

Article

# Automated Identification of Crop Tree Crowns from UAV Multispectral Imagery by Means of Morphological Image Analysis

Ricardo Sarabia <sup>1</sup>, Arturo Aquino <sup>2</sup>, Juan Manuel Ponce <sup>2,\*</sup>, Gilberto López <sup>1</sup> and José Manuel Andújar <sup>2</sup> 

<sup>1</sup> Instituto de Ingeniería Agrícola y Uso Integral del Agua, Chapingo Autonomous University, Texcoco 56230, Mexico; al17129576@chapingo.mx (R.S.); GLOPEZC@chapingo.mx (G.L.)

<sup>2</sup> Department of Electronic Engineering, Computer Systems and Automation, Higher Technical School of Engineering, University of Huelva, Fuerzas Armadas Ave, 21007 Huelva, Spain; arturo.aquino@diesia.uhu.es (A.A.); andujar@diesia.uhu.es (J.M.A.)

\* Correspondence: jmponce.real@diesia.uhu.es

Received: 11 December 2019; Accepted: 21 February 2020; Published: 25 February 2020



**Abstract:** Within the context of precision agriculture, goods insurance, public subsidies, fire damage assessment, etc., accurate knowledge about the plant population in crops represents valuable information. In this regard, the use of Unmanned Aerial Vehicles (UAVs) has proliferated as an alternative to traditional plant counting methods, which are laborious, time demanding and prone to human error. Hence, a methodology for the automated detection, geolocation and counting of crop trees in intensive cultivation orchards from high resolution multispectral images, acquired by UAV-based aerial imaging, is proposed. After image acquisition, the captures are processed by means of photogrammetry to yield a 3D point cloud-based representation of the study plot. To exploit the elevation information contained in it and eventually identify the plants, the cloud is deterministically interpolated, and subsequently transformed into a greyscale image. This image is processed, by using mathematical morphology techniques, in such a way that the absolute height of the trees with respect to their local surroundings is exploited to segment the tree pixel-regions, by global statistical thresholding binarization. This approach makes the segmentation process robust against surfaces with elevation variations of any magnitude, or to possible distracting artefacts with heights lower than expected. Finally, the segmented image is analysed by means of an ad-hoc moment representation-based algorithm to estimate the location of the trees. The methodology was tested in an intensive olive orchard of 17.5 ha, with a population of 3919 trees. Because of the plot's plant density and tree spacing pattern, typical of intensive plantations, many occurrences of intra-row tree aggregations were observed, increasing the complexity of the scenario under study. Notwithstanding, it was achieved a precision of 99.92%, a sensibility of 99.67% and an F-score of 99.75%, thus correctly identifying and geolocating 3906 plants. The generated 3D point cloud reported root-mean square errors (RMSE) in the X, Y and Z directions of 0.73 m, 0.39 m and 1.20 m, respectively. These results support the viability and robustness of this methodology as a phenotyping solution for the automated plant counting and geolocation in olive orchards.

**Keywords:** aerial imagery; image analysis; multispectral imagery; crop tree; phenotyping; plant population; UAV

## 1. Introduction

Currently, global food demands entail one of the most challenging problems addressed by society. Indeed, as a consequence of the population growth expectations, the demand for crop production

is estimated to increase on the order of 100% in 2050, when compared to 2005 reports [1]. This scenario forces society to develop agricultural and food systems prone to proactively satisfy such a demand while being capable of minimizing the environmental impact. In this sense, crop phenotyping constitutes a crucial tool in order to achieve this balance.

Indeed, deep knowledge about observable crop traits and the way the genotype of plants expresses in relationship with the environmental factors comprise a relevant and valuable information for farmers [2]. Within this context, individual plant counting is a key factor, not only regarding to crop phenotyping, but also providing valuable information, supporting farmers when planning breeding strategies and another agricultural tasks. Thus, the plant population determines the crop density, defined as the number of plants per cultivated hectare. This statistic is closely related to different aspects, such as the efficiency of water and fertilizer resources, or pathogen susceptibility [3]. In addition, it plays a key role when estimating crop yield in tree-based cultivation, and it helps farmers when designing watering and/or fertilization schemes [4]. The importance of the plant population does not stop here, as it is a significant indicator when applying for public subsidies [5], pricing plantations [6], or assessing losses after any kind of extraordinary event, such as fire damage, pest infestations or other natural disasters. However, traditional counting methods are usually based on in-field human visual inspections, so as happens with other phenotyping activities [7,8], it implies tedious, time consuming and prone-to-error tasks, especially when it comes to large-scale plantations [3]. Due to these difficulties, there is a pressing need for the development of new techniques aimed at carrying out plant counting in an accurate, efficient and automated way.

Nowadays, Unmanned Aerial Vehicles (UAVs) have popularised as part of the remote sensing technologies incorporated into precision agriculture, and they have become widely used in crop phenotyping research [9,10]. This is mostly due to the advantages they offer over traditional aerial imaging systems already tested within this application, such as those based on manned airplanes or satellites. When compared to them, UAV-based imaging implies lower operational costs, less weather constraints and the possibility of operating under cloudy conditions [9,11–13]. Furthermore, the growth that the market related to UAVs and remote sensing equipment is experiencing nowadays makes this technology increasingly accessible and affordable. Hence, they are definitely promising tools within the scope of smart farming and precision agriculture, with potential uses in crop phenotyping tasks [9,14].

In fact, when focusing on plant detection and counting, a considerable amount of research where crop tree identification is realised from UAV-based imagery can be found already. Images acquired are usually processed, generating representative data structures of the study sites which are subsequently analysed in order to detect and count the plants. Hence, Malek et al. [5] approached palm tree detection, by analysing a set of candidates, previously computed using the scale-invariant feature transform (SIFT), with an extreme learning machine (ELM) classifier. Candidates categorised as trees were post-processed by means of a contour method based on level sets (LS) and local binary patterns (LBP), in order to identify the shapes of their crowns. In Miserque-Castillo et al. [15], a framework for counting oil palms was developed, where a sliding window-based technique procured a set of candidates. After processing with LBP, they were classified by a logistic regression model. Primicerio et al. [16] studied plant detection within vine rows. The segmentation of the plant mass was carried out on the basis of dynamic segmentation, Hough space clustering and total least squares regression. After individual plant identifications were estimated, a multi-logistic model for the detection of missing plants was applied. Jiang et al. [17] introduced a GPU-accelerated scale-space filtering methodology for detecting papaya and lemon trees in UAV images. To that end, initial captures were converted to a Lab-based colour space, mostly exploiting the information contained in the channel *a* (representative of the colour values from red to green) to differentiate the plants from the ground. Koc-San et al. [18] undertook citrus trees location and counting from UAV multispectral imagery. To that end, they proposed a set of procedures based on sequential thresholding and the Hough transform. In the same vein, Csillik et al. [19] focused on citrus crops, intending the identification of trees by using convolutional neural networks (CNNs). In addition, they used a simple linear iterative clustering (SLIC)

algorithm for classification refinement. CNNs were also used by Ampatzidis and Partel [20] in order to detect citrus trees. Specifically, the CNN model was trained by using a YOLOv3 object detection algorithm. Furthermore, they implemented a normalised difference vegetation index (NDVI)-based image segmentation method for estimating the canopy area. In Selim et al. [21], approached orange tree detection from high-resolution images, by applying an object-based classification methodology, using a multi-resolution segmentation of the data derived from aerial imagery. Deep learning and CNN technology was exploited by Aparna et al. [4]. In this case, coconut palm tree detection was the aim. Initial captures were transformed into an HSV colour representation, and then binarized and conveniently cropped in sub-images, with which the CNN classifier was trained. In Kestur et al. [22], an ELM methodology was proposed for detecting tree crowns from aerial images captured in the visible spectrum. Thus, the developed ELM spectral classifier was applied in order to segment the tree crowns-pixel areas from the rest of the image. The methodology was validated by studying banana, mango and coconut palm trees. Marques et al. [23] focused on the detection of chestnut trees. They considered different kinds of sensorics for acquiring aerial images. Thus, RGB and Colour Infrared (CIR) images were used in their research, where different segmentation techniques were explored in order to properly isolate the tree-belonging pixel-regions to subsequently carry out the eventual identification of the trees.

Regarding olive plantations, which constitute the study case considered throughout the experimentation developed here, several studies where olive tree phenotyping is approached by using UAV-based aerial imagery can be found. Thus, Díaz-Varela et al. [24] attempted the estimation of the height and crown diameter of olive trees by means of structure-from-motion (SfM) image reconstruction and geographical object-based image analysis (GEOBIA). Along the same line, Torres-Sánchez et al. [25] also proposed a methodology for the estimation of different olive tree features. Particularly, height, crown volume and canopy area were addressed. This was accomplished by generating digital surface models (DSMs) from aerial imagery, and object-based image analysis (OBIA). This study was extended in [26], where different flight altitudes and overlapping degrees were tested in order to optimise the DSM generation, in terms of computational cost. In Salamí et al. [6], olive trees counting was approached by using a UAV equipped with a small embedded computer. This device was aimed at processing captures on board, and to provide via cloud services, nearly real-time plant count estimations to the end-user.

