Postharvest Biology and Technology 206 (2023) 112587

Contents lists available at ScienceDirect

Postharvest Biology and Technology

journal homepage: www.elsevier.com/locate/postharvbio

Fruit sizing using AI: A review of methods and challenges

Juan C. Miranda [a], Jordi Gené-Mola [a,b], Manuela Zude-Sasse [c], Nikos Tsoulias [c], Alexandre Escolà [a], Jaume Arnó [a], Joan R. Rosell-Polo [a], Ricardo Sanz-Cortiella [a], José A. Martínez-Casasnovas [a], Eduard Gregorio [a,*]

[a] Research Group in AgroICT & Precision Agriculture, Universitat de Lleida (UdL) - Agrotecnio CERCA Center, 25198 Lleida, Catalonia, Spain

[b] Efficient Use of Water in Agriculture Program, Institute of AgriFood, Research and Technology (IRTA), Parc Científic i Tecnològic Agroalimentari de Gardeny (PCiTAL), Fruitcentre, 25003 Lleida, Catalonia, Spain.

[c] Department of Horticultural Engineering, Leibniz Institute for Agricultural Engineering and Bioeconomy (ATB), Max-Eyth-Allee, 14469 Potsdam, Germany

ARTICLE INFO

Keywords:
Artificial intelligence
Fruit detection
Fruit measure
Image processing
Deep learning
Fruit quality

ABSTRACT

Fruit size at harvest is an economically important variable for high-quality table fruit production in orchards and vineyards. In addition, knowing the number and size of the fruit on the tree is essential in the framework of precise production, harvest, and postharvest management. A prerequisite for analysis of fruit in a real-world environment is the detection and segmentation from background signal. In the last five years, deep learning convolutional neural network have become the standard method for automatic fruit detection, achieving F1-scores higher than 90 %, as well as real-time processing speeds. At the same time, different methods have been developed for, mainly, fruit size and, more rarely, fruit maturity estimation from 2D images and 3D point clouds. These sizing methods are focused on a few species like grape, apple, citrus, and mango, resulting in mean absolute error values of less than 4 mm in apple fruit. This review provides an overview of the most recent methodologies developed for in-field fruit detection/counting and sizing as well as few upcoming examples of maturity estimation. Challenges, such as sensor fusion, highly varying lighting conditions, occlusions in the canopy, shortage of public fruit datasets, and opportunities for research transfer, are discussed.

1. Introduction

Agricultural production of fresh fruit and vegetables must substantially increase to address the food demand due to the growing population, which is expected to reach 9.7 billion people by mid-century (UN, 2022). However, the production needs to respect the environment and meet the requirements of social and economic sustainability (FAO, 2017). Avoiding food waste is a major concern in all these aspects. Food waste in the fruit supply chain can be caused by lack of fruit safety and decay of fruit, as well as rejection of fruit on the market due to insufficient product quality (Nicastro and Carillo, 2021). Product decay is addressed by postharvest technologies to keep fruit at marketing quality. On the other hand, achieving the desired fruit quality is left to the farmers, but demanded by the actors of the value chain (Saitone and Sexton, 2017). If the market value is considered, the appearance represents the most important quality parameter that needs to be achieved (Musacchi and Serra, 2018). The main variable of complex appearance is the fruit size. Size as well as other variables of appearance such as colour, shape or absence of defects are addressed by means of inline grading, sorting fruit according to the different market needs. However, too large fruit are difficult to market, since they frequently show reduced storability (Paul and Pandey, 2014), whereas consumers prefer large, but not uncommonly large fruit (Iwanami, 2011). Similarly, it is difficult to find an economically reasonable market for small fruit, due to unfavourable ratio of edible to residual parts of the fruit.

Size is considered a “search characteristic” that can be assessed before purchasing (Yeo and Edwards, 2006). It plays a role in the consumer’s decision to buy. Consumers preference for apples varies depending on country, region, type of market, gender, family income, education, age, food safety factors (i.e. pesticide use) and memory of previous eating experiences (Harker et al., 2003; Bonany et al., 2013; Bavay et al., 2013; Favre et al., 2022). A study conducted in Canadian territories reported the ideal size for dessert apple ranges between 74 and 76 mm considering various ages (Hampson et al., 2002). Sorting machines in packing lines are commonly dividing fruit in size classes and the class obtained influences the fruit price. Apple size is frequently

* Corresponding author.
E-mail address: eduard.gregorio@udl.cat (E. Gregorio).

https://doi.org/10.1016/j.postharvbio.2023.112587
Received 28 February 2023; Received in revised form 4 September 2023; Accepted 19 September 2023
Available online 23 September 2023

included in trading standards or regional producers' standards. In the European market, cultivar-specific fruit size of 60 mm has been requested (OECD, 2021). However, such requests have been becoming less binding due to introduction of kids' apples and other specific products. Nevertheless, for sweet cherry, but also European plum and other stone fruits, the size of fruit is directly affecting the market price, e.g. in sweet cherry cultivar 'Celeste' even selective harvesting appears as economically vital. It was shown that at red ripening stage 55 days after full bloom (DAFB), 35 % of cherries reached > 28 mm, whereas after 60 DAFB another 40 % more passed this value-creating threshold. Selective harvesting was reasonable in this case, since the higher size class gained 20 % increased market price (Heim and Zude-Sasse, 2014; He et al., 2015). However, for drawing harvest conclusions, the information on fruit size is requested in real-time.

Strategies to produce the desired fruit size are captured in the concept of crop load management (Robinson et al., 2017), which requires the feedback on actual fruit number and fruit size in the field (Delong et al., 2006). High crop load, beyond the fruit bearing capacity may result in small fruit considering the MaluSim approach from Lakso (Penzel et al., 2020). Low crop load situation obviously results in reduced yield per area, however, yield can be even further reduced due to an increased risk of storage disorders appearing in large fruit. In control and 1-MCP treated 'Gala' apples flesh breakdown in storage increased with enhanced fruit size (Lee et al., 2013). In nectarine, fruit size served as an input variable to model fruit development and storability (Casagrande et al., 2021). Furthermore, in non-destructive quality sensing of citrus fruit, it was shown that fruit size affects the non-destructive spectral-optical analysis in the short-wave near infrared (NIR) wavelength range (Miller and Zude-Sasse, 2004; Sun et al., 2021). Fraser et al. (2003) has shown that the distribution of light varies within citrus fruit. Consequently, information on fruit size and fruit size distribution in the canopy may support development of more robust sensor calibrations.

Automatic detection, location and sizing of fruit in the field are agricultural problems in which computer vision and geo-positioning play a fundamental role. Fruit detection consists of finding a candidate region of interest (ROI) in a given image, point cloud or other type of data, and classifying it as fruit or background. The fruit location problem goes even one step further by locating the fruit in a local or global coordinate system (e.g. the position of the fruit in an image or on the Earth, respectively), and making a coordinate conversion to transfer fruit position detected in images onto the real world coordinate system. In-field fruit sizing consists of measuring the fruit (e.g. diameter, length, volume, etc.) on the tree to obtain morphological data.

Systems applied to fruit detection and sizing must deal with data acquired under a variety of lighting conditions (Chaivivatrakul and Dailey, 2014), and their performance may be affected by factors such as shadows, reflections, backlights, background colour, inclusion and occlusions (Fig. 1a,b). Other factors such as coinciding structures or the slope in an orchard (Fig. 1c) affect the open view to the fruit object and require geometric correction of sensor raw data. The combination of these factors influences the accuracy of the detection result. According to the lighting conditions, it is useful to mention work carried out in night-time conditions (Fig. 1d), with the help of artificial lighting rigs used both to lighten the scene and to reduce the undesirable effects of variable lighting. Fruit clustering, occlusions and shading (Fig. 1e,f) are other factors that need to be taken into consideration in fruit detection (Jarvinen et al., 2019).

(a) (b) (c)

(d) (e) (f)

Fig. 1. Examples of fruit detection challenges under several conditions. (a) Fuji apples with shadows (Gené-Mola et al., 2020e). (b) Golden Delicious apples (colour similar to the background) with reflections and backlights. (c) Slope in an apple tree orchard. (d) Apple tree in night-time conditions (Gené-Mola et al., 2019b). (e) Cluster of grapes (Arnó, 2008). (f) Peaches, occlusions and shading.

Regarding fruit growing stages, changes in colour and shape along the growing season affect the performance of the fruit detection system. Depending on the purpose of the study, measurements are taken at different stages: mapping of flower and fruit distributions for yield prediction (Underwood et al., 2016), or detection of fruit at several stages of growth to monitor the evolution of the orchard (Tian et al., 2019; Tsoulias et al., 2022). Nevertheless, most studies have been carried out at harvest time, when the objective is, for example, to predict the yield, map the production or for automatic harvesting purposes (Gené-Mola et al., 2019a; Wang et al., 2019). In this stage, the fruit has reached maximum size and frequently changed colour from green to yellow or red, which is less challenging compared to detection of small, green objects in green foliage (Tsoulias et al., 2020, 2023).

With respect to the algorithms used, two aspects need to be considered: 1) obtaining a high-performance fruit detection, which means having high detection rates and a low number of false positives; 2) the development of computationally efficient algorithms to achieve low processing times (Häni et al., 2020a). In this regard, the application of deep learning in computer vision and the use of 3D sensors have revolutionized fruit detection (Koirala et al., 2019a). However, the shortage of public fruit datasets, as well as the diversity of lighting conditions and capture devices, makes it difficult to compare the fruit detection algorithms that have been published (Qureshi et al., 2017). Nonetheless, efforts have been carried out to collect and classify specialized agricultural datasets that include different sensor types, fruit varieties, and field conditions (Lu and Young, 2020). In addition, it should be noted that fruit detection and sizing systems usually deal with complex, unstructured and changing agricultural environments, in contrast to the generally clearly defined targets that detection systems work on in industrial applications (Bechar and Vigneault, 2016; Zhao et al., 2016a). Although promising results have been achieved in industry environments, it is still cumbersome to determine the fruit load when these techniques are implemented in the field. All the reasons set out above explain why fruit detection and sizing appears an interesting application of AI and is currently a focal point of interest.

This work presents a review of the state-of-the-art of computer vision-based fruit detection and sizing methods. The present review work is structured in six main sections. Section 1 comprises this introduction. Sections 2 and 3 deal with fruit detection, reviewing the handcrafted computer vision and deep learning methods, respectively. Section 4 covers the field of fruit size analysis and maturity estimation that can be later applied in crop load management and yield estimation. In Section 5, the challenges to be faced when applying fruit sizing are discussed. Final conclusions are presented in Section 6.

2. Fruit detection based on handcrafted features

2.1. Background

Before the advent of deep learning, most of the computer vision algorithms relied on the identification and extraction of image features such as corners, edges and blobs, and the subsequent classification of these features that defined the image or parts of the image. The design of the methodology to extract these features was done manually (handcrafted) based on human vision insights and intuitions (Nanni et al., 2017); this is why these algorithms are known as handcrafted feature-based methods. Previous reviews of these methods are thoroughly described in Gongal et al. (2015) and in Zhao et al. (2016a). While there is no single recipe to frame all the handcrafted methods, the aim of this section is to provide an up-to-date review.

Fruit detection algorithms, as a special case of general object detection (Fig. 2), can follow approaches based on two main steps (Wang and Zheng, 2019; Ward et al., 2019): (1) candidate region proposals generation; and (2) detection and recognition.

2.2. Candidate region proposals

The generation of candidate region proposals is the step of the process in which potential regions of interest are identified from data received by sensors. As sub-tasks, this can be divided into region selection and region description. Thresholding has been one of the most commonly used methods to classify fruit and background regions

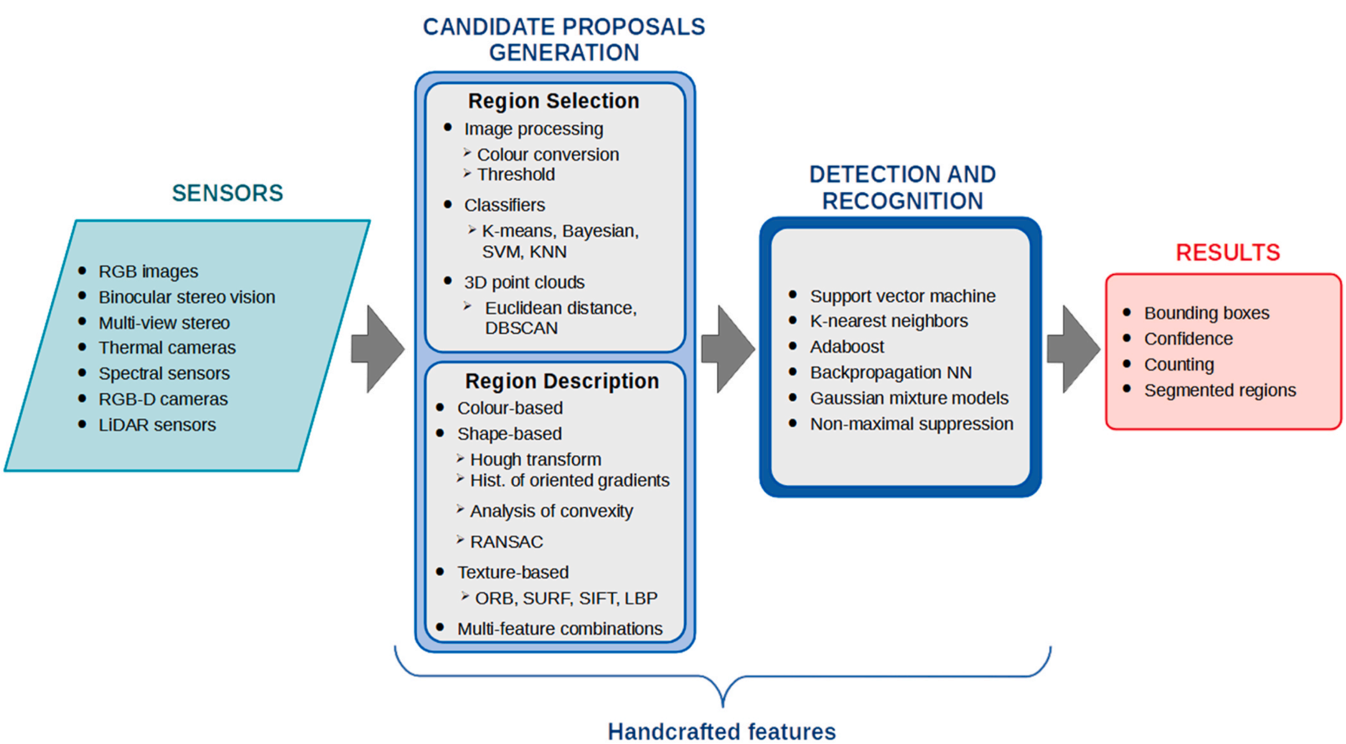

Fig. 2. Pipeline for fruit detection based on handcrafted features.

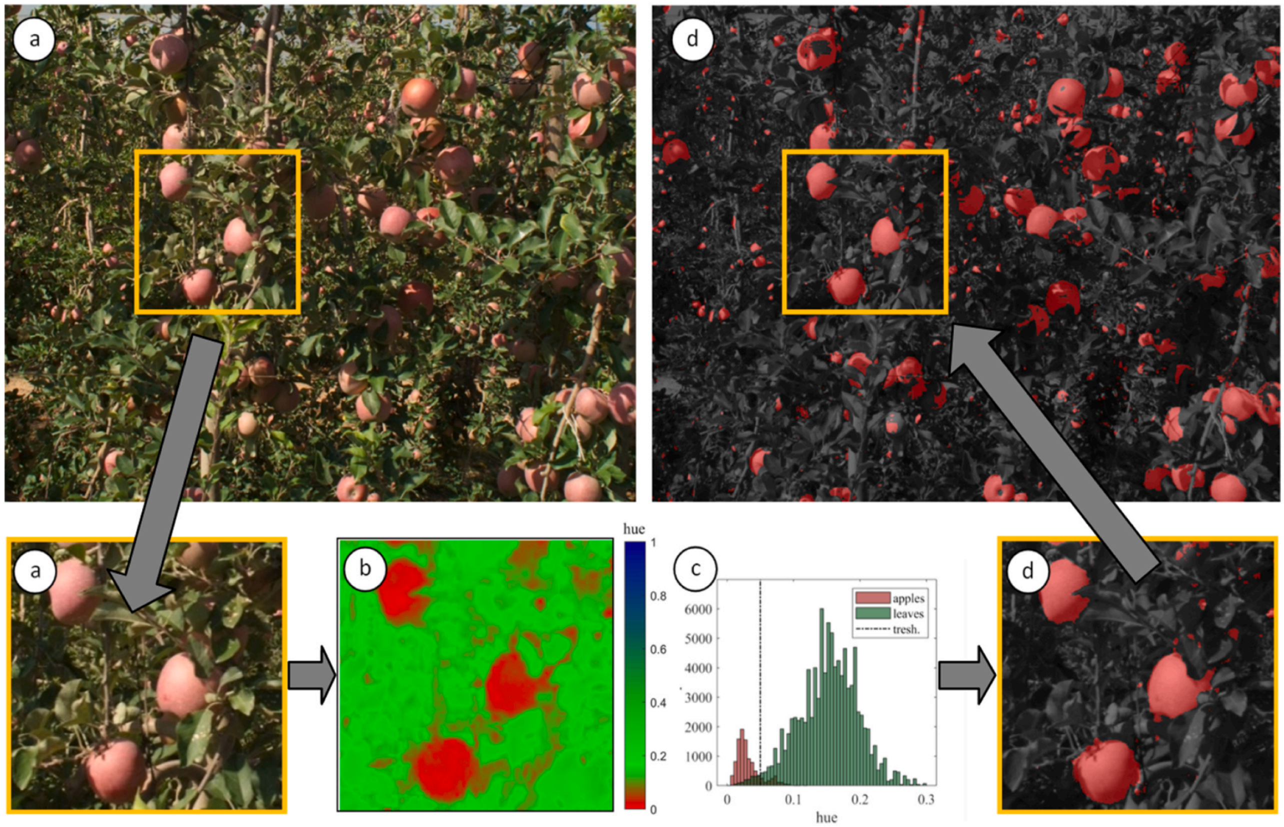

Fig. 3. Example of colour conversion and intensity thresholding applied to fruit segmentation. (a) RGB image. (b) Image converted to the HSV colour space. Colour scale corresponds to the hue (H) value. (c) Histogram of hue values for apple (red) and background (green) pixels. The vertical dash-dotted line corresponds to the selected threshold. (d) Segmented apples after applying the hue threshold.

