Evaluation of CHM Morphological Processing and U-Net Deep Learning for Citrus Tree Canopy Delineation from UAV Imagery

Daiaeddine, M. J.,[1*] Badrouss, S.,[1] Jinnou, H.,[2] El Harti, A.,[1] Bachaoui, E. M.,[1] Biniz, M.[3] and Mouncif, H.[1]
[1]Geomatics, Georesources and Environment Laboratory, Faculty of Sciences and Techniques, Sultan Moulay Slimane University, Beni Mellal, Morocco, E-mail: mohamedjibril.daiaeddine@usms.ma*
ORCID ID: https://orcid.org/0009-0006-5525-7956*
[2]Landscape Dynamics, Risks and Heritage Laboratory, Faculty of Letters and Human Sciences, Sultan Moulay Slimane University of Beni Mellal, Morocco
[3]Information Processing and Decision Support Laboratory, Faculty of Sciences and Techniques, Sultan Moulay Slimane University, Beni Mellal, Morocco.
Corresponding Author

DOI: https://doi.org/10.52939/ijg.v22i7.5077

Abstract
Accurate delineation of individual tree crowns from UAV imagery is essential for precision citrus management but remains difficult in orchards with overlapping crowns and inter-row weeds. This study compared two contrasting approaches for citrus crown delineation: a canopy height model (CHM)-based morphological pipeline and a U-Net deep-learning model applied to RGB imagery. Both used an identical marker-controlled Watershed step for individual crown separation, so that the comparison reflected the segmentation stage alone. RGB imagery was acquired with a DJI Mavic 3M UAV over two structurally contrasting citrus orchards in the Beni Mellal-Khenifra region of Morocco, planted with the Maroc Late and Sidi Aissa varieties, and performance was assessed at the pixel and object levels. In the structurally simple Maroc Late orchard the two methods were effectively equivalent, both achieving high object-level agreement (detection score, the Jaccard index of the detected and reference crown sets, about 94%) and recovering crown area almost perfectly. The methods diverged sharply in the dense, weed-affected Sidi Aissa orchard: the U-Net sustained robust performance (detection score 93.6%; crown-area $R^2 = 0.93$), whereas the CHM-based pipeline, although it still detected roughly two-thirds of the reference crowns, was heavily penalised by false positives from weeds and merged crowns, lowering its detection score to 46.9% (crown-area $R^2 = 0.64$). Because the separation step was held constant, this divergence is attributable to the quality of the upstream mask rather than to the delineation algorithm: learned spectral and textural features proved decisive where the height threshold lost discriminative power. The results indicate that method choice should be matched to orchard structure, a lightweight height-threshold approach being sufficient in clean orchards and an RGB-trained model preferable under structural complexity. These findings, obtained from two citrus varieties at a single site and date, require multi-site validation before broader generalization.

Keywords: Canopy Height Model, Individual Tree Crown Delineation, Marker, Controlled Watershed, UAV imagery, U-Net

1. Introduction

The detection, delineation, and counting of tree canopies in orchards are integral to advancing precision agriculture, enhancing environmental sustainability, and improving orchard management practices. These methodologies provide a foundation for accurate yield estimation, as the size and health of tree canopies are directly correlated with fruit production [1]. Precise canopy mapping facilitates better forecasting of yield, enabling efficient planning for harvesting, storage, and market supply chains. Additionally, these techniques are instrumental in identifying gaps and dead trees within orchards, enabling targeted replanting and ensuring orchard uniformity over time. Beyond yield prediction, canopy detection plays a critical role in resource management by optimizing the application

ISSN: 1686-6576 (Printed) | ISSN 2673-0014 (Online) |

of water, fertilizers, and pesticides [2]. Reducing both costs and environmental impact through site-specific interventions. Canopy mapping further informs the spatial arrangement and density of trees, allowing for the refinement of orchard designs to enhance light penetration and air circulation, thereby fostering healthier trees and higher productivity [3]. Furthermore, the accurate assessment of tree populations has significant economic implications, including applications for public subsidies [4], valuation of plantations [5], and evaluation of damages from extraordinary events such as fires, pest infestations, or natural disasters.

Traditional methods for detecting, delineating, and counting tree canopies, often based on in-field human visual inspections, present significant challenges, particularly for large-scale agricultural operations. These approaches are not only labor-intensive and time-consuming but also prone to errors [6], limiting their reliability and scalability. This underscores the need for innovative techniques that enable accurate, efficient, and automated solutions for tree canopy mapping. Recent advancements in UAV imagery have emerged as a promising alternative to traditional methods across various agricultural applications. UAVs have been effectively employed for tasks such as yield estimation [7], weed detection [8], health assessment [9], and identification of crop tree crowns [10]. Their increasing popularity stems from their flexibility, cost-effectiveness, capability to operate under diverse weather conditions, and high operational efficiency, even in cloudy environments [11].

Numerous studies have investigated classical detection techniques for identifying individual trees using UAV imagery. For example, [12] developed a methodology utilizing UAV-acquired multispectral imagery to generate a Digital Surface Model (DSM) for tree detection, geolocation, and counting. Their approach incorporated mathematical morphology techniques, such as morphological reconstruction and H-maxima transforms, for segmenting tree crowns from the DSM. To refine segmentation, Otsu's thresholding was applied, addressing challenges like overlapping canopies. [13] proposed an automatic citrus tree extraction method using UAV-acquired RGB imagery and DSMs. Their workflow combined sequential thresholding, Canny edge detection, and Circular Hough Transform to effectively delineate tree crowns. Similarly, [14] employed an orientation-based radial symmetry transform for detecting potential tree locations, followed by the Chan–Vese active contour model to delineate tree crowns. Their methodology used UAV-based DSMs, avoiding the need for terrain normalization via Digital Terrain Models (DTMs) or CHMs. [15] introduced a methodology combining Connected Components Labeling (CCL) and morphological image operations for processing UAV-acquired multispectral imagery, applying techniques like grayscale conversion, histogram equalization, and morphological filtering to count individual citrus trees. [16] proposed a segmentation method using UAV-acquired RGB imagery, combining selective region intensity histogram equalization and region of interest (ROI) extraction via chromatic aberration and Otsu thresholding, with a support vector machine (SVM) trained on color, texture, and local binary pattern (LBP) features. Across this family, accuracy depends on a representation derived either from scene geometry (DSM/CHM height) or from hand-crafted radiometric and textural descriptors, and on thresholds or morphological operators configured for a particular orchard.

Despite their effectiveness on the scenes for which they were tuned, these classical methods rely heavily on manually configured parameters tailored to specific targets or image characteristics, requiring expert knowledge and limiting their adaptability across diverse scenarios [17]. Their performance is most fragile precisely where orchards are most demanding: under canopy overlap, where geometric or morphological operators merge touching crowns, and under inter-row vegetation, where height- or radiometry-based thresholds cannot separate understory growth from low canopy. Deep learning algorithms, rooted in artificial intelligence, offer a promising solution to overcome these limitations in remote sensing applications [18]. These techniques have shown remarkable progress across domains, garnering significant attention from both academia and industry [19]. Unlike conventional approaches, deep learning models automatically extract hierarchical features, eliminating the need for manual parameter selection [20]. Convolutional Neural Networks (CNNs), in particular, are widely used in vegetation remote sensing studies [21].

Studies integrating UAV technology with deep learning have demonstrated significant advances in tree-crown detection. [22]compared local maxima (LM), marker-controlled watershed segmentation (MCWS), and Mask R-CNN for individual-tree detection in young plantation forests; LM and MCWS operated on CHMs whereas Mask R-CNN used multi-band imagery and a ResNet-50 backbone, achieving superior accuracy. [23]combined Faster R-CNN and U-Net architectures to detect and segment apple trees and to extract morphological parameters such as crown width and area. [10] used Mask R-CNN with fused RGB and CHM data to detect and delineate orange-tree crowns in dense plantations,

ISSN: 1686-6576 (Printed) | ISSN 2673-0014 (Online) | © Geoinformatics International

addressing overlapping canopies. [24] applied the U²-Net model to UAV RGB imagery for olive-crown segmentation, outperforming HRNet, U-Net, and DeepLabv3+, and [25]integrated multispectral imagery and vegetation indices (NDVI, RVI, SAVI) within a YOLO-Fi model for canopy detection in complex orchards. These works establish that learned representations can sustain accuracy under structural complexity, but they share three recurring constraints: they require annotated training data and non-trivial computation; they are typically validated on a single orchard or structural condition, without spatially independent or cross-site testing, so generalization is asserted rather than demonstrated; and where paradigms are compared, the comparison is between end-to-end pipelines that differ at several stages at once input representation, network, and post-processing so an observed performance difference cannot be attributed to any single component. Taken together, the literature presents two paradigms with complementary and largely unexamined trade-offs (Table 1).

Table 1: Methods, data, strengths, and limitations of UAV-based individual tree-crown detection and delineation approaches

Ref.	Method / paradigm	Data	Target	Key strength	Principal limitation(s)
[10]	Mask R-CNN, RGB and RGB + CHM fusion (CHM as 4th band) (deep learning)	RGB, CHM	Citrus (orange)	High accuracy in dense plantations; CHM fusion improves results (91.7% to 97.0%; $R^2 \approx 0.97$)	Requires both CHM and training; heavy instance segmentation
[12]	Morphological reconstruction + H-maxima + Otsu, with connected-component seeding (classical)	Multispectral, DSM	Generic crop trees (orange grove)	Training-free; geolocation, counting, and CC-level separation of touching crowns	Threshold/operator tuning; degrades under heavy crown overlap
[13]	Sequential thresholding + Canny + Circular Hough Transform (classical)	Multispectral, DSM	Citrus	Explicit geometric (circular) crown priors; delineation accuracy > 80%	Assumes near-circular crowns; parameter-sensitive
[14]	Orientation-based radial symmetry transform + Chan–Vese active contour (classical)	DSM	Citrus	No height threshold / DTM–CHM normalization required; overall F1 ≈ 91.2%	Contour initialization and parameter sensitivity
[15]	Connected Components Labeling (CCL) + morphological filtering (classical)	Multispectral (5-band)	Citrus	Simple, fast counting; accuracy and precision > 95% (recall 0.97)	Merges touching/overlapping crowns
[16]	Histogram equalization + chromatic-aberration/ Otsu ROI + SVM on color/texture/LBP (shallow learning)	RGB	Citrus	Learned shallow features; robust to brightness and weed variability (≈ 85% accuracy)	Hand-crafted features; threshold-dependent
[22]	LM / MCWS / Mask R-CNN (paradigm comparison)	CHM (LM, MCWS); multi-band imagery (Mask R-CNN)	Young plantation forest (Chinese fir)	Direct paradigm comparison; Mask R-CNN multi-band best (F1 ≈ 94.7%)	Whole-pipeline comparison confounds stages; CHM methods underperform; forest, not orchard
[23]	Faster R-CNN + U-Net (deep learning)	UAV RGB imagery	Apple	Detection + segmentation; extracts morphological parameters (crown width, area)	Two-stage; annotation and compute cost
[24]	U²-Net (deep learning)	RGB	Olive	Outperforms HRNet, U-Net, DeepLabv3+; crown-area $R^2 > 0.93$	Single structural condition; no cross-site test
[25]	YOLO-Fi (detect + localize + segment) + fused vegetation indices (RVI/NDVI/SAVI via mRMR) (deep learning)	Multispectral	Apple	Robust detection/ segmentation in complex orchards ($mAP_{50-95} = 0.72$, mIoU = 0.75); enables variable spraying	Requires multispectral sensor and index engineering

ISSN: 1686-6576 (Printed) | ISSN 2673-0014 (Online) | © Geoinformatics International

Classical height- and morphology-based methods are training-free and computationally light but parameter-bound and brittle under overlap and understory vegetation; learned methods are robust to such complexity but carry annotation and computational costs and are rarely tested for generalization. Two questions therefore remain open. First, when a height-based pipeline and a learned RGB pipeline differ in performance, why do they differ because of the delineation algorithm, or because of the information the upstream mask encodes? Existing comparisons cannot answer this, because they vary the delineation step together with the representation. Second, under which orchard conditions should each paradigm be preferred? Because most studies report a single structural setting, the literature offers little condition-dependent guidance to a practitioner choosing a method for a given orchard. This study is motivated by these two gaps: it isolates the contribution of the upstream representation by holding the delineation step fixed, and it stratifies the same comparison across two orchards of deliberately contrasting structural complexity to derive a condition-dependent recommendation rather than a single performance ranking.

