Article

Cherry Tree Crown Extraction from Natural Orchard Images with Complex Backgrounds

Zhenzhen Cheng [1], Lijun Qi [1,*] and Yifan Cheng [2]

1   College of Engineering, China Agricultural University, No.17 Qing Hua Dong Lu, Haidian District, Beijing 100083, China; zhenzhenc@cau.edu.cn
2   College of Horticulture, Xinyang Agriculture and Forestry University, No. 1 Beihuan Road, Pingqiao District, Xinyang 464007, China; cyf9405@163.com
*   Correspondence: qilijun@cau.edu.cn

Abstract: Highly effective pesticide applications require a continual adjustment of the pesticide spray flow rate that attends to different canopy characterizations. Real-time image processing with rapid target detection and data-processing technologies is vital for precision pesticide application. However, the extant studies do not provide an efficient and reliable method of extracting individual trees with irregular tree-crown shapes and complicated backgrounds. This paper on our study proposes a Mahalanobis distance and conditional random field (CRF)-based segmentation model to extract cherry trees accurately in a natural orchard environment. This study computed Mahalanobis distance from the image's color, brightness and location features to acquire an initial classification of the canopy and background. A CRF was then created by using the Mahalanobis distance calculations as unary potential energy and the Gaussian kernel function based on the image color and pixels distance as binary potential energy. Finally, the study completed image segmentation using mean-field approximation. The results show that the proposed method displays a higher accuracy rate than the traditional algorithms K-means and GrabCut algorithms and lower labeling and training costs than the deep learning algorithm DeepLabv3+, with 92.1%, 94.5% and 93.3% of the average P, R and F1-score, respectively. Moreover, experiments on datasets with different overlap conditions and image acquisition times, as well as in different years and seasons, show that this method performs well under complex background conditions, with an average F1-score higher than 87.7%.

Keywords: agricultural computer vision; tree-crown segmentation; complex scene; natural orchard environment

1. Introduction

Precision agriculture is a management strategy that uses modern science and technology to obtain required agricultural information for efficient precision crop management, such as formula fertilization, precision seeding, pest control, weed removal and water management [1–3]. Precision-spraying technology is vital for prevention and pest control. However, although precision-spraying technologies have been widely used in precision agricultural production, their efficient application on cherry orchards remains a big challenge [4]. For instance, implementing established spraying strategies to tree crops with different canopy characteristics, such as irregular sizes and shapes, may lead to spray drift and pesticide overdosing, posing a great risk to farmers and the environment [5,6]. The situation may be worsening when the crown's size and shape change significantly in different growth stages [7]. To reduce the negative impact of pesticide application, it is necessary to develop a canopy extraction technology that provides accurate tree-canopy data for precision-spraying systems.

Proximal sensing vehicle-mounted technologies are defined as the use of sensors and traction systems to identify and detect agricultural parameters [8,9]. At present, field-based sensors are widely adopted for automatic tree identification, such as Visible–Near-Infrared

imaging [10,11], stereo imaging [12,13] and thermal imaging [14,15]. Target extraction based on RGB (Red Green Blue) digital cameras has seen a wide application in precision farming due to its low cost and non-contact data collecting [16–18]. In this process, color index-based segmentation techniques are mostly applied to complete the crucial background removal. Several studies summarized the well-performing color indices in distinguishing plants from backgrounds [19–21], including excessive green index (ExG) [22], excess green minus excess red index (ExR) in RGB color space [23] and spectral vegetation indices such as the Normalized Difference Vegetation Index (DVI) [24] and the Green–Red Vegetation Index (GRV) [25]. However, such color index-based methods will be ineffective when the background and plant share similar colors, e.g., green weeds and canopies. Omair extracted canopies from artificial turf in high-altitude communities using color and texture features [26], which is not applicable to a natural orchard environment where the physiological features will be unstable, especially when the canopy overlaps the weeds [27]. Past studies showed that thresholding and filtering technology based on the grayscale or edge characteristics are often used in image pre-processing or combined with other segmentation methods [28,29], which means that the method based on a single feature is difficult to remove complex backgrounds.

By contrast, the statistics-based machine learning (ML) method can overcome the limitations of feature-based segmentation [30–32]. ML methods can be divided into two categories: unsupervised and supervised learning algorithms. Unsupervised learning algorithms typically adopt clustering methods, such as Fuzzy C-Means (FCM), K-means and Gaussian Mixture Model (GMM). Liu et al. [33] used the Type-2 FCM to extract the Ginkgo and Platanus canopies from the UAV(Unmanned Aerial Vehicle) image without a complex background. Qi et al. [34] proposed an effective fruit tree segmentation method based on K-means clustering and color features to separate the background from the canopy. This method, however, is not ideal for the input images that contain weed background. Abdalla et al. [35] employed GMM, self-organizing maps, Fuzzy C-means and K-means algorithms calculated with the highest color features from ten color models to segment oilseed rape images, but the results cannot be generalized to other complex situations as the features between oilseed rape and background are obviously different and easy to distinguish. On the other hand, the unsupervised learning algorithms do not require model training and are simple to use, but they cannot process images with complex backgrounds.

Supervised learning algorithms, which can be divided into two types, i.e., traditional supervised learning algorithms and deep learning, can perform well in the segmentation task of complex scenes [36–38]. Chen et al. [39] proposed a citrus canopy segmentation method based on the SVM (Support Vector Machines) segmentation model trained with 14 color features and five statistical textures. Mattos et al. [40] used the CNN (Convolutional Neural Network) algorithm to segment the citrus canopy from the background, which has an overall accuracy of 94% in seven different orchards. Wu et al. [41] proposed the deep learning model to extract apple tree's canopies and parameters with an over 90% of segmentation accuracy and recall rate. This body of research has exemplified the application of artificial intelligence in the farming industry. However, traditional supervised learning algorithms rely on complex feature engineering (FE), while deep learning requires large-size labeled data and high-performance computers. Lu et al. [42] summarized 34 available deep learning datasets in the agricultural field, including weeds, fruits and common ground crops, but no open-source datasets on fruit trees are included. Hence, supervised learning is not always the best choice.

