agronomy

MDPI

*Article*

# A Methodology for the Automated Delineation of Crop Tree Crowns from UAV-Based Aerial Imagery by Means of Morphological Image Analysis

**Juan Manuel Ponce [1,*], Arturo Aquino [1], Diego Tejada [1], Basil Mohammed Al-Hadithi [2,3] and José Manuel Andújar [1]**

[1] Research Center in Technology, Energy and Sustainability (CITES), University of Huelva, Huelva-Palos de la Frontera Rd., Palos de la Frontera, 21819 Huelva, Spain; arturo.aquino@diesia.uhu.es (A.A.); diego.tejada@diesia.uhu.es (D.T.); andujar@diesia.uhu.es (J.M.A.)
[2] Centre for Automation and Robotics, UPM-CSIC. J., Gutierrez Abascal, 2., 28006 Madrid, Spain; basil.alhadithi@upm.es
[3] Department of Electrical, Electronics, Control Engineering and Applied Physics, School of Industrial Design and Engineering, Technical University of Madrid, Ronda de Valencia, 3., 28012 Madrid, Spain
* Correspondence: jmponce.real@diesia.uhu.es

**Abstract:** The popularisation of aerial remote sensing using unmanned aerial vehicles (UAV), has boosted the capacities of agronomists and researchers to offer farmers valuable data regarding the status of their crops. This paper describes a methodology for the automated detection and individual delineation of tree crowns in aerial representations of crop fields by means of image processing and analysis techniques, providing accurate information about plant population and canopy coverage in intensive-farming orchards with a row-based plant arrangement. To that end, after pre-processing initial aerial captures by means of photogrammetry and morphological image analysis, a resulting binary representation of the land plot surveyed is treated at connected component-level in order to separate overlapping tree crown projections. Then, those components are morphologically transformed into a set of seeds with which tree crowns are finally delineated, establishing the boundaries between them when they appear overlapped. This solution was tested on images from three different orchards, achieving semantic segmentations in which more than 94% of tree canopy-belonging pixels were correctly classified, and more than 98% of trees were successfully detected when assessing the methodology capacities for estimating the overall plant population. According to these results, the methodology represents a promising tool for automating the inventorying of plants and estimating individual tree-canopy coverage in intensive tree-based orchards.

**Keywords:** aerial imagery; canopy cover; morphological image analysis; crop tree; unmanned aerial vehicle (UAV)



**Citation:** Ponce, J.M.; Aquino, A.; Tejada, D.; Al-Hadithi, B.M.; Andújar, J.M. A Methodology for the Automated Delineation of Crop Tree Crowns from UAV-Based Aerial Imagery by Means of Morphological Image Analysis. *Agronomy* **2022**, *12*, 43. https://doi.org/10.3390/agronomy12010043

Academic Editor: Yanbo Huang

Received: 22 November 2021
Accepted: 22 December 2021
Published: 25 December 2021



## 1. Introduction

Modern agricultural practices developed around the precision agriculture (PA) paradigm, demand data-collecting systems for assembling information regarding the spatial and temporal variability of those factors of influence in agricultural production [1]. This data-driven approach is aimed at developing decision-making frameworks to help farmers with their daily tasks [2], thus supporting the eventual optimisation of the farming-inputs management and favouring the improvement of the overall agricultural activity in terms of crop productivity, sustainability, and profitability [3].

Hence, non-invasive data acquisition regarding crops condition has become of great interest to the agricultural sector [4,5]. Thus, extensive related research has been carried out on the development of solutions aimed at improving farming processes, by developing decision support models by means of remotely sensed data [6–11]. In this sense, aerial

remote sensing has had a pivotal role in agriculture over time. Indeed, its capacities for systematically assessing large tracts of land soon caught the attention of researchers, so the first studies on the use of aerial images in agriculture date back to the 1950s [12]. Notwithstanding, it is in the 1970s when their use began to stand out when satellite images started to become popular as data sources for agronomic researchers [1]. Initial applications, which focused mainly on the land cover classification [13,14] distinguishing between types of crops and vegetation present in the surveyed fields, soon became more complex as technology advances allowed higher spatial resolution and increased computational power, necessary for processing and analysing the collected data.

However, in recent years, there has been an increasing interest in aerial remote sensing technologies within modern agriculture, mostly due to the popularisation of unmanned aerial vehicles (UAV). Its advantages over other forms of aerial imaging, although debated, have caught the attention of farmers and agronomists, and UAV-based remote sensing has strengthened as a data-acquisition technology for PA [15–18], with a vast amount of related research published, that has consistently underlined its usefulness in agronomy, and its applicability in diverse related domains. Among them, UAV-based remotely sensed imagery has shown huge potential and applicability in crop field mapping and characterisation [19]. Low-altitude aerial imagery can offer some advantages when compared to satellite-based imaging, in terms of both spatial and temporal resolution, or when performing image acquisition under cloudy conditions [20,21]. By conveniently processing these high-spatial resolution data by means of diverse image analysis techniques, it can be yielded accurate representations of the fields surveyed [22], thus enabling the possibility to discriminate relevant information regarding the orography of the cultivation area, biological characteristics of the vegetation, dendrometric features of the canopies, or phenological plant status, among others [23–26].

Amongst the different issues that can be approached using UAV-based imagery, individual tree detection and characterization are of special interest. Indeed, accurate knowledge about plant population, and the geometric characteristics of the cultivation areas, regarding the spatial relationship between plants and their individual canopy-related features, can offer useful information for farmers. On the one hand, accurate data about tree populations can lead to meaningful statistics, such as crop density, directly involved in the development of early yield estimation models, or in the enhancement of watering/fertilization strategies [27,28]. Equally, geometric features of plant canopies can offer relevant indicators, also related to the yield capacities of the crop, or the water and nutritional status of the plants [29]. Due to their high-throughput within PA context, and given the difficulties in collecting these data in traditional manners, all mostly based on in-field visual inspections and costly labour-intensive manual measurements [30], the use of UAV-based imagery in conjunction with computer vision methodologies has gained relevance within the research community, seeking solutions for obtaining this type of information in the most autonomous and accurate way possible. Thus, Zortea et al. [31] proposed a method for detecting citrus-trees from aerial images collected with a quadcopter. The authors undertook the detection of problem processing those images by means of deep learning, with a double convolutional neural network (CNN) configuration, first identifying the tree rows, and then classifying the set of candidates of tree-occurrences computed. A similar approach was followed by Csillik et al. [32], using CNNs and simple linear iterative clustering (SLIC) for the purpose of detecting citrus-trees. In the same vein, Ampatzidis and Partel [33] also drew on deep learning techniques for automating citrus-tree detection from aerial imagery and additionally estimating different features of the plants identified, as the area of their canopy projections. Recio et al. [34] focused on citrus crops too, proposing an image analysis procedure which, from an initial k-means clustering-based pixel classification of the aerial captures, allowed individual plant identification and tree crown-canopy coverage estimations. Ok and Ozdarici-Ok [35] presented a method for identifying individual tree canopies, applicable in citrus tree-based crops. They proposed an ad-hoc orientation-based radial symmetry transform for detecting tree candidates from digital

surface models (DSM) of crop plots, and then individually delineating them by means of active contour models. Modica et al. [36] proposed a geographic object-based image analysis (GEOBIA) approach for segmenting crop vegetation from UAV-based multispectral images. To that end, they developed and compared different solutions based on four different machine-learning object-classification algorithms: K-nearest neighbour, support vector machines, random forests, and normal bayes, all tested in a citrus-tree orchard and in an onion field. Dong et al. [37] combined digital height models (DHM) with vegetation indices for individually segmenting apple and pear-tree crowns from aerial representations of cultivation areas. Likewise, Marques et al. [38] proposed the generation of a model based on elevation information, obtained from initial aerial captures, and the use of vegetation indices for detecting chestnut trees, and subsequently extracting different morphological characteristics from each of them. Salamí et al. [39] developed a framework for detecting and counting olive trees by processing aerial captures on-the-fly, using and comparing two different segmentation approaches: colour-based and stereo vision-based segmentations. Jiang et al. [40] used UAV imagery with the aim of automating the identification of papaya and lemon trees. To that end, they proposed a detection algorithm that processes images, previously transformed into the Lab colour space, on the basis of scale-space filtering (SSF). Johansen et al. [41] explored the impact of the height at which aerial imagery acquisition is performed when extracting geometrical characteristics of lychee-trees, for assessing the effects of pruning in the corresponding orchard.

Although extensive research has been carried out on this subject, there is still room for improvement, and some critical related issues have only been partially resolved. On one hand, most studies to date tend to focus on specific crops, and the generalisation of the proposed solutions and their usefulness with different types of crop trees remains uncertain. On the other hand, it can be appreciated a lack of research conducted in high density orchards, where plant arrangement is prone to the occurrences of overlapping canopies between adjacent trees, thus hindering the task of automatically detecting the plants at an individual level by means of computer vision techniques. Indeed, this issue could be of great interest given the trend towards intensification of agriculture in recent years, and the increasingly common adoption of super-intensive planting schemes [42–44].

Within this context, this paper presents a novel methodology for identifying crop trees in low-altitude UAV-based aerial images by means of image analysis techniques, with the ability to automatically delineate those tree crowns detected, thus enabling the estimation of features related to the size and morphology of each individual canopy. This solution is aimed at being applicable on high-density orchards with a regular row-based planting arrangement. Moreover, it was designed to be applicable to scenarios where plant arrangement is conducive to the occurrence of overlapping between the canopies of adjacent trees. Even in those situations in which the degree of overlapping was so extreme that lead to the loss of the shape of the tree crowns, in their corresponding pixel representations, the methodology showed outstanding performance. In addition, whereas most of the related research to date has tended to focus on specific crops, the developed methodology was designed to work independently of the type of trees under study. To this end, it was conceived dispensing with the use of vegetation indices, or any kind of machine learning-based approach, which would have to be modelled from any fixed training configuration, obtained from each specific crop. Therefore, a comprehensive framework based on morphological image analysis is proposed, which after transforming multispectral aerial images of an orchard according to the procedure proposed by Sarabia et al. [45], is capable of individually segmenting the pixel area of each crown projection, despite the presence of intra-row overlapping tree crowns, yielding estimations of the overall plant population as well as of individual tree canopy coverage.

## 2. Materials and Methods

### *2.1. Case Study Site*

The developed methodology has been tested with three different types of tree-based crops: olive (*Olea europaea* L.), lemon (*Citrus limon* L.), and orange (*Citrus sinesis* L.) trees. To that end, three different orchards, one per crop, were surveyed.

In the case of the olive trees, the grove under study is in Gibraleón, province of Huelva (Andalusia, Southwest Spain). This land plot, centred in DMS coordinates 7°02′48.44″ W and 37°20′39.80″ N, spreads over an area of approximately 17.5 ha, with a total of 3916 olive trees planted. On the other hand, the two citrus-tree orchards can also be found in this province, in the surroundings of the village of La Redondela, Isla Cristina. The lemon-tree orchard is located centred in the coordinates 7°17′06.93″ W and 37°14′07.24″ N, and it has an approximate extent of 1.37 ha and 552 trees planted. The orange-tree orchard can be found in the coordinates 7°18′06.75″ W and 37°13′49.55″ N, with an area of 1.94 ha and 781 trees. Figure 1 shows aerial images of these three orchards. It should be remarked that none of these aerial images properly belong to the present study, as they were taken from the Spanish National Geographic Institute ("Instituto Geográfico Nacional") [46]. They were not used at any moment throughout the development, nor the validation of the methodology described in this paper, so they are proposed just for illustrative purposes.


**Figure 1.** Aerial images of the case study sites: (**a**) lemon-tree orchard; (**b**) orange-tree orchard; (**c**) olive grove. Note in each case the corresponding plot under study enclosed in yellow.

It should be underscored that they all show a regular row-based planting pattern. In this respect, it is worth noting that in the case of the orange grove, there is a very high degree of overlap between intra-row adjacent trees. This is the reason it was selected for the study, with the aim of carrying out a sort of stress test with which to evaluate the developed methodology.

### *2.2. Image Acquisition Equipment*

Aerial images of the land plots under study were collected with a multispectral sensor MicaSense RedEdge-M™ (MicaSense, Inc., Seattle, WA, USA). This device is able to simultaneously capture in five different discrete spectral bands, and its main features are summarised in Table 1.

**Table 1.** Main features of the MicaSense RedEdge-M™ multispectral camera.

| Spectral Bands | Centre Wavelength (nm) | Bandwidth (nm) |
|---|---|---|
| Blue | 475 | 20 |
| Green | 560 | 20 |
| Red | 668 | 10 |
| Red-Edge | 717 | 10 |
| NIR | 840 | 40 |
| Ground Sample Distance (GSD) | 8 cm per pixel (per band) at 120 m AGL [1] | |
| Max Capture Rate | 1 capture per second (all bands), 12-bit RAW | |
| Field of View | 47.2° HFOV [2] | |
| Imager Resolution | 1280 × 960 pixels | |

[1] Above Ground Level. [2] Horizontal Field of View.

The sensing equipment was completed with a GPS module, directly connected to the camera for georeferencing the images as they were captured, and a 5-band downwelling light sensor (DLS). From the data collected by this latter sensor, it is possible to make corrections in flight regarding global lighting changes. In addition, a MicaSense Reflectance Panel was included in the setup. Captures of this panel provide accurate information regarding the amount of light reaching the ground when performing image acquisition, so these data are used for transforming raw pixels values to absolute reflectance when later processing the collected images.