This study accordingly compares two contrasting, paradigmatically opposed strategies for segmenting and delineating individual citrus tree canopies from UAV imagery: (1) a CHM-based morphological pipeline that extracts crowns from canopy-height variation, and (2) a U-Net deep-learning model that exploits spectral, textural, and contextual features in RGB imagery. These two were selected as baselines because they represent the two dominant and information-theoretically distinct representations used in operational orchard mapping scene geometry versus learned appearance and because rendering them comparable through a shared, parameter-identical Watershed delineation step converts what is usually a confounded end-to-end comparison into a controlled one. The novelty of this work lies in its controlled experimental design and the analysis it enables: by fixing the delineation stage and stratifying by structural complexity across the well-spaced Maroc Late orchard and the dense, weed-affected Sidi Aissa orchard, the study (i) attributes the performance divergence to the upstream representation rather than to the separation algorithm, (ii) provides a mechanistic explanation of why the height representation loses discriminative power as weed cover and crown overlap increase, and (iii) translates these findings into a defensible, condition-dependent decision rule for method selection. In doing so it contributes practical guidance to precision agriculture while delineating the scope and the failure modes within which each approach is appropriate.

Individual-crown separation in both pipelines is performed with a distance-transform-based, marker-controlled Watershed [26]. This variant was chosen deliberately. An unseeded Watershed applied to an image gradient over-segments severely, which is precisely the problem that marker (seed) control was introduced to solve [27]; region-growing and graph-cut formulations, by contrast, require seed selection, parameter or energy-function tuning, or substantial computation [28] and [29], and learning-based instance-segmentation heads such as Mask R-CNN [30] would have to be trained separately for each pipeline reintroducing exactly the multi-stage confound this study sets out to remove. Marker-controlled Watershed seeded by local maxima of the Euclidean distance transform is, by contrast, unsupervised, governed by a single interpretable parameter (the minimum inter-marker distance), and well suited to separating touching, convex, compact objects a property exploited both in cell-cluster segmentation [31] and in individual tree-crown delineation from canopy height models [32] and it applies in identical form to the binary masks produced by both pipelines. Its principal limitation is over-segmentation of large or irregular crowns bearing several distance-transform maxima [33] and [34]; we constrain this with an empirically selected minimum-distance criterion, analogous to the crown-size constraints used in CHM-based detection [32], applied identically to both methods and both orchards, and we report and quantify the residual cases, which motivate the end-to-end instance-segmentation directions discussed later.

2. Materials and Methods

2.1 Study Area

The study area, illustrated in Figure 1(a), consists of two citrus orchards situated in the Béni Mellal-Khénifra region of Morocco. The first orchard, located at 32°19′28″N and 6°29′36″W, covers 6.11 ha and is dedicated to the Maroc Late citrus variety. It was established in 2006/2007 and contains 2409 trees, corresponding to a density of approximately 394 trees/ha. The second orchard, located at 32°19′25″N and 6°29′50″W, spans approximately 12.3 ha and is cultivated with the Sidi Aissa citrus variety. It was established in 2005/2006 and contains 4,782 trees, corresponding to a density of approximately 389 trees/ha. Both orchards follow an approximately regular planting pattern, with an average spacing close to 5 m between trees.

ISSN: 1686-6576 (Printed) | ISSN 2673-0014 (Online) | © Geoinformatics International

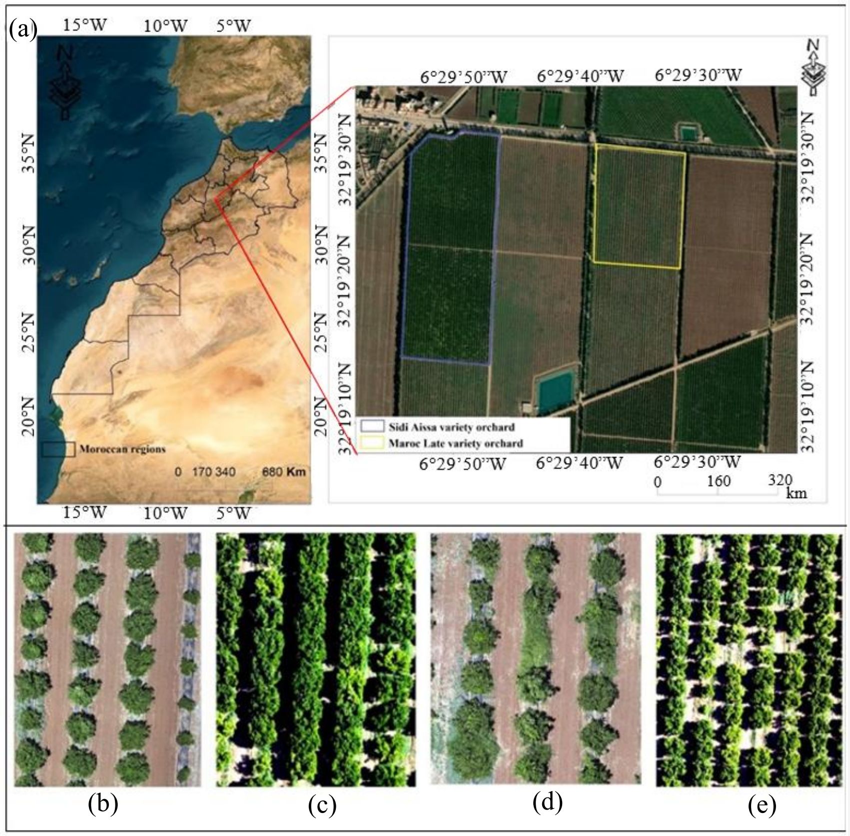

Figure 1: Geographic location and structural variability of the study area:
(a) Location of the two citrus orchards in the Béni Mellal-Khénifra region of Morocco, with orchard boundaries for the Maroc Late and Sidi Aissa varieties, (b) Variation in tree age caused by replanting, (c) Crown overlap between adjacent trees, (d) Weed growth between trees, (e) Missing trees caused by tree removal, disrupting orchard uniformity

However, tree age is heterogeneous because some trees have been removed and replaced over time. Therefore, the establishment year does not strictly reflect the age of every individual tree. Individual canopy diameters range from approximately 0.6 to 4.4 m in the Maroc Late orchard and from 0.8 to 5.6 m in the Sidi Aissa orchard. The larger maximum canopy diameter observed in the Sidi Aissa orchard suggests more developed crowns and a greater likelihood of crown overlap.

The orchards exhibit marked structural variability caused by differences in tree age due to replanting (Figure 1(b)), variation in canopy size and crown development resulting in overlapping crowns (Figure 1(c)), and weed growth between trees, which may introduce spectral confusion during image segmentation (Figure 1(d)). In addition, missing trees resulting from tree removal disrupt the regular spatial arrangement of the orchards (Figure 1(e)). These conditions make individual tree crown delineation challenging by increasing canopy-size variability, reducing the separability of adjacent crowns, introducing vegetation-related spectral confusion, and disrupting the regular planting pattern. Robust segmentation methods are therefore required to accurately extract individual citrus tree crowns.

2.2 UAV Data Acquisition

Aerial RGB imagery was acquired on 8 July 2024 using a DJI Mavic 3M UAV equipped with a 20 MP RGB camera, a 4/3 CMOS sensor, an image resolution of 5280 × 3956 pixels, and a 24 mm equivalent focal length. Automated flight missions were planned and executed for both orchards using DJI Pilot 2, following a parallel grid flight pattern designed to ensure complete and homogeneous image coverage of the study areas (Figure 2). Image acquisition was conducted under clear-sky and low-wind conditions to reduce atmospheric disturbance, illumination variability, and motion-induced image degradation.

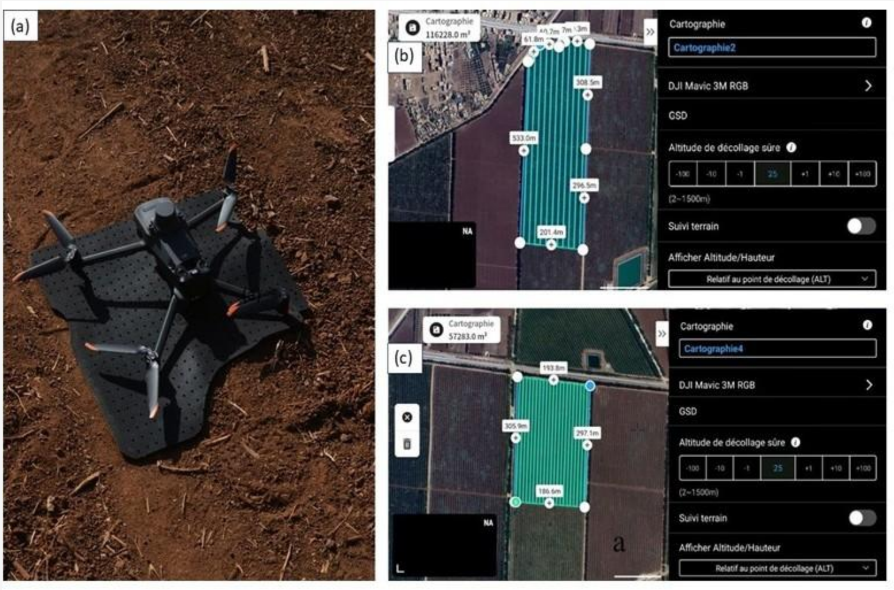

Figure 2: UAV platform and flight planning for RGB data acquisition:
(a) DJI Mavic 3M UAV used for image acquisition, (b) DJI Pilot 2 flight plan for the Sidi Aissa orchard,
(c) DJI Pilot 2 flight plan for the Maroc Late orchard

The UAV was flown at a nominal altitude of 25 m relative to the take-off point, with the gimbal pitch set to −90° for nadir image acquisition. The flight speed was fixed at 5 $m \cdot s^{-1}$. Forward and side overlaps were set to 85% and 75%, respectively, to provide sufficient image redundancy for tie-point extraction, Structure-from-Motion/Multi-View Stereo photogrammetric reconstruction, and subsequent generation of high-resolution surface models.

The flight altitude of 25 m was selected as a compromise between spatial resolution, photogrammetric reconstruction quality, operational efficiency, and safe clearance above the orchard canopy. At this altitude, the theoretical Ground Sampling Distance (GSD) of the RGB imagery was approximately 0.7 $cm \cdot pixel^{-1}$, corresponding to 0.007 $m \cdot pixel^{-1}$. This spatial resolution enabled detailed representation of canopy structure, inter-crown gaps, foliage texture, and transition zones between adjacent trees. Considering that canopy diameters ranged from approximately 0.6 to 5.6 m across the two orchards, the acquired imagery provided approximately 85 to 800 pixels across individual crown diameters, depending on tree size. This level of spatial detail is suitable for both CHM-based morphological image analysis and deep learning-based citrus canopy segmentation. The final image datasets consisted of 361 RGB images for the Maroc Late orchard and 742 RGB images for the Sidi Aissa orchard. These images were subsequently used for photogrammetric processing, orthomosaic generation, canopy height model production, canopy segmentation, and individual tree crown delineation.