Compared to the structured environments, a natural orchard environment poses more challenges to image segmentation. For example, the images taken from cherry orchards may include various non-target elements such as sky, land, cover films (Figure 1a) and houses (Figure 1b). Moreover, the tree's physiological properties, especially the porous media, lead to an uneven distribution of light within the canopy (Figure 1c). It is also difficult to differentiate the canopy from the weeds when there is an overlap between the two (Figure 1d).
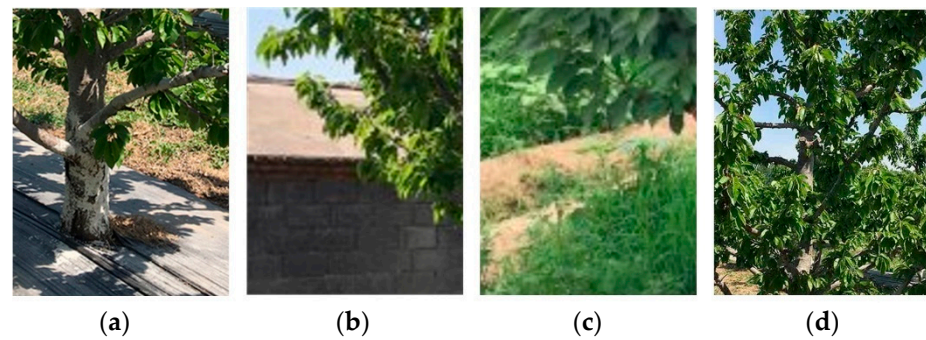
Figure 1. Challenges of crown segmentation in the unstructured environment: (**a**) sky, land and cover films; (**b**) house; (**c**) uneven light distribution within the canopy images; and (**d**) weeds.

The objective of the current study was, therefore, to propose a method for extracting cherry-tree canopies from the complex background that has a higher accuracy rate than the traditional supervised learning methods and lower computational costs than unsupervised algorithms, such as deep learning. This study provides a new way to characterize image features of tree crown by computing Mahalanobis distance from the image's color, brightness, and location features to acquire an initial classification of the canopy and background. Moreover, the tree-crown area features and global image features were considered by using the conditional random field (CRF). It was created by using the Mahalanobis distance calculations as unary potential energy and the Gaussian kernel function based on the image color and pixels distance as binary potential energy. Finally, the study completed image segmentation using mean-field approximation. This proposed work will contribute to future machine-vision-based tree-crop extraction.

2. Materials Acquisition

2.1. Test Site and Image Acquisition

The study was conducted on a cherry orchard of Zhongnong Futong Company in Tongzhou, Beijing (116°48′32.1″ E and 39°51′46.37″ N) (Figure 2a). The test area has representative climatic characteristics in local orchards, with weeds germinating and flourishing in May and June. Cherry trees are spaced 4 m per plant with a 5 m path between rows. The average height of the canopy was about 1.7 m. An untrained tree experimental field was selected, with a total area of 10,660 m^2 (82 × 130 m).

Figure 2. Test site, and the illustrations of images under different lighting conditions and weeds densities: (**a**) cherry orchard in Tongzhou Beijing, (**b**) cloudy day with high-density weeds and (**c**) sunny day with normal-density weeds.

The trees were photographed with a digital camera (LICE-7M2, Sony Inc., Tokyo, Japan), which features a 24.3-megapixel sensor that enables high-resolution images, 5-Axis Steady Shot INSIDE Stabilization, ISO sensitivity of 400, a focal length of 135 mm and a body size of 126.9 × 95.7 × 59.7 mm. The camera was mounted on Beno IT15 gimba-

land positioned 1 to 1.5 m above the ground and 3 to 3.5 m in front of the tree trunk. In order to execute the image processing program, the input image was adjusted to the following unified parameters: png format (lossless compression), 900:1600-pixel aspect ratio, 24 bits (three channels in total, and 8 bits for each channel). The images were collected under different light conditions and in different growth stages from April to October, each year, between 2018 and 2020. Figure 2 illustrates images captured under different lighting and weedy conditions in the field. Figure 2b exemplifies a cloudy day with high-density weeds, whereas Figure 2c displays a sunny day with normal-density weeds.

Template and Ground Truth Generation

The template and the ground truth are images of manually delineating the canopy area, with the former used for image classification and the latter for algorithm performance comparisons. In order to minimize subjectivity in image labeling, we invited an expert with experience in agricultural image processing to use Photoshop to generate ground truth and the template. The template includes two manually segmented images, which only capture the crown area and are representative in terms of the colors, light and shooting angles in the image datasets. They were captured on a sunny and cloudy day, respectively. The template was created with Adobe Photoshop 2018 as an image-labeling tool [43], following three steps: selecting the lasso tool in Adobe Photoshop 2018 to outline the edge of the tree crown, taking the inside region of the closed curve as the image foreground and using the fill function to set the area outside the foreground to black (Figure 3b). In effect, the ground truth processing for algorithm performance comparison follows the same workflow but adds one step of setting the image foreground to white (Figure 3c).
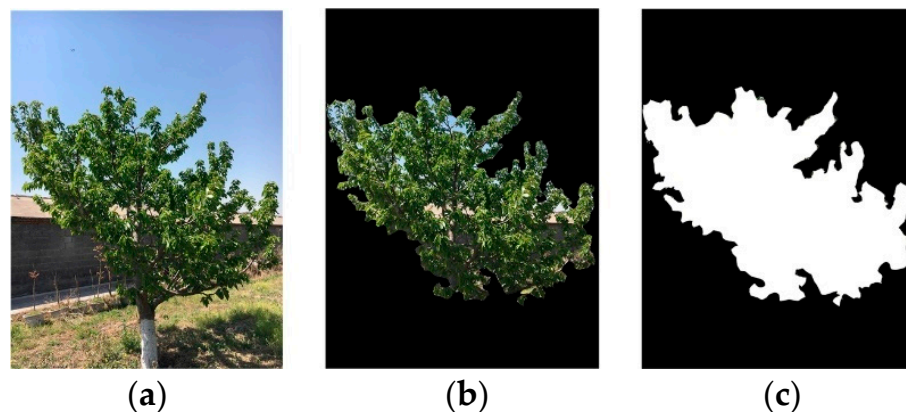

(a)　　　　　　　　(b)　　　　　　　　(c)

Figure 3. Image annotation schematic: (**a**) original canopy image, (**b**) standard image and (**c**) ground truth image.

