European Journal of Agronomy 115 (2020) 126030

Contents lists available at ScienceDirect

European Journal of Agronomy

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

Deep learning techniques for estimation of the yield and size of citrus fruits using a UAV

O.E. Apolo-Apolo[a], J. Martínez-Guanter[a], G. Egea[a], P. Raja[b], M. Pérez-Ruiz[a,*]

[a] Universidad de Sevilla, Área de Ingeniería Agroforestal, Dpto. de Ingeniería Aeroespacial y Mecánica de Fluidos, Spain
[b] School of Mechanical Engineering, SASTRA Deemed University, Tamil Nadu, India

ARTICLE INFO

Keywords:
Fruit detection
Machine learning
Citrus
Neural networks
Harvest
Fruit size
Yield estimation

ABSTRACT

Accurate and early estimation of citrus yields is important for both producers and agricultural cooperatives to be competitive and make informed decisions when selling their products. Yield estimation is key for predicting stock volumes, avoiding stock ruptures and planning harvesting operations. Visual yield estimations have traditionally been employed, resulting in inaccurate and misleading information. The main goal of this study was to develop an automated image processing methodology to detect, count and estimate the size of citrus fruits on individual trees using deep learning techniques. During 3 consecutive annual campaigns, a total of 20 trees from a commercial citrus grove were monitored using images captured from an unmanned aerial vehicle (UAV). These trees were harvested manually, and fruit sizes were measured. A Faster R-CNN Deep Learning model was trained using a custom dataset to detect oranges in the obtained images. An average standard error (SE) of 6.59 % was obtained between visual counting and the model's fruit detection. Using the detected fruits, fruit size estimation was also performed. The promising results obtained indicate that this size estimation method can be employed for size discrimination prior to harvest. A model based on Long Short-term Memory (LSTM) was trained for yield estimation per tree and for a total yield estimation. The actual and estimated yields per tree were compared, resulting in an approximate error of SE = 4.53 % and a standard deviation of SD = 0.97 Kg. The actual total yield, the estimated total yield and the total yield estimated by an expert technician were compared. The error in the estimation by the technician was SE = 13.74 %, while the errors in the model were SE = 7.22 % and SD = 4083.58 Kg. These promising results demonstrate the potential of the present technique to provide yield estimates for citrus fruits or even other types of fruit.

1. Introduction

Citrus fruits are one of the most important types of crops in the Mediterranean productive area, with a harvested area of 513,602 ha, 57 % of which (295,000 ha) is located in Spain and approximately 30 % of which (154,000 ha) is located in Italy. Therefore, the agro-industrial citrus system is extremely important in Spain (Beltrán-Esteve and Reig-Martínez, 2014). In this context, Spain is currently the largest producer and exporter of fresh-market orange fruit in Europe despite the strong market competitiveness of countries outside the European Union (Cardeñosa et al., 2015; García-Tejero et al., 2010).

The harvest operations for fresh-market oranges are among the most expensive tasks associated with Spanish citrus crops because such operations are manually performed (Castro-Garcia et al., 2017; Torregrosa et al., 2009). Fruit harvesting usually accounts for 25 % of the total production costs and requires nearly 50 % of the crop labour because mechanized harvesting is not yet widespread for this crop (Agustí Fonfría, 2012; Castro-Garcia et al., 2019). In addition, choosing the optimum harvest time depends on the maturity index, fruit colour and size, estimation of juice content and market demands, among other factors (Lado et al., 2014). However, although these parameters are important, the major challenge with this type of crop for both individual farmers and producer organizations (i.e., cooperatives) is estimating the abundance and size of fruits as early as possible to predict the volume of stock required at the supply chain level and to organize harvesting operations (Chinchuluun et al., 2006). For citrus and many other crops, fruit size is an important quality criterion (Gongal et al., 2018). Currently, yield information and fruit size are available only after the fruits are harvested, weighed and sorted using commercial grading machines in processing centres (Bulanon et al., 2009). Manual measurements of fruit traits are time-consuming and prone to measurement and recording errors (Rahnemoonfar and Sheppard, 2017).

* Corresponding author at: Ctra. Utrera, km 1, Seville, 41013, Spain.
E-mail address: manuelperez@us.es (M. Pérez-Ruiz).

https://doi.org/10.1016/j.eja.2020.126030
Received 12 August 2019; Received in revised form 20 February 2020; Accepted 21 February 2020
Available online 29 February 2020

Throughout history, estimates of crop yields have been based on the agronomic conditions of the crop, historical crop yield data and visual observations of the crop (Dorj et al., 2017). In Spanish orange orchards, visual techniques are often used to estimate yields; however, these methods are highly subjective because they depend on technician knowledge and experience. An error between the actual yield and yield prediction of 15–25% can occur (Castro-Garcia et al., 2019). Therefore, developing a highly accurate method to estimate crop yield and fruit size before harvest is the key to helping farmers make decisions.

In recent years, due to advances in computers, cameras and image analysis techniques, a wide range of methodologies based on counting fruits numbers have been developed to estimate crop yields (Gongal et al., 2015). For example, some earlier studies regarding fruit recognition were conducted on apples (Aggelopoulou et al., 2011; Gongal et al., 2018; Zhou et al., 2012), citrus fruits (Blasco et al., 2003; Gong et al., 2013; Kurtulmus et al., 2011; Lin et al., 2020, 2019a; Okamoto and Lee, 2009), tomatoes (Yamamoto et al., 2014; Zhao et al., 2016) and guava (Lin et al., 2019b). However, most of these techniques use algorithms based on the spectral response among pixels as a unique feature to detect and count fruits (Burnett and Blaschke, 2003). As image pixels are very sensitive to illumination changes under unstructured light conditions, using such algorithms can reduce the accuracy of fruit detection (Lin et al., 2019a; Rosebrock, 2016).

As an alternative to the above techniques, artificial intelligence (AI) methods, especially artificial neural networks (ANNs), have been recently used for yield estimation based on fruit detection (Kamilaris and Prenafeta-Boldú, 2018; Sa et al., 2016; Zhu et al., 2018). These methods have markedly improved over the past decade due to advances in computer technology, which has led to remarkable results in different areas, including agriculture (Krizhevsky et al., 2012; Lecun et al., 2015). For instance, highly successful yield estimations for mango (Kestur et al., 2019; Koirala et al., 2019; Wang et al., 2019), citrus (Zhuang et al., 2018), tomato (Rahnemoonfar and Sheppard, 2017) and apples (Cheng et al., 2017) have been achieved using ANNs. All these tools are part of what experts define as deep learning (DL), a subset of machine learning (ML), a non-linear information processing technique based on feature extraction and pattern analysis (Deng & Yu, 2014). In this scenario, scientists hope that these technologies can replace the knowledge and intuition that farmers have always had (Wolfert et al., 2017).

A previous literature review showed that while many algorithms and ANNs have been developed to estimate the yields of different types of crops, most of these methods use pictures taken at ground level, such as the method presented by Koirala et al. (2019). Although these methods obtained accurate yield results, some are time-consuming, labour intensive, costly and unviable for orange orchards due the size of the cultivated area. Unmanned aerial vehicles (UAVs), which have previously been used successfully in precision farming, are an inexpensive alternative that can be rapidly implemented (Martínez-Guanter et al., 2019). The main advantages of these vehicles are their flexibility, their low price and the repeatability of the results that they can obtain (Martínez et al., 2017). Sensors onboard these vehicles can generate a large amount of data, generally in the form of images or videos (Csillik et al., 2018; Ziliani et al., 2018). In this context, the combination of the imagery from UAVs with DL techniques provides unique perspectives and information that would otherwise be either impossible or very expensive to obtain using traditional techniques involving human effort (Yin et al., 2018).

Regarding yield estimation, only 20 % of ANN applications are used for this purpose in agricultural fields according to Liakos et al. (2018). Moreover, Moltó et al. (1992) and Jiménez et al. (2000) suggested that only approximately 60–70 % of crop production is visible from the outside of a tree. Therefore, a challenge in this type of study is to generate complex models that can address occluded fruits, changing light conditions or different sizes, shapes and positions that may be present in the field.

Thus, the main goal of this novel study was to develop and evaluate the performance and robustness of an automated citrus fruit detection and size estimation system using red-green-blue (RGB) images obtained by flying a UAV in combination with a developed ANN model. Moreover, a model based on the number of fruits detected to estimate the full yield of a commercial citrus orchard was developed. With this new tool to accurately estimate the yield and size of orange fruits, we plan to reduce the error associated with visual methods. In this manner, citrus producers can be assisted to increase their economic profitability and reduce uncertainty.