2.3 Photogrammetric Processing: Orthomosaic and CHM Generation

UAV RGB images acquired over the Maroc Late and Sidi Aissa citrus orchards were processed using Agisoft Metashape Professional Edition v2.1.1 (Agisoft LLC, St. Petersburg, Russia) following a Structure-from-Motion/Multi-View Stereo (SfM-MVS) photogrammetric workflow. All photogrammetric products were generated in the WGS 84 / UTM Zone 29N coordinate reference system (EPSG:32629). Direct georeferencing was supported by the onboard RTK positioning system of the DJI Mavic 3M UAV in combination with a DJI D-RTK 2 mobile base station installed on a tripod within each orchard during image acquisition (Figure 3).

Figure 3: DJI D-RTK 2 base station used to support RTK-based direct georeferencing during UAV image acquisition

ISSN: 1686-6576 (Printed) | ISSN 2673-0014 (Online) | © Geoinformatics International

This configuration provides nominal camera positioning accuracies of 1.5 cm + 1 ppm horizontally and 3 cm + 1 ppm vertically. These values refer to the nominal RTK positioning accuracy of the acquisition system and were used to support the spatial consistency of the photogrammetric reconstruction. Image alignment was performed using the high-accuracy setting in Agisoft Metashape, with key point and tie point limits set to 40,000 and 4,000, respectively. All 361 images from the Maroc Late orchard and all 742 images from the Sidi Aissa orchard were successfully aligned. The alignment produced RMS reprojection errors of 0.43 px for the Maroc Late orchard and 0.48 px for the Sidi Aissa orchard, indicating stable internal camera alignment and reliable image block geometry. Dense point clouds were then generated using high-quality Multi-View Stereo reconstruction with mild depth filtering. Digital Surface Models (DSMs) were interpolated from the full dense point clouds at a spatial resolution of 0.007 m·pixel^{-1}, corresponding to approximately 0.7 cm·pixel^{-1}. Digital Terrain Models (DTMs) were derived using the built-in ground point classification tool in Agisoft Metashape, with the following parameters: maximum angle of 15°, maximum distance of 0.20 m, and cell size of 1.00 m. The DTM was subsequently interpolated from the classified ground points.

In the Sidi Aissa orchard, classified ground points were manually reviewed before DTM interpolation. Points corresponding to weed-covered areas that were incorrectly classified as ground were reassigned to the non-ground class. This correction was performed to reduce artificial elevation of the terrain surface in areas affected by weed growth between trees, which could otherwise lead to underestimation of canopy height after DSM–DTM subtraction. Orthomosaics were generated by projecting the calibrated images onto the DSM using mosaic blending. The final orthomosaics were exported at a ground sampling distance of approximately 0.7 cm·pixel^{-1}, consistent with the theoretical RGB image resolution obtained at the nominal flight altitude of 25 m. The Canopy Height Model (CHM) was computed by subtracting the DTM from the DSM as defined in Equation 1:

$$CHM(x,y) = DSM(x,y) - DTM(x,y)$$

Equation 1

Where:

CHM(x,y) represents the canopy height at pixel location (x,y)

DSM(*x,y*) represents the surface elevation including tree crowns and ground objects

DTM(*x,y*) represents the estimated bare-ground elevation.

Negative CHM values were clipped to zero to remove non-physical height values, and a 3 × 3 median filter was applied to reduce isolated elevation outliers while preserving local canopy structure. A preliminary CHM quality assessment was conducted using representative non-tree areas. CHM pixel values were extracted from 20 bare-soil polygons in the Maroc Late orchard and 30 polygons in the Sidi Aissa orchard, including 15 bare-soil polygons and 15 weed-covered polygons between trees. Zonal statistics were calculated using ArcGIS 10.7.1. Bare-soil areas produced mean CHM values of 0.03 ± 0.04 m in the Maroc Late orchard and 0.04 ± 0.05 m in the Sidi Aissa orchard, suggesting limited residual DSM–DTM offsets in non-vegetated areas. In contrast, weed-covered areas between trees in the Sidi Aissa orchard showed higher CHM values, with a mean of 0.29 ± 0.15 m. This result indicates that elevated weed patches may contribute to false positive detections in CHM-based canopy segmentation, particularly in areas where weed height is spectrally or structurally similar to low citrus vegetation. The resulting orthomosaics and CHMs for both orchards are shown in Figure 4. These products provided the spatial and structural information required for subsequent canopy extraction, tree crown segmentation, and comparative analysis of morphological and deep learning-based methods. The complete set of Agisoft Metashape processing parameters used for photogrammetric reconstruction is summarized in Table 2.

2.4 Creation of the Ground Truth Dataset

Following the generation of the orthomosaic and Canopy Height Model (CHM) for each orchard, individual citrus tree canopies were manually delineated in ArcGIS 10.7.1 to construct an expert-annotated reference dataset (Figure 5). All annotations were performed by a single trained operator with experience in UAV image interpretation and citrus canopy structure to maintain annotation consistency. However, the potential subjectivity associated with single-operator delineation, particularly in areas with dense canopy overlap, is acknowledged as a limitation of the present study. The delineation procedure was based on the joint visual interpretation of the RGB orthomosaic and the CHM.

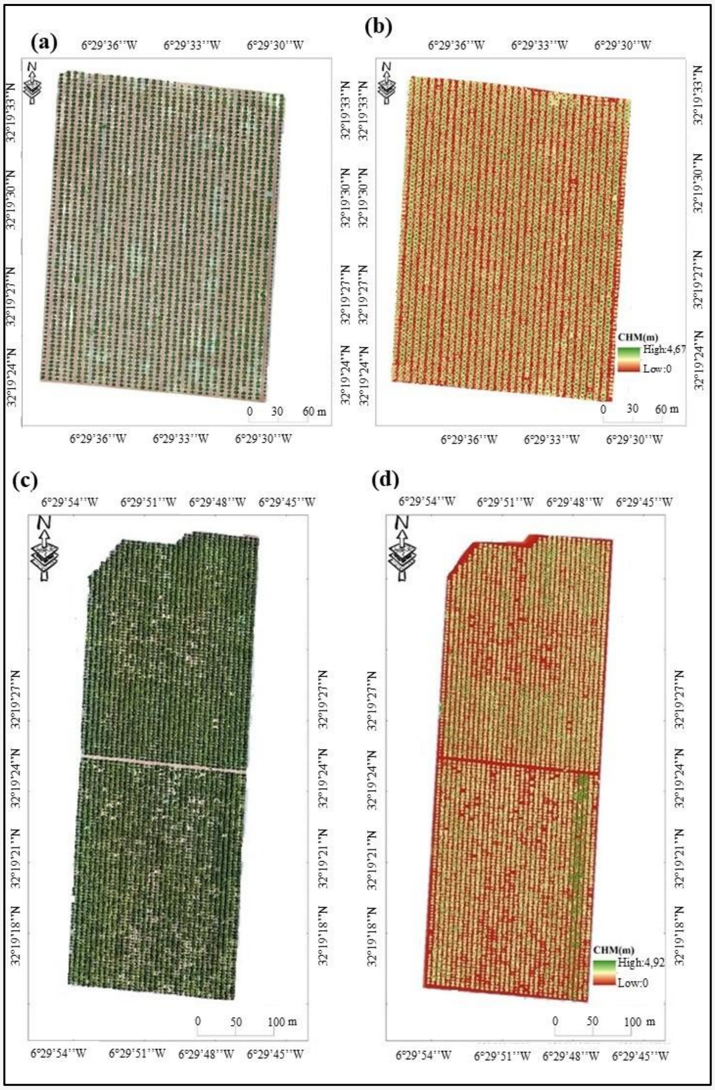

Figure 4: Orthomosaics and Canopy Height Models of the study orchards:
(a) Orthomosaic of the Maroc Late orchard, (b) CHM of the Maroc Late orchard,
(c) Orthomosaic of the Sidi Aissa orchard, (d) CHM of the Sidi Aissa orchard

Table 2: Agisoft Metashape processing parameters used for photogrammetric reconstruction

Stage	Settings
Software	Agisoft Metashape Professional v2.1.1
Coordinate reference system	WGS 84 / UTM Zone 29N (EPSG:32629)
Geometric control	Onboard RTK + DJI D-RTK 2 base station
Nominal RTK accuracy	1.5 cm + 1 ppm horizontal; 3 cm + 1 ppm vertical
Image alignment	High accuracy
Key/tie point limits	40,000 / 4,000
RMS reprojection error	0.43 px, Maroc Late; 0.48 px, Sidi Aissa
Dense point cloud	High quality; mild depth filtering
DSM	Interpolated from dense cloud; resolution 0.007 $m \cdot pixel^{-1}$, approximately 0.7 $cm \cdot pixel^{-1}$
DTM	Ground classification; max angle 15°, max distance 0.20 m, cell size 1.00 m
Manual correction	Weed-covered points between trees reassigned to non-ground class in Sidi Aissa
Orthomosaic	Mosaic blending on DSM; GSD approximately 0.7 $cm \cdot pixel^{-1}$
CHM	DSM − DTM; negative values clipped to zero; 3 × 3 median filter

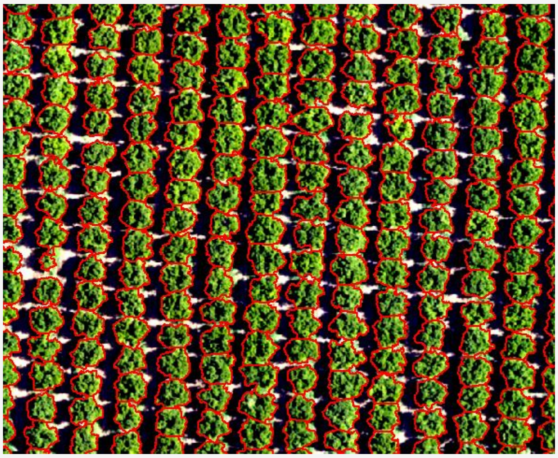

Figure 5: Manual delineation of individual citrus tree canopies in ArcGIS 10.7.1. Red polygons represent expert-annotated individual tree crown boundaries overlaid on the RGB orthomosaic

The RGB orthomosaic was used to identify visible canopy attributes, including crown shape, foliage texture, shadow patterns, canopy edges, visible gaps, and concave boundaries between adjacent crowns. The CHM provided complementary structural information related to canopy height and crown morphology. Adjacent trees were separated by identifying distinct crown units based on visible crown morphology and local height variations. Boundaries between neighboring crowns were placed along lower-height transition zones between adjacent canopy peaks, hereafter referred to as saddle zones, when these zones were also supported by visible gaps, concave crown edges, or texture discontinuities in the RGB orthomosaic. These saddle zones were interpreted visually rather than identified using an automatic numerical threshold, allowing adaptation to heterogeneous canopy shapes while introducing a degree of operator-dependent interpretation.

Ambiguous cases were resolved according to predefined annotation rules. Missing trees and dead trees without recognizable canopy structure were excluded from the reference dataset. Very small or partially occluded trees were delineated only when their canopy boundaries could be clearly distinguished from the background or from neighboring crowns. For trees located along the orchard boundary, only the canopy portion within the study area was delineated. After delineation, the polygon layer was visually inspected to correct topology errors, remove duplicate features, and ensure consistency between crown boundaries and the orchard planting pattern. The final reference dataset contained 2,409 individual tree canopy polygons for the Maroc Late orchard and 4,782 individual tree canopy polygons for the Sidi Aissa orchard. These polygons represent visible and recognizable citrus tree canopies within the study area. Although the combined use of RGB and CHM information provided complementary visual support for crown delineation, the absence of independent validation by a second annotator or inter-annotator agreement assessment remains a limitation of the reference dataset.

