



Article

Citrus Tree Segmentation from UAV Images Based on Monocular Machine Vision in a Natural Orchard Environment

Yayong Chen ^{1,†}, Chaojun Hou ^{1,†}, Yu Tang ^{1,*} , Jiajun Zhuang ¹, Jintian Lin ¹, Yong He ² ,
Qiwei Guo ^{1,*}, Zhenyu Zhong ³, Huan Lei ³ and Shaoming Luo ¹

¹ Academy of Contemporary Agriculture Engineering Innovations, Zhongkai University of Agriculture and Engineering, Guangzhou 510225, China; 441200923@163.com (Y.C.); houchaojun@zhku.edu.cn (C.H.); zhuangjiajun@zhku.edu.cn (J.Z.); linjtian@163.com (J.L.); smluo@gdut.edu.cn (S.L.)

² College of Biosystems Engineering and Food Science, Zhejiang University, Hangzhou 310058, China; yhe@zju.edu.cn

³ Guangdong Key Laboratory of Modern Control Technology, Guangdong Institute of Intelligent Manufacturing, Guangzhou 510070, China; zy.zhong@giim.ac.cn (Z.Z.); huan.l@giim.ac.cn (H.L.)

* Correspondence: ty2008@zhku.edu.cn (Y.T.); guoqiwei@zhku.edu.cn (Q.G.); Tel.: +86-020-8900-3790 (Y.T. & Q.G.)

† These two authors contributed equally to this work and should be considered co-first authors.

Received: 18 November 2019; Accepted: 12 December 2019; Published: 16 December 2019



Abstract: The segmentation of citrus trees in a natural orchard environment is a key technology for achieving the fully autonomous operation of agricultural unmanned aerial vehicles (UAVs). Therefore, a tree segmentation method based on monocular machine vision technology and a support vector machine (SVM) algorithm are proposed in this paper to segment citrus trees precisely under different brightness and weed coverage conditions. To reduce the sensitivity to environmental brightness, a selective illumination histogram equalization method was developed to compensate for the illumination, thereby improving the brightness contrast for the foreground without changing its hue and saturation. To accurately differentiate fruit trees from different weed coverage backgrounds, a chromatic aberration segmentation algorithm and the Otsu threshold method were combined to extract potential fruit tree regions. Then, 14 color features, five statistical texture features, and local binary pattern features of those regions were calculated to establish an SVM segmentation model. The proposed method was verified on a dataset with different brightness and weed coverage conditions, and the results show that the citrus tree segmentation accuracy reached $85.27\% \pm 9.43\%$; thus, the proposed method achieved better performance than two similar methods.

Keywords: agricultural unmanned aerial vehicles; monocular computer vision; tree crown segmentation; circumstance brightness; weed environment orchard

1. Introduction

In recent years, fully autonomous agricultural unmanned aerial vehicles (UAVs) were widely applied in orchards, and target detection (e.g., of fruit trees) constitutes one of the key technologies for autonomous operation [1–3]. However, unstructured factors (e.g., brightness condition (BC) and weed coverage condition (WCC)) in a complex orchard environment affect the fruit tree detection accuracy [4–8]. Currently, methods employed for the detection of fruit trees with agricultural UAVs consist mainly of spectral imaging and machine vision technologies; as such, the accuracy of these fruit tree detection methods depends greatly on the tree segmentation accuracy [9–11].

Many detailed research studies were performed to improve tree segmentation accuracy. Csillik et al. [12] segmented fruit tree crowns with a convolutional neural network (CNN) and

a simple linear iteration classification algorithm and detected citrus trees in complex agricultural environments within spectral citrus image datasets obtained with a multi-spectral camera onboard a UAV. Johansen et al. [13] segmented canopies of trees based on a three-dimensional (3D) model containing the height, geometry, and size of the trees and evaluated the pruning effect of crops in multi-spectral images collected by a multi-spectral UAV system on ornamental plants at multiple altitudes. Srestasathiern et al. [14] segmented oil palm trees using a vegetation index through a non-maximum suppression algorithm and monitored the number of trees in high-resolution spectral images. Roope et al. [11] segmented separate spruce trees by establishing a 3D hyper-spectral tree model with hyper-spectral UAV images and a watershed algorithm and estimated the degree of damage wrought by longhorn beetles at the single-tree level with an estimation spectral curve. Malambo et al. [15] segmented crop canopies by establishing a 3D structure from motion (SfM) model with high-resolution UAV images and evaluated corn height parameters in the field. Torres-Sanchez et al. [16] segmented a planted forest by generating a 3D ground model and automatically monitored the forest by extracting geometric parameters through an object-based extraction method in a high-resolution UAV imaging system. Juan et al. [17] segmented individual tree canopies in a digital surface model reconstructed by SfM technology and estimated the canopy position, height, and diameter. Guo et al. [18] segmented a forest canopy in a geometric point cloud model established by a UAV–light detection and ranging (LiDAR) system and captured the forest canopy height, canopy coverage and terrestrial biomass of three ecosystems, namely, coniferous broad-leaved mixed, evergreen broad-leaf, and mangrove forests. Pedro et al. [19] segmented chestnut trees in image data from a red/green/blue (RGB) + infrared camera UAV system through an object clustering extraction method and automatically monitored the trees according to their geometric features and the canopy coverage rate of the resulting area. Wu et al. [20] segmented the area of trees with a watershed method and a polynomial fitting method in a 3D forest model established by a UAV–LiDAR system and calculated the canopy coverage of a planted forest with a canopy height model and multiple linear regression model. Li et al. [21] segmented the crowns of separate trees in a geometric point cloud image with a 3D forest model established by an airborne laser scanner and achieved the maintenance and management of forest ecology with a global navigation satellite system and an inertial measurement unit auxiliary structure with SfM technology. Omair et al. [22] segmented trees on urban grass with color parameters, a gray-level co-occurrence matrix (GLCM) parameter and a clustering algorithm after eliminating lens distortion and counted the trees automatically with an RGBUAV camera; the segmentation technique achieved an accuracy of 70% and demonstrated good applicability for estimating forest degradation. Maciel et al. [23] detected the center line of citrus trees in a high-density orchard with a CNN algorithm in a sliding window and segmented single trees with a CNN; the results showed an overall accuracy of 94% in seven different test orchards. Ramesh et al. [24] segmented single fruit trees with an extreme learning machine, a geometry filtering threshold, and a watershed separation algorithm and detected the numbers of banana, mango, and coconut trees in different orchards with high-resolution RGB cameras onboard a fixed-wing UAV and a multi-rotor UAV with an accuracy of 85%. Lin et al. [25] segmented the areas of single-tree canopies in oblique UAV images with k-means clustering in the $L^*a^*b^*$ color space (defined by the International Commission on Illumination (CIE) in 1976) and a threshold method with pseudo-NDVI color mapping and texture mapping technology. Carlos et al. [26] applied four segmentation methods (k-means, artificial neural network (ANN), random forest (RForest), and spectral indices (SI)) to grape tree canopies and found that the SI+ANN and RForest methods were superior, with an accuracy of approximately 0.98 in high-resolution UAV images of trees under different shade and soil conditions, useful for the exquisite management of commercial vineyards. Zhao et al. [27] segmented regions of pomegranate trees using the U-net and a region-CNN with a high-resolution visible imaging and multi-spectral imaging UAV system and calculated the water stress parameter and nutritional status of the trees in multi-spectral data. Ultimately, they found that the region-CNN provided better segmentation results. Corey et al. [28] segmented seedling trees in cut-down forest images acquired by an automatically controlled UAV with a chromatic aberration

segmentation algorithm in combination with a vertical take-up and take-off algorithm and satellite positioning technology. In addition, they counted the number of seedling trees using a classification and regression tree machine learning model to estimate the forest regeneration rate. Sabzi et al. [29] developed a computer vision method to detect apples in trees and estimated their ripeness based on the most effective color features and ANN classifier; the proposed method achieved an accuracy of 97.88%. Li et al. [30] presented a corn classification method based on computer vision and the maximum-likelihood estimator for classifying normal and damaged corn, and the results showed an overall accuracy of 96.67%.

Although many tree segmentation methods were studied for UAV images in recent years, those methods were mainly applied in structured environments. In contrast, the use of segmentation methods in unstructured environments, especially in citrus orchards under different brightness and weed coverage conditions (e.g., some weeds in a citrus orchard can improve the physical and chemical properties of the soil and attract beneficial insects [31]), is yet to be reported. Considering the cost of equipment, the coverage of the algorithm, the detection accuracy, and the portability of the system, this paper proposes a citrus tree segmentation method based on monocular machine vision in a natural orchard environment. Therefore, the main processes used to segment citrus trees in unstructured environments were as follows: (1) development of an illumination compensation algorithm to pre-process the under-lit foreground in citrus orchard images, thereby reducing the sensitivity of the algorithm to the environmental brightness, (2) extraction of the potential regions of interest (ROIs, i.e., the citrus trees) from weed elements as accurately as possible by combining chromatic aberration technology and the Otsu threshold method, and (3) accurate segmentation of trees from orchard images with a binary detection support vector machine (SVM) established based on the calculated values of 14 color features, five texture features, and local binary pattern (LBP) features.

2. Materials

To debug and test the proposed method, a monocular RGB camera was outfitted on an agricultural UAV and used to acquire image datasets with in the height range of the agricultural UAV during operation. The equipment and the data are briefly introduced in this section.

2.1. The Study Area and the UAV Equipment

The study area is a commercial orange orchard of Conghua Hualong Fruit and Vegetable Preservation Co., LTD, located in Conghua District in Guangzhou, China (113°31'15.5" east (E), 23°38'42.4" north (N)). The orchard covers an area of 0.6 km², and the center of its terrain is 8.9 m higher than its edge. There are 453 "Shatangju citrus" (*Citrus reticulata* Blanco) trees with tree ages of 4.5 years in the orchard. Furthermore, the tree height is 2.7 ± 0.9 m, and the crown diameter is 2.8 ± 0.8 m (minimum diameter of 0.9 m and maximum diameter of 3.4 m). An overall image of the orchard is shown in Figure 1a.

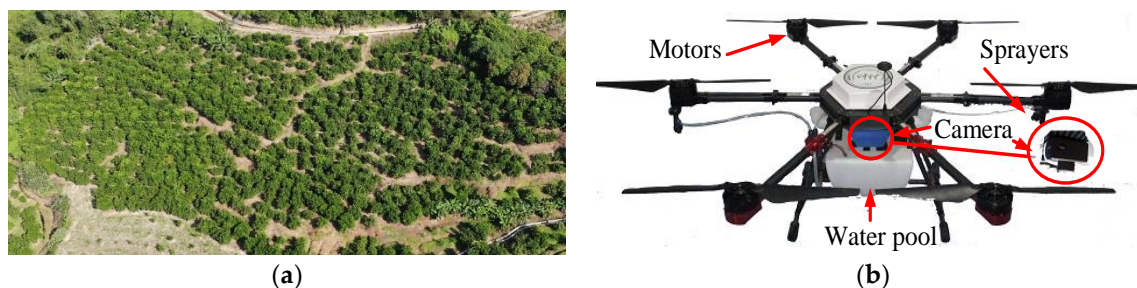


Figure 1. The general view of the orchard and the physical photo of the unmanned aerial vehicle (UAV). (a) General view of the citrus orchard captured by a UAV with an operation height of 100 m. (b) The physical photo of the agricultural UAV.