(Fig. 3). This method aims to binarize data by setting a numerical threshold into a discriminative feature that describes the object of interest. A common feature used has been the colour (Maldonado and Barbosa, 2016; Qian et al., 2018), although other types of data have been considered such as the area of pixels (Liu et al., 2019), the depth (Tao and Zhou, 2017), and the temperature using thermal imaging (Pedraza et al., 2019). Fruit reflectance and geometric features are also applied to define the fruit ROI (Gene-Mola et al., 2019a; Tsoulias et al., 2020).

Another possibility for region selection is to apply machine learning classifiers. Classification methods allow objects within a space to be distinguished by specific features. The most used classifiers include the unsupervised k-means algorithm (Wang et al., 2018a; Shi et al., 2020), and different supervised methods such as Bayesian (Lin et al., 2019), the k-nearest neighbours (KNN) (Qureshi et al., 2017) and support vector machine (SVM) procedures (Zhang et al., 2020). With regard to 3D point clouds, there are many methods that allow their segmentation (Grilli et al., 2017). Two of the most commonly used methods in fruit detection are Euclidean clustering (Nguyen et al., 2016) and density-based spatial clustering of applications with noise (DBSCAN) (Eizentals and Oka, 2016).

Region description is a step prior to detection and recognition in which the identified regions are described with features to refine the selection according to their appearance and geometry. The result can be a multi-dimensional numeric vector or a set of pixels or point cloud with candidate labels (fruit, background, etc.). Colour, shape, texture and multiple features are used to describe regions.

Colour-based radiometric features mostly comprise the statistical data about channels in colour spaces. For example, in Syal et al. (2014) features were extracted by using the mean colour of the ‘a′ and ‘b′ components in L*a*b space. Using light detection and ranging (LiDAR) at 660 nm provides information on the chlorophyll content of fruit, which can support the segmentation (Tsoulias et al., 2023).

Shape-based techniques are useful in cases where the fruit and the background have the same colour. These are ideal for detecting fruits whose shape differs from leaves and branches. In fruit detection, the most relevant shape-based techniques are the Hough transform (HT) and the histogram of oriented gradients (HOG). One of the main variants of HT is the circular Hough transform (CHT), which has been widely used to locate spherical fruit in orchards (Wang et al., 2018a; Chen et al., 2021). Other shape-based techniques include analysis of convexity (Kelman and Linker, 2014), three-point circle fitting (Sun et al., 2019), or random sample consensus (RANSAC) (Nguyen et al., 2016).

Textures are small patterns with fluctuations of the intensity between groups of neighbouring pixels. Texture-based methods are used to detect fruit of the same colour as the background, taking advantage of the invariant characteristics of textures to changes in lighting and the smoother surfaces of the fruits. Among the texture-based methods used in fruit detection can be cited oriented FAST and rotated BRIEF (ORB), speeded-up robust features (SURF), scale-invariant feature transform (SIFT) and local binary patterns (LBP) (Chaivivatrakul and Dailey, 2014; Wang et al., 2018a). Multiple feature combinations have been preferred by some authors to improve the detection success rate. Li et al. (2016) and Qureshi et al. (2017) present examples of this approach for the detection of immature citrus and mango fruit, respectively.

2.3. Detection and recognition

Once a set of candidate regions and a list of features that describe each of these regions have been obtained, the next step is to classify them into true (fruit) or false (background) detection. For this purpose, a variety of classifiers have been used such as SVM (Gené-Mola et al., 2020a; Wu et al., 2020), KNN (Li et al., 2016; Nyarko et al., 2018), Adaboost (Wang et al., 2018a; Mekhalfi et al., 2020), random forest (Yu et al., 2021), backpropagation neural network (BPNN) (Cheng et al., 2017), and Gaussian mixture model (GMM) (Roy et al., 2019).

A common source of error in fruit counting systems is the presence of multiple detections (more than one detection of a single fruit), which results in an over-counting error. To prevent multiple detections, some authors have applied the non-maximum suppression (NMS) algorithm, which consists of discarding the overlapped detections with non-maximum confidence values (Yu et al., 2021).

Handcrafted detection algorithms are still applied due to their lower use of resources (computer power and memory) (Zhang et al., 2020) and the relatively minor amount of data required to train them compared to blackbox methods such as deep neural networks. They are used in cases where the object has a high contrast with the background and can be easily distinguished. These methods have also been implemented on platforms where high computational power is not available (Fu et al., 2018; Habib et al., 2020). The disadvantage of handcrafted detection algorithms is the lack of generalization in detecting fruit in other acquisition conditions for which specific algorithms were not designed. In addition to this, the functions need to be optimized manually, which is time consuming (Farjon et al., 2020).

3. Fruit detection based on deep learning

3.1. Background

Deep learning has meant a breakthrough in computer vision and, consequently, in fruit detection. Koirala et al. (2019a) reviewed the use of deep neural networks for fruit detection. Prior to 19/01/2019 they found a total of 9 papers in the Scopus data base (www.scopus.com) using the keywords: 'deep' + 'learning' + 'fruit' + 'detection'. Four years later (on 31/07/2023), a total of 347 articles were found in Scopus on the same search basis, showing that the use of deep learning for fruit detection is a highly active research field with a rapid increase in scientific production (Fig. 4).

The most commonly used deep neural networks in computer vision are the so-called convolutional neural networks (CNN), where the neurons of each unit are organized in three-dimensional matrices (feature maps). Consecutive units are connected by means of convolutional layers, pooling layers and fully connected layers used to process the input data and extract features at different scales (LeCun et al., 2015).

CNNs have demonstrated a level of performance similar to that of the human eye in tasks such as image classification, object detection, and semantic and instance segmentation (Voulodimos et al., 2018). Image classification refers to the problem of classifying the whole image in a specific class, for instance an image of a fruit in the fruit class or variety (Fig. 5a). Object detection refers to the problem of identifying the regions (bounding boxes) that contain the objects of interest, for instance locating the fruit that appear in an image (Fig. 5b). Semantic segmentation refers to the problem of classifying each pixel in the image, for instance labelling each pixel as fruit, trunk, branch or background (Fig. 5c). Finally, instance segmentation combines object detection and semantic segmentation: first objects of interest are located in the image and then the objects are segmented, identifying which pixels of the image correspond to each detected object (Fig. 5d).

A comparative table of the results reported in different deep learning-based fruit detection works is shown in Table 1. The F1-score metric was selected as it is the most commonly used in fruit detection papers. Other metrics such as average precision (AP) or accuracy (ACC) are reported when the F1-score results were not available. It should be noted that the reported results depend not only on the CNN structure and its parameters but also on the difficulty of the dataset. Thus, works assessing different structures with different datasets are not comparable.

3.2. Fruit detection using image classification CNNs

The structure of image classification CNNs is based on an input layer (the image to classify) connected with a group of convolutional layers that act as feature extractors (feature maps), ending with a group of fully connected layers that act as classifiers. The convolutional layers encode image features into more discriminative features by convolving the feature maps with filters (learned weights). Finally, fully connected layers are placed at the end of the CNN to classify feature maps in one of the classes of the output layer.

The use of classification CNNs for fruit counting is marginal because these architectures classify the entire image in a unique class and do not locate the objects inside images. Wang et al. (2021) proposed a modified version of the VGG16 network (Simonyan and Zisserman, 2014) to count the number of apple flowers in an image. The total number of flowers in the image was considered the image class, and the network was trained to directly estimate the number of flowers visible in the image, without locating them. A similar approach was used in Bhattarai and Karkee (2022), who modified the classification block of the VGG16 architecture to regress the number of flowers or fruits in apple tree images.

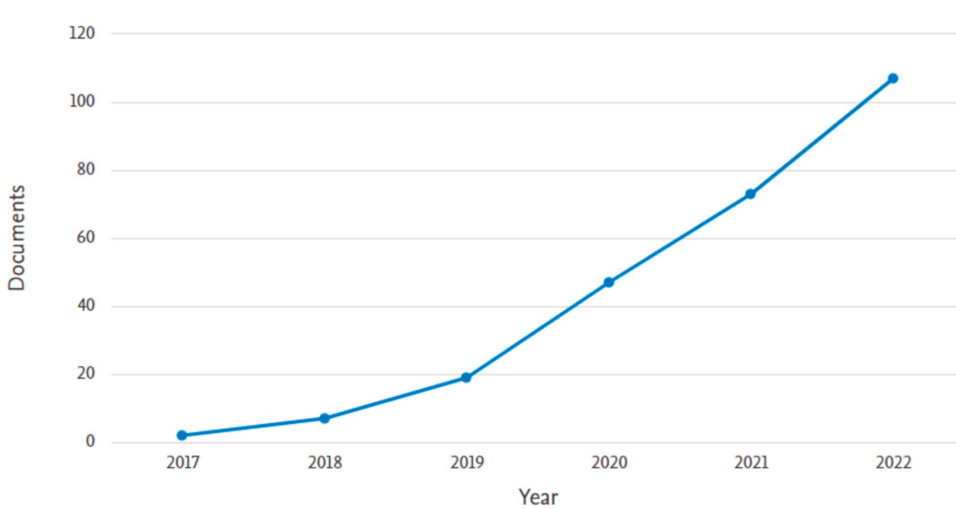

Fig. 4. Number of articles (conference proceedings not included) published per year in Scopus data-base containing keywords 'deep' + 'learning' + 'fruit' + 'detection'.

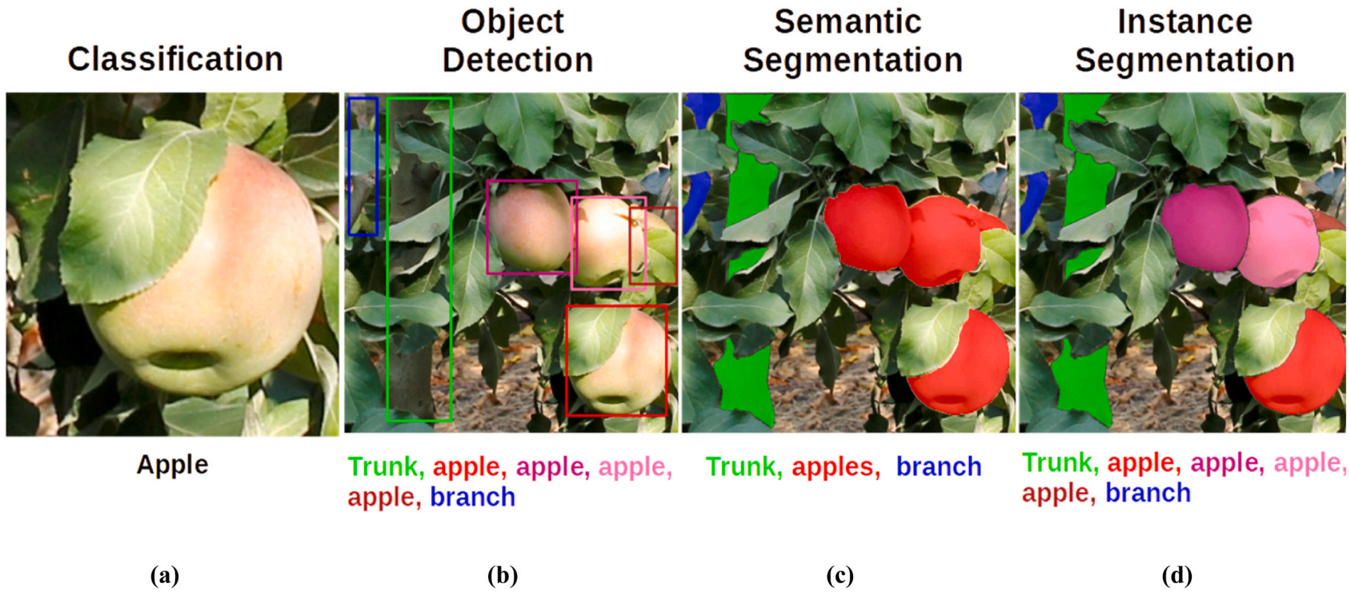

Fig. 5. Common computer vision tasks. Examples of apples on trees in field conditions: (a) classification, (b) object detection, (c) semantic segmentation, (d) instance segmentation.

3.3. Fruit detection using object detection CNNs

Object detection CNNs are formed with two main structures: backbone and head. The backbone usually uses the first layers of an image classification CNN as feature extractor to encode the data into feature maps. Then, the head structure uses the feature maps provided by the backbone to predict the object locations and their class. Depending on the head structure, object detection networks can be classified as one- or two-stage networks.

The first CNNs used for fruit detection were two-stage networks type, with a structure based on two main modules: (1) a region proposal module used to propose ROIs likely to contain a fruit; (2) a classification branch used to classify the proposed regions into fruit or background and refine the detection bounding box. The most commonly used two-stage CNN for fruit detection is the Faster-RCNN (Ren et al., 2017), which has been used to detect apples (Apolo-Apolo et al., 2020b; Kang and Chen, 2020; Tian et al., 2019), oranges (Apolo-Apolo et al., 2020a, Biffi et al., 2021), mangos (Bargoti and Underwood, 2017a; Koirala et al., 2019b), kiwis (Gan et al., 2018), and strawberries (Chen et al., 2019), among others.

One-stage CNNs (or single shot detectors) simultaneously predict object class and bounding box without the need of a region proposal branch. Single shot detectors (SSD) used for fruit detection include the single shot multibox detector (Liu et al., 2016) and the You Only Look Once (YOLO) (Redmon and Farhadi, 2018) and its variants v2, v3, v4 and v5. The SSD was used in Vasconez et al. (2020) with the MobileNet backbone for detection of apples, avocadoes and lemons. YOLOv2, YOLOv3 and YOLOv4 were used with DarkNet-19 and DarkNet-53 backbones, respectively, in different fruit detection works for apples, pears, kiwis, mangoes, bananas and grapes (Bresilla et al., 2019; Fu et al., 2021, 2022; Koirala et al., 2019b; Santos et al., 2020; Tian et al., 2019).

To enhance the performance of fruit detection systems, some authors have proposed the use of multi-modal deep neural networks to fuse different image modalities such as colour (RGB), depth or infrared (IR) intensity. Using a red-green-blue-depth (RGB-D) camera, Gené-Mola et al. (2019c) showed an increase of 4.46 % in the F1-score when combining colour, range-corrected IR intensity and depth images for apple detection with Faster-RCNN. Similarly, colour and depth images were combined for passion fruit detection in Tu et al. (2020), and colour and thermal images were combined in Gan et al. (2018) for orange detection. More recently, Sun et al. (2022) developed a new multi-modal network termed noise-tolerant feature fusion network (NT-FFN) which merged colour and depth features by means of attention modules, resulting in a better fruit detection performance: from F1-score of 0.910 (using RGB) to 0.934 (fusing RGB-D through NT-FFN).

The introduction of edge computing applications and the need of deploying real-time fruit detection in embedded computers, such as NVIDA Jetson products, has shifted the attention of researchers to the development of smaller object detection CNNs (Roy and Bhaduri, 2022; Zhang et al., 2021b; Zhang et al., 2022a). In consequence, many fruit detection papers published during the last two years are focused on achieving faster inference speeds in low power devices by means of light-weight and fast CNNs such as YOLOv5s and other YOLO-based tiny variants (Gai et al., 2021; Wang and He, 2021; Yan et al., 2021).

3.4. Fruit detection using semantic and instance segmentation CNNs

The fully convolutional network (FCN) (Long et al., 2015) is one of the most used architectures for fruit segmentation. FCN uses the first convolutional layers of CNN image classification as a backbone to encode data in discriminative feature maps. Then, the last feature map from the backbone is up-scaled by means of skip connections that combine information from shallower layers (finer but less discriminative) and deeper layers (coarser but more discriminative). FCN was used to detect kiwi fruits (Williams et al., 2019), oranges and apples (Chen et al., 2017; Liu et al., 2018).

Other authors have opted to develop new architectures specifically designed for fruit segmentation. The MangoNet architecture, developed by Kestur et al. (2019), replaced the last 3 convolution layers of FCN with a single convolution layer, obtaining a similar performance to that of FCN but reducing network complexity. Wang et al. (2020) developed a new architecture to adapt ResNet-50 for apple edge segmentation. Bargoti and Underwood (2017b) proposed a sliding window approach, which classified each pixel by means of a self-developed multilayer perceptron (MLP) and a CNN.

