

Article

Identification of Citrus Trees from Unmanned Aerial Vehicle Imagery Using Convolutional Neural Networks

Ovidiu Csillik ¹, John Cherbini ², Robert Johnson ³, Andy Lyons ³ and Maggi Kelly ^{3,4,*}

¹ Department of Geoinformatics, Z_GIS, University of Salzburg, 5020 Salzburg, Austria; ovidiu.csillik@sbg.ac.at

² UAV/Remote Sensing Subject Matter Expert, Berkeley, CA 94705, USA; cherbini@gmail.com

³ University of California Division of Agriculture and Natural Resources, Davis, CA 95618, USA; robjohnson@ucanr.edu (R.J.); andlyons@ucanr.edu (A.L.)

⁴ Department of Environmental Sciences, Policy and Management, University of California, Berkeley, CA 94720, USA

* Correspondence: maggi@berkeley.edu

Received: 24 October 2018; Accepted: 14 November 2018; Published: 20 November 2018



Abstract: Remote sensing is important to precision agriculture and the spatial resolution provided by Unmanned Aerial Vehicles (UAVs) is revolutionizing precision agriculture workflows for measurement crop condition and yields over the growing season, for identifying and monitoring weeds and other applications. Monitoring of individual trees for growth, fruit production and pest and disease occurrence remains a high research priority and the delineation of each tree using automated means as an alternative to manual delineation would be useful for long-term farm management. In this paper, we detected citrus and other crop trees from UAV images using a simple convolutional neural network (CNN) algorithm, followed by a classification refinement using superpixels derived from a Simple Linear Iterative Clustering (SLIC) algorithm. The workflow performed well in a relatively complex agricultural environment (multiple targets, multiple size trees and ages, etc.) achieving high accuracy (overall accuracy = 96.24%, Precision (positive predictive value) = 94.59%, Recall (sensitivity) = 97.94%). To our knowledge, this is the first time a CNN has been used with UAV multi-spectral imagery to focus on citrus trees. More of these individual cases are needed to develop standard automated workflows to help agricultural managers better incorporate large volumes of high resolution UAV imagery into agricultural management operations.

Keywords: CNN; deep learning; superpixels; precision agriculture; UAS; feature extraction; citrus; tree identification

1. Introduction

Precision agriculture plays an important role in agricultural sustainability [1,2] and since at least the 1980s, remote sensing has been a valuable component of these efforts [3]. Remote sensing has been important in soil management [4,5], pest management [6,7], weed detection [8–11], vegetation health and vigor [12], among other needs. An important research focus for efficient management of tree plantations and orchards has been developing accurate methods to delineate and enumerate individual trees from high resolution optical imagery. The combination of spectral data from individual tree canopies with field data can lead to improved yield prediction, understanding the growing characteristics or detection of anomalies in the growing process of trees [13,14].

Algorithms for feature extraction or segmentation from high resolution imagery have increased in the last decades with the proliferation of high spatial resolution imagery [15]. Many algorithms exist for

agricultural tree crown detection and enumeration using high-resolution remote sensing imagery and Lidar (Light Detection and Ranging) [16,17] and high resolution imagery alone [18–23]. For example, Santoro, et al. [24] proposed an automatic four-step procedure for individual fruit tree identification using GeoEye-1 imagery by minimizing typical background noise from orchards. Ozdarici-Ok [25] detected and delineated individual citrus trees from GeoEye-1 images using vegetation extraction, fast radial symmetry transform and simple object-based hierarchical operations. The authors showed that an object-based approach achieved higher accuracy than a pixel-based approach in the detection and delineation rates. Özcan, et al. [26] used probabilistic voting for crown detection and watershed segmentation for crown delineation in order to help forecasting the crop yield in an olive, peach, pear and citrus orchard using Google Earth images. Enumeration and delineation of oil palm trees from high-resolution satellite imagery has also received important attention [27,28].

The advent and use of unmanned aerial vehicles (UAVs) in precision agriculture represents an important milestone in the acquisition and analysis of high resolution imagery on demand [2,29,30]. UAVs represent a low-cost method for image acquisition with successful and promising usage in tree identification analysis, particularly in agricultural settings. For example, Malek, et al. [31] identified palm trees from UAV images using an extreme learning machine classifier. Jiang, et al. [32] proposed an improved scale-space filtering algorithm in a GPU-environment for papaya tree detection, achieving high accuracies with an increase in the speed of computation. In an agrosilvopastoral environment, Surovy, et al. [33] successfully estimated position and heights of tree plantations based on a 3D point cloud and orthophoto derived from UAV images. Goldbergs, et al. [34] used the canopy maxima and watershed segmentation for individual tree detection in Australian Savannas based on structure-from-motion (SfM) 3D point clouds from a low-cost UAV. Dıaz-Varela et al. [35] used a combination of DSM local maxima and object-based image analysis (OBIA) method using UAV orthomosaic imagery for the estimation of olive tree crown parameters (tree height and crown diameter). A similar approach was used by Torres-Sanchez et al. 2015 [14]. There is also extensive literature about the use of UAVs and individual plant detection to map nuisance weeds in agricultural fields [8,30,36].

Deep-learning (DL) algorithms are increasingly used in remote sensing applications [37]. Deep Learning, as a subset of Machine Learning (ML) and Artificial Intelligence (AI), includes a variety of methods that automatically learn features from large amounts of data and perform well with large numbers of training samples [38]. And while AI, ML and DL algorithms were introduced into the geoscience and remote sensing community some time ago [39,40], we are experiencing their increased use and appreciation [41–44] for their ability to accurately solve common remote sensing analysis problems including classification and segmentation. Common DL algorithms include various neural networks and other unsupervised and supervised feature-learning algorithms. Deep-learning techniques play an important role in agriculture applications, especially for plant enumeration and identification [45].

Recently, many researchers have been focusing on convolutional neural networks (CNNs) for the segmentation of high resolution imagery. CNNs are hierarchical architectures that can be trained on large-scale datasets to perform object recognition and detection [46]. For example, Guirado, et al. [47] compared convolutional neural networks (CNNs) with OBIA methods for scattered shrub detection from Google Earth images, achieving higher performance with less human intervention and transferability of the CNN model used. Li, et al. [45] detected and counted oil palm trees in Malaysia from QuickBird images using a CNN framework. They detected more than 96% of the oil palm trees, outperforming three other state-of-the-art methods for tree detection (i.e., local maxima, template matching and artificial neural networks). With a more detailed analyses, Chen, et al. [48] counted apples and orange fruits in an unstructured environment based on fully convolutional networks and Wang et al. [49] and Sa et al. [50] implemented Faster Region-based CNN (Faster R-CNN) algorithms for mango fruit flower detection and fruit detection (sweet pepper and melon), respectively.

The exploitation of deep learning techniques for tree identification from UAV imagery is still under development and, to our knowledge, a CNN classification has not been applied yet for multi-age citrus tree identification in an agricultural orchard. In this study, we provide a simple cases study of citrus trees detection from UAV images using a combination of a CNN framework with post-processing using object-based image analysis. Our intent was not to test this methods' performance in all conditions or with numerous targets: this is a case study using imagery collected in ideal conditions, as is common practice with agricultural managers, to examine the performance of a novel algorithm, to provide the impetus for further exploration. The study area is located in one of the world's most productive agricultural areas (San Joaquin Valley, California) and is part of the Lindcove Research and Extension Center of the University of California.