3. Methodology

Figure 4 is a flowchart of the proposed method: feature extraction, Mahalanobis distance computation and conditional random field (CRF) building. All algorithms were developed in MathWorks MATLAB R2018a and Python3.6 software on a PC equipped with an Intel®Core ™ i7-6700 central processing unit (CPU) and 16 GB of random-access memory (RAM).
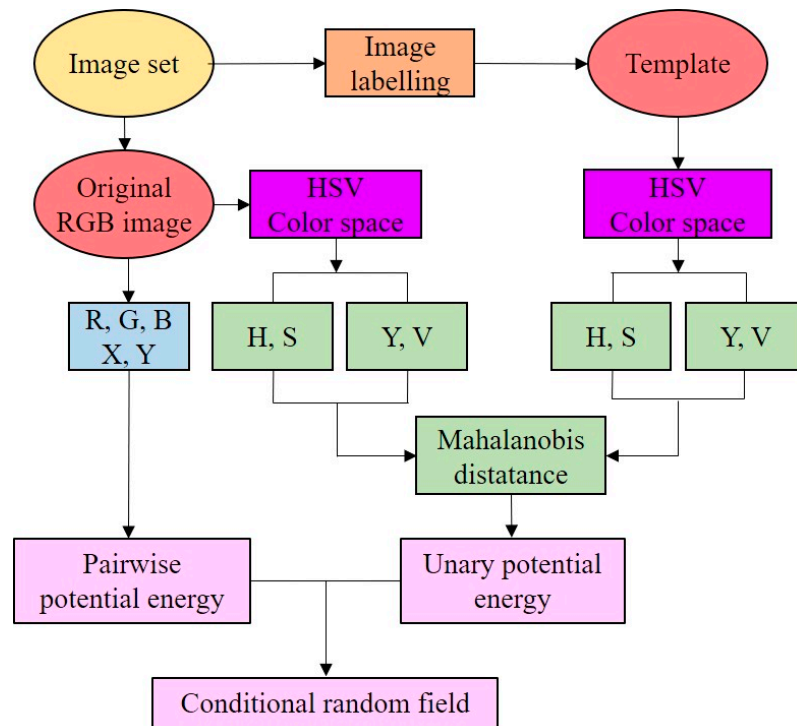
Figure 4. Overall flowchart of the proposed method.

3.1. Feature Construction of Tree Crown

Feature extraction consists of two steps: extracting color features to remove the non-green background and extracting brightness and height features to separate canopies from weeds.

3.1.1. Color Feature Extraction

Since the plants were distinctively green, the study removed non-green background regions based on color features [44]. HSV (hue, saturation and value) color model was adopted to extract color features, where H represents hue, S represents saturation, and *V* represents brightness. In this model, only the hue (*H*) and saturation (*S*) channels describe color information, while brightness (*V*) is a separate channel [45]. Thus, the HSV model will effectively deal with lighting changes or uneven lighting in the orchard image, which is an unachievable outcome in RGB color space. Converting RGB color space to HSV follows the calculations below [45]:

$$MAX = \max(R', G', B'), MIN = \min(R', G', B') \tag{1}$$

$$H = \begin{cases} 0, & \text{if } R' = G' = B' \\ 60 \times \left(0 + \dfrac{G' - B'}{MAX - MIN}\right), & \text{if } MAX = R' \\ 60 \times \left(2 + \dfrac{B' - R'}{MAX - MIN}\right), & \text{if } MAX = G' \\ 60 \times \left(4 + \dfrac{R' - G'}{MAX - MIN}\right), & \text{if } MAX = B' \end{cases} \tag{2}$$

$$S = \left\{ \begin{array}{cc} 0, & \text{if } R' = G' = B' \\ else, & \frac{MAX - MIN}{MAX} \end{array} \right\} \times 100\% \tag{3}$$

$$V = MAX \times 100\% \tag{4}$$

where we have the following:

$$R' = \frac{r}{r+g+b}, \quad G' = \frac{g}{r+g+b}, \quad B' = \frac{b}{r+g+b}$$

where r, g and b are the red, green and blue channels in RGB color space; R', G' and G' are the normalized red, green and blue channels; H represents the type of color; S indicates the degree of color saturation; and V is the value of brightness.

Figure 5 is an example of the variation of spectral components at the exacted location (marked as a red dotted line) under different lighting conditions. Figure 5b shows that the G component in the RGB color model changes sharply, while the H and S components in the HSV color space are stable. Therefore, the hue (H) and saturation (S) are selected as the color features for non-green background removal.
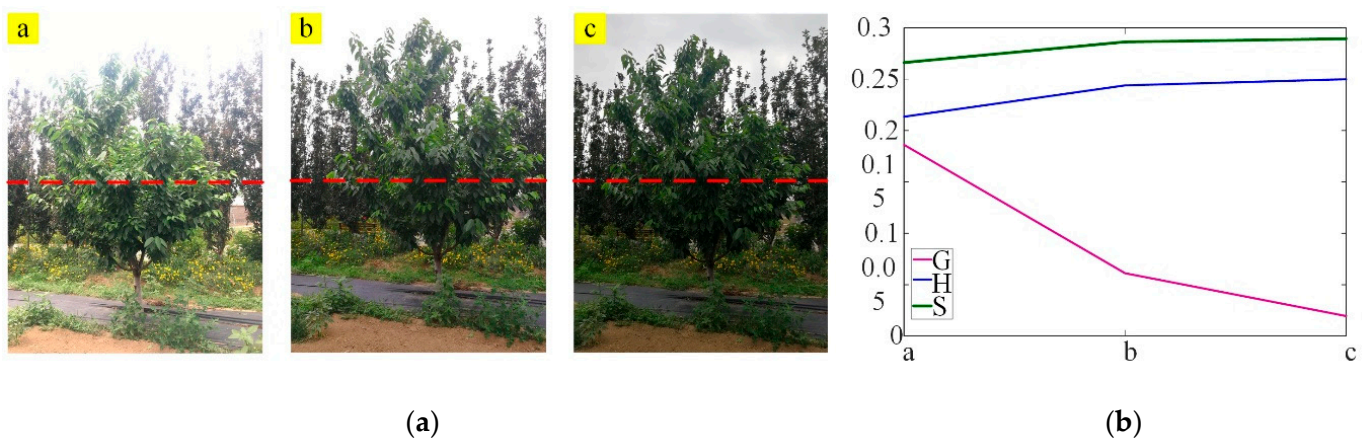

(**a**)   (**b**)

Figure 5. Influence of lighting conditions on spectral components: (**a**) an image of a fruit tree with fading light from left to right and (**b**) the curve of green, hue and saturation under different lighting conditions.