A disadvantage of semantic segmentation CNNs is that it is not possible to directly count fruits from a segmented image because all fruits appearing in an image are segmented under the same class. Instance segmentation CNNs overcame this issue by combining object detection and semantic segmentation. The most popular instance segmentation CNN used for fruit detection is Mask-RCNN (He et al., 2017),

Table 1
A comparative table of results reported in different deep learning-based fruit detection works. Results are reported in terms of F1-score and processing time per image. Accuracy (ACC), Pearson's R value and Average Precision (AP) are provided when the F1-score value is not available.

Approach	Data type	Method	Backbones	Crop	F1-score	Processing time (seconds per image)	Reference
Image classification	RGB	ResNet50	N/A	Apples	0.978 (ACC)	N/M	(Häni et al., 2020a)
	RGB	CountNet	VGG-16	Apples	0.962 (R)	N/M	(Bhattarai and Karkee, 2022)
Object detection	RGB	Faster-RCNN	VGG-16	Mangoes	0.881	N/M	(Stein et al., 2016)
	RGB	Faster-RCNN	VGG-16	Mangoes	0.908	0.13	(Bargoti and Underwood, 2017a)
	RGB	CNN + WS	Self-developed	Apples	0.861	0.24	(Bargoti and Underwood, 2017b)
	RGB	Faster-RCNN	ResNet-50	Strawberries	0.842	0.113	(Chen et al., 2019)
	RGB	MangoYOLO(pt)	N/M	Mangoes	0.968	0.015	(Koirala et al., 2019b)
	RGB	YOLOv2-M1	Darknet-19	Apples, pears	0.79	0.05	(Bresilla et al., 2019)
	RGB	YOLOv3dense	Darknet-53	Apples	0.864	0.304	(Tian et al., 2019)
	RGB	Faster-RCNN	Inception v2	Avocados	0.84 (AP)	0.217	(Vasconez et al., 2020)
	RGB	Faster-RCNN	Inception v2	Cherries	0.733	N/M	(Villacrés and Auat Cheein, 2020)
	RGB	Faster-RCNN	ResNet v2 Atrous	Apples	0.919	N/M	(Apolo-Apolo et al., 2020b)
	RGB	DY3TNet	Darknet-53	Kiwis	0.903 (AP)	0.034	(Fu et al., 2021)
	RGB	ATSS	ResNet50	Apples	0.925 (AP)	N/M	(Biffi et al., 2021)
	RGB	YOLOv4dense	DenseNet	Cherries	0.947	0.467	(Gai et al., 2021)
	RGB	YOLOv5s-pruned	Modified CSPDarknet	Apple fruitlets	0.915	0.008	(Wang and He, 2021)
	RGB	YOLOv5s-attention	Modified CSPDarknet	Apples	0.875	0.015	(Yan et al., 2021)
	RGB	YOLOv4	CSPDarknet-53	Bananas	0.941	0.045	(Fu et al., 2022)
	RGB+Thermal	Faster-RCNN + CHT	VGG-16	Oranges	0.929	N/M	(Gan et al., 2018)
	RGB+ NIR_C+Depth	Faster-RCNN	VGG-16	Apples	0.898	0.074	(Gené-Mola et al., 2019c)
	RGB+Depth	MS-FRCNN	ResNet101	Passion fruits	0.946	0.175	(Tu et al., 2020)
	RGB+Depth	NT-FFN	Self-developed	Citrus	0.934	0.026	(Sun et al., 2022)
Semantic segmentation	HyperSpectral	Hyperspectral CNN	N/A	Mangoes	0.989	N/M	(Wendel et al., 2018)
	RGB	FCN-8S	VGG-16	Kiwis	0.878	0.25	(Williams et al., 2019)
	RGB	MangoNet+CCL	N/A	Mangoes	0.844	N/M	(Kestur et al., 2019)
Fruit edge segmentation	RGB	Self-developed	ResNet 50	Apples	0.531	0.075	(Wang et al., 2020)
Instance segmentation	RGB	Mask-RCNN	ResNet50 +FPN	Strawberries	0.956	0.125	(Yu et al., 2019)
	RGB	Mask-RCNN	ResNet101	Grapes	0.847	N/M	(Santos et al., 2020)
	RGB	Mask-RCNN	ResNet101-FPN	Apples	0.858	0.15	(Gené-Mola et al., 2020d)
	RGB	Mask-RCNN-suppression	ResNet101-FPN	Apples	0.905	0.25	(Chu et al., 2021)
	RGB	Mask-RCNN-attention	ResNet50 +FPN	Apples	0.964	0.25	(Wang and He, 2022)
Multitask	RGB	DaSNet-v1	ResNet-101	Apples	0.832	0.072	(Kang and Chen, 2019)
	RGB	DaSNet-v2	Darknet-53	Apples	0.873	0.070	(Kang and Chen, 2020)
Point cloud segmentation	Point cloud	PointNet	PointNet	Grapes	0.91 (ACC)	N/A	(Kurtser et al., 2020a)
	Point cloud	LFPNet	PointNet	Apples, pears, grapes	0.802 (ACC)	N/A	(Yu et al., 2022a)
	Point cloud	Mask-RCNN	F-PointNet	Pomegranates	0.845	N/A	(Yu et al., 2022b)

N/A = not applicable. N/M = not mentioned. NIR_C = Near-infrared (range-corrected intensity).

which is an extension of Faster-RCNN that includes a segmentation branch to mask detected objects. Mask-RCNN was used with VGG-19 backbone for apple detection (Kang and Chen, 2020), with ResNet-50 backbone for strawberry detection (Yu et al., 2019) and with ResNet-101 backbone for apple (Gené-Mola et al., 2020d) and grape detection (Santos et al., 2020). More recently, some authors have proposed modifications in the Mask-RCNN architecture in order to achieve a better fruit detection performance (Chu et al., 2021; Wang and He, 2022) or a faster inference speed (Jia et al., 2021).

Kang and Chen (2019) developed a multi-task architecture termed "Detection and Segmentation Network" (DaSNetv1). This architecture combines a segmentation branch used to segment apples, trunks and branches, and a detection branch to locate fruits. This network was specifically designed for harvesting robots, allowing the detection of fruits and obstacles (branches) in a single network. Later, the same authors presented an improved version (DaSNetv2) (Kang and Chen, 2020) which replaced the previous detection branch with an instance segmentation architecture, allowing detection and instance segmentation to be performed on fruits, and semantic segmentation on branches in one step.

So far, the reviewed architectures were designed for working with image data. However, the evolution of photonics has led to the deployment of 3D sensors for robotic applications and, thus, to an increasing interest in using deep learning architectures to work with 3D data such as point clouds. Kurtser et al. (2020a) proposed the use of PointNet (Qi et al., 2017) for grape segmentation in 3D point clouds acquired with RGB-D sensors. The best results were obtained when combining RGB and XYZ data, reporting an average accuracy of 65 % in field conditions. Inspired by PointNet, Yu et al. (2022a) developed a new lightweight architecture for apple, pear and lemon point cloud segmentation that reported a mean accuracy of 80.2 %. Recently, Yu et al. (2022b) have tested the F-PointNet (Qi et al., 2018), a variant of the PointNet in which the frustrum between the camera shooting and the detected fruits is used for point cloud segmentation. An F1-score of 0.845 and AP score of 0.952 were obtained for mature pomegranate fruit detection.

3.5. Datasets for training fruit detection CNNs

The main disadvantage of using deep learning methods is the high amount of annotated data required for training models. The existence of large datasets such as ImageNet (Deng et al., 2009), Pascal VOC (Everingham et al., 2010) or COCO (Lin et al., 2014) enables CNN pre-training with publicly available data and fine-tuning of the network for fruit detection with new annotated images, reducing significantly the amount of images required to train the CNN. Nevertheless, the annotation of new data continues to be an intensive time-consuming task (Koirala et al., 2019b).

Some authors have analysed the correlation between dataset size and CNN performance. Tian et al. (2019) reported that performance improved with the number of fruit training images, reaching convergence around 3000 images. A similar analysis was performed by Koirala et al. (2019b) and Bargoti and Underwood (2017a), who reached convergence at around 400 training images and 500,000 annotated instances, while Wang et al. (2022) showed that 2500 annotated objects were sufficient for single-class fruit training.

Lu and Young (2020) reviewed publicly available datasets that could be of interest for training future fruit detection CNNs. Table 2 provides details of the ten datasets included in Lu and Young (2020) and seven additional datasets for fruit detection, classification and segmentation.

When a CNN model trained with a given dataset do not generalize well with new data, semi-automatic labelling is an option for better annotation efficiency. Semi-automatic labelling consists of automatically detecting fruits in new images with a pre-trained network (or an unsupervised method) and generating the ground truth by manually correcting the detections (dos Santos Ferreira et al., 2019). Another option is to use weakly supervised methods: Bellocchio et al. (2019, 2020) proposed a deep learning approach that only required a simple image binary labelling; Biffi et al. (2021) proposed a deep learning approach based on an adaptive training sample selection (ATSS) method that only requires annotation of the centre point of the objects; while Bhattarai and Karkee (2022) proposed a regression network (CountNet) which only requires the ground truth of the number of fruits per image.

When the number of empirical data is limited, different strategies have been applied to increase the capability of the network to generalize. Data augmentation techniques use annotated images to create new images by means of image transformations such as image flipping, rotation, and colour perturbations. This is a common practice employed in fruit detection works (Koirala et al., 2019a). Another option is to use synthetic data. Bresilla et al. (2019) generated synthetic images with random elliptic dark-green shapes (leaves) and light-green and light-red circles (fruit). More recently, the introduction of cycle generative adversarial networks has shown an improvement of the realism of synthetic images (Zhang et al., 2021a), increasing the performance of trained networks by more than 8 % (Barth et al., 2020).

3.6. Fruit tracking and counting

A common source of error when estimating fruit load is the double counting of fruit. The easiest method to prevent this is to acquire data

Table 2
Publicly available datasets for fruit detection. Data can be accessed by clicking on the corresponding title (highlighted in blue).

Title	Year	Image type	Images * (instances)	Image size * *	Annotation type	Crops	References
ACFR-orchard fruit dataset	2016	RGB	3704	308 × 202/ 500 × 500	Bounding boxes	Almonds, apples, mangoes	(Bargoti and Underwood, 2017a)
DeepFruits	2016	RGB	586	Different sizes	Bounding boxes	7 different fruits	(Sa et al., 2016)
MangoNet semantic dataset	2018	RGB	49	4000 × 3000	Segmentation masks	Mangoes	(Kestur et al., 2019)
Date fruit dataset	2019	RGB Videos	8079 15	Different sizes	Length, weight, maturity	Dates	(Altaheri et al., 2019)
Embrapa WGISD	2019	RGB	300 (4432)	2048 × 1365/ 5184 × 3456	Instance segmentation	Grapes	(Santos et al., 2020)
ISARLab_counting_dataset	2019	RGB	1560	300 × 300/ 606 × 403	Fruit number per image	Almonds, olives, apples	(Bellocchio et al., 2019)
Kfuji-RGB-DS dataset	2019	RGB+Depth+NIR	967 (12,839)	548 × 373	Bounding boxes	Apples	(Gené-Mola et al., 2019b)
MangoYOLO data set	2019	RGB	1730	612 × 512	Bounding boxes	Mangoes	(Koirala et al., 2019b)
MinneApple	2019	RGB	1000 (41,000)	1280 × 720	Instance segmentation	Apples	(Häni et al., 2020b)
WSU apple dataset	2019	RGB	2298	Different sizes	Bounding boxes	Apples	(Bhusal et al., 2019)
Apple_detect dataset	2020	RGB	5969	1024 × 1024	Apple centre point	Apples	(Biffi et al., 2021)
FruitsGB: Top Indian fruits with quality	2020	RGB	12,000	256 × 256	Quality label	6 different fruits	(Meshram et al., 2020)
Fuji-SfM dataset	2020	RGB Point cloud	288 (1749) 1 (1455)	1024 × 1024 10.5 Mpts	Segmentation masks 3D bounding boxes	Apples	(Gené-Mola et al., 2020e)
LFuji-air dataset	2020	Point cloud	88 (1444)	235 kpts	3D bounding boxes	Apples	(Gené-Mola et al., 2020b)
Scifresh-apple-RGB-images	2020	RGB	800	1920 × 1080	Bounding boxes	Apples	(Gao et al., 2020)
Mango fruit on tree image collection	2021	RGB	250	4752 × 3168	Fruit number per image	Mangoes	(Walsh et al., 2021)
PFuji-Size dataset	2021	Point cloud	4 (615)	9.1 Mpts	3D instance segmentation + fruit centre location + diameters	Apples	(Gené-Mola et al., 2021b)

In the case of point cloud based datasets: *number of point clouds provided in the dataset, * * number of points per point cloud (average).

along the orchard without overlap between consecutive frames (Bargoti and Underwood, 2017a; Apolo-Apolo et al., 2020a). However, since the ratio of visible to occluded fruit is not always constant, the use of multi-view approaches is sometimes required to increase fruit detectability (Hemming et al., 2014). Hence, to prevent double counting, fruit need to be tracked during scanning.

Two different strategies have been applied to track fruit across consecutive frames: video multi-object tracking (MOT) and the use of 3D data to locate the position of detections in the 3D space (Fig. 6). So far, the method most commonly used for video fruit tracking has been the Kalman filter (Anderson et al., 2021a; Itakura et al., 2021; Liu et al., 2018; Wang et al., 2019). Alternatively, Das et al. (2015) used optical flow, while Stein et al. (2016) used epipolar geometry by projecting the epipolar lines of the detected fruits centre onto consecutive frames. Recently, deep learning has demonstrated a good performance for solving MOT tasks (Dendorfer et al., 2020), being DeepSORT (Wojke et al., 2017) the deep learning-based tracking method most used for fruit counting in videos (Osman et al., 2021; Parico and Ahamed, 2021; Villacrés et al., 2023). A different approach was implemented in Roy and Isler (2016), who used calibrated cameras to register images through affine transformation. Similarly, Apolo-Apolo et al. (2020b) and Chen et al. (2019), applied affine transformations to build an orthomosaic of the entire orchard and subsequently detected fruits. Another method involves utilizing image stitching, as demonstrated by Zhang et al. (2022b), who used a SIFT-based image matching technique to form unique panoramic image of the captured fruit trees.

An alternative approach to reduce double-counting issues is the detection of fruit in the 3D space by means of RGB-D cameras, LiDAR sensors, or structure-from-motion (SfM). Wang et al. (2013) used a stereovision system synchronized with two global navigation satellite system (GNSS) receivers in order to transform apple locations into the global coordinate system. Then, fruit detected in consecutive frames closer than 0.16 m were automatically merged. Other works proposed the use of SfM to merge fruit detected from different camera positions (Gené-Mola et al., 2020d; Häni et al., 2020a; Liu et al., 2018, 2019; Santos et al., 2020). Taking advantage of the 3D data generated with SfM photogrammetry, fruits are previously detected in images and subsequently projected onto the 3D space for pair-wise association (Table 3).

4. Fruit size and maturity estimation

4.1. Size estimation from 2D images

This group includes the set of works carried out by Stajnko et al., where apple fruit diameters were estimated throughout their growing season using RGB (Stajnko and Čmelik, 2005; Stajnko et al., 2009) and thermal images (Stajnko et al., 2004). A high coefficient of determination was obtained when comparing the estimated fruit diameter growing curves with the actual ones (R^2 of 0.89 and 0.96 for RGB and thermal images, respectively). The tests with thermal cameras also showed that it is more difficult to detect the thermal gradient of the fruits inside the crown; this is because they heat up less than fruit located on the outside part. Likewise, Wang et al. (2020) used a spherical video camera for monitoring the apple growth from fruit thinning to their ripening. Estimates of the horizontal diameter of apples were made by applying ellipse and circular fitting methods and with the help of calibration balls. Ellipse fitting estimates yielded a mean average absolute error of 0.90 mm, much less than the 2.80 mm error obtained using a circular fitting. The size of citrus fruit was also estimated by Apolo-Apolo et al. (2020a), in this case using images taken from an unmanned aerial vehicle (UAV) and considering a wood ruler of known dimensions as the calibration object. Recently, Lu et al. (2022) proposed a near real-time apple fruit detection and sizing method from images taken by a low-cost smartphone in various growth stages. To estimate the fruit size, a red artificial apple was placed as a reference in the middle of the target area during the data collection stage. Estimated fruit sizes achieved R^2 values of 0.68 and 0.66 in fruit height and fruit width, respectively.