2. Materials and Methods

2.1. Study area

The study area is located at the Lindcove Research and Extension Center (LREC) in Tulare County, CA (36.360981°N, −119.062048°W) (Figure 1). The Center has been in operation since 1959 and covers 75 ha (190 acres) in the San Joaquin Valley. The Center grows nearly 600 tree crop varieties, mostly citrus species, with trees of various ages and sizes. The San Joaquin Valley has hot, dry summers and cool rainy winters and is highly productive agriculturally, contributing the majority of the California's agricultural production. Citrus is an important crop planted on over 100,000 ha (250,000 acres) in California and worth nearly \$2B annually [51]. Field and laboratory research at LREC focuses on the citrus crop but also includes projects on avocado, olive, walnut and pomegranate crops [52]. LREC research includes the development of new citrus varieties, determining the best rootstock and fruit producing (citrus, avocado, walnut and pomegranate) combinations for San Joaquin Valley conditions and developing and implementing novel methods for pest and disease control that protect natural enemies and honey bees. Experimental design includes planting of research trees in fields of common or mixed crop varieties, with same or varying rootstock and of similar or varying age. A reference tree database of 2,912 individual trees was created previously by manually locating individual trees on high resolution (1 m) National Aerial Imagery Program (NAIP) imagery from 2012. Each tree in the point file was attributed with its location, variety, rootstock and planting date. This database was used for subsequent validation.

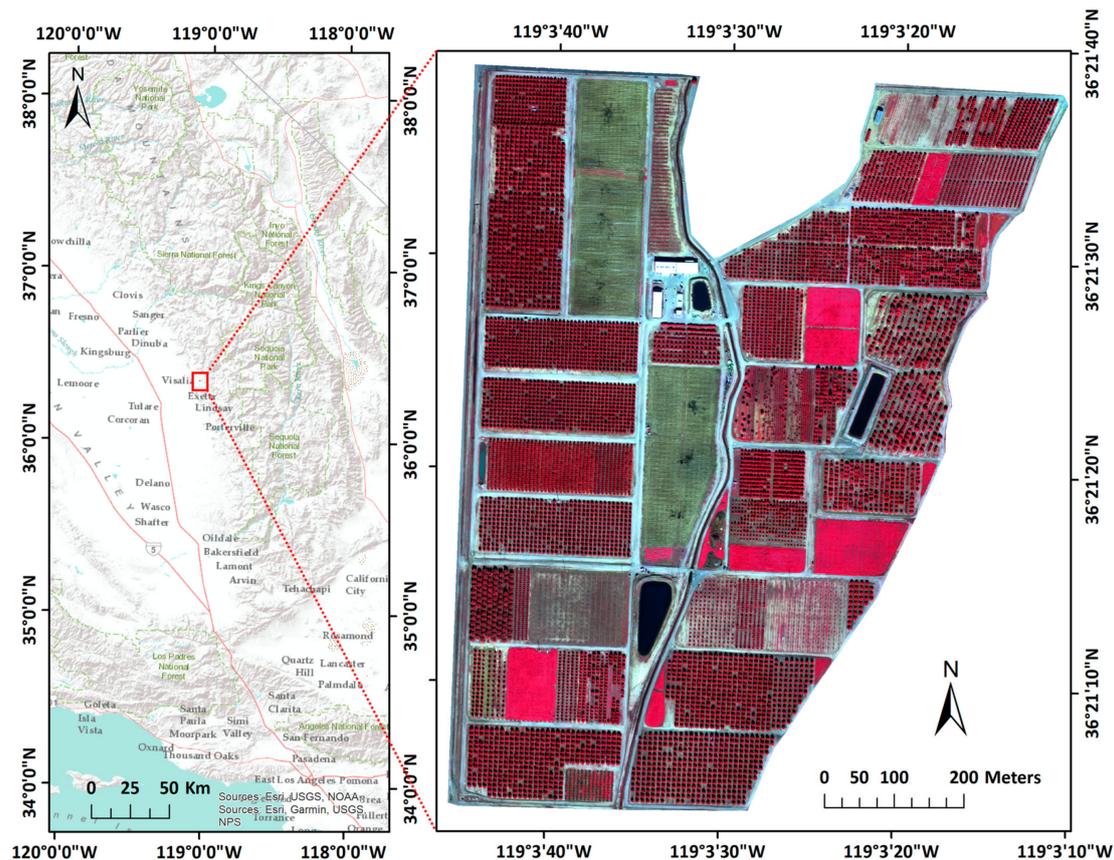


Figure 1. Study area location near Visalia, California (left) and a false color image as acquired by the UAV with a spatial resolution of 0.12 m and covering 64.6 ha (right).

2.2. UAV Imagery Collection

Imagery was acquired January 31st, 2017 in clear conditions between the hours of 13:00 and 14:00 to minimize effects of shadowing on the images. The imagery was collected as a proof-of-concept, as an alternative to NAIP as a high resolution image source for subsequent management and research planning. A senseFly eBee fixed-wing UAV was used (wingspan of 109.7 cm) with a Parrot Sequoia multi-spectral camera with average along- and across-track overlaps of 75% and 75%, respectively. To maximize resolution and coverage area, the UAS was operated at a consistent 104m AGL using SRTM terrain data to ensure consistent altitudes across the captured area. The Parrot Sequoia camera has four sensors (green: 530–570 nm, red: 640–680 nm, red edge: 730–740 nm and near-infrared: 770–810 nm), a RGB composite sensor and an external irradiance sensor with GPS and IMU placed on top of the UAV to capture sensor angle, sun angle, location and irradiance for each image during flight. This additional sensor data were primarily used for image calibration. Two image acquisition missions were performed following a standard flight protocol where physical radiometric targets were imaged prior to flight and a systematic gridded flight pattern was used. 4,574 multispectral images were collected (flight 1 = 3,985 and flight 2 = 589) (Figure 2). Flight planning and mission control software was SenseFly's eMotion software [53].

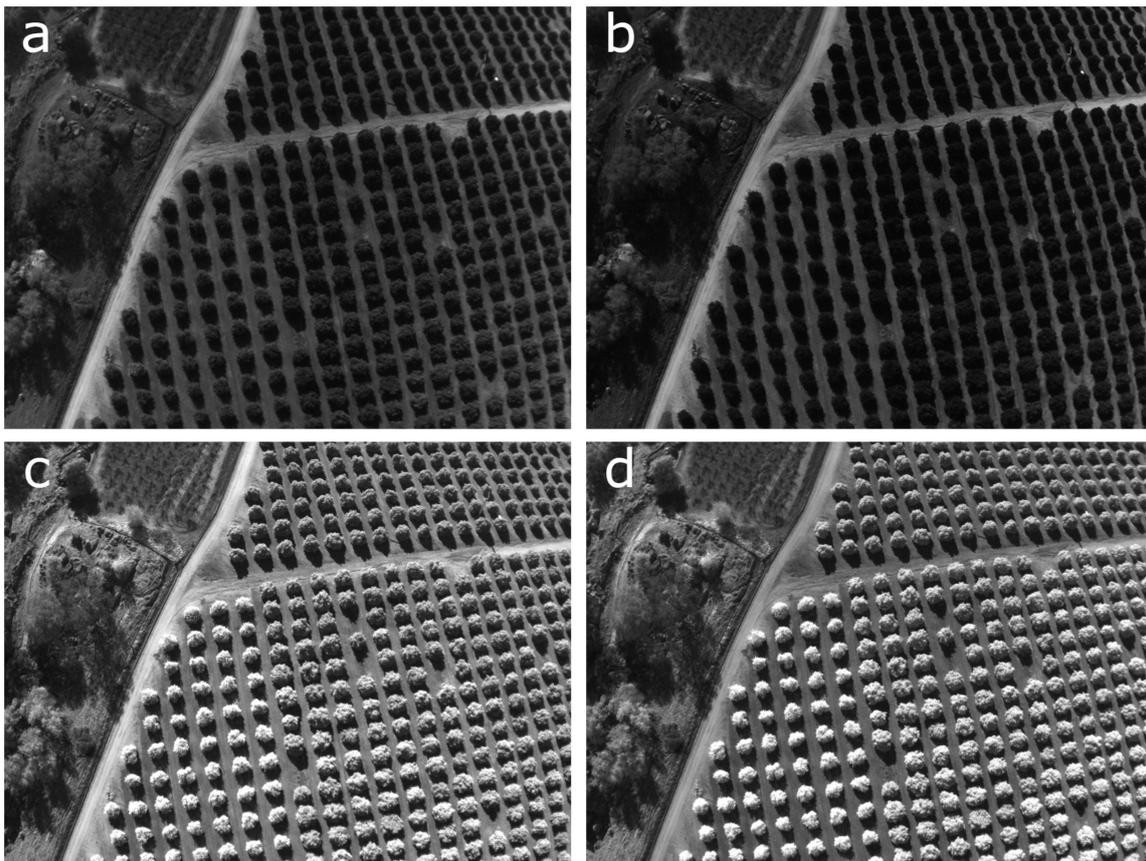


Figure 2. Examples of green (a), red (b), red edge (c) and near infrared (d) bands as acquired by the senseFly eBee fixed-wing UAS using a Parrot Sequoia multi-spectral camera.

