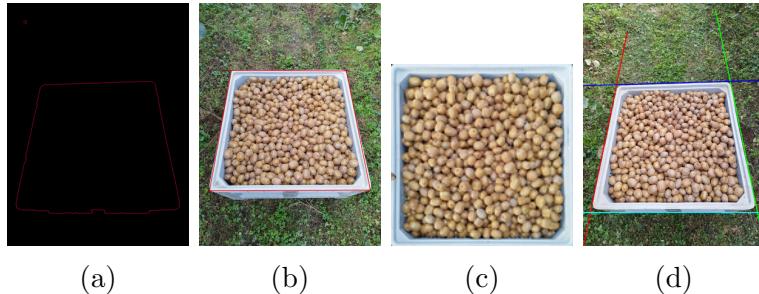


Figure 4: Bin detection methods in detail: (a): Detected contours after image processing. (b): Approximated bin edges. (c): Transformed Output. (d): Straight line methods.

5.3 Instance Segmentation

Mask-RCNN [10], a state-of-the-art framework for instance segmentation, renowned for its speed, versatility and ease of use, is employed for precise instance segmentation, extracting individual masks for each kiwi fruit. This step is crucial as it makes possible individual fruits size estimation.

5.3.1 Fine-tuning Mask-RCNN

The fine-tuning process of Mask-RCNN involved manual labeling of four images of bins using the VGG Image Annotator [12]. Additionally, kiwi images from the Fruit-Images-Dataset were processed to conform to the same data structure. Training was

conducted with the images split into the conventional 80-10-10 sets (training, validation, testing), following the guidelines outlined in the Mask-RCNN documentation [11].

5.4 Kiwi instances selection

Due to the amount of data used to train Mask-RCNN, the model generalization capabilities are not the best. Also kiwi-fruit instances need to be filtered so that only non-obstructed fruits positioned flat wrt the camera are kept for measurement. Following steps in this section are done sequentially, instance by instance. If a fruit is excluded at one step, it doesn't reach the next step. This is done for obvious optimization reasons.

5.5 Selection of Kiwi Instances

The substantial volume of data used to train Mask-RCNN has implications for the model's generalization capabilities. To enhance precision, the selection process for kiwi-fruit instances is crucial. Instances are filtered to retain only those that are non-obstructed and positioned flat with respect to the camera for accurate measurement. The steps outlined in this section are executed sequentially, handling each instance individually. If a fruit is excluded at any step, it does not proceed to the subsequent steps, optimizing the process for efficiency. Figure 5 shows as example of detected kiwi fruit instances inside fruit bins.

5.5.1 Score and Area

Mask-RCNN provides masks along with their respective scores. Masks with a score less than 70% are discarded to ensure reliable results. To address potential generalization errors, masks with an area less than 5% of the total image area are excluded, helping eliminate inaccurately selected instances. Additionally, polygons are approximated on each mask contour, and masks with multiple contours are excluded from the measurements.

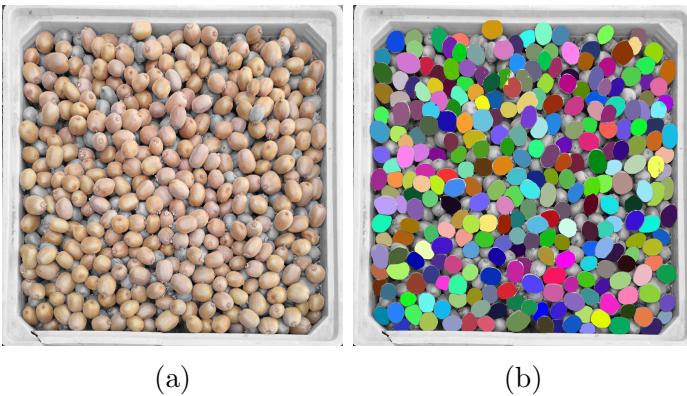


Figure 5: (a): detected fruits are colored, the background is in black and white. (b): detected fruit masks are randomly colored.