2. Materials and methods

2.1. Field tests and UAV imagery acquisition

The research was conducted during three consecutive seasons (2017, 2018 and 2019) and occurred over a 4-ha area of a commercial citrus orchard (Citrus sinensis L. cv. Navelina) near Seville, in southwestern Spain (latitude: 37.512574; longitude: -5.956659) with a tree spacing of 5.5 m × 4.5 m. Sampling was conducted from 20 individual trees in groups of five in a row randomly selected from a total of 1654 trees. Two weeks before the harvesting, two pictures per tree, including one for each face (left and right), were taken using a UAV (Fig. 1). A total of 40 RGB high-resolution images were acquired for each season under natural lighting conditions. The UAV used was a commercial quadcopter DJI Phantom 3 Professional (DJI Technology Co., Ltd., Shenzhen, China). The onboard sensor used was the RGB Sony Exmor 1/2.3″ CMOS camera with a lens FOV of 94°, a focal length of 20 mm, a focal ratio of f/2.8 f, and a focus to infinity. The image resolution of the camera was 4000 × 3000 pixels (JPG format). The UAV was flown between rows of trees at low altitude (between 5 and 6 m above the crop canopy). The flight altitude was configured according to previous experience such that the pixel size (0.26 cm/pixel) was suitable for fruit detection. The manual flight option was used, with a constant preserved distance for each image. A wooden ruler calibrated in centimetres (cm) was placed next to each tree during image acquisition to serve as a scale reference when determining the average fruit size.

The images acquired throughout the different campaigns by means of the UAV included in many cases some non-targeted citrus trees. Therefore, a delimitation of the region of interest (ROI) was performed to isolate the study entities. To do this, a mask was generated on the original image around the studied tree using a script developed with the open-source computer vision library OpenCV. The first step was to compute the centre of the image by dividing the width (4000 px) and height (3000 px) by two; hence, the central pixel of the image could be accessed by their (x, y) coordinates. Then, a white rectangle (width: 2200 px; height: 1650 px) was drawn using the OpenCV function *cv2.rectangle* to focus only on the target tree. The remaining pixels were turned off using the function *cv2.bitwise and* are shown as a black colour on each image (Fig. 2).

2.2. Obtention of the ground truth for yield estimation

The sampled citrus trees were harvested manually during the three seasons to quantify the actual yield in kilograms per tree (Y_{apt}). The fruit was picked and placed into picking sacks. Then, the fruit was weighed and dumped from the picking sacks into boxes (20 kg per box), which were manually loaded onto a field transport vehicle. A representative sample of 20 randomly selected fruits per tree was used to estimate the average fruit size. A gauge calibrated in millimetres and a balance with gram-level accuracy were used to measure the fruit sizes and weights, respectively. Historical total actual yield (T_{ayield}) data after the harvest for the whole citrus orchard from 2011 to 2018 are reported in Table 1.

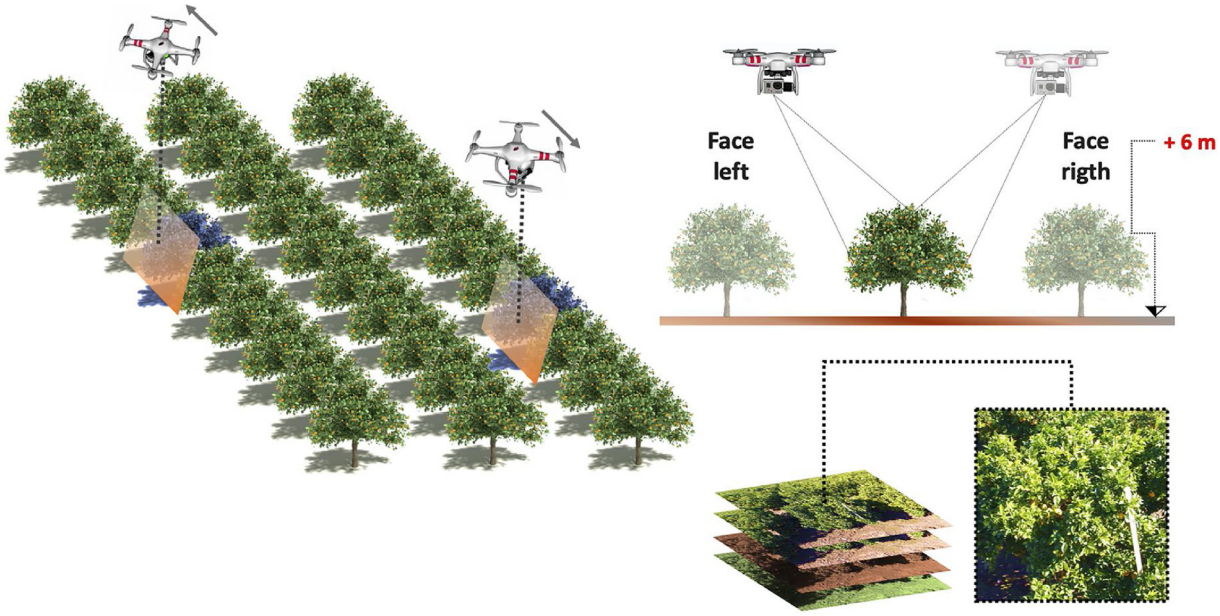

Fig. 1. Workflow in the field tests (left) and process for UAV image acquisition on both sides of the trees (upper right); obtained citrus tree canopy image with a wooden rule to scale the image (lower right).

2.3. Building the image dataset

To properly train the developed DL model, a considerable amount of orange images had to be obtained as a starting dataset. Without sufficient and representative training data, the model cannot learn the underlying discriminative patterns required to make robust classifications of fruits (Sonka et al., 1993). In this case, the characteristics of the oranges on the trees may dramatically differ (e.g., green fruits, fruits of different sizes and shapes, fruits occluded by branches and leaves, and overlapping fruits). Thus, additional flights were made over the rest of the trees of the crop to acquire a sufficient number of images to train the ANN. A total of 300 high-resolution images were taken from the UAV during the studied years.

However, in practice, the formation of a custom model using a single dataset is a difficult challenge because tens of thousands of images are often needed (Chollet, 2017; Lecun et al., 2015) to achieve high accuracy in object detection, but overfitting problems can occur. Accordingly, some researchers apply data augmentation (Krizhevsky et al., 2012; Simonyan and Zisserman, 2015). This methodology consists of increasing the size of a dataset using techniques to transform pictures, such as rotation, adding filters, and changing colour channels, among others (Ma et al., 2019). This process increases the network's ability to generalize and reduces overfitting (Rosebrock, 2018). In this work, pictures were rotated by 90, 180 and 270 degrees and resized to 600 × 600 px using a Python script developed by the authors (Fig. 3). Consequently, a dataset containing a total of 900 pictures was used to

Table 1
Historical yield data after the harvest of the whole citrus orchard.

Season	*T_{ayield} (kg)
2011	156,564.00
2012	161,291.00
2013	173,412.00
2014	117,200.00
2015	141,980.00
2016	193,715.00
2017	179,549.00
2018	146,210.00
Mean	158,740.12

* T_{ayield} = Total actual yield.

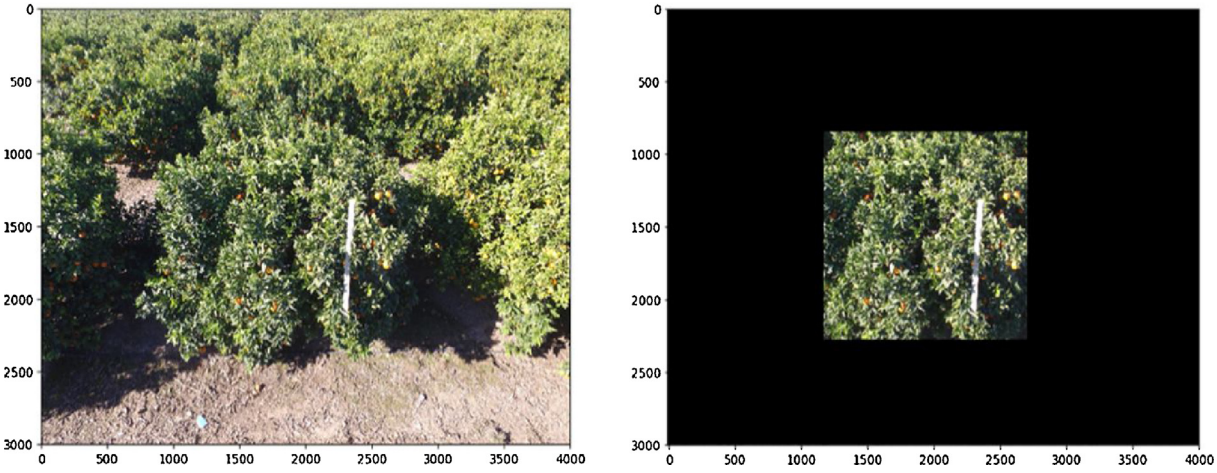

Fig. 2. Original image of an orange tree (left) and a masked image (right) where the masked pixels with values greater than zero are shown.