2.5 CHM-Based Canopy Extraction Using Thresholding and Morphological Refinement

A CHM-based image processing approach was applied to extract citrus canopy regions from the surrounding background. Unlike deep learning-based methods, this approach does not require labeled training data; however, it requires the selection of a height threshold to separate elevated canopy structures from low-height background elements. The processing workflow consisted of four steps: CHM preparation, threshold-based binarization, hole filling, and binary mask normalization. Let $C(x,y)$ denote the CHM value, in meters, at pixel location (x,y). NoData pixels were excluded from the analysis. No additional smoothing or filtering was applied beyond the CHM post-processing described in Section 2.3; the processed CHM was used directly for canopy extraction. The initial canopy mask was obtained by applying a fixed height threshold T. Pixels with CHM values greater than or equal to T were classified as canopy, while pixels below T were classified as background as defined in Equation 2:

$$I(x,y)=\begin{cases}1, & C(x,y)\geq T\\ 0, & C(x,y)<T\end{cases}$$

Equation 2

Where:

I(x,y) is the initial binary canopy mask,
C(x,y) is the *CHM* value at pixel location (x, y),
T is the height threshold.

The threshold was evaluated separately for each orchard using the CHM height distribution and visual inspection of representative areas containing bare soil, weeds, young trees, mature trees, and overlapping crowns. The same threshold value was selected for both orchards: 0,5m.

This value provided a practical compromise between retaining citrus canopy pixels and excluding low-height background elements such as bare soil and low vegetation. The selected threshold is consistent with the CHM quality assessment reported in Section 2.3, where bare-soil areas showed near-zero CHM values of 0.03-0.04 m. In the Sidi Aissa orchard, weed-covered areas had higher CHM

ISSN: 1686-6576 (Printed) | ISSN 2673-0014 (Online) |

values, with a mean of 0.29 ± 0.15 m. Therefore, some taller weed patches may have exceeded the 0.5 m threshold and been retained as canopy pixels, potentially contributing to false positive detections in the CHM-based method. After thresholding, the binary mask sometimes contained internal gaps within tree crowns due to local canopy height variation, occlusion, or residual CHM noise. A hole-filling operation was therefore applied using the binary_fill_holes() function. This operation converts background pixels fully enclosed by canopy pixels into canopy pixels [35], producing more continuous canopy regions as defined in Equation 3:

$$H = I \cup F(I)$$

Equation 3

Where:

H is the binary mask after hole filling,

I is the initial binary canopy mask, and

F(I) is the enclosed background regions within detected canopy objects.

Although hole filling improves mask continuity, it may also fill small enclosed gaps within merged canopy clusters, particularly in areas with overlapping crowns. Finally, the output mask was explicitly converted to binary values as defined in Equation 4:

$$B(x,y) = \begin{cases} 1, & H(x,y) > 0 \\ 0, & H(x,y) = 0 \end{cases}$$

Equation 4

Where:

B(x,y) is the final binary canopy mask.

This mask was used as the input for the subsequent individual crown separation step and for comparison with the deep learning-based segmentation results.

2.6 Deep Learning Approach: U-Net Model for Tree Canopy Segmentation Using RGB Imagery

2.6.1 Data preprocessing

Preprocessing and data augmentation procedures were applied to the RGB orthomosaics to prepare the dataset for U-Net-based citrus canopy segmentation. The manually delineated canopy polygons were first rasterized to generate binary reference masks with the same spatial resolution, extent, and coordinate alignment as the RGB orthomosaics. Canopy pixels were assigned a value of 1, whereas background pixels were assigned a value of 0. Each RGB orthomosaic and its corresponding binary reference mask were cropped into 512 × 512 pixel patches using a 64-pixel overlap between adjacent patches, corresponding to a stride of 448 pixels in both horizontal and vertical directions. At the orthomosaic GSD of approximately 0.7 cm·pixel^{-1}, each 512 × 512 pixel patch covered approximately 3.58 m × 3.58 m on the ground, while the 64-pixel overlap corresponded to approximately 0.45 m. The stride corresponded to approximately 3.14 m. This limited overlap was used to reduce boundary effects for crowns located near patch edges while avoiding excessive redundancy among neighboring patches. However, because some mature or overlapping crowns exceeded the spatial extent of a single patch, large crowns could be partially represented across adjacent patches.

The extracted patches were divided into training, validation, and testing subsets using approximate proportions of 70%, 20%, and 10%, respectively. The final number of patches in each subset is presented in Table 3. Since the partitioning was performed at the patch level rather than through spatially independent zones, spatial autocorrelation may remain between neighboring patches. This limitation should be considered when interpreting the generalization performance of the U-Net model. Data augmentation was applied only to the training subset to reduce overfitting and promote model robustness to spatial and radiometric variability. The same geometric transformations were applied simultaneously to each RGB patch and its corresponding mask to preserve pixel-level alignment. Random horizontal and vertical flips were each applied with a probability of 0.5. Random rotations were applied within a range of ±30°. During geometric transformations, bilinear interpolation was used for RGB images, whereas nearest-neighbor interpolation was used for masks to preserve binary class labels. Color augmentation was applied only to the RGB images and not to the corresponding masks. Specifically, ColorJitter was used with brightness = 0.1 and contrast = 0.1, corresponding to random adjustment factors sampled from the interval [0.9, 1.1] for both brightness and contrast.

Table 3: Distribution of image patches for training, validation, and testing across the two orchards

Orchard	Total Sub-Images	Training (70%)	Validation (20%)	Test (10%)
Maroc Late	1,632	1,143	326	163
Sidi Aissa	4,326	3,029	865	432

ISSN: 1686-6576 (Printed) | ISSN 2673-0014 (Online) | © Geoinformatics International

Before model input, RGB pixel values were rescaled to the [0, 1] range and normalized using the mean and standard deviation calculated from the training subset. The U-Net model was trained to perform binary semantic segmentation of citrus canopy pixels from RGB imagery. The resulting predicted canopy masks were subsequently used for comparison with the CHM-based canopy extraction approach and for later individual crown separation.

2.6.2 U-Net model architecture

The segmentation network used in this study was based on the U-Net architecture introduced by [36] for biomedical image segmentation and subsequently widely adopted for semantic segmentation in remote sensing. The model was designed for binary semantic segmentation of citrus canopy pixels from RGB UAV imagery. The overall architecture, together with the internal composition of each convolutional block, is illustrated in Figure 6.

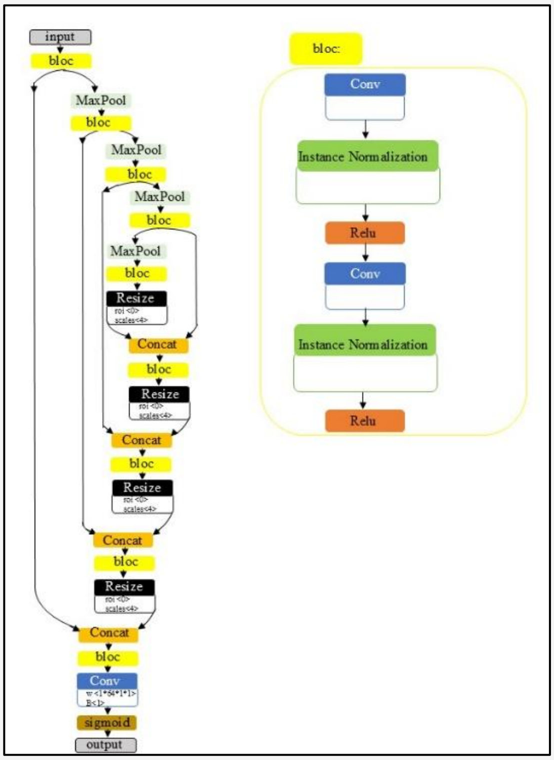

Figure 6: Overall architecture of the U-Net model and detailed structure of each block

The encoder consisted of four successive convolutional blocks with feature-map depths of 64, 128, 256, and 512 channels. Each block included two consecutive 3 × 3 convolutional layers, each followed by Instance Normalization [37] and a ReLU activation[38]. All 3 × 3 convolutions used padding to preserve feature-map dimensions within each block. A 2 × 2 max-pooling operation was applied after each encoder block, reducing the spatial resolution by a factor of two while increasing the level of semantic abstraction. The bottleneck was implemented as a convolutional block with 1024 feature channels.

The decoder mirrored the encoder using four convolutional blocks with 512, 256, 128, and 64 channels. Spatial resolution was progressively restored using nearest-neighbor upsampling with a scale factor of two. At each decoder stage, the upsampled feature maps were concatenated with the corresponding encoder feature maps through skip connections, allowing the decoder to recover fine spatial details lost during downsampling. The final layer consisted of a 1 × 1 convolution that reduced the output to a single channel. During inference, a Sigmoid activation was applied to generate a per-pixel canopy probability map. Instance Normalization was selected to improve robustness to patch-level radiometric variability in the UAV orthomosaics. Because the training samples were extracted from large RGB orthomosaics, individual patches could differ in local illumination, shadow intensity, canopy brightness, and background appearance. By normalizing each sample independently, Instance Normalization reduces dependence on batch-level statistics and helps limit the influence of local radiometric variations between patches. This choice was therefore intended to support stable feature learning across heterogeneous orchard conditions rather than being motivated solely by batch-size limitations. The overall encoder–decoder structure and skip-connection design were inspired by the original U-Net, while Instance Normalization and nearest-neighbor upsampling were introduced as implementation-specific adaptations. No dropout layers were used; regularization was provided through data augmentation and validation-based checkpoint selection.

2.6.3 Experimental details

All deep learning experiments were carried out on the Google Colab Pro+ platform using an NVIDIA A100 GPU with 40 GB of dedicated memory and 54 GB of system RAM. Model implementation, training, and evaluation were performed in a PyTorch-based environment, with rasterio and scikit-image used for geospatial data handling and image processing. Two U-Net models were trained independently, one for the Maroc Late orchard and one for the Sidi Aissa orchard, using only RGB orthomosaic patches as input. Therefore, the reported performance reflects within-orchard segmentation rather than cross-

ISSN: 1686-6576 (Printed) | ISSN 2673-0014 (Online) | © Geoinformatics International

orchard generalization. Optimization was performed using the Adam algorithm with its standard parameters: $\beta_1 = 0.9$, $\beta_2 = 0.999$, and $\varepsilon = 1 \times 10^{-8}$. The task was formulated as binary canopy/background segmentation and optimized using binary cross-entropy loss. Convolutional layers were initialized using the Kaiming uniform initialization scheme, which is suitable for networks using ReLU activations. A fixed learning rate of 1×10^{-4} was used without a learning-rate scheduling strategy. This value was selected empirically from preliminary experiments and was supported by stable training and validation curves. The batch size was set to 16 to balance GPU memory usage, training stability, and computational efficiency when processing 512 × 512 pixel patches. Each model was trained for 100 epochs. Although formal early stopping was not implemented, model selection was performed using a best-checkpoint strategy, in which the weights corresponding to the highest validation IoU were retained for final testing. During inference, predicted probability maps were converted into binary canopy masks using a threshold of 0.5.

The reproducibility of the patch-level train/validation/test split was supported by fixing the random_state parameter of the scikit-learn splitting function to 42. However, full training reproducibility was not guaranteed because torch.manual_seed, numpy.random.seed, and CUDA-level random operations were not explicitly fixed. These controls should be incorporated in future implementations to improve reproducibility of the complete training pipeline.