As UAV, a DJI™ Matrice 100 (SZ DJI Technology Co., Ltd., Shenzhen, Guangdong, China) quadcopter was used for the olive case study. On the other hand, for the two remaining cases corresponding to the citrus-tree orchards, image acquisition was performed with a DJI™ Phantom 3 (SZ DJI Technology Co., Ltd., Shenzhen, Guangdong, China). Because of the dimensions of both orchards, this task could be carried out with a smaller UAV with lesser capabilities. The main characteristics of both platforms are listed in Table 2. The UAVs, along with the sensing equipment specified in Table 1, are also shown in Figure 2.

**Table 2.** Main features of the UAVs used for image acquisition.

| | DJI™ Matrice 100 | DJI™ Phantom 3 |
|---|---|---|
| Diagonal Wheelbase | 650 mm | 350 mm |
| Max Take-off Weight | 3600 g | 1280 g [1] |
| Max Speed | 17 m/s (GPS mode, no payload, no wind)<br>22 m/s (ATTI mode, no payload, no wind) | 16 m/s (ATTI mode, no payload, no wind) |
| Max Wind Resistance | 10 m/s | 10 m/s |
| Operating Temperature | 10° to 40° | 0° to 40° |

[1] Value not provided by the manufacturer. The indicated one corresponds to the weight of the aircraft, batteries and propellers included.


(**a**) (**b**)

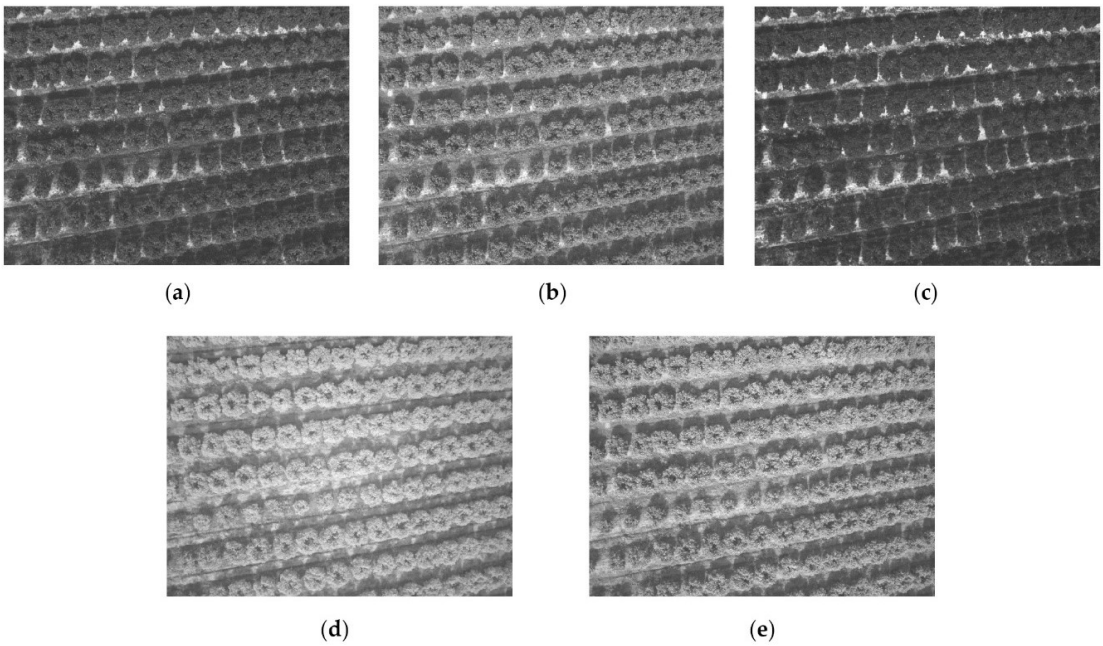
**Figure 2.** Aerial image acquisition equipment: (**a**) DJI Matrice 100-based setup; (**b**) DJI Phantom 3 Advance-based setup.

### *2.3. Flights Planning and Execution*

All those flights carried out for image acquisition were planned with the DJIFlightPlanner™ (AeroScientific-Spatial Scientific Pty. Ltd., Adelaide, Australia; Photometrix, Kew, Australia; Dechant Consulting Services Inc., Beaux Arts Village, WA, USA) software. They all shared the same configuration, aiming at capturing aerial images with a forward overlap of 85%, a lateral overlap of 65%, and a ground sample distance (GSD) of 4.8 cm. Flight missions were performed with a height of 70 m with regard to the take-off point, a cruising speed of 4.2 m/s, and time between captures of 1.5 s. They were executed autonomously in GPS mode and monitored with the ground control software Litchi (VC Technology, Ltd., London, UK).

The olive grove was flown on 13 June 2019, approximately between 11 a.m. and 1 p.m. 8865 images were taken per spectral band, yielding a total of 44,325 captures. In the case of citrus trees, two different flights were made, one for each type of crop. Both flights took place on 1 March 2020, in the same time slot as in the olive grove case.

On the other hand, a total of 1520 images, 304 per spectral band, were taken in the orange grove. 1580 (316 per spectral band) images were acquired in the case of the lemon-tree orchard. In Figure 3, some examples of images captured during those flights are shown. Specifically, they correspond to the orange-tree orchard.



(**a**) (**b**) (**c**)

(**d**) (**e**)

**Figure 3.** Examples of aerial images acquired for the case study corresponding to the orange-tree orchard, captured in different spectral bands: (**a**) Blue band; (**b**) Green band; (**c**) Red band; (**d**) NIR band; (**e**) Red-Edge band.

### *2.4. Developed Methodology*

#### 2.4.1. Image Preprocessing

In order to accomplish the main goal of the developed methodology, which is the individual delineation of the tree crowns boundaries in aerial images, it is necessary to transform initial captures into a data structure suitable for that purpose. The first pre-processing stage was carried out by the application of the procedure proposed by Sarabia et al. [45]. It puts forward a comprehensive method to automatically process multispectral aerial images of a land plot, with the objective of obtaining a unique binary image representative of it, where crop tree-belonging pixel regions are segmented from the ground, and location points of each individual plant are estimated. To that end, drawing from the set of aerial captures, an initial photogrammetric imaging procedure generates a 3D-point representation of the plot to survey. In this case, PIX4Dmapper$^{TM}$ (Pix4D, Prilly, Switzerland) was used as photogrammetry software for the purpose of yielding a high-density point cloud for each case study addressed. It should be noted that this software was configured such that each resulting 3D point could be correctly re-projected in a minimum of 3 different images, and no image scaling was applied. Once the 3D-point representations were obtained, corresponding digital height models (DHM) were computed by applying an inverse distance weighing (IDW) interpolation [47]. This task was conducted using ArcGis$^{TM}$ 10.3 (Esri, Inc., Redlands, CA, USA) along with its Geostatistical Analyst Tools extension. For this procedure, the software was configured for a fixed neighbourhood search, with a number of input neighbours up to 4, a search radius of 10 m, and a weighting exponent of 2. Then, each DHM was approached as a greyscale image. Height information at each point was treated as a pixel-intensity value, thus the models were processed by means of morphological analysis and statistical thresholding with the aim of segmenting those pixel-areas belonging to crop-tree canopies from the rest of that image. To that end, a background estimation was computed for each DHM-image, aiming at filtering each tree-crown projection by replacing the grey values of their corresponding pixels with the minimum intensity value found in the set of ground pixels present in their nearest surroundings. For that purpose, the process is carried out by iteratively applying a set of morphological opening operations [48], using a disk-shaped structuring element whose radius, $\beta$, is increased at each step, in an attempt to offer the optimum filtering for each tree canopy. The stop condition occurs when this kernel is large enough to completely contain the biggest tree-crown projection in the image. Thus, the maximum value to be reached by the radius of the structuring element corresponds to one of the two parameters that need to be set in order to apply this methodology, so it should be configured to guarantee that the corresponding structuring element is able to cover the largest individual canopy. The second required parameter relates to the height below which regional maxima of the image are discarded, assuming that tree crowns are represented at higher elevations, thus favouring their segmentation. Hence, it should be configured in accordance with the height expected for the tree crops. Table 3 shows the values of both configuration parameters for each case study addressed.

**Table 3.** Configuration of the parameter values of the preprocessing procedure, taken for each case study.

| Case Study | Filtering-Kernel Maximum Radius-Size, $r_\beta$ (pixels/cm [1]) | Relevant Maxima-Threshold, h (m) |
|---|---|---|
| Lemon-tree | 42.5/204 | 1.5 |
| Orange-tree | 55/264 | 1.1 |
| Olive | 70/336 | 1 |

[1] Based on a GSD of 4.8 cm.

After subtracting the background estimations from the corresponding grayscale representations of the DHMs, Otsu's method [49] was applied for eventually binarising those images, thus yielding the eventual canopy mass-ground segmentation. Finally, the estimation of the tree location points was undertaken by analysing the size of each of the connected

components in the resulting segmentations, with regard to the dominant tree crown size in the binary image. These connected components can be defined as those groupings of pixels with the same intensity value and a neighbourhood relationship (8-connectivity was considered in this study) [50], which, in this case, potentially represent the projections of tree canopies. In Figure 4, the workflow followed by this methodology is illustrated.

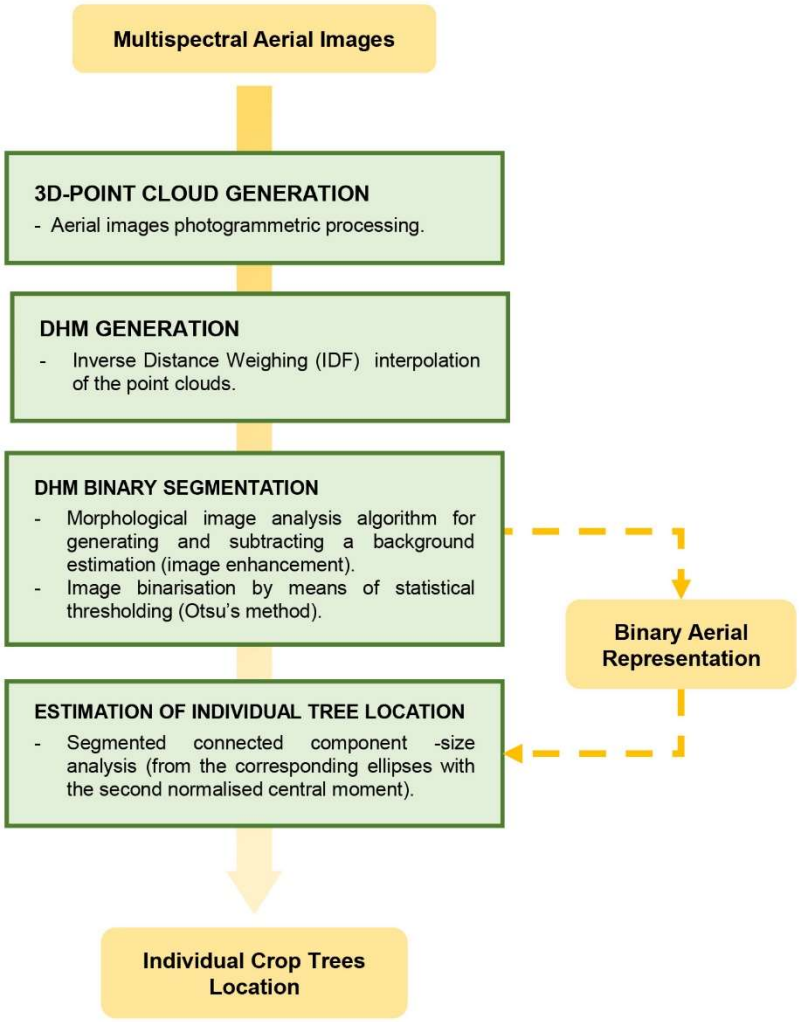



**Figure 4.** Flowchart of the methodology proposed by Sarabia et al. [45], aimed at generating a binary image from aerial multispectral captures of a land plot, where canopy coverage is shown segmented from the ground, and the set of location points of the crop trees in the plot is generated.

### 2.4.2. Image Processing

The starting point of the procedure here presented, which comprises the core of the methodology, is a binary image representative of the orchard understudy, and where tree canopies are segmented from the ground, along with a set of potential location points for the individual crop plants in that image. Figure 5 shows, as an example, the binary image resulting from pre-processing the aerial captures collected in one of the study sites considered in this work. Additionally, the estimated tree location points are represented on it.