In this paper, a new methodology for the identification of crop trees located in intensive farming-based orchards, by means of the analysis of aerial images, is proposed. To that end, we start from a set of aerial captures acquired by a UAV equipped with a multispectral camera while flying over the land plot under study. These multispectral images are processed in order to yield a DSM, following standard image matching and photogrammetry techniques. The core of the novel proposal of the methodology is comprised by an image analysis-based algorithm, aimed at identifying the trees by exploiting the elevation information contained in this data structure. To that end, the DSM is converted into as a greyscale image, where elevation information is approached as grey level values. Then, this image is transformed by means of mathematical morphology, in order to individually segment the tree-belonging pixels from the ground, by a statistical global thresholding-based binarization. Eventually, that resulting segmentation is analysed by an ad-hoc procedure to detect intra-row tree aggregations, consisting in studying the second central moment of the tree pixel-regions. The whole methodology was tested in an intensive olive orchard, obtaining results that highlight its effectiveness as a full-automated solution for crop trees detection and counting, and its robustness against complex scenarios, as intra-row tree aggregations and a strong ground elevation variability were present in the study plot.

Hereafter, the present manuscript is structured as follows: Section 2 focuses on the experimental design. Thus, Section 2.1 describes the characteristics of the olive orchard in which, as study case, images were acquired for the purpose of testing the methodology. Section 2.2 exposes all the aspects related to how aerial image acquisition was performed. In Section 2.3, the image analysis methodology for

trees detection, counting and geolocation is developed, addressing the stages of image pre-processing (Section 2.3.1), the generation of a DSM as a base data structure (Section 2.3.2), and the image segmentation and analysis (Sections 2.3.3 and 2.3.4, respectively). Then, in Section 2.4, the set of metrics computed to assess the performance of the methodology is proposed. Section 3 presents the results obtained, which are then discussed in Section 4. Section 5 concludes the manuscript, giving a brief summary of the main findings achieved and identifying aspects that might be approached in further investigations. Finally, Appendix A formally defines all the morphological operators used throughout the developed image analysis methodology.

## 2. Materials and Methods

### 2.1. Study Case Site

The olive grove where the testing aerial imagery was acquired is located in Gibraleón, province of Huelva (Andalusia, Southwest Spain). In particular, the area under study, centred in the coordinates  $7^{\circ}02'48.44''\text{W}$  and  $37^{\circ}20'39.80''\text{N}$ , corresponds to an orchard with an approximate extent of 17.5 ha, where an intensive cultivation system is applied, with a plant spacing pattern of  $5.5 \times 7$  m; the *Olea europaea* L. cultivated variety is Picual. It should be noted that this orchard shows a notable variability in terms of soil composition, crown size of the trees and altitude, varying from around 54 m to around 96 m above sea level. A third-party aerial capture of the study site, obtained by manned flight-based imaging, is shown in Figure 1. It should be underscored that this third-party image is only offered for the purpose of illustrating the study plot, so it was not used at all throughout the experimentation.



**Figure 1.** Third-party aerial capture of the case study site shown to illustrate the study plot, highlighted in red.

### 2.2. Image Acquisition

#### 2.2.1. Aerial Imaging Equipment

Aerial imaging was conducted using a DJI<sup>TM</sup> Matrice 100 UAV (SZ DJI<sup>TM</sup> Technology Co., Ltd., Shenzhen, Guangdong, China). This device is propelled by four rotors (quadcopter), enabling its vertical take-off and landing. With a diagonal wheelbase of 650 mm and a maximum take-off weight of 3600 g, it can reach a maximum cruise speed of 22 m/s, withstanding a wind resistance up to 10 m/s. It is controlled in an operating frequency varying from 5725 to 5825 GHz, with a maximum transmission distance of 5 km.

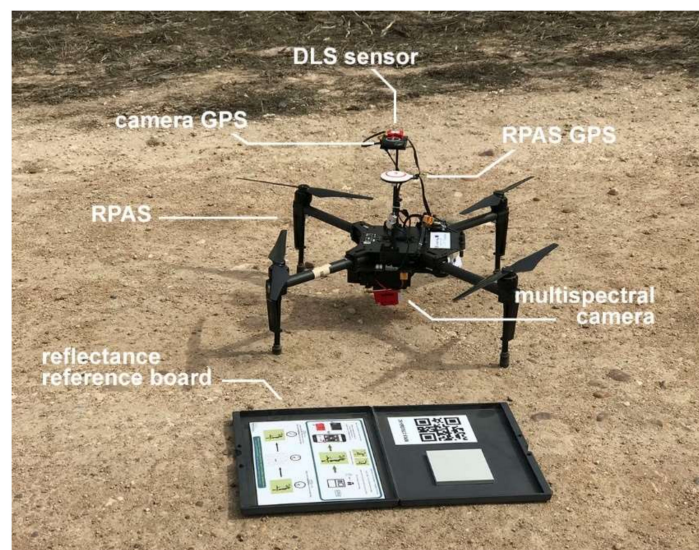
Images were taken with the multispectral camera MicaSense RedEdge-M<sup>TM</sup> (MicaSense, Inc., Seattle, WA, USA), installed on the UAV. This sensing device is capable of capturing information in five different spectral bands within the visible and the infrared spectrum. Table 1 summarises the most relevant features related to these bands.



**Table 1.** Features of the spectral bands captured by the multispectral camera MicaSense RedEdge-M™.

Band Number	Band Name	Centre Wavelength (nm)	Bandwidth (nm)
1	Blue	475	20
2	Green	560	20
3	Red	668	10
4	Near Infrared	840	40
5	Red Edge	717	10

The camera was mounted together with a dedicated GPS device for the purpose of georeferencing each captured image. A downwelling light sensor (DLS) was also included into the setup, in order to calibrate the images according to the changing conditions of ambient light. Finally, for accurate ground reflectance calibration, a reference board (grey reference) was used by imaging it during both the take-off and landing. In Figure 2, the UAV is shown together with all the equipment described above.

**Figure 2.** Equipment used to capture the aerial imagery used in this paper.

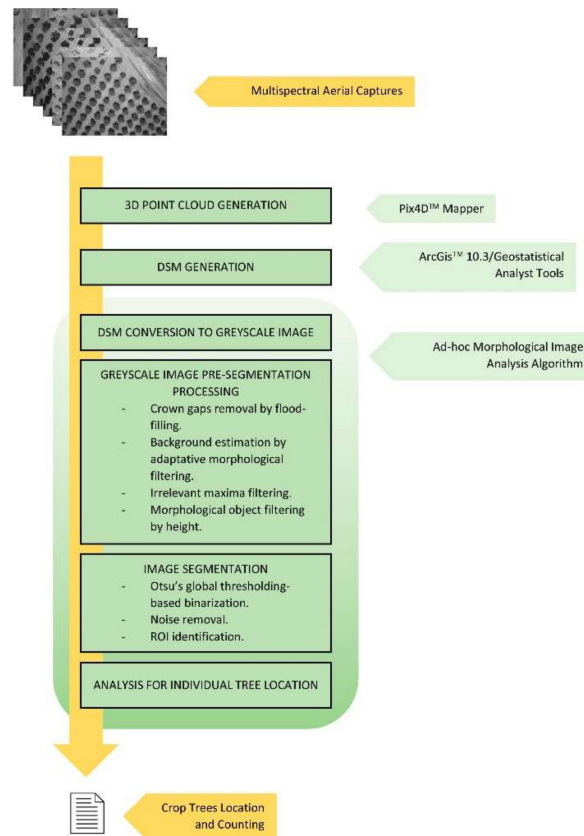
### 2.2.2. Flight Planning and Development

The flight mission planning was set with the DJI™ Flight Planner software, by drawing the polygon delimiting the study plot (highlighted in red in Figure 1). Within the study plot, the mission was planned according to the criterion of minimising the number of turns to be made by the UAV to cover it entirely. Thus, the flight was configured to be performed autonomously, at an altitude of 70 m and at a cruise speed of 15 km/h. The multispectral camera was configured with a time period between captures of 1.5 s. With these settings, it was intended to capture images with forward and lateral overlaps of 85% and 65% respectively, and with a desired GSD of 0.05 m/pixel. The flight took place on June 13, 2019, approximately between 11 a.m. and 1 p.m. Litchi software (VC Technology, Ltd. ©, London, UK) was used for operating and monitoring the mission. A total of 44,325 images were acquired during the flight, 8865 per each of the five spectral bands in which the multispectral camera can capture information.

### 2.3. Image Analysis Methodology for Olive Trees Detection, Geolocation and Counting

The main objective pursued in this investigation is the development of a procedure able to perform olive tree detection, location and counting from aerial captures by means of image analysis. To that end, a methodology has been designed under those principles to, first, transform images acquired into a DSM, as a representative data structure of the whole orchard under study; and then, to exploit the information contained in it in order to carry out a binary segmentation, in which tree-belonging

pixels could be differentiated from the rest of the image. Eventually, the result of this segmentation is analysed to detect intra-row aggregations, thus finally yielding the individual tree locations and the accurate plant population estimation. The flowchart shown in Figure 3 illustrates the different stages comprising the developed methodology, which are deeply detailed throughout the next subsections.

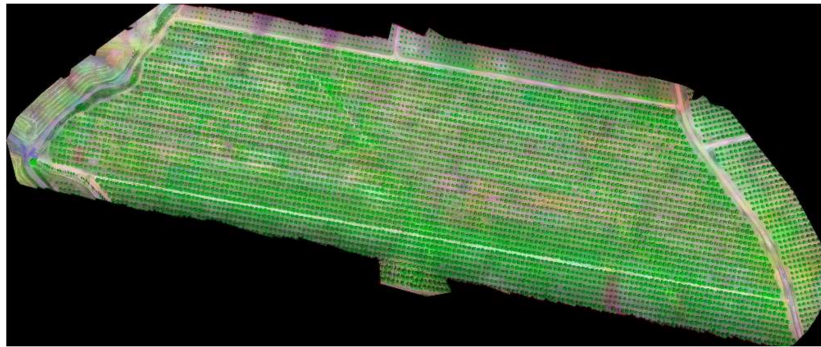


**Figure 3.** Representative diagram of the methodology proposed for detecting and counting crop trees from multispectral aerial images.