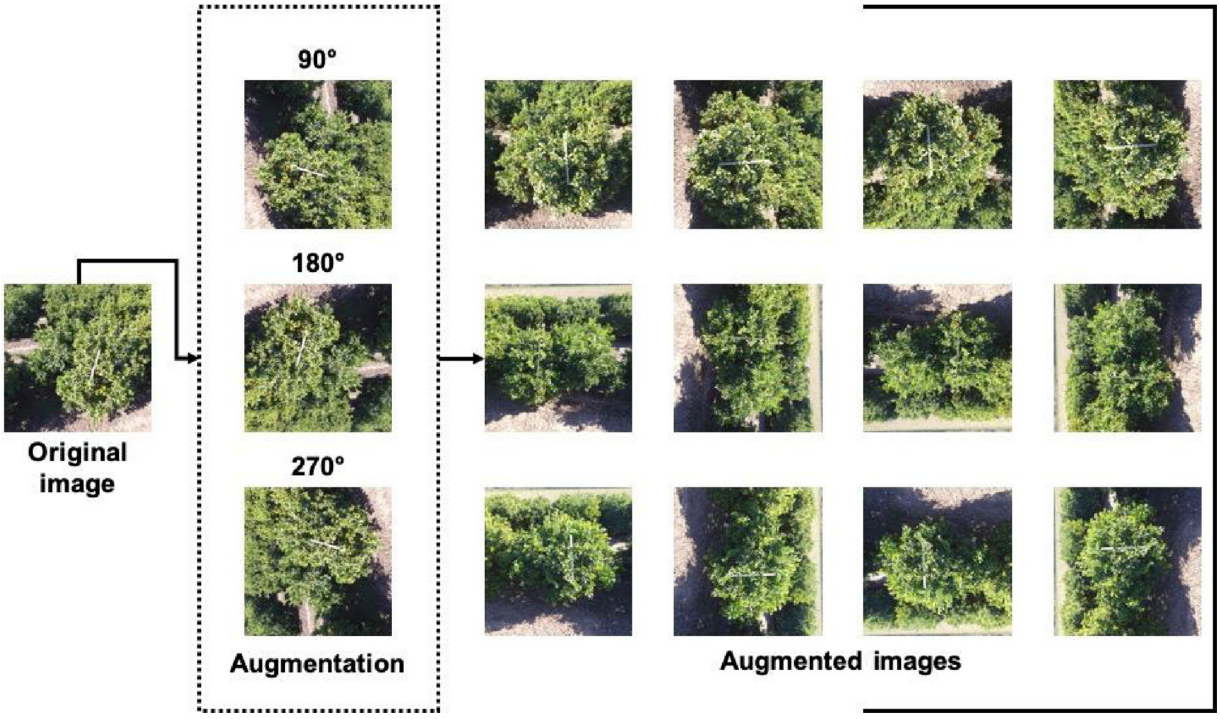

Fig. 3. Dataset augmentation process; original images were rotated 90, 180 and 270 degrees and resized to 600 × 600 px for use as input data.

train the CNN.

2.4. Deep learning architecture

An ANN model architecture was developed to identify the key features of the input images and ultimately find orange fruits in them. Among the entire range of existing ANNs, a convolutional neural network (CNN) was used, with the intention of mimicking the human eye in fruit detection and performing automatic fruit counting in the acquired images. CNNs consider an image to be a matrix of pixels whose size (kernel) is (height x width x depth), where the depth is the number of image channels (3 for our RGB crop images). CNNs have several specialized hidden layers with a hierarchical structure (Lecun et al., 2015); the first layers can detect lines, corners and simple shapes, whereas deeper layers can recognize complex shapes (Rosebrock, 2018). A common CNN architecture consists of three types of neural layers: a convolutional layer, a pooling layer and a fully connected layer (Fig. 4). Feature extraction, non-linearity operations and dimension reduction were performed with this common architecture. Additionally, a fully connected layer was used to classify data from images (Guo et al., 2016) while a softmax function assigned the probability of belonging to a class.

Despite the advances in computational processes and available power offered by graphics processing unit (GPU), training a neural network from scratch is still highly computationally expensive and requires large datasets for learning. To overcome these obstacles, the technique called transfer learning (Gu et al., 2018) was used. The main objective of this technique to transfer the knowledge from one model trained on a large dataset such as ImageNet (Deng et al., 2009) to another model to solve a specific task (Talukdar et al., 2018). Several popular pretrained networks using transfer learning, such as VGG-16, ResNet 50, DeepNet and AlexNet Inception V2, are described in the literature (Rosebrock, 2018).

The Faster-R-CNN model was selected since this network can use several architectures, such as ResNet, Inception and Atrous, thus increasing the efficiency and precision of fruit detection (Dias et al., 2018). The Faster R-CNN Inception Resnet V2 Atrous Coco (Ren et al., 2017) pre-trained model with a TensorFlow object detection application programming interface (API) was used. TensorFlow is an open-source software library for numerical computations (Kamilaris and Prenafeta-Boldú, 2018) and was used because of its flexibility and the ability of deploy network computations in multiple central processing

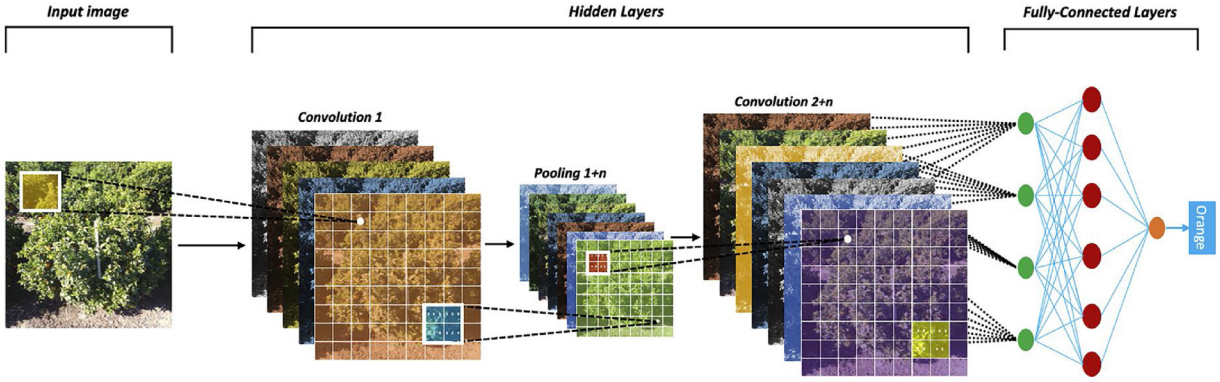

Fig. 4. CNN architecture for fruit detection.

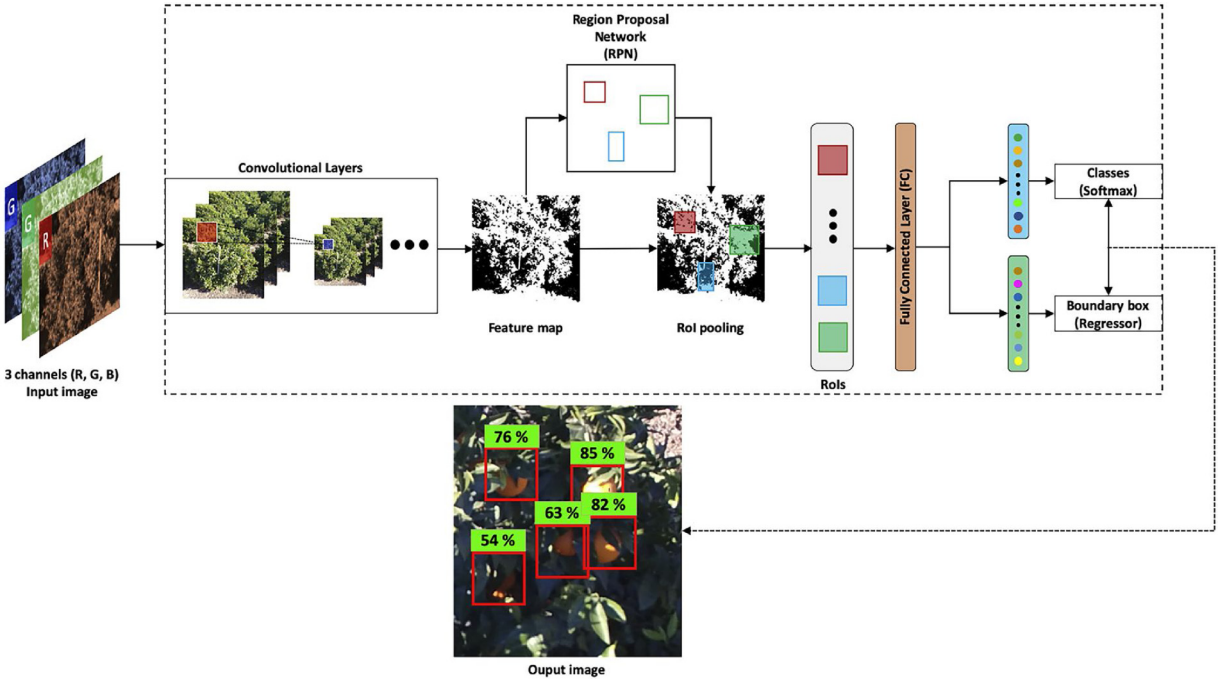

Fig. 5. The Faster R-CNN architecture.

units (CPUs), GPUs and servers. The model comprises three steps: with a citrus tree image as the input, Faster R-CNN extracts feature maps from the image using a CNN and then passes these maps through a region proposal network (RPN), which returns object proposals (Rosebrock, 2018). Finally, these maps are classified and the bounding boxes enclosing the orange fruits are predicted (Fig. 5).

