



## Data Article

# AmodalAppleSize\_RGB-D dataset: RGB-D images of apple trees annotated with modal and amodal segmentation masks for fruit detection, visibility and size estimation



Jordi Gené-Mola<sup>a,b,\*</sup>, Mar Ferrer-Ferrer<sup>b</sup>, Jochen Hemming<sup>c</sup>, Pieter van Daltsen<sup>c</sup>, Dirk de Hoog<sup>c</sup>, Ricardo Sanz-Cortiella<sup>b</sup>, Joan R. Rosell-Polo<sup>b</sup>, Josep-Ramon Morros<sup>d</sup>, Verónica Vilaplana<sup>d</sup>, Javier Ruiz-Hidalgo<sup>d</sup>, Eduard Gregorio<sup>b</sup>

<sup>a</sup> Efficient Use of Water in Agriculture Program, Institute of AgriFood Research and Technology (IRTA), Fruitcentre, Parc Científic i Tecnològic Agroalimentari de Gardeny (PCiTAL), 25003 Lleida, Catalonia, Spain

<sup>b</sup> Research Group in AgroICT& Precision Agriculture - GRAP, Department of Agricultural and Forest Sciences and Engineering, Universitat de Lleida (UdL) - Agrotecnio-CERCA Center, Lleida, Catalonia, Spain

<sup>c</sup> Wageningen University and Research, 6700 AA Wageningen, the Netherlands

<sup>d</sup> Department of Signal Theory and Communications, Universitat Politècnica de Catalunya, Barcelona, Catalonia, Spain

## ARTICLE INFO

## Article history:

Received 14 December 2023

Accepted 18 December 2023

Available online 30 December 2023

Dataset link: [AmodalAppleSize\\_RGB-D dataset \(Original data\)](#)

## ABSTRACT

The present dataset comprises a collection of RGB-D apple tree images that can be used to train and test computer vision-based fruit detection and sizing methods. This dataset encompasses two distinct sets of data obtained from a Fuji and an Elstar apple orchards. The Fuji apple orchard sub-set consists of 3925 RGB-D images containing a total of 15,335 apples annotated with both modal and amodal apple segmentation masks. Modal masks denote the visible portions of the apples, whereas amodal masks encompass both visible and occluded apple regions. Notably, this dataset is the first public resource to incorporate on-tree fruit amodal masks. This pioneering inclusion addresses a critical gap in

DOI of original article: [10.1016/j.compag.2023.107854](https://doi.org/10.1016/j.compag.2023.107854)

\* Corresponding author at: Efficient Use of Water in Agriculture Program, Institute of AgriFood Research and Technology (IRTA), Fruitcentre, Parc Científic i Tecnològic Agroalimentari de Gardeny (PCiTAL), 25003 Lleida, Catalonia, Spain.

E-mail address: [jordi.gene@irta.cat](mailto:jordi.gene@irta.cat) (J. Gené-Mola).

Social media: [@GrapUdL](#) (J. Gené-Mola)

<https://doi.org/10.1016/j.dib.2023.110000>

2352-3409/© 2023 Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>)

**Keywords:**

Instance segmentation  
 Modal segmentation  
 Amodal segmentation  
 Yield prediction  
 Depth image  
 Fruit measurement  
 Fruit visibility  
 Agricultural robotics

existing datasets, enabling the development of robust automatic fruit sizing methods and accurate fruit visibility estimation, particularly in the presence of partial occlusions. Besides the fruit segmentation masks, the dataset also includes the fruit size (calliper) ground truth for each annotated apple. The second sub-set comprises 2731 RGB-D images capturing five Elstar apple trees at four distinct growth stages. This sub-set includes mean diameter information for each tree at every growth stage and serves as a valuable resource for evaluating fruit sizing methods trained with the first sub-set. The present data was employed in the research paper titled "Looking behind occlusions: a study on amodal segmentation for robust on-tree apple fruit size estimation" [1].

© 2023 Published by Elsevier Inc.