Another alternative is based on knowing the distance to the camera of each of the fruit that appear in the image. In their pioneering work, Regunathan and Lee (2005) combined colour images with distance information obtained with ultrasound sensors and, using trigonometry, estimated the dimensions of citrus fruit. In this line, Wang et al. (2017) used the images and depth data provided by an RGB-D camera to

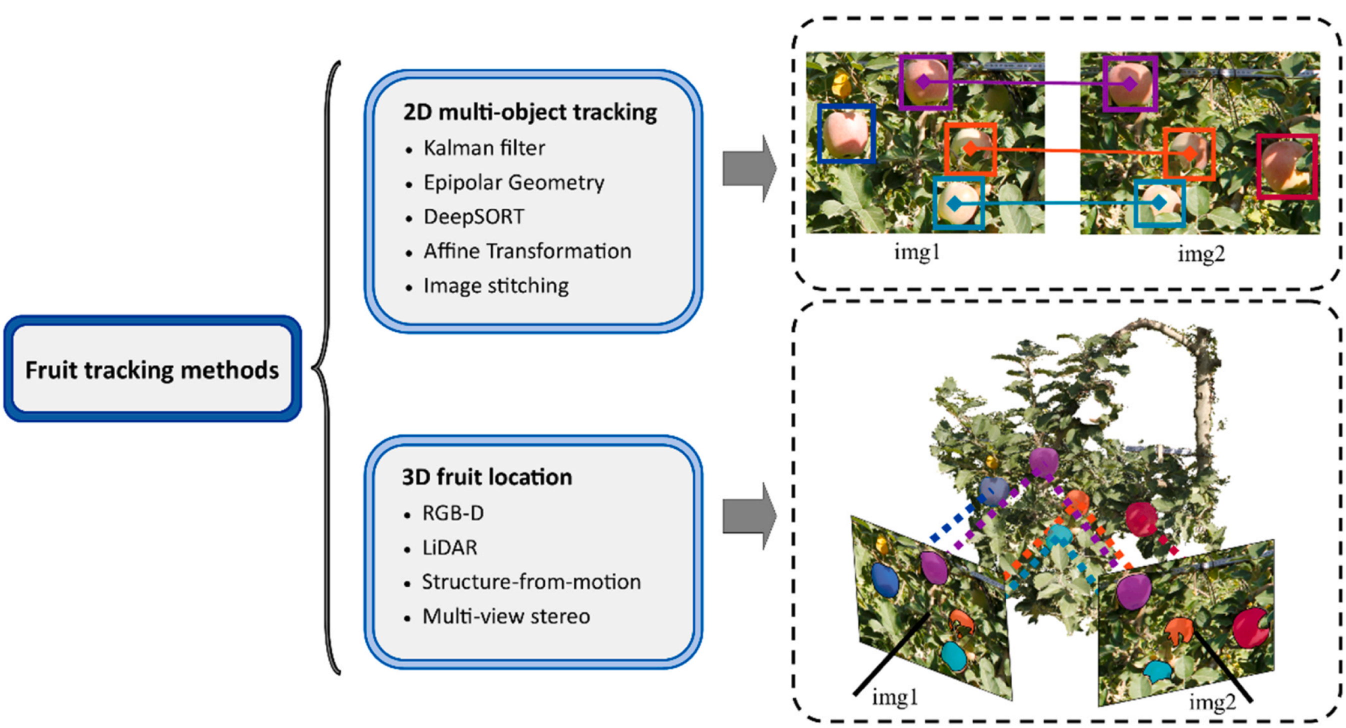

Fig. 6. Summary of methods used to prevent fruit double counting. Multi-object tracking (top) and 3D projection (bottom) procedures.

Table 3
A comparative table of results reported in different fruit counting works. Results are reported in terms of R^2.

Tracking method	Sensors	Fruit detection method	Backbone	Crops	R^2 *	Reference
Images without overlap	RGB	MLP + CHT	N/A	Apples	0.83	(Bargoti and Underwood, 2017b)
Epipolar geometry	RGB + LiDAR	Faster-RCNN	VGG-16	Mangoes	0.90	(Stein et al., 2016)
Kalman	RGB	SSD	Mobilenet	Avocado, apples, lemons	0.77	(Vasconez et al., 2020)
	RGB	MangoYolo	Not specified	Mangoes	0.62 (DR)	(Wang et al., 2019)
DeepSORT	RGB	YOLOv4	CSPDarknet53	Pears	0.755 (MOTA)	(Parico and Ahamed, 2021)
Orthomosaic	RGB	Faster-RCNN	Resnet V2 Atrous	Apples	0.80	(Apolo-Apolo et al., 2020b)
	RGB	Faster-RCNN	Resnet-50	Strawberries	0.84 (DR)	(Chen et al., 2019)
3D projection	RGB	HSV thresholding	N/A	Apples	0.12 (ADRE)	(Wang et al., 2013)
	RGB	Faster-RCNN	Not specified	Mangoes	0.78	(Liu et al., 2019)
	RGB	Mask-RCNN	ResNet-101-FPN	Apples	0.80	(Gené-Mola et al., 2020d)

* Average detection rate error (ADRE), detection rate (DR) and MOT accuracy (MOTA) are provided when the R^2 coefficient is not available.

estimate the size of mango fruits in trees by applying the thin lens theory. Root-mean-square errors (RMSE) of 4.9 and 4.3 mm were obtained in the fruit length and width estimates, respectively. Gongal et al. (2018) estimated apple sizes from distances provided by a time-of-flight (ToF) camera using a regression model that converts pixels (digital camera) to millimetres. The mean absolute percentage error (MAPE) was 15.2 %, lower than the 30.9 % obtained when the size was derived from the point clouds provided by a ToF camera. Another approach to estimate the fruit size from two images taken at different positions was presented by Rakun et al. (2019). This procedure uses image registration and similar triangles, known the distance between the two camera positions. Average diameter errors of 7 and 8 mm were obtained for peach and apple fruits, respectively.

4.2. Size estimation from 3D point clouds

As mentioned, fruit size estimation based on 2D images require the use of calibration targets or to merge the image data with ancillary distance information that adds complexity and computational costs to the processing. These limitations can be overcome using 3D sensing techniques (Rosell and Sanz, 2012; Gregorio and Llorens, 2021), such as LiDAR, structured-light, binocular stereo vision, multi-view stereo (MVS) or RGB-D cameras, among others, which allow the generation of three-dimensional reconstructions of the fruits.

Regarding LiDAR-based techniques, Méndez et al. (2019) used a 3D laser scanner with RGB data and applied the *k*-means algorithm to estimate the number and size of oranges. The computed diameters did not show significant differences in relation to those measured manually ($p = 0.35$). As the authors point out, this is a time-consuming method, but given its high accuracy it can serve as a reference for other faster and more economical methods. For their part, Tsoulias et al. (2020) used a mobile terrestrial LiDAR scanner to monitor apples at different growth stages. Fruit diameter was estimated from each point cluster identified as apple and the resulting R^2 with RMSE was 0.46 with 10.8 % and 0.67 with 7.7 %, 42 DAFB and at harvest, respectively.

Structured-light principle was used by Rist et al. (2019), who tested a hand-held high-resolution scanner for 3D phenotyping of grape bunches under field conditions. These are high cost, high precision devices with acquisition speeds of about 1 million points/s. The authors achieved R^2 values of 0.70–0.91 in the prediction of several phenotypic traits (number and diameter of grapes; bunch width, length and volume). The RMSE values were 13.51 and 19.24 mm for bunch width and length, respectively, and 28.09 mL for the volume.

Binocular stereo vision was applied in harvesting robots under outdoor conditions by Luo et al. (2016). The authors proposed a method to detect the cut-off point of the peduncle and estimate the volume of grape bunches. As a result of their work, they obtained errors of less than 17 mm and 19 mm in bunch height and diameter, respectively. Herrero-Huerta et al. (2015) applied MVS for vineyard phenotyping and determined the grape bunch volume using an automatic method that fit a convex hull to the point cloud and a semi-automatic method that generated a CAD model. In both methods, similar coefficients of determination were obtained (0.76 and 0.77) when comparing the estimates with the actual bunch volumes. MVS was also applied in vineyards by Rose et al. (2016), who determined berry diameter by fitting spheres to point clouds. Estimates were highly accurate with differences of about 2 mm with respect to manual measurements. Recent studies (Gené-Mola et al., 2021a; Grilli et al., 2021) have applied MVS and SfM to carry out in-field diameter estimation of apple fruit. Gené-Mola et al. (2021a) compared the performance of four different size estimation methods under several fruit visibility/occlusion levels (Fig. 7). The least squares method was concluded to be the most efficient in terms of computational cost, while the MAE ranged from 4.5 to 7.8 mm depending on the visibility. These errors were lower than those obtained with the largest segment method and similar to those obtained with the M-estimator sample and consensus (MSAC) method and template matching. For their part, Grilli et al. (2021) developed a procedure for on-tree automatic apple fruit counting and sizing using videos acquired with a smartphone. Apple size estimation was performed by fitting spheres (RANSAC method) on the point cloud.

RGB-D sensors have been applied in vineyard yield estimation by Hacking et al. (2019). In their study, RGB-D measurements were used to create one mesh per grape bunch and determine its volume and mass. Also in vineyards, Kurtser et al. (2020b) used point clouds generated by an RGB-D camera to detect the grape clusters. These authors proposed three methods based on fitting geometric shapes to estimate the grape cluster size, obtaining the best results using percentile bounding boxes. Also using RGB-D sensors, Yu et al. (2022b) performed 3D sphere fitting to estimate the position and size of pomegranate fruits. Estimates of the fruit radius presented an RMSE of 2.35 mm and R^2 of 0.826 when compared to the actual radius, while the position error was less than 5 mm.

As seen in this section, fruit size estimations have been carried out in a limited number of studies, many of them focusing on a few species. It is difficult to compare the performance of the different techniques due to the diversity of metrics used for their evaluation (Table 4). It is therefore advisable that future works include, at least, the mean absolute error (MAE) and the coefficient of determination (R^2) when comparing estimated and actual size values.

4.3. Advancing fruit maturity estimation

In addition to the fruit size, knowing their maturity is essential for proper crop load management as well as for subsequent postharvest processes. Although automatic methods for fruit maturity estimation are less developed than sizing methods, some pioneering works have been carried out. Since many fruit species exhibit specific change of shape during fruit development, the fruit maturity can be estimated by means of the shape of the singularized, segmented fruit data. In apple, the shape of fruit was modelled by means of statistical approach (Danckaers

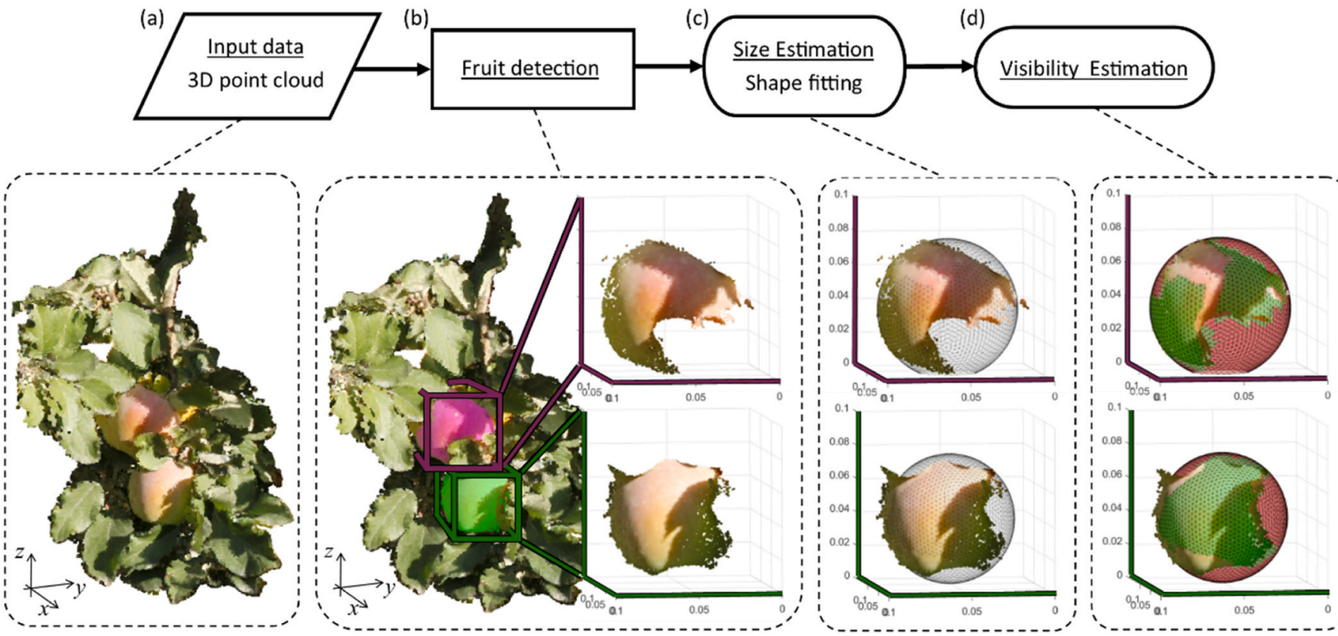

Fig. 7. Pipeline proposed by Gené-Mola et al. (2021a) to obtain in-field diameter estimation of apple fruit.

et al., 2017) or Fourier signature (Rogge et al., 2015; Tapia Zapata et al., 2022). Such approaches provide the next step of extracting information describing the maturity of fruit. In mango, the change around the shoulder of the fruit indicates maturity, which was analysed by means of RGB imaging (Sahu and Potdar, 2017).

Beside the shape of fruit also the pigment content and distribution provide information on the fruit maturity. The pigment contents have been addressed by means of colour analysis and spectral-optical data with enhanced resolution providing information on the reflectance intensity altered by absorption of pigments at their specific wavebands (Merzlyak et al., 2003; Walsh et al., 2020). Measurements were carried out in contact to the fruit to avoid stray light or with passive RGB sensors being not reliable in varying lighting conditions. However, the intensity measured by means of RGB-D and LiDAR sensors was employed previously to analyse the pigments of whole canopies employing LiDAR sensors emitting at 532 nm (green) or 660 nm (red), the latter to measure the leaf chlorophyll content (Eitel et al., 2010). Accordingly, 3D fruit segmentation and chlorophyll analysis were recently shown on apples in the orchard (Tsoulias et al., 2023) and banana fruit in postharvest (Saha and Zude-Sasse, 2022). Employing the return signal strength intensity of LiDAR sensor requests the radiometric calibration referencing the lowest and highest measurable intensity as well as curvature correction (Saha and Zude-Sasse, 2022). Classification of different measuring dates during fruit development were shown for apple as well as banana fruit, providing an interesting alternative to multispectral 2D readings.

5. Discussion and future trends

5.1. The importance of data acquisition

Sensors are the first stage of detection/counting and size estimation of fruit and, thus, are critical for the performance of the entire process. Limitations of up-to-date available sensors are transferred to the obtained measurements which feed up the subsequent applied algorithms, thus limiting their effectiveness. Changes in environmental lighting affects RGB and RGB-D cameras' performance (Gené-Mola et al., 2020c; Fu et al., 2020). In addition, structured light sensors usually fail to characterize the contours, what is especially problematic in small objects, such as fruit. Also, in contours of objects, LiDAR's mixed pixels phenomenon (Sanz-Cortiella et al., 2011) leads to distorted points clouds and filtering is often required.

Some more advanced and affordable new sensors, which are expected to achieve great advances in this field, are being already tested for fruit detection and sizing. Thus, multi-beam as well as solid state LiDAR sensors are a great step forward. There are also LiDAR + RGB systems that allow obtaining coloured point clouds, although the correct colour assignment needs further improvement, especially in the outlines of small objects. The possibility of using two LiDAR systems at different wavelengths to simultaneously determine the normalized difference vegetation index (NDVI) of fruit in addition to the fruit number and size has recently been demonstrated (Tsoulias et al., 2023). Some companies have developed systems that merge different sensing principles (sensor fusion) with AI and post processing algorithms in the same product (Zheng et al., 2021). Also, the use of smartphones' embedded sensors (GNSS, RGB cameras, LiDAR …) allows much more compact systems with post processing capabilities in the same hardware.

Some ideas for next steps towards fruit's detecting and sizing systems can be outlined, such as the combination of multiple sensors, with the same or different sensing principles (sensor fusion). In addition, the application of stereo vision, SfM and MVS principles to thermal cameras can also be assessed, in order to obtain point clouds in the thermal range of the electromagnetic spectrum. The same can be applied to multispectral (MS) cameras, which are commercially available at affordable prices: systems similar to those based on RGB cameras but with MS cameras can be developed, allowing to obtain 2D images of fruits/trees/crops including IR and, from these, obtain 2D images of vegetative indices (VI), such as NDVI or the normalized difference red edge index (NDRE), among others. Likewise, progress can be made in obtaining 3D point clouds in the IR and in these VI. In the case of MS cameras with multiple optical objectives (one for each spectral band), it is necessary to address the effect on the measurements of not having a single common optical objective.

Apart from the sensors themselves, it is also necessary to delve into the measurement system as a whole: vehicle, supports and optimal location of the sensors (Xie et al., 2022). The GNSS receiver system associated with the sensors is also important, since the more precise it is, the better, because it affects the accuracy of the measurements, especially with regard to the location of the fruit, which must allow mapping the size of fruit both in the tree and in the plot.

Table 4
Sensing techniques and methods for in-field fruit size estimation reporting coefficient of determination (R^2), absolute error (AE), mean absolute error (MAE), mean bias error (MBE), root mean square error (RMSE), mean average percentage error (MAPE).