3.1.2. Brightness and Height Feature Crossing

Weeds have a similar green color to the canopy; thus, we are not able to remove them from the image by using color features only. Weeds are annual herbaceous plants usually growing 0.5 m tall, while tree crops can be up to 2.5–4.0 m. However, this significant height difference may be ineffective in distinguishing them when the weeds and the tree canopy may come into contact and overlap in the image. Therefore, the target region can be divided into two parts by plant height: the upper canopy area and the overlapping area between the canopy and the weeds. In the isolated canopy region, it is easy to distinguish the canopy from weeds based on height. In the overlapping area, the distinguishment can be based on light intensity. The light intensity from the upper crown layer to the lower is gradually weakened [46]. Thus, the lower crown layer is mainly in the shadow due to insufficient light, while the weed area is in the sun (Figure 6). This study selected the brightness distribution in the vertical direction of the image as the feature to distinguish between the canopy and weeds. The height and brightness features can be extracted from vertical pixel coordinates of the image and the V component of the HSV space, respectively. Pixel coordinates represent the position of pixels in the image. For an image with height H, each pixel height is (Y_j = 1,2, H). The V component was obtained by Equation (4).
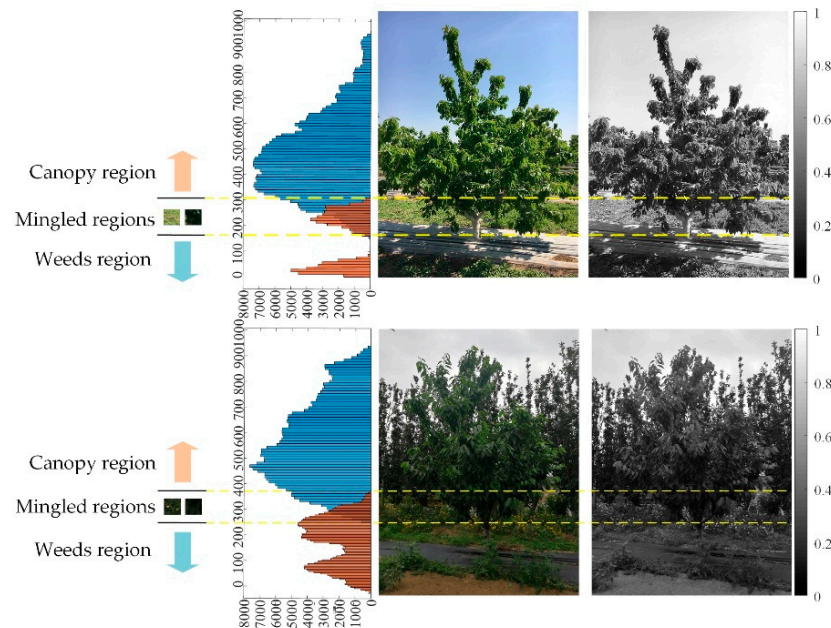
Figure 6. The illustrations of the brightness and height distribution of canopy under different lighting conditions. The yellow dotted lines indicate the mingled areas, and the color bar manifests brightness value. From left to right are the crown- and weed-height distribution, the original image and the brightness image; from top to bottom are the sunny fruit-tree images and cloudy fruit-tree images.

3.2. Mahalanobis Distance Computation

Mahalanobis distance is a distance criterion that assigns each pixel into prediction groups, i.e., tree crown and background, by measuring pixel similarity [47]. This study utilized Mahalanobis distance rather than other measures, e.g., Euclidean distance, because Mahalanobis distance considers the correlation between features. The Mahalanobis distance classified the canopy and background pixels by measuring the feature similarity between the original image and the template. The corresponding calculation follows two steps:

(a) Computing mean vectors and covariance matrix: The mean vectors are the average value of the feature, commonly referring to the centroid of data distribution. The feature is a four-dimensional vector (*H*, *S*, *Y* and *V*) which is extracted from the template and sample images. The mean vectors are calculated as follows:

$$\mu = f\left(\overline{H}, \ \overline{S}, \ \overline{V}, \ \overline{Y}\right) \tag{5}$$

$$\overline{H} = \frac{1}{n}\sum_{i=1}^{n} H_i, \overline{S} = \frac{1}{n}\sum_{i=1}^{n} S_i, \overline{V} = \frac{1}{n}\sum_{i=1}^{n} V_i, \overline{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i \tag{6}$$

where *H*, *S* and *V* are the hue, saturation and brightness components of HSV color space, respectively; *Y* is pixel height; n is the number of pixels; *i*=1,2, 3, *n*; and *f* is a vector composed of *H*, *S*, *V* and *Y*. The covariance matrix is a square and symmetric matrix containing the variances and covariances associated with components of feature *f* (*H*, *S*, *V* and *Y*). The formula to compute the covariance between two variables is as follows:

$$Cov(f, \mu) = \frac{1}{n}(f - \mu)^{\mathrm{T}}(f - \mu) \tag{7}$$

where *f* is a pair of variables with the four components (*H*, *S*, *V*, *Y*); *μ* is the mean vectors obtained by Equation (5); *n* is the number of pixels.

(b) Computing the Mahalanobis distance: Mahalanobis distance will divide each pixel into two groups described by different mean vectors and covariances. Its formula Equation (8) is as follows:

$$M = \sqrt{(f - \mu)^T Cov^{-1}(f - \mu)}, f = (H, S, V, Y) \tag{8}$$

where f is four-dimensional vectors containing the H, S, V and Y values of each pixel; μ is the mean vectors calculated by Equation (5); the Cov is the covariance matrix calculated by Equation (7).