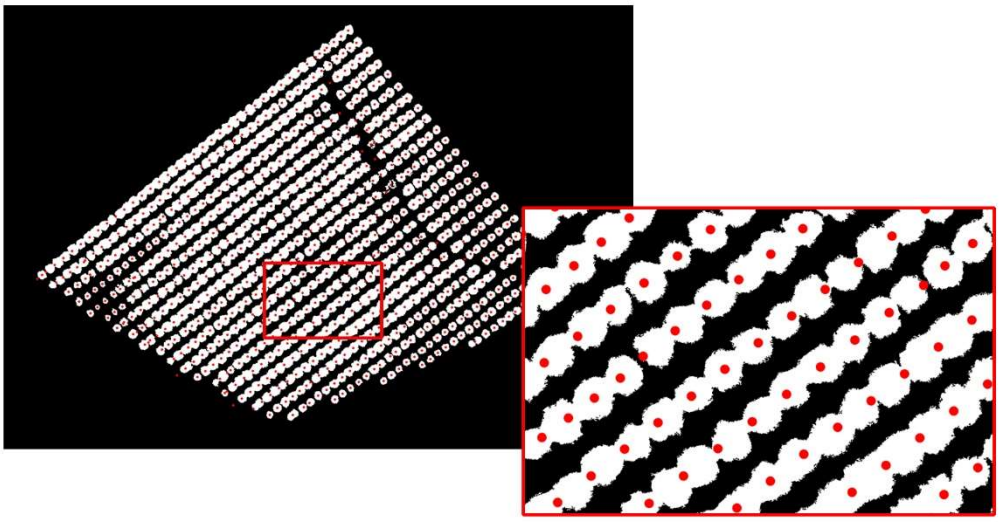


**Figure 5.** Binary image resulting from pre-processing aerial captures of the orange grove-case study, according to the methodology proposed by Sarabia et al. [45], represented together with the estimated tree location points (red circles). Note in the zoomed area, highlighted in red, the occurrences of overlapping tree crowns.

Hence, given both pieces of information, the general idea behind the developed methodology goes through the individual processing of each of the connected components. It focuses on transforming those components into small disjoint groupings of pixels, somehow representative of each individual tree contained within them. These sorts of pixel-seeds are aimed at backing the eventual splitting of the connected component according to the plants contained. Hence, they are used to support the final individual tree coverage segmentation, by a Watershed transform-based operation. This process of transforming and segmenting the connected components is mainly based on the application of different morphological operators, which are formally defined in Appendix A, so the reader is encouraged to consult it for further information. The whole developed methodology that is presented in detail hereafter is briefly described in the flowchart shown in Figure 6.

Given the binary image $I_{BIN}$ yielded by the preprocessing, each of its connected components is individually processed, as mentioned above. First, it is checked the number of trees in the component, estimated during preprocessing. Given a connected component of $I_{BIN}$, $cc_i$, if the estimated number of trees contained in it, $nT_i$, is just one, it is not necessary to conduct any processing at all. When $nT_i$ is greater than one, its corresponding pixel region is initially isolated in an alternative intensity matrix, embedding it in the smallest all-black-pixel box that can contain it. For this auxiliary image, its Euclidean distance transform [51] is calculated. Thus, given this binary image $I_{CCi}$, which encloses $cc_i$ in its proper bounding box, its distance transform, $I_{DTi}$, is calculated in accordance with the following definitions:

$$I_{DTi} = DT(I_{CCi}), \tag{1}$$

$$[DT(I_{CCi})](p) = \min\{d_\varepsilon(p,q) | f(q) = 0\}, \tag{2}$$

$$d_\varepsilon(p,q) = \sqrt{(p_x - q_x)^2 + (p_y - q_y)^2} \tag{3}$$

where, given the binary image $f$ and the pixel $p$, *[DT(f)(p)]* refers to the Euclidean distance from $p$ to its closest background pixel. The resulting matrix, $I_{DTi}$, can be seen and subsequently treated as a greyscale image, as illustrated in Figure 7.

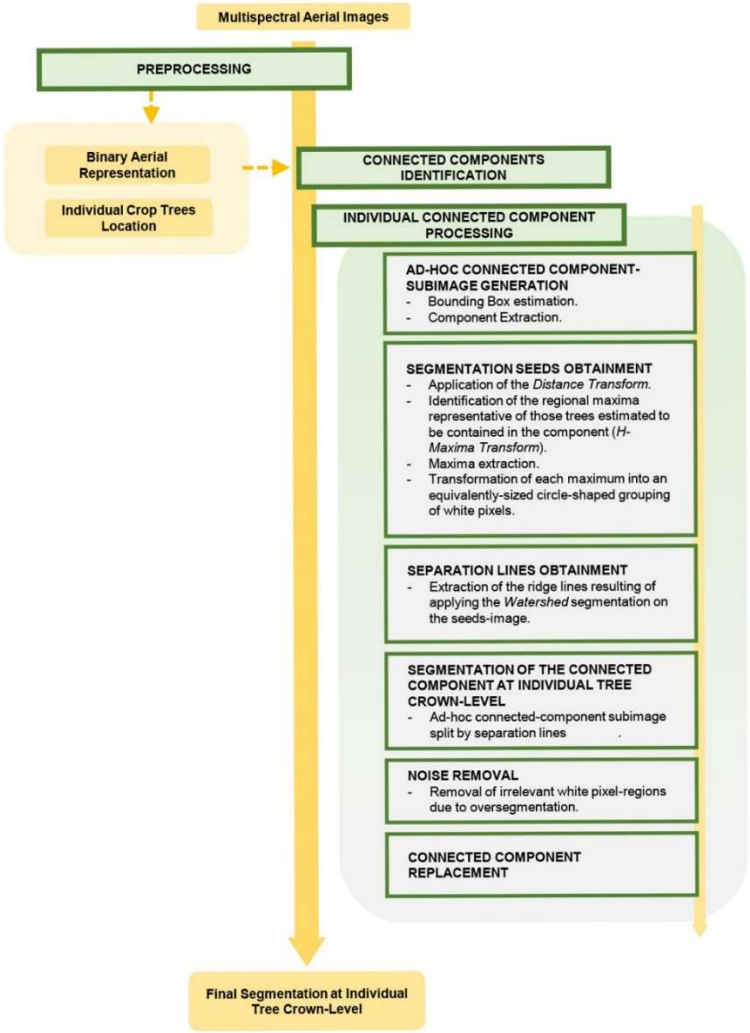



**Figure 6.** Flowchart of the morphological image analysis algorithm for individually segmenting the crowns of the trees in the image aimed at processing.

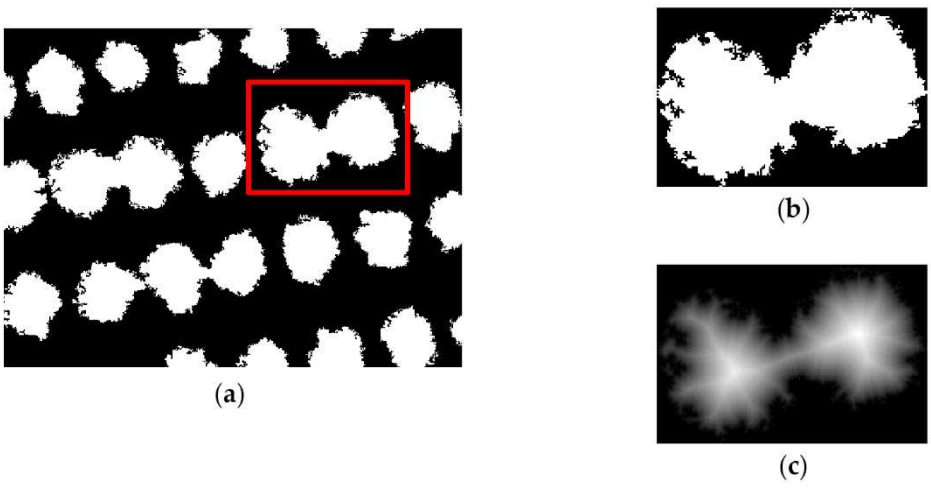



**Figure 7.** (**a**) Sub-image of the initial binary image, $I_{BIN}$; (**b**) the resulting image of enclosing the connected component highlighted in red in (**a**) in its corresponding bounding box; (**c**) intensity matrix, $I_{DTi}$, resulting from the computing of the Euclidean distance transform of the binary image represented in (**b**).

Subsequently, the regional maxima of the resulting intensity matrix $I_{DTi}$ are explored as tree representative. The goal here is to isolate those more representative regional maxima, as they can be expected to take place where the centres of mass of the trees within the component are potentially located. This is carried out on the basis of the *h-maxima* transform, which allows us to remove from $I_{DTi}$ all maxima with depth lower or equal than scalar $h$. It can be computed as the morphological reconstruction by dilation ($R^{\delta}$) of $I_{DTi}$ from the marker image resulting from subtracting $h$ from the intensity value of each pixel in $I_{DTi}$. Thus, irrelevant maxima are firstly removed from $I_{DTi}$ by computing:

$$I_{DTi,h} = HMAX_h(I_{DTi}) = R^{\delta}_{I_{DTi}}(I_{DTi} - h) \quad (4)$$

The surviving regional maxima in $I_{DTi,h}$ are those considered representatives and, therefore, are extracted from the image. It is achieved by subtracting from $I_{DTi,h}$ the result of suppressing from it all the remaining regional maxima:

$$I_{RMAXi,h} = I_{DTi,h} - R^{\delta}_{I_{DTi,h}}(I_{DTi,h} - 1) \quad (5)$$

Note that the rightmost term of the subtraction is the *h-maxima* transform setting the value of $h$ to 1. This value enables the removal of all remaining regional maxima from $I_{DTi,h}$, as all of them will have, at least, that height. Then, by thresholding $I_{RMAXi,h}$, it can be achieved a binary mask containing those maxima representing the centre of mass of the trees:

$$[I_{RMAXBINi,h}](p) = \begin{cases} 255 \; if \; I_{RMAXi,h}(p) > 0 \\ 0 \quad in \; any \; other \; case \end{cases} \quad (6)$$

The result of this process depends on the definition of the $h$ value in Equation (4). The irregularity of the shape of trees, which can be corroborated in Figure 7, may force the appearance of peripheral regional maxima of reduced height. To avoid the selection of any of these non-representative maxima, the procedure described in Equations (4)–(6) is iteratively performed for increasing $h$ values starting from 1; being the value selected that fulfilling to be the maximum $h$ value, such that the number of segmented maxima contained in $I_{RMAXBINi,h}$ is equal to the number of estimated trees $nT_i$:

$$I_{MAXi} = I_{RMAXBINi,h} \left| \#CC(I_{RMAXBINi,h}) = nT_i \wedge \#CC(I_{RMAXBINi,h+1}) < nT_i \right. \quad (7)$$

where for a given binary image $\boldsymbol{f}$, *#CC(f)* refers to the number of connected components existing in it. Figure 8 shows the maxima extracted at different $h$ values for the connected component previously proposed in Figure 7.

(**a**) (**b**) (**c**) (**d**) (**e**)

**Figure 8.** Maxima extracted from the distance-transform image corresponding to the connected component proposed in Figure 7, computed at different heights *h*: (**a**) $h$ = 1; (**b**) $h$ = 5; (**c**) $h$ = 10; (**d**) h = 15; (**e**) $h$ = 16. Note how the proper maxima are obtained for a height of 15, since the number of trees previously estimated for the component is 2, and this value for $h$ is the highest with which it is achieved a maxima segmentation with 2 connected components.

From this binary representation of the maxima, $I_{MAXi}$, an ad-hoc image is automatically generated, and eventually used as a marker image to perform the individual segmentation of each tree crown within the initial component $cc_i$. This new image, named hereafter $I_{SEEDSi}$, results from transforming $I_{MAXi}$ such that each of its components is replaced by an all-white pixels-circle with the same area. This is, for each connected component belonging

to $I_{MAXi}$, a corresponding circle-shaped connected component can be found in $I_{SEEDSi}$, with the same pixel-area and centre of mass. Figure 9 shows an example of this ad-hoc image, obtained in this specific case from the initially connected component presented in Figure 7.

(**a**) (**b**)

**Figure 9.** Segmentation markers for finally splitting the connected component proposed in Figure 7: (**a**) binary image corresponding to the representative maxima of those trees enclosed in the component, $I_{MAXi}$; (**b**) binary image generated from (**a**), where each connected component is replaced by a disc-shaped pixel-region with the same area and location, and which comprises the final segmentation markers.

As stated before, $I_{SEEDSi}$ contains the set of markers with which the initially connected component $cc_i$ is finally segmented at the individual tree crown level. $I_{SEEDSi}$ has been generated according to the hypothesis that the isolated maxima, from which it is obtained, are somehow representative of the size and positions of the overlapped treetops in the component. Therefore, the goal at this point is the achievement of an artifact of separation between the area of influence of each circle-shaped pixel-region, assuming that it will subsequently provide a proper segmentation for the tree crowns. This approach will be exploited in the application of the Watershed transform. Indeed, given a greyscale image, this segmentation tool approaches the intensity levels as values of altitude, so the darkest regions can be seen as catchment basins. By "flooding" them, the lines of water convergence can be detected. These ridgelines are determined by the morphological features of those catchment basins, such as their shape or depth, so they can adequately split the zones of influence of the basins. Hence, the idea is to approach the circled areas as those catchment basins by employing the distance transform formulated in Equation (2). Notwithstanding, in this case, the pixel-regions corresponding to the circles are expected to be the darkest (the deepest) areas in the transformed image, so the distance transform must be redefined to compute here the distance of each pixel to the closest white one. Thus, given $I_{SEEDSi}$, the grayscale image $I_{DTSEEDSi}$, over which the Watershed segmentation is subsequently performed, can be mathematically defined as follows:

$$I_{DTSEEDSi} = DT'(I_{SEEDSi}), \tag{8}$$

$$\left[DT'(f)\right](p) = \min\{d_\varepsilon(p,q) | f(q) = 1\}, \tag{9}$$

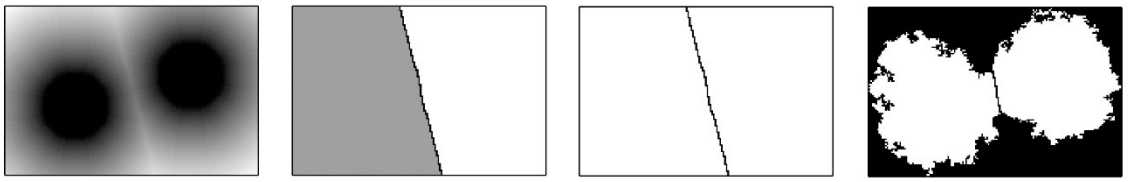
Next, the Watershed transform is applied on $I_{DTSEEDSi}$. The mathematical formulation of this morphological transform can be quite extensive, so the reader is encouraged to consult related literature already published for further study [52,53]. In this sense, it should be underscored that the Watershed transform implementation used in this study is based on an algorithm proposed in [54]. Thus, after its application on $I_{DTSEEDSi}$, those ridges establishing the boundaries of the basins can be extracted as follows:

$$[I_{RIDGESi}](p) = \begin{cases} 0 \; if \; [WS(I_{DTSEEDSi})](p) = 0 \\ 1 \; in \; any \; other \; case \end{cases} \tag{10}$$

where $WS$ denotes the Watershed transform. Finally, the binary image $I_{RIDGESi}$ is used to separate the tree crown projections contained in the initially connected component $cc_i$:

$$I_{FINALi} = \min\{I_{CCi}, I_{RIDGESi}\} \tag{11}$$

Figure 10 illustrates this final stage of segmentation based on the Watershed transform.