An agricultural UAV developed in the laboratory of Zhongkai University of Agriculture and Engineering (ZHKU0606-1510-CS) was used to collect the dataset of orchard images of the study area. As shown in Figure 1b, the UAV structure includes a kinetic section, an image collection section, and a spraying section. In particular, the kinetic section provides the impetus for flight movement and carrying equipment and consists of six high-speed motors, a control circuit, and a control module. The kinetic section provides a horizontal positioning accuracy of 0.5 m and a vertical positioning accuracy of 0.1 m when controlled by a global positioning system (GPS) unit and remote controller. The onboard image collection section consists of a high-resolution RGB camera (Hawkeye Firefly 8SE Action), which saves images on a secure digital memory card and is fixed by an angle fixer that reduces the high-frequency noise generated by the six high-speed motors. The spraying section, which will be used for future spraying tests, includes nozzles, a water pool, a pump, and pipes.

Considering the spraying height [32], the topographic relief of the orchard, and the height of the fruit trees, the camera distance for the UAV was set to 10–15 m. During each image acquisition process, the image collection section was set vertically downward, and the UAV was suspended steadily in the air until each image was fully collected. The data were stored as joint photographic expert group (JPEG) format to implement a debugging algorithm on a personal computer (PC). The parameters of the UAV and the onboard camera are shown in Table 1.

Table 1. The parameters of the unmanned aerial vehicle (UAV) and the onboard camera. CMOS—complementary metal-oxide-semiconductor; FOV—field of view; JPEG—joint photographic expert group.

System Parameter	Value
Vertical accuracy	±0.1 m
Camera sensor	1/2.3-inch CMOS
Image resolution	4032 × 3024
FOV	85°
Focal length	24 mm
Diaphragm size	f/2.8
Time of exposure	1/1000 s
Camera model	FC2103
Format/type of images	JPEG

2.2. The Dataset

The dataset comprised 7148 trees in 334 images (the parent dataset, dataset 0) collected in seven batches from November 2018 to March 2019. The images in the natural orchard environment were classified into six categories, namely, two BC categories, insufficient brightness (IB) and sufficient brightness (SB), and three WCC categories, namely, small weed coverage rate (SWCR), medium weed coverage rate (MWCR), and large weed coverage rate (LWCR). The WCCs were defined as follows: SWCR, $A < 35\%$; MWCR, $35\% \leq A < 60\%$; LWCR, $A \geq 60\%$, where A is the ratio of the foreground (the area was $[\pi/2, \pi]$ in H of the hue/saturation/intensity (HSI) color space) to the image area. Furthermore, IB indicates that $I_A < 0.3$, whereas SB indicates that $I_A \geq 0.3$, where I_A is the average illumination value (I of the HSI color space) of area A . Table 2 shows the number of images and the number of fruit trees in each of the six (2×3 ; (IB, SB) \times (SWCR, MWCR, LWCR)) categories of the data.

Table 2. The six categories of imagedata ¹. BC—brightness condition; WCC—weed coverage condition; IB—insufficient brightness; SB—sufficient brightness; SWCR—small weed coverage rate; MWCR—medium weed coverage rate; LWCR—large weed coverage rate.

Dataset	BC	WCC			Total
		SWCR	MWCR	LWCR	
Dataset 0	IB	64 (1344) *	29 (589)	89 (1869)	182 (3802)
	SB	57 (1351)	32 (483)	63 (1512)	152 (3346)
	Total	121 (2695)	61 (1072)	152 (3381)	334 (7148)
Training Set	IB	32 (701)	14 (222)	44 (1024)	90 (1947)
	SB	29 (641)	16 (280)	32 (780)	77 (1701)
	Total	61 (1342)	30 (502)	76 (1804)	167 (3648)
Test Set	IB	32 (643)	15 (347)	45 (845)	92 (1855)
	SB	28 (710)	16 (203)	31 (732)	75 (1645)
	Total	60(1353)	31(570)	76(1577)	167(3500)

¹ n(m); * n is the number of images in each category, and m is the total number of fruit trees in that category.

Half of the data in each category of dataset 0 were randomly selected as the training set, and the remaining images were set aside as the test set. Each training area in the training set was manually marked as rectangles of 200×200 pixels, including 1000 positive samples of fruit trees, 200 negative samples of soil or withered grass (in winter or after the application of herbicide), and 800 negative samples of weeds. Figure 2 shows some examples of images within dataset 0 under IB and SB conditions.

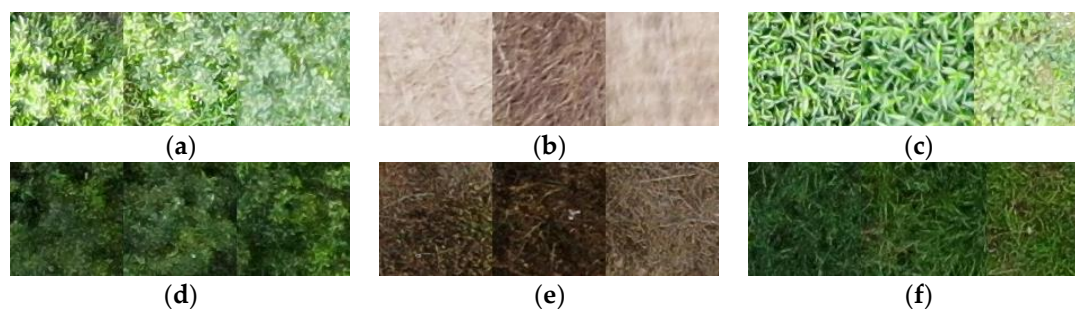


Figure 2. In sufficient brightness (IB) and sufficient brightness (SB) examples of dataset 0. (a) Positive area sample of fruit trees under SB conditions. (b) Negative area sample of soil or withered grass under SB conditions. (c) Negative area sample of weeds under SB conditions. (d) Positive area sample of fruit trees under IB conditions. (e) Negative area sample of soil or withered grass under IB conditions. (f) Negative area sample of weeds under IB conditions.

The training set and test set described in Section 2.2 were used to train and test the SVM segmentation model. The RoIs of the training set were manually extracted after the illumination compensation; the color and texture features were calculated in the completely marked areas. In addition, the RoIs of the test set were extracted with the method introduced in Sections 3.1 and 3.2; these RoIs were complete when the color features were calculated, and these RoIs were the maximum inner rectangles when the texture features were calculated.

3. The Citrus Tree Segmentation Method

Figure 3 shows the three main processes of the method proposed in this paper: image pre-processing, RoI extraction, and fruit tree segmentation. In the image pre-processing stage, the selective region intensity histogram equalization method (SRIHE) was used to reduce the sensitivity of the algorithm to the brightness. Then, to extract the RoIs from the orchard images, chromatic aberration technology was combined with the Otsu method. Finally, the color and texture features of the RoIs were calculated to establish the SVM segmentation model for the citrus

trees. The data described in Section 2.2 were used to evaluate the proposed segmentation method. All experiments were performed using MathWorks MATLAB R2018a software on a PC equipped with an Intel(R) Pentium(R) G4600 (3.60 GHz) central processing unit (CPU) and 16 GB of random-access memory (RAM).

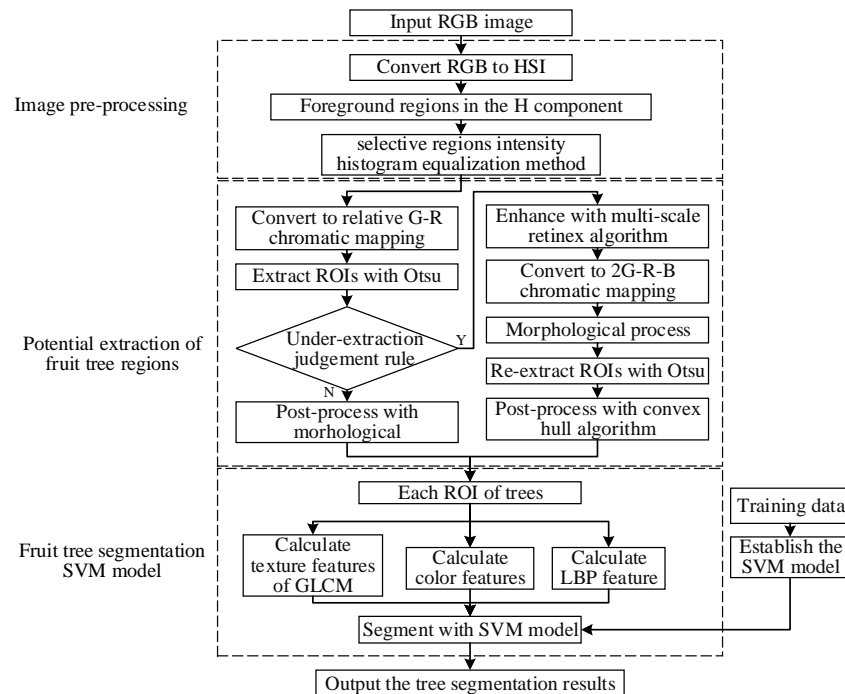


Figure 3. Flow chart of the proposed segmentation method.

3.1. Image Pre-Processing

In the orchard images collected by the agricultural UAV, an insufficiently illuminated environment would affect the brightness contrast of fruit trees and reduce the extraction effect of the segmentation method. Therefore, the SRIHE was proposed to pre-process the images, thereby compensating for the brightness contrast of the foreground. The SRIHE calculates the brightness histogram (the I component in the HSI color space) of the image foreground, the areas of which are selected in the H component. The illumination contrast of the foreground is then highlighted and changed to the H and S components of the HSI color space, which is helpful for the subsequent extraction of ROIs and segmentation of fruit trees. Compared to the traditional histogram equalization method in illumination (HE), the SRIHE selects the foreground to compensate for the brightness contrast even when the foreground is in darkness, while the HE takes the global histogram for the adjustment and potentially lacks a successful effect, as it is affected by the background [33], which is consistent with the findings of Tan et al. [34] and Kim et al. [35]

As inputs to the SRIHE, the images were firstly transformed from RGB color space to HSI color space to obtain three independent components: H, S, and I. The range of each pixel in the H component was $[0, 2\pi]$, reflecting the color of each pixel perceived by the human eye, where the range $[\pi/2, \pi]$ denotes the color green. The green area was marked, and the set of pixels in the green range of H was denoted as S_g , which was expressed as $S_g = \{(s, y) | \pi/2 \leq H(x, y) \leq \pi\}$. The histogram of I_g was calculated and equalized to I_g' , where S_g is the I component in HSI color space. The pixels outside of S_g remained and were recombined with I_g' to form a new brightness mapping I' . Then, I' and the original H and S mapping were compiled to form new HSI' images, which were then transformed into new RGB images.