Figure 7 exemplifies the Mahalanobis distance calculation results. Figure 7b is the Mahalanobis distance obtained only by the H and S features where the non-green background was removed based on color feature, but the weeds were remained in the image. Figure 7c is the Mahalanobis distance based on H, S, V and Y features. From Figure 7, the Mahalanobis distance of the canopy regions is small, the gray value is low and the color is close to black. When the background area has a low similarity to the canopy, the Mahalanobis distance value becomes larger and the color is close to white.
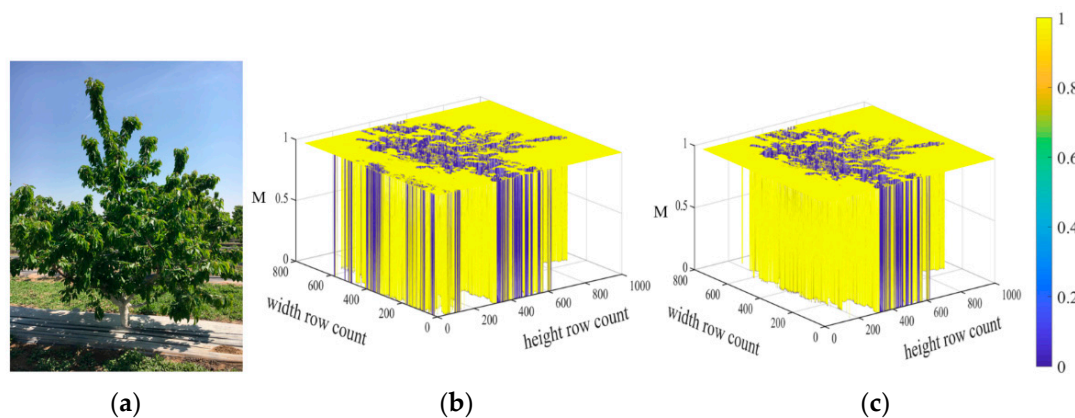

Figure 7. Examples of Mahalanobis distance computing: (**a**) original image; (**b**) three-dimensional image of Mahalanobis distance based on H and S features; and (**c**) three-dimensional image of Mahalanobis distance based on H, S, V and Y features.

3.3. Conditional Random Field for Image Segmentation

After a pre-classification of the image based on Mahalanobis distance, this section discusses conditional random field (CRF) modeling for image segmentation.

3.3.1. Energy Function Construction

Conditional Random Field (CRF) is a conditional probability distribution model that outputs random variables with a set of random input variables [48]. In the image segmentation task, CRF treats pixels or pixel features as random input variables with a probability distribution and pixel label as output variables. If the definition of the random variable $Y_i = (y_1, y_2, y_3, y_n;)$ obeys the Markov property, the distribution of Y_i constitutes a conditional random field. Each pixel i is assigned a corresponding label Y_i through observable variable $X_i = (x_1, x_2, x_3, x_n)$ in this random field. Figure 8 illustrates the overall operation of the CRF model.
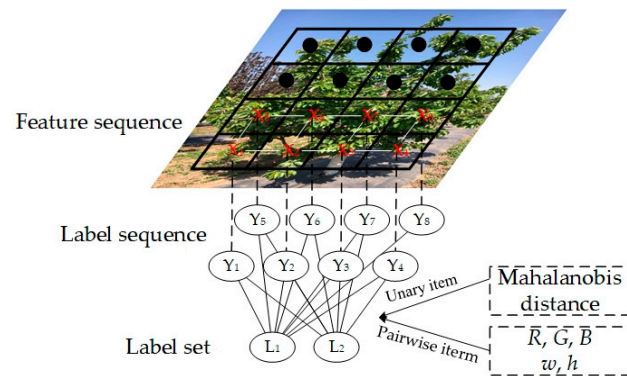
Figure 8. The CRF operation chart. *X* is the feature sequence, *Y* is the label sequence, L1 indicates the tree-crown class, and L2 indicates the background class. *R*, *G* and *B* are the red, green and blue channels in RGB color space; *w* and *h* represent the vertical and horizontal coordinate of the image, respectively.

The specific steps are as follows:

(1) Model construction: establishing the mapping relationship between *X* and *Y* through the conditional probability distribution P(*Y* | *X*). In the fully connected conditional random field model, P(*Y* | *X*) is expressed in the form of Gibbs distribution:

$$P(Y|X) = \frac{1}{Z(X)} \exp\{-E(Y|X)\} \tag{9}$$

where *X* indicates the feature set f, and *Y* corresponds to the class labels, $Y \in \{L1, L2\}$. L1 represents the tree crown, and L2 is the background. *Z* is a normalization term that ensures the distribution *P* sums to 1 and is defined as follows:

$$Z(X) = \sum_Y \exp\{E(Y|X)\} \tag{10}$$

where E(*Y* | *X*) denotes the Energy function.

(2) The Energy function minimization: CRF aims to find the output *Y* with the maximum conditional probability P(*Y* | *X*). According to Equation (9), the problem of conditional probability maximization is the problem of energy minimization, which can be expressed as follows:

$$y^* = \arg\ \min_y E(Y|X) \tag{11}$$

where y* is the minimization of Energy function E(*Y* | *X*) The E(*Y* | *X*) consists of two types of potential energy: unary potentials and pairwise potentials:

$$E(Y|X) = \sum_{i=1}^N \psi_u(y_i) + \sum_{i,j=1}^N \psi_p(y_i, y_j) \tag{12}$$

where $\psi_u(y_i)$ is the unary potential for the probability of pixel *i* taking the label y_i, denoting the pixel's local information; $\psi_p(y_i, y_j)$ is the pairwise potential, representing the label class similarity relationship between nearby pixels *i* and *j*, including inter-pixel global information; and *i*, $j \in \{1, 2, 3, N\}$ are the pixel indices.

Equation (12) shows that unary and pairwise potential functions are the crux of conditional random field modeling. Their respective definition follows.

(3) The unary potential construction: The unary potential is the probability that a pixel obtains the corresponding label, indicating the category information of the current observation point. The study employed the Mahalanobis distance classifier results

described in Section 3.2 to construct the unary potential energy. The unary potential takes the negative logarithm to provide a framework that unifies energy minimization:

$$\psi_u(y_i) = -\log(1 - P_M(y_i))\tag{13}$$

$P_M(y_i)$ is the label assignment probability for each pixel by the Mahalanobis distance classifier, which is calculated by Equation (8). The smaller the Mahalanobis distance, the greater the probability that the pixel is assigned to the canopy category. When the probability that the pixel i takes the label y_i is large, the unary potential and energy are small.