(**a**) (**b**) (**c**) (**d**)

**Figure 10.** Final segmentation at individual tree crown level of the connected component $cc_i$: (**a**) distance transform of $I_{SEEDSi}$, based on the Euclidean distance of each pixel to the closest white one; (**b**) Watershed transform of (**a**); (**c**) ridge line obtained from the Watershed transform-image (**b**); (**d**) the image resulting from splitting the initial component $cc_i$ with the ridge line in (**c**). It should be noted that, for the sake of facilitating its visualization, the image display range has been modified for subimages (**a**,**b**), conveniently increasing their contrast.

Each connected component $cc_i$ containing multiple trees in the original binary image $I_{BIN}$, is replaced by those in its corresponding processed sub-image $I_{FINALi}$, resulting in a final binary image, $I_{SEG}$, where every tree crown appears individually segmented.

#### 2.4.3. Image Postprocessing

When undertaking the last stage of segmentation, whereby initially connected components are properly split according to the ridgelines previously yielded, over segmentation phenomena can occur [55–57]. Indeed, given the irregularity in terms of the silhouette of the tree crowns, it is possible for non-significant parts of a plant to be outside the boundaries established by the Watershed transform for its corresponding projection. This would potentially provoke the generation of small connected components wrongly disconnected (according to the 8-connectivity considered) from the main ones representing tree crowns (see Figure 1a). These anomalous pixel-regions are removed by applying a morphological opening followed by a morphological reconstruction to exactly recover the surviving greater connected components corresponding to tree crowns:

$$I_{FINALPOSTi} = R^{\delta}_{I_{FINALi}}(\gamma_{\beta}(I_{FINALi})) \tag{12}$$

where $R^{\delta}$ refers to the reconstruction by dilation, and $\gamma_{\beta}$ is the morphological opening by a disc-shaped structuring element $\beta$. The size of $\beta$ must be large enough to remove the anomalous small connected components, but also smaller than tree crowns to preserve them and enable their reconstruction. The considerable difference in terms of size between the anomalous and tree-crown components meant this decision was not critical, due to a wide range of values being valid. In this context, the radius of $\beta$ was set to 5.

The explained postprocessing aimed at removing irrelevant components is illustrated in Figure 11.

### *2.5. Evaluation of the Developed Methodology*

As mentioned in Section 2.1, the developed methodology was tested on images acquired in three different crop fields. The approach chosen for assessing its performance was based on the comparison of the binary images resulting from each case study with handmade binary segmentation of the tree crowns. These latter were obtained from ad-hoc images resulting from orthomosaics representatives of each land plot under study, which were in turn generated from the aerial captures collected during the image acquisition stage. Indeed, given the set of multispectral aerial captures collected for each case study, they were processed with photogrammetry software (Pix4D$^{TM}$ Mapper), in order to achieve high-resolution aerial representations of the corresponding land plots. Resulting in five different orthomosaics after processing, one per spectral band of capture, three of them (blue, red-edge, and near-IR) were conveniently combined to finally yield a representative

aerial image of the crop, wherein it became rather easy to differentiate the individual tree crowns for a human observer. Then, by using an image editor, a manual delineation of each tree crown was carried out, producing a binary image which, as a gold-standard, was eventually compared at pixel level with the corresponding segmentation provided by the developed methodology. Figure 12 shows a proposal of the colour images generated for each case study.


**Figure 11.** Postprocessing aimed at filtering anomalous small connected components generated after segmentation: (**a**) initial subimage $I_{CCi}$ of the connected component to be segmented; (**b**) ridge line-image $I_{RIDGESi}$ obtained with the Watershed transform for segmenting (**a**); (**c**) resulting segmentation after splitting the component in (**a**) using (**b**) -note in the zoomed area the generation of anomalous disconnected artifacts; (**d**) segmentation-image $I_{FINALPOSTi}$ obtained after postprocessing-note in the zoomed area that all elements are connected, considering 8-connectivity, and how the shape of the crowns are not affected as a consequence of the postprocessing.

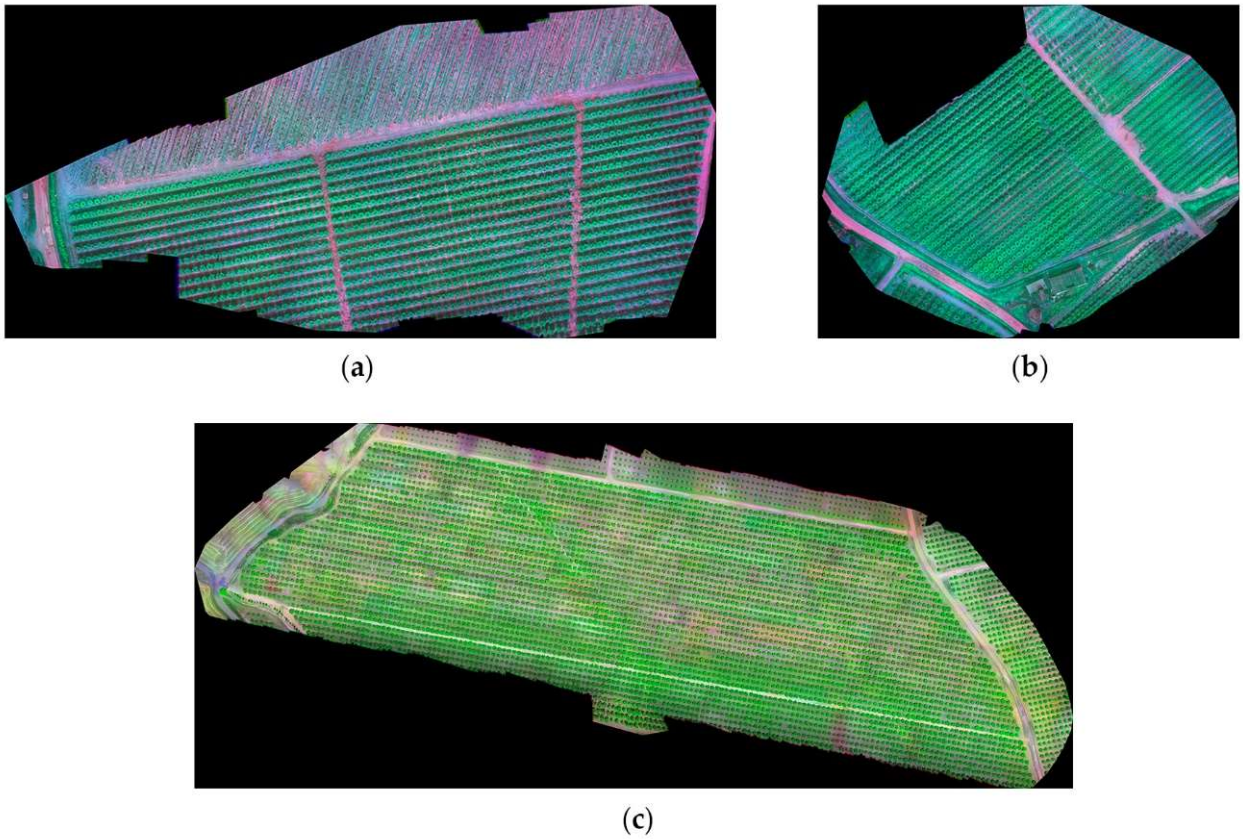



**Figure 12.** Ad-hoc colour representations of each case study, obtained from the combination of the Blue, Red-Edge, and NIR spectral bands, and used to generate the binary ground-truth images. (**a**) Lemon-tree orchard; (**b**) orange-tree orchard; (**c**) olive grove.

It should be noted that any other combination of the available orthomosaics may be valid for this purpose. In this sense, the sole objective was to obtain an image comfortable to work with when editing it for manually segmenting the tree crowns using an image editor.

Hence, for each case study, given the binary image resulting from applying the developed methodology, and its corresponding ground-truth image, every pixel in the first one was labelled in accordance with the next four categories:

- True Positive (*TP*): that foreground pixel (white pixel) for which its analogous one in the ground-truth image was categorised as tree crown-belonging pixel.
- False Positive (*FP*): that foreground pixel which was labelled as a non-tree crown-belonging pixel in the ground-truth image.
- True Negative (*TN*): that background pixel (black pixel) such that the corresponding one in the ground-truth image was categorised as a non-tree crown-belonging pixel.
- False Negative (*FN*): that background pixel for which its analogous one was labelled as tree crown-belonging pixel in the ground-truth image.

Based on these pixel categories, the following metrics are proposed for the purpose of assessing the accuracy of the segmentation resulting from the application of the developed methodology:

- Precision (*PR*): this metric refers to the probability with which a given foreground pixel was correctly categorised. It can be formulated as follows:

$$PR = \frac{tp}{tp + fp} \tag{13}$$

where *tp* and *fp* are, respectively, the number of pixels categorised as *TP* and the number of them labelled as *FP*.

- Recall (*RC*): it represents the ratio between the number of foreground pixels correctly classified and the whole set of instances of actual tree-belonging pixels in the image. Mathematically:

$$RC = \frac{tp}{tp + fn} \tag{14}$$

where *fn* is the number of pixels labelled as *FN*.

- $F_{\text{-}score}$: as the harmonic mean of these two metrics just proposed:

$$F_{-score} = 2 \times \frac{PR \times RC}{PR + RC} \tag{15}$$

- Overall Accuracy (*OA*): it proposes the percentage of pixels correctly classified.

$$OA = \frac{tp + tn}{tp + tn + fp + fn} \tag{16}$$

- Intersection-over-Union (*IoU*): also known as the Jaccard index, this metric signifies the similarity between the resulting segmentation and its corresponding ground-truth image. It can be defined with the following expression:

$$IoU = \frac{|I_{RESUL} \cap I_{GT}|}{|I_{RESUL} \cup I_{GT}|} \tag{17}$$

where $I_{RESUL}$ and $I_{GT}$ are the final segmentation and its binary ground-truth image, respectively. They should be approached as the sets of their corresponding foreground instances, so this metric is computed as the ratio between the number of elements in their intersection and the cardinality of their union.

In addition, a secondary analysis of the results was undertaken, focused on the capacities of the proposed solution for individual tree crown identification, by comparing

the connected components yielded in the resulting binary images and the actual tree crowns, represented throughout the connected components in the corresponding ground-truth image. The categories defined below can be instantiated on the basis of this comparison:

- True Positive, at tree-level ($TP_t$): that connected component in the final segmentation corresponding to an actual tree crown in the binary ground-truth image.
- False Positive, at tree-level ($FP_t$): that connected component in the final segmentation for which it cannot be found an analogous component, representative of a tree crown, in the ground-truth image.
- False Negative, at tree-level ($FN_t$): that actual tree crown in the ground-truth segmentation not represented in the final segmentation. In other words, those tree crowns are not detected by the algorithm.

Thus, metrics equivalent to those proposed in Equations (13)–(15) were computed for the defined cases.

On the other hand, it should be noted that methodology was not only tested according to the parametrisation proposed in Section 2.4.1, but also with different configurations regarding the maximum size of the filtering kernel used during preprocessing, in order to evaluate the influence of this value on its performance.

## 3. Results

As mentioned in Section 2.5, the assessment of the developed methodology was approached in two ways: first, analysing at pixel level the semantic segmentation provided by the final binary images; next, by quantifying its performance in terms of individual tree crown detection. All this, for each of the three different cases considered for testing.

Regarding the capacities of the developed methodology for providing accurate segmentation of individual tree crowns in a binary representation of the field plot, Table 4 summarises the results calculated for the metrics proposed for that purpose.