3.2. Potential Fruit Tree Region Extraction

To extract RoIs from different WCC backgrounds, the RG chromatic extraction method (ERGCM) was proposed to preliminarily extract RoIs from orchard images. Because the ERGCM is strongly sensitive to weeds and under-extracts RoIs from a weedy background, an under-extraction judgement rule (UEJR) was established to judge whether the preliminary result was under-extracted. If so, the under-extracted images would be re-extracted using the chromatic mapping method presented herein, namely, the EMSRCM, which was proposed based on the combination of multi-scale retinex (MSR) and chromatic technology.

3.2.1. RG Chromatic Tree Extraction Method

Since the trees and background in SWCR images have distinguishable color characteristics, the proposed ERGCM is an appropriate technology for extracting RoIs, as shown in Figure 4. A screenshot of a single tree in an SWCR image is shown in Figure 4a,b, the R curve (the red component in RGB color space), G curve (the green component), G–R curve (the traditional G–R chromatic value), and im_p curve (the relative G–R chromatic value) of each pixel of the marked red line in Figure 4a. Comparing regions I and II (region I is the fruit tree area, and region II is the background) of the R and G curves in Figure 4b, the color differences in the R and G values between the trees and background can be quantified, as R is larger than G in II and smaller in I. Therefore, the G–R value is larger in region I than in region II, confirming that chromatic mapping is preferable for the extraction of trees in SWCR images. However, in region III, the G–R value in region III is smaller than that in region I; region III exhibits a weak local brightness, while region I highlights one tree. This might be because the traditional G–R chromatic value is the absolute difference between G and R and is, thus, affected by the local brightness. To reduce the effect of the local brightness, im_p , which is relative to the pixel brightness I (I component in HSI color space), was introduced. The transformation process is expressed by Equation (1).

$$im_p = \begin{cases} G_p - R_p / I_p, & I_p > 0 \\ 0, & I_p = 0 \end{cases} \quad (1)$$

where p is a single pixel in an image, R_p is the R component of pixel p after the SRIHE processing step, G_p is the G component of pixel p , and I_p is the I component of pixel p in HSI color space. Comparing the G–R and im_p curves in region III (the curve sections within the two red rectangles in Figure 4b), the G–R value is severely decreased, which might reduce the extraction and segmentation accuracy, but the im_p value is still acceptable. This is because im_p takes the local brightness into account; in addition, the RG chromatic mapping is measured relative to the brightness of the pixels.

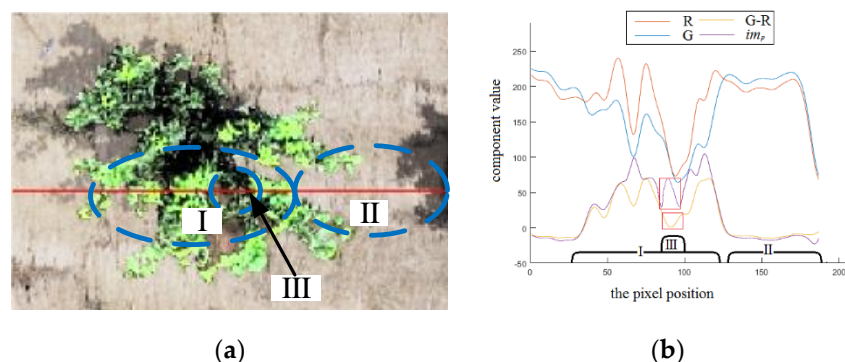


Figure 4. R (the red component in RGB color space), G (the green component), G–R (the traditional G–R chromatic value), and im_p (the relative G–R chromatic value) curves of a tree in a small weed coverage rate (SWCR) image. (a) Example of the I_a of a tree in an SWCR image. (b) R, G, R–G, and im_p curves of the line in I_a .

In im_p mapping, the illumination of a tree is relatively larger than that of the background, which is obviously different, creating a histogram with a bimodal structure. To minimize the misclassification probability of extracting trees under SWCR conditions, the Otsu method [36] was adopted to calculate the optimal threshold to extract the RoIs (fruit trees) by maximizing the between-cluster variance. Subsequently, to remove areas of small interference outside, inside, and around (at the edges) the RoIs and to improve the RoI extraction accuracy, the morphological method was adopted. Firstly, the small disturbances outside the RoIs might be caused by an impure influence within the background, especially for some small areas of green disturbance, such as weeds. These interferences usually cover small areas and can, thus, be processed by an area exclusion method with a threshold of 0.05% of the image, while the minimum area of trees (with a diameter of 0.9 m, as introduced in Section 2.2) in the image of dataset 0 is 0.10%. Secondly, some jagged interferences along the edges of the RoIs might be caused by unclear boundaries resulting from shaking of the camera onboard the agricultural UAV; these jagged boundaries can be removed by expansion and erosion with a circular structure. Although a morphological treatment with a small radius could remove only small interference, the radius of expansion and erosion was set to five pixels in this paper; this radius is considered reasonable because excessive erosion and expansion would change the edges and contours, reducing the RoI extraction accuracy. Finally, the interference due to holes inside the RoIs, which could be caused by color singularities inside the fruit trees, were filled by a hole-filling algorithm.

3.2.2. Chromatic Mapping Extraction Method Enhanced with MSR

Since the ERGCM considers only the im_p of an image, this method is strongly sensitive to weeds. For subsequent processing, a UEJR was proposed based on the area and number of RoIs to determine whether the primary RoI result is under-extracted. When the results of the ERGCM in a weedy environment (such as environments exhibiting MWCR or LWCR conditions) are under-extracted, the weeds between RoIs might be erroneously extracted, as shown in Figure 10e,h. Following such under-extraction, independent fruit trees can be connected, significantly increasing the area of RoIs and significantly decreasing the number of independent RoIs. Let A denote the ratio of the total RoI area to the image area, N denote the number of individual RoIs in the image, A_t denote the true ratio of the RoI area to the image area, and N_t denote the true number of trees in an image. The UEJR could be defined based on the relationships between these factors; when $A > T_A$ or $N < T_N$, the preliminary result is under-extracted, where T_A is the threshold of the RoI area and T_N is the threshold of the RoI number.

To re-extract the under-extracted images, EMSRCM based on MSR technology was proposed to reduce the influence of weeds, especially under MWCR and LWCR conditions. The MSR method was used to separate the reflection image (M) from the source image (S) by transforming the RGB image into a frequency mapping; this approach is typically used to enhance the foreground in images under darkness or fog [37,38]. Therefore, the multi-scale reflection sub-image weighted fusion method proposed by Rahman et al. and Ojala et al. [39,40] was adopted to separate M from S to further extract RoIs from M through chromatic aberration technology. The algorithm process is as follows: The R, G, and B color channels of images are processed with three separate scales, and the high-frequency area of fruit trees is highlighted to improve the separability between fruit trees and weeds. This algorithm can be expressed as Equation (2).

$$R_i(x, y) = \sum_{j=1}^3 W_j \{ \log S_j(x, y) - \log [S_i(x, y) \otimes G_j(x, y)] \}, \quad (2)$$

where $S_i(x, y)$ is the i -th channel of the image, in which $i = 1, 2, 3$ correspond to the R, G, and B color channels of the image, respectively, W_j is the scale weight of each channel, $W_1 = W_2 = W_3 = 1/3$, \otimes denotes a convolution operation, and $G_j(x, y)$ represents Gaussian kernel functions on different scales. Finally, the result of the i -th channel is $M_i(x, y) = \exp[R_i(x, y)]$. Figure 5 shows the enhanced effect of MSR on the trees against a background containing weeds. Figure 5a presents the S image

of the partial example under LWCR conditions, Figure 5b presents the M image of S , and Figure 5c shows the statistical results of the tree and weed areas in the S and M images, including the G–R, G–B, and 2G–R–B values. The G–R, G–B, and 2G–R–B chromatics of the fruit trees in S are all close to those of the weed background, making it difficult to extract the trees directly. However, the G–R and G–B chromatics of the fruit trees in M are more separable than those in S , enhancing the separability of the trees in the 2G–R–B (the superposition of G–R and G–B) chromatic and improving the ROI extraction accuracy.

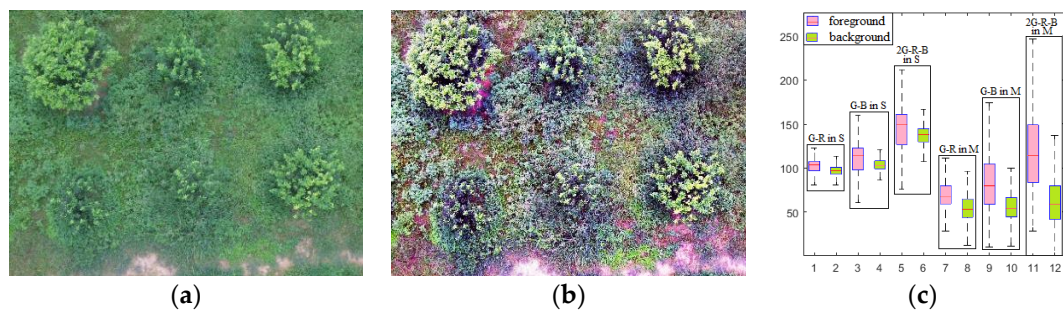


Figure 5. Reflection (M) and source (S) images of an example image under large weed coverage rate (LWCR) conditions and their chromatic results. (a) An example image under LWCR conditions of S mapping S_1 . (b) M_1 of S_1 , processed by multi-scale retinex (MSR). (c) The G–R, G–B, and 2G–R–B values of the M_1 of S_1 .

Based on these conclusions, the first step of the EMSRCM is to separate M from S and to transform M to a 2G–R–B mapping, thereby enhancing the trees in the weedy background. Then, a large-radius closed filter operation is applied to the image to remove small interference areas around the trees potentially attributable to some weeds in the background that have similar frequency characteristics as the citrus trees in the image. In addition, a top-hat filter is used to eliminate some interference at different positions in the local background of the image possibly caused by differences in the frequency among different background regions exposed by MSR. Then, the Otsu method and an open filter are applied to the image to extract the RoIs from the weedy background. As a large-radius morphological treatment, such as an open filter or a closed filter, is inevitably used to reduce the weedy background, the edge information of the trees is important. To minimize the loss of information during the application of the large-radius morphological method intended to reduce the influence of weeds, a convex hull transform algorithm for edge convex filling is used to improve the extraction accuracy. The entire EMSRCM process is described in Table 3.

Table 3. The pseudo-code of the EMSRCM (combination of multi-scale retinex (MSR) and chromatic technology). UEJR—under-extraction judgment rule; RGB—red, green, and blue; M—reflection.

Input:	The Under-Extracted Image I_u Judged with the UEJR
Step 1	Separate the R, G, and B channels of I_u and calculate R_i by Equation (2) and M images (M_u).
Step 2	Convert M_u into the 2R-G-B chromatic mapping I_c .
Step 3	Apply a closed filter I_c with a 20-pixel radius to obtain I_{cc} .
Step 4	Apply a top-hat filter I_{cc} with a 24-pixel radius to obtain I_{cct} .
Step 5	Apply an open filter I_{cct} with a 20-pixel radius to obtain I_{ccto} .
Step 6	Extract I_{cct} with the Otsu method to obtain the binary image I_{bw} .
Step 7	Use a convex hull transform to convert I_{bw} into I_{cn-bw} .
Step 8	Exclude small areas of interference in I_{cn-bw} with a 0.05% threshold and generate I_{roi} .
Output	The extracted foreground areas of the image, I_{roi} .