(4)　The pairwise potential computation: The pairwise potential pixels are constraints of the final label assignment. Its goal is to assign adjacent labeled pixels with similar characteristics to the same category. The punishment strength is positively correlated to the feature difference between adjacent pixels under the same label, thereby restricting the classifier's misclassification behavior. The general form of the paired potential function is a linear combination of Gaussian kernel functions:

$$\psi_P(y_i, y_j) = u(y_i, y_j) \sum_{m=1}^{K} \omega^{(m)} K^{(m)}(f_i, f_j)\tag{14}$$

where $u(y_i, y_j)$ is a constant symmetric label compatibility function between the labels y_i, and y_j to punish the similar pixels with different class labels. When the classifier assigns different labels to adjacent pixels, the greater the difference between pixel features, the smaller the penalty is, which is consistent with Gibbs energy minimization. Moreover, $\omega^{(m)}$ is the coefficient weight of the given kernels; $m = (1, 2, 3, N)$ is the number of kernel $K^{(m)}$; $K^{(m)}(f_j, f_j)$ is the kernel potential function on feature vectors; and f_j is feature vectors of pixels i, while f_j is feature vectors of pixels j.

This study used two Gaussian kernels to construct the K, which is primarily composed of the pixels' spectral and distance information ($m = 2$):

$$K^{(2)}(f_i, f_j) = \omega^{(1)} \exp\left(-\frac{|p_i - p_j|^2}{2\theta_\alpha^2} - \frac{|C_i - C_j|^2}{2\theta_\beta^2}\right) + \omega^{(2)} \exp\left(-\frac{|p_i - p_j|^2}{2\theta_\gamma^2}\right)\tag{15}$$

where the first item is an appearance kernel based on RGB color and distance information; C is a three-dimensional vector composed of R, G and B components; P is a two-dimensional position vector composed of vertical and horizontal directions; C_i and C_j are color vectors of the pixel on positions p_i and p_j. The second item is a smooth kernel used to remove small isolated areas; $\omega^{(1)}$ and $\omega^{(2)}$ are the coefficient weights of each kernel; θ_α, θ_β and θ_γ are the parameters of Gaussian kernel.

3.3.2. CRF Inference

The average field approximation theory is an efficient inference method that approximates the conditional probability distribution $P(Y)$ with a simple distribution $Q(Y)$, thereby simplifying the calculation process [49]. The average field approximation is calculated with Equation (16):

$$Q(Y) = \prod_i Q_i(y_i)\tag{16}$$

where $Q_i(y_i)$ is the independent marginal distribution of the random variable y_i. For ease of exposition, assume that $Q(Y)$ is the product of multiple independent distributions.

To make the distribution $Q(Y)$ approximate the true distribution $P(Y)$, this study used the Kullback–Leibler (KL) distance as a metric, which is defined as follows:

$$D(Q\|P) = \sum_{i,y} Q_i(Y_i) \log\left(\frac{Q_i(y_i)}{P_i(y_i)}\right)\tag{17}$$

where D is the KL distance between the Q distribution and the P distribution. The Q distribution can be calculated by taking the minimum KL distance as the convergence criterion. The iteration flow includes five steps: message passing, weighting filter output, compatibility conversion, the unary potentials adding and probability normalizing. Liu has introduced this iteration process in detail [50].

3.4. Evaluation Indices and Competing Segmentation Methods for Segmentation Performance

3.4.1. Competing Segmentation Methods

To validate the proposed method's crown segmentation effects on cherry trees, this study compared it with the K-means clustering algorithm, Convolutional Neural Networks (CNN) and GrabCut algorithm, which were widely used in tree image segmentation [51–53]. K-means is an unsupervised machine learning algorithm that does not require labeled datasets and model training [54]. To extract crowns by using the K-means clustering algorithm, the Elbow methods were employed to determine the optimal number of clusters (k) by computing the sum of squared errors (SSE). Figure 9a shows the relationship between k and SSE. When the k value increases, the SSE value drops sharply. However, the SSE value will not change significantly if the k number continues to increase. Therefore, the K value at the bending position identifies the optimal number of clusters. CNN is a supervised machine learning algorithm that relies on labeled data and model training. DeepLabV3+ is one of the best CNN-based semantic segmentation models at present [55]. It improves the Xception network and adopts an Encoder–Decoder structure, thereby optimizing boundary details by restoring the ow-level features (Figure 9d). DeepLabV3+ retrains the Atrous Spatial Pyramid Pooling (ASPP) module to acquire multi-scale information. To segment canopies using DeeplabV3+, this study created 500 single-channel labeled images (see Figure 9c for an example), of which 400 were used for training and 50 for verification and testing, respectively. The model was trained on the Ubuntu 18.04 operating system of NVIDIA RTX 2080TI. GrabCut is an interactive algorithm that requires user interaction to implement image segmentation [56]. GrabCut segments images by creating a new pixel distribution that is close to the foreground's pixel distribution. The foreground is the area inside the red bounding box, which is manually drawn by experts with image processing experience (see Figure 9b).
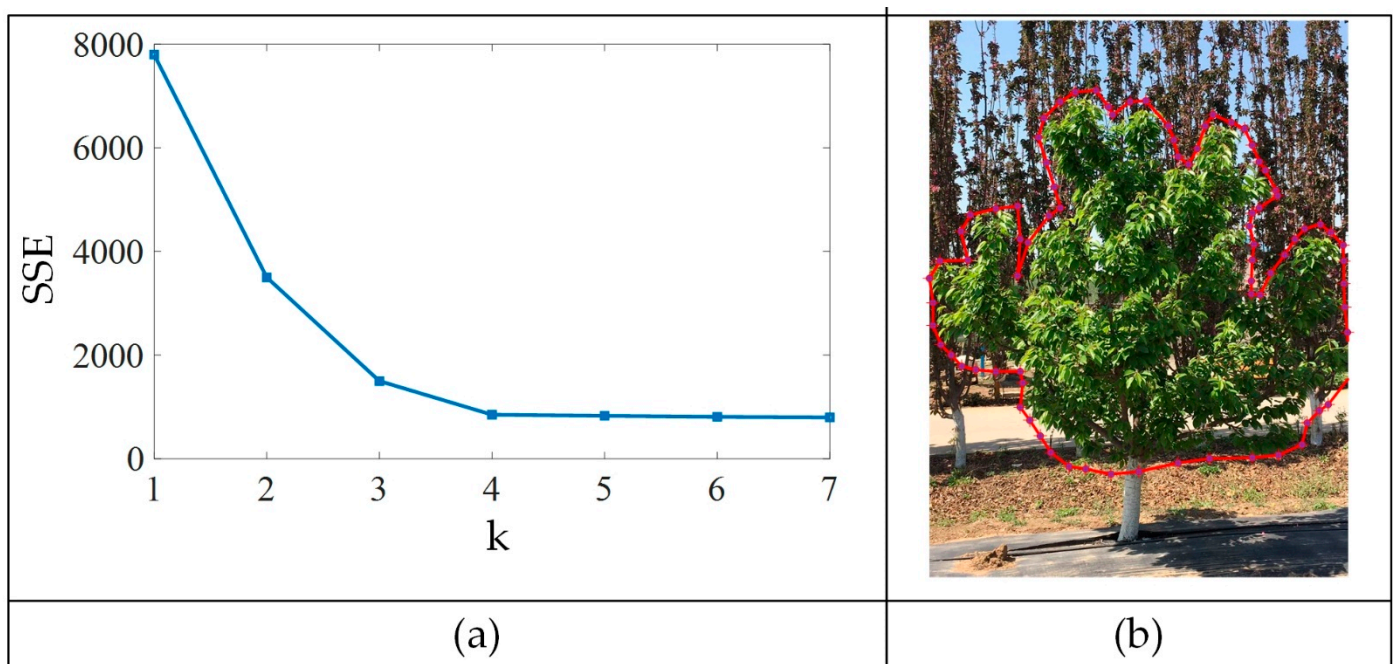