2.7 Delineation of Individual Tree Crowns

The binary masks produced by both the U-Net model and the CHM-based morphological processing were post-processed to delineate individual tree crowns using a distance transform-based marker-controlled Watershed segmentation. This approach is well suited for separating touching, convex-shaped objects such as citrus crowns and, unlike region-growing or graph-cut methods, operates automatically without manual seed selection, energy functions, or additional training data. First, the Euclidean Distance Transform [39] was computed from the binary mask using the *distance_transform_edt* function of *scipy.ndimage*, assigning each foreground pixel its distance to the nearest background pixel as defined in Equation 5:

$$D(p) = \min_{q \in B} \|p - q\|$$

Equation 5

where:

D(p) is Euclidean distance transform value at foreground pixel p

p is the foreground (canopy) pixel

q is the nearest background pixel to p

B is the set of background pixels

So that the highest values lie near crown centers. Local maxima [40], interpreted as crown centers, were then detected with the peak_local_max function of scikit-image as defined in Equation 6:

$$M = \{p \in I \mid D(p) \geq D(q), \forall q \in N(p)\}$$

Equation 6

Where:

M is the set of detected local maxima (crown center candidates)

D(p) is the distance transform value at pixel p

N(p) is the local neighborhood of pixel p

To prevent over-segmentation from multiple maxima within a single crown, a minimum-distance constraint of 50 pixels was imposed, selected empirically and applied identically to both orchards and both segmentation outputs. At the working GSD of 0.7 cm·pixel^{-1}, this corresponds to ≈ 35 cm comparable to the radius of the smallest crowns preventing two maxima within a single small crown while preserving the separation of distinct adjacent trees. Each maximum was assigned a unique label to form the marker image, and the Watershed transform [41] was applied to the inverse distance map, constrained to the binary mask as defined in Equation 7:

$$L = Watrshed(-D, Marker, I)$$

Equation 7

Where:

L is the labeled output raster (each crown assigned a unique label)

$-D$ is the inverse of the distance transform map

Marker is the marker image of labeled local maxima (seed points)

I is the binary canopy mask constraining the segmentation

Yielding a labeled raster in which each crown is uniquely identified.

2.8 Evaluation Metrics

The segmentation pipeline was assessed at two complementary levels: pixel-level and object-level. Pixel-level metrics quantify the agreement between predicted binary canopy masks and reference masks, whereas object-level metrics assess the detection and delineation of individual tree crowns. Both levels are necessary, because a method can achieve high pixel-level overlap yet still fail to separate adjacent crowns into distinct objects.

ISSN: 1686-6576 (Printed) | ISSN 2673-0014 (Online) | © Geoinformatics International

2.8.1 Detection metrics (pixel and object levels)

Precision (Equation 8), recall (Equation 9), F1-score (Equation 10), and Intersection over Union (IoU) (Equation 11) provided the common basis for evaluation at both levels and were derived from counts of true positives (TP), false positives (FP), and false negatives (FN):

$$Precision=\frac{TP}{TP+FP}$$

Equation 8

$$Recall=\frac{TP}{TP+FN}$$

Equation 9

$$F1\text{-}Score=\frac{2\times Precision\times Recall}{Precision+Recall}$$

Equation 10

$$IoU=\frac{TP}{TP+FP+FN}$$

Equation 11

The two levels differ only in what constitutes a detection unit. At the pixel level, the binary canopy masks produced by the CHM-based morphological approach and the U-Net model were compared with the manually delineated reference masks, with canopy treated as the positive class and background as the negative class; a TP is a canopy pixel correctly labelled as canopy, an FP a background pixel labelled as canopy, and an FN a canopy pixel labelled as background. Overall pixel accuracy was not emphasized, because it is biased by the dominance of background pixels in orchard imagery. At the object level, each crown polygon obtained after Watershed separation was matched to an expert-annotated reference polygon under a one-to-one overlap rule: every predicted crown was assigned to the reference crown with which it shared the largest overlap, and each predicted and each reference crown could enter only one match, which prevents several detections from being assigned to the same reference. Here a TP is a predicted crown matched to a single reference crown, an FP a predicted crown that overlaps no reference crown, and an FN a reference crown with no matching prediction. For object-level reporting, the detection rate was additionally expressed as a percentage as defined in Equation 12:

$$Detection\ Score=\frac{TP}{TP+FP+FN}\times 100$$

Equation 12

2.8.2 Geometric agreement of matched crowns

Detection metrics show whether crowns were correctly identified but not how closely predicted and reference boundaries agree. Two further evaluations were therefore applied to the n true-positive crowns, both restricted to matched pairs. Area agreement compared the predicted canopy area ($\hat{A}_i$) with the reference area (A_i) through the coefficient of determination (R^2) (Equation 13), root mean square error (RMSE) (Equation 14), mean absolute error (MAE) (Equation 15), and mean absolute percentage error (MAPE) (Equation 16):

$$R^2=1-\frac{\sum_{i=1}^{n}\left(A_i-\hat{A}_i\right)^2}{\sum_{i=1}^{n}\left(A_i-\bar{A}\right)^2}$$

Equation 13

$$RMSE=\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(A_i-\hat{A}_t\right)^2}$$

Equation 14

$$MAE=\frac{1}{n}\sum_{i=1}^{n}\left|A_i-\hat{A}_i\right|$$

Equation 15

$$MAPE(\%)=\frac{100}{n}\sum_{i=1}^{n}\left|\frac{A_i-\hat{A}_t}{A_i}\right|$$

Equation 16

Where:

A_i is the reference canopy area of the i-th crown,
$\hat{A}_i$ is the predicted area
$\bar{A}$ is the mean reference area
n is the number of true-positive crowns.

Spatial agreement was then evaluated for each crown individually. Whereas Equation 11 expresses global mask overlap, the per-tree IoU expresses the overlap of a single crown as defined in Equation 17:

$$IoU_{tree,i}=\frac{\left|P_i\cap G_i\right|}{\left|P_i\cup G_i\right|}$$

Equation 17

Where:

P_i and G_i are the predicted and reference polygons of the i^{th} crown.

The mean per-tree *IoU* summarizes crown-level spatial agreement. Both geometric measures pertain to true positives only and were interpreted jointly with the object-level detection results.

ISSN: 1686-6576 (Printed) | ISSN 2673-0014 (Online) | © Geoinformatics International

3. Results

3.1 Tree Canopy Segmentation Performance

3.1.1 CHM-based morphological processing

The CHM-based pipeline segments canopy pixels from height information alone, applying the fixed 0.5 m threshold defined in Section 2.5, followed by hole filling and binary normalization. Because it uses no labeled training data, its accuracy is governed entirely by the separability of canopy and non-canopy heights within each orchard. Pixel-level performance, computed on the held-out test set, is reported in Table 4. In the Maroc Late orchard the method performed strongly, with a precision of 0.93, a recall of 0.99, an F1-score of 0.96, and an IoU of 0.92; the IoU remained below both precision and recall, as required by their definitions (Section 2.8.1). This success, despite the simplicity of a single height threshold, is explained by the wide height margin separating background from canopy in this orchard. The bare-soil inter-rows produced a mean CHM of only 0.03 ± 0.04 m (Section 2.3), so the 0.5 m threshold sits roughly 0.47 m more than ten standard deviations above the soil distribution; consequently, almost no background pixel is misclassified, which accounts for the high precision. The high recall follows from the fact that mature citrus crowns rise well above 0.5 m, so few canopy pixels fall below the threshold. With a mean spacing of about 5 m and a maximum canopy diameter of 4.4 m, crowns do not reach their neighbours, leaving open inter-crown gaps that the hole-filling step does not bridge; merging of adjacent crowns is therefore rare, and the residual 7% of false positives is attributable mainly to slight over-extension of the mask along crown margins.

Performance was markedly lower in the Sidi Aissa orchard (precision = 0.66, recall = 0.81, F1-score = 0.73, IoU = 0.57; IoU again below precision and recall). To identify the dominant factor behind this decline, the three candidate causes weeds, crown overlap, and tree spacing were examined against the structural data of the two orchards. Tree spacing can be ruled out as a driver: the planting densities of the two orchards are almost identical (394 versus 389 trees ha^{-1}) with comparable ~5 m spacing, so the gap cannot be explained by layout. The decline is instead governed by weeds and overlap. Weed-covered inter-row areas in Sidi Aissa reached a mean CHM of 0.29 ± 0.15 m (Section 2.3); although this mean lies below the 0.5 m threshold, its upper tail exceeds it. Assuming an approximately normal weed-height distribution, the fraction of weed pixels above the threshold is $P(z > (0.5 - 0.29)/0.15) = P(z > 1.40) \approx 0.08$, i.e. roughly 8% of the weed-covered area is retained as canopy and registered as false positives the principal cause of the precision collapse. Crown overlap is the second contributor: in Sidi Aissa the maximum canopy diameter (5.6 m) exceeds the inter-tree spacing (5 m), so the largest crowns necessarily touch, and the hole-filling step then closes the low-CHM gaps between them, converting genuinely non-canopy inter-crown pixels into canopy and further degrading precision and boundary fidelity. This contrasts with Maroc Late, where the maximum crown diameter (4.4 m) remains below the spacing. The lower recall (0.81) reflects a complementary error source on the canopy side: low-height canopy elements young replanted trees and thin crown edges whose CHM falls below 0.5 m are classified as background and missed.

Taken together, these results show that the height-threshold approach is well suited to orchards with clean inter-rows and crowns smaller than the planting spacing, but that its pixel-level accuracy degrades substantially in dense, weed-affected stands where (i) the upper tail of the weed-height distribution crosses the threshold and (ii) crowns exceed the spacing and are merged by hole filling. This behaviour is also visible qualitatively in Figure 7 crowns in well-separated configurations (Figure 7(a)) are delineated accurately, whereas under inter-row weeds and overlapping canopies (Figure 7(b)) the method incorporates weed regions and merges neighbouring crowns into single connected components.

3.1.2 U-Net model

Two U-Net models were trained independently on the RGB orthomosaic patches of the two orchards under the configuration described in Section 2.6.3 (Adam optimizer, binary cross-entropy loss, fixed learning rate of 1×10^{-4}, batch size 16, and 100 epochs, with final weights selected by the best-checkpoint strategy on the highest validation IoU). On the A100 GPU, training required approximately 6–7 min per epoch (≈10–12 h per model).

Table 4: Pixel-level evaluation of the CHM-based morphological processing on the held-out test set (the same test patches used to evaluate the U-Net) for both orchards

Orchard	Precision	Recall	F1-Score	IoU
Maroc Late	0.93	0.99	0.96	0.92
Sidi Aissa	0.66	0.81	0.73	0.57

ISSN: 1686-6576 (Printed) | ISSN 2673-0014 (Online) |

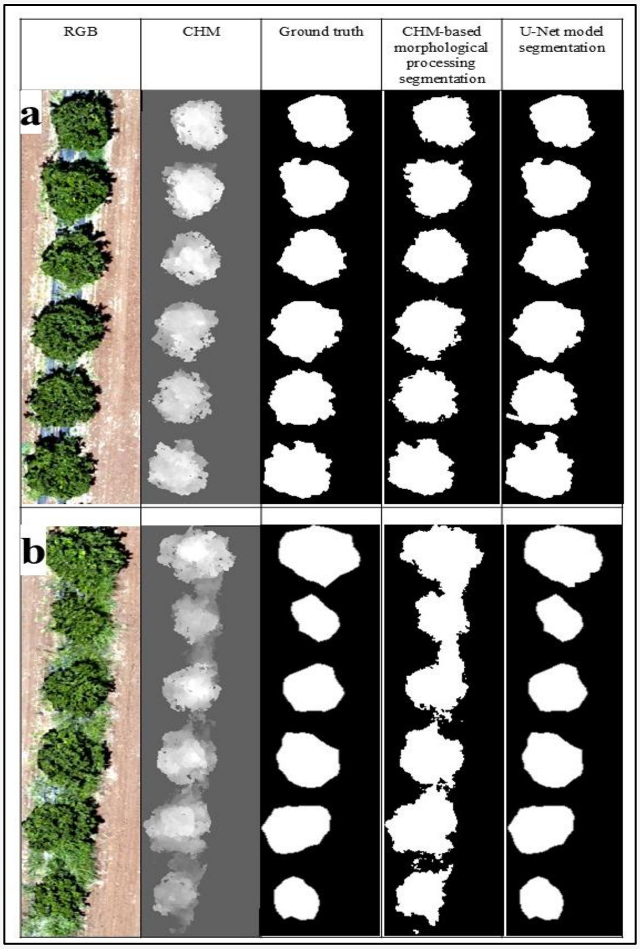

Figure 7: Canopy segmentation results under different orchard conditions: (a) segmentation of separate tree canopies, where individual crowns are well-defined, and (b) segmentation in a complex scenario with weeds between trees, affecting canopy separation