The advantage of using pre-trained models is that these models are trained on huge image datasets such as the Common Objects in Context (COCO) dataset developed by Microsoft (Microsoft, Redmond, WA, USA). This dataset contains 330,000 labelled images with 80 object categories or classes in which the label "orange" can be found (Lin et al., 2014). At this point, the model was able to automatically detect oranges, but the dataset images were completely different from the acquired UAV images. Therefore, the transfer learning technique can be used to train a model to detect and count orange fruits in the UAV images.

2.5. Training the deep learning model (Faster R-CNN)

A manual labelling process to identify where the fruits were located in each image from the training dataset (900 images) was performed (Fig. 6). The open-source tool LabelImg (Tzutalin, 2015) was used for this process. Once all fruits were labelled, an Extensible Markup Language (XML) file with label data and the coordinates of the bounding rectangles for each fruit in the image was generated.

Once the labelling process was complete, the configuration details about the model and labels were implemented in the TensorFlow API. The computing hardware used for all the processes was a MacBook Pro (MacOs High Sierra 10.13.4) with a 2.5 GHz Intel Core i7 processor, 16 GB of RAM and a Graphics AMD Radeon R9 M370 × 2048 MB Intel Iris Pro 1536 MB. The Open-Source Computer Vision (OpenCV) library (http://opencv.org/), which includes several hundred computer vision algorithms, was used to process images (Rosebrock, 2016). The Keras (Chollet, 2015) open-source library was used in combination with TensorFlow backend tools to build and deploy the DL architecture. Finally, when the model was trained, the 40 images taken by the UAV

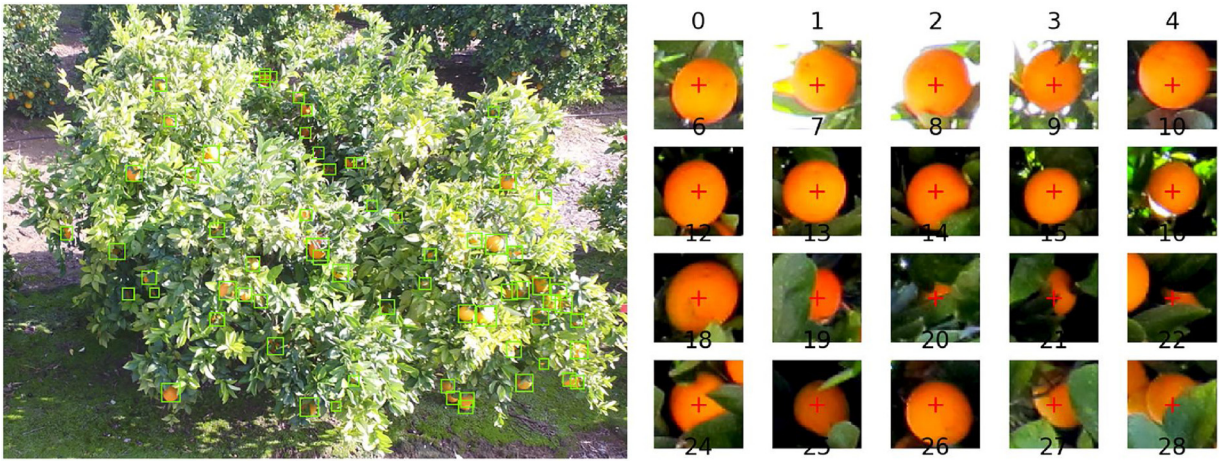

Fig. 6. The labelling process for the test images (left) and different types of fruits labelled (right).

over the 20 trees were used for fruit detection.

2.6. Modelling yield estimation based on fruit detection

Yield estimation from fruit counts is still a challenge. In production systems with large tree crowns, it is assumed that the images do not reflect the total yield but rather a percentage of it (Gongal et al., 2018). Additionally, the yield can be affected by several factors (climate, fertilization, diseases, etc.) and does not conform to a classical linear model. Currently, neural networks such as Long Short-Term Memory (LSTM) for recurrent neural networks can understand and recognize complex patterns in data (Fischer and Krauss, 2018). In this work, a model based on LSTM and encoding with Keras was used to estimate yield based on the fruit count of each tree. To train the model, the total number of fruits per tree (N_{ft}) for each face with the actual yield per tree (Y_{apt}) and the total actual yield (T_{ayield}) for each season were analysed. The data were assembled in pairs of two years for training, and the remaining one was used for testing. Finally, the model was trained for yield estimation per tree (Y_{ept}) and for total yield estimation (T_{eyield}) separately.

2.7. Image processing to estimate the fruit size

Size determination of fruits in an image is similar to computing the distance from a camera to an object; in both cases, a ratio that measures the number of pixels per a given scale or area must be defined (Rosebrock et al., 2016). In this project, together with the fruit detection and yield estimation using ANNs, fruit size estimation based on a "calibration" process using a reference object (wood ruler) was performed. The reference object has two important properties: (1) the dimensions (in terms of width and height) in a measurable unit (cm) are known and (2) the reference is easy to find and identifiable either in terms of the location of the object or its appearance.

The wood ruler used as a reference object has a known width (5 cm) and height (100 cm), with a total area of 500 cm^2. This object was measured in pixels in each image using a Python script with OpenCV. The script performs an exhaustive process from an RGB image to be able to measure fruit size. The first step is converting the image to greyscale since with this transformation, only eight bits are required to store and resolve image segmentation problems (Saravanan, 2010). Then, a Gaussian smoothing and a median blur filter using a kernel with a size of 5 × 5 pixels were used to remove the noise in the images. Next, a list of boundaries in the RGB colour was defined since OpenCV represents images as NumPy arrays in reverse order. Each entry in the list was defined as a range with two values: a list of lower limits and a list of upper limits. The lower and upper limits for both fruits and wood ruler were [(57, 125, 126), (119, 214, 236)] and [(131, 126, 117), (255, 255, 255)], respectively. These values allowed detection of the different colour ranges for each fruit on the canopy. Finally, once the limits were defined, the *cv2.inRange* built-in function was used. This method returns a binary image (mask), specifying which pixels fall into the specified upper and lower ranges (Fig. 7).

With the mask as an input, a process to identify the contours of the objects was carried out (Fig. 8). These contours were automatically examined using the *cv2.findContours* function. If the contour was not sufficiently large, the region was presumed to be noise left over from the previous segmentation process and discarded. Once contours were detected, the bounding box for each object in the image was computed using the *cv2.boxPoints* function. This function assign coordinates in top-left, top-right, bottom-right, and bottom-left order for each object detected. Knowing the coordinates for each corner of the bounding box, the midpoint (M) between any two points was calculated using Eq. (1):

$$M = \left(\frac{x_1 + x_2}{2}, \frac{y_1 + y_2}{2}\right), \quad (1)$$

where (x_1, y_1) and (x_2, y_2) are the coordinate values for a given line segment.

In the output image, an outline of the object detected in green colour with the vertices as small red circles and the values in pixels for both the width and length of the bounding box, with two purple lines connecting the midpoints, is shown (Fig. 8c). A width of 50.50 ± 1.43 pixels and a height of 747.06 ± 3.81 pixels for the wood ruler were obtained based on the associated bounding box (Fig. 8b). Thus, there were approximately 37,753.80 pixels per every 500 cm^2 in each image.

On the other hand, a fruit is a relatively circular object, and the total area (A) can be calculated based on the fruit diameter (D) using the expression $A = \frac{\pi D^2}{4}$. Hence, considering the diameter of a circle and the dimensions mentioned above, the fruit size (F_z) can be approximated using Eq. (2):

$$F_z = \sqrt{\frac{4x0.013xA_p}{\pi}} \quad (2)$$

where the value 0.013 is the relationship between the actual dimensions of the wood ruler and the average value of the dimension in pixels.

2.8. Statistical analysis for model performance evaluation

The total number of fruits per tree (N_{ft}) in each of the acquired pictures was manually counted using the Photoshop count tool (Adobe Systems Inc., San Jose, United States) as suggested by Payne et al. (2014). Therefore, with these data, the precision (P), recall (R) and F1-score were used as the evaluation metrics for fruit detection. These model evaluation metrics are defined as follows:

$$Precision\ (P) = \frac{TP}{TP + FP} \quad Recall\ (R) = \frac{TP}{TP + FN} \quad F_1 Score = 2 \cdot \left(\frac{P \cdot R}{P \cdot R}\right)$$

where TP corresponds to true positives, i.e., when the algorithm correctly detected a fruit with a bounding box, FP indicates false positives, i.e., when a box is computed in a location where a fruit is not located, and FN denotes false negatives, i.e. when a target fruit is not detected.

To assess the suitability of the method proposed, all fruit sizes were clustered into categories according to the Codex Standard for Oranges (CODEX STAN 245-2004) and compared with manual fruits size measurements. The set of categories in which oranges with a minimum dimension of less than 53 mm are excluded is reported in Table 2.

Finally, the actual yield per tree (T_{apt}) and the yield estimated per tree (Y_{ept}) were compared. Then, the actual total yield (T_{ayield}), the total estimated yield (T_{eyield}) and the estimates obtained by an expert technician were compared. The standard error (SE) and standard deviation (SD) were calculated to determine the accuracy of the model's estimates.