Techniques	Size estimation method	Fruits	Size parameters	Performance	References
Binocular stereo vision	Calibration object. Regression model to predict pixel sizes.	Grape	Bunch diameter/ height	AE < 18.6 mm / 16.2 mm	(Luo et al., 2016)
High resolution 3D scanner	Sphere fitting.	Grape	Bunch length / width / volume	R^2 = 0.70 / 0.71 /0.91	(Rist et al., 2019)
LiDAR-based sensor	Sphere fitting.	Apple	Fruit diameter	R^2 = 0.38–0.95 RMSE: 4.1–15.8 % MAE= 3.5–12.4 mm MBE= −10.7 to 7.5 mm.	(Tsoulias et al., 2020)
MVS	Pixel conversion. Image registration.	Apple Peach	Fruit diameter Fruit diameter	MAE= 8 mm MAE= 7 mm[1]	(Rakun et al., 2019)
	Convex hull. CAD generation.	Grapes	Bunch volume	R^2 = 0.77 / 0.76 (convex hull / CAD model)	(Herrero-Huerta et al., 2015)
MVS, SfM	Sphere fitting.	Apple	Fruit diameter	R^2 = 0.91 RMSE = 5.1 mm MAE= 3.7 mm MBE = −1.9 mm MAPE= 5.9 %	(Gené-Mola et al., 2021a)
RGB + ultrasonic sensor	Distances with ultrasonic sensors. Pixel conversion.	Apple	Fruit diameter	RMSE= 0.4 cm.	(Regunathan and Lee, 2005)
RGB camera	Calibration object. Pixel conversion.	Avocado Mandarin Navel orange Apple Mango Mango	Fruit diameter Fruit diameter Fruit diameter Fruit diameter Fruit length Fruit width	RMSE= 3.4 mm RMSE= 3.8 mm RMSE= 2.4 mm RMSE= 2.0 mm RMSE= 5.3 mm / 5.5 mm (controlled/ ambient light) RMSE= 3.7 mm / 4.6 mm. (controlled/ ambient light)	(Wang et al., 2018b)
	Calibration object. Pixel conversion.	Apple	Fruit diameter growing curve	R^2 = 0.96	(Stajnko et al., 2009)
	Pixel conversion.	Apple	Fruit diameter	r = 0.55 – 0.80 (harvest stage)	(Stajnko and Čmelik, 2005)
	Calibration object. Pixel conversion. Ellipse/ circle fitting.	Apple	Fruit diameter	MAE= 0.90 mm / 2.80 mm (ellipse/ circle fitting)	(Wang et al., 2020)
	Calibration object. Pixel conversion.	Grapes	Berry diameter	R^2 = 0.88	(Roscher et al., 2014)
RGB camera + ToF	Calibration object. Regression model to predict pixel sizes.	Apple	Fruit diameter	MAPE: 15.2 % (RGB) / 30.9 % (ToF)	(Gongal et al., 2018)
RGB-D camera	Bounding box/ ellipsoid/ cylinder fitting.	Grapes	Bunch length / width	MAE= ~2.9 cm /~3.6 cm.	(Kurtser et al., 2020b)
	Pixel conversion units.	Mango	Fruit length / width	RMSE= 4.9 mm / 4.3 mm.	(Wang et al., 2017)
	Sphere fitting.	Pomegranate	Fruit radius	RMSE= 2.35 mm R^2 = 0.826	(Yu et al., 2022b)
Thermal cameras	Pixel conversion units.	Apple	Fruit diameter growing curve	R^2 = 0.89	(Stajnko et al., 2004)

Another aspect is the optimization of the resolution of the images and point clouds obtained by the sensors, so that they do not compromise the processing speed and allow progress towards real-time detection. Progresses must also be made in the implementation of systems that are increasingly plug and play, to facilitate their effective implementation in the sector, without the need to be an ICT expert. Finally, more studies are needed to know how external variables - apart from lighting conditions - such as temperature, dust, fog, vibrations etc., influence sensors' performance.

5.2. Fruit counting

Fruit growers still often use manual fruit counts on trees sampled within the plot to estimate orchard fruit yield. The fruit grower's experience helps to make this task more efficient, but since fruit counting is manual, it is always very laborious and expensive. AI application is expected to become the new paradigm in providing fruit growers with fast and reliable fruit counting methods for yield estimation. However, different strategies can be proposed, some of which still require new advances to be applied in a practical way. As demonstrated in this review, getting to automate fruit counting is not a simple matter, when occlusions, varying background, changing lighting exposure, unstructured canopies, and variable crop-load level are some of the challenges to face (Bhattarai and Karkee, 2022). If this were not enough, different steps must be addressed in the overall fruit counting process adding even more computational complexity. To give an example of the difficulty involved, it is known that only object-level annotation takes a substantial amount of manual annotation hours to create large labelled datasets (Pawara et al., 2020). The regression-based fruit counting approach has also been raised in some research (Bhattarai and Karkee, 2022; Pawara et al., 2020). In contrast to detection-based, annotation of only the total number of fruits at image level is used to train a neural network for counting. Thus, there is no need for explicit individual detection and localization resulting in a computationally simpler process. Another particularly interesting strategy is the one mentioned in Hobbs et al. (2021), where a deep learning-based density estimation approach is applied to count the number of flowering pineapple plants. By combining the latest advances in remote sensing and computer vision, counting is then affordable in orchards with high planting density. Indirect yield prediction, more than fruit count, has also been implemented for years by developing models that relate yield to features from environment (meteorological information) and/or features from canopy or tree physiology (management mode, plant growth state) (He et al., 2022).

The automatic fruit detection with computer vision algorithms is a key task for fruit counting systems. From 2016, the introduction of deep learning stablished object detection convolutional neural networks as the standard method to detect fruit in images, achieving F1-scores higher than 90 %, similar to the human eye. The tendency of the used CNN architectures is being the following. First (from 2016 to 2018, approximately), importance was given to improve accuracy. In this period, architectures such as Faster-RCNN demonstrated to be more accurate than the previous methods. Later (from 2018 to 2021, approximately), efforts were focused on improving efficiency and speed (Table 1). During this period, one-stage networks such as YOLO and its variance became the most popular, demonstrating real-time processing speeds and similar accuracy than the previous architectures. The current trend is to develop lightweight CNN to be implemented for edge computing purposes processed in embedded computers. It is expected that this will facilitate the deployment of commercial and affordable fruit counting devices (Zhang et al., 2021b).

Future research in fruit detection is expected to introduce emerging machine learning methods such as vision transformers (Carion et al., 2020), which are promising deep learning architectures based on attention mechanisms which could be more efficient and accurate than the popular CNNs. On the other hand, it is also expected an advance on point cloud-based object detection algorithms, which so far, have shown a lower performance compared with 2D algorithms based on CNNs. In this regard, further research should be done in order to test 3D machine learning methods that have not been applied for fruit detection such as the use of graph neural networks (Zhou et al., 2020) or PointPillars (Lang et al., 2019).

Based on the revised literature, authors consider the detection problem a quite mature problem at the level of computer vision. The handcrafted methods have been largely superseded by those based on deep learning. The lack of generalization in the detection of fruits together with variable conditions of the acquisition process (lighting changes, noise, background colours, etc.) are the main factors that directly affect handcrafted methods. Nevertheless, there are still environments (e.g. high-contrast fruit compared to the background) where methods based on handcrafted features could continue to be advantageous given their low computational cost. In order to advance to the development and deployment of commercial devices, future works should apply and evaluate the methods for in-field counting. Having high detection rates in the images does not ensure a high performance of these systems for yield estimation and mapping, since there are other factors that affect the systems performance such as the structure of trees, the amount of fruit occlusions, the use of multi-view methods, the strategies to prevent fruit double counting (such fruit tracking), etc. During the literature review we found that there are many works that evaluate the detection performance in images, but few works comparing the number of fruits detected in the images with respect to the actual number of fruits on trees or orchards. In addition, very few fruit counting and mapping methods were evaluated at different growth stages and different scanning conditions to ensure that systems generalize well at different environments. Thus, further efforts should be done to confront the challenges not related with the detection and evaluate the generalization of the models on larger datasets including different orchards scanned at different conditions.

Finally, authors involved in future research should also consider making publicly available the codes and the datasets with detailed explanations about how to implement and use them. This will facilitate that the scientific community advances efficiently and collaboratively. In other research fields such as Computer Science it is a common practice to make codes and data available. This, for example, explains the rapid advances in deep learning during the last decade. However, in the field of fruit detection and counting there is still a reluctance for the open science, which makes difficult to reproduce the methods and makes it difficult collaborative and additive research. Counting fruits is a task in which AI has allowed great advances. But, thinking of applying sensors and processes punctually within the plots, the combined use of AI together with efficient sampling cannot be ruled out, this being a still pending issue.

5.3. Fruit sizing and characterization

The fruit sizing task has not received as much attention as fruit detection, but several advances have been achieved during the last decade. Most of the revised works measure the fruit size in pixels in images and then apply a conversion from pixels to millimeters. Methods based on 2D images require the usage of calibration targets placed at the same distance to the cameras than fruit, which limits the efficiency of the data acquisition process. However, more advanced methods are based on 3D data, which can directly measure the fruit size in millimeters, or on RGB-D data, which allow the conversion from pixels to millimeters by applying the pinhole camera model.

So far, the dimensions of fruit at advanced ripening stages have been estimated, but there is also an interest in earlier maturity stages. For instance, the measurement of apple fruitlets is of interest for precisely adjusting the dose when applying chemical thinning. Thus, while accurate fruit sizing results have been reported in the revised literature, further research should be done for measuring young fruit to take appropriate actions in crop load management.

One of the major challenges when measuring fruit size with sensors is the presence of fruit occlusions. Although a high percentage of fruits are partially occluded, some fruit sizing works found in the literature limit the evaluation of their methods on fully visible fruits. From the authors' opinion, to transfer the fruit sizing research methods needs to deal with occlusions. Consequently, future works should provide fruit sizing results at different degrees of occlusion. In addition, further research should be carried out to automatically estimate the percentage of visibility of detected fruits, which would allow to identify the most occluded detections and limit the measurement to the most visible fruits, which are likely to be better measured.

The authors consider that future works will involve the development of methods capable of real-time monitoring the fruit temperature (sunburn risk), estimating fruit maturity (Section 4.3), early detecting its defects and diseases, etc. Ultimately, current size estimation methods should evolve towards fruit characterization methods to allow a more complete knowledge of the different variables that affect fruit growth.

5.4. Opportunities for research transfer

Considering in-field fruit size estimation (necessary to monitor growth and estimate fruit weight), new research and commercial opportunities are emerging with the priority of developing robust and low-cost systems, and also under the premise of having to process large amounts of data when applied to large farms with large number of trees. Currently, there are a few companies that detect and count blossom and fruits and, in some cases, estimate fruit dimensions to estimate yield (Anderson et al., 2021b). Fruit detection and sizing would unlock the possibility to estimate per tree crop load and adjust the number of fruits on a tree and quality bases. It is well-known that crop load influences fruit quality (Serra et al., 2016; Embree et al., 2007). Together with crop load, irrigations strategies but also fruit location in the canopy, mainly according to height, are also affecting quality parameters (Alcobendas et al., 2013). In-field fruit location and sizing systems allow fruits to be georeferenced on a per tree basis and also to register their position in the canopy. An early detection and sizing solution would allow the farmers to apply thinning strategies within the same season considering the number of fruitlets per tree or even per branch or according to height. Several measurements along the season would provide him/her feedback about fruit growth uniformity and expected fruit quality. Late measurements, right before harvesting, would provide feedback on the applied strategy and information for the next season. When use in a whole farm approach instead of in a sampling approach, those systems would allow farms to better plan their logistics (labor force hiring, distribution within the field, transport and storage capacity, etc.) and also accurately estimate their yield and benefit according to fruit size classifications. In addition, when fruit size distribution is obtained for a whole plot, real-time or even map-based selective harvesting strategies could also be applied after a cost-benefit analysis according to fruit size or even to fruit colour (when such information is also gathered).

Summarizing, having information on the number of fruit, their size and their location within the trees and throughout the plot on a continuous, non-discreet, bases, will allow farms apply fruit quality and or cost-benefit -oriented strategies and make more informed decisions in the framework of Precision Agriculture or Precision Fructiculture resulting in enhanced fruit quality and reduced fruit size distribution.

6. Conclusions

From the analysis of previous research in fruit detection and sizing, it can be concluded that, although very significant advances have been achieved in the recent past (in particular since the development of deep learning algorithms), it remains as an open field of study, which is currently a focal point of great interest.

Fruit load management and yield estimation in fruit orchards is still usually done by manual/visual fruit counting and sizing. However, this is always a costly task in terms of time and labour. For this reason, automatic counting using fruit detection systems is becoming a feasible option for the fruit sector.

Actually, both leaf area and fruit size can be estimated with LiDAR, RGB and RGB-D sensors-based systems, enabling the tree-individual analysis of fruit bearing capacity. Such precise management avoids errors and, therefore, can contribute to more sustainable fruit production. In postharvest, the fruit size determines the fruit value in some crops such as sweet cherry. In other crops such as apples, the storability of fruit can be affected by fruit size.

Both active and passive electromagnetic (EM) radiation-based sensors are being used for detection and sizing of fruits, most of them in the visible, IR or thermal region of the EM spectrum. Hopeful future advances are expected from new emerging sensor, electronics and post processing systems as well as sensor fusion, which should lead to achieving this goal in a practical and affordable way in a few years. Optimizing measurements' size files, GNNS accuracy and systems' simplicity of use no doubt will help greatly to the adoption of the commercially products that will gradually appear in the coming years.

Hand in hand with these advances in sensors, a key point has also been the developments made in the applied algorithms, taking special relevance those based on artificial intelligence techniques and specifically deep learning based on convolutional neural networks, CNNs. Most fruit detection and sizing recent approaches use image classification, object detection or semantic and instance segmentation CNNs. However, near future advances are also linked to the availability of high quality and size datasets to train the algorithms that will be developed from now on.

CRediT authorship contribution statement

Juan C. Miranda: Conceptualization, Data curation, Investigation, Methodology, Writing – original draft. **Jordi Gené-Mola:** Conceptualization, Data curation, Investigation, Methodology, Supervision, Writing – original draft, Writing – review & editing. **Manuela Zude-Sasse:** Conceptualization, Methodology, Supervision, Writing – original draft, Writing – review & editing. **Nikos Tsoulias:** Conceptualization, Methodology, Writing – original draft, Writing – review & editing. **Alexandre Escolà:** Conceptualization, Methodology, Project administration, Writing – original draft, Writing – review & editing. **Jaume Arnó:** Conceptualization, Methodology, Project administration, Supervision, Writing – original draft, Writing – review & editing. **Joan R. Rosell-Polo:** Conceptualization, Methodology, Project administration, Writing – original draft, Writing – review & editing. **Ricardo Sanz-Cortiella:** Conceptualization, Methodology, Writing – original draft, Writing – review & editing. **José A. Martínez-Casasnovas:** Conceptualization, Methodology, Project administration, Writing – review & editing. **Eduard Gregorio:** Conceptualization, Investigation, Methodology, Project administration, Supervision, Writing – original draft, Writing – review & editing.

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.

Data availability

No data was used for the research described in the article.

Acknowledgement

This work was partly funded by the Department of Research and Universities of the Generalitat de Catalunya (grants 2017 SGR 646 and 2021 LLAV 00088) and by the Spanish Ministry of Science and Innovation / AEI/10.13039/501100011033 / FEDER (grants RTI2018-094222-B-I00 [PAgFRUIT project] and PID2021-126648OB-I00 [PAgPROTECT project]). The Secretariat of Universities and Research of the Department of Business and Knowledge of the Generalitat de Catalunya and European Social Fund (ESF) are also thanked for financing Juan Carlos Miranda's pre-doctoral fellowship (2020 FI_B 00586). 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.

References

Alcobendas, R., Mirás-Avalos, J.M., Alarcón, J.J., Nicolás, E., 2013. Effects of irrigation and fruit position on size, colour, firmness and sugar contents of fruits in a mid-late maturing peach cultivar. Sci. Hortic. 164 (17), 340–347. https://doi.org/10.1016/j.scienta.2013.09.048.

Altaheri, H., Alsulaiman, M., Muhammad, G., Amin, S.U., Bencherif, M., Mekhtiche, M., 2019. Date fruit dataset for intelligent harvesting. Data Brief 26, 104514. https://doi.org/10.1016/j.dib.2019.104514.

Anderson, N.T., Walsh, K.B., Koirala, A., Wang, Z., Amaral, M.H., Dickinson, G.R., Sinha, P., Robson, A.J., 2021a. Estimation of fruit load in Australian mango orchards using machine vision. Agronomy 11 (9), 1711. https://doi.org/10.3390/agronomy11091711.

Anderson, N.T., Walsh, K.B., Wulfsohn, D., 2021b. Technologies for forecasting tree fruit load and harvest timing—From ground, sky and time. Agronomy 11 (7), 1409. https://doi.org/10.3390/agronomy11071409.

Apolo-Apolo, O.E., Martínez-Guanter, J., Egea, G., Raja, P., Pérez-Ruiz, M., 2020a. Deep learning techniques for estimation of the yield and size of citrus fruits using a UAV. Eur. J. Agron. 115, 126030 https://doi.org/10.1016/j.eja.2020.126030.

Apolo-Apolo, O.E., Pérez-Ruiz, M., Martínez-Guanter, J., Valente, J., 2020b. A cloud-based environment for generating yield estimation maps from apple orchards using UAV imagery and a deep learning technique. Front. Plant Sci. 11, 1086. https://doi.org/10.3389/fpls.2020.01086.

Arnó, J., 2008. Una Investigación sobre la Variabilidad Intraparcelaria en Viña y el Uso de Sensores Láser en Viticultura de Precisión (Spanish) (Ph.D. Thesis), Universitat de Lleida, Catalonia, Spain. ⟨https://www.tesisenred.net/handle/10803/8150#page=1⟩.

Bargoti, S., Underwood, J., 2017a. Deep fruit detection in orchards. In: Proceedings of the 2017 IEEE Int. Conf. Robot. Autom. (ICRA), 3626–3633. https://doi.org/10.1109/ICRA.2017.7989417.

Bargoti, S., Underwood, J.P., 2017b. Image segmentation for fruit detection and yield estimation in apple orchards. J. Field Robot. 34, 1039–1060. https://doi.org/10.1002/rob.21699.