This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>)

---

**Specifications Table**

Subject	Agronomy and Crop Science, Horticulture, Agriculture Engineering, Computer Vision and Pattern Recognition.
Specific subject area	Precision Agriculture, Fruit Detection, Fruit Size Estimation, Agricultural Robotics, Machine Learning, Deep Learning
Data format	Raw images: JPG Processed RGB images: JPG Processed depth images: NPY (numpy) Modal and amodal segmentation masks: JSON Fruit size ground truth: TXT Images focal length: TXT
Type of data	RGB-D images; modal and amodal segmentation masks; fruit size ground truth.
Data collection	Fuji and Elstar apple trees were imaged using a Canon EOS 60 DSLR and a Nikon Z6 camera respectively. All Fuji apples were manually measured with a Vernier calliper and linked to an apple ID, while 15 Elstar apples per tree were measured providing mean tree diameters. Depth images were generated via Structure-from-Motion (SfM) and Multi-View-Stereo (MVS) using Agisoft Metashape. A total of 3925 image patches of 1024 × 1024-pixel were annotated with modal and amodal masks linked to the apple IDs. Modal masks were annotated using VIA annotator software. Amodal masks were generated based on the apples' spherical shape and diameter-to-pixel conversion using the pinhole camera model.
Data source location	Fuji Data: Agramunt, Catalonia, Spain (latitude: 41° 44' 47.07" N; longitude: 1° 01' 52.23" E; 312 m a.s.l. ETRS89) Elstar Data: Randwijk, The Netherlands (latitude: 51° 56' 18.25" N; longitude: 5° 42' 24.33" E; 312 m a.s.l. ETRS89)
Data accessibility	Repository name: Dataverse Data identification number: <a href="https://doi.org/10.34810/data916">https://doi.org/10.34810/data916</a> [2] Direct URL to data: <a href="https://dataverse.csuc.cat/dataset.xhtml?persistentId=doi:10.34810/data916">https://dataverse.csuc.cat/dataset.xhtml?persistentId=doi:10.34810/data916</a>
Related research article	Looking behind occlusions: A study on amodal segmentation for robust on-tree apple fruit size estimation. Gené-Mola J, Ferrer-Ferrer M, Gregorio E, Blok PM, Hemming J, Morros JR, Rosell-Polo JR, Vilaplana V, Ruiz-Hidalgo J. 2023. Computers and Electronics in Agriculture 209 (2023) 107,854. DOI: <a href="https://doi.org/10.1016/j.compag.2023.107854">https://doi.org/10.1016/j.compag.2023.107854</a> [1]

---

**1. Value of the Data**

- **Groundbreaking Amodal Segmentation:** This dataset is pioneering by being the first publicly available dataset for on-tree fruit detection and sizing that includes amodal segmentation masks ground truth. The inclusion of amodal masks, covering both visible and occluded apple regions, offers a unique advantage for on-tree fruit sizing even in the

presence of partially occlusions. This feature enables the development of robust automatic fruit sizing methods and accurate fruit visibility estimation, filling a gap in existing datasets.