3.3. Fruit Tree SVM Segmentation Model

To segment the RoIs accurately, an SVM segmentation model was established by calculating the color features and texture features of the RoIs. The color features of the RoIs include each component

of the RGB, his, and La*b* color spaces, excluding the I component in the HSI color space and the L component in the La*b* color space, which represent the image brightness and are easily affected by the environment after the images are pre-processed. Therefore, the color features of the RoIs comprise the averages and variances of the R, G, B, H, S, a*, and B* components and constitute 14 dimensions in total.

For the texture features of the RoIs, the statistical texture features of the GLCM and the LBP features were adopted to reflect the regional features of the tree structure. Firstly, the GLCMs of the RoIs were calculated as $G(i, j)$, and the five texture statistics of $G(i, j)$ were calculated. Firstly, the contrast was calculated as $CON = \sum_{n=0}^{k-1} n^2 \sum_{i-j=n} G(i, j)$, where $k=16$ indicates the gray scale of the image. CON measures the local change and the image matrix distribution and reflects the image clarity and texture grooving depth. A deeper grooving depth generates a greater value of CON and of the texture. Secondly, the energy was calculated as $ASM = \sum_{i=0}^k \sum_{j=0}^k (G(i, j)^2)$, which reflects the uniformity coefficient of the gray-level distribution and the texture thickness; a thicker texture generates a smaller energy. Thirdly, the entropy was calculated as $ENT = - \sum_{i=0}^k \sum_{j=0}^k G(i, j) \lg G(i, j)$, which reflects the complexity of the image gray-level distribution; a more complex image generates a greater entropy. Fourthly, the inverse variance was calculated as $IDM = \sum_{i=0}^k \sum_{j=0}^k \frac{G(i, j)}{1+(i-j)^2}$, which reflects the local change in the image texture; a smaller change generates a larger IDM . Lastly, the correlation was calculated as $COR = \sum_{i=0}^n \sum_{j=0}^n \frac{(i, j)G(i, j) - u_i v_j}{s_i s_j}$, where $u_i = \sum_{i=0}^k \sum_{j=0}^k i G(i, j)$ and $v_j = \sum_{i=0}^k \sum_{j=0}^k j G(i, j)$, which reflects the local gray-level correlation; a stronger correlation generates a greater value of COR . In addition, the LBP texture feature possesses 59-dimensional distribution characteristics that can be calculated by the method introduced in Reference [41] and has the advantage of grayscale invariance.

A linear kernel-based SVM segmentation model was established based on the above 80-dimensional features, and the effect was evaluated using the following eight indicators: (1) the intersection over union (IoU), $IoU = \frac{n(r) \cap n(h)}{n(r) \cup n(h)} \times 100\%$, where $n(r)$ is the area of tree output by the segmentation algorithm and $n(h)$ is the manually segmented area, which represents the positioning accuracy of the image segmentation algorithm; (2) the precision of segmentation pixels (P_I), $P_I = TP_I / (TP_I + FP_I)$, which represents the correctness of the segmentation, where TP_I is the true area of segmented trees, FP_I is the area in which fruit trees are erroneously segmented as background, and FN_I is the area where the background is erroneously segmented as fruit trees; (3) the recall rate (R_I), $R_I = TP_I / (TP_I + FN_I)$, which indicates the correctness of the segmentation of fruit trees; (4) the F1-score ($F1_I$), $F1_I = 2 \times P_I \times R_I / (P_I + R_I)$, representing the general segmentation effect; (5) the correct segmentation rate of fruit trees (CTR) [12], which represents the segmentation accuracy of the number of trees and is used to represent the final segmentation effect in this paper; (6) the tree segmentation precision (P_C), $P_C = TP_C / (TP_C + FP_C)$, which represents the correctness of the tree segmentation accuracy, where TP_C is the true number of segmented trees, FP_C is the erroneously segmented number of trees when the trees are segmented into the background, and FN_C is the erroneously segmented number of trees when the background is segmented into the foreground; (7) the recall rate of fruit tree segmentation (R_C), $R_C = TP_C / (TP_C + FN_C)$, which represents the effect of truly segmenting the fruit trees; (8) the F1-score of trees ($F1_C$), $F1_C = 2 \times P_C \times R_C / (P_C + R_C)$, which represents the comprehensive effect of segmenting the trees.

4. Results and Discussion

4.1. Evaluation of the Influence of the Image Brightness Condition on the Segmentation Results

Figure 6 shows two example images under IB conditions and the pre-processing results of these images by the HE and SRIHE. Figure 6a shows an image with weak brightness in the foreground and background, Figure 6b shows the processed result of Figure 6a using HE, and Figure 6c shows the processed result of Figure 6a using SRIHE. Figure 6c is superior to Figure 6b, because the SRIHE takes the foreground region for illumination compensation, thereby avoiding any adjustment of the background, while the HE takes the global histogram for the adjustment. In addition, when the brightness of the foreground is lower than that of the background in an image, as shown in Figure 6d, the adjustment

performed with the HE might be insufficient to effectively compensate for the foreground brightness when the background and foreground are adjusted at the same time, as displayed in Figure 6e. However, the adjustment process of the SRIHE is more relevant to the foreground, as potential foreground areas are selected for compensation; thus, the effect is better than the effect of the HE, as shown in Figure 6f. Therefore, pre-processing with the SRIHE reduces the difference in tree illumination among images to some extent and reduces the sensitivity of the algorithm to the environmental BC.

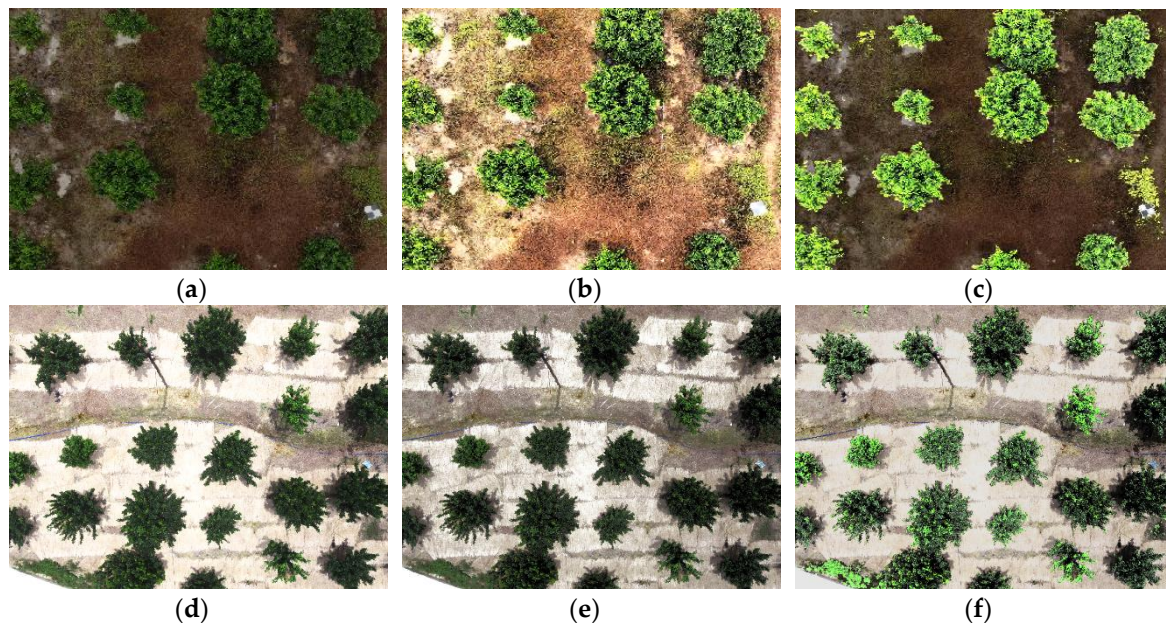


Figure 6. Results of two illumination compensation examples. (a) An example under IB conditions, I_1 . (b) I_1 after applying the histogram equalization (HE). (c) I_1 after applying the selective region intensity histogram equalization (SRIHE). (d) An example under IB conditions, I_2 . (e) I_2 after applying the HE. (f) I_2 after applying the SRIHE.

To evaluate the effect of applying the SRIHE, the images in the test set introduced in Section 3.2.2 were processed with the SRIHE, and their RGB components were obtained, as shown in the box plot in Figure 7. The results demonstrate that the brightness differences among the trees in different images were reduced; for example, the average values of the R components were 38 under IB and SWCR conditions and 115 under SB and SWCR conditions, but these values changed to 128 and 145, respectively, after applying the SRIHE, with the differences reduced by 60. The statistical results of other channels were also analyzed, and the standard deviation of the trees in each image band was reduced from 41 to 21. Therefore, we assert that the SRIHE reduces the difference among each component between different trees in images and makes trees more consistent under different BCs, which ensures the accuracy of the ROI extraction under different illumination environments.

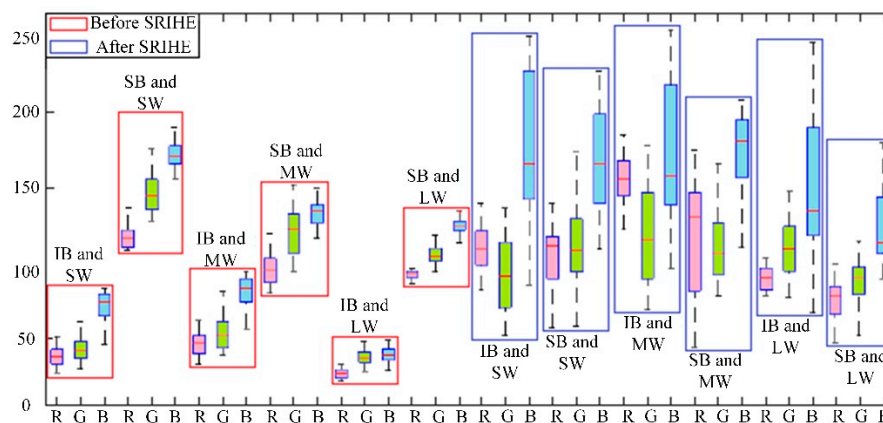


Figure 7. Data representation of the fruit tree areas in RGB color space before and after pre-treatment (SW, MW, and LW mean SWCR, medium weed coverage rate (MWCR), and LWCR conditions, respectively).

A comparison of the extraction results of images under different BCs reveals that the tree segmentation method proposed in this paper is not sensitive to image brightness, as shown in Figure 11a,d. The SRIHE compensates for the illumination and reduces the variation in the influence on different images, thereby improving the extraction and segmentation accuracy. In addition, the relative chromatic aberration mapping im_p used in the ERGCM and the MSR technology used in the EMSRCM are insensitive to the local brightness of images; these results are similar to the findings of Guo et al. [41] and Kyung et al. [42].