2.3. Photogrammetric and Image Processing

Imagery was photogrammetrically processed using Pix4D Mapper software [54]. To ensure absolute measurements, radiometric target data was used to calibrate the reflectance values prior to the photogrammetric process. Image tie-points were then generated for locations with two or more overlapping images and used to produce a 4-band ortho-imagery mosaic with GSD of 12.841 cm in WGS1984, UTM Zone 11N projection (Figure 2). The red and near infrared bands were used to create a Normalized Difference Vegetation Index (NDVI) image using the formula $(NIR - Red) / (NIR + Red)$.

2.4. CNN Workflow

The analysis was developed and tested using a 64-bit operating system, with Intel® Core™ i5-6500 CPU @ 3.20 GHz and 16 GB RAM. We applied the CNN workflow using Trimble's eCognition Developer 9.3 [55], which is based on the Google TensorFlow API [56]. Trimble's eCognition Developer is one of the most popular software for object-based image analysis and the application of the CNN using this platform gave the opportunity of integrating the CNN approach with object-based post-processing of the results, thus performing the entire analysis in one software. Application of the CNN in eCognition consisted of three steps: (1) derivation of 4,000 training samples of 40×40 pixels for the three classes used (5 minutes), (2) training the CNN model (13 minutes) and (3) applying the trained CNN model to the validation area not used in training (2 minutes).

2.4.1. CNN Training and Classification

Finding the best architecture for the CNN is an ongoing debate in the world of deep learning. It usually starts from a simple model and the hyper-parameters are tuned iteratively until a good model is found for a specific application. Our study area was first divided into a training area in

the north and a validation area in the south. An initial CNN was trained with three classes - trees, bare soil and weeds - with 4,000 samples per class. The samples were derived based on the dataset of manually identified individual trees (trees in the north used for training and trees in the south used for validation) and randomly generated samples in areas without trees, which were classified as bare soil or weeds. To increase the number of training samples and the robustness of CNN, we derived a 3×3 pixel buffer area around our training point samples. In this way, every tree is now represented by 9 pixels around the center of the tree and the algorithm will randomly choose locations out of these 9 pixels. Training samples were patches of 40×40 pixels because this size best matched the size of the targeted trees. Trial-and-error approaches with sample sizes smaller or bigger than 40×40 pixels were tried; values smaller than this increased the multiple-crown detection errors while values bigger than 40×40 missed some of the small trees. A size of 40×40 was also in line with most of the tree sizes in our study areas (Figure 3).

All 4 spectral bands from the UAV dataset were used in the training step (i.e. green, red, near infrared and red edge). Examples of the training samples are shown in Figure 3. We chose a simple CNN model that uses one hidden layer that convolve the input layers using different kernels and generating different feature maps [57]. For this hidden layer, a kernel size of $4 \times 11 \times 11$ (4 bands and 11×11 pixels) was used for convolution and 40 distinct feature maps were generated. Max pooling was applied to reduce the resolution of the feature maps using a 2×2 filter with a stride of 2, both in horizontal and vertical directions. During the training of the CNN, the learning rate was set to 0.0015 after trial-and-error tests. This parameter dictates the amount by which weights are adjusted during the statistical gradient descent optimization. Lower values of the learning rate can slow down the process of training or finding suboptimal weights by finding local minima, while higher values will improve speed but increase the risk of missing the optimal minimum [57]. We used 5,000 training steps with 50 training samples used at each training step.

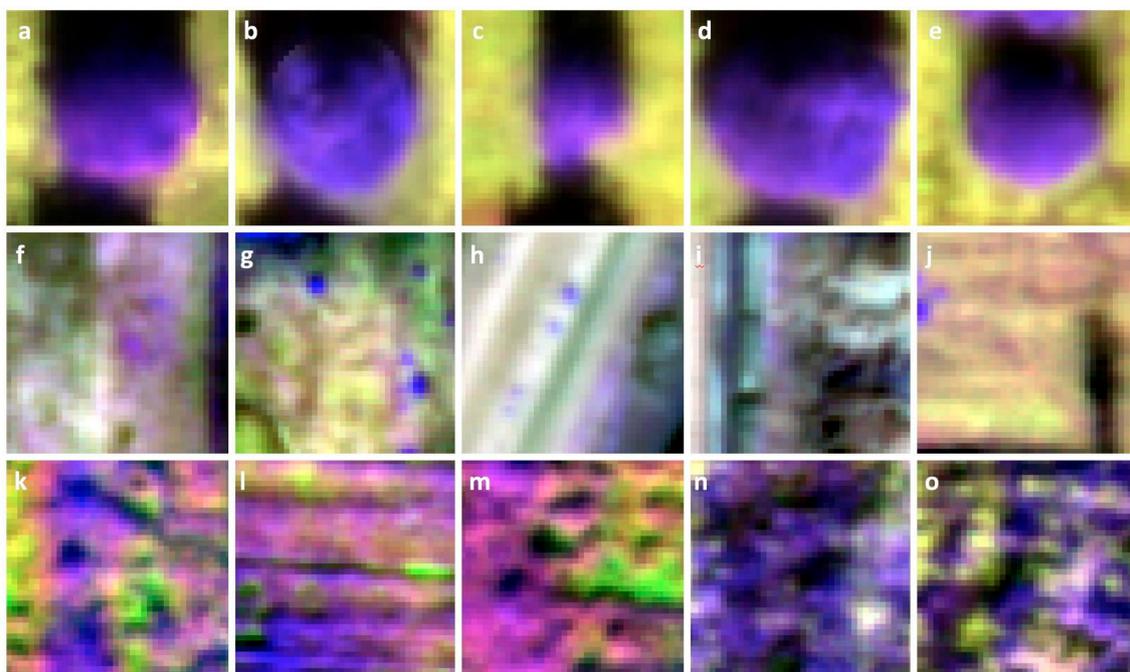


Figure 3. Example of training sets used. Each image is an example of one 40×40 pixel training for: (a–e) trees, (f–j) bare ground and (k–o) weeds.

The initial CNN model ran quickly (13 minutes) and produced a heatmap showing probability values of tree detection - closer to 1, better the likelihood of target presence. To extract individual tree locations, we further smoothed the resulted heatmap with a 15×15 Gaussian filter and detected the

trees by finding the local maxima in a 9×9 pixel neighborhood, with values higher than a certain threshold (0.5).

This initial classification resulted in some confusion between weeds located at the edges of parcels and trees and there was some difficulty in distinguishing small trees from large trees. The first problem was solved by removing detected “trees” with low NDVI (<0.7). The second challenge required additional processing, detailed in Section 2.4.2.