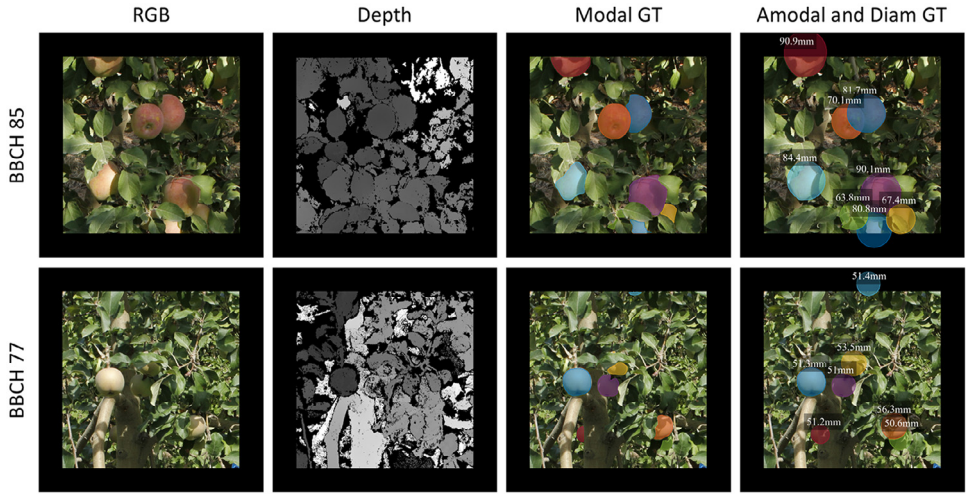
- **Benefits Across the Agricultural Ecosystem:** The use of this data will benefit researchers in the field of computer vision applied to fruit monitoring. Consequently, this will also benefit the rest of stakeholders of the value chain, such as technology companies offering fruit monitoring solutions, as well as agricultural consultants and farmers that need to obtain fruit production maps and monitor the fruit growth to manage the orchards efficiently.
- **Algorithm Development and Reproducibility:** The provided dataset will facilitate the development and training of modal and amodal segmentation algorithms capable of detecting both visible and occluded parts of apples. The dataset is prepared to be applied to train and fine-tune existing instance segmentation neural networks such as Mask R-CNN [3]. Additionally, the dataset includes ground truth fruit sizes, enabling the implementation and reproduction of results obtained in the fruit sizing methods proposed in [1] and [4].
- **Rich and Diverse Dataset:** With an extensive collection of RGB-D images featuring both modal and amodal segmentation masks, researchers can delve into advanced fruit detection and sizing techniques. The dataset encompasses apples at different maturity stages from both Fuji and Elstar orchards, enabling comparative studies across different apple varieties and growth stages. This multi-orchard perspective broadens the dataset's applications and provides a comprehensive context for researching fruit detection and sizing algorithms.

## 2. Background

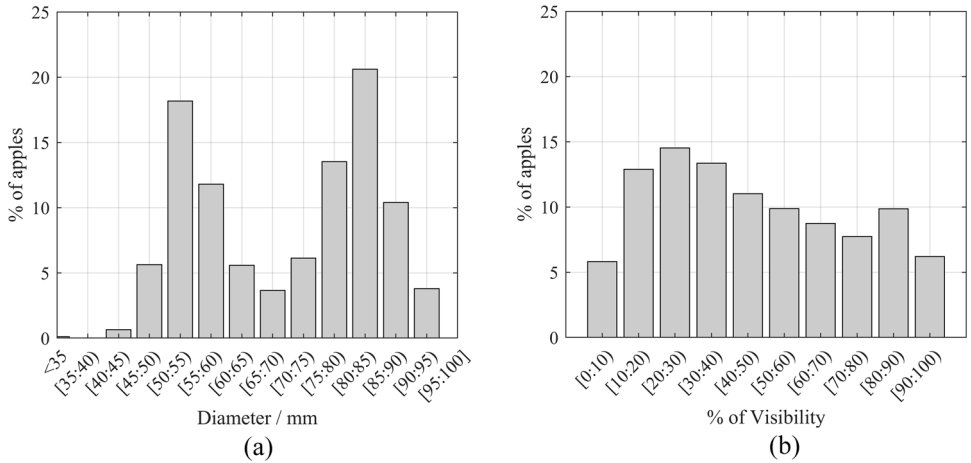
The development of automatic fruit sizing methods is of interest for orchard management, with applications in fruit thinning, yield prediction, mapping, and automated harvesting [5]. Traditional manual measurements using Vernier callipers are limited in precision and efficiency, prompting the exploration of image-based approaches. Currently, many image-based fruit sizing methods use modal instance segmentation methods to predict the visible regions of fruits and subsequently measure their size [6]. However, these methods encounter challenges in effectively handling partially occluded fruits. Recently, the authors proposed a solution to this challenge introducing the use of amodal instance segmentation, predicting both visible and occluded regions of fruits [1]. The present data article adds significant value to the published research paper by providing a comprehensive and publicly available resource for training and testing the proposed modal-amodal segmentation algorithm. Researchers can utilize this dataset to replicate experiments, validate findings, and further advance the field of fruit detection and sizing in orchard management.