5.5.2 Obstruction Removal

For accurate size measurements of kiwi fruits, they are approximated to ellipses, a method known for its superior accuracy [13, 14, 15]. Two types of ellipses are estimated: one based on the contours of the fruit mask and another on the convex hull (the smallest convex shape enclosing all given points) of the fruit mask. The areas of these ellipses are compared, and obstructed fruits, deviating substantially from an elliptical form, are consequently removed.

5.5.3 Stem Removal

Due to the varying positions of kiwifruits inside the fruit bin, we aim to retain only those positioned optimally for measurement (flat towards the camera). As kiwis typically have the stem attachment on the top and the flower attachment on the bottom, their presence is verified. This is achieved by applying erosion and dilation to remove noise and smooth surfaces. Subsequently, contours within the masked kiwi are searched, and if any are detected, the corresponding fruit is excluded from the measurement. An example of this process can be seen in Figure 6

In future iterations, the addition of a layer to detect defects and unusual shapes could further enhance measurement accuracy.

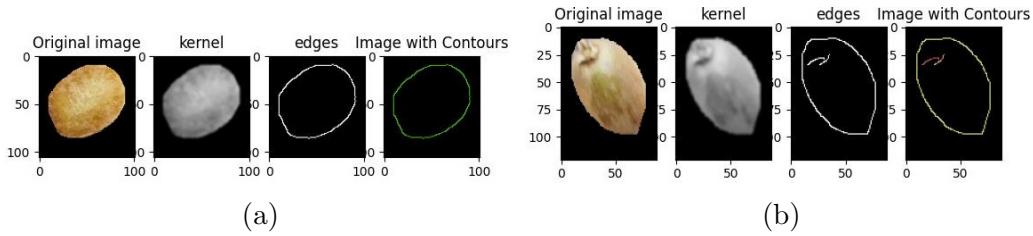


Figure 6: Stem removal process. **(a)**: no stem presence, **(b)** the stem was detected, note the color difference as multiple edges were detected

6 Size Estimation and Classification

The dimensions of the ellipses for each kiwi instance, excluding occluded objects, are averaged to mitigate overestimation. Subsequently, the size is converted from pixels to centimeters, and the volume is approximated to that of an ellipsoid. It's important to note that this introduces approximation errors due to the calculation using the equatorial size of the ellipse, while kiwi fruits exhibit two different diameters at equatorial sections. For this reason a $bias = 0.8$ parameter was introduced to reduce approximation errors.

Finally, the kiwis are classified, and the results are graphically presented in comparison to the reference data, as shown in Figure 8.