For simplicity purposes, all morphological operators involved in the methodology described throughout this section, are formally defined in Appendix A.

### 2.3.1. Image Pre-Processing

As a first step, captures obtained by aerial imaging are radiometrically corrected using the illumination information provided by the camera's DLS sensor and the reflectance measured in the images captured of the reference board. Then, the corrected images are processed to yield the orthomosaics corresponding to each of the five spectral bands considered. In Figure 4, a colour image resulting from the combination of the blue, near infrared (NIR) and red edge bands is presented. It should be underlined that this ad-hoc image was exclusively generated for the purpose of supporting the assessment of the methodology's performance by a human observer, as detailed in Section 2.4. Therefore, they were chosen so as to obtain a proper visual tree differentiation, being other combinations of bands surely also suitable for this purpose.



**Figure 4.** Colour image generated from the information provided by the orthomosaics of the Blue, Red Edge and NIR spectral bands.

In addition, a 3D point cloud is generated as well, to later develop the representative DSM of the overflowed land plot. To that end, every point in the cloud is determined by its re-projection in at least three images; then, it is characterised with a triplet of coordinates, where the two first ones determine its relative location within the cloud and the third one refers to its elevation. Thus, a high-density 3D point cloud with a total number of 205,998,922 points is reached.

The task of creating both the set of orthomosaics and the 3D point cloud was carried out using the photogrammetry Pix4D™ Mapper software. As representative indicators of the errors committed during the pre-processing stage, the software reported root-mean square errors (RMSE) in the  $X$ ,  $Y$  and  $Z$  directions of 0.73 m, 0.39 m and 1.20 m, respectively. Note that these errors do not correspond to the quality of the point cloud, but to the error between the initial and the computed image positions.

### 2.3.2. Digital Surface Model (DSM)

The DSM is generated by deterministic spatial analysis, from the 3D point cloud yielded above, by applying an inverse distance weighting (IDW) interpolation [27]. According to this method, the attribute value (the elevation in this case) of an unsampled point is decided from the attribute value of its surrounding known points. The influence of the known sampled points decreases as their distance from the targeted point increase, so the unsampled point value is computed on the basis of the attribute values of the surrounding points observations, inversely weighted according to their distance. So, being  $S_0$  a targeted point, its interpolation value  $\hat{Z}(S_0)$  can be mathematically defined as follows:

$$\hat{Z}(S_0) = \sum_{i=1}^N \lambda_i Z(S_i). \quad (1)$$

where  $Z(S_i)$  is the observed value for the  $i$ -th surrounding point  $S_i$  of  $N$  points;  $\lambda_i$  is the weight assigned to  $S_i$ , according to its distance  $d_{i0}$  to  $S_0$ . Hence,  $\lambda_i$  can be defined as follow:

$$\lambda_i = \frac{d_{i0}^{-p}}{\sum_{i=1}^N d_{i0}^{-p}} \quad (2)$$

where  $p$  is a weighting exponent that controls the way in which weight decreases with distance; the weights  $\lambda_i$  vary between 0 and 1 for each point and the total sum of them is the unit:  $\sum_{i=1}^N \lambda_i = 1$ .

For computing the DSM using this IDW spatial interpolation, ArcGis™ 10.3 (Esri, Inc., Redlands, CA, USA) and its Geostatistical Analyst Tools extension were used. The size of the cell was matched to the cell size of the orthomosaics computed before. In the same vain, interpolation output raster was also restricted by the dimensions of these orthomosaics. In addition, it should be noted that, during the analysis, it was established a fixed neighbourhood search, using a circular radius distance of 10 m and a maximum number of neighbourhood points of 4, with a weighting exponent  $p$  of 2.

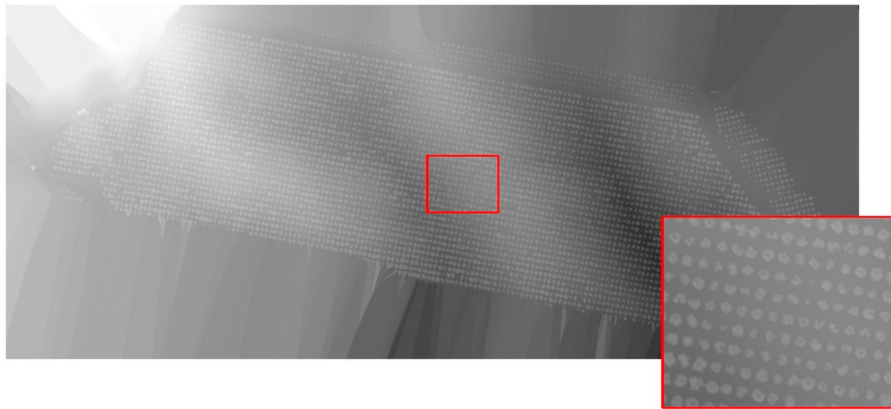
### 2.3.3. Image Segmentation Algorithm

The DSM, obtained after processing the initial aerial captures, is used as fundamental data to eventually perform crop trees detection and the subsequent location and counting. Every voxel (3D pixel) in the DSM is defined by its  $x$  and  $y$  position within the map, and its altitude with respect to the sea level. This altitude information is exploited in such a way that trees are segmented by considering their absolute height with respect to their local neighbourhood.

First, the DSM is approached as a 2D greyscale image by taking the voxels' elevation information as the intensity values of their corresponding pixels in this greyscale image. Thus, given  $DSM$  as the representative matrix of the DMS previously computed, the intensity matrix,  $GS_{DSM}$ , which approaches this model as an 8-bit greyscale image, can be defined as follows:

$$GS_{DSM}(x, y) = \begin{cases} DSM(x, y), & \text{if } DSM(x, y) > 0 \\ 0, & \text{in any other case} \end{cases} \quad (3)$$

where  $DSM(x, y)$  is the elevation value with respect to the sea level provided by the DSM for the point  $(x, y)$ . Figure 5 shows a representation of the DSM as a greyscale image.



**Figure 5.** Representation of the computed DSM as the intensity image  $GS_{DSM}$ . Note in the zoomed area, highlighted in the red square, the differences in terms of grey level between those pixel regions which apparently belong to olive trees, and those from the surrounding ground. Then, given that each pixel intensity value is assigned according to its elevation in the DSM, higher pixel values indicate higher altitudes with respect to the sea level. It should be noted that, for the sake of facilitating its visualisation, the image display range has been established between the minimum bigger-than-zero value from the DMS, and its maximum.

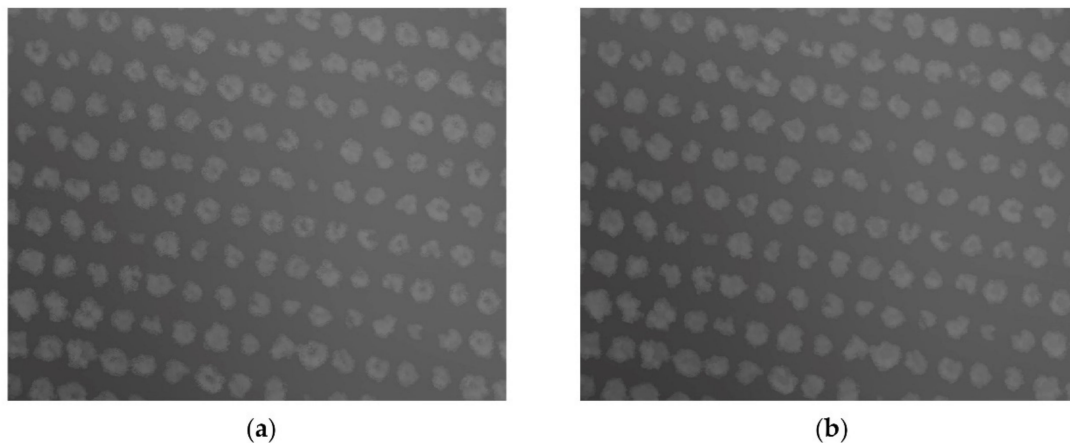
Once this greyscale image is obtained, a filling operation is performed to homogenise the grey level values of the tree crowns which, in some cases, showed darker areas potentially related to hollows in the foliage. Mathematically, it can be defined from a morphological reconstruction as follows:

$$I_{GS1} = R_{GS_{DSM}}^{\varepsilon}(GS'_{DMS}),$$

$$GS'_{DMS}(x, y) = \begin{cases} GS_{DMS}, & \text{if } (x, y) \text{ is a border pixel} \\ 255, & \text{in any other case} \end{cases} \quad (4)$$

where  $GS'_{DMS}$  is a border image of  $GS_{DMS}$ , and  $R^{\varepsilon}$  refers to the morphological reconstruction by erosion ( $\varepsilon$ ) of  $GS_{DMS}$  from marker  $GS'_{DMS}$  until idempotence. Figure 6 shows the effect of this operation on the zoomed area of Figure 5.





**Figure 6.** Filling gaps illustration: (a) same zoomed area of  $GS_{DMS}$  shown in Figure 5; (b) result of the filling gaps operation applied to (a).