## 3. Data Description

This dataset comprises two distinct subsets of data. The first subset consists of 3925 RGB-D Fuji apple images, encompassing a total of 15,335 apple instances. Each apple instance is annotated with modal and amodal segmentation masks, as well as apple diameter ground truth. The modal mask represents the visible regions of apples, whereas the amodal mask includes both visible and occluded regions (Fig. 1). This subset is suitable for training, validating and testing fruit detection and sizing methods, as demonstrated in previous studies [1,4]. The availability of modal and amodal masks also enables the estimation of apples visibility, which is of interest in fruit sizing and harvesting robotics [7]. Fig. 2 shows the distribution of diameters and visibility of all annotated apples in this subset. The diameter distribution shows two peaks in the ranges of [50:55] mm and [80:85] mm, which correspond to the mean diameter of apples measured in the BBCH77 and BBCH85 growth stages [5] from which this data was acquired.



**Fig. 1.** Fuji images and annotations. Each row corresponds to an image acquired at a different apple maturity stage: BBCH 85 and BBCH 77. The first column shows the RGB image, the second column displays the Depth image, the third column illustrates the modal segmentation masks ground truth, and the fourth column presents the amodal segmentation masks along with the apple diameters ground truth.



**Fig. 2.** a) Bar plot showing the percentage of apples based on diameter. b) Bar plot showing the percentage of apples based on different levels of visibility.

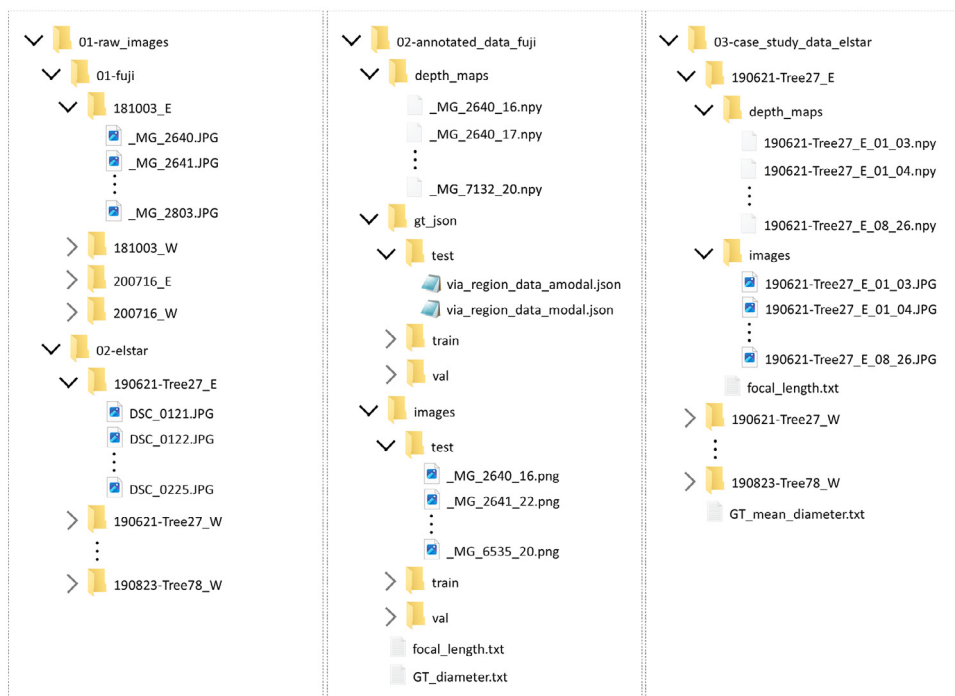
The second subset of data comprises 2731 RGB-D images captured in an Elstar apple orchard. These images were obtained from five apple trees on four different dates, corresponding to the apple growth stages of BBCH75, BBCH77, BBCH79 and BBCH85, in the BBCH growth scale [8]. This subset includes the mean diameter of apples on each tree at every growth stage (Table 1). It serves as a valuable resource for evaluation of fruit sizing methods trained with the aforementioned subset, under diverse data and environmental conditions different from those used during the training phase.

The dataset can be downloaded at <https://doi.org/10.34810/data916> [2]. Once it is unpacked, the dataset contains three main folders: '01-raw\_images'; '02-annotated\_data\_fuji'; and '03-case\_study\_data\_elstar' (Fig. 3). The '01-raw\_images' directory contains all raw images used to

**Table 1**

Mean calliper of apples grown in the 5 imaged Elstar trees on the four different dates that data was collected. All diameter values are in millimetres.