2.4.2. Classification refinement

Trees with large canopies were identified multiple times by the CNN workflow. In order to aggregate the targets representing the same tree, we segmented the heatmap produced by CNN (the probability of tree detection) and the NDVI layer into superpixels. In this case, the process constrained the segmentation towards circular objects in the heatmap and constrained the segmentation to differentiate between green and bare soil in the NDVI layer. Superpixels have been used as a pre-processing stage to over-segment high resolution imagery into low-level groups, thus simplifying later computation [58]. We used the Simple Linear Iterative Clustering (SLIC) algorithm [58,59] because of its simplicity and applicability. SLIC has only one parameter, k , which is the desired number of equally sized superpixels to be generated. Since we were interested in larger trees identified multiple times, we used larger superpixels sizes than the CNN training patch size (i.e., 40×40 pixels), because we assumed that trees larger than this size were generally detected multiple times. Superpixels sizes were 40×40 , 41×41 , 42×42 , . . . 50×50 pixels, values higher than these did not improve accuracy. For each iteration, superpixels that contained multiple points detected as tree were selected. Only those with shape close to a circle (Roundness less than 0.5; values closer to 0 indicate a perfect circle) and with a symmetric shape (asymmetry less than 0.5; this feature computes the variance in x-direction and y-direction and values closer to 0 indicate a symmetric shape), thus avoiding creating an object from two consecutive trees, were selected, the points merged and the centroid of the superpixels was selected as the newly identified tree (Figure 4).

2.5. Validation

We used the manually delimited reference tree dataset ($n = 2,912$) for validation. As the reference dataset included only trees, our validation process focused on trees alone and ignored weeds and bare ground. Each tree identified by the CNN method was compared to its closest manually located tree. We used common evaluation statistics for binary classification [60–62] and calculated True Positives (TP) (a tree is correctly identified), False Positives (FP) (a tree is incorrectly identified; a commission error) and False Negatives (FN) (a tree is missed; an omission error). TP, FN and FP indicate perfect identification, under-identification and over-identification, respectively. We then calculated Precision (P), Recall (R) and F-score (F). *Precision* (i.e., positive predictive value) describes the correctness of detected trees and how well the algorithm dealt with false positives (Equation (1)), *Recall* (i.e., sensitivity) describes the tree detection rate and how well the algorithm dealt with false negatives (Equation (2)) and the *F-score* is the harmonic mean of Recall and Precision and reports the overall accuracy taking both commission and omission errors into consideration (Equation (3)).

$$P = TP / (TP + FP) \quad (1)$$

$$R = TP / (TP + FN) \quad (2)$$

$$F\text{-score} = 2 * ((P * R) / (P + R)) \quad (3)$$

3. Results

The first pass at classification using CNN performed well but resulted in some confusion between weeds located at the edges of parcels and trees and was challenged in distinguishing small trees from large trees and in many cases, multiple crowns were detected within single large trees (Figure 4d).

The second classification phase iteratively segmented the smoothed heatmap and NDVI layer using SLIC superpixels of sizes larger than 40×40 pixels (Figure 4e,f) reduced the number of mistaken tree identifications results and resulted in better matching with the ground reference samples (Figure 4i).

The final classified image is shown in Figure 5. The CNN approach accurately detected many similar sized trees in the southern portion of LREC (the test region) (Figure 5a) and effectively dealt with medium size trees (Figure 5b), similar sized trees (Figure 5d) and heterogeneous tree canopy sizes (Figure 5f) that were correctly classified. Some refinements were needed where large canopy trees were present to reduce the effect of multiple crown detection (Figure 5c). In the final classification, a total of 3015 individual trees were detected, compared to 2912 reference trees. The TP (truly detected trees) was 2852 FN (missed trees) was 60 and FP (falsely added trees) was 163. Higher value of falsely detected trees was influenced by former existing trees with in-between weeds that left a pattern in the landscape similar with a tree canopy (e.g., 72 false detected trees in southern part of the test area). Missed trees are mostly related to newly planted trees with very small canopies that are not detectable by the algorithm.

By combining an object-based post-processing approach for the CNN results of tree detection, the accuracy of the results was improved significantly. Without refinement, the CNN approach achieved an F-score of 78%, Precision was 65% and Recall 98%. After reducing the effect of multiple crown detection, the overall accuracy (F-score) for the final classification was 96.24%, with a Precision of 94.59% and Recall of 97.94%.

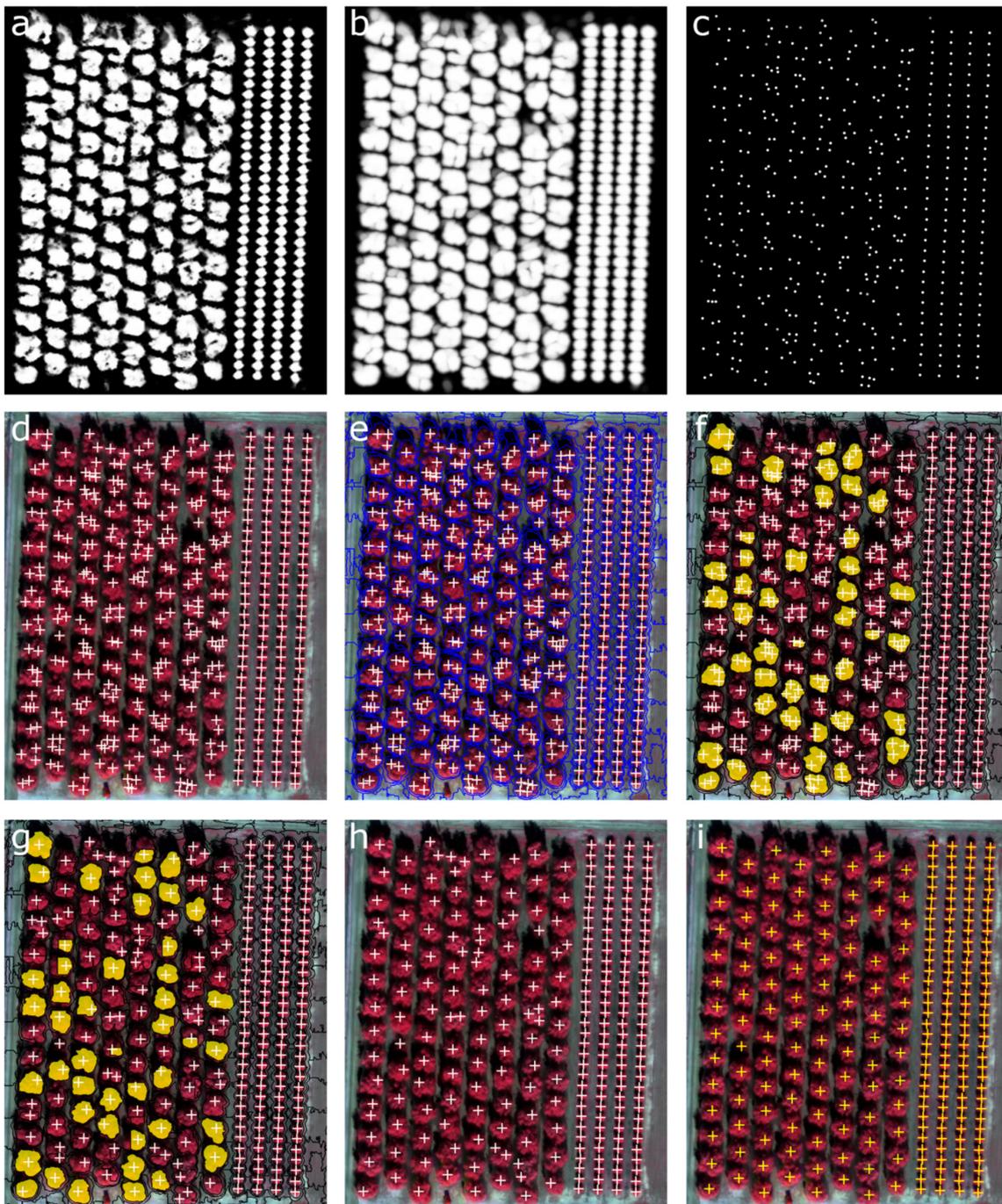


Figure 4. Probability heatmap of tree presence resulted from CNN with values between 0 and 1 (black to white) (a). To reduce the effects of possible noise, the heatmap was smoothed (b) and trees were identified using a threshold of 0.5 (c). The subset depicted here demonstrates the problem of multiple crown detection for large trees (d) that was solved by iteratively segmenting the smoothed heatmap and NDVI layer using SLIC superpixels of sizes larger than 40x40 pixels (e). Multiple crown detection were identified (f) and a single hit per tree was recomputed (g). The results after the refinement (h) better matched the ground reference samples (i).