Afterwards, a homogenisation of the grey level values of  $I_{GS1}$  is performed aimed at favouring its later optimum binarization. Since  $I_{GS1}$  directly derives from the DSM, its pixel values represent altitude magnitudes expressed in meters with respect to the sea level. Consequently, disturbing cases when binarizing may appear, such as that in which ground pixels have higher grey level values than those of tree pixels (when the former are at higher altitudes than the latter). Hence, in order to avoid this difficulty,  $I_{GS1}$  is homogenised by subtracting from it an accurate background estimate. This is calculated by iteratively opening  $I_{GS1}$  with a circular structuring element of increasing radius, taking at each step the minimum value between the opening results at the  $i$ -th and the  $i-1$ -th iteration. Mathematically:

$$\begin{aligned}
 & I_{BE_{DEF}} = I_{BE_n}, \\
 \text{being} \quad & I_{BE_i} = \begin{cases} I_{GS1}, & \text{if } i = 0 \\ \text{MIN}(I_{BE_{i-1}}, \gamma_{\beta_i}(I_{BE_{i-1}})) & \text{in any other case} \end{cases} \quad (5) \\
 & i = 1, \dots, n.
 \end{aligned}$$

where  $\gamma_{\beta_i}$  is the morphological opening operation using a disc-shaped structuring element  $\beta$  of radius  $i \times 5$ . For a given tree crown in the image, its optimum filtering takes place when its grey level values are substituted with the minimum value existing in its closest background neighbourhood. It happens when the opening operation is performed using a structuring element with the minimum radius allowing the element to completely contain the tree crown. Therefore, note that the formulated approach provides a flexible framework favouring the accurate filtering of every tree independently from its size. The number of iterations has been fixed to  $n = 14$ , which corresponded to a maximum radius value of the structuring element equal to 70. This value has been set to ensure the accurate filtering of the greater trees, being adaptable to different image capturing conditions deriving in other maximum tree crown sizes. Once  $I_{BE_{DEF}}$  is computed, the homogenisation of  $I_{GS1}$  is obtained by:

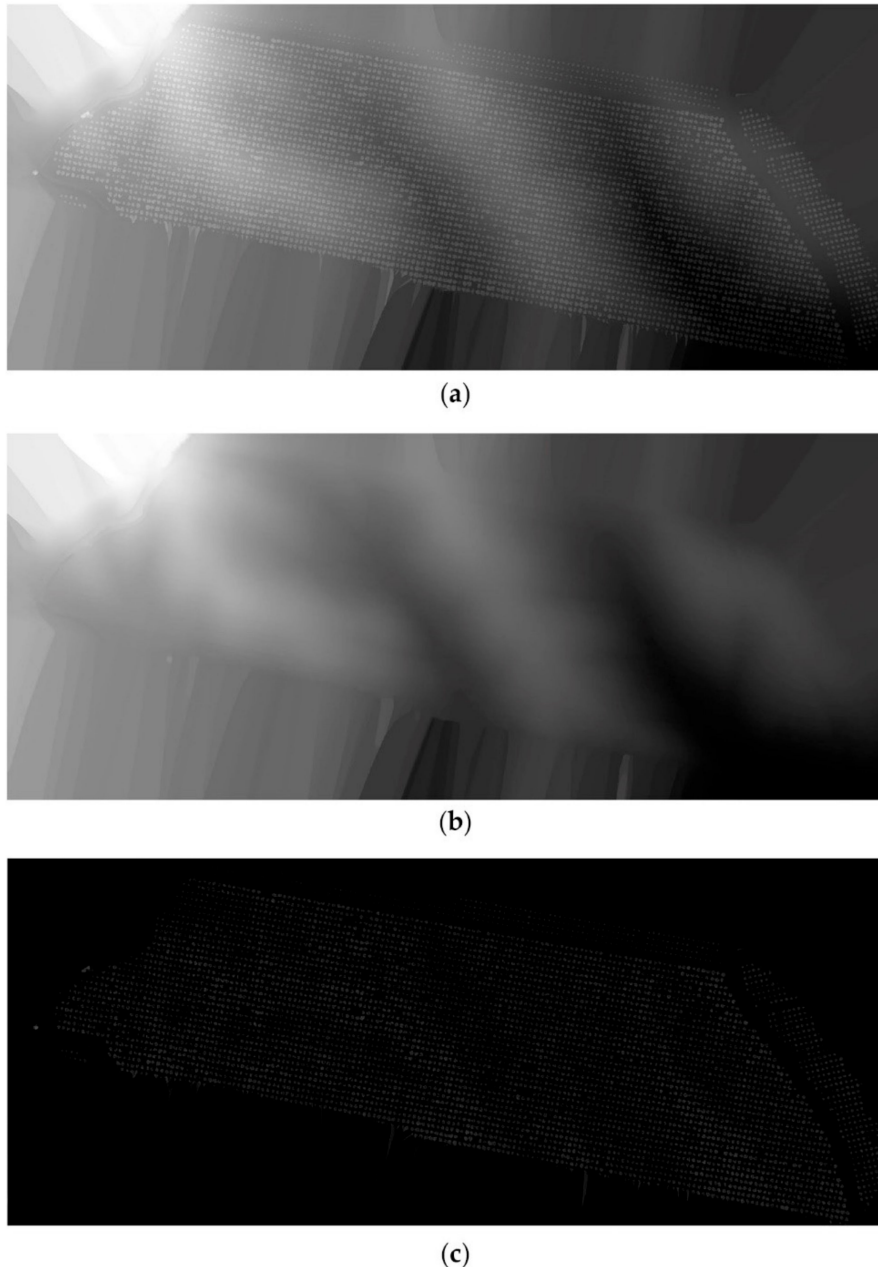
$$I_{GS2} = I_{GS1} - I_{BE_{DEF}}. \quad (6)$$

Figure 7 illustrates the described process to yield a homogenised version of  $I_{GS1}$ . At this point,  $I_{GS2}$  is a homogenised image in which its grey level values represent absolute altitudes. In other words, the previous processing made the effect of ideally flattening the original surface and placing its background at the sea level. In this context, tree crowns are the elements expected to be at higher levels, so the next processing is intended for removing from the image irrelevant maxima, which were

considered to be those pixels with altitude values lower or equal than 1 m. This effect is achieved by applying the  $H$ -maxima transform to image  $I_{GS2}$ :

$$I_{GS3} = HMAX_h(I_{GS2}) = R_{I_{GS2}}^\delta(I_{GS2} - h), h = 1. \quad (7)$$

where  $R^\delta$  refers to the morphological reconstruction by dilation ( $\delta$ ) of  $I_{GS2}$  from marker  $I_{GS2} - h$ . In this case, artefacts with elevation values greater than 1 m were retained for being of interest; note that this criterion can be easily modified by adjusting the  $h$  parameter.



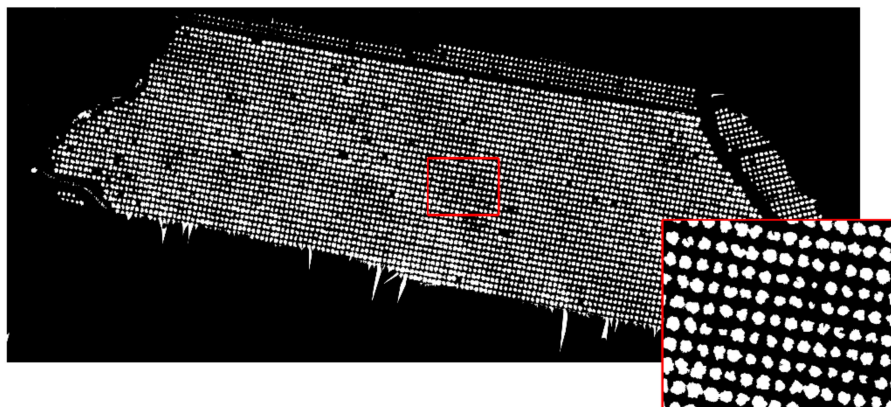
**Figure 7.** (a) Greyscale image  $I_{GS1}$  resulting from filling gaps in the image shown in Figure 5,  $GS_{DMS}$ ; (b) background estimation of (a),  $I_{BEDEF}$ ; (c) resulting image  $I_{GS2}$  after subtracting (b) to (a).

Next, the elements surviving the previous filtering by height are segmented by binarizing image  $I_{GS3}$  using the Otsu's method [28]. This approach assumes that the population of grey level values of the image is made up of two dominant groups or classes, corresponding to the foreground and the

background pixels, respectively. Hence, it determines the grey value maximising the separability of both classes, which results in the greater median distance between them, or analogously, the minimum intra-class variance. Therefore, given the threshold  $thresh$  resulting from applying the Otsu's method to image  $I_{GS3}$ , its binarization can be defined as follows:

$$I_{BIN1}(x, y) = \begin{cases} 255, & \text{if } I_{GS3}(x, y) > thresh \\ 0, & \text{in any other case} \end{cases} \quad (8)$$

Figure 8 illustrates  $I_{BIN1}$ , resulting from the binarization of image  $I_{GS3}$ , shown in Figure 7c.



**Figure 8.** Image  $I_{BIN1}$  resulting from the binarization of  $I_{GS3}$ , shown in Figure 6c. Note in the zoomed area, in the red square, how potential plants have been accurately segmented from the background.