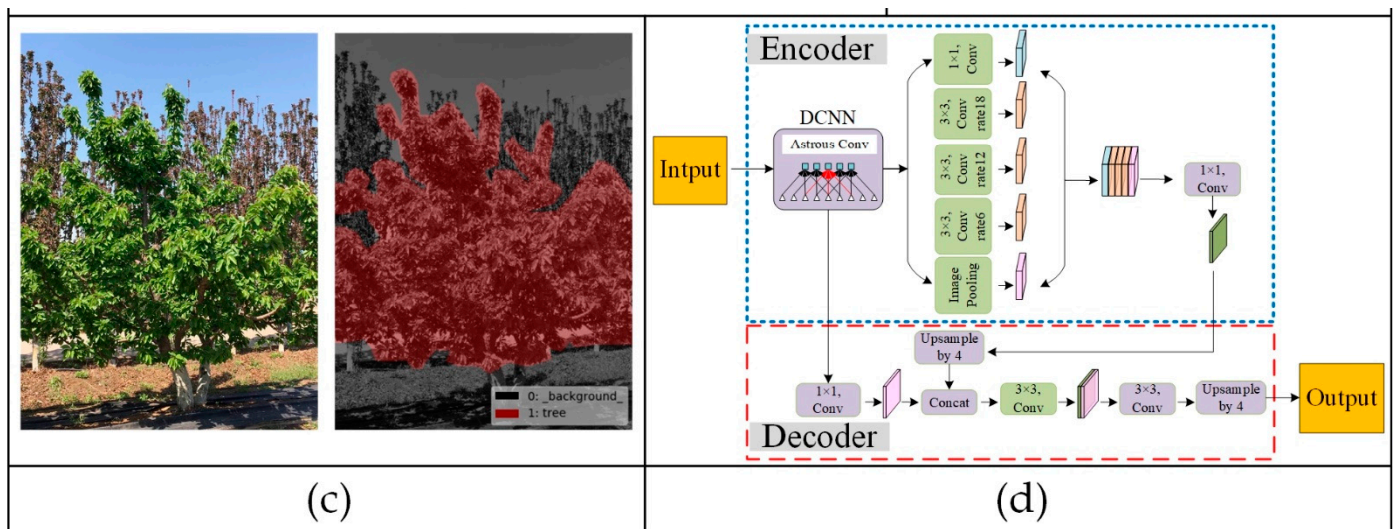(a)　　　　　　　　　　　　　　　　　　　　　　　　(b)

Figure 9. *Cont.*

Figure 9. The illustrations of the four competing segmentation. (**a**) Relationship between k and SSE. (**b**) Image foreground labeling. (**c**) Labeled data for model training. (**d**) The structure of the DeepLabV3+ network.

3.4.2. Evaluation Indices

Three performance measures, namely Precision (*P*), Recall (*R*) and F1-score (*F1*), are introduced to evaluate the cherry canopy segmentation results. *P* is the correctly extracted percentage of the canopy pixels with the segmentation method, indicating the segmentation accuracy. *R* is the percentage of the missing canopy pixels, measuring the segmentation completeness. *F1* is the compromised mean of the *P* and *R*, representing the global metric of canopy segmentation accuracy. The actual trees area was manually counted using Adobe Photoshop 2018. These metrics are defined as follows:

$$P = TP/(TP + FP) \times 100\% \tag{18}$$

$$R = TP/(TP + FN) \times 100\% \tag{19}$$

$$F_1 = 2PR/(P + R) \times 100\% \tag{20}$$

where *TP* is the number of canopy pixels correctly produced by the segmentation algorithm. *FP* represents the number of background pixels that are misidentified as trees. *FN* represents the number of tree pixels that are misidentified as background. A higher value of these three metrics indicates the segmentation method's better performance.