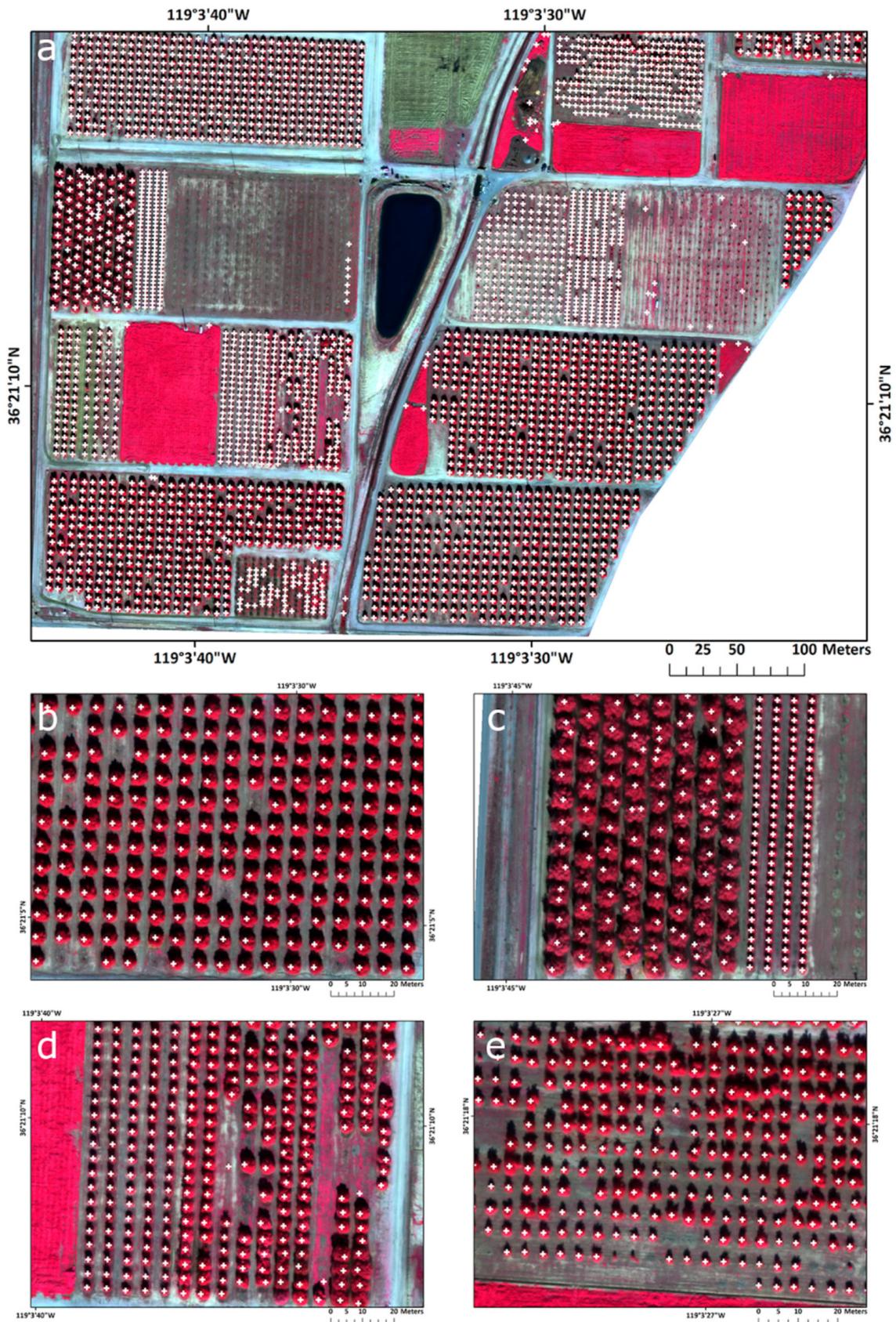


Figure 5. Final tree detection from the test area with white crosses indicating the location of trees: (a) the southern portion of LREC; (b) medium size trees correctly classified; (c) large canopy trees with reduced effect of multiple crown detection, after the object-based refinement; (d) similar sized trees were correctly classified; and (e) heterogeneous tree canopy sizes were correctly classified.

4. Discussion

In the last decade, there has been a proliferation in literature reviewing the application of UAV imagery collection and analysis for ecological applications, natural resource monitoring and agricultural management [29,63,64]. Literature concerning the use of UAVs for agriculture application are increasing rapidly, particularly in the arena of precision agriculture. UAV workflows are increasingly being added to agricultural management for precise observation, measurement and monitoring of crop condition over the growing season [14,65–68]; for estimation and measurement of crop yields [69–71], for identifying and monitoring weeds [8,30,36] and for the precise application of pesticides or fertilizers [72]. Core to these needs is the fundamental task of feature (or object) extraction.

In an agricultural setting, feature extraction involves the ability to extract from high resolution optical imagery individual plants (e.g., trees and weeds) from the image mosaic. And while agricultural settings can be more uniform than natural ones, the case in this paper focused on a complex orchard environment in a research setting, with multiple varieties of tree crops of different ages grown, resulting in high scene spatial variability. In such cases, the complexity of spatial and structural patterns of targets in high resolution imagery can make feature extraction difficult. In our case, we used a two-step method that employed CNN as an initial classifier and then a Simple Linear Iterative Clustering (SLIC) algorithm to refine results and successfully deal with the multi-scale problem of varying tree age.

Our accuracy results are high (Overall Accuracy, F-score = 96.24) and comparable with other studies that used UAV imagery for tree detection. Torres-Sánchez et al. used DSMs from UAV-based multi-spectral data with OBIA methods to calculate the geometric features of individual olive trees ($n = 135$) resulting in 97% accuracy [14]. Using different methods, Wallace et al. investigated the ability to detect individual trees in boreal forests using UAV-based hyperspectral imagery and the associated point cloud [73]. They were examining a comparable number of trees (4151 reference trees) as in our study but in a forested environment. In a smaller study (367 reference trees) using a local-maxima based algorithm on UAV-derived Canopy Height Models, Mohan et al. reported an F-score of 86% in an open canopy mixed conifer forest in Wyoming, USA [74]. Not surprisingly, UAV-based LiDAR data yields very high overall accuracies (typically around 98%) for individual tree detection [75,76]. Ours is the first study to use CNN on the imagery data alone to achieve highly accurate individual tree detection.

UAV imagery collection for agricultural applications is increasing globally and more of these individual cases are needed to develop more standard workflows that will help field and research managers deal with large volumes of high resolution imagery. UAV collected imagery can be extremely high resolution: ground sample distance (GSD) can range from a few cm [14,33] to much higher, depending on flying height and camera. At these fine resolutions, image sizes will add up. For example, a 4-band image mosaic of a 100 ha (250 acre) study area at 10 cm GSD can be moderate in size (e.g., 0.5 Gb) but the complete data package containing the multiple input images and final photogrammetric products can be 5 Gb in size. With repeat missions over the growing season and over larger study areas, data volumes will rise.

While this work highlighted one case study using a simple convolutional neural network, deeper CNNs with multiple hidden layers or ability to regionalize data should be tested to understand improvements. For example, newer CNN algorithms called Region Based CNNs (e.g., Fast R-CNN, Faster R-CNN) that detect objects by identifying regions that are most likely to contain the object to be identified [77,78] are proving to be fast and efficient in detecting objects. These have most often been used in detection of cars (e.g., [79]) but there is a growing literature around detection of agricultural features such as flowers and fruit [49,50].