Date	Tree 27	Tree 48	Tree 63	Tree 70	Tree78
21/06/19	36.05	44.39			
03/07/19	43.97	51.61	40.12	43.98	50.92
16/07/19	57.33	60.58	49.15	51.88	57.64
23/08/19	73.49	76.19	66.76	69.77	75.49



**Fig. 3.** Dataset files and folder structure for each data subset: row data, annotated Fuji data and Elstar data.

build the dataset. The Fuji apple images are structured with subfolders depending on the date and side of the row of trees from which the images were captured. These subfolders are named with the following coding: `YYMMDD_S`, where `YYMMDD` refers to the date (year, month and day), and `S` refers to the side of the row of trees imaged (East and West). All raw Fuji images have a size of  $5184 \times 3456$  pixels and were saved in JPG format. The raw Elstar apple images are structured with subfolders depending on the date, the tree number and the side of the row of trees that was imaged. These subfolders are named with the coding `YYMMDD-TreeXX_S`, where `XX` is the imaged tree number. All raw Elstar images have a size of  $6000 \times 4000$  pixels and were saved in JPG format.

The '02-annotated\_data\_fuji' directory contains all annotated Fuji RGB-D images. Images of this subset have a size of  $1300 \times 1300$  pixels and were obtained by cropping raw image patches containing the Fuji apples that were manually measured in the field. This subset was split in training (2304 images), validation (814 images) and test (807 images) sets. Training images include images acquired from the west side, while validation and test images were acquired from the east side. This criterion was applied because raw consecutive images had a certain percentage of overlap. By using the images acquired from one side for training and from the other side for validation and testing, the authors aimed to ensure that there is no overlap between train-

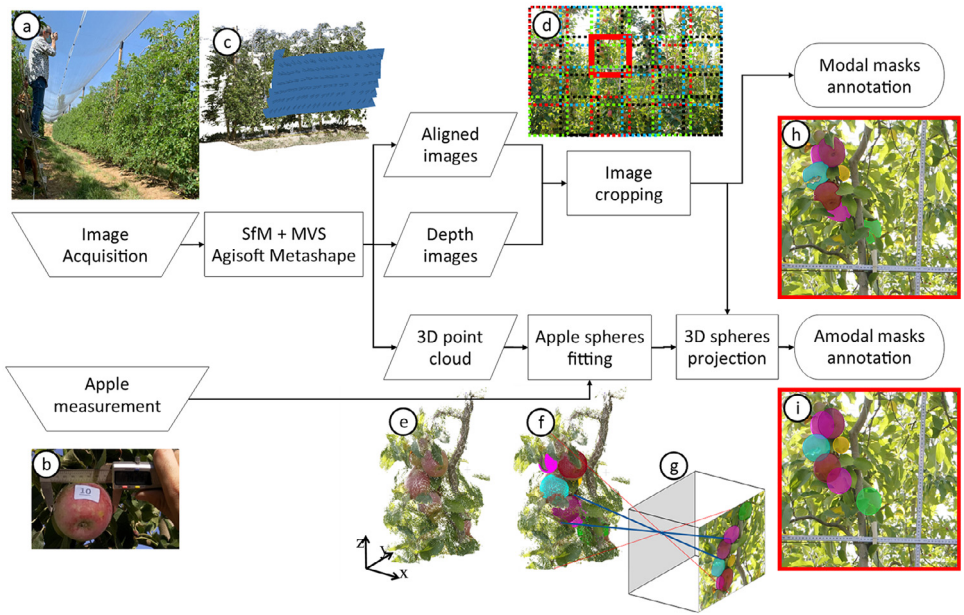


Fig. 4. Data curation and annotation outline.