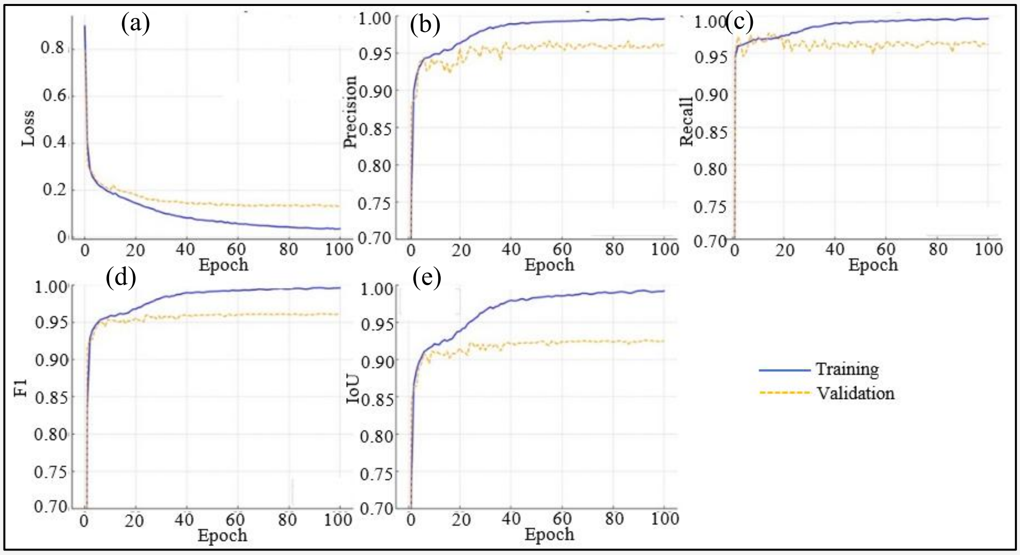

Figure 8: Training and validation performance curves for the Maroc Late orchard: (a) loss, (b) precision, (c) recall, (d) F1 score and (e) IOU

ISSN: 1686-6576 (Printed) | ISSN 2673-0014 (Online) | © Geoinformatics International

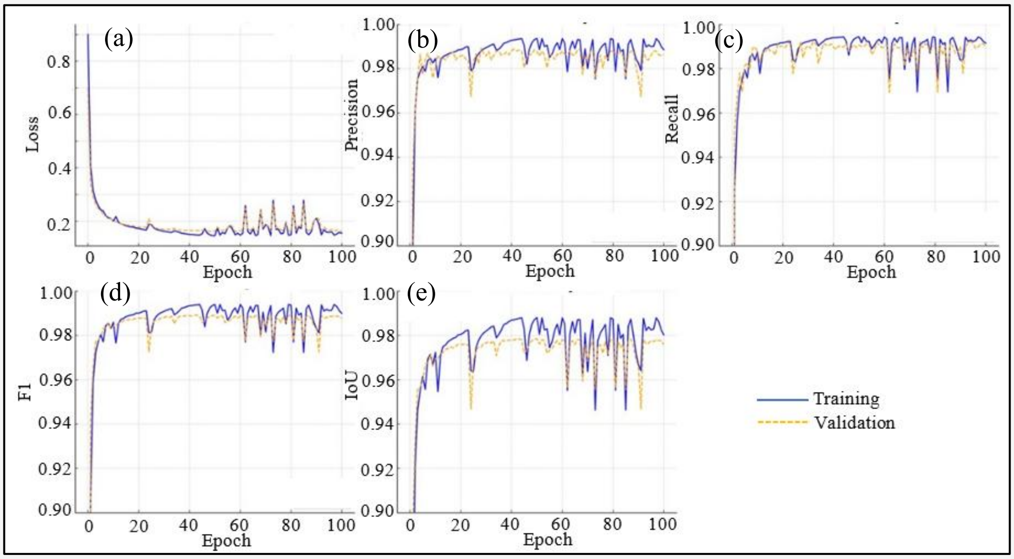

Figure 9: Training and validation performance curves for the Sidi Aissa orchard: (a) loss, (b) precision, (c) recall, (d) F1 score and (e) IOU

Table 5: Evaluation metrics for citrus tree canopy segmentation using the U-Net model in two citrus orchards

Orchard	Precision	Recall	F1-Score	IoU
Maroc Late	0.993	0.994	0.993	0.987
Sidi Aissa	0.964	0.964	0.964	0.931

The training and validation dynamics are shown in Figure 8 (Maroc Late) and Figure 9 (Sidi Aissa); these curves describe the training and validation subsets, whereas the final metrics in Table5 are computed on the independent test subset. In both orchards the training loss decreased steadily and the validation loss followed the same downward trend without diverging. To assess overfitting explicitly, the train–validation gap was quantified at the selected checkpoint: the IoU gap was approximately [≈0.06] for Maroc Late and [≈0.02] for Sidi Aissa (Figure 8(e)), Figure 9(e)). A gap below ~0.10 IoU, combined with non-diverging loss curves, indicates limited overfitting rather than a failure to generalize. The smaller gap for the structurally more complex Sidi Aissa orchard should not be read as easier segmentation; it is consistent with that orchard's higher canopy-pixel fraction (larger, more overlapping crowns), which lowers the pixel-averaged loss and the apparent gap independently of crown-boundary accuracy.

The Sidi Aissa validation curves nonetheless exhibited pronounced epoch-to-epoch oscillations, most visible in the validation IoU (Figure 9(e)). These fluctuations indicate a degree of training instability near convergence: with a fixed learning rate of 1×10^{-4} and the greater structural heterogeneity of this orchard (overlapping crowns, irregular crown morphology, and weed interference), successive mini-batches produced larger and less consistent gradient updates. The instability did not propagate to the final result, because the best-checkpoint strategy retained the highest-validation-IoU epoch rather than the last one; nevertheless, a decaying or cosine-annealing learning-rate schedule, or a longer run combined with early stopping, would likely have damped these oscillations and is recommended for future work.

The trained models were evaluated on the held-out test patches of each orchard (163 patches for Maroc Late and 432 for Sidi Aissa; Section 2.6.1). As reported in Table 5, the U-Net achieved a precision of 0.993, a recall of 0.994, an F1-score of 0.994, and an IoU of 0.987 on Maroc Late, with corresponding values of 0.964, 0.965, 0.964, and 0.931 on Sidi Aissa. In both orchards the IoU remained below the precision and recall, consistent with the metric definitions (Section 2.8.1). Performance was again slightly lower for Sidi Aissa, in line with its more complex canopy structure, and the gap relative to Maroc Late was considerably smaller than for the CHM-based pipeline: the Sidi Aissa IoU deficit fell from about 0.35 under the CHM method (Table 4) to about 0.06 under the U-Net (Table 5). The qualitative behaviour of the U-Net under complex conditions is consistent with these metrics (Figure 7). Whereas the CHM-based output incorporates weed regions and merges overlapping crowns, the U-Net preserves the separation of individual crowns even in the presence of inter-row weeds and canopy overlap (Figure 7(b)), reflecting its capacity to exploit spectral and textural cues rather than relying on height alone.

3.2 Individual Tree Crown Delineation Results

The binary canopy masks from both pipelines were separated into individual crowns using the marker-controlled Watershed procedure of Section 2.7 and evaluated at the object level for detection and geometric agreement. The detection metrics and mean per-tree IoU are reported in Table 6, the area-agreement regressions in Figure 10, and the workflow in Figure 11. Because the separation step and its parameters were identical across both methods and orchards, any object-level difference reflects the quality of the input mask rather than the Watershed algorithm, so the comparison isolates the contribution of the upstream segmentation.

Table 6: Evaluation of individual tree canopy delineation performance

Orchard	round ruth rees	Segmentation Approach	Detected Trees	TP	FP	FN	Mean per-tree IoU	Precision	Recall	F1	Detection Score (%)
Maroc Late	2409	CHM-Based Morphological Processing	2503	2382	121	27	0.88	0.95	0.99	0.97	94.15
		Deep Learning Approach: U-Net Model	2518	2396	122	13	0.92	0.95	0.99	0.97	94.67
Sidi Aissa	4782	CHM-Based Morphological Processing	5012	3125	1887	1657	0.59	0.62	0.65	0.64	46.86
		Deep Learning Approach: U-Net Model	4947	4703	244	79	0.84	0.95	0.98	0.97	93.57

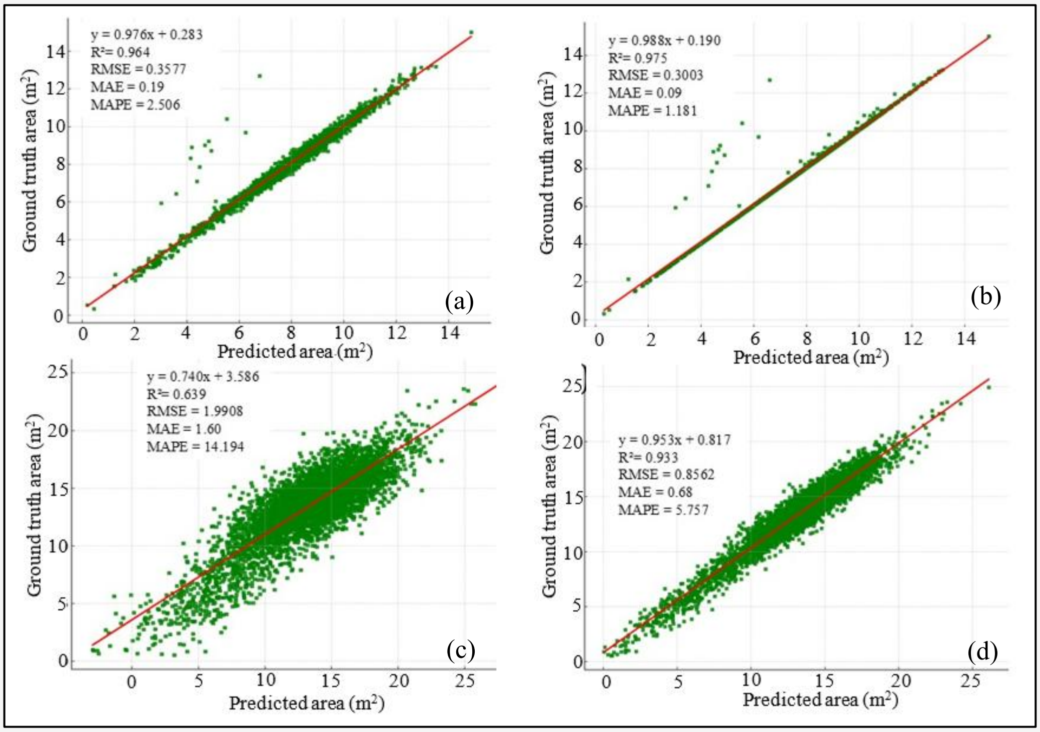

Figure 10: Linear regression analysis of canopy area estimation for true positive tree crowns using Watershed segmentation: (a, b) Maroc Late orchard results with CHM-based processing and U-Net model, respectively, and (c, d) Sidi Aissa orchard results with CHM-based processing and U-Net model, respectively

ISSN: 1686-6576 (Printed) | ISSN 2673-0014 (Online) | © Geoinformatics International

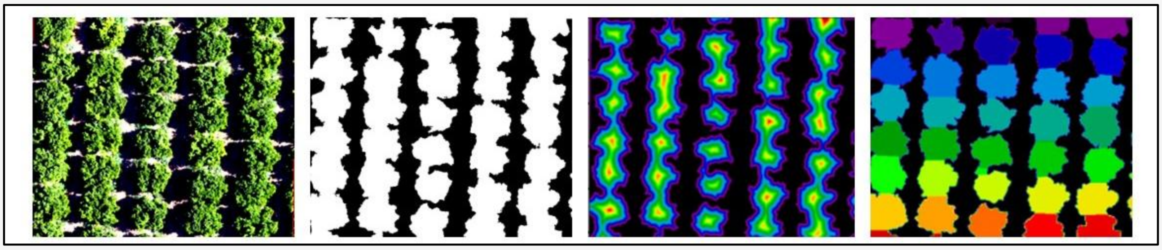

Figure 11: Individual tree canopy delineation process: (a) RGB image, (b) Binary mask representing segmented tree canopies, (c) Distance transform computation used for tree crown separation and (d) Final tree canopy delineation using Watershed segmentation

This is borne out by the Maroc Late orchard, where the two methods were effectively indistinguishable: when the mask is clean, the Watershed step alone individualises crowns and the segmentation method is largely immaterial. The pipelines diverged only in Sidi Aissa, where the U-Net sustained a detection score close to that of Maroc Late while the detection score of the CHM-based pipeline fell sharply. This fall does not mean that few crowns were found the CHM pipeline still recovered roughly two-thirds of the reference crowns (recall ≈ 0.65) but that a large number of spurious detections from weeds and merged crowns drove its detection score down. With the separation algorithm held constant, this decline is a direct consequence of the degraded CHM mask under dense, weed-affected conditions, consistent with the pixel-level loss of precision reported in Section 3.1.1.