As a result of this binarization, image pixels are segmented into two classes, background (black pixels) and foreground (white pixels); this latter being potentially formed by pixels belonging to olive trees. Next, in order to remove the spurious connected components (set of neighbour foreground pixels) abnormally small, a morphological opening is applied on the binary image  $I_{BIN1}$ :

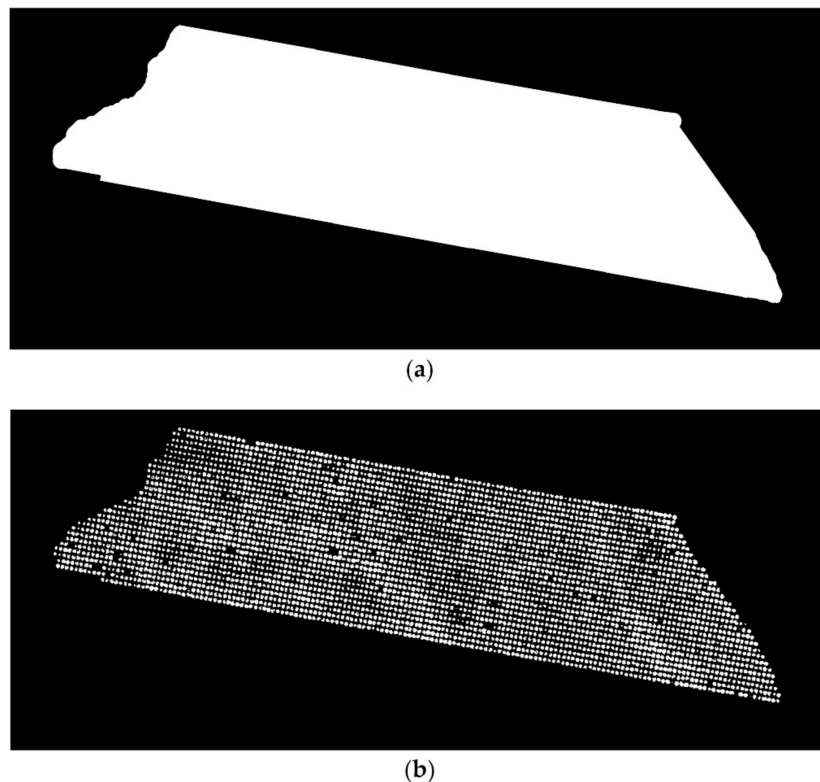
$$I_{BIN2} = \gamma_{\beta}(I_{BIN1}). \quad (9)$$

where  $\gamma_{\beta}$  stands for the morphological opening performed by using a disk-shaped structuring element  $\beta$  of 5 pixels in radius. To exactly recover the shape of the connected components surviving to this noise filtering,  $I_{BIN1}$  is morphologically reconstructed by dilation ( $R^{\delta}$ ) from marker  $I_{BIN2}$ , which leads to  $I_{BIN3}$ :

$$I_{BIN3} = R_{I_{BIN1}}^{\delta}(I_{BIN2}). \quad (10)$$

After noise removal, the polygon drawn to delimit the region of interest (ROI) for flight planning (Section 2.2.2) is used as a mask image,  $I_{ROI}$ , to constrain the area of interest within the image for the rest of the analysis. Figure 9 illustrates  $I_{ROI}$ , together with the result of its application to  $I_{BIN3}$ , which can be mathematically formulated as:

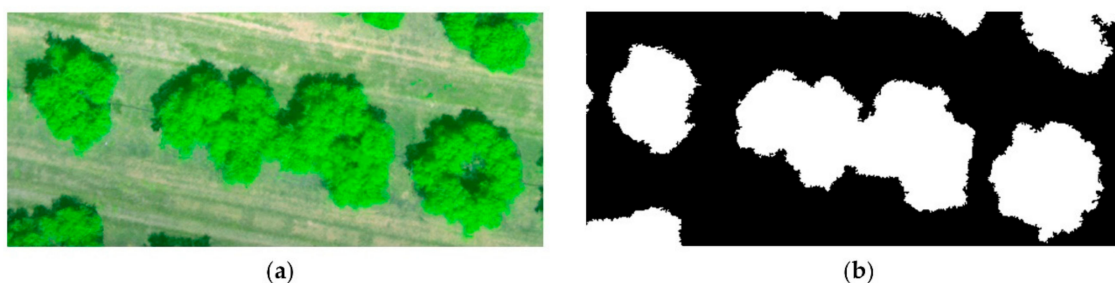
$$I_{BINdef}(x, y) = \begin{cases} I_{BIN3}(x, y), & \text{if } I_{ROI}(x, y) > 0 \\ 0 & \text{in any other case} \end{cases} \quad (11)$$



**Figure 9.** (a) ROI mask image  $I_{ROI}$ ; (b) image  $I_{BINdef}$  resulting from filtering the binary image  $I_{BIN3}$  (visually very similar to  $I_{BIN3}$ , shown in Figure 8), with image  $I_{ROI}$ .

#### 2.3.4. Image Analysis Algorithm for Tree Counting and for the Estimation of Tree Locations

As it can be seen in Figure 9,  $I_{BINdef}$  provides a segmentation of the olive trees from the background. A first approach to count the number of plants might just consider the number of connected components in that binary image. Nevertheless, the possibility of existing connected components not exactly corresponding to a sole olive tree has to be considered. Indeed, because of the variability in terms of crown size shown by the trees of the plot under study, their foliage may appear overlapped within the same row, thus resulting in wrongly merged connected components in the binary image; overlapping of tree crowns from different rows is not expected as it is prevented by pruning; Figure 10 illustrates this phenomenon. Therefore, the image analysis procedure described below has been developed intended to accurately provide plant population, despite intra-row tree aggregations.



**Figure 10.** (a) Sub-image of the study plot orthomosaic represented in Figure 4, where it can be observed a couple of trees with overlapping foliage; (b) sub-image of the binary image resulted from the segmentation performed,  $I_{BINdef}$ , corresponding to the area represented in (a). Note how the two olive trees share the same connected component.



The procedure is based on analysing the morphology of the segmented connected components of  $I_{BINdef}$ , in order to determine the estimated number of trees contained in them. To that end, the components of the binary image are firstly approached with the ellipses that share the same normalised second central moment [29]. Thus, for a given connected component  $cc_i$ , its representing ellipse  $E_i$  is defined by the following set of elements:

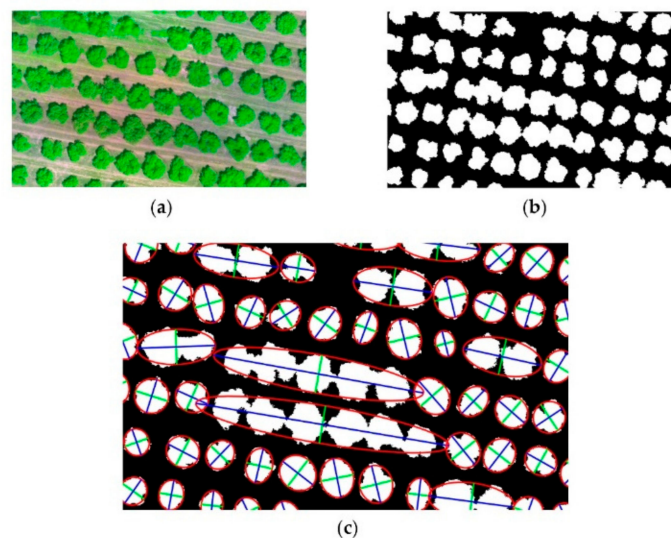
$$E_i = \{c_x E_i, c_y E_i, d_1 E_i, d_2 E_i, \alpha E_i\}. \quad (12)$$

being  $c_x E_i$  and  $c_y E_i$  the coordinates of its centre,  $d_1 E_i$  and  $d_2 E_i$  the length in pixels of its two axes and  $\alpha E_i$  the angle formed by its longer axis and a horizontal imaginary axis. Consequently, the length of the major and minor axes of the ellipse can be defined as:

$$\begin{aligned} MajAx(cc_i) &= MAX\{d_1 E_i, d_2 E_i\}, \\ MinAx(cc_i) &= MIN\{d_1 E_i, d_2 E_i\}. \end{aligned} \quad (13)$$

As can be seen in Figure 11, whilst the minor axes keep comparable length values throughout the whole population of ellipses, regardless of the number of plants contained in the corresponding components, the length of the major axes show a strong dependence with this number. To exploit this feature, next, the maximum length value of all the computed minor axes was calculated as a reference to subsequently be used throughout the rest of the analysis. Thus:

$$MAX_{MinAX} = MAX\{MinAx(cc_i)\}. \quad (14)$$



**Figure 11.** (a) Sub-image of the study plot orthomosaic represented in Figure 4; (b) sub-image of the binary image resulted from the segmentation performed, corresponding to the area represented in (a); (c) representation of the ellipses (in red) computed for each connected component in the image (b), with their corresponding major (in blue) and minor (in green) axes.

Then, counting of trees was conducted by comparing  $MAX_{MinAX}$  to the length of the major axis of each connected component  $cc_i$ , by computing:

$$TreeNumber(cc_i) = \begin{cases} 1, & \text{if } MajAx(cc_i) \leq MAX_{MinAX} \times 1.20 \\ \lfloor MajAx(cc_i) / MAX_{MinAX} \rfloor, & \text{in any other case} \end{cases} \quad (15)$$

As the shapes of tree crowns are irregular, the  $MAX_{MinAX}$  value, computed on the population of minor axes, might be slightly lower than the length of the major axis of an eventual connected

component representing a sole tree. Additionally, we note that tree spacing in agricultural plantations, such as that considered in this study, is ideally regular, so the potential for overlapping trees are greater. Therefore, as Equation (15) shows, increasing  $MAX_{MinAX}$  by 20% provides flexibility to the former situation while it respects the latter assumption, as the aggregation of great tree crowns is not expected to enlarge the resulting object by only a 20%. The concrete value has been decided empirically, being not critical, as values moderately higher and lower were also found to provide comparable results. Finally, once the number of plants per connected component is estimated, the total number of trees is calculated by the addition of these partial results:

$$TotalTreePopulation = \sum_{i=1}^n TreeNumber(cc_i), \quad (16)$$

being  $n$  the number of connected components.