ing and validation/test sets. The 'depth\_maps' subfolder includes all image depth maps. These files have the same filename as the colour images but with NPY extension. These NPY files can be loaded in Python (recommended) or other programming languages (Matlab, C++, ...) with the corresponding NumPy libraries or packages. Once loaded, they form a ndarray matrix of size  $1300 \times 1300$  pixels (equal to colour image size) with float numbers corresponding to the depth values in meters. The modal and amodal segmentation masks are provided in the 'gt\_json' subfolder with JSON extension files. These files can be loaded with Python (recommended) or other programming languages (matlab, C++, ...) with the corresponding JSON libraries. Once the JSON file is loaded, it creates a dictionary of size equal to the number of images in the set. For each image, there is a subdictionary with 'region' attributes containing the segmentation mask annotations for each instance (apples) appearing in the image. The 'shape\_attributes' includes the polygon shapes that encloses the masks, defined by its vertices points with x and y pixel coordinates, while the 'apple\_ID' attribute specifies the identification number of each annotated instance. Finally, the focal length of all annotated images is provided in the 'focal\_length.txt' file, while the diameter ground truth of all annotated apple\_ID's is provided in millimeters in the 'GT\_diameter.txt'.

The '03-case\_study\_data\_elstar' directory contains the RGB-D Elstar images. Similar to the Fuji images, the images in this subset also have a size of  $1300 \times 1300$  pixels and were obtained by cropping raw image patches. The depth maps are also saved in NPY extension and have the same size as the corresponding colour images. Besides the RGB-D images, this subset also provides information of the focal length of all images and the mean apples diameter measured on each tree at each measured date.

#### 4. Experimental Design, Materials and Methods

The methodology pipeline used to create and annotate the Fuji dataset is illustrated in Fig. 4. Raw images were captured in the field (Fig. 4a) following the procedure described in [9]. These

images were acquired in a Fuji apple orchard situated in the province of Lleida, Catalonia, Spain. Fuji apple trees were photographed using a handheld EOS 60D DSLR camera (Canon Inc. Tokyo, Japan) on two different dates to ensure variability in apples diameters. The first data acquisition was carried out on October 3rd, 2018, when the apples were at the BBCH85 growth stage. The second data acquisition took place on July 16th, 2020, with the apples at the BBCH77 growth stage.

A total of 615 apples grown on the imaged trees were manually measured and annotated with an apple\_ID. This measurement was carried out with a Vernier calliper by measuring the maximum horizontal diameter, as depicted in Fig. 4.b. The consecutive images had a substantial overlap (more than 75%) to facilitate the application of Structure-from-Motion (SfM) and Multi-View-Stereo (MVS) techniques (Fig. 4.c). These photogrammetric techniques were applied using Agisoft Professional Metashape software (v1.6.4, Agisoft LLC, St. Petersburg, Russia). This software was instrumental in obtaining intrinsic and extrinsic camera parameters (aligned images), depth maps of the images, and a 3D point cloud of the imaged trees.

The images and the depth maps were cropped into 24 image patches of  $1100 \times 1100$  pixels each (Fig. 4.d). This cropping was performed to enhance the ratio between apple size (in pixels) and image size, as the raw images contained numerous small apples in comparison to the overall image size. After cropping the raw images, the resulting image patches exhibited a more favourable apple/image size ratio for object detection neural networks, as recommended in [10].

The resulting image patches were annotated with modal masks (Fig. 4.h) using a semi-automatic annotation procedure. Initially, the Mask RCNN pretrained with the Fuji-SfM dataset [11] was used to generate initial modal segmentation masks. Subsequently, these annotations were manually revised and corrected using the VIA annotation software [12]. Additionally, all apple segmentation masks were associated with the corresponding apple\_IDs to establish a link between the masks annotations and the respective ground truth apple diameters.

To derive the amodal segmentation masks, the 3D point clouds obtained from the SfM and MVS processes were used (Fig. 4.e). All measured apples were identified in the 3D point cloud, and spheres with diameters equal to the apple diameters were fitted to the 3D models of all apples (Fig. 4.f). This sphere fitting procedure was executed using CloudCompare software (Cloud Compare [GPL software] v2.9 Omnia). For each image patch, the 3D spheres of the apples within the image field of view were projected onto the image plane following the pinhole camera model (Fig. 4.g). As the spheres are considered complete representations of apples (including both visible and occluded parts), their projection onto the image plane resulted in the amodal segmentation masks (Fig. 4.i). To account for the possibility of amodal apple masks extending beyond image edges, the image patches were padded with 100 blank pixels around the images, as indicated by the black zone around images in Fig. 1.