Individual case studies are critical in a new and growing field and integration of multiple missions over larger study areas and with different targets can provide elements for larger projects focused on: generalizable results across conditions and targets [77]; camera, platform and protocol benchmarking (e.g., as has been done for airborne LiDAR [80]), data fusion and scaling [81] and multi-temporal analysis. For example, there is a burgeoning field of remote sensing analysis dedicated

to multi-temporal data analysis. The increase in multitemporal scenes acquired by the Landsat and Sentinel satellites and cubesats such as Planet have catalyzed growth in multi-temporal and multi-scalar image fusion algorithms [82–84] that assist in understanding land cover and agricultural dynamics. UAVs will play a role here: the ability to acquire high-resolution imagery on demand through the growing season will necessarily mean that these data will be mined for their ability to provide multi-temporal insights to agricultural plant phenology, yield, vigor and so forth. However, UAV images do have additional challenges when compared to satellite based platforms. For example, multi-sensor fusion is challenged by poor georegistration of multi-date UAV images that can limit the effectiveness of any subsequent analysis of agricultural dynamics [85]. Additionally, flying height, camera specifications and lack of spectral calibration can influence the final accuracy of resulting image mosaics [86].

5. Conclusions

Methods to delineate, enumerate and monitor individual trees in an agricultural setting from high resolution optical imagery are required for efficient and precise crop management. Monitoring of individual trees for growth, fruit production and pest and disease occurrence remains a high research and operational priority and the delineation of each tree using automated methods as an alternative to manual delineation will be useful for future long-term crop management. There are many methods for feature extraction, for example, object-based image analysis (OBIA) [8,15,20] and watershed segmentation from DSM [14,35] but these methods can struggle to deal effectively with multi-scale objects and landscapes with complex spatial patterning. Machine Learning methods in remote sensing for classification and segmentation are on the increase for their ability to accurately solve common remote sensing analysis problems in large datasets given sufficient training. In this paper, we showed a simple ML method (convolutional neural networks) performed well on a relatively complex (multiple targets, multiple size trees, etc.) tree crop research property with high accuracy (overall accuracy = 96.24%, Precision (positive predictive value) = 94.59%, Recall (sensitivity) = 97.94%). To our knowledge, this is the first time a CNN has been used with UAV imagery to focus on the citrus crop. As UAV imagery collection for agricultural applications is increasing globally, more of these individual cases are needed to develop more standard workflows that will help field and research managers better incorporate large volumes of high resolution imagery into their management operations.

Author Contributions: Conceptualization, O.C. and M.K.; Data curation, O.C., J.C., R.J. and A.L.; Formal analysis, O.C.; Methodology, O.C., J.C. and M.K.; Project administration, M.K.; Validation, M.K.; Writing—original draft, O.C. and M.K.; Writing—review & editing, O.C., A.L. and M.K.

Funding: The work of Ovidiu Csillik was supported by the Austrian Science Fund (FWF) through the Doctoral College GIScience (DK W1237-N23).

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

References

1. Bongiovanni, R.; Lowenberg-Deboer, J. Precision Agriculture and Sustainability. *Precis. Agric.* **2004**, *5*, 359–387. [[CrossRef](#)]
2. 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](#)]
3. 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](#)]
4. Ali, I.; Greifeneder, F.; Stamenkovic, J.; Neumann, M.; Notarnicola, C. Review of Machine Learning Approaches for Biomass and Soil Moisture Retrievals from Remote Sensing Data. *Remote Sens.* **2015**, *7*, 16398–16421. [[CrossRef](#)]

5. Carrão, H.; Russo, S.; Sepulcre-Canto, G.; Barbosa, P. An empirical standardized soil moisture index for agricultural drought assessment from remotely sensed data. *Int. J. Appl. Earth Obs. Geoinf.* **2016**, *48*, 74–84. [[CrossRef](#)]
6. Kelly, M.; Guo, Q. Integrated agricultural pest management through remote sensing and spatial analyses. In *General Concepts in Integrated Pest and Disease*; Springer: Dordrecht, The Netherlands, 2007.
7. Acharya, M.C.; Thapa, R.B. Remote sensing and its application in agricultural pest management. *J. Agric. Environ.* **2015**, *16*, 43–61. [[CrossRef](#)]
8. 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](#)] [[PubMed](#)]
9. Thorp, K.R.; Tian, L.F. A Review on Remote Sensing of Weeds in Agriculture. *Precis. Agric.* **2004**, *5*, 477–508. [[CrossRef](#)]
10. Lamb, D.W.; Brown, R.B. PA—Precision Agriculture: Remote-Sensing and Mapping of Weeds in Crops. *J. Agric. Eng. Res.* **2001**, *78*, 117–125. [[CrossRef](#)]
11. López-Granados, F.; Jurado-Expósito, M. Using remote sensing for identification of late-season grass weed patches in wheat. *Weed Sci.* **2006**, *54*, 346–353. [[CrossRef](#)]
12. Berni, J.A.J.; Zarco-Tejada, P.J.; Suarez, L.; Fereres, E. Thermal and Narrowband Multispectral Remote Sensing for Vegetation Monitoring From an Unmanned Aerial Vehicle. *IEEE Trans. Geosci. Remote Sens.* **2009**, *47*, 722–738. [[CrossRef](#)]
13. Costes, E.; Lauri, P.E.; Regnard, J.L. Analyzing fruit tree architecture: Implications for tree management and fruit production. *Hortic. Rev.* **2006**, *32*, 1–61.
14. 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](#)]
15. Blaschke, T.; Hay, G.J.; Kelly, M.; Lang, S.; Hofmann, P.; Addink, E.; Queiroz Feitosa, R.; van der Meer, F.; van der Werff, H.; van Coillie, F.; et al. Geographic Object-Based Image Analysis - Towards a new paradigm. *ISPRS J. Photogramm. Remote Sens.* **2014**, *87*, 180–191. [[CrossRef](#)] [[PubMed](#)]
16. Verma, N.K.; Lamb, D.W.; Reid, N.; Wilson, B. Comparison of Canopy Volume Measurements of Scattered Eucalypt Farm Trees Derived from High Spatial Resolution Imagery and LiDAR. *Remote Sens.* **2016**, *8*, 388. [[CrossRef](#)]
17. Jakubowski, M.K.; Li, W.; Guo, Q.; Kelly, M. Delineating Individual Trees from Lidar Data: A Comparison of Vector- and Raster-based Segmentation Approaches. *Remote Sens.* **2013**, *5*, 4163–4186. [[CrossRef](#)]
18. Larsen, M.; Eriksson, M.; Descombes, X.; Perrin, G.; Brandtberg, T.; Gougeon, F.A. Comparison of six individual tree crown detection algorithms evaluated under varying forest conditions. *Int. J. Remote Sens.* **2011**, *32*, 5827–5852. [[CrossRef](#)]
19. Erikson, M.; Olofsson, K. Comparison of three individual tree crown detection methods. *Mach. Vis. Appl.* **2005**, *16*, 258–265. [[CrossRef](#)]
20. Ke, Y.; Quackenbush, L.J. A review of methods for automatic individual tree-crown detection and delineation from passive remote sensing. *Int. J. Remote Sens.* **2011**, *32*, 4725–4747. [[CrossRef](#)]
21. Tao, S.; Wu, F.; Guo, Q.; Wang, Y.; Li, W.; Xue, B.; Hu, X.; Li, P.; Tian, D.; Li, C.; et al. Segmenting tree crowns from terrestrial and mobile LiDAR data by exploring ecological theories. *ISPRS J. Photogramm. Remote Sens.* **2015**, *110*, 66–76. [[CrossRef](#)]
22. Yu, Q.; Gong, P.; Clinton, N.; Biging, G. Object-based detailed vegetation classification with airborne high spatial resolution remote sensing imagery. *Photogram. Eng. Remote Sens.* **2006**, *72*, 799–811. [[CrossRef](#)]
23. Guo, Q.; Kelly, M.; Gong, P.; Liu, D. An Object-Based Classification Approach in Mapping Tree Mortality Using High Spatial Resolution Imagery. *GISci. Remote Sens.* **2007**, *44*, 24–47. [[CrossRef](#)]
24. Santoro, F.; Tarantino, E.; Figorito, B.; Gualano, S.; D’Onghia, A.M. A tree counting algorithm for precision agriculture tasks. *Int. J. Digit. Earth* **2013**, *6*, 94–102. [[CrossRef](#)]
25. Ozdarici-Ok, A. Automatic detection and delineation of citrus trees from VHR satellite imagery. *Int. J. Remote Sens.* **2015**, *36*, 4275–4296. [[CrossRef](#)]
26. Özcan, A.H.; Hisar, D.; Sayar, Y.; Ünsalan, C. Tree crown detection and delineation in satellite images using probabilistic voting. *Remote Sens. Lett.* **2017**, *8*, 761–770. [[CrossRef](#)]