3. Results and discussion

3.1. Accuracy of the model in fruit detection

Achieving a high accuracy in fruit detection is crucial for good yield estimation. In this study, the 900 images labelled for orange detection were sufficient to explain the wide variability in the dataset. The workflow from an input image until the fruits are detected is shown in Fig. 9. Over each detected orange fruit, a red bounding box with the probability of containing the fruits is shown. The number of fruits detected was high, presumably due to the application of data augmentation, which helped to overcome the common problems in object detection caused by illumination conditions, the distance to the fruit, and fruit clustering, among other reasons, as suggested by Voulodimos et al. (2018).

In many cases, due to outdoor conditions, the pictures taken suffered notable changes in colour, thus causing difficulties in object detection. Therefore, in Table 3, an analysis of the precision of fruit

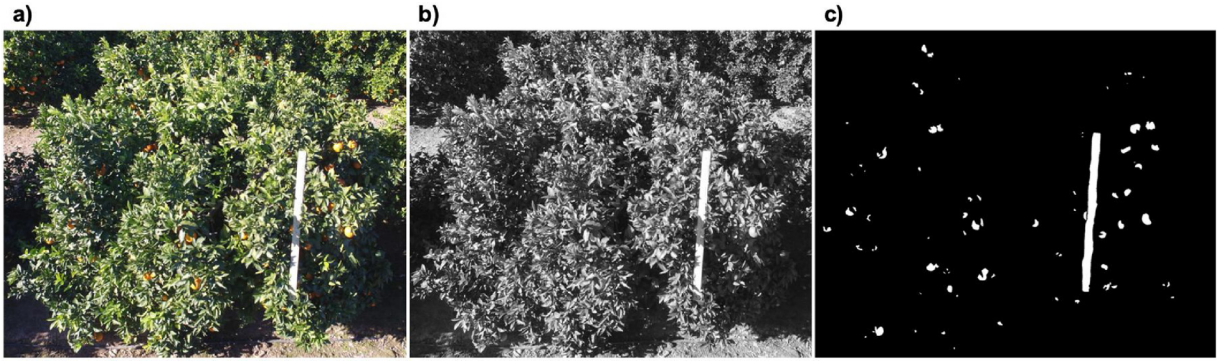

Fig. 7. Flowchart of the previous image processing steps in Python. The original RGB image (a) is first transformed into a greyscale image (b). Then, taking the lower and upper limits for both for fruits and the wood ruler, a binary image is obtained (c).

detection is presented. Overall, the values for each one of the metrics used to assess obtained results greater than 90 % in terms of precision (P). False positives were observed in the pictures in which the number of immature fruits (green) and the brightness of sunlight were slightly greater. These results can be significantly improved by taking pictures at several times throughout the day, as suggested by Ma et al. (2019). Finally, the F1-score (F1) exhibited values greater than 89 %, indicating the high robustness of the trained model.

The models based on object detection for fruit detection have limitations since these models cannot detect invisible fruits that are occluded by foliage or other fruits (Kamilaris and Prenafeta-Boldú, 2018). Fruits (oranges) are perceived within a tree canopy on the basis of colour, texture and contours, which are highly variable. Therefore, the model cannot detect all the fruits, but it can detect most of the visible fruit. In Table 4, the fruits manually counted and those detected by the model are reported. Considering manual counting to be the most reliable method, an error of approximately SE = 11 % was obtained. The errors between manual counts and object detection were similar to those obtained by Rolim et al. (2015). These results demonstrate that the use of simple data augmentation techniques such as image rotation and transfer learning can facilitate building tools with high potential for yield estimation.

3.2. Yield estimation

As previously discussed, most studies focused on fruit detection have tried to detect the number of fruits per tree as a means to estimate yields (Gongal et al., 2015), but this method is not sufficient to build a model that can predict the total yield of a commercial citrus orchard.

Table 2
Fruit categories according to CODEX STAN 245-2004.

Min (mm)	Max (mm)	Category
53	60	13
56	63	12
58	66	11
60	68	10
62	70	9
64	73	8
67	76	7
70	80	6
73	84	5
77	88	4
81	92	3
84	96	2
87	100	1
92	110	0

Hence, in Table 5, the actual yield per tree (Y_{apt}) and the estimated yield per tree (T_{ept}) are reported. The errors (SE values) obtained for each year are less than 3%, indicating good accuracy of the prediction. The errors with greater values occurred in 2018, probably because the actual yield for that year was less than expected in an average year.

According to the actual historical data (2011–2018) for the commercial citrus orchard, the yield can vary by approximately 11.5 % 12,530 kg between years, which is a high volume of production. In other words, farmers may experience profit or loss depending on the yield estimations. In this context, accurate yield estimations are challenging for both farmers and expert technicians. In Table 6, the

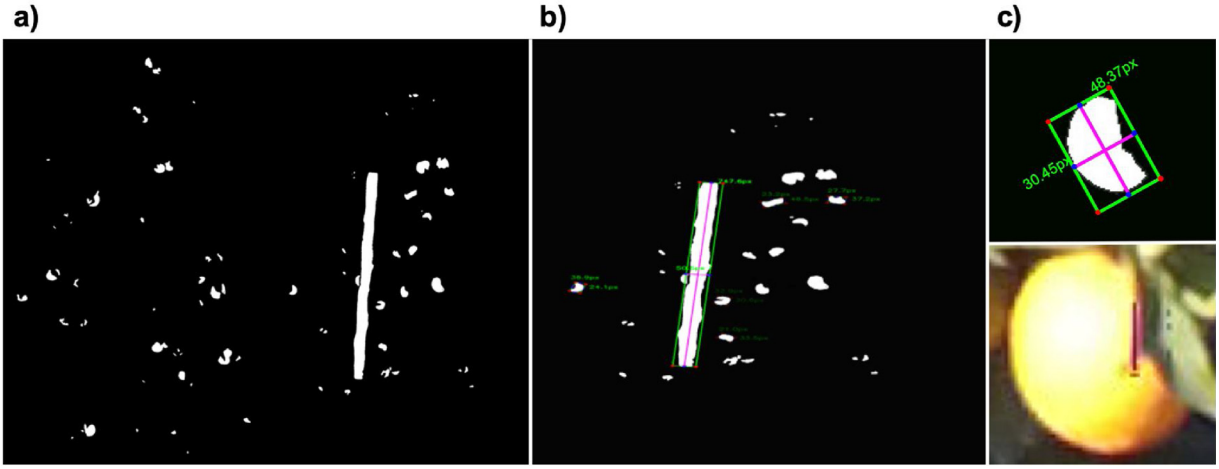

Fig. 8. Process to determine the number of pixels of each object from a mask (a). Objects detected with the bounding box and the number of pixels for both the width and height (b). Extended details of the bounding box for a fruit with its actual RGB values (c).

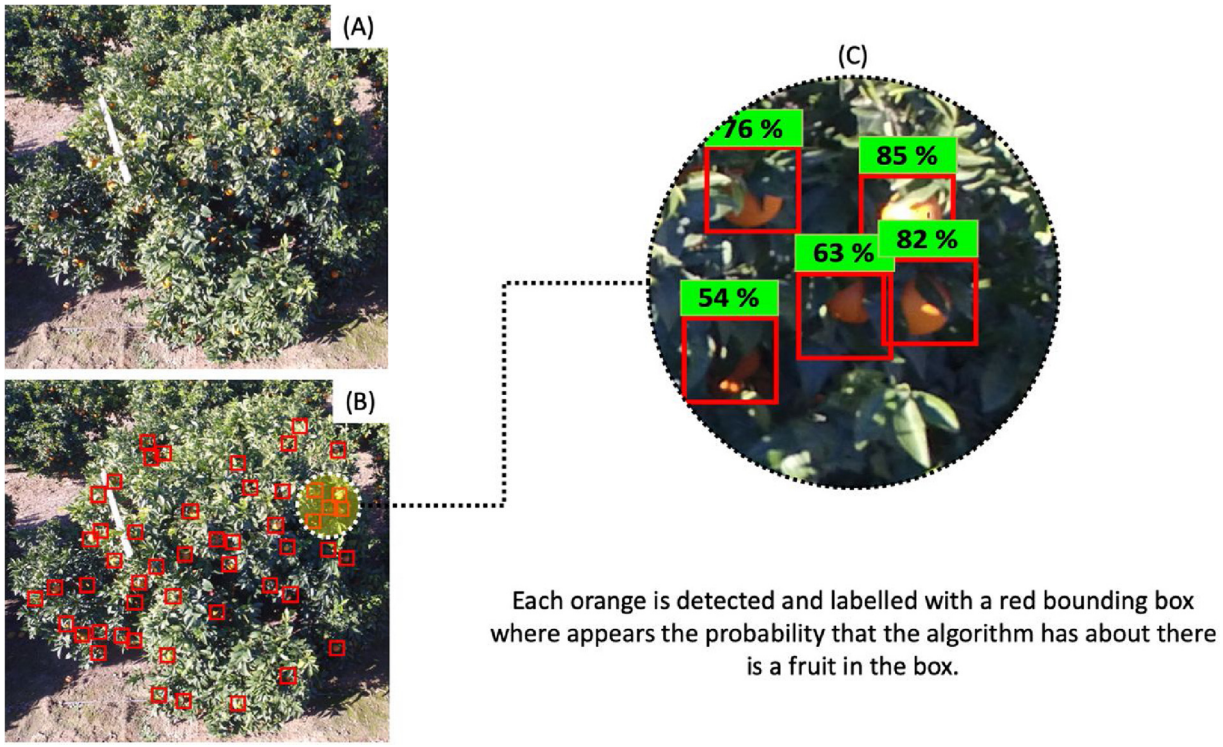

Fig. 9. (A) Original RGB image taken from an UAV; (B) orange fruits automatically detected and bounding boxes; (C) detail of the fruit detection and probabilities obtained by the Mask-R CNN model.

estimations for each year assessed are reported. The greatest error (SE = 11.52 %) occurred during 2016, probably due to a major volume of production, in contrast to the other years in which the error was greater than 5%. The best result (SE = 4.34 %) occurred for 2017 with the training data for (2016–2018), presumably because of the substantial differences in the volume of production. These differences cause a better learning by the model in the training process, which leads to more accurate estimation. However, the most important result is the estimate for all years, which had a mean error of 7.22 %. The error for the expert technician was 13.74 %. Although these results are slightly dissimilar, they may help to farmers in make better decisions.