The CHM degradation in Sidi Aissa reflects concurrent over- and under-detection, as indicated by the comparable false-positive and false-negative counts in Table 6. Weed patches exceeding the 0.5 m threshold are retained in the mask and partitioned into spurious objects, inflating false positives; overlapping crowns bridged by hole filling coalesce into single components that yield one marker per cluster, leaving the remaining reference crowns as false negatives; and low or recently replanted trees below the threshold are omitted entirely. The first mechanism raises the false-positive count and the latter two the false-negative count, which explains the simultaneous prevalence of both. The U-Net mask suppresses the weed signal and preserves inter-crown boundaries, attenuating all three mechanisms and bringing the Sidi Aissa detection score close to the Maroc Late level. Detection is summarised by the detection score, which is the Jaccard index of the predicted and reference crown sets and therefore penalises false positives and false negatives jointly rather than measuring the fraction of crowns recovered; object-level detection admits no true-negative term, as the background is not an enumerable population of non-crown objects.

Geometric fidelity, assessed for matched crowns through the mean per-tree IoU and the predicted–reference area regression (Figure 10), reproduced the detection ranking. In the Maroc Late orchard both pipelines recovered crown area almost perfectly, with near-unity slopes and closely comparable coefficients of determination (R^2 = 0.964 for the CHM method and R^2 = 0.975 for the U-Net), together with low relative errors (MAPE = 2.5% and 1.2%, respectively; Figure 10(a) and figure 10(b). The two methods separated sharply only in Sidi Aissa: the U-Net retained a strong, near-1:1 agreement (slope = 0.953, R^2 = 0.933, MAPE = 5.8%; Figure 10(d), whereas the CHM-based regression was both weaker and biased (slope = 0.740, intercept = 3.59 m^2, R^2 = 0.639, MAPE = 14.2%; Figure 10(c), its depressed slope and large intercept indicating a systematic departure from the reference areas where merged clusters and retained weed regions distort the segmented crown size. The regression was restricted to true-positive crowns by design, since false positives have no reference area and false negatives no predicted area; their inclusion would require artificial zero values and would convert the analysis from one of crown-area agreement among matched trees into a composite detection-and-segmentation measure already provided by the detection score. The two measures are thus complementary, and the regression statistics should be read together with the detection score, particularly for the CHM pipeline in Sidi Aissa, where the low detection score means that even this R^2 = 0.639 characterises only the correctly detected subset and would be lower still over all crowns. The full delineation sequence RGB input, binary mask, distance transform, and final Watershed crowns is shown for a representative subset in Figure 11.

4. Discussion

Because the two pipelines are identical at every stage after the canopy mask is formed, their divergence measures the value of the information each mask carries as encoded by these two specific

ISSN: 1686-6576 (Printed) | ISSN 2673-0014 (Online) | © Geoinformatics International

implementations. The morphological pipeline evaluated here reduces every pixel to a single quantity, its height, and applies a single global threshold; in this form it succeeds only where canopy and non-canopy occupy distinct height ranges and it makes no use of the spectral and textural cues that separate citrus foliage from other objects of similar height. The U-Net instead learns spectral, textural, and contextual features from the RGB imagery the colour and texture that distinguish foliage from weeds and the context that fixes inter-crown boundaries and it is this that underlies its robustness. The performance gap therefore originates not in the delineation algorithm but in what each representation, as implemented, encodes. This contrast should be read as a comparison between a deliberately minimal height-threshold pipeline and a learned RGB model, not as a general verdict on CHM- or height-based segmentation, since point-cloud, multi-scale, or locally adaptive height methods exploit the same data differently and need not share these limitations.

This framing also clarifies why the present morphological approach fails specifically as canopy density and weed cover rise, and which of its failures are intrinsic and which are design choices. One failure is intrinsic to a height-only representation: where the height ranges of inter-row vegetation and of low or young citrus canopy overlap, height alone cannot separate them, so a height-based mask is fundamentally ambiguous in those regions. The other two failures, however, follow from specific design decisions rather than from height data as such. The use of a single global threshold forces one cutoff across the whole orchard, so that recovering low or young canopy necessarily admits more weeds; a locally adaptive or variable-window threshold could relax this trade-off. Likewise, the hole-filling step assumes isolated crowns and, by construction, bridges the genuine low-height saddles between overlapping neighbours, erasing the boundaries the separation step depends on; a method that did not impose this assumption would not merge crowns in the same way. The breakdown observed here is therefore structural for this pipeline, but it reflects the simplicity of the chosen height-thresholding scheme as much as any limit of height information itself.

The methods also fail in distinct ways. The crowns missed by the U-Net cluster among very small or recently replanted trees with faint signatures, crowns in the deep shadow of taller neighbours, and individuals so entangled with surrounding canopy or weeds that no boundary is visible even to a human interpreter; the morphological errors are dominated instead by weeds mistaken for crowns and by merged clusters. A third mode is common to both and independent of the mask: the Watershed step splits a large or irregular crown bearing several distance-transform maxima into separate objects (Figure 12), which motivates future end-to-end instance-segmentation models. In deployment terms, the morphological route front-loads its cost into the photogrammetric DTM and CHM, which must be rebuilt for every acquisition, whereas the U-Net carries a one-off training cost and then runs on RGB tiles alone, making it the more economical option for repeated monitoring.

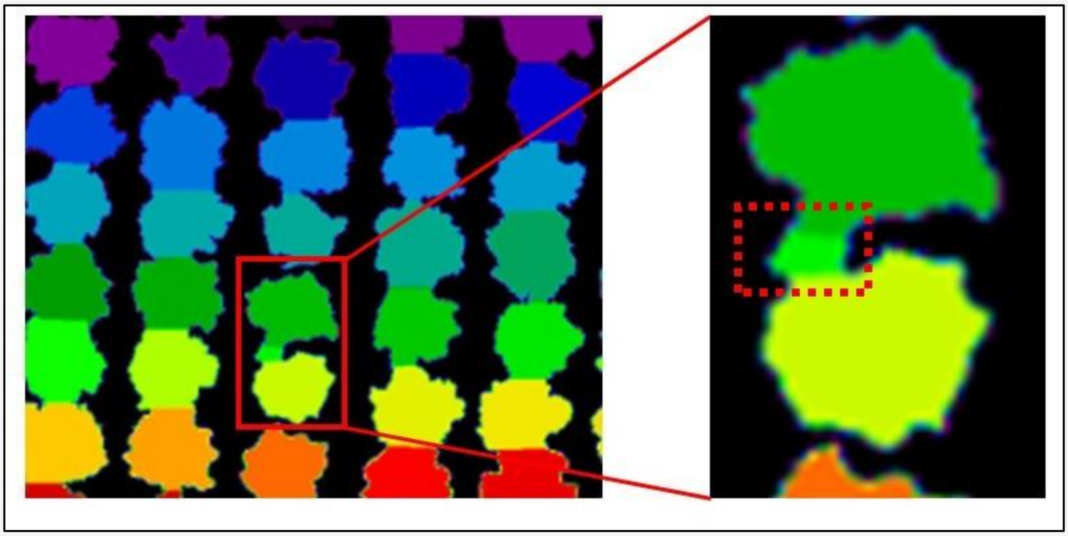

Figure 12: Example of over-segmentation where part of a citrus tree crown is incorrectly detected as a separate tree

The comparison with earlier orchard-delineation studies in Table 1 is qualitative rather than a performance ranking, since the cited works differ in species, GSD, platform, annotation protocol, evaluation metric, and location, with no shared benchmark; their absolute percentages are therefore not commensurable and no claim to outperform them is made. The recurring pattern they show, and which this study reproduces internally by holding the delineation step fixed, is that learned segmentation retains accuracy under structural complexity where simple height- or maxima-based methods decline, even as more elaborate height-based pipelines may narrow that gap. These conclusions are bounded by the study design. The work used one site, one date, and one UAV platform, and a single species citrus represented by two varieties, Maroc Late and Sidi Aissa; the two orchards lie within a few hundred metres and were imaged under identical illumination, so the agreement between them cannot be extrapolated to other species, growth stages, sensors, or regions. The morphological pipeline, in turn, is only as reliable as the photogrammetric reconstruction beneath its CHM and, as noted above, represents a deliberately simple baseline rather than the best achievable height-based method. The broader inference that learned models suit complex orchards is accordingly offered as a hypothesis consistent with these two citrus varieties and with the particular methods compared, to be confirmed by multi-site, multi-date, and multi-species evaluation that also includes more advanced height-based segmentation.

5. Conclusion

This study compared two paradigmatically opposed strategies for delineating individual citrus tree crowns from UAV imagery a CHM-based morphological pipeline and a U-Net trained on RGB data under a shared, parameter-identical marker-controlled Watershed step, so that any observed difference could be attributed to the upstream representation rather than to the separation algorithm. Examined across two orchards of deliberately contrasting structure, the two approaches proved effectively equivalent where crowns were well spaced and inter-rows were clear, and diverged only as canopy overlap and weed cover increased, conditions under which the learned representation remained reliable while the height-based one did not. The principal outcome is therefore not a ranking of methods but a characterisation of the regime in which they differ.

The central contribution of the work is methodological as much as practical. By holding the delineation stage fixed and stratifying the comparison by structural complexity, the study converts a comparison that is usually confounded across several pipeline stages into a controlled one, and in doing so provides a defensible, condition-dependent basis for method selection. For structurally simple, weed-free orchards, the morphological pipeline is the rational default, since it attains delineation accuracy comparable to the deep-learning model without annotated training data; for heterogeneous orchards with overlapping crowns and inter-row vegetation, the RGB-trained model is the more dependable choice. The appropriate recommendation is thus to match the method to the orchard rather than to adopt a single approach across all conditions.

These conclusions are necessarily bounded by the scope of the experiment, which drew on a single acquisition over two neighbouring orchards of two varieties. The agreement observed between them should accordingly be treated as a hypothesis to be tested rather than a general property, and broader validation across additional sites, acquisition dates, citrus species, and UAV platforms using spatially independent training and test partitions and repeated training runs remains necessary before the findings can be generalised. Future work could also build on the complementary behaviour of the two representations by integrating canopy-height and RGB information within a single network rather than treating them as competing pipelines, while end-to-end instance-segmentation architectures offer a route to removing the residual over-segmentation that the shared Watershed step introduces.

Acknowledgments

We sincerely thank the anonymous reviewers for their constructive comments and valuable suggestions.