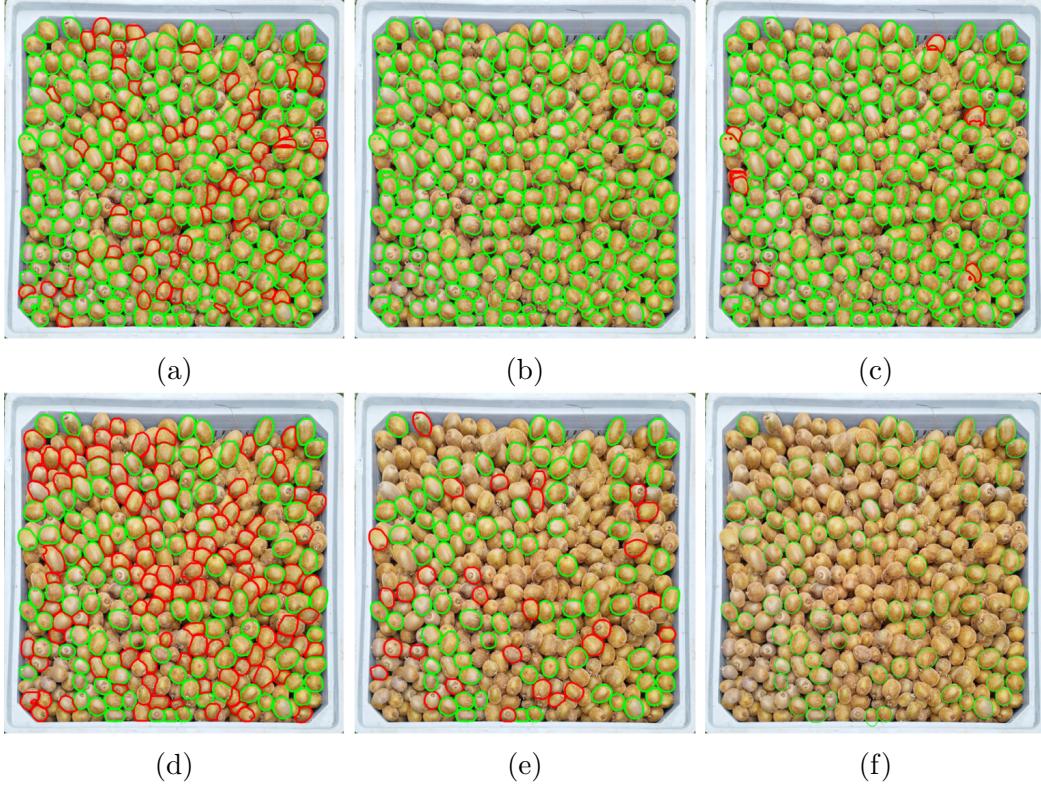


Figure 7: Green indicates fruits advancing to the next stage, while red denotes exclusions. (a): Score selection step; (b): Area selection step; (c): Polygons overlap step; (d): Occlusion filter step; (e): Presence of the stem; (f): Final result.

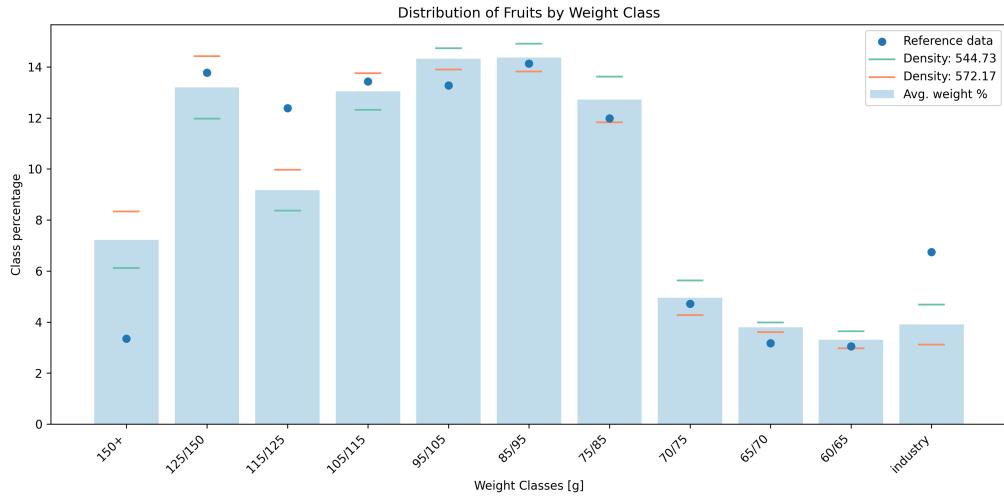


Figure 8: Distribution of kiwis by weight class, with *bias* parameter.

7 Results and Conclusions

The results, illustrated in Figure 8, showcase promising outcomes. A notable observation is the discernible difference between the 1st category (150+ grams) and the industry category. This disparity could be attributed to off-grade kiwis, especially

the flat ones, which, when positioned to the side, are misclassified as large kiwis (150+ grams) due to their geometric properties.

An intriguing aspect of the project lies in its methodology, where measurements are not conducted kiwi-by-kiwi, but rather estimated through sampling the entire population. This approach allows for quick data acquisition and provides directly applicable insights for orchard management.

8 Future Works

Although the project has shown promising outcomes, there exist avenues for further enhancement and development. To optimize execution times, future work involves refining Mask-RCNN to handle high mask volumes in larger images.