As was discussed in the introduction, yield estimates with high accuracy are invaluable, especially to small farmers, since these data give to farmer negotiating power during the sale of their products. In many cases, exporters penalized farmers when the difference between the actual and yield estimate is high. Consequently, farmers lose money and sometimes need to find other buyers for their products or even face the risk of loss of the harvest. According to the Observatory for Prices and Markets of Andalusia the prices that farmers received for their orange production during 2016, 2017 and 2018 were 0.30, 0.28 and 0.26

Table 3
Fruit detection analyses for each tree during the three seasons analysed.

	Year 2016						Year 2017						Year 2018					
Tree	TP	FP	FN	P	R	F1	TP	FP	FN	P	R	F1	TP	FP	FN	P	R	F1
1	61	2	2	0.97	0.97	0.97	52	1	1	0.98	0.98	0.98	37	1	2	0.97	0.95	0.96
1	35	1	3	0.97	0.91	0.94	77	2	5	0.97	0.94	0.96	128	7	5	0.95	0.96	0.95
1	37	1	1	0.97	0.98	0.98	92	4	8	0.96	0.92	0.94	66	0	2	1.00	0.97	0.98
1	66	3	5	0.96	0.93	0.94	66	1	1	0.98	0.98	0.98	104	6	4	0.95	0.96	0.95
1	54	1	7	0.98	0.89	0.93	57	2	5	0.97	0.92	0.94	113	8	2	0.93	0.98	0.96
1	58	2	1	0.97	0.98	0.97	110	5	11	0.96	0.91	0.93	35	2	3	0.95	0.92	0.93
1	64	3	6	0.96	0.91	0.93	107	3	4	0.97	0.96	0.97	51	3	3	0.94	0.94	0.94
1	60	2	5	0.97	0.92	0.94	95	3	12	0.97	0.89	0.93	57	4	7	0.93	0.89	0.91
1	60	1	7	0.98	0.89	0.93	64	2	5	0.97	0.93	0.95	113	8	7	0.93	0.94	0.94
1	58	1	6	0.98	0.90	0.94	57	1	7	0.98	0.89	0.93	40	3	1	0.93	0.97	0.95
1	56	2	7	0.97	0.89	0.93	48	3	6	0.94	0.89	0.91	77	2	9	0.97	0.90	0.94
1	30	1	1	0.97	0.97	0.97	80	1	8	0.99	0.91	0.95	43	1	2	0.98	0.96	0.97
1	46	4	3	0.92	0.94	0.93	60	2	8	0.97	0.88	0.92	78	4	4	0.95	0.95	0.95
1	39	1	4	0.97	0.90	0.94	98	3	10	0.97	0.91	0.94	39	2	1	0.95	0.98	0.97
1	41	2	4	0.95	0.91	0.93	53	2	5	0.96	0.92	0.94	33	1	3	0.97	0.93	0.95
1	26	2	2	0.93	0.92	0.92	46	2	3	0.96	0.93	0.94	88	2	6	0.98	0.94	0.96
1	36	3	1	0.92	0.97	0.95	51	3	2	0.94	0.97	0.96	88	2	9	0.98	0.91	0.94
1	51	1	2	0.98	0.96	0.97	26	1	2	0.96	0.94	0.95	68	5	4	0.93	0.94	0.94
1	32	1	4	0.97	0.90	0.93	83	2	4	0.98	0.95	0.96	62	2	7	0.97	0.90	0.93
20	45	1	3	0.98	0.94	0.96	44	5	1	0.90	0.98	0.94	102	5	7	0.95	0.94	0.95
Mean	47.75	1.75	3.70	0.96	0.93	0.95	68.30	2.40	5.40	0.96	0.93	0.95	71.10	3.40	4.40	0.96	0.94	0.95

TP = true positives, FP = false positives, FN = false negatives, P = precision, R = recall, F1 = F1-score.

Table 4
The number of fruits manually counted and the number of fruits detected with the model.

	Year 2016			Year 2017			Year 2018		
Tree	N_{ft}	TP	Error (%)	N_{ft}	TP	Error (%)	N_{ft}	TP	Error (%)
1	63	61	3.17	53	52	1.89	39	37	5.13
2	38	35	7.89	82	77	6.10	133	128	3.76
3	38	37	2.63	100	92	8.00	68	66	2.94
4	71	66	7.04	67	66	1.49	108	104	3.70
5	61	54	11.48	62	57	8.06	115	113	1.74
6	59	58	1.69	121	110	9.09	38	35	7.89
7	70	64	8.57	111	107	3.60	54	51	5.56
8	65	60	7.69	107	95	11.21	64	57	10.94
9	67	60	10.45	69	64	7.25	120	113	5.83
10	64	58	9.38	64	57	10.94	41	40	2.44
11	63	56	11.11	54	48	11.11	85	77	9.41
12	31	30	3.23	88	80	9.09	45	43	4.44
13	49	46	6.12	68	60	11.76	82	78	4.88
14	43	39	9.30	108	98	9.26	40	39	2.50
15	45	41	8.89	58	53	8.62	36	33	8.33
16	28	26	7.14	49	46	6.12	94	88	6.38
17	37	36	2.70	53	51	3.77	97	88	9.28
18	53	51	3.77	28	26	7.14	72	68	5.56
19	35	32	8.57	87	83	4.60	69	62	10.14
20	48	45	6.25	45	44	2.22	109	102	6.42
Mean	51.40	47.75	6.85	73.70	68.30	7.07	75.45	71.10	5.86

TP = true positives, N_{ft} = the number of fruits per tree counted manually.

Table 5
Actual weights after harvesting for each tree in the three years.

	Year 2016				Year 2017				Year 2018			
Tree	Y_{apt}	Y_{ept}	Stdev	Error (%)	Y_{apt}	Y_{ept}	Stdev	Error (%)	Y_{apt}	Y_{ept}	Stdev	Error (%)
1	135.70	130.24	1.11	4.02	164.50	155.70	1.69	5.65	53.10	51.17	0.21	3.64
2	101.30	97.95	0.95	3.31	128.30	121.08	2.49	5.97	100.30	96.82	0.21	3.47
3	101.00	96.99	0.75	3.97	174.10	163.89	1.92	6.23	8.70	8.39	0.01	3.53
4	68.60	65.99	0.19	3.80	105.00	99.90	2.25	5.10	133.40	128.68	0.31	3.54
5	129.90	124.56	0.25	4.11	180.20	169.63	1.98	6.23	162.20	156.50	0.73	3.51
6	101.20	97.49	0.81	3.67	142.80	136.89	1.54	4.32	97.30	94.03	0.32	3.36
7	105.90	101.77	0.85	3.90	153.30	144.69	2.29	5.95	117.00	112.69	0.33	3.69
8	114.40	110.70	1.06	3.23	108.20	102.87	1.48	5.18	73.80	71.15	0.23	3.59
9	91.40	88.38	0.58	3.30	116.00	108.99	1.16	6.44	91.20	88.05	0.28	3.45
10	92.40	88.60	1.03	4.11	168.10	158.62	2.92	5.98	36.40	35.03	0.16	3.77
11	121.10	117.22	1.14	3.20	116.30	112.02	0.96	3.82	125.40	120.35	0.21	4.03
12	129.80	124.76	0.41	3.88	137.40	130.79	0.67	5.06	124.80	120.41	0.42	3.52
13	82.00	79.11	0.56	3.53	126.30	119.59	0.99	5.61	31.50	30.33	0.03	3.73
14	128.10	123.85	0.79	3.32	132.40	127.05	1.51	4.21	57.30	55.13	0.15	3.79
15	109.30	105.52	0.86	3.46	80.60	76.97	0.88	4.72	155.80	149.97	0.64	3.74
16	70.20	67.34	0.66	4.07	112.80	107.17	1.63	5.26	125.90	121.15	0.33	3.77
17	125.40	121.19	1.03	3.36	146.40	138.45	1.43	5.75	61.10	58.89	0.19	3.62
18	121.70	116.73	0.64	4.08	85.70	81.40	1.74	5.28	97.90	94.47	0.36	3.50
19	121.00	116.85	0.41	3.43	145.20	136.16	1.15	6.64	95.80	92.12	0.27	3.84
20	***	***	***	***	110.10	104.73	2.43	5.13	69.20	66.74	0.28	3.55
Mean	***	***	***	***	131.69	124.83	1.66	5.43	90.91	87.60	0.28	3.63

*** The Y_{apt} in tree number 20 in 2016 was missed during the harvest.

Table 6
Yield estimation for each year.