Barth, R., Hemming, J., Van Henten, E.J., 2020. Optimising realism of synthetic images using cycle generative adversarial networks for improved part segmentation. Comput. Electron. Agric. 173, 105378 https://doi.org/10.1016/j.compag.2020.105378.

Bavay, C., Symoneaux, R., Maître, I., Kuznetsova, A., Bruun Brockhoff, P., Mehinagic, E., 2013. Importance of fruit variability in the assessment of apple quality by sensory evaluation. Postharvest Biol. Technol. 77, 67–74. https://doi.org/10.1016/j.postharvbio.2012.11.005.

Bechar, A., Vigneault, C., 2016. Agricultural robots for field operations: concepts and components. Biosyst. Eng. 149, 94–111. https://doi.org/10.1016/j.biosystemseng.2016.06.014.

Bellocchio, E., Ciarfuglia, T.A., Costante, G., Valigi, P., 2019. Weakly supervised fruit counting for yield estimation using spatial consistency. IEEE Robot. Autom. Lett. 4, 2348–2355. https://doi.org/10.1109/LRA.2019.2903260.

Bellocchio, E., Costante, G., Cascianelli, S., Fravolini, M.L., Valigi, P., 2020. Combining domain adaptation and spatial consistency for unseen fruits counting: a quasi-unsupervised approach. IEEE Robot. Autom. Lett. 5, 1079–1086. https://doi.org/10.1109/LRA.2020.2966398.

Bhattarai, U., Karkee, M., 2022. A weakly-supervised approach for flower/fruit counting in apple orchards. Comput. Ind. 138, 103635 https://doi.org/10.1016/j.compind.2022.103635.

Bhusal, S., Karkee, M., Zhang, Q., 2019. Apple dataset benchmark from orchard environment in modern fruiting wall [WWW Document]. Res. Exch. https://doi.org/10.7273/000001752.

Biffi, L.J., Mitishita, E., Liesenberg, V., Santos, A.A. dos, Gonçalves, D.N., Estrabis, N.V., Silva, J., de, A., Osco, L.P., Ramos, A.P.M., Centeno, J.A.S., Schimalski, M.B., Rufato, L., Neto, S.L.R., Marcato Junior, J., Gonçalves, W.N., 2021. ATSS deep learning-based approach to detect apple fruits. Remote Sens 13 (1), 54. https://doi.org/10.3390/rs13010054.

Bonany, J., Buehler, A., Carbó, J., Codarin, S., Donati, F., Echeverría, G., Egger, S., Guerra, W., Hilaire, C., Höller, I., Iglesias, I., Jesionkowska, K., Konopacka, D., Kruczynska, D., Martinelli, A., Pitiot, C., Sansavini, S., Etehr, R., Schoorl, F., 2013. Consumer eating quality acceptance of new apple varieties in different European countries. Food Qual. Prefer. 30 (2), 250–259. https://doi.org/10.1016/j.foodqual.2013.06.004.

Bresilla, K., Perulli, G.D., Boini, A., Morandi, B., Corelli Grappadelli, L., Manfrini, L., 2019. Single-shot convolution neural networks for real-time fruit detection within the tree. Front. Plant Sci. 10, 611. https://doi.org/10.3389/fpls.2019.00611.

Carion, N., Massa, F., Synnaeve, G., Usunier, N., Kirillov, A., Zagoruyko, S., 2020. End-to-end object detection with transformers. arXiv:2005.12872. ⟨http://arxiv.org/abs/2005.12872⟩.

Casagrande, E., Génard, M., Lurol, S., Charles, F., Plénet, D., Lescourret, F., 2021. A process-based model of nectarine quality development during pre- and post-harvest. Postharvest Biol. Technol. 175, 111458 https://doi.org/10.1016/j.postharvbio.2020.111458.

Chaivivatrakul, S., Dailey, M.N., 2014. Texture-based fruit detection. Precis. Agric. 15, 662–683. https://doi.org/10.1007/s11119-014-9361-x.

Chen, J., Wu, J., Wang, Z., Qiang, H., Cai, G., Tan, C., Zhao, C., 2021. Detecting ripe fruits under natural occlusion and illumination conditions. Comput. Electron. Agric. 190, 106450 https://doi.org/10.1016/j.compag.2021.106450.

Chen, S.W., Shivakumar, S.S., Dcunha, S., Das, J., Okon, E., Qu, C., Taylor, C.J., Kumar, V., 2017. Counting apples and oranges with deep learning: a data-driven approach. IEEE Robot. Autom. Lett. 2, 781–788. https://doi.org/10.1109/LRA.2017.2651944.

Chen, Y., Lee, W.S., Gan, H., Peres, N., Fraisse, C., Zhang, Y., He, Y., 2019. Strawberry yield prediction based on a deep neural network using high-resolution aerial orthoimages. Remote Sens. 11 (13), 1584. https://doi.org/10.3390/rs11131584.

Cheng, H., Damerow, L., Sun, Y., Blanke, M., 2017. Early yield prediction using image analysis of apple fruit and tree canopy features with neural networks. J. Imaging 3 (1), 6. https://doi.org/10.3390/jimaging3010006.

Chu, P., Li, Z., Lammers, K., Lu, R., Liu, X., 2021. Deep learning-based apple detection using a suppression mask R-CNN. Pattern Recognit. Lett. 147, 206–211. https://doi.org/10.1016/j.patrec.2021.04.022.

Danckaers, F., Huysmans, T., Van Dael, M., Verboven, P., Nicolaï, B., Sijbers, J., 2017. Building 3D statistical shape models of horticultural products. Food Bioprocess. Technol. 10, 2100–2112. https://doi.org/10.1007/s11947-017-1979-z.

Das, J., Cross, G., Qu, C., Makineni, A., Tokekar, P., Mulgaonkar, Y., Kumar, V., 2015. Devices, systems, and methods for automated monitoring enabling precision agriculture. In: Proceedings of the 2015 IEEE International Conference on Automation Science and Engineering (CASE), 462–469. https://doi.org/10.1109/CoASE.2015.7294123.

Delong, M.J., Prange, K.R., Harrison, A.P., Embree, G.C., Nichols, S.D., Wright, A.H., 2006. The influence of crop-load, delayed cooling and storage atmosphere on post-storage quality of 'Honeycrisp'™ apples. J. Hortic. Sci. Biotechnol. 81 (3), 391–396. https://doi.org/10.1080/14620316.2006.11512078.

Dendorfer, P., Ošep, A., Milan, A., Schindler, K., Cremers, D., Reid, I., Roth, S., Leal-Taixé, L., 2020. MOTChallenge: A Benchmark for Single-Camera Multiple Target Tracking. Int. J. Comput. Vis. 129, 845–881. https://doi.org/10.1007/s11263-020-01393-0.

Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., Fei-Fei, L., 2009. ImageNet: A large-scale hierarchical image database In: Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition, 248–255. https://doi.org/10.1109/CVPRW.2009.5206848.

dos Santos Ferreira, A., Freitas, D.M., da Silva, G.G., Pistori, H., Folhes, M.T., 2019. Unsupervised deep learning and semi-automatic data labeling in weed discrimination. Comput. Electron. Agric. 165, 104963 https://doi.org/10.1016/j.compag.2019.104963.

Eitel, J.U.H., Vierling, L.A., Long, D.S., 2010. Simultaneous measurements of plant structure and chlorophyll content in broadleaf saplings with a terrestrial laser scanner. Remote Sens. Environ. 114 (10), 2229–2237. https://doi.org/10.1016/j.rse.2010.04.025.

Eizentals, P., Oka, K., 2016. 3D pose estimation of green pepper fruit for automated harvesting. Comput. Electron. Agric. 128, 127–140. https://doi.org/10.1016/j.compag.2016.08.024.

Embree, C.G., Myra, M.T.D., Nichols, D.S., Wright, A.H., 2007. Effect of blossom density and crop load on growth, fruit quality, and return bloom in 'Honeycrisp' apple. HortScience 42 (7), 1622–1625. https://doi.org/10.21273/HORTSCI.42.7.1622.

Everingham, M., Van Gool, L., Williams, C.K.I., Winn, J., Zisserman, A., 2010. The Pascal visual object classes (VOC) challenge. Int. J. Comput. Vis. 88, 303–338. https://doi.org/10.1007/s11263-009-0275-4.

FAO, 2017. The Future of Food and Agriculture – Trends and Challenges, Rome. ⟨http://www.fao.org/3/i6583e/i6583e.pdf⟩. (Accessed 16 September 2022).

Farjon, G., Krikeb, O., Hillel, A.B., Alchanatis, V., 2020. Detection and counting of flowers on apple trees for better chemical thinning decisions. Precis. Agric. 21 (3), 503–521. https://doi.org/10.1007/s11119-019-09679-1.

Favre, L., Hunter, D.A., O'Donoghue, E.M., Erridge, Z.A., Napier, N.J., Somerfield, S.D., Hunt, M., McGhie, T.K., Cooney, J.M., Saei, A., Chen, R.K.Y., McKenzie, M.J., Brewster, D., Martin, H., Punter, M., Carr, B., Tattersall, A., Johnston, J.W., Gibon, Y., Heyes, J.A., Lill, R.E., Brummell, D.A., 2022. Integrated multi-omic analysis of fruit maturity identifies biomarkers with drastic abundance shifts spanning the harvest period in 'Royal Gala' apple. Postharvest Biol. Technol. 193, 112059 https://doi.org/10.1016/j.postharvbio.2022.112059.

Fraser, D.G., Jordan, R.B., Künnemeyer, R., McGlone, V.A., 2003. Light distribution inside mandarin fruit during internal quality assessment by NIR spectroscopy. Postharvest Biol. Technol. 27 (2), 185–196. https://doi.org/10.1016/S0925-5214(02)00058-3.

Fu, L., Feng, Y., Wu, J., Liu, Z., Gao, F., Majeed, Y., Al-Mallahi, A., Zhang, Q., Li, R., Cui, Y., 2021. Fast and accurate detection of kiwifruit in orchard using improved YOLOv3-tiny model. Precis. Agric. 22, 754–776. https://doi.org/10.1007/s11119-020-09754-y.

Fu, L., Gao, F., Wu, J., Li, R., Karkee, M., Zhang, Q., 2020. Application of consumer RGB-D cameras for fruit detection and localization in field: a critical review. Comput. Electron. Agric. 177, 105687 https://doi.org/10.1016/j.compag.2020.105687.

Fu, L., Liu, Z., Majeed, Y., Cui, Y., 2018. Kiwifruit yield estimation using image processing by an Android mobile phone. IFAC Pap. 51 (17), 185–190. https://doi.org/10.1016/j.ifacol.2018.08.137.

Fu, L., Wu, F., Zou, X., Jiang, Y., Lin, J., Yang, Z., Duan, J., 2022. Fast detection of banana bunches and stalks in the natural environment based on deep learning. Comput. Electron. Agric. 194, 106800 https://doi.org/10.1016/j.compag.2022.106800.

Gai, R., Chen, N., Yuan, H., 2021. A detection algorithm for cherry fruits based on the improved YOLO-v4 model. Neural Comput. Appl. https://doi.org/10.1007/s00521-021-06029-z.

Gan, H., Lee, W.S., Alchanatis, V., Ehsani, R., Schueller, J.K., 2018. Immature green citrus fruit detection using color and thermal images. Comput. Electron. Agric. 152, 117–125. https://doi.org/10.1016/j.compag.2018.07.011.

Gao, F., Fu, L., Zhang, X., Majeed, Y., Li, R., Karkee, M., Zhang, Q., 2020. Multi-class fruit-on-plant detection for apple in SNAP system using Faster R-CNN. Comput. Electron. Agric. 176, 105634 https://doi.org/10.1016/j.compag.2020.105634.

Gené-Mola, J., Gregorio, E., Auat Cheein, F., Guevara, J., Llorens, J., Sanz-Cortiella, R., Escolà, A., Rosell-Polo, J.R., 2020a. Fruit detection, yield prediction and canopy geometric characterization using LiDAR with forced air flow. Comput. Electron. Agric. 168, 105121 https://doi.org/10.1016/j.compag.2019.105121.

Gené-Mola, J., Gregorio, E., Auat Cheein, F., Guevara, J., Llorens, J., Sanz-Cortiella, R., Escolà, A., Rosell-Polo, J.R., 2020b. LFuji-air dataset: annotated 3D LiDAR point clouds of Fuji apple trees for fruit detection scanned under different forced air flow conditions. Data Brief 29, 105248. https://doi.org/10.1016/j.dib.2020.105248.

Gené-Mola, J., Gregorio, E., Guevara, J., Auat, F., Sanz-Cortiella, R., Escolà, A., Llorens, J., Morros, J.R., Ruiz-Hidalgo, J., Vilaplana, V., Rosell-Polo, J.R., 2019a. Fruit detection in an apple orchard using a mobile terrestrial laser scanner. Biosyst. Eng. 187, 171–184. https://doi.org/10.1016/j.biosystemseng.2019.08.017.

Gené-Mola, J., Llorens, J., Rosell-Polo, J.R., Gregorio, E., Arnó, J., Solanelles, F., Martínez-Casasnovas, J.A., Escolà, A., 2020c. Assessing the performance of RGB-D sensors for 3d fruit crop canopy characterization under different operating and lighting conditions. Sensors 20 (24), 7072. https://doi.org/10.3390/s20247072.

Gené-Mola, J., Sanz-Cortiella, R., Rosell-Polo, J.R., Escolà, A., Gregorio, E., 2021a. In-field apple size estimation using photogrammetry-derived 3D point clouds: Comparison of 4 different methods considering fruit occlusions. Comput. Electron. Agric. 188, 106343 https://doi.org/10.1016/j.compag.2021.106343.

Gené-Mola, J., Sanz-Cortiella, R., Rosell-Polo, J.R., Escolà, A., Gregorio, E., 2021b. 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, 107629. https://doi.org/10.1016/j.dib.2021.107629.

Gené-Mola, J., Sanz-Cortiella, R., Rosell-Polo, J.R., Morros, J.-R.R., Ruiz-Hidalgo, J., Vilaplana, V., Gregorio, E., 2020d. Fruit detection and 3D location using instance segmentation neural networks and structure-from-motion photogrammetry. Comput. Electron. Agric. 169, 105165 https://doi.org/10.1016/j.compag.2019.105165.

Gené-Mola, J., Sanz-Cortiella, R., Rosell-Polo, J.R., Morros, J.R., Ruiz-Hidalgo, J., Vilaplana, V., Gregorio, E., 2020e. Fuji-SfM dataset: a collection of annotated images and point clouds for Fuji apple detection and location using structure-from-motion photogrammetry. Data Brief 30, 105591. https://doi.org/10.1016/j.dib.2020.105591.

Gené-Mola, J., Vilaplana, V., Rosell-Polo, J.R., Morros, J.-R., Ruiz-Hidalgo, J., Gregorio, E., 2019b. KFuji RGB-DS database: Fuji apple multi-modal images for fruit detection with color, depth and range-corrected IR data. Data Brief 25, 104289. https://doi.org/10.1016/j.dib.2019.104289.

Gené-Mola, J., Vilaplana, V., Rosell-Polo, J.R., Morros, J.-R.R., Ruiz-Hidalgo, J., Gregorio, E., 2019c. Multi-modal deep learning for Fuji apple detection using RGB-D cameras and their radiometric capabilities. Comput. Electron. Agric. 162, 689–698. https://doi.org/10.1016/j.compag.2019.05.016.

Gongal, A., Amatya, S., Karkee, M., Zhang, Q., Lewis, K., 2015. Sensors and systems for fruit detection and localization: a review. Comput. Electron. Agric. 116, 8–19. https://doi.org/10.1016/j.compag.2015.05.021.

Gongal, A., Karkee, M., Amatya, S., 2018. Apple fruit size estimation using a 3D machine vision system. Inf. Process. Agric. 5 (4), 498–503. https://doi.org/10.1016/j.inpa.2018.06.002.