Finally, once trees are counted, a representative location for each of them within the image is attempted. Hence, the following definition was established:

$$TreeLocation_{cc_i} = \left\{ \begin{array}{l} (x_{ij}, y_{ij}) \\ (x_{ik}, y_{ik}), \end{array} \right. \left. \begin{array}{l} (x_{i1}, y_{i1}) = (c_x E_i, c_y E_i), TreeNumber(cc_i) = 1 \vee \\ TreeNumber(cc_i) = t, k = 1, \dots, t \end{array} \right\}. \quad (17)$$

This is, for a given connected component  $cc_i$  containing a sole tree, the location of the latter is decided as the location of the centre of the ellipse  $E_i$  representing the former. Conversely, for aggregated components, the location of the contained multiples trees is estimated by equally spacing them throughout the major axis of its representing ellipse, taking as reference the centre of this last. With this approach, two situations must be considered. The first case refers to when  $cc_i$  contains an odd number of trees. In this scenario, the location of the central tree matches with the centre of its ellipse, being the resting locations estimated by displacements to the left and to the right of this reference. Mathematically:

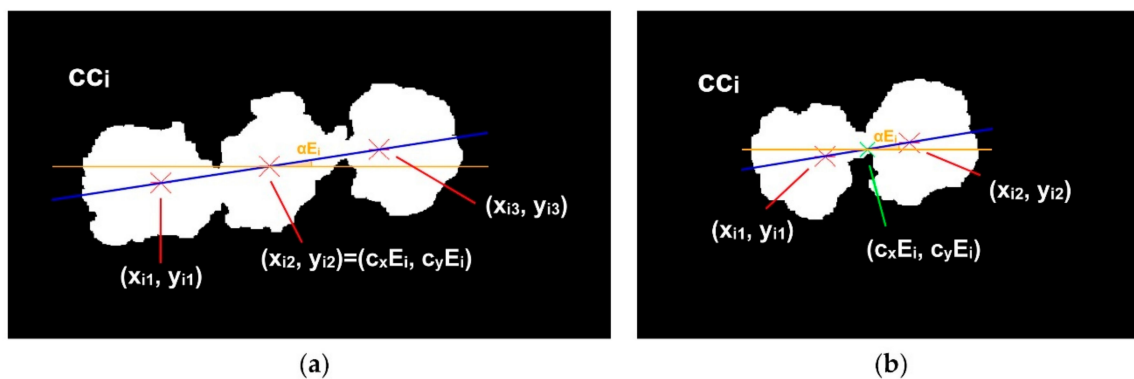
$$\begin{aligned} x_{ik} &= \begin{cases} c_x E_i + k \times jump_i \times \cos(\alpha E_i + \pi), & \text{if } k < t/2 \\ c_x E_i, & \text{if } k = t/2 \\ c_x E_i + (k - t/2) \times jump_i \times \cos(\alpha E_i), & \text{if } k > t/2 \end{cases}, \\ y_{ik} &= \begin{cases} c_y E_i + k \times jump_i \times \sin(\alpha E_i + \pi), & \text{if } k < t/2 \\ c_y E_i, & \text{if } k = t/2 \\ c_y E_i + (k - t/2) \times jump_i \times \sin(\alpha E_i), & \text{if } k > t/2 \end{cases}, \\ &k = 1, \dots, t, \quad t = TreeNumber(cc_i), \\ &jump_i = MajAx(cc_i) / (t + 1). \end{aligned} \quad (18)$$

where  $jump_i$  represents the magnitude of the displacements among the estimated tree centres. Note that the first case models the estimated locations placed at the left of the central tree, the third case models those placed at the right, and the second one defines the case of the central tree. The second scenario occurs when  $cc_i$  contains an even number of trees. For this case, the centre of its representing ellipse does not match with the expected centre of a tree, but with the overlapping zone of two of them. Hence, this location is not assigned to any tree, but it is only taken as a reference:

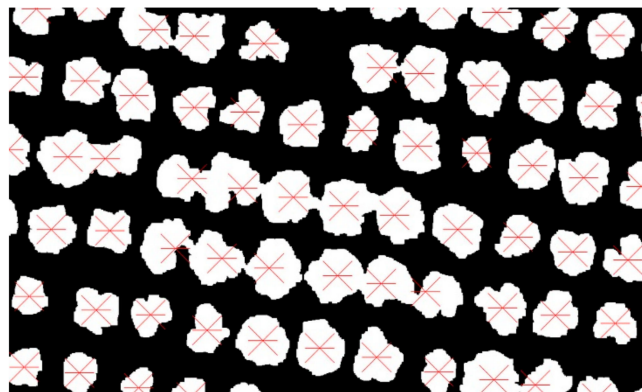
$$\begin{aligned} x_{ik} &= \begin{cases} c_x E_i + (k - 0.5) \times jump_i \times \cos(\alpha E_i + \pi), & \text{if } k \leq t/2 \\ c_x E_i + (k - t/2 - 0.5) \times jump_i \times \cos(\alpha E_i), & \text{if } k > t/2 \end{cases}, \\ y_{ik} &= \begin{cases} c_y E_i + (k - 0.5) \times jump_i \times \sin(\alpha E_i + \pi), & \text{if } k \leq t/2 \\ c_y E_i + (k - t/2 - 0.5) \times jump_i \times \sin(\alpha E_i), & \text{if } k > t/2 \end{cases}, \\ &k = 1, \dots, t, \quad t = TreeNumber(cc_i), \\ &jump_i = MajAx(cc_i) / (t + 1). \end{aligned} \quad (19)$$

The first case models the estimated locations at the left of the centre of the ellipse  $E_i$ , while the second case models those placed at its right. Figure 12 graphically describes the formulated procedure

to estimate tree locations. In Figure 13, the result of computing the tree potential location points is illustrated. The yielded locations are marked in red in the binary sub-image shown in Figure 11.



**Figure 12.** Illustration of the process to estimate a representative location for aggregated trees, formulated in Equations (17)–(19). Examples for odd (a) and even (b) number of aggregated trees are given.



**Figure 13.** Result of the estimation of the individual tree location points.

#### 2.4. Performance Evaluation of the Image Analysis Methodology

In order to assess the performance of the methodology proposed, it was firstly necessary to locate and determine the exact number of olive trees in the land plot under study. This was carried out by a human observer, by inspecting, labelling and counting the tree crowns appearing in the ad-hoc orthomosaic of the study site proposed in Figure 4.

The performance assessment of the methodology was approached by comparing the actual number of plants, and their distribution, to the estimations yielded by the image analysis algorithm.

In order to quantitatively evaluate this comparison, the set of metrics defined here below are proposed:

- *Precision*: it gives the hit ratio for the trees found by the algorithm. Mathematically:

$$Precision = \frac{TP}{TP + FP} \quad (20)$$

where *TP* (true positives) is the number of trees correctly identified, and conversely, *FP* (false positives) refers to the number of instances wrongly proposed by the algorithm as potential olive trees. A tree is considered to be correctly identified only when the algorithm placed its estimated location within its crown.

- *Sensitivity*: it provides the ratio of actual trees found by the algorithm:

$$Sensitivity = \frac{TP}{TP + FN} \quad (21)$$

where *FN* (false negatives) is defined as the number of actual olive trees not detected by the algorithm.

- *F1<sub>score</sub>*: it is the harmonic mean of the two metrics described above, being mathematically defined as:

$$F1_{score} = 2 \times \frac{(Precision \times Sensitivity)}{(Precision + Sensitivity)} \quad (22)$$

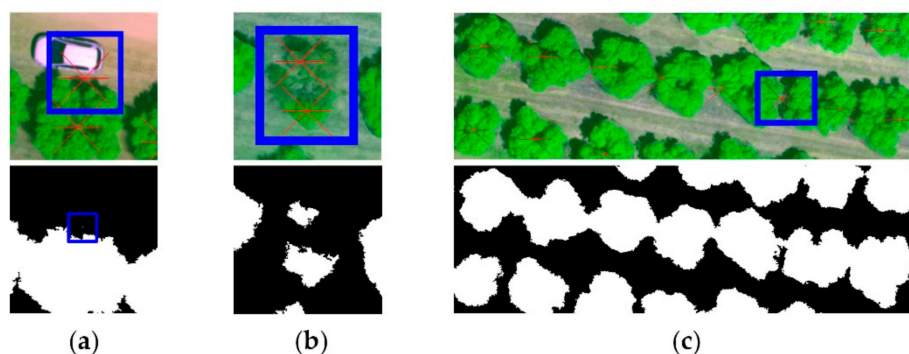
### 3. Results

According to the metrics proposed, the results provided by the presented methodology for crop trees detection, location and counting are exposed in Table 2. As it can be observed, 99.92% of tree proposals were correct, and 99.67% of the actual trees were found.

**Table 2.** Performance assessment of the automated trees detection and counting methodology, expressed in terms of the metrics defined to that purpose.