				Yield estimation using LSTM		
Year	T_{ayield}	Technical estimation	Error (%)	T_{eyield}	Stdev	Error (%)
2016	193715	220000	13.57	171400.06	7795.95	11.52
2017	179549	200000	11.39	171750.59	2861.67	4.34
2018	146210	170000	16.27	137724.95	1593.11	5.8
Mean	173158	196666.67	13.74	160291.87	4083.58	7.22

T_{ayield} = total actual yield, T_{eyield} = total estimated yield.

euros/kg respectively. Hence, considering the average selling price and the error made by the technician, an amount of 6582.43 euros/season was lost. However, by employing the proposed yield estimation technique based on LSTM, the farmer would save approximately 2979.91 euros/season.

3.3. Fruit size estimation

As an important parameter to determine fruit quality, market price and optimal time for harvesting (Gongal et al., 2018), fruit size was estimated throughout the three seasons by means of the automated image analysis method presented in Section 2. Fig. 10 presents the distributions of the actual fruit sizes and estimated fruit sizes. Most of the actual fruit sizes measured over the two years used as ground-truth

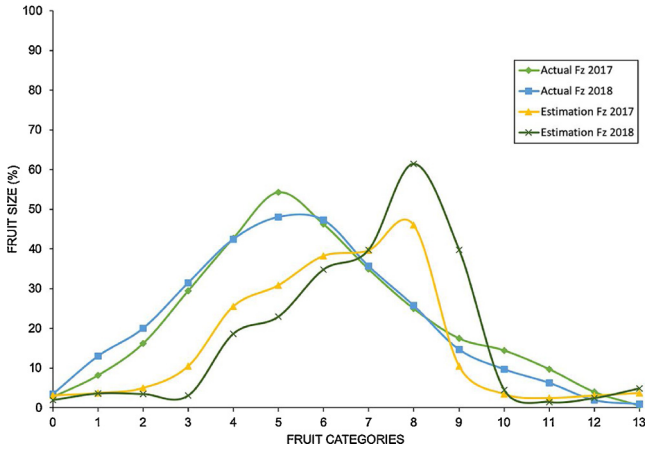

Fig. 10. Distributions of actual fruit sizes measured manually and estimated fruit sizes.

(2016-17 and 2017-18) were between categories 4 and 6. However, the fruit sizes estimated in the third season (2018-19) were approximately between categories 7 and 9. These differences are presumably due to the number of fruits analysed in each case. In other words, for the actual fruit sizes, only 20 fruits per tree were measured, while for fruit size estimation, all detected fruits whose surface were completely visible were used. Therefore, although this method only provides approximate values, the results can be applied to orange fruit management to optimize the resources required to harvest, transport, store, and, manufacture oranges as needed.

4. Conclusions

The results of this project in terms of the number of fruits and fruit size are promising. The fruit size data indicate that the average size can be effectively estimated; however, this approach is generic in nature because the sizes of all fruits detected cannot be computed. For example, some fruits are covered by leaves, branches and other elements of the orange tree canopy. Nevertheless, these results confirm that the developed model can be used to estimate the fruit sizes in an orange orchard.

The yield estimates were very accurate because the model results were closer to the actual yield than the results of visual estimations by a professional technician. These results support the development of another model to estimate the yields with even greater accuracy. With a system that generates adjusted results in this manner, using flight platforms such as UAVs represents a leap towards the implementation of this type of model covering increasingly larger areas and at a low cost. The flexibility and ease of handling and imaging of this equipment makes them a tool with great potential for close detection of fruits and analysis of crops. Future works will involve detection in real time from UAV videos, which can reduce the processing time required to obtain accurate results. Moreover, a new dataset with more images can be built to achieve more accurate performance in fruit size estimation.

In productive systems with large crown trees such as the one studied, fruit detection is greatly influenced by the presence of foliage and the structure of the crop itself. In the coming years, production systems will probably have to adapt to geometries that allow the majority of fruits to be visible from the outside. This is already happening, with a significant increase in trellised crops and narrow rows. In this manner, the automation of tasks such as detection will greatly benefit.

Author contributions

For research articles with several authors, a short paragraph specifying their individual contributions must be provided. The following statements should be used “Conceptualization, MPR and GE; methodology, OEAA, MPR, JMG and GE; software, OEAA and JMG; data analysis, OEAA and JMG; writing—original draft preparation OEAA, MPR, JMG and GE; writing—review and editing, MPR and OEAA; supervision, MPR and PR; project administration MPR and GE; funding acquisition, MPR and GE. All authors have read and agreed to the published version of the manuscript.

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.

Acknowledgments

The authors would like to thank Pablo Agüera, a UAV pilot from Soluciones Agrícolas de Precisión, S.L., for help and support during image acquisition. Additionally, the authors want to thank the Predoctoral Research Fellowship for the development of the University of Seville R&D&I program (IV.3 2017) granted to OEAA and the Torres-Quevedo contract (PTQ-17-09506) granted to JMG by the Spanish Ministry of Economy.

References

Aggelopoulou, A.D., Bochtis, D., Fountas, S., Swain, K.C., Gemtos, T.A., Nanos, G.D., 2011. Yield prediction in apple orchards based on image processing. Precis. Agric. 12, 448–456. https://doi.org/10.1007/s11119-010-9187-0.

Agustí Fonfría, M., 2012. Citricultura. Madrid. .

Beltrán-Esteve, M., Reig-Martínez, E., 2014. Comparing conventional and organic citrus grower efficiency in Spain. Agric. Syst. 129, 115–123. https://doi.org/10.1016/j.agsy.2014.05.014.

Blasco, J., Aleixos, N., Moltó, E., 2003. Machine vision system for automatic quality grading of fruit. Biosyst. Eng. 85, 415–423. https://doi.org/10.1016/S1537-5110(03)00088-6.

Bulanon, D.M., Burks, T.F., Alchanatis, V., 2009. Image fusion of visible and thermal images for fruit detection. Biosyst. Eng. 103, 12–22. https://doi.org/10.1016/j.biosystemseng.2009.02.009.

Burnett, C., Blaschke, T., 2003. A multi-scale segmentation/object relationship modelling methodology for landscape analysis. Ecol. Modell. 168, 233–249. https://doi.org/10.1016/S0304-3800(03)00139-X.

Cardeñosa, V., Medrano, E., Lorenzo, P., Sánchez-Guerrero, M.C., Cuevas, F., Pradas, I., Moreno-Rojas, J.M., 2015. Effects of salinity and nitrogen supply on the quality and health-related compounds of strawberry fruits (Fragaria × ananassa cv. Primoris). J. Sci. Food Agric. 95, 2924–2930.

Castro-Garcia, S., Blanco-Roldán, G.L., Ferguson, L., González-Sánchez, E.J., Gil-Ribes, J.A., 2017. Frequency response of late-season ‘Valencia’ orange to selective harvesting by vibration for juice industry. Biosyst. Eng. 155, 77–83. https://doi.org/10.1016/j.biosystemseng.2016.11.012.

Castro-Garcia, S., Aragon-Rodriguez, F., Sola-Guirado, R.R., Serrano, A.J., Soria-Olivas, E., Gil-Ribes, J.A., 2019. Vibration monitoring of the mechanical harvesting of citrus to improve fruit detachment efficiency. Sensors (Switzerland) 19, 1–10. https://doi.org/10.3390/s19081760.

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, 6. https://doi.org/10.3390/jimaging3010006.

Chinchuluun, R., Lee, W.S., Burks, T.F., Florida State Horticultural, S., 2006. Machine vision-based citrus yield mapping system. Proc. 119th Annu. Meet. Florida State Hortic. Soc. 119. pp. 142–147.

Csillik, O., Cherbini, J., Johnson, R., Lyons, A., Kelly, M., 2018. Identification of citrus trees from unmanned aerial vehicle imagery using convolutional neural networks. Drones 2, 39. https://doi.org/10.3390/drones2040039.

Dias, P.A., Tabb, A., Medeiros, H., 2018. Apple flower detection using deep convolutional networks. Comput. Ind. 99, 17–28. https://doi.org/10.1016/j.compind.2018.03.010.

Dorj, U.O., Lee, M., Yun, seok, S., 2017. An yield estimation in citrus orchards via fruit detection and counting using image processing. Comput. Electron. Agric. 140, 103–112. https://doi.org/10.1016/j.compag.2017.05.019.

Fischer, T., Krauss, C., 2018. Deep learning with long short-term memory networks for financial market predictions. Eur. J. Oper. Res. 270, 654–669. https://doi.org/10.1016/j.ejor.2017.11.054.

García-Tejero, I., Jiménez-Bocanegra, J.A., Martínez, G., Romero, R., Durán-Zuazo, V.H., Muriel-Fernández, J.L., 2010. Positive impact of regulated deficit irrigation on yield and fruit quality in a commercial citrus orchard [Citrus sinensis (L.) Osbeck, cv. salustiano]. Agric. Water Manag. 97, 614–622. https://doi.org/10.1016/j.agwat.2009.12.005.