Regarding the Elstar images, they were captured in an Elstar apple orchard located in Randwijk, the Netherlands. This set of images includes five different trees that were photographed at four dates (June 21st, July 3rd, July 16th, and August 23rd), corresponding to the growth stages of BBCH75, BBCH77, BBCH78 and BBCH85. The camera used was a Nikon Z6 camera. The diameter of apples grown in 10 marked fruit clusters per tree was measured with a Vernier calliper on all the dates the trees were imaged, providing the mean apple diameter per tree at each growth stage (refer to Table 1). The measurement dates corresponding to apple growth stages of 75 and 85 did not coincide with the date of the image data. Therefore, the mean diameter was obtained by interpolating the mean apples diameters between the two closest measurement dates: June 13th and 25th for the BBCH75, and August 16th and 27th for the BBCH85. Regarding to the image acquisition, the procedure to obtain the RGB-D images and the focal length of the Elstar apple set was identical to the procedure outlined for Fuji set, including steps a to d of the data curation methodology described in the preceding paragraphs and illustrated in Fig. 4.

## Limitations

None.

## Ethics Statement

Authors have read and follow the [ethical requirements](#) for publication in Data in Brief and confirm that the current work does not involve human subjects, animal experiments, or any data collected from social media platforms.

## Declaration of generative AI and AI-Assisted Technologies in the Writing Process

During the preparation of this work the author(s) used ChatGPT-3.5 to improve readability and language. After using this tool, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

## Data Availability

[AmodalAppleSize\\_RGB-D dataset \(Original data\)](#) (Dataverse)

## CRedit Author Statement

**Jordi Gené-Mola:** Conceptualization, Methodology, Investigation, Data curation, Writing – original draft, Visualization, Supervision; **Mar Ferrer-Ferrer:** Methodology, Data curation, Writing – review & editing; **Jochen Hemming:** Supervision, Funding acquisition, Writing – review & editing; **Pieter van Dalfsen:** Methodology, Investigation, Data curation, Writing – review & editing; **Dirk de Hoog:** Methodology, Investigation, Data curation, Writing – review & editing; **Ricardo Sanz-Cortiella:** Investigation, Resources, Data curation, Writing – review & editing; **Joan R. Rosell-Polo:** Investigation, Supervision, Project administration, Funding acquisition, Writing – review & editing; **Josep-Ramon Morros:** Methodology, Supervision, Writing – review & editing; **Verónica Vilaplana:** Methodology, Supervision, Writing – review & editing; **Javier Ruiz-Hidalgo:** Methodology, Supervision, Writing – review & editing; **Eduard Gregorio:** Conceptualization, Methodology, Investigation, Supervision, Project administration, Funding acquisition, Writing – review & editing.

## Acknowledgments

This work was partly funded by the Departament de Recerca i Universitats de la Generalitat de Catalunya (grant [2021 LLAV 00088](#)), the Spanish Ministry of Science, Innovation and Universities (grants [RTI2018-094222-B-I00](#) [PAgFRUIT project], [PID2021-1266480B-I00](#) [PAgPROTECT project] and [PID2020-117142GB-I00](#) [DeeLight project] by MCIN/AEI/10.13039/501100011033 and by “ERDF, a way of making Europe”, by the European Union). Data presented in this paper is also part of a Public Private Partnership project Precisie Tuinbouw, WP Fruit 4.0 (PPS KV 1604-025) and financed by Topsector Tuinbouw & Uitgangsmateriaal and various private companies. The work of Jordi Gené Mola was supported by the Spanish Ministry of Universities through a Margarita Salas postdoctoral grant funded by the European Union - NextGenerationEU. We would also like to thank Nufri (especially Santiago Salamero and Oriol Morrerres) for their support during data acquisition.