27. Srestasathiern, P.; Rakwatin, P. Oil Palm Tree Detection with High Resolution Multi-Spectral Satellite Imagery. *Remote Sens.* **2014**, *6*, 9749–9774. [[CrossRef](#)]
28. Shafri, H.Z.M.; Hamdan, N.; Saripan, M.I. Semi-automatic detection and counting of oil palm trees from high spatial resolution airborne imagery. *Int. J. Remote Sens.* **2011**, *32*, 2095–2115. [[CrossRef](#)]
29. Hogan, S.D.; Kelly, M.; Stark, B.; Chen, Y. Unmanned aerial systems for agriculture and natural resources. *Calif. Agric.* **2017**, *71*, 5–14. [[CrossRef](#)]
30. Torres-Sánchez, J.; López-Granados, F.; De Castro, A.I.; Peña-Barragán, J.M. Configuration and specifications of an Unmanned Aerial Vehicle (UAV) for early site specific weed management. *PLoS ONE* **2013**, *8*, e58210. [[CrossRef](#)] [[PubMed](#)]
31. 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](#)]
32. 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](#)]
33. Surový, P.; Almeida Ribeiro, N.; Panagiotidis, D. Estimation of positions and heights from UAV-sensed imagery in tree plantations in agrosilvopastoral systems. *Int. J. Remote Sens.* **2018**, *39*, 4786–4800. [[CrossRef](#)]
34. Goldbergs, G.; Maier, S.W.; Levick, S.R.; Edwards, A. Efficiency of Individual Tree Detection Approaches Based on Light-Weight and Low-Cost UAS Imagery in Australian Savannas. *Remote Sens.* **2018**, *10*, 161. [[CrossRef](#)]
35. Díaz-Varela, R.A.; de la Rosa, R.; León, L.; Zarco-Tejada, P.J. 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](#)]
36. Peña, J.M.; Torres-Sánchez, J.; Serrano-Pérez, A.; de Castro, A.I.; López-Granados, F. Quantifying efficacy and limits of unmanned aerial vehicle (UAV) technology for weed seedling detection as affected by sensor resolution. *Sensors* **2015**, *15*, 5609–5626. [[CrossRef](#)] [[PubMed](#)]
37. Kussul, N.; Lavreniuk, M.; Skakun, S.; Shelestov, A. Deep Learning Classification of Land Cover and Crop Types Using Remote Sensing Data. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 778–782. [[CrossRef](#)]
38. Liu, T.; Abd-Elrahman, A.; Morton, J.; Wilhelm, V.L. Comparing fully convolutional networks, random forest, support vector machine, and patch-based deep convolutional neural networks for object-based wetland mapping using images from small unmanned aircraft system. *GISci. Remote Sens.* **2018**, *55*, 243–264. [[CrossRef](#)]
39. Estes, J.E.; Sailer, C.; Tinney, L.R. Applications of artificial intelligence techniques to remote sensing. *Prof. Geogr.* **1986**, *38*, 133–141. [[CrossRef](#)]
40. Myneni, R.B.; Maggion, S.; Jaquinta, J.; Privette, J.L.; Gobron, N.; Pinty, B.; Kimes, D.S.; Verstraete, M.M.; Williams, D.L. Optical remote sensing of vegetation: Modeling, caveats, and algorithms. *Remote Sens. Environ.* **1995**, *51*, 169–188. [[CrossRef](#)]
41. Ball, J.E.; Anderson, D.T.; Chan, C.S. Comprehensive survey of deep learning in remote sensing: Theories, tools, and challenges for the community. *JARS* **2017**, *11*, 042609. [[CrossRef](#)]
42. Zhang, L.; Zhang, L.; Du, B. Deep Learning for Remote Sensing Data: A Technical Tutorial on the State of the Art. *IEEE Geosci. Remote Sens. Mag.* **2016**, *4*, 22–40. [[CrossRef](#)]
43. Kamilaris, A.; Prenafeta-Boldú, F.X. Deep learning in agriculture: A survey. *Comput. Electron. Agric.* **2018**, *147*, 70–90. [[CrossRef](#)]
44. Han, X.; Zhong, Y.; Zhang, L. An Efficient and Robust Integrated Geospatial Object Detection Framework for High Spatial Resolution Remote Sensing Imagery. *Remote Sens.* **2017**, *9*, 666. [[CrossRef](#)]
45. Li, W.; Fu, H.; Yu, L.; Cracknell, A. Deep Learning Based Oil Palm Tree Detection and Counting for High-Resolution Remote Sensing Images. *Remote Sens.* **2016**, *9*, 22. [[CrossRef](#)]
46. Hu, F.; Xia, G.-S.; Hu, J.; Zhang, L. Transferring Deep Convolutional Neural Networks for the Scene Classification of High-Resolution Remote Sensing Imagery. *Remote Sens.* **2015**, *7*, 14680–14707. [[CrossRef](#)]
47. Guirado, E.; Tabik, S.; Alcaraz-Segura, D.; Cabello, J.; Herrera, F. Deep-Learning Convolutional Neural Networks for scattered shrub detection with Google Earth Imagery. *arXiv* **2017**.
48. Chen, S.W.; Shivakumar, S.S.; Dcunha, S.; Das, J.; Okon, E.; Qu, C.; Taylor, C.J.; Kumar, V. Counting Apples and Oranges With Deep Learning: A Data-Driven Approach. *IEEE Robot. Autom. Lett.* **2017**, *2*, 781–788. [[CrossRef](#)]