Expanding the dataset to include a diverse array of images, encompassing various perspectives, sizes, and weights, holds potential for improving precision, reducing false positives, and refining the overall model. The pursuit of zero false positives is a crucial objective [16], ensuring the calculated data's reliability for practical field use. This, in turn, contributes to ongoing research endeavors and the continual improvement of orchard productivity. An alternative avenue for exploration includes utilizing short video recordings of fruit bins, analyzing them frame-by-frame, and averaging the results for more robust measurements.

Looking ahead, the prospect of adapting the project to run on mobile devices, such as smartphones, emerges as a valuable consideration. This extension would extend the utility of the application to farmers and researchers working in orchards without reliable telecom connections, enhancing accessibility and usability in real-world settings.

References

- [1] Reddy, P.P. (2015). *Impacts of Climate Change on Agriculture*. In: Climate Resilient Agriculture for Ensuring Food Security. Springer, New Delhi. https://doi.org/10.1007/978-81-322-2199-9_4
- [2] Bhattacharyya, P., Pathak, H., Pal, S. (2020). *Impact of Climate Change on Agriculture: Evidence and Predictions*. In: Climate Smart Agriculture. Green Energy and Technology. Springer, Singapore. https://doi.org/10.1007/978-981-15-9132-7_2
- [3] Spreer, W., Müller, J. (2011). *Estimating the mass of mango fruit (*Mangifera indica*, cv. Chok Anan) from its geometric dimensions by optical measurement*. Computers and Electronics in Agriculture, 75, 125–131.
- [4] Seyed Razavi and Maryam Bahram-Parvar, “Some Physical and Mechanical Properties of Kiwifruit,” *International Journal of Food Engineering*, vol. 3, pp. 1-16, 2007. <https://doi.org/10.2202/1556-3758.1276>
- [5] Fu, L., Sun, S., Li, R., & Wang, S. (2016). Classification of Kiwifruit Grades Based on Fruit Shape Using a Single Camera. *Sensors (Basel)*, 16(7), 1012. <https://doi.org/10.3390/s16071012>. PMID: 27376292; PMCID: PMC4970062.
- [6] Horea Muresan, Mihai Oltean, *Fruit recognition from images using deep learning*, Acta Univ. Sapientiae, Informatica Vol. 10, Issue 1, pp. 26-42, 2018.
- [7] Horea-Bogdan Muresan, *An Automated Algorithm for Fruit Image Dataset Building*, 2022 17th Conference on Computer Science and Intelligence Systems (FedCSIS), 04-07 September, 2022, doi:<https://doi.org/10.15439/2022F58>.
- [8] Git Repo, <https://github.com/Horea94/Fruit-Images-Dataset/tree/master>
- [9] OpenCV documentation about camera calibration, https://docs.opencv.org/4.x/dc/dbb/tutorial_py_calibration.html
- [10] HE, Kaiming, et al. Mask r-cnn. In: Proceedings of the IEEE international conference on computer vision. 2017. p. 2961-2969.
- [11] Mask-RCNN docs, <https://web.archive.org/web/20180321023221/https://engineering.matterport.com/color-instance-segmentation-with-mask-r-cnn-and-tensorflow-7c761e238b46>
- [12] VGG Image Annotator (VIA) Abhishek Dutta, Ankush Gupta and Andrew Zisserman <https://annotate.officialstatistics.org/>
- [13] V. Voss, H. Suesse, R. Neubauer, *Moment-Based Invariant Fitting of Elliptical Segments*, Proceedings of the International Conference on Computer Analysis of Images and Patterns, 1995, pp. 562–567.
- [14] P.L. Rosin, *Measuring Shape: Ellipticity, Rectangularity, and Triangularity*, Proceedings of the 15th International Conference on Pattern Recognition, 2000, pp. 952–955.

- [15] M. Stojmenovic, A. Nayak, *Direct Ellipse Fitting and Measuring Based on Shape Boundaries*, Proceedings of the Pacific-Rim Symposium on Image and Video Technology, 2007, pp. 221–235.
- [16] Wang, Z.; Walsh, K.B.; Verma, B. *On-Tree Mango Fruit Size Estimation Using RGB-D Images*. Sensors 2017, 17, 2738. <https://doi.org/10.3390/s17122738>