4.2. Evaluation of the Influence of the Weed Coverage Condition on the Segmentation

In this paper, the ERGCM and EMSRCM were combined to extract RoIs, and the results show that the combination of these two methods is insensitive to natural weedy environments. The ERGCM was used to preliminarily extract the RoIs, which are suitable for the SWCR conditions of the test set, and Figure 8 shows the results. Figure 8a shows an example image in RGB color space under SWCR conditions. Figure 8b shows the transformed result mapping of Figure 8a using the relative G–R chromatic aberration mapping, which results in an obvious difference between the foreground and background. Figure 8c shows the separated result Figure 8b using the Otsu method, where the RoIs are completely separated from the image. Figure 8d shows the ERGCM result after the morphological post-treatment, in which erosion, dilation, and hole-filling operations were used to reduce the interference regions corresponding to minor misclassifications. The results demonstrate that the ERGCM is highly precise for RoI extraction and accurately extracts the edges of trees even in an SWCR background.

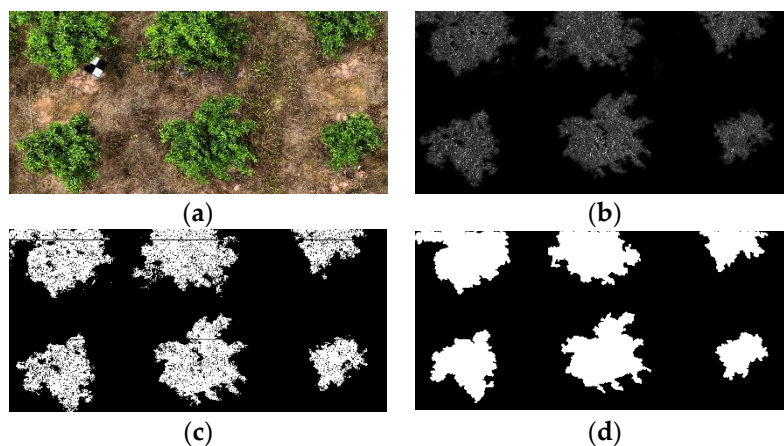


Figure 8. Extracted result of fruit trees in SWCR images using the RG chromatic extraction method (ERGCM). (a) An example SWCR image in RGB color space. (b) The transformed result of (a) using relative chromatic mapping. (c) Separated result of (b) using the Otsu method. (d) The result of (c) using morphological treatments.

Because of the sensitivity of ERGCM to weeds, the UEJR was proposed based on the area and number of RoIs; thus, the values of A , N , A_t , and N_t in the test set were calculated. A_t , N_t , A , and N in the test set were extracted by the ERGCM, and A_t and N_t were obtained using a manual method (the manual results are regarded as a reference). The results are shown in Table 4. When comparing A and N under the same BC and different WCCs, A increased significantly, and N decreased significantly with an increasing weed coverage rate. For example, A decreased from 39.1 ± 10.2 to 1.3 ± 0.6 and A increased from 29 ± 8.5 to 76 ± 19.1 under IB conditions because the ERGCM is sensitive to background weed coverage. However, the true values of the area and number of RoIs (A_t and N_t) did not change significantly. For example, the values of A_t were 30 ± 6.5 , 31 ± 8.5 , and 32 ± 5.8 under IB conditions with SWCR, MWCR, and LWCR backgrounds, respectively, and the corresponding values of N_t were 37.7 ± 8.5 , 41.7 ± 14.4 , and 32.4 ± 7.1 , suggesting that the true area and number of fruit trees did not change substantially. Based on this conclusion, the UEJR could be defined as follows: when $A > T_A = 45 \approx 34 + 11.5$ (%) or $N < T_N = 20 \approx 27.8 - 7.7$ (tree) for the RoIs in the image, the preliminary result is under-extracted.

Table 4. The values of A , N , A_t , and N_t for the test set.

BC and WCC		A^a (%)	A_t^b (%)	N^c (Trees)	N_t^d (Trees)
IB	SWCR	29 ± 8.5	30 ± 6.5	39.1 ± 10.2	37.7 ± 8.5
	MWCR	45 ± 9.1	31 ± 8.5	12.1 ± 4.4	41.7 ± 14.4
	LWCR	76 ± 19.1	32 ± 5.8	1.3 ± 0.6	32.4 ± 7.1
SB	SWCR	32 ± 7.2	31 ± 5.2	27.5 ± 17.3	25.5 ± 12.1
	MWCR	41 ± 10.9	36 ± 6.9	16.5 ± 2.7	30.1 ± 4.7
	LWCR	81 ± 17.3	38 ± 9.2	1.5 ± 0.4	35.3 ± 10.6
Average		57 ± 21.1	34 ± 11.5	17 ± 10.5	27.8 ± 7.7

^a A is the ratio of the total region of interest (RoI) area to the image area; ^b A_t is the true ratio of the RoI area to the image area; ^c N is the number of individual RoIs in the image. ^d N_t is the true number of trees in the image.

In addition, Figure 9 shows the RoI re-extraction result with the EMSRCM, which is suitable for the under-extracted images of the ERGCM in the MWCR and LWCR images of the test set. Figure 9a shows an example image in an MWCR environment, and Figure 9b shows Figure 9a transformed to the 2G–R–B chromatic mapping after being enhanced with MSR technology. Figure 9c shows Figure 9b filtered by an open filter and a top-hat filter, revealing that interference due to the weed background is clearly decreased and that the tree area is evidently increased. Figure 9d displays Figure 9c separated by the Otsu method, which processes most of the weeds in the extracted RoIs. Figure 9e shows

Figure 9d post-processed by applying the morphological treatment, which is used to exclude areas of interference. Finally, Figure 9f shows Figure 9e after applying the convex hull transform, which reduces the loss of information along the edges of the image. Therefore, the EMSRCM is an effective method for re-extracting RoIs in under-extracted images, such as the images under MWCR and LWCR conditions in the test set.

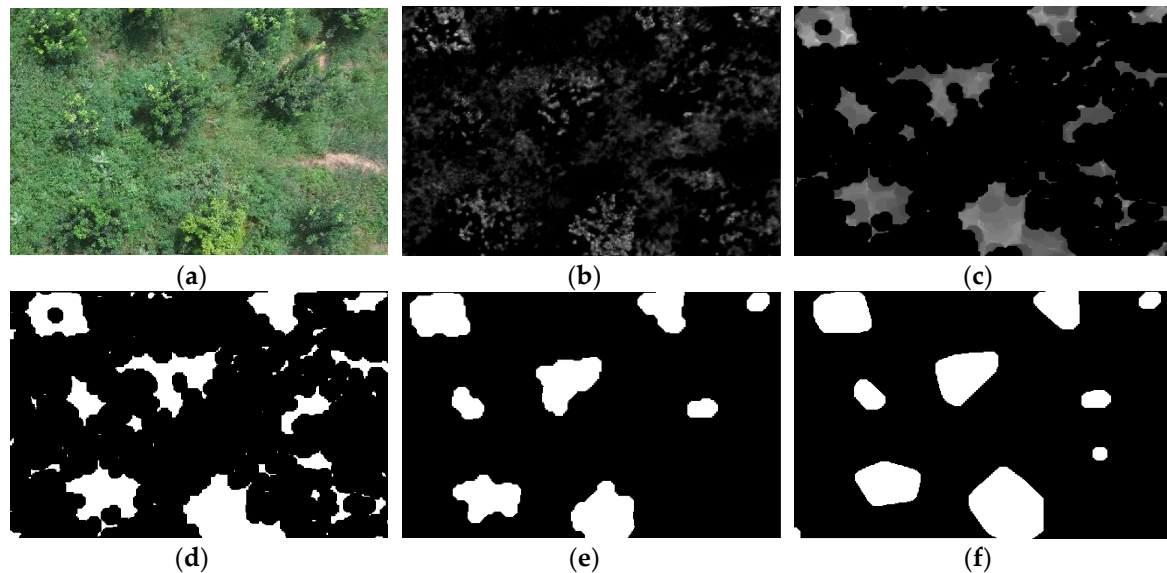


Figure 9. Re-extracted results of the EMSRCM. (a) An example image in RGB color space. (b) Grayscale images transformed by MSR and 2G–R–B chromatic mapping on (a). (c) Filtered result obtained by applying an open filter and a top-hat filter to (b). (d) Separated results obtained by applying the Otsu method to (c). (e) Post-processed result of (d) using morphological treatments. (f) Transformed result of (e) using a convex hull method.

Figure 10 shows the RoIs extracted with the ERGCM and EMSRCM under different WCCs (including SWCR, MWCR, and LWCR conditions), which illustrate the extraction effects. Figure 10b,c show the results of the ERGCM and EMSRCM under SWCR conditions, respectively; ERGCM produces better results, as this method achieves a higher accuracy along the edges of the extracted trees. This may be because the large-radius morphology used in the EMSRCM changes the edge information in the RoIs that even the convex hull transform algorithm cannot repair. Figure 10e,f,h,i show the results of the ERGCM and EMSRCM for trees under MWCR and LWCR conditions; we found that the EMSRCM results are acceptable, whereas the ERGCM results are under-segmented in the weedy environment. In conclusion, the ERGCM and EMSRCM have unique advantages and disadvantages; for example, the ERGCM boasts high accuracy but is sensitive to the WCC, while the EMSRCM is insensitive to the presence of weeds but forfeits edge information (but is more accurate than the ERGCM). To optimize the RoI extraction effect, the proposed UEJR was used to combine the ERGCM and EMSRCM; accordingly, some images whose RoI results were judged as being under-extracted were re-extracted by the EMSRCM. Therefore, the whole extraction method is highly accurate under SWCR conditions and insensitive to weeds because the ERGCM avoids losing edge information and the EMSRCM is insensitive to weedy environments.

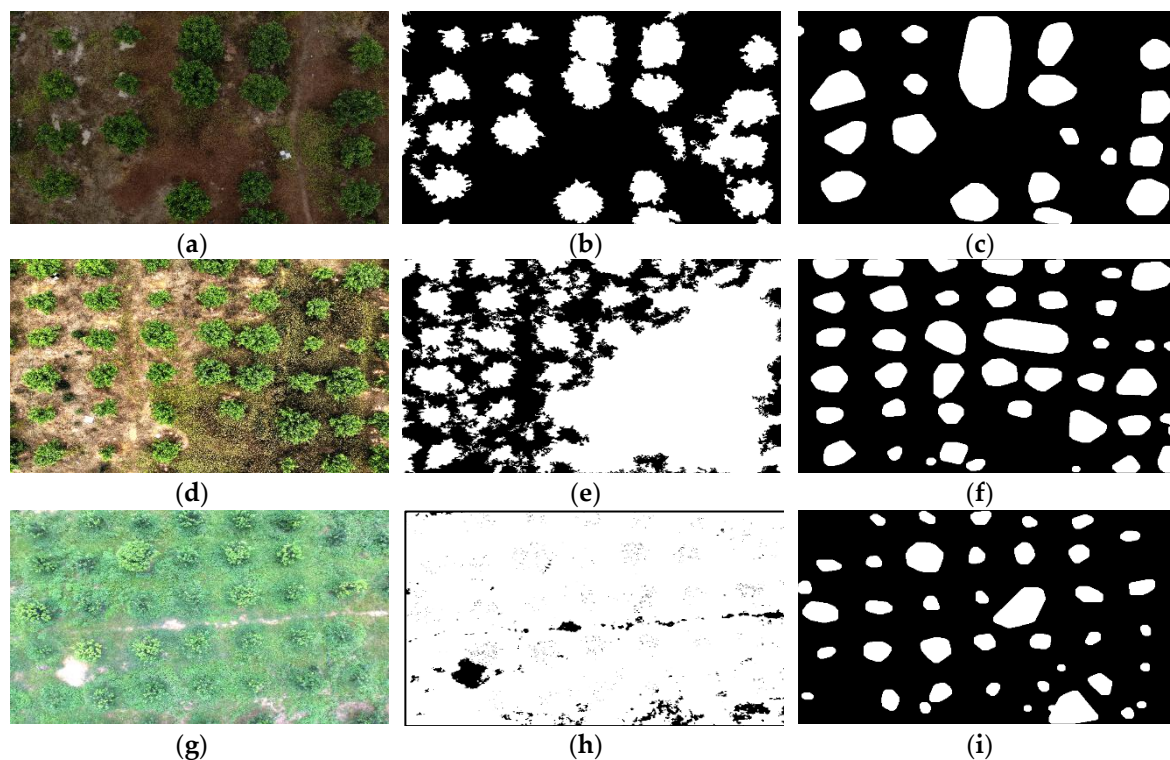


Figure 10. ERGCM and EMSRCM results under different WCCs. (a) An example image of SWCR I_a . (b) I_a processed by the ERGCM. (c) I_a processed by the EMSRCM. (d) An example image of MWCR I_b . (e) I_b processed by the ERGCM. (f) I_b processed by the EMSRCM. (g) An example image of LWCR I_c . (h) I_c processed by the ERGCM. (i) I_c processed by the EMSRCM.