**Table 4.** Accuracy of the semantic segmentation provided by the developed methodology.

| Case Study | *PR* [1] | *RC* [2] | $F_{\text{-score}}$ [3] | *OA* [4] | *IoU* [5] |
|---|---|---|---|---|---|
| Lemon-tree | 0.91993 | 0.89750 | 0.90858 | 0.98986 | 0.83247 |
| Orange-tree | 0.93169 | 0.94280 | 0.93721 | 0.98291 | 0.88185 |
| Olive | 0.88918 | 0.98249 | 0.93351 | 0.98185 | 0.87530 |

[1] Precision; $PR = tp/(tp + fp)$. [2] Recall; $RC = tp/(tp + fn)$. [3] $F_{\text{-score}} = 2 \times (PR \times RC)/(PR + RC)$. [4] Overall Accuracy; $OA = (tp + tn)/(tp + tn + fp + fn)$. [5] Intersection-over-Union; $IoU = |I_{RESUL} \cap I_{GT}| / |I_{RESUL} \cup I_{GT}|$.

As it can be observed in Table 4, the results obtained are rather similar in the three cases tested, achieving $F_{\text{-score}}$ values over 90% for all of them. Assuming that this latter measure along with *IoU* may be the most significant metrics to quantify the performance of the developed methodology in terms of pixel classification, the overall lower result corresponded to the lemon-tree orchard case study. Curiously, because of its characteristics regarding plant spatial arrangement and the average tree crown size, the occurrences of overlapping tree crowns were significantly lower in this case. This is in contrast to the orange grove case, where a high level of overlapping between adjacent canopies implied the processing of a considerable number of large connected components, in order to split the individual plant pixel-regions which appeared fused. Indeed, this case study was considered for stress testing because of the extreme conditions of its planting arrangement. However, and despite that, a subtle better accuracy was achieved in the segmentation when compared with the test performed for the lemon-tree orchard, yielding promising results when delineating individual tree crowns in very large, connected components, as can be appreciated in Figure 13. Notwithstanding, the differences are so insignificant for the three crops that they are hardly attributable to the developed methodology, and they may be related to the ground-truth images. The generation of those ground-truth images was performed by hand, involving an arduous task in which it was often difficult to visually differentiate the tree crowns from other kinds of vegetation when manually delineated by a human

observer. In this sense, given the nature of the images under study and the complexity of this process, it is not possible to guarantee that the different case studies were represented with the same degree of accuracy by their corresponding ground-truth images.

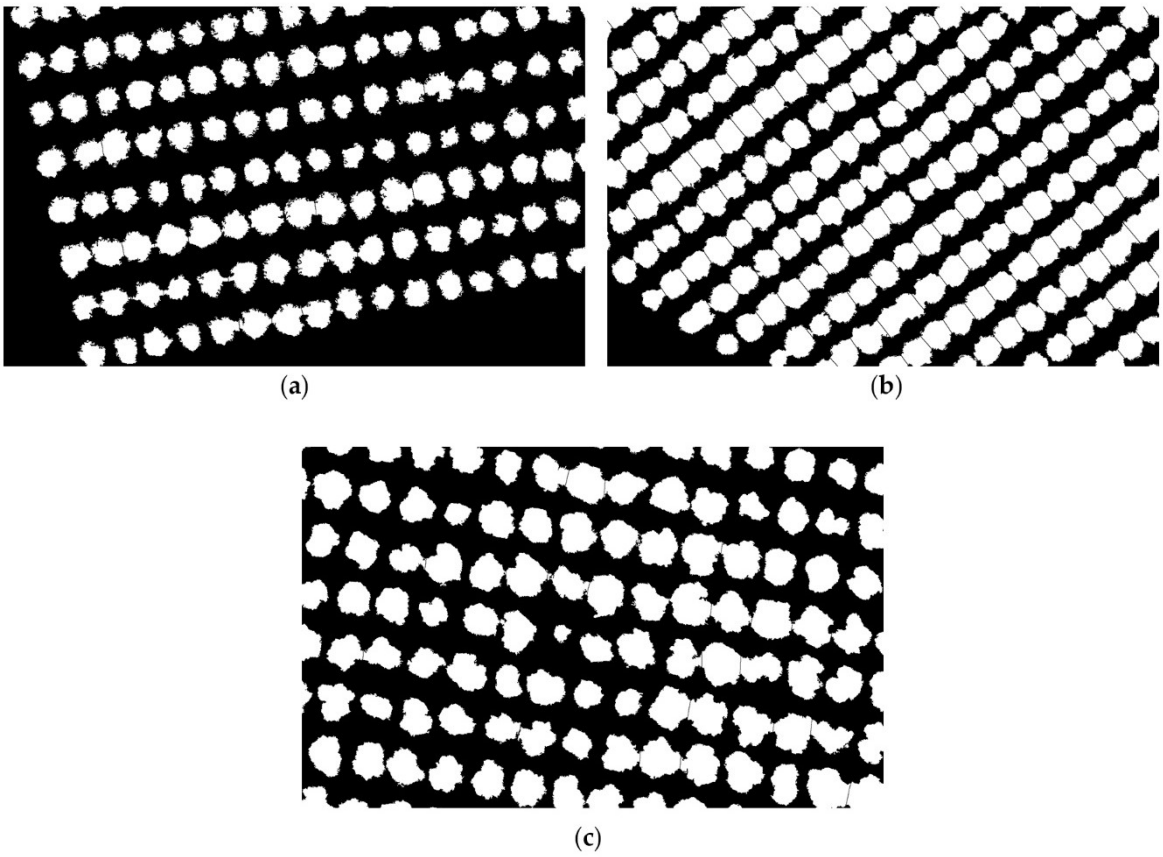



**Figure 13.** Sub-images of the final segmentations obtained for each considered case study: (**a**) lemon-tree orchard; (**b**) orange-tree orchard; (**c**) olive grove.

On the other hand, Table 5 presents the performance of the developed methodology when detecting individual tree crowns projections, and subsequently providing an estimation of the plant population. As can be corroborated, the performance is consistent for the three crops studied, showing only slight deviations.

**Table 5.** Results for the assessment of the developed methodology in terms of individual crop tree identification.

| Case Study | $tp_t$ [1] | $fp_t$ [2] | $fn_t$ [3] | $PR_t$ [4] | $RC_t$ [5] | $F_{\text{-score}t}$ [6] |
|---|---|---|---|---|---|---|
| Lemon-tree | 548 | 0 | 4 | 1 | 0.99275 | 0.99636 |
| Orange-tree | 758 | 2 | 23 | 0.99737 | 0.97055 | 0.98377 |
| Olive | 3906 | 5 | 10 | 0.99872 | 0.99744 | 0.99808 |

[1] Number of CCs categorised as $TP_t$. [2] Number of CCs categorised as $FP_t$. [3] Number of CCs categorised as $FN_t$. [4] Precision; $PR_t = tp_t/(tp_t + fp_t)$. [5] Recall; $RC_t = tp_t/(tp_t + fn_t)$. [6] $F_{\text{-score}t} = 2 \times (PR \times RC)/(PR + RC)$.

At this point, it should be remembered that the initial estimation of the trees embedded in the connected components is realised on the basis of a dominant tree crown size within the whole population of tree projections in the binary representation to process. This reference measure, obtained during preprocessing, is used as a basis for partitioning the major diameter of each component, yielding such an estimation. Specifically, this value is obtained from the initial binary image representative of the orchard, computing for each connected component the ellipse with its same normalised second central moment [58]; it is

taken from the whole set of minor axis lengths of all the ellipses obtained, as the maximum one. This reference value is increased by 20% before being used for considering certain tolerance. Finally, the number of trees in a given connected component is determined by dividing its major axis length by this value. There are some related phenomena that may occur during this subprocess, that can condition the result of the individual tree crown delineation. Thus, small tree crowns that are actually overlapped may be detected as unique plants. It should be emphasised that these small tree-aggregations happen very occasionally, and mostly at the end of the planting rows, where for some reason there is a certain tendency to plant the trees closer to each other. Likewise, regular adjacent tree crowns with a high degree of canopy overlapping may exhibit this problem. This may provoke an underestimation, which is subsequently translated to the eventual individual segmentation, as the number of predicted trees is used as a threshold when retrieving relevant maxima during the procedure. Consequently, individual tree crown-pixel regions may not be properly split when underestimating the number of crop trees in a given connected component. This will cause an increase in false negatives at the tree level ($FN_t$). Furthermore, the opposite phenomenon can occur, i.e., the overestimation of the number of potential tree projections embedded in each component. This may happen mostly with large connected components, with occurrences of rather big individual tree canopies. This overestimation involves the generation of a larger number of segmentation seeds than should be expected, and a subsequent oversegmentation of the corresponding connected components, which in turn results in the appearance of new false positives ($FP_t$). Both issues are illustrated in Figure 14.

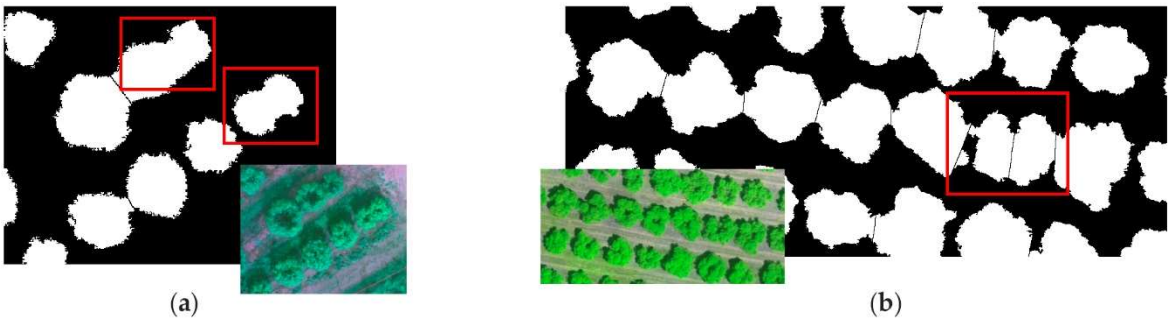



**Figure 14.** Problems with the individual tree crown delineation: (**a**) non-segmentation of adjacent tree-crown projections due to underestimation of the number of trees contained in the connected component; (**b**) oversegmentation of adjacent tree-crown projections due to overestimation of the number of trees contained in the connected component. Note highlighted in red those subcomponents involved with the issues reported.

Nevertheless, prior estimations of the trees contained in each connected component were highly and comparably accurate in all the three case studies, thus limiting the specific occurrences of the above-mentioned segmentation problems, and the overall performance of the methodology may be concluded as highly positive.

Finally, Table 6 collects results of applying the whole procedure for different sizes of the filtering-kernel radius, all in a range of plus/minus 30 pixels with respect to the initial parameter-value fixed for each case study (Table 3), for the purpose of assessing its sensibility and impact on the overall performance.

**Table 6.** Results of the developed methodology in terms of accuracy of semantic segmentation and individual crop-tree identification, regarding the maximum radius size of the preprocessing filtering kernel, $r_\beta$.

| | | Semantic Segmentation Accuracy | | | | | Individual Tree Detection | | |
|---|---|---|---|---|---|---|---|---|---|
| **Case Study** | $r_\beta$ **(pixels)** | $PR$ [1] | $RC$ [2] | $F_{\text{-score}}$ [3] | $OA$ [4] | $IoU$ [5] | $PR_t$ [6] | $RC_t$ [7] | $F_{\text{-score}t}$ [8] |
| Lemon-tree | 12.5 | 0.90357 | 0.30466 | 0.45567 | 0.95915 | 0.29506 | 0.64498 | 0.86232 | 0.73798 |
| | 22.5 | 0.90498 | 0.56947 | 0.69905 | 0.97248 | 0.53734 | 0.75211 | 0.96739 | 0.84627 |
| | 32.5 | 0.92030 | 0.89636 | 0.90817 | 0.98983 | 0.83179 | 1 | 0.99275 | 0.99636 |
| | 42.5 | 0.91993 | 0.8975 | 0.90858 | 0.98986 | 0.83247 | 1 | 0.99275 | 0.99636 |
| | 52.5 | 0.91993 | 0.89751 | 0.90858 | 0.98986 | 0.83248 | 1 | 0.99275 | 0.99636 |
| | 62.5 | 0.91980 | 0.89767 | 0.90860 | 0.98986 | 0.83250 | 1 | 0.99275 | 0.99636 |
| | 72.5 | 0.91965 | 0.89775 | 0.90857 | 0.98986 | 0.83245 | 1 | 0.99275 | 0.99636 |
| Orange-tree | 25 | 0.89027 | 0.3977 | 0.54980 | 0.91187 | 0.37912 | 0.60945 | 0.94110 | 0.73981 |
| | 35 | 0.91334 | 0.56325 | 0.69679 | 0.93367 | 0.53467 | 0.96741 | 0.95006 | 0.95866 |
| | 45 | 0.93170 | 0.92755 | 0.92962 | 0.98100 | 0.86850 | 0.99217 | 0.97311 | 0.98255 |
| | 55 | 0.93169 | 0.94280 | 0.93721 | 0.98291 | 0.88185 | 0.99737 | 0.97055 | 0.98378 |
| | 65 | 0.93167 | 0.94284 | 0.93722 | 0.98291 | 0.88186 | 0.99737 | 0.97055 | 0.98378 |
| | 75 | 0.93152 | 0.94335 | 0.93740 | 0.98295 | 0.88217 | 0.99737 | 0.97055 | 0.98378 |
| | 85 | 0.93245 | 0.94131 | 0.93686 | 0.98283 | 0.88121 | 0.99607 | 0.97311 | 0.98446 |
| Olive | 40 | 0.89326 | 0.92540 | 0.90905 | 0.97599 | 0.83326 | 0.96850 | 0.99719 | 0.98264 |
| | 50 | 0.89090 | 0.98005 | 0.93335 | 0.98185 | 0.87503 | 0.99872 | 0.99719 | 0.99796 |
| | 60 | 0.88918 | 0.98249 | 0.93351 | 0.98185 | 0.87530 | 0.99847 | 0.99745 | 0.99796 |
| | 70 | 0.88918 | 0.98249 | 0.93351 | 0.98185 | 0.87530 | 0.99872 | 0.99745 | 0.99808 |
| | 80 | 0.88868 | 0.98274 | 0.93335 | 0.98180 | 0.87502 | 0.99872 | 0.99770 | 0.99821 |
| | 90 | 0.88810 | 0.98304 | 0.93316 | 0.98174 | 0.87470 | 0.99923 | 0.99796 | 0.99859 |
| | 100 | 0.88823 | 0.98296 | 0.93320 | 0.98175 | 0.87476 | 0.99923 | 0.99770 | 0.99847 |

[1] $PR = tp/(tp + fp)$. [2] $RC = tp/(tp + fn)$. [3] $F_1 = 2 \times (PR \times RC)/(PR + RC)$. [4] $OA = (tp + tn)/(tp + tn + fp + fn)$. [5] $IIoU = |I_{RESUL} \cap I_{GT}| / |I_{RESUL} \cup I_{GT}|$. [6] $PR_t = tp_t/(tp_t + fp_t)$. [7] $RC_t = tp_t/(tp_t + fn_t)$. [8] $F_{\text{-score}t} = 2 \times (PR_t \times RC_t)/(PR_t + RC_t)$.