Actual Tree Population	Estimated Tree Population	TP	FP	FN	Precision	Sensitivity	F1 <sub>score</sub>
3919	3909	3906	3	13	0.9992	0.9967	0.9975

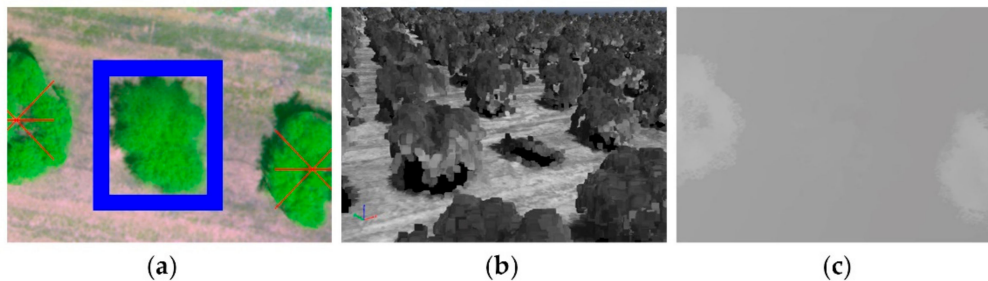
Regarding the failures detected, and focusing on the false positives (*FP*) reported, each of them can be justified by a different reason. Thus, one of them was caused by a car that was parked very close to the study site. Because of its height, it could not be discarded during image processing, neither filtered when the image was cropped according to the specified region of interest. As a result, a very small residual connected component, corresponding to this vehicle, was inevitably considered when analysing the ultimate binary image. A second false positive resulted from a tree with an anomalously damaged crown, so it was detected by the algorithm as two different plants. Finally, a last false tree detection was obtained when processing a large connected component containing seven aggregated olive trees. Due to the morphology and disposition of the overlapped tree pixel regions, the number of plants contained was overestimated in one unit. The different issues related to the false positives detected during the assessment of the methodology are illustrated in Figure 14.



**Figure 14.** False positives detected during performance assessment: (a) case related to a car parked next to the study site; (b) case consequence of wrongly splitting one tree into two different connected components because of its damaged condition; (c) case obtained after overestimating the number of trees contained in an aggregated connected component.



With respect to false negatives (*FN*), one of them was detected to come from the absence of information in the DSM, this probably due to not having enough matching points from different captures when reconstructing this part of the image. As a result, the elevation information in those corresponding points, provided by the DSM, was not significant enough to enable the discrimination of this plant (the phenomenon is illustrated in Figure 15).



**Figure 15.** False negative resulting from a lack of information in the point cloud: (a) aerial sub-image where the tree wrongly discarded by the algorithm is represented; (b) 3D point cloud-based representation of the area shown in (a); (c) elevation information provided by the DSM, represented as a grayscale image, corresponding to the area shown in (a).

In this respect, it should be noted that the density and quality of the 3D point cloud used to generate the DSM, is directly related to the overlapping with which aerial imagery is captured [25]. As commented in Section 2.2.2, the image acquisition flight and the multispectral camera setup were planned for the purpose of achieving a forward overlap of 85%. By increasing this overlapping, results could be virtually improved. However, since 99.97% of the trees were properly reconstructed, i.e., 3918 among 3919, it seems plausible to consider the setup proposed for image capture as valid. Being discarded defects in flight and image capture parametrisation, it is difficult to determine the reasons that provoked this issue, but it might be probably related to problems when capturing the aerial images, either due to weather conditions that could occasionally affect the stability of the UAV, or due to problems with the operability of the camera. Meanwhile, the rest of false negative cases detected were related to small trees, most of them in growth stage, which did not reach the minimum height (1 m) to be properly segmented from the background.

#### 4. Discussion

Table 3 compares the results of the methodology presented in this paper to those of the main published works also aimed at automated crop tree detection in orchards. A first aspect to be highlighted is that the present work outperforms the other proposals, despite the fact it was tested on a considerably greater plant population when compared to most of the reported research. Consequently, this surely included a wider variability in terms of the individual characteristics of the trees, and the way they are disposed throughout the land plot under study. Also, it should be underscored that, contrary to most of those works, this study considered challenging conditions related to overlapping intra-row tree crowns, aspect with a special impact on the accuracy with which plant population can be estimated in intensive orchards.

**Table 3.** Comparison of performance of different methods for crop trees counting published in the bibliography, and the present work.

Method	Actual Tree Population	Precision	Sensitivity	F1 <sub>score</sub>
Torres-Sánchez et al., 2015 [25]	135	–	0.945–0.969	–
Torres-Sánchez et al., 2018 [26]	–	–	0.970	–
Salamí et al., 2019 [6]	332	0.9939	0.9909	0.9924
Malek et al. [5]	617	0.9009	0.9440	0.9219
Csillik et al. [19]	2912	0.9459	0.9794	0.9624
Ampatzidis and Partel [20]	4931	0.9990	0.9970	0.9980
Selim et al. [21]	105	–	0.8286	–
Marques et al. [23]	1092	0.9944	0.9780	0.9861
This work	3918	0.9992	0.9967	0.9975

Thus, focusing on the case of the olive, a crop around which the proposed methodology was validated, counting of trees based on aerial imagery was attempted in Salamí et al. [6], obtaining a remarkable average precision of 99.84%. Nevertheless, plant detection was approached by using a circular template, imposing the prerequisite of only considering isolated trees, thus preventing that their crowns could appear overlapped in aerial captures. Contrary, the methodology presented here was able to deal with the individual location and counting of 385 trees configuring 293 aggregated connected components. Only in the case shown in Figure 14, the number of trees contained in such a component was not properly estimated. Moreover, the replicability of the methodology presented in [6] is questionable, as trees segmentation was attempted by colour discrimination. Indeed, it is very probable that any kind of natural or artificial artefact with similar colour to that of the olive tree canopies, could generate false positives. In this case, the precision of the colour segmentation approach, and consequently of the subsequent trees detection and counting, is compromised. A segmentation also based in pixel reflectance, although not only in the visible bands, followed by OBIA analysis was used in Torres-Sánchez et al. [25]. Concretely, a multi-resolution segmentation was firstly performed using the DSM and the green and NIR bands, considering colour, shape, smoothness and compactness, for which threshold values were manually adjusted. The manual decision of such key segmentation parameters questions concerns about its replicability in different situations. Furthermore, the approach requires a subsequent OBIA analysis to filter the first segmentation results. Conversely, the methodology described here proposes an analytical solution to the segmentation problem, only making use of the  $h$  parameter (equation (7)) in the segmentation step. In addition, this is a comprehensible parameter as it represents the minimum desired height in meters for the trees to be segmented. Then,  $h$  is more likely to be seen as a configuration parameter rather than a performance one. On a set of 135 olive trees, the study presented in Torres-Sánchez et al. [25] yielded sensitivity values ranging from 0.945 to 0.969, not considering the case of overlapping tree crowns. Later, the same main author and others assessed the influence of image overlap in the quality of the resulting DSM [26]. The methodology described in [25] was slightly modified and tested on an indeterminate number of trees, corroborating in the best scenario of those tested a sensibility of 0.97 in olive trees counting. The case of overlapping trees was not faced either.

Beyond the olive case, in Malek et al. [5], an overall precision of 0.9009 when detecting palm trees was achieved. They proposed a method based on training an ELM classifier on a set of key points, potentially representative of the occurrences of the trees, extracted from the initial captures. Csillik et al. [19] made also use of machine learning, concretely CNNs, for detecting citrus trees. Ampatzidis and Partel [20] also focused their research on citrus orchards, and also using CNN-based tree location. Despite the fact all these studies reported solid results, it should be noted that these kinds of machine learning solution tend to be strongly linked to the visual features of the crown trees with which they are trained. This fact makes their direct application to different kinds of crops difficult, but it surely implies the generation of new training sets and models. Contrary, the method proposed in this paper comprises an analytical solution, based on the morphological analysis and characterisation

of the general features of trees within the frame of an intensive cultivation, thus not being linked to a concrete type of crop. In Selim et al. [21], it was proposed a method for detecting orange trees from aerial imagery. The problem was undertaken in this case by means of object-based image analysis, correctly detecting 87 out of the 105 trees visible in the orthomosaic of the study case. Nevertheless, as with other researches previously referenced, difficulties were reported when dealing with overlapping tree crowns. In Kestur et al. [22], tree detection was faced on the basis of ELM- spectral and spatial classification. Despite promising results were reported about identifying trees belonging to different crops, it was not clearly specified how the training set was generated, hindering the replicability of the methodology. Marques et al. [23] proposed an effective method for detecting chestnut trees, clustering the plants by exploiting elevation data and vegetation indices (VI) information. About this latter, it should be noted that VI-based segmentations are strongly dependent of the spectral reflectivity features of the vegetation cover present in the study sites. Indeed, depending of its nature, the coverage may be confused with the plants aimed to be identified, thus potentially increasing the number of false positives yielded. This phenomenon may affect the generality of the proposed solution.

## 5. Conclusions

This investigation was undertaken in order to design and evaluate a framework for the automated identification, geolocation and counting of crop trees in intensive cultivation areas by means of UAV-based aerial imagery, multispectral sensing and image analysis techniques. The results reported support the viability of the methodology proposed as a valuable tool in phenotyping tasks, within the scope of the precision agriculture.

After testing in an olive orchard with 3919 trees, 99.67% of the plants were rightly identified, outperforming the results given by previous published work. Indeed, the algorithm designed for segmenting and analysing the data structure obtained from aerial captures, based on morphological image processing principles and the statistical analysis of the moments of tree-corresponding pixel artefacts, showed a remarkable performance in terms of tree discrimination, achieving very high detection rates. In addition, the solution also showed to be solid when dealing with multiple intra-row overlapping tree crowns. These findings should also be framed within the context of the complexity of the considered scenario, since the study plot was outstandingly larger than those used in most of previous studies, and it presented a remarkable variability in terms of soil composition, elevation and amount of weed.

Future work will test the application of the presented methodology to other types of orchards. In addition, it would be interesting to assess the performance of the algorithms when dealing with different plant spacing patterns, all of this for the sake of increasing confidence in the generality of the proposed solution.