49. Wang, Z.; Underwood, J.; Walsh, K.B. Machine vision assessment of mango orchard flowering. *Comput. Electron. Agric.* **2018**, *151*, 501–511. [[CrossRef](#)]
50. Sa, I.; Ge, Z.; Dayoub, F.; Upcroft, B.; Perez, T.; McCool, C. DeepFruits: A Fruit Detection System Using Deep Neural Networks. *Sensors* **2016**, *16*, 1222. [[CrossRef](#)] [[PubMed](#)]
51. *Citrus Fruits 2018 Summary*; USDA, National Agricultural Statistics Service: Washington, DC, USA, 2018.
52. White, H. Lindcove REC: Developing citrus varieties resistant to huanglongbing disease. *Calif. Agric.* **2017**, *71*, 18–20. [[CrossRef](#)]
53. senseFly—eMotion Ag. Available online: <https://www.sensefly.com/software/emotion-ag/> (accessed on 17 October 2018).
54. Pix4D. Available online: <https://pix4d.com/> (accessed on 7 August 2018).
55. eCognition | Trimble. Available online: <http://www.ecognition.com/> (accessed on 17 October 2018).
56. API Documentation | TensorFlow. Available online: https://www.tensorflow.org/api_docs/ (accessed on 17 October 2018).
57. Trimble. *eCognition 9.3 Reference Book*; Trimble Inc: Munich, Germany, 2018.
58. Csillik, O. Fast Segmentation and Classification of Very High Resolution Remote Sensing Data Using SLIC Superpixels. *Remote Sens.* **2017**, *9*, 243. [[CrossRef](#)]
59. Achanta, R.; Shaji, A.; Smith, K.; Lucchi, A.; Fua, P.; Süsstrunk, S. SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE Trans. Pattern Anal. Mach. Intell.* **2012**, *34*, 2274–2282. [[CrossRef](#)] [[PubMed](#)]
60. Saito, T.; Rehmsmeier, M. The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLoS ONE* **2015**, *10*, e0118432. [[CrossRef](#)] [[PubMed](#)]
61. Li, W.; Guo, Q.; Jakubowski, M.K. A new method for segmenting individual trees from the lidar point cloud. *Photogram. Eng. Remote Sens.* **2012**, *78*, 75–84. [[CrossRef](#)]
62. Goutte, C.; Gaussier, E. A Probabilistic Interpretation of Precision, Recall and F-Score, with Implication for Evaluation. In *Advances in Information Retrieval*; Springer: Berlin/Heidelberg, Germany, 2005; pp. 345–359.
63. Anderson, K. Gaston Lightweight unmanned aerial vehicles will revolutionize spatial ecology. *Front. Ecol. Environ.* **2013**, *11*, 138–146. [[CrossRef](#)]
64. Kislik, C.; Dronova, I.; Kelly, M. UAVs in Support of Algal Bloom Research: A Review of Current Applications and Future Opportunities. *Drones* **2018**, *2*, 35. [[CrossRef](#)]
65. Zhang, C.; Kovacs, J.M. The application of small unmanned aerial systems for precision agriculture: A review. *Precis. Agric.* **2012**, *13*, 693–712. [[CrossRef](#)]
66. Honkavaara, E.; Saari, H.; Kaivosoja, J.; Pölönen, I.; Hakala, T.; Litkey, P.; Mäkynen, J.; Pesonen, L. Processing and Assessment of Spectrometric, Stereoscopic Imagery Collected Using a Lightweight UAV Spectral Camera for Precision Agriculture. *Remote Sens.* **2013**, *5*, 5006–5039. [[CrossRef](#)]
67. Gonzalez-Dugo, V.; Zarco-Tejada, P.; Nicolás, E.; Nortes, P.A.; Alarcón, J.J.; Intrigliolo, D.S.; Fereres, E. Using high resolution UAV thermal imagery to assess the variability in the water status of five fruit tree species within a commercial orchard. *Precis. Agric.* **2013**, *14*, 660–678. [[CrossRef](#)]
68. 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](#)]
69. Hall, O.; Dahlin, S.; Marstorp, H.; Archila Bustos, M.; Öborn, I.; Jirstrom, M. Classification of Maize in Complex Smallholder Farming Systems Using UAV Imagery. *Drones* **2018**, *2*, 22. [[CrossRef](#)]
70. Wahab, I.; Hall, O.; Jirstrom, M. Remote Sensing of Yields: Application of UAV Imagery-Derived NDVI for Estimating Maize Vigor and Yields in Complex Farming Systems in Sub-Saharan Africa. *Drones* **2018**, *2*, 28. [[CrossRef](#)]
71. Maresma, Á.; Ariza, M.; Martínez, E.; Lloveras, J.; Martínez-Casasnovas, J. Analysis of Vegetation Indices to Determine Nitrogen Application and Yield Prediction in Maize (*Zea mays* L.) from a Standard UAV Service. *Remote Sens.* **2016**, *8*, 973. [[CrossRef](#)]
72. Hentschke, M.; Pignaton de Freitas, E.; Hennig, C.; Girardi da Veiga, I. Evaluation of Altitude Sensors for a Crop Spraying Drone. *Drones* **2018**, *2*, 25. [[CrossRef](#)]

73. Nevalainen, O.; Honkavaara, E.; Tuominen, S.; Viljanen, N.; Hakala, T.; Yu, X.; Hyypä, J.; Saari, H.; Pölonen, I.; Imai, N.N.; et al. Individual Tree Detection and Classification with UAV-Based Photogrammetric Point Clouds and Hyperspectral Imaging. *Remote Sens.* **2017**, *9*, 185. [[CrossRef](#)]
74. Mohan, M.; Silva, C.; Klauberg, C.; Jat, P.; Catts, G.; Cardil, A.; Hudak, A.; Dia, M. Individual Tree Detection from Unmanned Aerial Vehicle (UAV) Derived Canopy Height Model in an Open Canopy Mixed Conifer Forest. *For. Trees Livelihoods* **2017**, *8*, 340. [[CrossRef](#)]
75. Wallace, L.; Lucieer, A.; Watson, C.S. Evaluating Tree Detection and Segmentation Routines on Very High Resolution UAV LiDAR Data. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 7619–7628. [[CrossRef](#)]
76. Zhen, Z.; Quackenbush, L.; Zhang, L. Trends in Automatic Individual Tree Crown Detection and Delineation—Evolution of LiDAR Data. *Remote Sens.* **2016**, *8*, 333. [[CrossRef](#)]
77. Patrício, D.I.; Rieder, R. Computer vision and artificial intelligence in precision agriculture for grain crops: A systematic review. *Comput. Electron. Agric.* **2018**, *153*, 69–81. [[CrossRef](#)]
78. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. In *Advances in Neural Information Processing Systems 28*; Cortes, C., Lawrence, N.D., Lee, D.D., Sugiyama, M., Garnett, R., Eds.; Curran Associates, Inc.: Red Hook, NY, USA, 2015; pp. 91–99.
79. Tang, T.; Zhou, S.; Deng, Z.; Zou, H.; Lei, L. Vehicle Detection in Aerial Images Based on Region Convolutional Neural Networks and Hard Negative Example Mining. *Sensors* **2017**, *17*, 336. [[CrossRef](#)] [[PubMed](#)]
80. Wang, Y.; Hyypä, J.; Liang, X.; Kaartinen, H.; Yu, X.; Lindberg, E.; Holmgren, J.; Qin, Y.; Mallet, C.; Ferraz, A.; et al. International Benchmarking of the Individual Tree Detection Methods for Modeling 3-D Canopy Structure for Silviculture and Forest Ecology Using Airborne Laser Scanning. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 5011–5027. [[CrossRef](#)]
81. 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](#)]
82. Houborg, R.; McCabe, M.F. High-resolution NDVI from Planet’s constellation of earth observing nano-satellites: A new data source for precision agriculture. *Remote Sens.* **2016**, *8*, 768. [[CrossRef](#)]
83. Houborg, R.; McCabe, M.F. Daily Retrieval of NDVI and LAI at 3 m Resolution via the Fusion of CubeSat, Landsat, and MODIS Data. *Remote Sens.* **2018**, *10*, 890. [[CrossRef](#)]
84. Belgiu, M.; Csillik, O. Sentinel-2 cropland mapping using pixel-based and object-based time-weighted dynamic time warping analysis. *Remote Sens. Environ.* **2018**, *204*, 509–523. [[CrossRef](#)]
85. Wei, Z.; Han, Y.; Li, M.; Yang, K.; Yang, Y.; Luo, Y.; Ong, S.-H. A Small UAV Based Multi-Temporal Image Registration for Dynamic Agricultural Terrace Monitoring. *Remote Sens.* **2017**, *9*, 904. [[CrossRef](#)]
86. Fraser, B.T.; Congalton, R.G. Issues in Unmanned Aerial Systems (UAS) Data Collection of Complex Forest Environments. *Remote Sens.* **2018**, *10*, 908. [[CrossRef](#)]



© 2018 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/>).