4.3. Evaluation of Fruit Tree Segmentation Results

To evaluate the fruit tree segmentation results, the training set introduced in Section 2.2 was used to train the SVM segmentation model; then, the effect of the model was tested with a test set. All the fruit trees of the test set were segmented automatically by the algorithm, and the trees were artificially marked as a reference area (the true area) to calculate the model accuracy. The evaluation of the tree segmentation accuracy included the IoU , which represents the correctly segmented area, and CTR , which represents the correct number of trees. The P_I , R_I , and $F1_I$ parameters of the IoU the P_C , R_C , and $F1_C$ parameters describing the CTR were also calculated. Table 5 shows the calculated results with and without the SVM model for the test set.

Table 5. Statistical table of segmentation results.

Method	BC	WCC	IoU^a (%)	P_I^b (%)	R_I^c (%)	FI_I^d (%)	CTR^e (%)	P_C^f (%)	R_C^g (%)	FI_C^h (%)	
Without SVM model	IB	S	86.02	93.28	91.71	92.49	90.97	97.47	93.17	95.27	
		M	71.61	86.55	80.58	83.46	78.38	93.89	82.59	87.88	
		L	64.89	86.52	72.19	78.71	75.92	94.22	79.63	86.31	
	Average		72.52 ± 15.34	88.90 ± 5.51	79.74 ± 13.99	84.07 ± 9.96	81.16 ± 11.42	95.32 ± 2.80	84.53 ± 10.08	89.61 ± 6.77	
	SB	S	88.19	92.98	94.47	93.72	91.66	97.78	93.60	95.65	
		M	73.78	87.50	82.47	84.91	79.88	91.93	85.91	88.82	
		L	68.99	87.30	75.46	80.95	77.27	92.63	82.33	87.18	
	Average		76.01 ± 14.14	89.47 ± 4.56	83.47 ± 13.69	86.37 ± 9.28	83.09 ± 10.84	94.61 ± 4.53	87.23 ± 8.18	90.77 ± 6.36	
	With SVM model	IB	S	90.64	92.34	98.01	95.09	92.94	95.02	97.70	97.02
			M	77.17	84.49	89.92	87.11	80.71	93.04	85.893	87.982
L			75.28	84.16	87.70	85.89	77.60	94.22	81.48	85.07	
Average			80.68 ± 11.85	87.09 ± 6.55	91.64 ± 7.68	89.31 ± 7.07	82.97 ± 11.48	94.32 ± 1.46	87.3 ± 11.89	90.10 ± 8.83	
SB		S	90.78	91.58	99.05	95.17	93.60	96.37	96.34	96.70	
		M	78.32	84.69	91.25	87.84	81.02	91.09	89.32	89.51	
		L	76.39	87.25	85.94	86.61	79.19	91.96	87.39	88.38	
Average			81.85 ± 11.06	88.36 ± 5.00	91.73 ± 9.41	90.01 ± 6.55	84.89 ± 9.43	93.62 ± 4.02	90.69 ± 6.67	91.83 ± 6.38	

^a IoU is the intersection over union, which represents the positioning accuracy of the image segmentation algorithm;

^b P_I is the precision of segmentation pixels, which represents the correctness of the segmentation; ^c R_I is the recall rate of segmentation; ^d FI_I is the F1-score, representing the general segmentation effect; ^e CTR is the correct segmentation rate of fruit trees and the final segmentation effect in this paper; ^f P_C is the tree segmentation precision; ^g R_C is the recall rate of fruit tree segmentation; ^h FI_C is the F1-score of tree segmentation.

Comparing the results under IB and SWCR conditions, the IoU with and without the SVM model was 90.64% and 86.02%, respectively, with a 4.62% IoU increase (the maximum increase in IoU was 10.39% under IB and LWCR conditions). Similarly, to analyze the accuracy under MWCR and LWCR conditions, we know that segmentation with the SVM model can improve the segmentation accuracy. Color and texture features were used to establish the SVM model, thereby reducing the segmentation errors and improving the segmentation accuracy. However, the SVM model reduced the R_I slightly and increased P_I ; for example, P_I and R_I under IB and SWCR conditions changed from 93.73% to 92.34% and from 91.71% to 98.01%, respectively, with an increase of 6.30% in P_I and a reduction of 1.39% in R_I (the minimum reduction in P_I was 0.05% under SB and LWCR conditions, and the maximum increase in R_I was 15.51% under IB and LWCR conditions). This might be because the SVM model confirms each RoI one by one and effectively excludes the erroneously segmented regions, improving P_I considerably. However, the SVM model did not improve R_I ; rather, it reduced R_I slightly when the SVM was affected by the BCs and WCCs. In summary, although the SVM model slightly reduced R_I , FI_I was still enhanced when P_I was obviously improved. Furthermore, when comparing the results under different WCCs, the CTR was also reduced; for example, the CTR values were 90.97%, 78.38%, and 75.92% under IB and SWCR, MWCR, and LWCR conditions, respectively, without the SVM model and were changed to 92.94%, 80.71%, and 77.60%, respectively, with the SVM model. The presence of weeds in the background might affect the ERGCM and EMSRCM in the segmentation method, thereby reducing the CTR .

To illustrate the final fruit tree segmentation results, Figure 11 shows examples under each condition, including different BCs and WCCs. The red circles represent correctly segmented trees in the images, where each circle center is the center of the RoI, and its area is equal to that of the segmented tree. In contrast, the blue circles denote erroneously segmented regions, and the green squares indicate missing segmented trees. Figure 9 shows that the proposed method is insensitive to the environmental brightness and the presence of weeds, as the same WCC results under IB and SB conditions verify that the illumination condition does not greatly affect the segmentation, while the results under different WCCs and the same BC show that the proposed method is applicable in weedy orchards.

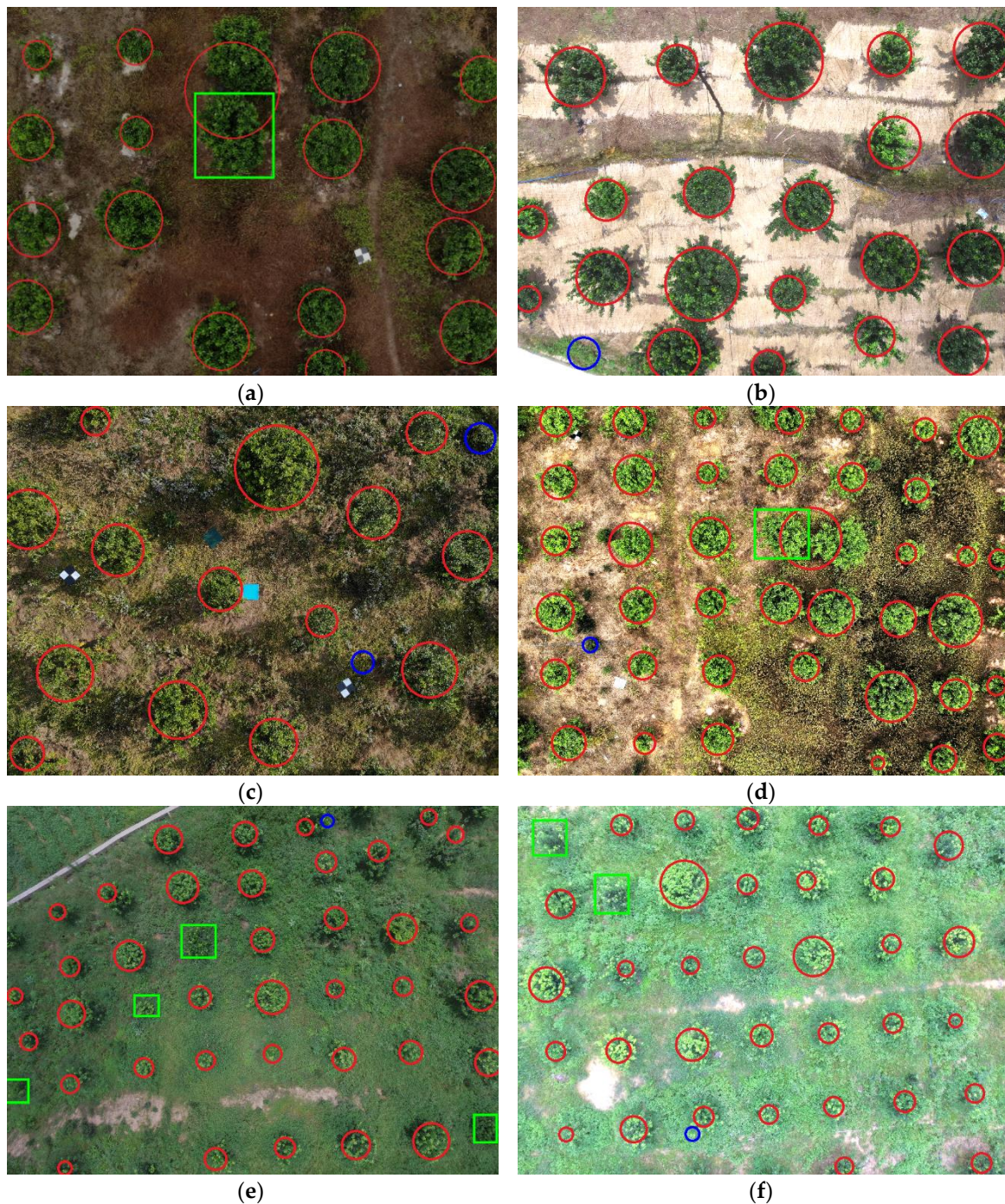


Figure 11. Citrus tree segmentation results for the agricultural UAV. (a) Segmentation results under IB and SWCR conditions. (b) Segmentation results under SB and SWCR conditions. (c) Segmentation results under IB and MWCR conditions. (d) Segmentation results under SB and MWCR conditions. (e) Segmentation results under IB and LWCR conditions. (f) Segmentation results under SB and LWCR conditions. The red circles represent correct results in the images, blue circles denote erroneously segmented regions, and green squares indicate missing trees.