As commented in Section 2.4.1, the maximum size to be reached by the structuring element used for image filtering should ideally be the smallest possible to cover the largest tree crown in the image. However, this parameter is manually selected, and estimating an appropriate value for it can be complex, especially when the number of trees under study is large. As expected, the poorest results were obtained when underestimating the maximum size of the kernel. Indeed, if the largest kernel is not able to cover some of the individual canopies, these may not be properly filtered, so resulting residual maxima, being then part of the background estimation, are wrongly subtracted from the initial DHM intensity matrix, thus affecting the subsequent canopy-ground segmentation, and consequently the overall performance of the methodology. The accuracy of the segmentation at pixel level is compromised by wrongly discarding those regions of the canopies that could not be properly filtered out, increasing the number of pixels classified as FN. On the other hand, tree crowns can hardly be uniquely represented by a single connected component due to oversegmentation, impoverishing the capabilities of the proposed solution to automate plant detection as the number of false positives ($FP_t$) explodes. Indeed, some additional tests were carried out, using extremely small values as the limit for the kernel radius, with the aim of confirming this issue. Thus, for the olive case study, the methodology was applied with a maximum radius of 10 pixels, resulting in a final segmentation with a total number of 8795 connected components. Given an actual population of 3916 trees in the olive groove, this segmentation would reach 4879 $FP_t$ in the best-case scenario, implying a very poor performance in terms of individual tree identification. For the purpose of illustrating this issue, Figure 15 shows a canopy-ground segmentation of the orange-tree orchard, carried out with a maximum size for the filtering kernel too small to satisfy the requirements of the methodology (15 pixels), as an extreme example of the problems arising from underestimating this parameter, where it can be visualised some of the issues just commented.

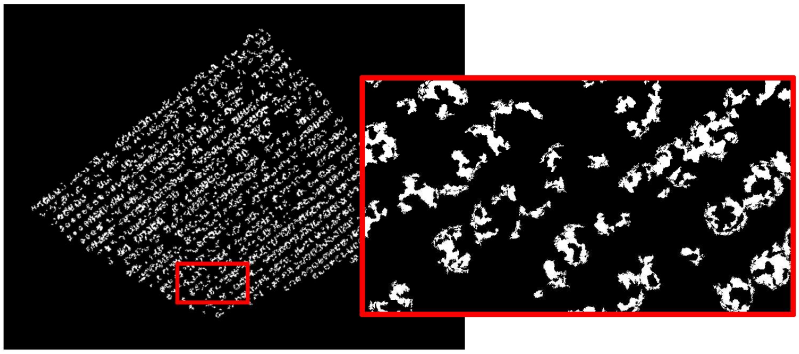


**Figure 15.** Resulting segmentation for the orange-tree case study, executed with a maximum size of 15 pixels for the radius of the filtering kernel used during preprocessing for estimating the background of the corresponding DHM-greyscale image.

Contrary, when overestimating the maximum size of the kernel it was not observed significant impoverishment of the results. In fact, even when performing stress tests with extremely large radius values, in the order of several hundred pixels, the results remained within similar ranges. However, caution must be exercised in drawing conclusions in this respect, as it may be possible that other case studies with important differences in terms of ground elevation or significant variability of trees' height would yield inadequate results if too large structuring elements are used. In general, based on the results reported, there is a sweet spot for this parameter, which it may be found around the average radius of the population of tree-crowns projections under study. The criterion suggested by the authors of the procedure, based on reaching a size large enough to cover the largest of the trees, seems appropriate, given that according to the results, the performance of the procedure is mainly compromised when the estimation of the background is carried out with filtering structuring elements that are too small, and not necessarily when this parameter is overestimated.

## 4. Discussion

Reported results regarding the assessment of the developed methodology, suggest its viability as a tool for the automated counting and delineation of individual tree canopy projections in aerial representations. Furthermore, its performance was comparable in the three cultivars tested, reinforcing its potential applicability in cultivation areas regardless of the type of crop. Notwithstanding, despite these compelling results, there are some points that are noteworthy.

In accordance with results reported, the number of undetected trees after individual canopy delineation is quite low, even in the orange orchard-case study, in which it can be observed large connected components where the pixel subregions to be split appear deeply fused, comprising a scenario prone to the underestimation of the tree crowns enclosed in those components during preprocessing; and by extension, to the occurrence of false negatives ($FN_t$). Despite this, it must not be overlooked that any inaccuracy of the initial estimation of the number of trees contained in each connected component, which is performed at preprocessing stage, may affect the quality of the final segmentation. In this vein and as already mentioned, the number of trees in the aerial binary representation is estimated on the basis of a reference value during this preprocessing stage. This value, after automatically computed, is used to rather split the major diameter of each connected component predicting the number of plants contained. The obtention of this reference value might be approached in different ways for the purpose of slightly improving the overall performance. Indeed, different possibilities were considered during the experimentation process, as individual reference values were calculated from each connected component instead of a global one, but no tangible improvements were observed with any of the alternatives proposed. That said, despite there being little room for improvement in this

sense, further research may focus on this topic, exploring new ways to predict the potential number of tree crowns embedded in the components to split.

With regard to the generation of ad-hoc segmentation seeds representatives of each of the pixel subregions to be split, there are also some aspects to be discussed. Indeed, in a near-ideal scenario, with a high level of homogeneity in terms of form and size of the tree crowns and a regular spacing pattern between plants, the location points provided by the methodology used during preprocessing, approached as single white pixels, could properly work as representative of the centre of mass of the corresponding tree projections. They could subsequently be used as proper segmentation seeds from which to determine the boundaries of each subcomponent, thus making it unnecessary to calculate ad-hoc segmentation seeds, as representative artifacts of the spatial influence of each plant within each component. However, this can hardly be expected in a real setting, and occurrences of adjacent overlapping tree crowns with unbalanced sizes are usual. In those cases, the corresponding estimated location points may appear displaced from the hypothetical centres of mass of the tree projections. This is because of the way they are calculated from the aforementioned global reference value, somehow assuming every crop tree has the same size within a given plot. Hence, they cannot properly define the pixel-region influenced by each tree, and subsequently, they cannot ensure an accurate individual tree crown segmentation.

Another issue to be addressed is related to the fact that the solution here proposed does not use vegetation indices, nor any kind of machine learning technology. As already mentioned, one of the main goals pursued during the research has been the applicability of the methodology regardless of the type of tree-based crop to be studied. In this sense, machine learning approaches are based on the development of classification models from training sets, which in turn are fed with representative data of the visual features of the elements to be identified; in this case, tree canopies. So, there is a dependency between the classification models achieved and the kinds of plants with which the corresponding training sets have been generated, thus restricting the scope of application, and hence compromising the generality of the solution. Notwithstanding, this does not invalidate these kinds of approaches. Indeed, proposals based on this form of artificial intelligence for detecting and delineating tree crowns from aerial representations are quite common and effective, always under the assumption that their replicability with different types of crops may be limited. Thus, Csillik et al. [32] detected citrus-trees from multispectral imagery with *precision* and *recall* of 94.59% and 97.94%, respectively. Another example of the use of CNNs for individual tree identification is proposed by Ampatzidis and Partel [33]. They undertook the detection of citrus-trees by means of CNN, achieving a *precision* of 99.90% and a recall of 99.70%, and overall accuracy of 85.5% when segmenting the canopy cover in the grove tested. In this latter work, after tree detection, an alternative approach for estimating the individual canopy projections is proposed based on the normalized difference vegetation index (NDVI). In fact, vegetation indices are a common tool for extracting plant features, included dendrometric characteristics, as it is the case. Thus, Marques et al. [38] used a CIR (Colour-InfraRed) imagery-based vegetation index along with elevation data for individually detecting chestnut-trees, reporting a precision of 99.44%, a recall of 97.80%, and an $F_{\text{-score}}$ of 98.61%, and segmenting the canopy coverage with accuracies from 93% to 99% in all case studies addressed. Despite being proven effective in certain scenarios, this kind of approach, based on vegetation indices, can be penalised in terms of performance when a high presence of weed occurs, as their spectral reflectance signature is hardly distinguishable from that of the plants to identify [59–62]. Indeed, for the purpose of illustrating this issue, vegetation index-based segmentations for two of the case studies addressed in this work are proposed in Figure 16. For each citrus-tree orchard, its corresponding NDVI information, computed from the initial multispectral images, is presented as a grey-scale image, along with the binary representation resulting from the attempt to segment the tree-belonging canopies by exploiting this information. It should be

noted that this segmentation has been approached by using statistical thresholding [49] with the corresponding NDVI data.

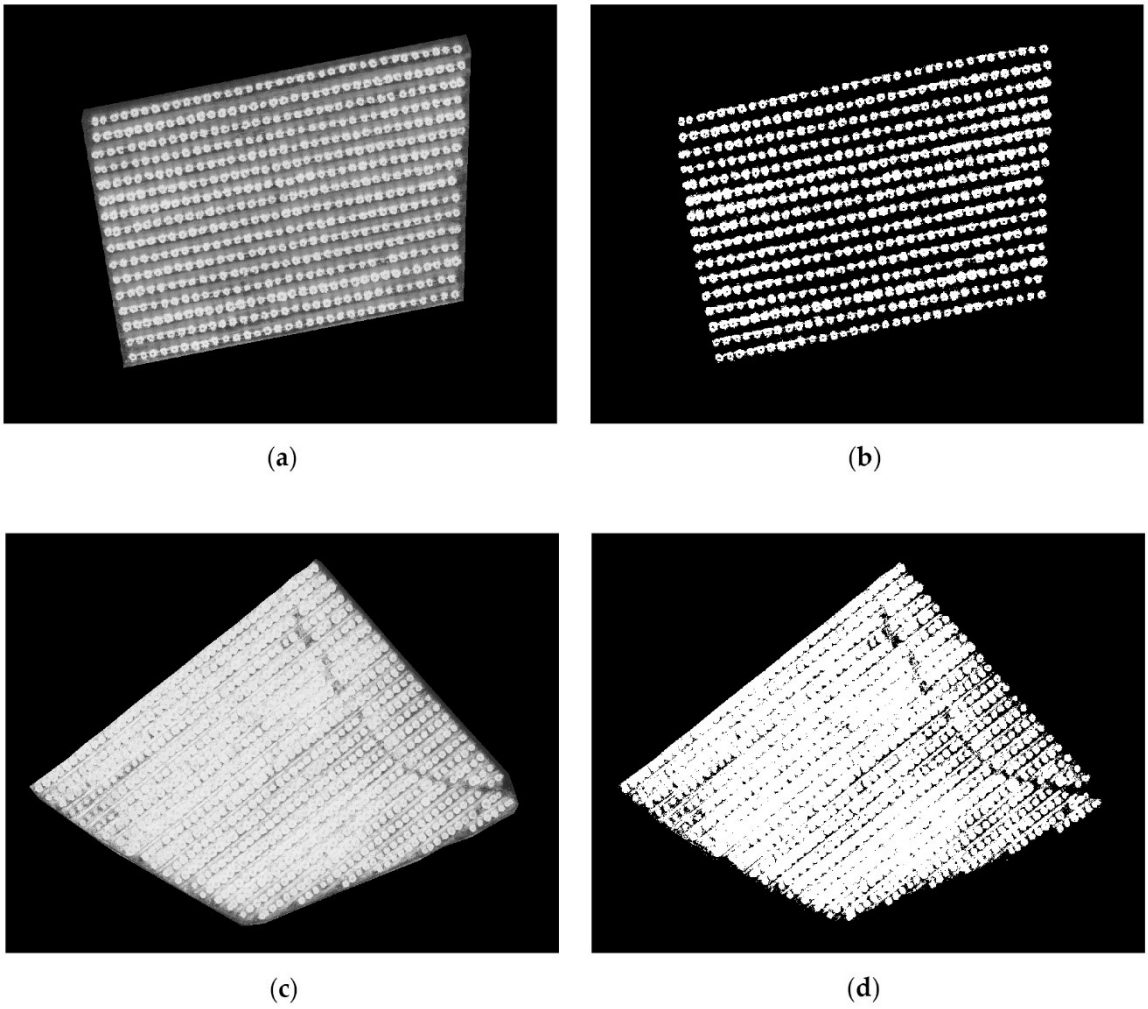


**Figure 16.** NDVI-based segmentation of the canopy coverage in the citrus-tree orchards: (**a**) lemon-tree orchard NDVI data, represented as a greyscale image; (**b**) binary representation of the lemon-tree orchard resulting from NDVI-based segmentation; (**c**) orange-tree orchard NDVI data, represented as a greyscale image; (**d**) binary representation of the orange-tree orchard resulting from NDVI-based segmentation.