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

**Funding:** This research and APC were funded by the INTERREG Cooperation Program V-A SPAIN-PORTUGAL (POCTEP) 2014–2020, and co-financed with ERDF funds, grant number 0155\_TECNOLIVO\_6\_E, within the scope of the TecnOlivo Project.

**Acknowledgments:** Authors would like to thank “Virgen de la Oliva” olive-oil cooperative for generously offering their orchards to conduct this work. R.S. would also like to thank Mexican National Council of Science and Technology (CONACYT) for supporting the development of this investigation.

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

## Appendix A

Mathematical morphology is a non-linear image processing technique, built from the basis of the set theory, essentially aimed at analysing the relevant structures in the image by proving this with a set

called structuring element (SE), which has an a-priori known shape and size. This appendix briefly defines the morphological operators employed in this paper, suggesting the reader to consult [30,31] for a deeper study.

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

$$f : D_f \subset \mathbb{Z}^2 \rightarrow N_0 = \{0, \dots, t_{max}\} \subset \mathbb{Z}. \quad (A1)$$

where  $t_{max}$  is the maximum value allowing to reach the data type used (e.g., 1 for binary images, 256 for 8-bit images, etc.). Thus,  $f$  maps the correspondence element 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:

$$f(\rho) \wedge g(\rho) = \min[f(\rho), g(\rho)]. \quad (A2)$$

where  $\min$  refers to the minimum operation. Conversely, the union of those two images responds to the following equation:

$$f(\rho) \vee g(\rho) = \max[f(\rho), g(\rho)], \quad (A3)$$

being  $\max$  the maximum operation.

The SE is an essential tool in mathematical morphology, used to study the shape of the objects contained in an image. Mathematically, an SE element can be seen as a binary image  $\beta$ , defining a mapping of a subset  $D_\beta$  of  $\mathbb{Z}^2$  to the subset of integer binary values  $B_0$ :

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

With this definition,  $\beta$  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, being necessary for its application that  $\#(D_\beta) < \#(D_f)$ . Common shapes implemented with SEs include circles, lines, diamonds, etc. In practice, the SE is used as a kernel, with its origin in its central pixel. Hence, an image is proven pixel by pixel to this kernel, modifying at every step the pixel in the image matching with the central pixel of the kernel, according to a given operation.

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

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

Therefore, pixel  $\rho$  in image  $f$  is modified with the minimum value of its neighbourhood according to the filter implemented by SE  $\beta$ . 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:

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

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

Combining erosion and dilation, two new operators called opening ( $\gamma$ ) and closing ( $\varphi$ ) are obtained:

$$\gamma_\beta(f) = \delta_\beta(\varepsilon_\beta(f)), \quad (A7)$$

$$\varphi_\beta(f) = \varepsilon_\beta(\delta_\beta(f)). \quad (A8)$$

Opening removes those brighter objects in the image that can be completely covered by  $\beta$ . Dually, closing removes the darker objects in the image completely covered by the SE.



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

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

Based on (9), the geodesic erosion of marker  $f$  constrained by mask  $g$  is:

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

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. It is denoted by:

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

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

$$\begin{aligned} R_g^\varepsilon(f) &= \varepsilon_g^{(i)}(f), \\ \text{where : } i &\text{ is such that :} \\ \varepsilon_g^{(i)}(f) &= \varepsilon_g^{(i+1)}(f). \end{aligned} \quad (\text{A12})$$

## References

1. Tilman, D.; Balzer, C.; Hill, J.; Befort, B.L. Global food demand and the sustainable intensification of agriculture. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 20260–20264. [[CrossRef](#)] [[PubMed](#)]
2. Araus, J.L.; Cairns, J.E. Field high-throughput phenotyping: The new crop breeding frontier. *Trends Plant Sci.* **2014**, *19*, 52–61. [[CrossRef](#)] [[PubMed](#)]
3. 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](#)]
4. Aparna, P.; Ramachandra, H.; Harshita, M.P.; Harshitha, S.; 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, 25–28 April 2018; IEEE: Bangalore, India, 2018; pp. 127–129.
5. Malek, S.; Bazi, Y.; Alajlan, N.; AlHichri, H.; Melgani, F. Efficient Framework for Palm Tree Detection in UAV Images. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *7*, 4692–4703. [[CrossRef](#)]
6. 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](#)]
7. Shakoor, N.; Lee, S.; Mockler, T.C. High throughput phenotyping to accelerate crop breeding and monitoring of diseases in the field. *Curr. Opin. Plant Biol.* **2017**, *38*, 184–192. [[CrossRef](#)]
8. Furbank, R.T.; Tester, M. Phenomics – Technologies to relieve the phenotyping bottleneck. *Trends Plant Sci.* **2011**, *16*, 635–644. [[CrossRef](#)]

9. Sankaran, S.; Khot, L.R.; Espinoza, C.Z.; Jarolmasjed, S.; Sathuvalli, V.R.; Vandemark, G.J.; Miklas, P.N.; Carter, A.H.; Pumphrey, M.O.; Knowles, N.R.; et al. Low-altitude, high-resolution aerial imaging systems for row and field crop phenotyping: A review. *Eur. J. Agron.* **2015**, *70*, 112–123. [[CrossRef](#)]
10. 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](#)]
11. Tripicchio, P.; Satler, M.; Dabisias, G.; Ruffaldi, E.; Avizzano, C.A. Towards Smart Farming and Sustainable Agriculture with Drones. In Proceedings of the 2015 International Conference on Intelligent Environments, Prague, Czech Republic, 15–17 July 2015; IEEE: Prague, Czech Republic, 2015; pp. 140–143.
12. 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](#)]
13. Peña, J.M.; Torres-Sánchez, J.; de Castro, A.I.; Kelly, M.; López-Granados, F. Weed Mapping in Early-Season Maize Fields Using Object-Based Analysis of Unmanned Aerial Vehicle (UAV) Images. *PLoS ONE* **2013**, *8*, e77151. [[CrossRef](#)]
14. Floreano, D.; Wood, R.J. Science, technology and the future of small autonomous drones. *Nature* **2015**, *521*, 460–466. [[CrossRef](#)] [[PubMed](#)]
15. Miserque Castillo, J.Z.; Laverde Diaz, R.; Rueda Guzmán, C.L. Development of an aerial counting system in oil palm plantations. *IOP Conf. Ser. Mater. Sci. Eng.* **2016**, *138*, 012007. [[CrossRef](#)]
16. Primicerio, J.; Caruso, G.; Comba, L.; Crisci, A.; Gay, P.; Guidoni, S.; Genesio, L.; Ricauda Aimonino, D.; Vaccari, F.P. Individual plant definition and missing plant characterization in vineyards from high-resolution UAV imagery. *Eur. J. Remote Sens.* **2017**, *50*, 179–186. [[CrossRef](#)]
17. 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](#)]
18. Koc-San, D.; Selim, S.; Aslan, N.; San, B.T. Automatic citrus tree extraction from UAV images and digital surface models using circular Hough transform. *Comput. Electron. Agric.* **2018**, *150*, 289–301. [[CrossRef](#)]
19. 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](#)]
20. Ampatzidis, Y.; Partel, V. UAV-Based High Throughput Phenotyping in Citrus Utilizing Multispectral Imaging and Artificial Intelligence. *Remote Sens.* **2019**, *11*, 410. [[CrossRef](#)]
21. Selim, S.; Sonmez, N.K.; Coslu, M.; Onur, I. Semi-automatic Tree Detection from Images of Unmanned Aerial Vehicle Using Object-Based Image Analysis Method. *J. Indian Soc. Remote Sens.* **2019**, *47*, 193–200. [[CrossRef](#)]
22. Kestur, R.; Angural, A.; Bashir, B.; Omkar, S.N.; Anand, G.; Meenavathi, M.B. Tree Crown Detection, Delineation and Counting in UAV Remote Sensed Images: A Neural Network Based Spectral–Spatial Method. *J. Indian Soc. Remote Sens.* **2018**, *46*, 991–1004. [[CrossRef](#)]
23. 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](#)]
24. Díaz-Varela, R.; de la Rosa, R.; León, L.; Zarco-Tejada, P. High-Resolution Airborne UAV Imagery to Assess Olive Tree Crown Parameters Using 3D Photo Reconstruction: Application in Breeding Trials. *Remote Sens.* **2015**, *7*, 4213–4232. [[CrossRef](#)]
25. Torres-Sánchez, J.; López-Granados, F.; Serrano, N.; Arquero, O.; Peña, J.M. High-Throughput 3-D Monitoring of Agricultural-Tree Plantations with Unmanned Aerial Vehicle (UAV) Technology. *PLoS ONE* **2015**, *10*, e0130479. [[CrossRef](#)] [[PubMed](#)]
26. Torres-Sánchez, J.; López-Granados, F.; Borra-Serrano, I.; Peña, J.M. Assessing UAV-collected image overlap influence on computation time and digital surface model accuracy in olive orchards. *Precis. Agric.* **2018**, *19*, 115–133. [[CrossRef](#)]
27. 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.
28. Otsu, N. A Threshold Selection Method from Gray-Level Histograms. *IEEE Trans. Syst. Man Cybern.* **1979**, *9*, 62–66. [[CrossRef](#)]
29. Jain, A.K. *Fundamentals of Digital Image Processing*; Prentice Hall: Englewood Cliffs, NJ, USA, 1989; ISBN 0133361659.

30. Soille, P. *Morphological Image Analysis: Principles and Applications*; Springer-Verlag GmbH: Heidelberg, Germany, 2004; ISBN 9783662050880.
31. Serra, J. *Image Analysis and Mathematical Morphology, vol. I*; Academic Press Inc.: Cambridge, MA, USA, 1982; ISBN 9780126372427.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).