Although the proposed method is weakly affected by different WCCs and BCs, there are still a few errors, such as mis-segmentations and erroneous segmentations. For example, the green rectangles of trees in Figure 11a,d are mis-segmented, indicating that a tree overlaps with another correctly segmented tree; this condition might result in a mis-segmented image. Here, the segmentation method did not divide any trees, which will be studied in the future. In Figure 11e,f, some trees are

mis-segmented under LWCR conditions, possibly because the large-radius morphology treatment changed the edges considerably; thus, the small tree area affected the SVM segmentation model result. The rectangles in Figure 11b,c,f delineate erroneously segmented trees where the areas of weeds were segmented as trees, possibly because those weed areas exhibited color and texture characteristics similar to those of the trees; thus, the SVM model was affected.

The receiver operating characteristic (ROC) curve [43] was used to evaluate the segmentation performance of the proposed method, where the true positive rate (TPR) was plotted against the false positive rate (FPR) on different WCCs of the test set. As shown in Figure 12, the areas under the curve (AUCs) when using the resultant SVM classifier under different conditions (i.e., SWCR, MWCR, and LWCR) were 95.09%, 91.91%, and 85.03%, respectively. The results show that the SVM classifier under SWCR conditions is better than that under both MWCR and LWCR conditions. Because the color appearance of weeds is similar to that of citrus trees, the number of weeds would affect the segmentation performance; therefore, the proposed method obtained the best and worst results under SWCR and LWCR conditions, respectively. This result is consistent with the conclusion in Figure 11 and the orchard image features.

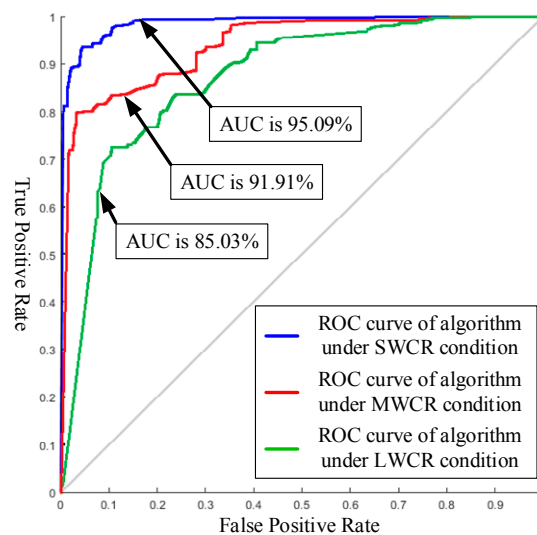


Figure 12. The receiver operating characteristic(ROC) curves of the support vector machine (SVM) under different WCCs on the test set.

The method presented by Lin et al. [25] (method 1) and that developed by Omair et al. [22] (method 2) were used for comparison; the data used in this part were images under SB conditions from the test set. Table 6 shows that the segmented results under different WCCs were accurate for the proposed method ($CTR = 85.27 \pm 9.43$), but the results of methods 1 and 2 were poor ($CTRs = 69.49 \pm 10.37$ and 64.04 ± 12.82) under MWCR and LWCR conditions. The poor result of method 1 might be due to the fact that the k-means model on the a^*b^* plane (a^* and b^* components in the La^*b^* color space) was greatly influenced, making it difficult to segment the trees under the different WCCs. In method 2, the entropy used to describe the texture was not sufficient for the weed images in the test set, and erosion was used to delete weeds around some fruit trees, leading to the removal of some small areas of trees. In conclusion, the proposed method achieved high accuracy in the natural orchard environment studied in this paper.

Table 6. The effects of different segmentation methods under different WCCs.

Method		WCC			Average
		SWCR	MWCR	LWCR	
Method 1	<i>IoU</i> (%)	84.23	67.25	51.11	66.06 ± 8.61
	<i>CTR</i> (%)	85.66	72.25	55.52	69.49 ± 10.37
Method 2	<i>IoU</i> (%)	79.81	63.18	39.24	58.31 ± 13.72
	<i>CTR</i> (%)	80.22	70.33	48.64	64.04 ± 12.82
Proposed method	<i>IoU</i> (%)	90.78	78.32	76.39	81.76 ± 11.06
	<i>CTR</i> (%)	93.60	81.02	79.19	85.27 ± 9.43

5. Conclusions

In this paper, a citrus tree segmentation method based on monocular machine vision in a natural orchard environment was proposed, and the main conclusions were as follows.

(1) After extracting the potential foreground areas in the H component (in HSI color space) with threshold extraction technology and calculating the brightness histogram (the I component in HSI color space), the SRIHE was proposed by equalizing the brightness histogram to compensate for the foreground illumination. The results show that the proposed illumination compensation method enhances the contrast of the foreground brightness and keeps the H and S unchanged, which can compensate for weak illumination and achieve a better effect than that of traditional histogram equalization.

(2) Based on this brightness compensation process, relative RG chromatic mapping and the Otsu threshold algorithm were used to extract the RoIs of citrus trees. The results on SWCR images show that the RoI extraction accuracy of trees reached 91.66%, and the edge information was preserved.

(3) The area and number of RoIs were used to judge whether the results were under-extracted. The under-extracted images were re-extracted with the MSR enhancement algorithm and chromatic technology, and the re-extraction method was tested under MWCR and LWCR conditions; the accuracies were 79.88% and 77.27%, respectively, which indicates an insensitivity to weeds.

(4) A fruit tree SVM segmentation model was established by calculating the color and texture features of the RoIs. The segmentation results of citrus trees in natural orchard images (under different BCs and WCCs) show an average accuracy of 85.27% ± 9.43%, which demonstrates that the proposed method can effectively suppress the false alarm rate.

Author Contributions: Y.C., C.H., J.Z., and Y.T. conceptualized and designed the experiments. Y.C., J.Z., J.L., and Q.G. performed the experiments. Y.C., J.Z., J.L., and Z.Z. analyzed the data and prepared the manuscript. Y.C., Y.T., H.L., Y.H., and S.L. prepared the experimental materials and contributed to the data analysis. All authors reviewed the manuscript.

Funding: The authors acknowledge support from the Project of Guangdong Province Support Plans for Top-Notch Youth Talents, China (No. 2016TQ03N704), the Planned Science and Technology Project of Guangdong Province, China (Nos. 2019A050510045 and 2019B020216001), the Planned Science and Technology Project of Guangzhou, China (Nos. 201704020076 and 201904010206), the Innovative Project for University of Guangdong Province (No. 2017KTSCX099), the Key Laboratory of Spectroscopy Sensing, Ministry of Agriculture and Rural Affairs, P.R. China (No. 2018ZJUGP001), the Guangdong Academy of Sciences Special Fund for the Building of First-Class Research Institutions in China (No. 2019GDASYL-0502007), the Special Funds for the Cultivation of Guangdong College Students' Scientific and Technological Innovation ("Climbing Program" Special Funds No. PDJH2019B0244), and the Science and Technology Innovation Fund for Graduate Students of Zhongkai University of Agricultural and Engineering (No. KJCX2019007). Thanks are extended to Conghua Hualong Fruit and Vegetable Preservation Co., LTD, for providing a commercial orange orchard for the experiment.

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

Abbreviations

The following abbreviations are used in this manuscript:

ANN	Artificial neural network
BC	Brightness condition
CNN	Convolutional neural network
CTR	Correct segmentation rate of fruit trees
EMSRM	Extraction on multi-scale retinex chromatic mapping method
ERGCM	RG chromatic extraction method
GLCM	Gray-level co-occurrence matrix
HE	Histogram equalization
HSI	Hue/saturation/intensity color space
IB	Insufficient brightness
im_p	Relative G–R chromatic value or mapping
IoU	Intersection over union
LBP	Local binary pattern
LiDar	Light detection and ranging
LWCR	Large weed coverage rate
MSR	Multi-scale retinex method
MWCR	Medium weed coverage rate
NDVI	Normalized difference vegetation index
RGB	Red/green/blue color space
RoIs	Regions of interest (potential area of fruit trees)
SB	Sufficient brightness
SfM	Structure from motion
SRIHE	Selective regions intensity histogram equalization method
SVM	Support vector machine
SWCR	Small weed coverage rate
UAV	Unmanned aerial vehicle
UEJR	Under-extraction judgement rule
WCC	Weed coverage condition

References

1. Katsigiannis, P.; Misopolinos, L.; Liakopoulos, V.; Alexandridis, T.K.; Zalidis, G. An autonomous multi-sensor UAV system for reduced-input precision agriculture applications. In Proceedings of the 2016 24th Mediterranean Conference on Control and Automation (MED), Athens, Greece, 21–24 June 2016; IEEE: Piscataway, NJ, USA, 2016; Volume 24, pp. 60–64. [\[CrossRef\]](#)
2. Becerra, V.M. Autonomous control of unmanned aerial vehicles. *Electronics* **2019**, *8*, 452. [\[CrossRef\]](#)
3. Abu, J.M.; Hossain, S.; Al-Masud, M.A.; Hasan, K.M.; Newaz, S.H.S.; Ahsan, M.S. Design and development of an autonomous agricultural drone for sowing seeds. In Proceedings of the 7th Brunei International Conference on Engineering and Technology (BICET) 2018, Bandar Seri Begawan, Brunei, 12–14 November 2018; Volume 7, pp. 1–4. [\[CrossRef\]](#)
4. Philipp, L.; Raghav, K.; Johannes, P.; Roland, S.; Cyrill, S. UAV-Based Crop and Weed Classification for Smart Farming. In Proceedings of the IEEE International Conference on Robotics and Automation (ICRA), Singapore, 29 May–3 June 2017; IEEE: Piscataway, NJ, USA, 2018; Volume 6, pp. 1–8. [\[CrossRef\]](#)
5. Dilek, K.; Serdar, S.; Nagihan, A.; Bekir, T.S. Automatic citrus tree extraction from UAV images and digital surface models using circular Hough transform. *Comput. Electron. Agric.* **2018**, *150*, 289–301. [\[CrossRef\]](#)
6. Shouyang, L.; Fred, B.; Bruno, A.; Philippe, B.; Hemmerlé, M. Estimates of plant density of wheat crops at emergence from very low altitude UAV imagery. *Remote Sens. Environ.* **2017**, *198*, 105–114. [\[CrossRef\]](#)
7. Zou, X.; Zhang, X.; Shi, J.; Li, Z.; Shen, T. Detection of chlorophyll content and distribution in citrus orchards based on low-altitude remote sensing and bio-sensors. *Int. J. Agric. Biol. Eng.* **2018**, *11*, 164–169. [\[CrossRef\]](#)
8. Zhuang, J.; Hou, C.; Tang, Y.; He, Y.; Guo, Q.; Zhong, Z.; Luo, S. Computer vision-based localisation of picking points for automatic litchi harvesting applications towards natural scenarios. *Biosyst. Eng.* **2019**, *187*, 1–20. [\[CrossRef\]](#)
9. Shao, Z.; Nan, Y.N.; Xiao, X.; Zhang, L.; Peng, Z. A Multi-View Dense Point Cloud Generation Algorithm Based on Low-Altitude Remote Sensing Images. *Remote Sens.* **2016**, *8*, 381. [\[CrossRef\]](#)