Gong, A., Yu, J., He, Y., Qiu, Z., 2013. Citrus yield estimation based on images processed by an Android mobile phone. Biosyst. Eng. 115, 162–170. https://doi.org/10.1016/j.

biosystemseng.2013.03.009.
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, 498–503. https://doi.org/10.1016/j.inpa.2018.06.002.
Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., Liu, T., Wang, X., Wang, G., Cai, J., Chen, T., 2018. Recent advances in convolutional neural networks. Pattern Recognit. 77, 354–377. https://doi.org/10.1016/j.patcog.2017.10.013.
Guo, Y., Liu, Y., Oerlemans, A., Lao, S., Wu, S., Lew, M.S., 2016. Deep learning for visual understanding: a review. Neurocomputing 187, 27–48. https://doi.org/10.1016/j.neucom.2015.09.116.
Jiménez, A.R., Ceres, R., Pons, J.L., 2000. A survey of computer vision methods for locating fruit on trees. Trans. ASAE 43, 1911–1920. https://doi.org/10.13031/2013.3096.
Kamilaris, A., Prenafeta-Boldú, F.X., 2018. Deep learning in agriculture: a survey. Comput. Electron. Agric. 147, 70–90. https://doi.org/10.1016/j.compag.2018.02.016.
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., 2019. Deep learning for real-time fruit detection and orchard fruit load estimation: benchmarking of 'MangoYOLO'. Precis. Agric. https://doi.org/10.1007/s11119-019-09642-0.
Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. ImageNet classification with deep convolutional neural networks alex. Adv. Neural Inf. Process. Syst. 25, 1097–1105. https://doi.org/10.1061/(ASCE)GT.1943-5606.0001284.
Kurtulmus, F., Lee, W.S., Vardar, A., 2011. Green citrus detection using "eigenfruit", color and circular Gabor texture features under natural outdoor conditions. Comput. Electron. Agric. 78, 140–149. https://doi.org/10.1016/j.compag.2011.07.001.
Lado, J., Rodrigo, M.J., Zacarías, L., 2014. Maturity indicators and citrus fruit quality. Stewart Postharvest Rev. 10.
Lecun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521, 436–444. https://doi.org/10.1038/nature14539.
Liakos, K.G., Busato, P., Moshou, D., Pearson, S., Bochtis, D., 2018. Machine learning in agriculture: a review. Sensors (Switzerland) 18, 1–29. https://doi.org/10.3390/s18082674.
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. Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics) 8693, 740–755. https://doi.org/10.1007/978-3-319-10602-1_48. LNCS.
Lin, G., Tang, Y., Zou, X., Li, J., Xiong, J., 2019a. 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, G., Tang, Y., Zou, X., Xiong, J., Li, J., 2019b. Guava detection and pose estimation using a low-cost RGB-D sensor in the field. Sensors (Switzerland) 19, 1–15. https://doi.org/10.3390/s19020428.
Lin, G., Tang, Y., Zou, X., Cheng, J., Xiong, J., 2020. Fruit detection in natural environment using partial shape matching and probabilistic Hough transform. Precis. Agric. 21, 160–177. https://doi.org/10.1007/s11119-019-09662-w.
Ma, J., Li, Y., Chen, Y., Du, K., Zheng, F., Zhang, L., Sun, Z., 2019. Estimating above ground biomass of winter wheat at early growth stages using digital images and deep convolutional neural network. Eur. J. Agron. 103, 117–129. https://doi.org/10.1016/j.eja.2018.12.004.
Martínez, J., Egea, G., Agüera, J., Pérez-Ruiz, M., 2017. A cost-effective canopy temperature measurement system for precision agriculture: a case study on sugar beet. Precis. Agric. 18, 95–110. https://doi.org/10.1007/s11119-016-9470-9.
Moltó, E., Plá, F., Juste, F., 1992. Vision systems for the location of citrus fruit in a tree canopy. J. Agric. Eng. Res. 52, 101–110. https://doi.org/10.1016/0021-8634(92)80053-U.
Okamoto, H., Lee, W.S., 2009. Green citrus detection using hyperspectral imaging. Comput. Electron. Agric. 66, 201–208. https://doi.org/10.1016/j.compag.2009.02.004.
Rahnemoonfar, M., Sheppard, C., 2017. Deep count: fruit counting based on deep simulated learning. Sensors (Switzerland) 17, 1–12. https://doi.org/10.3390/s17040905.
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, 1137–1149. https://doi.org/10.1109/TPAMI.2016.2577031.
Rolim, G., de, S., de Oliveira Aparecido, L.E., Ribeiro, R.V., 2015. Estimation of orange yield in citrus orchard based on digital photography. Afr. J. Agric. Res. 10, 3086–3092. https://doi.org/10.5897/ajar2015.9776.
Rosebrock, A., 2016. Practical Python and OpenCV + Case Studies. PyImageSearch.com, Baltimore. .
Rosebrock, A., 2018. Deep Learning for Computer Vision with Python. ImageNet Bundle. Pyimagesearch.com. PyImageSearch.com, Baltimore. .
Sa, I., Ge, Z., Dayoub, F., Upcroft, B., Perez, T., McCool, C., 2016. Deepfruits: a fruit detection system using deep neural networks. Sensors (Switzerland) 16. https://doi.org/10.3390/s16081222.
Saravanan, C., 2010. Color image to grayscale image conversion. Second International Conference on Computer Engineering and Applications 1–4. https://doi.org/10.1109/ICCEA.2010.192.
Simonyan, K., Zisserman, A., 2015. Very deep convolutional networks for large-scale image recognition. Published as a Conference Paper at ICLR 2015 1–14.
Sonka, M., Hlavac, V., Boyle, R., 1993. Image Processing, Analysis and Machine Vision. https://doi.org/10.1007/978-1-4899-3216-7.
Talukdar, J., Gupta, S., Rajpura, P.S., Hegde, R.S., 2018. Transfer learning for object detection using state-of-the-art deep neural networks. 2018 5th Int. Conf. Signal Process. Integr. Networks, SPIN 2018 78–83. https://doi.org/10.1109/SPIN.2018.8474198.
Torregrosa, A., Ortí, E., Martín, B., Gil, J., Ortiz, C., 2009. Mechanical harvesting of oranges and mandarins in Spain. Biosyst. Eng. 104, 18–24. https://doi.org/10.1016/j.biosystemseng.2009.06.005.
Voulodimos, A., Doulamis, N., Doulamis, A., Protopapadakis, E., 2018. Deep learning for computer vision: a brief review. Comput. Intell. Neurosci. https://doi.org/10.1155/2018/7068349. 2018.
Wang, Walsh, Koirala, 2019. Mango fruit load estimation using a video based MangoYOLO—Kalman filter—Hungarian algorithm method. Sensors 19, 2742. https://doi.org/10.3390/s19122742.
Wolfert, S., Ge, L., Verdouw, C., Bogaardt, M.J., 2017. Big data in smart farming – a review. Agric. Syst. 153, 69–80. https://doi.org/10.1016/j.agsy.2017.01.023.
Yamamoto, K., Guo, W., Yoshioka, Y., Ninomiya, S., 2014. On plant detection of intact tomato fruits using image analysis and machine learning methods. Sensors 14, 12191–12206. https://doi.org/10.3390/s140712191.
Yin, X., Lan, Y., Wen, S., Zhang, J., Wu, S., 2018. Natural UAV tele-operation for agricultural application by using Kinect sensor. Int. J. Agric. Biol. Eng. 11, 173–178. https://doi.org/10.25165/ijabe.v11i4.4096.
Zhao, Y., Gong, L., Zhou, B., Huang, Y., Liu, C., 2016. Detecting tomatoes in greenhouse scenes by combining AdaBoost classifier and colour analysis. Biosyst. Eng. 148, 127–137. https://doi.org/10.1016/j.biosystemseng.2016.05.001.
Zhou, R., Damerow, L., Sun, Y., Blanke, M.M., 2012. Using colour features of cv. "Gala" apple fruits in an orchard in image processing to predict yield. Precis. Agric. 13, 568–580. https://doi.org/10.1007/s11119-012-9269-2.
Zhu, N., Liu, X., Liu, Z., Hu, K., Wang, Y., Tan, J., Huang, M., Zhu, Q., Ji, X., Jiang, Y., Guo, Y., 2018. Deep learning for smart agriculture: concepts, tools, applications, and opportunities. Int. J. Agric. Biol. Eng. 11, 21–28. https://doi.org/10.25165/j.ijabe.20181103.4475.
Zhuang, J.J., Luo, S.M., Hou, C.J., Tang, Y., He, Y., Xue, X.Y., 2018. Detection of orchard citrus fruits using a monocular machine vision-based method for automatic fruit picking applications. Comput. Electron. Agric. 152, 64–73. https://doi.org/10.1016/j.compag.2018.07.004.
Ziliani, M.G., Parkes, S.D., Hoteit, I., McCabe, M.F., 2018. Intra-season crop height variability at commercial farm scales using a fixed-wing UAV. Remote Sens. 10, 1–25. https://doi.org/10.3390/rs10122007.