References

[1] Ye, X. and Sakai, K., (2012). *Fruit Yield Estimation through Multispectral Imaging.* In: A. K. Srivastava. Advances in Citrus Nutrition. Dordrecht, 453-473. https://doi.org/10.1007/978-94-007-4171-3_30.

[2] Jin, X., Liu, S., Baret, F., Hemerlé, M. and Comar, A., (2017). Estimates of Plant Density of Wheat Crops at Emergence from Very Low Altitude UAV Imagery. *Remote Sensing of Environment,* Vol.198. https://doi.org/10.1016/j.rse.2017.06.007.

[3] Anthony, B. M. and Minas, I. S., (2021). Optimizing Peach Tree Canopy Architecture for Efficient Light Use, Increased Productivity and Improved Fruit Quality. *Agronomy,* Vol. 11. https://doi.org/10.3390/agronomy11101961.

[4] Malek, S., Bazi, Y., Alajlan, N., AlHichri, H. and Melgani, F., (2014). Efficient Framework for Palm Tree Detection in UAV Images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,* Vol. 7. https://doi.org/10.1109/JSTARS.2014.2331425.

[5] Salamí, E., Gallardo, A., Skorobogatov, G. and Barrado, C., (2019). On-the-Fly Olive Tree Counting Using a UAS and Cloud Services. *Remote Sensing,* Vol. 11. https://doi.org/10.3390/rs11030316.

[6] Shakoor, N., Lee, S. and Mockler, T. C., (2017). High Throughput Phenotyping to Accelerate Crop Breeding and Monitoring of Diseases in the Field. *Current Opinion in Plant Biology,* Vol. 38. https://doi.org/10.1016/j.pbi.2017.05.006

[7] Daiaeddine, M. J., Badrouss, S., El Harti, A., Bachaoui, E. M., Biniz, M. and Mouncif, H., (2025). UAV Imagery, Advanced Deep Learning, and Yolov7 Object Detection Model in Enhancing Citrus Yield Estimation. *Foods and Raw Materials,* Vol.13. https://doi.org/10.21603/2308-4057-2025-2-650.

[8] El Imanni, H. S., El Harti, A., Bachaoui, E. M., Mouncif, H., Eddassouqui, F., Hasnai, M. A. and Zinelabidine, M. I., (2023). Multispectral UAV Data for Detection of Weeds in a Citrus Farm using Machine Learning and Google Earth Engine: Case Study of Morocco. *Remote Sensing Applications: Society and Environment,* Vol. 30. https://doi.org/10.1016/j.rsase.2023.100941.

[9] Jemaa, H., Bouachir, W., Leblon, B., LaRocque, A., Haddadi, A. and Bouguila, N., (2023). Uav-Based Computer Vision System for Orchard Apple Tree Detection and Health Assessment. *Remote Sensing,* Vol. 15. https://doi.org/10.3390/rs15143558.

[10] Lucena, F., Breunig, F. M. and Kux, H., (2022). The Combined Use of UAV-Based Rgb and Dem Images for the Detection and Delineation of Orange Tree Crowns with Mask R-Cnn: An Approach of Labeling and Unified Framework. *Future Internet,* Vol.14. https://doi.org/10.3390/fi14100275.

[11] Tripicchio, P., Satler, M., Dabisias, G., Ruffaldi, E. and Avizzano, C. A., (2015). Towards Smart Farming and Sustainable Agriculture with Drones. 2015 International Conference on Intelligent Environments. July 15–17, 2015, Prague, Czech Republic. https://doi.org/10.1109/IE.2015.29.

[12] Sarabia, R., Aquino, A., Ponce, J. M., López, G. and Andújar, J. M., (2020). Automated Identification of Crop Tree Crowns from Uav Multispectral Imagery by Means of Morphological Image Analysis. *Remote Sensing*, Vol.12. https://doi.org/10.3390/rs12050748.

[13] Koc-San, D., Selim, S., Aslan, N. and San, B. T., (2018). Automatic Citrus Tree Extraction from Uav Images and Digital Surface Models Using Circular Hough Transform. Computers and Electronics in Agriculture, Vol.150. https://doi.org/10.1016/j.compag.2018.05.001.

[14] Ok, A. O. and Ozdarici-Ok, A., (2018). 2-D Delineation of Individual Citrus Trees from Uav-Based Dense Photogrammetric Surface Models. *International Journal of Digital Earth,* Vol.11. https://doi.org/10.1080/17538947.2017.1337820.

[15] Donmez, C., Villi, O., Berberoglu, S. and Cilek, A., (2021). Computer Vision-Based Citrus Tree Detection in a Cultivated Environment Using Uav Imagery. *Computers and Electronics in Agriculture,* Vol.187. https://doi.org/10.1016/j.compag.2021.106273.

[16] Chen, Y., Hou, C., Tang, Y., Zhuang, J., Lin, J., He, Y., Guo, Q., Zhong, Z., Lei, H. and Luo, S., (2019). Citrus Tree Segmentation from Uav Images Based on Monocular Machine Vision in a Natural Orchard Environment. *Sensors,* Vol.19. https://doi.org/10.3390/s19245558.

[17] Dong, X., Zhang, Z., Yu, R., Tian, Q. and Zhu, X., (2020). Extraction of Information about Individual Trees from High-Spatial-Resolution Uav-Acquired Images of an Orchard. *Remote Sensing,* Vol.12. https://doi.org/10.3390/rs12010133.

[18] Partel, V., Kakarla, S. C. and Ampatzidis, Y., (2019). Development and Evaluation of a Low-Cost and Smart Technology for Precision Weed Management Utilizing Artificial Intelligence. *Computers and Electronics in Agriculture,* Vol.157. https://doi.org/10.1016/j.compag.2018.12.048.

[19] LeCun, Y., Bengio, Y. and Hinton, G., (2015). Deep Learning. *Nature,* Vol. 521. https://doi.org/10.1038/nature14539.

[20] On, C., Kanniah, K., and Lau, A. (2024). A Multiscale Quantification of Leaf Area Index with Unmanned Aerial Vehicle and SPOT-7 Satellite Imageries: A case of Kuala Lumpur, Malaysia. *International Journal of Geoinformatics*, Vol. 20(5), 54–68. https://doi.org/10.52939/ijg.v20i5.3231.

[21] Kattenborn, T., Leitloff, J., Schiefer, F. and Hinz, S., (2021). Review on Convolutional Neural Networks (CNN) in Vegetation Remote Sensing. *ISPRS Journal of Photogrammetry and Remote Sensing,* Vol.173. https://doi.org/10.1016/j.isprsjprs.2020.12.010.

[22] Yu, K., Hao, Z., Post, C. J., Mikhailova, E. A., Lin, L., Zhao, G., Tian, S. and Liu, J., (2022). Comparison of Classical Methods and Mask R-CNN for Automatic Tree Detection and Mapping using UAV Imagery. *Remote Sensing,* Vol. 14. https://doi.org/10.3390/rs14020295.

[23] Wu, J., Yang, G., Yang, H., Zhu, Y., Li, Z., Lei, L. and Zhao, C., (2020). Extracting Apple Tree Crown Information from Remote Imagery Using Deep Learning. Computers and Electronics in Agriculture, Vol. 174. https://doi.org/10.1016/j.compag.2020.105504.

[24] Ye, Z., Wei, J., Lin, Y., Guo, Q., Zhang, J., Zhang, H., Deng, H. and Yang, K., (2022). Extraction of Olive Crown Based on UAV Visible Images and the U2-Net Deep Learning Model. *Remote Sensing,* Vol. 14. https://doi.org/10.3390/rs14061523.

[25] Wei, P., Yan, X., Yan, W., Sun, L., Xu, J. and Yuan, H., (2024). Precise Extraction of Targeted Apple Tree Canopy with Yolo-Fi Model for Advanced Uav Spraying Plans. *Computers and Electronics in Agriculture,* Vol. 226. https://doi.org/10.1016/j.compag.2024.109425.

[26] Vincent, L. and Soille, P., (1991). Watersheds in Digital Spaces: An Efficient Algorithm Based on Immersion Simulations. *IEEE Transactions on Pattern Analysis & Machine Intelligence,* Vol.13. https://doi.org/10.1109/34.87344.

[27] Beucher, S. and Meyer, F., (2018). T*he Morphological Approach to Segmentation: The Watershed Transformation*. In: E. R. Dougherty. Mathematical Morphology in Image Processing. Boca Raton, FL. 433-481. https://doi.org/10.1201/9781482277234-12.

[28] Dalponte, M. and Coomes, D. A., (2016). Tree-Centric Mapping of Forest Carbon Density from Airborne Laser Scanning and Hyperspectral Data. Methods in Ecology and Evolution, Vol. 7. https://doi.org/10.1111/2041-210X.12575.

[29] Strîmbu, V. F. and Strîmbu, B. M., (2015). A Graph-Based Segmentation Algorithm for Tree Crown Extraction Using Airborne Lidar Data. *ISPRS Journal of Photogrammetry and Remote Sensing,* Vol. 104. https://doi.org/10.1016/j.isprsjprs.2015.01.018.

[30] He, K., Gkioxari, G., Dollár, P. and Girshick, R., (2017). Mask R-Cnn: Proceedings of the IEEE International Conference on Computer Vision (ICCV 2017). Venice, Italy. https://doi.org/10.1109/ICCV.2017.322.

[31] Malpica, N., De Solórzano, C. O., Vaquero, J. J., Santos, A., Vallcorba, I., García-Sagredo, J. M. and del Pozo, F., (1997). Applying Watershed Algorithms to the Segmentation of Clustered Nuclei. Cytometry: *The Journal of the International Society for Analytical Cytology,* Vol. 28. https://doi.org/10.1002/(sici)1097-0320(19970801)28:4%3C289::aid-cyto3%3E3.0.co;2-7.

[32] Chen, Q., Baldocchi, D., Gong, P. and Kelly, M., (2006). Isolating Individual Trees in a Savanna Woodland Using Small Footprint Lidar Data. *Photogrammetric Engineering & Remote Sensing,* Vol. 72. https://doi.org/10.14358/PERS.72.8.923.

[33] Meyer, F. and Beucher, S., (1990). Morphological Segmentation. *Journal of Visual Communication and Image Representation,* Vol. 1. https://doi.org/10.1016/1047-3203(90)90014-M.

[34] Jing, L., Hu, B., Noland, T. and Li, J., (2012). An Individual Tree Crown Delineation Method Based on Multi-Scale Segmentation of Imagery. *ISPRS Journal of Photogrammetry and Remote Sensing,* Vol. 70. https://doi.org/10.1016/j.isprsjprs.2012.04.003.

[35] Samet, H. and Tamminen, M., (2002). Efficient Component Labeling of Images of Arbitrary Dimension Represented by Linear Bintrees. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* Vol. 10. https://doi.org/10.1109/34.3918.

[36] Ronneberger, O., Fischer, P. and Brox, T., (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation: International Conference on Medical Image Computing and Computer-Assisted Intervention. October 5–9, 2015, Munich, Germany. Cham: Springer. https://doi.org/10.1007/978-3-319-24574-4_28

[37] Ulyanov, D., Vedaldi, A. and Lempitsky, V., (2016). Instance Normalization: The Missing Ingredient for Fast Stylization. arXiv preprint arXiv:1607. https://doi.org/10.48550/arXiv.1607.08022.

[38] Glorot, X., Bordes, A. and Bengio, Y., (2011). Deep Sparse Rectifier Neural Networks. *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics.* April 11–13, 2011, Fort Lauderdale, FL, USA. https://proceedings.mlr.press/v15/glorot11a.html.

[39] Gonzalez, R. C., (2009). *Digital Image Processing.* ed., Pearson Education India.

[40] Nixon, M. and Aguado, A. S., (2025). *Feature Extraction and Image Processing for Computer Vision.* ed., Academic Press,

[41] Soille, P., (1999). *Morphological Image Analysis: Principles and Applications.* ed., Springer, https://doi.org/10.1007/978-3-662-03939-7.