Gregorio, E., Llorens, J., 2021. Sensing crop geometry and structure. In: Kerry, R., Escolà, A. (Eds.), Sensing Approaches for Precision Agriculture. Springer, Cham, pp. 59–92. https://doi.org/10.1007/978-3-030-78431-7_3.
Grilli, E., Battisti, R., Remondino, F., 2021. An advanced photogrammetric solution to measure Apples. Remote Sens. 13 (1), 3960. https://doi.org/10.3390/rs13193960.
Grilli, E., Menna, F., Remondino, F., 2017. A review of point clouds segmentation and classification algorithms. Int. Arch. Photogramm., Remote Sens. Spat. Inf. Sci. - ISPRS Arch. 42 (2W3), 339–344. https://doi.org/10.5194/isprs-archives-XLII-2-W3-339-2017.
Habib, M.T., Majumder, A., Jakaria, A.Z.M., Akter, M., Uddin, M.S., Ahmed, F., 2020. Machine vision based papaya disease recognition. J. King Saud. Univ. Comput. Inf. 32 (3), 300–309. https://doi.org/10.1016/j.jksuci.2018.06.006.
Hacking, C., Poona, N., Manzan, N., Poblete-Echeverría, C., 2019. Investigating 2-D and 3-D proximal remote sensing techniques for vineyard yield estimation. Sensors 19 (17), 3652. https://doi.org/10.3390/s19173652.
Hampson, C., Sanford, K., Cline, J., 2002. Preferences of Canadian consumers for apple fruit size. Can. J. Plant Sci. 82 (1), 165–167. https://doi.org/10.4141/P01-059.
Häni, N., Roy, P., Isler, V., 2020a. A comparative study of fruit detection and counting methods for yield mapping in apple orchards. J. Field Robot. 37, 263–282. https://doi.org/10.1002/rob.21902.
Häni, N., Roy, P., Isler, V., 2020b. MinneApple: a benchmark dataset for apple detection and segmentation. IEEE Robot. Autom. Lett. 5 (2), 852–858. https://doi.org/10.1109/LRA.2020.2965061.
Harker, F.R., Gunson, F.A., Jaeger, S.R., 2003. The case for fruit quality: an interpretive review of consumer attitudes, and preferences for apples. Postharvest Biol. Technol. 28 (3), 333–347. https://doi.org/10.1016/S0925-5214(02)00215-6.
He, K., Gkioxari, G., Dollár, P., Girshick, R., 2017. Mask R-CNN. In: Proceedings of the 2017 IEEE Int. Conf. Comput. Vis (ICCV), 2980–2988. https://doi.org/10.1109/ICCV.2017.322.
He, L., Fang, W., Zhao, G., Wu, Z., Fu, L., Li, R., Majeed, Y., Dhupia, J., 2022. Fruit yield prediction and estimation in orchards: a state-of-the-art comprehensive review for both direct and indirect methods. Comput. Electron. Agric. 195, 106812 https://doi.org/10.1016/j.compag.2022.106812.
He, L., Zhou, J., Zhang, Q., Karkee, M., 2015. Evaluation of multipass mechanical harvesting on 'Skeena' sweet cherries trained to Y-trellis. HortScience 50 (8), 1178–1182. https://doi.org/10.21273/HORTSCI.50.8.1178.
Heim, S., Zude-Sasse, M., 2014. Analyse der Fruchtwachstumsrate für die Bewertung einer selektiven Ernte bei Prunus avium 'Celeste' und Prunus domestica 'Tophit plus'. In: Proceedings of the Annual Conference DGG. Deutsche Gartenbauwissenschaftliche Gesellschaft, Dresden, Germany.
Hemming, J., Ruizendaal, J., Willem Hofstee, J., van Henten, E.J., 2014. Fruit detectability analysis for different camera positions in sweet-pepper. Sensors 14, 6032–6044. https://doi.org/10.3390/s140406032.
Herrero-Huerta, M., González-Aguilera, D., Rodriguez-Gonzalvez, P., Hernández-López, D., 2015. Vineyard yield estimation by automatic 3D bunch modelling in field conditions. Comput. Electron. Agric. 110, 17–26. https://doi.org/10.1016/j.compag.2014.10.003.
Hobbs, J., Prakash, P., Paull, R., Hovhannisyan, H., Markowicz, B., Rose, G., 2021. Large-scale counting and localization of pineapple inflorescence through deep density-estimation. Front. Plant Sci. 11, 599705. ⟨https://www.frontiersin.org/articles/10.3389/fpls.2020.599705/full⟩.
Itakura, K., Narita, Y., Noaki, S., Hosoi, F., 2021. Automatic pear and apple detection by videos using deep learning and a Kalman filter. OSA Contin. 4, 1688–1695. https://doi.org/10.1364/osac.424583.
Iwanami, H., 2011. Breeding for fruit quality in apple. In: Jenks, M.A., Bebeli, P. (Eds.), Breeding for Fruit Quality, pp. 173–200. https://doi.org/10.1002/9780470959350.ch8.
Jarvinen, T.D., Choi, D., Heinemann, P., Schupp, J., Baugher, T.A., 2019. Tree trunk position estimation for accurate fruit counts in apple yield mapping. In: Proceedings of the 2019 ASABE International Meeting, 1900918. https://doi.org/10.13031/aim.201900918.
Jia, W., Zhang, Z., Shao, W., Hou, S., Ji, Z., Liu, G., Yin, X., 2021. FoveaMask: a fast and accurate deep learning model for green fruit instance segmentation. Comput. Electron. Agric. 191, 106488 https://doi.org/10.1016/j.compag.2021.106488.
Kang, H., Chen, C., 2019. Fruit detection and segmentation for apple harvesting using visual sensor in orchards. Sensors 19, 4599. https://doi.org/10.3390/s19204599.
Kang, H., Chen, C., 2020. Fruit detection, segmentation and 3D visualisation of environments in apple orchards. Comput. Electron. Agric. 171, 105302 https://doi.org/10.1016/j.compag.2020.105302.
Kelman, E., Linker, R., 2014. Vision-based localisation of mature apples in tree images using convexity. Biosyst. Eng. 118 (1), 174–185. https://doi.org/10.1016/j.biosystemseng.2013.11.007.
Kestur, R., Meduri, A., Narasipura, O., 2019. MangoNet: a deep semantic segmentation architecture for a method to detect and count mangoes in an open orchard. Eng. Appl. Artif. Intell. 77, 59–69. https://doi.org/10.1016/j.engappai.2018.09.011.
Koirala, A., Walsh, K.B., Wang, Z., McCarthy, C., 2019a. Deep learning – method overview and review of use for fruit detection and yield estimation. Comput. Electron. Agric. 162, 219–234. https://doi.org/10.1016/j.compag.2019.04.017.
Koirala, A., Walsh, K.B., Wang, Z., McCarthy, C., 2019b. Deep learning for real-time fruit detection and orchard fruit load estimation: benchmarking of 'MangoYOLO'. Precis. Agric. 20 (6), 1107–1135. https://doi.org/10.1007/s11119-019-09642-0.
Kurtser, P., Ringdahl, O., Rotstein, N., Andreasson, H., 2020a. PointNet and geometric reasoning for detection of grape vines from single frame RGB-D data in outdoor conditions. In: Proc. North. Light. Deep Learn. Work. 1–6.
Kurtser, P., Ringdahl, O., Rotstein, N., Berenstein, R., Edan, Y., 2020b. In-field grape cluster size assessment for vine yield estimation using a mobile robot and a consumer level RGB-D Camera. IEEE Robot. Autom. Lett. 5 (2), 2031–2038. https://doi.org/10.1109/LRA.2020.2970654.
Lang, A.H., Vora, S., Caesar, H., Zhou, L., Yang, J., Beijbom, O., 2019. Pointpillars: Fast encoders for object detection from point clouds. arXiv 1812.05784. ⟨https://arxiv.org/abs/1812.05784⟩.
LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521, 436–444. https://doi.org/10.1038/nature14539.
Lee, J., Mattheis, J.P., Rudell, D.R., 2013. Fruit size affects physiological attributes and storage disorders in cold-stored 'Royal Gala' apples. HortScience 48 (12), 1518–1524. https://doi.org/10.21273/HORTSCI.48.12.1518.
Li, H., Lee, W.S., Wang, K., 2016. Immature green citrus fruit detection and counting based on fast normalized cross correlation (FNCC) using natural outdoor colour images. Precis. Agric. 17 (6), 678–697. https://doi.org/10.1007/s11119-016-9443-z.
Lin, G., Tang, Y., Zou, X., Li, J., Xiong, J., 2019. In-field citrus detection and localisation based on RGB-D image analysis. Biosyst. Eng. 186, 34–44. https://doi.org/10.1016/j.biosystemseng.2019.06.019.
Lin, T.Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., Zitnick, C.L., 2014. Microsoft COCO: Common objects in context. In: Proceedings of the European Conference on Computer Vision, 740–755. https://doi.org/10.1007/978–3-319–10602-1_48.
Liu, T.H., Ehsani, R., Toudeshki, A., Zou, X.J., Wang, H.J., 2019. Identifying immature and mature pomelo fruits in trees by elliptical model fitting in the Cr–Cb color space. Precis. Agric. 20 (1), 138–156. https://doi.org/10.1007/s11119-018-9586-1.
Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.Y., Berg, A.C., 2016. SSD: Single shot multibox detector. In: Proceedings of the European Conference on Computer Vision, 21–37. https://doi.org/10.1007/978–3-319–46448-0_2.
Liu, X., Chen, S.W., Aditya, S., Sivakumar, N., Dcunha, S., Qu, C., Taylor, C.J., Das, J., Kumar, V., 2018. Robust fruit counting: combining deep learning, tracking, and structure from motion. In: Proceedings of the 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 1045–1052. https://doi.org/10.1109/IROS.2018.8594239.
Liu, X., Chen, S.W., Liu, C., Shivakumar, S.S., Das, J., Taylor, C.J., Underwood, J., Kumar, V., 2019. Monocular camera based fruit counting and mapping with semantic data association. IEEE Robot. Autom. Lett. 4 (3), 2296–2303. https://doi.org/10.1109/LRA.2019.2901987.
Long, J., Shelhamer, E., Darrell, T., 2015. Fully convolutional networks for semantic segmentation. In: Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 3431–3440. https://doi.org/10.1109/CVPR.2015.7298965.
Lu, S., Chen, W., Zhang, X., Karkee, M., 2022. Canopy-attention-YOLOv4-based immature/mature apple fruit detection on dense-foliage tree architectures for early crop load estimation. Comput. Electron. Agric. 193, 106696 https://doi.org/10.1016/j.compag.2022.106696.
Lu, Y., Young, S., 2020. A survey of public datasets for computer vision tasks in precision agriculture. Comput. Electron. Agric. 178, 105760 https://doi.org/10.1016/j.compag.2020.105760.
Luo, L., Tang, Y., Zou, X., Ye, M., Feng, W., Li, G., 2016. Vision-based extraction of spatial information in grape clusters for harvesting robots. Biosyst. Eng. 151, 90–104. https://doi.org/10.1016/j.biosystemseng.2016.08.026.
Maldonado, W., Barbosa, J.C., 2016. Automatic green fruit counting in orange trees using digital images. Comput. Electron. Agric. 127, 572–581. https://doi.org/10.1016/j.compag.2016.07.023.
Mekhalfi, M.L., Nicolò, C., Ianniello, I., Calamita, F., Goller, R., Barazzuol, M., Melgani, F., 2020. Vision system for automatic on-tree kiwifruit counting and yield estimation. Sensors 20 (15), 4214. https://doi.org/10.3390/s20154214.
Méndez, V., Pérez-Romero, A., Sola-Guirado, R., Miranda-Fuentes, A., Manzano-Agugliaro, F., Zapata-Sierra, A., Rodríguez-Lizana, A., 2019. In-field estimation of orange number and size by 3D laser scanning. Agronomy 9 (12), 885. https://doi.org/10.3390/agronomy9120885.
Merzlyak, M.N., Solovchenko, A.E., Gitelson, A.A., 2003. Reflectance spectral features and non-destructive estimation of chlorophyll, carotenoid and anthocyanin content in apple fruit. Postharvest Biol. Technol. 27 (2), 88–103. https://doi.org/10.1016/S0925-5214(02)00066-2.
Meshram, V., Thanomliang, K., Ruangkan, S., Chumchu, P., Patil, K., 2020. FruitsGB: top Indian fruits with quality (WWW Document). IEEE DataPort. https://doi.org/10.21227/gzkn-f379.
Miller, W.M., Zude-Sasse, M., 2004. NIR-based sensing to measure soluble solids content of Florida citrus. Appl. Eng. Agric. 20 (3), 321–327. https://doi.org/10.13031/2013.16052.
Musacchi, S., Serra, S., 2018. Apple fruit quality: overview on pre-harvest factors. Sci. Hortic. 234, 409–430. https://doi.org/10.1016/j.scienta.2017.12.057.
Nanni, L., Ghidoni, S., Brahnam, S., 2017. Handcrafted vs. non-handcrafted features for computer vision classification. Pattern Recognit. 71, 158–172. https://doi.org/10.1016/j.patcog.2017.05.025.
Nguyen, T.T., Vandevoorde, K., Wouters, N., Kayacan, E., De Baerdemaeker, J.G., Saeys, W., 2016. Detection of red and bicoloured apples on tree with an RGB-D camera. Biosyst. Eng. 146, 33–44. https://doi.org/10.1016/j.biosystemseng.2016.01.007.
Nicastro, R., Carillo, P., 2021. Food loss and waste prevention strategies from farm to fork. Sustainability 13 (10), 5443. https://doi.org/10.3390/su13105443.
Nyarko, E.K., Vidović, I., Radočaj, K., Cupec, R., 2018. A nearest neighbor approach for fruit recognition in RGB-D images based on detection of convex surfaces. Expert Syst. Appl. 114, 454–466. https://doi.org/10.1016/j.eswa.2018.07.048.
OECD, 2021. Apples, International Standards for Fruit and Vegetables. OECD Publishing, Paris. https://doi.org/10.1787/12ebba9f-en-fr.