As can be seen, the result obtained in the case of the lemon-tree orchard is more than acceptable. In fact, this segmentation could potentially be used as a basis for applying the methodology proposed here to delineate each tree crown individually. However, this is not the case for the orange grove, where the presence of a considerable amount of weeds, with a spectral response apparently similar to that offered by the orange trees, makes it difficult to discriminate between tree-canopy projections and soil. This does not disqualify the use vegetative indices within this area of research, since in this case it may be achieved another combination of spectral reflectance measurements that could improve segmentation results for the orange-tree orchard; but it does exemplify the problems that could be involved in using this type of indicators. On this basis, and because of the initial premise of yielding a methodology applicable without any adjustment, in as many diverse scenarios as possible, the use of any kind of vegetation indices was discarded in this work.

In fact, despite the use of vegetation indices and machine learning techniques may be considered as a trend in this field, several studies have addressed the problem of discriminating individual tree-canopies by dispensing this kind of approach, and still reporting significant results. Thus, as previously introduced, Salamí et al. [39] proposed

colour and stereo-vision -based segmentations to solve the automated detection of olive trees from aerial captures. They reported an $F_{\text{-score}}$ of 99.9% when detecting trees from a dataset with a total of 332 plants. Notwithstanding, it should be noted that the method was tested in olive orchards where plant arrangement guaranteed some spacing between trees, so the problem of overlapping adjacent canopies was not addressed. On the other hand, the study proposed by Ok and Ozdarici-Ok [35] may be of particular interest, given that their proposal was aimed at individually segmenting citrus-tree crowns, by means of morphological image analysis, and it also faced the overlapping canopies issue. After evaluating their methodology, they reported a precision of 91.1%, and a recall of 91.3% when assessing its performance at the pixel level. Positive results were also reported when evaluating the quality of the solution in terms of object detection, achieving an $F_{\text{-score}}$ value of 91.2%. In order to establish comparisons with this study, as well as with other works previously discussed, it can be underlined that the methodology here proposed achieved overall precision and recall values of 91.3% and 94.1%, respectively, when assessing the semantic segmentations in terms of pixel classification. Likewise, tests conducted for evaluating individual tree detection outperformed them yielding an $F_{\text{-score}}$ of 99.27%. In addition, it should be remarked that, contrary to most of the related published literature, the proposed methodology was conceived to be applicable as a generic solution regardless of the type of crop. Hence, it was tested with three different types of crop trees, obtaining promising results in all the threes case studies addressed.

## 5. Conclusions

In this paper, an image analysis methodology for the processing of aerial multispectral sensed data collected in crop tree-based intensive cultivation areas has been proposed, generating binary representations of them, and automatically delineating each tree crown that appears, thus providing individual plant canopy cover information.

Different case studies were set out to evaluate the effectiveness of the methodology, and to assess its viability as a solution independent from the type of tree-based crop cultivated, facing diverse scenarios in terms of the variability in the size of the tree population, or the characteristics of the plant's spatial arrangement. In all cases, including the borderline test realised in an orange-tree grove which presented challenging large intra-row canopy aggregations, results provided accurate segmentations at individual tree crown levels. Based on these results, along with the performance shown in terms of individual tree detection, the framework proposed here can be of broad use to farmers and agronomists for inventorying and monitoring crop trees in cultivated land plots.

Further research could be undertaken to investigate different approaches aimed at detecting and treating those exceptional cases reported, responsible for specific situations in which the under/overestimation of trees embedded in those connected components to process occurs. In addition, despite the fact that the methodology was tested with three different types of tree-based crops, it is assumable that undertaking new case studies focused on other additional crops would reinforce the applicability of the proposed solution, as independent of the kinds of trees farmed.

**Author Contributions:** Conceptualization, J.M.P. and A.A.; methodology, J.M.P. and A.A.; software, J.M.P., A.A. and B.M.A.-H.; validation, J.M.P., A.A., D.T. and B.M.A.-H.; formal analysis, J.M.P. and A.A.; investigation, J.M.P. and A.A.; resources, J.M.A.; data curation, J.M.P., D.T. and B.M.A.-H.; writing—original draft preparation, J.M.P.; writing—review and editing, A.A. and J.M.A.; visualization, J.M.P. and A.A.; supervision, J.M.A.; project administration, J.M.A.; funding acquisition, J.M.A. All authors have read and agreed to the published version of the manuscript.

**Funding:** Grant P18-RTJ-4539 funded by the Regional Government of Andalusia and by "PAIDI, Plan Andaluz de Investigación, Desarrollo e Innovación".

**Acknowledgments:** Authors would like to thank "Virgen de la Oliva" olive-oil cooperative for generously offering their orchards to conduct this work.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Appendix A

Mathematical morphology (MM) is a well-established technique for the nonlinear analysis of spatial structures, with remarkable applicability in image processing. Thus, morphological image analysis is rather based on the basic idea of transforming the image to process by its interaction with a simpler structure, called structuring element (SE) and for which its shape and size are known a-priori, resulting in more expressive images according to certain aimed characterization. The morphological operators used in the methodology proposed here are formally described below. In addition, literature related to morphological image analysis can be consulted for deeper study [48,63].

Let $f$ be a greyscale image representing a mapping from a subset $D_f$ of $\mathbb{Z}^2$, which defines its domain, to a bounded subset of nonnegative integers $N_0$:

$$f : D_f \subset \mathbb{Z}^2 \rightarrow N_0 = \{0, \ldots, t_{max}\} \subset \mathbb{Z}. \tag{A1}$$

where $t_{max}$ is the maximum reachable value according to the type of data used (e.g., 1 for binary images, 256 for 8-bit images, etc.). Hence, $f$ maps the correspondence by element between two sets, the first being composed of spatially ordered elements $\rho$ (pixels), $\rho \in D_f$ and denoted by a pair of coordinates $(x, y)$, while the second is built with an ordered set of possible values.

With the previous definitions, the intersection of two greyscale images, $f$ and $g$, is defined as:

$$f(\rho) \wedge g(\rho) = min[f(\rho), g(\rho)]. \tag{A2}$$

being $min$ the minimum operation. The union of those two images can be conversely defined as follows:

$$f(\rho) \vee g(\rho) = max[f(\rho), g(\rho)], \tag{A3}$$

where $max$ denotes the maximum operation.

As already mentioned, morphological operations basically consist of probing the image to process with a certain SE. Mathematically, a SE element can be seen as a binary image $\beta$, defining a mapping from a subset $D_\beta$ of $\mathbb{Z}^2$ to the subset of integer binary values $B_0$:

$$\beta : D_\beta \subset \mathbb{Z}^2 \rightarrow B_0 = \{0, 1\} \subset \mathbb{Z}. \tag{A4}$$

Therefore, $\beta$ maps the correspondence between the spatially ordered pixels $\rho$, $\rho \in D_\beta$ and referenced by a pair of coordinates $(x, y)$, and their values. This mapping must be designed so as to morphologically describe the object to be analysed, with $\#(D_\beta) < \#\left(D_f\right)$ being necessary for its application. Common shapes implemented with SEs include circles, lines, diamonds, etc. In practice, SE is used as a kernel, with its origin in its central pixel. Hence, an image is probed pixel by pixel with this kernel, modifying the pixel in the image at every step, matching with the central pixel of the kernel, according to a given operation.

The morphological erosion of image $f$ by an SE $\beta$ centred in pixel $\rho$, is given by the following expression:

$$\left[\varepsilon_\beta(f)\right](\rho) = min\{f(\rho + b) | b \in D_\beta\}. \tag{A5}$$

As a result, pixel $\rho$ in image $f$ is modified with the minimum value of its neighbourhood according to the filter implemented by SE $\beta$. Therefore, the effect of erosion is the expansion of darker regions, conditioned by the shape defined in SE.

The dual operator of erosion is dilation. The morphological dilation of image $f$ by a SE $\beta$, centred in pixel $\rho$, is formulated as:

$$\left[\delta_\beta(f)\right](\rho) = max\{f(\rho + b) | b \in D_\beta\}. \tag{A6}$$

Conversely, dilation expands brighter regions in $f$ according to the morphology of SE.

Combining erosion and dilation, two new operators called opening ($\gamma$) and closing ($\varphi$) can be defined:

$$\gamma_\beta(f) = \delta_\beta(\varepsilon_\beta(f)), \tag{A7}$$

$$\varphi_\beta(f) = \varepsilon_\beta(\delta_\beta(f)). \tag{A8}$$

Opening removes those brighter objects in the image that can be completely covered by $\beta$. Inversely, closing operation removes those darker objects in the image which are completely covered by this SE.

The operators described are complemented by geodesic transformations. The geodesic dilation is the iterative dilation of an image $f$, called marker, with respect to the mask image $g$, using a unitary SE. Marker $f$ must be contained within mask $g$. Mathematically, the operator is defined as:

$$\text{where : } \begin{array}{c} \delta_g^{(n)}(f) = \delta_g^{(1)}\left[\delta_g^{(n-1)}(f)\right], \text{ being } \delta_g^{(1)}(f) = \delta(f) \wedge g, \\ \#\left(D_f\right) = \#(D_g), \text{ and } f(\rho) \leq g(\rho),\ \forall \rho \in D_f, D_g. \end{array} \tag{A9}$$

Conversely, the geodesic erosion of image marker $f$ constrained by mask $g$ is:

$$\text{where : } \begin{array}{c} \varepsilon_g^{(n)}(f) = \varepsilon_g^{(1)}\left[\varepsilon_g^{(n-1)}(f)\right], \text{ being } \varepsilon_g^{(1)}(f) = \varepsilon(f) \vee g, \\ \#\left(D_f\right) = \#(D_g), \text{ and } f(\rho) \geq g(\rho),\ \forall \rho \in D_f, D_g. \end{array} \tag{A10}$$

Geodesic dilation and erosion are the basis for building morphological reconstructions. Indeed, the morphological reconstruction by dilation of mask $g$ by marker $f$, is the geodesic dilation of $f$ constrained by $g$ until idempotence, and it is denoted by:

$$\text{where } i \text{ is such that: } \begin{array}{c} R_g^\delta(f) = \delta_g^{(i)}(f), \\ \delta_g^{(i)}(f) = \delta_g^{(i+1)}(f). \end{array} \tag{A11}$$

Dually, the morphological reconstruction by erosion of mask $g$ by marker $f$, is the geodesic erosion of $f$ constrained by $g$ until idempotence:

$$\text{where } i \text{ is such that: } \begin{array}{c} R_g^\delta(f) = \delta_g^{(i)}(f), \\ \varepsilon_g^{(i)}(f) = \varepsilon_g^{(i+1)}(f). \end{array} \tag{A12}$$

## References


1. Mulla, D.J. Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps. *Biosyst. Eng.* **2013**, *114*, 358–371. [CrossRef]
2. Zhai, Z.; Martínez, J.F.; Beltran, V.; Martínez, N.L. Decision support systems for agriculture 4.0: Survey and challenges. *Comput. Electron. Agric.* **2020**, *170*, 105256. [CrossRef]
3. Srinivasan, A. *Handbook of Precision Agriculture: Principles and Applications*; CRC: New York, NY, USA, 2006; ISBN 9781482277968.
4. Liaghat, S.; Balasundram, S.K. A review: The role of remote sensing in precision agriculture. *Am. J. Agric. Biol. Sci.* **2010**, *5*, 50–55. [CrossRef]
5. Wójtowicz, M.; Wójtowicz, A.; Piekarczyk, J. Application of remote sensing methods in agriculture. *Commun. Biometry Crop Sci.* **2016**, *11*, 31–50.
6. Aquino, A.; Millan, B.; Diago, M.-P.; Tardaguila, J. Automated early yield prediction in vineyards from on-the-go image acquisition. *Comput. Electron. Agric.* **2018**, *144*, 26–36. [CrossRef]
7. Aquino, A.; Ponce, J.M.; Andújar, J.M. Identification of olive fruit, in intensive olive orchards, by means of its morphological structure using convolutional neural networks. *Comput. Electron. Agric.* **2020**, *176*, 105616. [CrossRef]