## 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.



## References

- [1] J. Gené-Mola, M. Ferrer-Ferrer, E. Gregorio, P.M. Blok, J. Hemming, J.R. Morros, J.R. Rosell-Polo, V. Vilaplana, J. Ruiz-Hidalgo, Looking behind occlusions: a study on amodal segmentation for robust on-tree apple fruit size estimation, *Comput. Electron Agric.* 209 (2023) 107854, doi:[10.1016/j.compag.2023.107854](https://doi.org/10.1016/j.compag.2023.107854).
- [2] J. Gené-Mola, M. Ferrer Ferrer, J. Hemming, P. van Dalsen, D. de Hoog, R. Sanz Cortiella, J.R. Rosell Polo, J.R. Morros Rubió, V. Vilaplana, J. Ruiz Hidalgo, E. Gregorio López, AmodalAppleSize\_RGB-D Dataset, CORA.Repositori de Dades de Recerca V1 (2023), doi:[10.34810/data916](https://doi.org/10.34810/data916).
- [3] K. He, G. Gkioxari, P. Dollár, R. Girshick, Mask R-CNN, in: *Proceedings of the IEEE International Conference on Computer Vision*, 2017, pp. 2980–2988, doi:[10.1109/ICCV.2017.322](https://doi.org/10.1109/ICCV.2017.322).
- [4] M. Ferrer-Ferrer, J. Ruiz-Hidalgo, E. Gregorio, V. Vilaplana, J.R. Morros, J. Gené-Mola, Simultaneous fruit detection and size estimation using multitask deep neural networks, *Biosyst. Eng.* 233 (2023) 63–75, doi:[10.1016/j.biosystemseng.2023.07.010](https://doi.org/10.1016/j.biosystemseng.2023.07.010).
- [5] N.T. Anderson, K.B. Walsh, D. Wulfsohn, Technologies for forecasting tree fruit load and harvest timing—From ground, sky and time, *Agronomy* 11 (2021), doi:[10.3390/agronomy11071409](https://doi.org/10.3390/agronomy11071409).
- [6] J.C. Miranda, J. Gené-Mola, M. Zude-Sasse, N. Tsoulias, A. Escolà, J. Arnó, J.R. Rosell-Polo, R. Sanz-Cortiella, J.A. Martínez-Casasnovas, E. Gregorio, Fruit sizing using AI: a review of methods and challenges, *Postharvest Biol. Technol.* 206 (2023) 112587, doi:[10.1016/j.postharvbio.2023.112587](https://doi.org/10.1016/j.postharvbio.2023.112587).
- [7] W. Au, H. Zhou, T. Liu, E. Kok, X. Wang, M. Wang, C. Chen, The Monash Apple retrieving system: a review on system intelligence and apple harvesting performance, *Comput. Electron Agric.* 213 (2023), doi:[10.1016/j.compag.2023.108164](https://doi.org/10.1016/j.compag.2023.108164).
- [8] U. Meier, Growth stages of mono- and dicotyledonous plants, 2001. [10.5073/bbch0515](https://doi.org/10.5073/bbch0515).
- [9] J. Gené-Mola, R. Sanz-Cortiella, J.R. Rosell-Polo, A. Escolà, E. Gregorio, PFuji-Size dataset: a collection of images and photogrammetry-derived 3D point clouds with ground truth annotations for Fuji apple detection and size estimation in field conditions, *Data Brief* 39 (2021) 107629, doi:[10.1016/j.dib.2021.107629](https://doi.org/10.1016/j.dib.2021.107629).
- [10] A. Koirala, K.B. Walsh, Z. Wang, C. McCarthy, Deep learning – method overview and review of use for fruit detection and yield estimation, *Comput. Electron Agric.* 162 (2019) 219–234, doi:[10.1016/j.compag.2019.04.017](https://doi.org/10.1016/j.compag.2019.04.017).
- [11] J. Gené-Mola, R. Sanz-Cortiella, J.R. Rosell-Polo, J.R. Morros, J. Ruiz-Hidalgo, V. Vilaplana, E. Gregorio, Fuji-SfM dataset: a collection of annotated images and point clouds for Fuji apple detection and location using structure-from-motion photogrammetry, *Data Br* 30 (2020), doi:[10.1016/j.dib.2020.105591](https://doi.org/10.1016/j.dib.2020.105591).
- [12] A. Dutta, A. Zisserman, The VIA annotation software for images, audio and video, in: *Proceedings of the 27th ACM International Conference on Multimedia*, ACM, 2019, doi:[10.1145/3343031.3350535](https://doi.org/10.1145/3343031.3350535).