Osman, Y., Dennis, R., Elgazzar, K., 2021. Yield estimation and visualization solution for precision agriculture. Sensors 21 (19), 6657. https://doi.org/10.3390/s21196657.
Parico, A.I.B., Ahamed, T., 2021. Real time pear fruit detection and counting using YOLOv4 models and Deep SORT. Sensors 21 (14), 4803. https://doi.org/10.3390/s21144803.
Paul, V., Pandey, R., 2014. Role of internal atmosphere on fruit ripening and storability - a review. J. Food Sci. Technol. 51, 1223–1250. https://doi.org/10.1007/s13197-011-0583-x.
Pawara, P., Boshchenko, A., Schomaker, L.R.B., Wiering, M.A., 2020. Deep learning with data augmentation for fruit counting. In: Rutkowski, L., et al. (Eds.), Artificial Intelligence and Soft Computing, 2020. ICAISC, pp. 203–214. ⟨https://link.springer.com/chapter/10.1007/978-3-030-61401-0_20⟩.
Pedraza, I.L.A., Diaz, J.F.A., Pinto, R.M., Becker, M., Tronco, M.L., 2019. Sweet citrus fruit detection in thermal images using fuzzy image processing. In: Proceedings of the IEEE Colombian Conference on Applications of Computational Intelligence (ColCACI 2019), 182–193. https://doi.org/10.1007/978–3-030–36211-9_15.
Penzel, M., Lakso, A.N., Tsoulias, N., Zude-Sasse, M., 2020. Carbon consumption of developing fruit and the fruit bearing capacity of individual RoHo 3615 and Pinova apple trees. Int. Agrophys. 34 (4), 409–423. https://doi.org/10.31545/intagr/127540.
Qi, C.R., Liu, W., Wu, C., Su, H., Guibas, L.J., 2018. Frustum PointNets for 3D object detection from RGB-D data. In: Proceedings of the Thirty First IEEE Conference on Computer Vision and PatternRecognition (CVPR 2018), pp. 918–927. https://doi:10.1109/CVPR.2018.00102.
Qi, C.R., Su, H., Mo, K., Guibas, L.J., 2017. PointNet: deep learning on point sets for 3D classification and segmentation. In: Proceedings of the Thirtieth IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2017), pp. 77–85. https://doi.org/10.1109/CVPR.2017.16.
Qian, J., Xing, B., Wu, X., Chen, M., Wang, Y., 2018. A smartphone-based apple yield estimation application using imaging features and the ANN method in mature period. Sci. Agric. 75 (4), 273–280. https://doi.org/10.1590/1678-992x-2016-0152.
Qureshi, W.S., Payne, A., Walsh, K.B., Linker, R., Cohen, O., Dailey, M.N., 2017. Machine vision for counting fruit on mango tree canopies. Precis. Agric. 18 (2), 224–244. https://doi.org/10.1007/s11119-016-9458-5.
Rakun, J., Stajnko, D., Zazula, D., 2019. Plant size estimation based on the construction of high-density corresponding points using image registration. Comput. Electron. Agric. 157, 288–304. https://doi.org/10.1016/j.compag.2018.12.053.
Redmon, J., Farhadi, A., 2018. YOLOv3: an incremental improvement. Tech Report, arXiv1804.02767. ⟨http://arxiv.org/abs/1804.02767⟩.
Regunathan, M., Lee, W.S., 2005. Citrus fruit identification and size determination using machine vision and ultrasonic sensors. In: Proceedings of the 2005 ASAE Annual International Meeting, 053017. https://doi.org/10.13031/2013.19821.
Ren, S., He, K., Girshick, R., Sun, J., 2017. Faster R-CNN: towards real-time object detection with region proposal networks. IEEE Trans. Pattern Anal. Mach. Intell. 39 (6), 1137–1149. https://doi.org/10.1109/TPAMI.2016.2577031.
Rist, F., Gabriel, D., Mack, J., Steinhage, V., Töpfer, R., Herzog, K., 2019. Combination of an automated 3D field phenotyping workflow and predictive modelling for high-throughput and non-invasive phenotyping of grape bunches. Remote Sens 11 (24), 2953. https://doi.org/10.3390/rs11242953.
Robinson, T.L., Lakso, A.N., Greene, D., 2017. Precision crop load management: the practical implementation of physiological models. Acta Hortic. 1177, 381–390. https://doi.org/10.17660/ActaHortic.2017.1177.55.
Rogge, S., Defraeye, T., Herremans, E., Verboven, P., Nicolaï, B.M., 2015. A 3D contour based geometrical model generator for complex-shaped horticultural products. J. Food Eng. 157, 24–32. https://doi.org/10.1016/j.jfoodeng.2015.02.006.
Roscher, R., Herzog, K., Kunkel, A., Kicherer, A., Töpfer, R., Förstner, W., 2014. Automated image analysis framework for high-throughput determination of grapevine berry sizes using conditional random fields. Comput. Electron. Agric. 100, 148–158. https://doi.org/10.1016/j.compag.2013.11.008.
Rose, J.C., Kicherer, A., Wieland, M., Klingbeil, L., Töpfer, R., Kuhlmann, H., 2016. Towards automated large-scale 3D phenotyping of vineyards under field conditions. Sensors 16 (12), 2136. https://doi.org/10.3390/s16122136.
Rosell, J.R., Sanz, R., 2012. A review of methods and applications of the geometric characterization of tree crops in agricultural activities. Comput. Electron. Agric. 81, 124–141. https://doi.org/10.1016/j.compag.2011.09.007.
Roy, A.M., Bhaduri, J., 2022. Real-time growth stage detection model for high degree of occultation using DenseNet-fused YOLOv4. Comput. Electron. Agric. 193, 106694 https://doi.org/10.1016/j.compag.2022.106694.
Roy, P., Isler, V., 2016. Surveying apple orchards with a monocular vision system. : IEEE Int. Conf. Autom. Sci. Eng. 916–921. https://doi.org/10.1109/COASE.2016.7743500.
Roy, P., Kislay, A., Plonski, P.A., Luby, J., Isler, V., 2019. Vision-based preharvest yield mapping for apple orchards. Comput. Electron. Agric. 164, 104897 https://doi.org/10.1016/j.compag.2019.104897.
Sa, I., Ge, Z., Dayoub, F., Upcroft, B., Perez, T., McCool, C., 2016. DeepFruits: a fruit detection system using deep neural networks. Sensors 16 (8), 1222. https://doi.org/10.3390/s16081222.
Saha, K.K., Zude-Sasse, M., 2022. Estimation of chlorophyll content in banana during shelf life using LiDAR laser scanner. Postharvest Biol. Technol. 192, 112011 https://doi.org/10.1016/j.postharvbio.2022.112011.
Sahu, D., Potdar, R.M., 2017. Defect identification and maturity detection of mango fruits using image analysis. Am. J. Artif. Intell. l 1, 5–14. https://doi.org/10.11648/j.ajai.20170101.12.
Saitone, T.L., Sexton, R.J., 2017. Agri-food supply chain: evolution and performance with conflicting consumer and societal demands. Eur. Rev. Agric. Econ. 44, 634–657. https://doi.org/10.1093/erae/jbx003.
Santos, T.T., de Souza, L.L., dos Santos, A.A., Avila, S., 2020. Grape detection, segmentation, and tracking using deep neural networks and three-dimensional association. Comput. Electron. Agric. 170, 105247 https://doi.org/10.1016/j.compag.2020.105247.
Sanz-Cortiella, R., Llorens-Calveras, J., Rosell-Polo, J.R., Gregorio-Lopez, E., Palacin-Roca, J., 2011. Characterisation of the LMS200 laser beam under the influence of blockage aurfaces. Influence on 3D acanning of tree orchards. Sensors 11, 2751–2772. https://doi.org/10.3390/s110302751.
Serra, S., Leisso, R., Giordani, L., Kalcsits, L., Musacchi, S., 2016. Crop load influences fruit quality, nutritional balance, and return bloom in 'Honeycrisp' apple. HortScience 51 (3), 236–244. https://doi.org/10.21273/HORTSCI.51.3.236.
Shi, G., Ranjan, R., Khot, L.R., 2020. Robust image processing algorithm for computational resource limited smart apple sunburn sensing system. Inf. Process. Agric. 7 (2), 212–222. https://doi.org/10.1016/j.inpa.2019.09.007.
Simonyan, K., Zisserman, A., 2014. Very deep convolutional networks for large-scale image recognition. arXiv 1409.1556. ⟨http://arxiv.org/abs/1409.1556⟩.
Stajnko, D., Čmelik, Z., 2005. Modelling of apple fruit growth by application of image analysis. Agric. Conspec. Sci. 70 (2), 59–64.
Stajnko, D., Lakota, M., Hočevar, M., 2004. Estimation of number and diameter of apple fruits in an orchard during the growing season by thermal imaging. Comput. Electron. Agric. 42 (1), 31–42. https://doi.org/10.1016/S0168-1699(03)00086-3.
Stajnko, D., Rakun, J., Blanke, M., 2009. Modelling apple fruit yield using image analysis for fruit colour, shape and texture. Eur. J. Hortic. Sci. 74 (6), 260–267.
Stein, M., Bargoti, S., Underwood, J., 2016. Image based mango fruit detection, localisation and yield estimation using multiple view geometry. Sensors 16, 1915. https://doi.org/10.3390/s16111915.
Sun, C., Aernouts, B., Saeys, W., 2021. Effects of harvest time, fruit size and cultivar on the bulk optical properties of Satsuma mandarin. Postharvest Biol. Technol. 175, 111412 https://doi.org/10.1016/j.postharvbio.2020.111412.
Sun, Q., Chai, X., Zeng, Z., Zhou, G., Sun, T., 2022. Noise-tolerant RGB-D feature fusion network for outdoor fruit detection. Comput. Electron. Agric. 198, 107034 https://doi.org/10.1016/j.compag.2022.107034.
Sun, S., Jiang, M., He, D., Long, Y., Song, H., 2019. Recognition of green apples in an orchard environment by combining the GrabCut model and Ncut algorithm. Biosyst. Eng. 187 (22), 201–213. https://doi.org/10.1016/j.biosystemseng.2019.09.006.
Syal, A., Garg, D., Sharma, S., 2014. Apple fruit detection and counting using computer vision techniques. In: Proceedings of the 2014 IEEE International Conference on Computational Intelligence and Computing Research (IEEE ICCIC 2014), 405–412. https://doi.org/10.1109/ICCIC.2014.7238364.
Tao, Y., Zhou, J., 2017. Automatic apple recognition based on the fusion of color and 3D feature for robotic fruit picking. Comput. Electron. Agric. 142, 388–396. https://doi.org/10.1016/j.compag.2017.09.019.
Tapia Zapata, N., Saha, K.K., Tsoulias, N., Zude-Sasse, M., 2022. Fourier analysis of LiDAR scanned 3D point cloud data for surface reconstruction and fruit size estimation. In: Proceedings of the 2022 IEEE Workshop on Metrology for Agriculture and Forestry (MetroAgriFor), 197–202. https://doi.org/10.1109/MetroAgriFor55389.2022.9964881.
Tian, Y., Yang, G., Wang, Z., Wang, H., Li, E., Liang, Z., 2019. Apple detection during different growth stages in orchards using the improved YOLO-V3 model. Comput. Electron. Agric. 157, 417–426. https://doi.org/10.1016/j.compag.2019.01.012.
Tsoulias, N., Paraforos, D.S., Xanthopoulos, G., Zude-Sasse, M., 2020. Apple shape detection based on geometric and radiometric features using a LiDAR laser scanner. Remote Sens. 12 (15), 2481. https://doi.org/10.3390/RS12152481.
Tsoulias, N., Saha, K.K., Zude-Sasse, M., 2023. In-situ fruit analysis by means of LiDAR 3D point cloud of normalized difference vegetation index (NDVI). Comput. Electron. Agric. 205, 107611 https://doi.org/10.1016/j.compag.2022.107611.
Tsoulias, N., Xanthopoulos, G., Fountas, S., Zude-Sasse, M., 2022. Effects of soil ECa and LiDAR-derived leaf area on yield and fruit quality in apple production. Biosyst. Eng. 223 (B), 182–199. https://doi.org/10.1016/j.biosystemseng.2022.03.007.
Tu, S., Pang, J., Liu, H., Zhuang, N., Chen, Y., Zheng, C., Wan, H., Xue, Y., 2020. Passion fruit detection and counting based on multiple scale faster R-CNN using RGB-D images. Precis. Agric. 21, 1072–1091. https://doi.org/10.1007/s11119-020-09709-3.
UN, 2022. World population prospects 2022. Department of Economic and Social Affairs. Population Division. https://population.un.org/wpp/_. (Accessed 16 September 2022).
Underwood, J.P., Hung, C., Whelan, B., Sukkarieh, S., 2016. Mapping almond orchard canopy volume, flowers, fruit and yield using lidar and vision sensors. Comput. Electron. Agric. 130, 83–96. https://doi.org/10.1016/j.compag.2016.09.014.
Vasconez, J.P., Delpiano, J., Vougioukas, S., Auat Cheein, F., 2020. Comparison of convolutional neural networks in fruit detection and counting: a comprehensive evaluation. Comput. Electron. Agric. 173, 105348 https://doi.org/10.1016/j.compag.2020.105348.
Villacrés, J.F., Auat Cheein, F., 2020. Detection and characterization of cherries: a deep learning usability case study in Chile. Agronomy 10 (6), 835. https://doi.org/10.3390/agronomy10060835.
Villacrés, J., Viscaino, M., Delpiano, J., Vougioukas, S., Auat Cheein, F., 2023. Apple orchard production estimation using deep learning strategies: a comparison of tracking-by-detection algorithms. Comput. Electron. Agric. 204, 107513 https://doi.org/10.1016/j.compag.2022.107513.
Voulodimos, A., Doulamis, N., Doulamis, A., Protopapadakis, E., 2018. Deep learning for computer vision: a brief review. Comput. Intell. Neurosci. 2018, 7068349. https://doi.org/10.1155/2018/7068349.
Walsh, K.B., Blasco, J., Zude-Sasse, M., Sun, X., 2020. Review: Visible-NIR 'point' spectroscopy in postharvest fruit and vegetable assessment: the science behind three

decades of commercial use. Postharvest Biol. Technol. 168, 111246 https://doi.org/10.1016/j.postharvbio.2020.111246.
Walsh, K., Payne, A., Subedi, P., Anderson, N., 2021. Mango Fruit on Tree Image Collection (WWW Document). CQUniversity. https://doi.org/10.4226/145/59c9b2448242c.
Wang, C., Lee, W.S., Zou, X., Choi, D., Gan, H., Diamond, J., 2018a. Detection and counting of immature green citrus fruit based on the Local Binary Patterns (LBP) feature using illumination-normalized images. Precis. Agric. 19 (6), 1062–1083. https://doi.org/10.1007/s11119-018-9574-5.
Wang, D., He, D., 2021. Channel pruned YOLO V5s-based deep learning approach for rapid and accurate apple fruitlet detection before fruit thinning. Biosyst. Eng. 210, 271–281. https://doi.org/10.1016/j.biosystemseng.2021.08.015.
Wang, D., He, D., 2022. Fusion of Mask RCNN and attention mechanism for instance segmentation of apples under complex background. Comput. Electron. Agric. 196, 106864 https://doi.org/10.1016/j.compag.2022.106864.
Wang, D., Li, C., Song, H., Xiong, H., Liu, C., He, D., 2020. Deep learning approach for apple edge detection to remotely monitor apple growth in orchards. IEEE Access 8, 26911–26925. https://doi.org/10.1109/ACCESS.2020.2971524.
Wang, H., Zheng, X., 2019. Survey of deep learning based object detection. In: Proceedings of the Second International Conference on Big Data Technologies (ICBDT2019), 149–153. https://doi.org/10.1145/3358528.3358574.
Wang, Q., Nuske, S., Bergerman, M., Singh, S., 2013. Automated crop yield estimation for apple orchards. In: Proceedings of the Thirteenth International Symposium on Experimental Robotics, 745–758. doi: 10.1007/978-3-319-00065-7_50.
Wang, X., Tang, J., Whitty, M., 2021. DeepPhenology: estimation of apple flower phenology distributions based on deep learning. Comput. Electron. Agric. 185, 106123 https://doi.org/10.1016/j.compag.2021.106123.
Wang, X., Tang, J., Whitty, M., 2022. Data-centric analysis of on-tree fruit detection: experiments with deep learning. Comput. Electron. Agric. 194, 106748 https://doi.org/10.1016/j.compag.2022.106748.
Wang, Z., Koirala, A., Walsh, K., Anderson, N., Verma, B., 2018b. In field fruit sizing using a smart phone application. Sensors 18 (10), 3331. https://doi.org/10.3390/S18103331.
Wang, Z., Walsh, K., Koirala, A., 2019. Mango fruit load estimation using a video based MangoYOLO—Kalman filter—Hungarian algorithm method. Sensors 19 (12), 2742. https://doi.org/10.3390/s19122742.
Wang, Z., Walsh, K.B., Verma, B., 2017. On-tree mango fruit size estimation using RGB-D images. Sensors 17 (12), 2738. https://doi.org/10.3390/s17122738.
Ward, I.R., Laga, H., Bennamoun, M., 2019. RGB-D Image-based object detection: from traditional methods to deep learning techniques. In: Rosin, P., Lai, Y.K., Shao, L., Liu, Y. (Eds.), RGB-D Image Analysis and Processing. Springer, Cham, pp. 169–201. https://doi.org/10.1007/978-3-030-28603-3_8.
Wendel, A., Underwood, J., Walsh, K., 2018. Maturity estimation of mangoes using hyperspectral imaging from a ground based mobile platform. Comput. Electron. Agric. 155, 298–313. https://doi.org/10.1016/j.compag.2018.10.021.
Williams, H.A.M., Jones, M.H., Nejati, M., Seabright, M.J., Bell, J., Penhall, N.D., Barnett, J.J., Duke, M.D., Scarfe, A.J., Seok, H., Lim, J., Macdonald, B.A., 2019. Robotic kiwifruit harvesting using machine vision, convolutional neural networks, and robotic arms. Biosyst. Eng. 181, 140–156. https://doi.org/10.1016/j.biosystemseng.2019.03.007.
Wojke, N., Bewley, A., Paulus, D., 2017. Simple online and realtime tracking with a deep association metric. In: Proceedings of the 2017 IEEE International Conference on Image Processing (ICIP), 3645–3649. https://doi.org/10.1109/ICIP.2017.8296962.
Wu, G., Li, B., Zhu, Q., Huang, M., Guo, Y., 2020. Using color and 3D geometry features to segment fruit point cloud and improve fruit recognition accuracys. Comput. Electron. Agric. 174, 105475 https://doi.org/10.1016/j.compag.2020.105475.
Xie, D., Chen, L., Liu, L., Chen, L., Wang, H., 2022. Actuators and sensors for application in agricultural robots: a review. Machines 10 (19), 913. https://doi.org/10.3390/machines10100913.
Yan, B., Fan, P., Lei, X., Liu, Z., Yang, F., 2021. A real-time apple targets detection method for picking robot based on improved YOLOv5. Remote Sens. 13 (9), 1619. https://doi.org/10.3390/rs13091619.
Yeo, S.T., Edwards, R.T., 2006. Encouraging fruit consumption in primary schoolchildren: a pilot study in North Wales, UK. J. Hum. Nutr. Diet. 19, 299–302. https://doi.org/10.1111/j.1365-277X.2006.00706.x.
Yu, L., Xiong, J., Fang, X., Yang, Z., Chen, Y., Lin, X., Chen, S., 2021. A litchi fruit recognition method in a natural environment using RGB-D images. Biosyst. Eng. 204, 50–63. https://doi.org/10.1016/j.biosystemseng.2021.01.015.
Yu, Q., Yang, H., Gao, Y., Ma, X., Chen, G., Wang, X., 2022a. LFPNet: Lightweight network on real point sets for fruit classification and segmentation. Comput. Electron. Agric. 194, 106691 https://doi.org/10.1016/j.compag.2022.106691.
Yu, T., Hu, C., Xie, Y., Liu, J., Li, P., 2022b. Mature pomegranate fruit detection and location combining improved F-PointNet with 3D point cloud clustering in orchard. Comput. Electron. Agric. 200, 107233 https://doi.org/10.1016/j.compag.2022.107233.
Yu, Y., Zhang, K., Yang, L., Zhang, D., 2019. Fruit detection for strawberry harvesting robot in non-structural environment based on Mask-RCNN. Comput. Electron. Agric. 163, 104846 https://doi.org/10.1016/j.compag.2019.06.001.
Zhang, C., Zou, K., Pan, Y., 2020. A method of apple image segmentation based on color-texture fusion feature and machine learning. Agronomy 10 (7), 972. https://doi.org/10.3390/agronomy10070972.
Zhang, W., Chen, K., Wang, J., Shi, Y., Guo, W., 2021a. Easy domain adaptation method for filling the species gap in deep learning-based fruit detection. Hortic. Res. 8, 119.
Zhang, W., Liu, Y., Chen, K., Li, H., Duan, Y., Wu, W., Shi, Y., Guo, W., 2021b. Lightweight fruit-detection algorithm for edge computing applications. Front. Plant Sci. 12, 740936 https://doi.org/10.3389/fpls.2021.740936.
Zhang, Y., Yu, J., Chen, Y., Yang, W., Zhang, W., He, Y., 2022a. Real-time strawberry detection using deep neural networks on embedded system (rtsd-net): an edge AI application. Comput. Electron. Agric. 192, 106586 https://doi.org/10.1016/j.compag.2021.106586.
Zhang, Y., Zhang, W., Yu, J., He, L., Chen, J., He, Y., 2022b. Complete and accurate holly fruits counting using YOLOX object detection. Comput. Electron. Agric. 198, 107062 https://doi.org/10.1016/j.compag.2022.107062.
Zhao, Y., Gong, L., Huang, Y., Liu, C., 2016a. A review of key techniques of vision-based control for harvesting robot. Comput. Electron. Agric. 127, 311–323. https://doi.org/10.1016/j.compag.2016.06.022.
Zheng, C., Abd-Elrahman, A., Whitaker, V., 2021. Remote sensing and machine learning in crop phenotyping and management, with an emphasis on applications in strawberry farming. Remote Sens. 13 (3), 531. https://doi.org/10.3390/rs13030531.
Zhou, J., Cui, G., Hu, S., Zhang, Z., Yang, C., Liu, Z., Wang, L., Li, C., Sun, M., 2020. Graph neural networks: a review of methods and applications. AI Open 1, 57–81. https://doi.org/10.1016/j.aiopen.2021.01.001.