8. Noguera, M.; Millán, B.; Pérez-Paredes, J.J.; Ponce, J.M.; Aquino, A.; Andújar, J.M. A New Low-Cost Device Based on Thermal Infrared Sensors for Olive Tree Canopy Temperature Measurement and Water Status Monitoring. *Remote Sens.* **2020**, *12*, 723. [CrossRef]
9. Qureshi, W.S.; Payne, A.; Walsh, K.B.; Linker, R.; Cohen, O.; Dailey, M.N. Machine vision for counting fruit on mango tree canopies. *Precis. Agric.* **2017**, *18*, 224–244. [CrossRef]
10. Zhang, J.; Tian, H.; Wang, D.; Li, H.; Mouazen, A.M. A Novel Approach for Estimation of Above-Ground Biomass of Sugar Beet Based on Wavelength Selection and Optimized Support Vector Machine. *Remote Sens.* **2020**, *12*, 620. [CrossRef]
11. Zhao, Y.; Gong, L.; Zhou, B.; Huang, Y.; Liu, C. Detecting tomatoes in greenhouse scenes by combining AdaBoost classifier and colour analysis. *Biosyst. Eng.* **2016**, *148*, 127–137. [CrossRef]
12. Colwell, R.N. Determining the prevalence of certain cereal crop diseases by means of aerial photography. *Hilgardia* **1956**, *26*, 223–286. [CrossRef]
13. Bauer, M.; Cipra, J. Identification of Agricultural Crops by Computer Processing of ERTS MSS Data. *LARS Tech. Rep.* **1973**, *20*, 205–212.
14. Hoffman, R.O.; Edwards, D.M.; Eucker, C.C. Identifying and Measuring Crop Type Using Satellite Imagery. *Trans. ASAE* **1976**, *19*, 1066–1070. [CrossRef]
15. Colomina, I.; Molina, P. Unmanned aerial systems for photogrammetry and remote sensing: A review. *ISPRS J. Photogramm. Remote Sens.* **2014**, *92*, 79–97. [CrossRef]
16. Hunt, E.R.; Daughtry, C.S.T. What good are unmanned aircraft systems for agricultural remote sensing and precision agriculture? *Int. J. Remote Sens.* **2018**, *39*, 5345–5376. [CrossRef]
17. Yang, G.; Liu, J.; Zhao, C.; Li, Z.; Huang, Y.; Yu, H.; Xu, B.; Yang, X.; Zhu, D.; Zhang, X.; et al. Unmanned Aerial Vehicle Remote Sensing for Field-Based Crop Phenotyping: Current Status and Perspectives. *Front. Plant Sci.* **2017**, *8*, 1111. [CrossRef]
18. Zhang, C.; Kovacs, J.M. The application of small unmanned aerial systems for precision agriculture: A review. *Precis. Agric.* **2012**, *13*, 693–712. [CrossRef]
19. Hassler, S.C.; Baysal-Gurel, F. Unmanned Aircraft System (UAS) Technology and Applications in Agriculture. *Agronomy* **2019**, *9*, 618. [CrossRef]
20. Gevaert, C.M.; Tang, J.; García-Haro, F.J.; Suomalainen, J.; Kooistra, L. Combining hyperspectral UAV and multispectral Formosat-2 imagery for precision agriculture applications. In *Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*; IEEE: Piscataway, NJ, USA, 2014. [CrossRef]
21. Matese, A.; Toscano, P.; Di Gennaro, S.F.; Genesio, L.; Vaccari, F.P.; Primicerio, J.; Belli, C.; Zaldei, A.; Bianconi, R.; Gioli, B. Intercomparison of UAV, Aircraft and Satellite Remote Sensing Platforms for Precision Viticulture. *Remote Sens.* **2015**, *7*, 2971–2990. [CrossRef]
22. He, F.; Xiong, W.; Habib, A. A Structure-from-Motion Approach Using UAV-based Imagery for Precision Agriculture Applications. In Proceedings of the 10th International Conference on Mobile Mapping Technology, Cairo, Egypt, 6–8 May 2017.
23. Chapman, S.C.; Merz, T.; Chan, A.; Jackway, P.; Hrabar, S.; Dreccer, M.F.; Holland, E.; Zheng, B.; Ling, T.J.; Jimenez-Berni, J. Pheno-Copter: A Low-Altitude, Autonomous Remote-Sensing Robotic Helicopter for High-Throughput Field-Based Phenotyping. *Agronomy* **2014**, *4*, 279–301. [CrossRef]
24. Marino, S.; Alvino, A. Detection of Spatial and Temporal Variability of Wheat Cultivars by High-Resolution Vegetation Indices. *Agronomy* **2019**, *9*, 226. [CrossRef]
25. Zaman-Allah, M.; Vergara, O.; Araus, J.L.; Tarekegne, A.; Magorokosho, C.; Zarco-Tejada, P.J.; Hornero, A.; Albà, A.H.; Das, B.; Craufurd, P.; et al. Unmanned aerial platform-based multi-spectral imaging for field phenotyping of maize. *Plant Methods* **2015**, *11*, 1–10. [CrossRef]
26. Roth, L.; Streit, B. Predicting cover crop biomass by lightweight UAS-based RGB and NIR photography: An applied photogrammetric approach. *Precis. Agric.* **2018**, *19*, 93–114. [CrossRef]
27. Jin, X.; Liu, S.; Baret, F.; Hemerlé, M.; Comar, A. Estimates of plant density of wheat crops at emergence from very low altitude UAV imagery. *Remote Sens. Environ.* **2017**, *198*, 105–114. [CrossRef]
28. Aparna, P.; Ramachandra, H.; Sounder, H.; Harshita, M.P.; Nandkishore, K.; Vinod, P.V. CNN Based Technique for Automatic Tree Counting Using Very High Resolution Data. In Proceedings of the 2018 International Conference on Design Innovations for 3Cs Compute Communicate Control (ICDI3C), Bangalore, India, 25–26 April 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 127–129.
29. Rosell, J.R.; Sanz, R. A review of methods and applications of the geometric characterization of tree crops in agricultural activities. *Comput. Electron. Agric.* **2012**, *81*, 124–141. [CrossRef]
30. Ma, Q.; Su, Y.; Guo, Q. Comparison of Canopy Cover Estimations from Airborne LiDAR, Aerial Imagery, and Satellite Imagery. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2017**, *10*, 4225–4236. [CrossRef]
31. Zortea, M.; Macedo, M.M.G.; Mattos, A.B.; Ruga, B.C.; Gemignani, B.H. Automatic Citrus Tree Detection from UAV Images based on Convolutional Neural Networks. In Proceedings of the 31th Sibgrap/WIA—Conference on Graphics, Patterns and Images, Foz do Iguaçu, Brazil, 29 October–1 November 2018.
32. Csillik, O.; Cherbini, J.; Johnson, R.; Lyons, A.; Kelly, M. Identification of citrus trees from unmanned aerial vehicle imagery using convolutional neural networks. *Drones* **2018**, *2*, 39. [CrossRef]
33. Ampatzidis, Y.; Partel, V. UAV-Based High Throughput Phenotyping in Citrus Utilizing Multispectral Imaging and Artificial Intelligence. *Remote Sens.* **2019**, *11*, 410. [CrossRef]
34. Recio, J.A.; Hermosilla, T.; Ruiz, L.A.; Palomar, J. Automated extraction of tree and plot-based parameters in citrus orchards from aerial images. *Comput. Electron. Agric.* **2013**, *90*, 24–34. [CrossRef]

35. Ok, A.O.; Ozdarici-Ok, A. 2-D delineation of individual citrus trees from UAV-based dense photogrammetric surface models. *Int. J. Digit. Earth* **2017**, *11*, 583–608. [CrossRef]
36. Modica, G.; De Luca, G.; Messina, G.; Praticò, S. Comparison and assessment of different object-based classifications using machine learning algorithms and UAVs multispectral imagery: A case study in a citrus orchard and an onion crop. *Eur. J. Remote Sens.* **2021**, *54*, 431–460. [CrossRef]
37. Dong, X.; Zhang, Z.; Yu, R.; Tian, Q.; Zhu, X. Extraction of Information about Individual Trees from High-Spatial-Resolution UAV-Acquired Images of an Orchard. *Remote Sens.* **2020**, *12*, 133. [CrossRef]
38. Marques, P.; Pádua, L.; Adão, T.; Hruška, J.; Peres, E.; Sousa, A.; Sousa, J.J. UAV-Based Automatic Detection and Monitoring of Chestnut Trees. *Remote Sens.* **2019**, *11*, 855. [CrossRef]
39. Salamí, E.; Gallardo, A.; Skorobogatov, G.; Barrado, C. On-the-Fly Olive Trees Counting Using a UAS and Cloud Services. *Remote Sens.* **2019**, *11*, 316. [CrossRef]
40. Jiang, H.; Chen, S.; Li, D.; Wang, C.; Yang, J. Papaya Tree Detection with UAV Images Using a GPU-Accelerated Scale-Space Filtering Method. *Remote Sens.* **2017**, *9*, 721. [CrossRef]
41. Johansen, K.; Raharjo, T.; McCabe, M. Using Multi-Spectral UAV Imagery to Extract Tree Crop Structural Properties and Assess Pruning Effects. *Remote Sens.* **2018**, *10*, 854. [CrossRef]
42. Diez, C.M.; Moral, J.; Cabello, D.; Morello, P.; Rallo, L.; Barranco, D. Cultivar and tree density as key factors in the long-term performance of super high-density olive orchards. *Front. Plant Sci.* **2016**, *7*, 1226. [CrossRef] [PubMed]
43. Menzel, C.M.; Le Lagadec, M.D. Increasing the productivity of avocado orchards using high-density plantings: A review. *Sci. Hortic.* **2014**, *177*, 21–36. [CrossRef]
44. Majid, I.; Khalil, A.K.; Nazir, N. Economic Analysis of High Density Orchards. *Int. J. Adv. Res. Sci. Eng.* **2018**, *7*, 821–829. [CrossRef]
45. Sarabia, R.; Aquino, A.; Ponce, J.M.; López, G.; Andújar, J.M. Automated Identification of Crop Tree Crowns from UAV Multispectral Imagery by Means of Morphological Image Analysis. *Remote Sens.* **2020**, *12*, 748. [CrossRef]
46. Instituto Geográfico Nacional. Available online: http://www.ign.es/web/ign/portal (accessed on 19 November 2021).
47. Shepard, D. A two-dimensional interpolation function for irregularly-spaced data. In Proceedings of the 1968 23rd ACM National Conference, New York, NY, USA, 27–29 August 1968; ACM Press: New York, NY, USA, 1968; pp. 517–524.
48. Soille, P. *Morphological Image Analysis: Principles and Applications*; Springer: Berlin/Heidelberg, Germany, 2004; ISBN 9783662050880.
49. Otsu, N. A Threshold Selection Method from Gray-Level Histograms. *IEEE Trans. Syst. Man. Cybern.* **1979**, *9*, 62–66. [CrossRef]
50. Rosenfeld, A. Connectivity in Digital Pictures. *J. ACM* **1970**, *17*, 146–160. [CrossRef]
51. Maurer, C.R.; Qi, R.; Raghavan, V. A linear time algorithm for computing exact Euclidean distance transforms of binary images in arbitrary dimensions. *IEEE Trans. Pattern Anal. Mach. Intell.* **2003**, *25*, 265–270. [CrossRef]
52. Beucher, S.; Lantuejoul, C. Use of watersheds in contour detection. In Proceedings of the International Workshop on Image Processing, Real-Time Edge and Motion Detection, Rennes, France, 17–21 September 1979; pp. 17–21.
53. Beucher, S.; Meyer, F. The Morphological Approach to Segmentation: The Watershed Transformation. In *Mathematical Morphology in Image Processing*, 1st ed.; Dougherty, E., Ed.; CRC: New York, NY, USA, 1992; Volume I, pp. 433–481. ISBN 9780824787240.
54. Meyer, F. Topographic distance and watershed lines. *Signal Process.* **1994**, *38*, 113–125. [CrossRef]
55. Gies, V.; Bernard, T.M. Statistical solution to watershed over-segmentation. *Proc. Int. Conf. Image Process. ICIP* **2004**, *3*, 1863–1866. [CrossRef]
56. Frucci, M. From Segmentation to Binarization of Gray-level Images. *J. Pattern Recognit. Res.* **2008**, *1*, 1–13. [CrossRef]
57. Najman, L.; Schmitt, M. Geodesic saliency of watershed contours and hierarchical segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **1996**, *18*, 1163–1173. [CrossRef]
58. Jain, A.K. *Fundamentals of Digital Image Processing*; Prentice Hall: Upper Saddle River, NJ, USA, 1989; ISBN 0133361659.
59. Pérez, A.J.; López, F.; Benlloch, J.V.; Christensen, S. Colour and shape analysis techniques for weed detection in cereal fields. *Comput. Electron. Agric.* **2000**, *25*, 197–212. [CrossRef]
60. Thorp, K.R.; Tian, L.F. A Review on Remote Sensing of Weeds in Agriculture. *Precis. Agric.* **2004**, *5*, 477–508. [CrossRef]
61. López-Granados, F. Weed detection for site-specific weed management: Mapping and real-time approaches. *Weed Res.* **2011**, *51*, 1–11. [CrossRef]
62. Torres-Sánchez, J.; Peña-Barragán, J.M.; Gómez-Candón, D.; De Castro, A.I.; López-Granados, F. Imagery from unmanned aerial vehicles for early site specific weed management. *Precis. Agric.* **2013**, *13*, 193–199. [CrossRef]
63. Serra, J. *Image Analysis and Mathematical Morphology*; Academic Press Inc.: Cambridge, MA, USA, 1982; Volume I, ISBN 9780126372427.