10. Qin, W.C.; Qiu, B.J.; Xue, X.Y.; Chen, C.; Xu, Z.F.; Zhou, Q.Q. Droplet deposition and control effect of insecticides sprayed with an unmanned aerial vehicle against plant hoppers. *Crop Prot.* **2016**, *85*, 79–88. [[CrossRef](#)]
11. Näsi, R.; Viljanen, N.; Kaivosoja, J.; Alhonoja, K.; Hakala, T.; Markelin, L.; Honkavaara, E. Estimating Biomass and Nitrogen Amount of Barley and Grass Using UAV and Aircraft Based Spectral and Photogrammetric 3D Features. *Remote Sens.* **2018**, *10*, 1082. [[CrossRef](#)]
12. Csillik, O.; Cherbini, J.; Johnson, R.; Lyons, A.; Kelly, M. Identification of Citrus Trees from Unmanned Aerial Vehicle Imagery Using Convolutional Neural Networks. *Drones* **2018**, *2*, 39. [[CrossRef](#)]
13. Johansen, K.; Raharjo, T.; McCabe, M. Using Multi-Spectral UAV Imagery to Extract Tree Crop Structural Properties and Assess Pruning Effects. *Remote Sens.* **2018**, *10*, 854. [[CrossRef](#)]
14. Srestasathiern, P.; Rakwatin, P. Oil Palm Tree Detection with High Resolution Multi-Spectral Satellite Imagery. *Remote Sens.* **2014**, *6*, 9749–9774. [[CrossRef](#)]
15. Malambo, L.; Popescu, S.C.; Murray, S.C.; Putmana, E.; Pugh, N.A.; Horne, D.W.; Richardson, G.; Sheridan, R.; Rooney, W.L.; Avant, R.; et al. Multitemporal field-based plant height estimation using 3D point clouds generated from small unmanned aerial systems high-resolution imagery. *Int. J. Appl. Earth Obs. Geoinf.* **2018**, *64*, 31–42. [[CrossRef](#)]
16. Torres-Sanchez, J.; De Castro, A.I.; Pena-Barragan, J.M.; Jimenez-Brenes, F.M.; Arquero, O.; Lovera, M.; Lopez-Granados, F. Mapping the 3D structure of almond trees using UAV acquired photogrammetric point clouds and object-based image analysis. *Biosyst. Eng.* **2018**, *176*, 172–184. [[CrossRef](#)]
17. Juan, G.H.; Eduardo, G.F.; Alexandre, S.; João, S.; Alexandra, N.; Alexandra, C.C.; Luis, F.; Margarida, T.; Ramón, D.V. Using high resolution UAV imagery to estimate tree variables in Pinus pinea plantation in Portugal. *For. Syst.* **2016**, *25*, 16. [[CrossRef](#)]
18. Guo, Q.; Su, Y.; Hu, T.; Zhao, X.; Wu, F.; Li, Y.; Liu, J.; Chen, L.; Xu, G.; Lin, G.; et al. An integrated UAV-borne LiDar system for 3D habitat mapping in three forest ecosystems across China. *Int. J. Remote Sens.* **2017**, *38*, 2954–2972. [[CrossRef](#)]
19. Pedro, M.; Luis, P.; Telmo, A.; Jonas, H.; Emanuel, P.; Antonio, S.; Joaquim, J.S. UAV-based automatic detection and monitoring of chestnut trees. *Remote Sens.* **2019**, *11*, 855. [[CrossRef](#)]
20. Wu, X.; Shen, X.; Cao, L.; Wang, G.; Cao, F. Assessment of Individual Tree Detection and Canopy Cover Estimation using Unmanned Aerial Vehicle based Light Detection and Ranging (UAV-LiDar) Data in Planted Forests. *Remote Sens.* **2019**, *11*, 908. [[CrossRef](#)]
21. Li, J.; Yang, B.; Cong, Y.; Cao, L.; Fu, X.; Dong, Z. 3D Forest Mapping Using a Low-Cost UAV Laser Scanning System: Investigation and Comparison. *Remote Sens.* **2019**, *11*, 717. [[CrossRef](#)]
22. Omair, H.; Nasir, A.K.; Roth, H.; Khan, M.F. Precision forestry: Trees counting in urban areas using visible imagery based on an unmanned aerial vehicle. *IFAC-PapersOnLine* **2016**, *49*, 16–21. [[CrossRef](#)]
23. Zortea, M.; Macedo, M.M.G.; Mattos, A.B.; Ruga, B.C.; Gemignani, B.H. Automatic citrus tree detection from UAV images based on convolutional neural networks. In Proceedings of the 2018 31th SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI), Foz do Iguaçu, Brazil, 29 October–1 November 2018; Volume 11, pp. 1–7.
24. Ramesh, K.; Akanksha, A.; Bazila, B.; Omkar, S.N.; Gautham, A.; Meenavathi, M.B. Tree crown detection, delineation and counting in uav remote sensed images: A neural network based spectral–spatial method. *J. Indian Soc. Remote Sens.* **2018**, *46*, 991–1004. [[CrossRef](#)]
25. Lin, Y.; Lin, J.; Jiang, M.; Yao, L.; Lin, J. Use of UAV oblique imaging for the detection of individual trees in residential environments. *Urban For. Urban Green.* **2015**, *14*, 404–412. [[CrossRef](#)]
26. Poblete-Echeverría, C.; Guillermo, O.; Ben, I.; Matthew, B. Detection and segmentation of vine canopy in ultra-high spatial resolution RGB imagery obtained from unmanned aerial vehicle (UAV): A case study in a commercial vineyard. *Remote Sens.* **2017**, *9*, 268. [[CrossRef](#)]
27. Zhao, T.; Yang, Y.; Niu, H.; Wang, D.; Chen, Y.Q. Comparing U-Net convolutional network with mask R-CNN in the performances of pomegranate tree canopy segmentation. *SPIE* **2018**, *107801J*, 1–9. [[CrossRef](#)]
28. Feduck, C.; McDermid, G.J.; Castilla, G. Detection of Coniferous Seedlings in UAV Imagery. *Forests* **2018**, *9*, 432. [[CrossRef](#)]
29. Sabzi, S.; Abbaspour-Gilandeh, Y.; García-Mateos, G.; Ruiz-Canales, A.; Molina-Maetinez, J.M.; Ignacio Arribas, J. An Automatic Non-Destructive Method for the Classification of the Ripeness Stage of Red Delicious Apples in Orchards Using Aerial Video. *Agronomy* **2019**, *9*, 84. [[CrossRef](#)]

30. Li, X.; Dai, B.; Sun, H.; Li, W. Corn classification system based on computer vision. *Symmetry* **2019**, *11*, 591. [[CrossRef](#)]
31. Blaix, C.; Moonen, A.C.; Dostatny, D.F.; Izquierdo, J.; Le Corff, J.; Morrison, J.; Von Redwitz, C.; Schumacher, M.; Westerman, P.R. Quantification of regulating ecosystem services provided by weeds in annual cropping systems using a systematic map approach. *Weed Res.* **2018**, *58*, 151–164. [[CrossRef](#)]
32. Tang, Y.; Hou, C.; Luo, S.; Lin, J.; Yang, Z.; Huang, W. Effects of operation height and tree shape on droplet deposition in citrus trees using an unmanned aerial vehicle. *Comput. Electron. Agric.* **2018**, *148*, 1–7. [[CrossRef](#)]
33. Nam, M.; Rhee, P.K. An efficient face recognition for variant illumination condition. *Intell. Signal Process. Commun. Syst. (ISPACS)* **2004**, *12*, 111–115. [[CrossRef](#)]
34. Tan, S.F.; Isa, N.A.M. Exposure Based Multi-Histogram Equalization Contrast Enhancement for Non-Uniform Illumination Images. *IEEE Access* **2019**, *7*, 70842–70861. [[CrossRef](#)]
35. Kim, J.Y.; Kim, L.S.; Hwang, S.H. An advanced contrast enhancement using partially overlapped sub-block histogram equalization. *IEEE Trans. Circuits Syst. Video Technol.* **2002**, *11*, 475–484. [[CrossRef](#)]
36. Otsu, N.A. Threshold Selection Method from Gray-Level Histograms. *IEEE Trans. Geosci. Remote Sens.* **1979**, *9*, 62–66. [[CrossRef](#)]
37. Pu, Y.F.; Siarry, P.; Chatterjee, A.; Wang, Z.N.; Zhang, Y.; Liu, Y.G.; Zhong, J.L.; Wang, Y. A Fractional-Order Variational Framework for Retinex: Fractional-Order Partial Differential Equation-Based Formulation for Multi-Scale Nonlocal Contrast Enhancement with Texture Preserving. *IEEE Trans. Geosci. Remote Sens.* **2018**, *27*, 1214–1229. [[CrossRef](#)] [[PubMed](#)]
38. Wu, Q.S.; Luo, X.L.; Li, H.; Liu, P.Z. An Improved Multi-Scale Retinex Algorithm for Vehicle Shadow Elimination Based on Variational Kimmel. In Proceedings of the 7th International Conference on Ubiquitous Intelligence and Computing and 7th International Conference on Autonomic and Trusted Computing (UIC/ATC), Xi'an, China, 26–29 October 2010; IEEE Computer Society: Washington, DC, USA, 2010; pp. 31–34. [[CrossRef](#)]
39. Yu, T.; Meng, X.; Zhu, M.; Han, M. An Improved Multi-scale Retinex Fog and Haze Image Enhancement Method. In Proceedings of the 2016 International Conference on Information System and Artificial Intelligence (ISAI), Hong Kong, China, 24–26 June 2016; IEEE: Piscataway, NJ, USA, 2016; pp. 557–560. [[CrossRef](#)]
40. Ojala, T.; Pietikäinen, M.; Mäenpää, T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans. Pattern Anal. Mach. Intell.* **2002**, *24*, 971–987. [[CrossRef](#)]
41. Guo, Q.; Chen, Y.; Tang, Y.; Zhuang, J.; He, Y.; Hou, C.; Chu, X.; Zhong, Z.; Luo, S. Lychee Fruit Detection Based on Monocular Machine Vision in Orchard Environment. *Sensors* **2019**, *19*, 4091. [[CrossRef](#)] [[PubMed](#)]
42. Kyung, W.; Kim, D.; Ha, Y. Real-time multi-scale Retinex to enhance night scene of vehicular camera. In Proceedings of the 17th Korea-Japan Joint Workshop on Frontiers of Computer Vision (FCV), Ulsan, Korea, 9–11 February 2011; Volume 17, pp. 1–4. [[CrossRef](#)]
43. Fawcett, T. An introduction to ROC analysis. *Pattern Recognit. Lett.* **2006**, *27*, 861–874. [[CrossRef](#)]