4. Results and Discussion

4.1. Segmentation Results of the Four Competing Methods

To evaluate the performance of the proposed method, the study selected 200 images with complex backgrounds, including bare soil, weeds, sidewalks, houses, plastic films and shelter, and under different lighting conditions. Figure 10 shows the original images of cherry trees. There were four rows of images that were taken on different weather conditions, i.e., sunny or cloudy and with different densities of weeds, i.e., high-density or low-density. Figure 10 exemplifies the four competing methods' segmentation results. From left to right are: the original image, the result of the proposed method, the K-means algorithm's result, the DeepLabV3+ algorithm's result, the GrabCut algorithm's result and the ground truth.
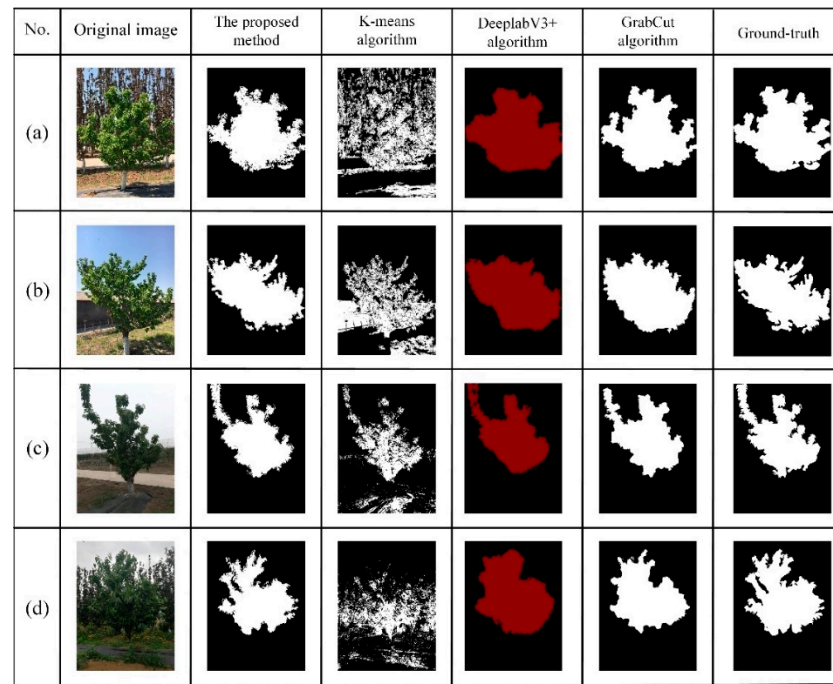
Figure 10. Comparison of different segmentation results: (**a**) sunny day and low-density weeds; (**b**) sunny day and high-density weeds; (**c**) cloudy day and low-density weeds; and (**d**) cloudy day and low-density weeds. From left to right are the original image, the result of the proposed method, the K-means algorithm's result, the DeepLabV3+ algorithm's result, the GrabCut algorithm's result and the ground truth.

Figure 10 shows that the K-means algorithm failed to discriminate the canopy from the background. Lighting conditions have a greater impact on the K-means algorithm than the densities of weeds, as this method over-segmented all shaded areas as tree crowns. The remaining three methods performed well in canopy identification, robust under different light conditions and with different densities of weeds. However, the DeepLabV3+ algorithm results are not satisfactory, which is ineffective in crown branch recognition due to the ignored local information in convolution and upsampling. The overall segmentation results using the GrabCut method are better than that of K-means and DeepLabV3+ algorithm. However, the GrabCut method lost many image details, resulting in smooth tree-crown edges. This experiment shows that the proposed method could accurately identify tree crowns and obtain more image details than other algorithms.

Table 1 shows the average segmentation result and the computational cost of 200 test images using different segmentation methods. The average P, R and F1 values of K-means are 58.1%, 79.7% and 68.9%, respectively. The segmentation accuracy of DeepLabV3+ is higher than that of K-means, with the average P increased by 24.3% and the average R reduced by 5.8%. The GrabCut algorithm's average P, R and F1 values have increased by 4.4%, 6.5% and 5.7% compared with K-means and 28.2%, 0.6% and 14.9% compared with DeepLabV3+. The results also show that the proposed method performs better than the remaining three methods with 92.1%, 94.5% and 93.3% of the average P, R and F1 values. K-means algorithm takes the lowest computational cost and does not require labeling data or model training. In contrast, the DeepLabV3+ algorithm needs massive model training and image annotations. GrabCut requires labeling all testing images, thereby taking the longest time. Hence, the proposed method in the study has a higher accuracy rate than traditional unsupervised algorithms and a lower computational cost than interactive algorithms and supervised algorithms.

Table 1. Average results for 200 images using different algorithms.

Method	Segmentation Evaluation Index			Computational Cost Assessment Index		
	Average P (%)	Average R (%)	Average F1 (%)	Average Time (s)	The Number of Labeled Images	Training Time-Consuming
K-means	58.1	79.7	68.9	0.366	-	-
DeaplabV3+	82.4	73.8	78.1	0.554	500	8 h
Grabcut	86.3	80.3	83.8	0.978	200	-
Proposed Algorithm	92.1	94.5	93.3	0.736	2	-

4.2. Performance Results under Different Overlapping Conditions and at Different Day Times

Different shooting angles may lead to two types of image samples, i.e., some crowns heavily overlapping with weeds, and crowns and weeds barely touching each other. Moreover, orchard images taken at different times of the day may have different exposure. For instance, images captured at noon, under strong sunlight, have the overexposure problem. This section compares the segmentation results of images taken under different overlapping conditions and in different times of the day. Figure 11 exemplifies cherry-tree images taken under a slightly overlapping condition (Figure 11a), partially overlapping condition (Figure 11b) and highly overlapping condition (Figure 11c), or in the morning (Figure 11d), at noon (Figure 11e) and in the evening (Figure 11f).
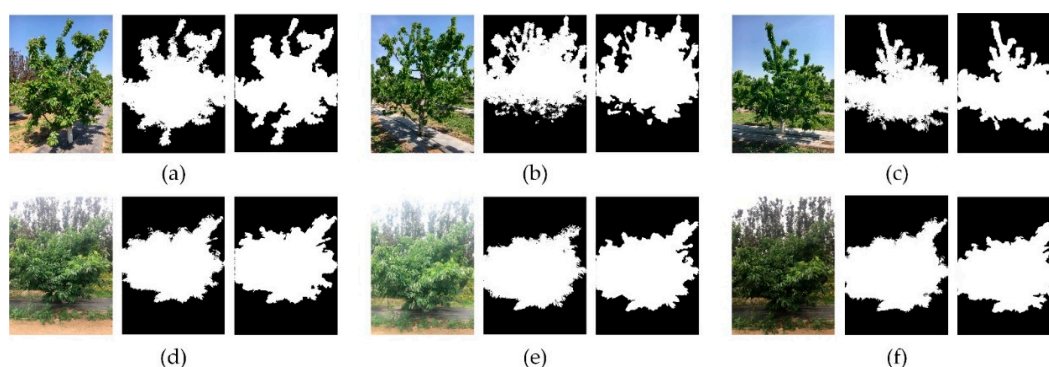

(a)　　　　　　　　　　(b)　　　　　　　　　　(c)

(d)　　　　　　　　　　(e)　　　　　　　　　　(f)

Figure 11. Comparison of segmentation results under different overlapping conditions and at different day times: (**a**) samples under a slightly overlapping condition; (**b**) samples under a partially overlapping condition; (**c**) samples under a highly overlapping condition; (**d**) samples collected in the morning; (**e**) samples collected at noon; and (**f**) samples collected in the evening.

Figure 11 shows that the tree crown was accurately extracted using the proposed method. Although the proposed method would miss some treetop leaves under high-exposure conditions (see Figure 11e) and lose some crown details under low-exposure conditions (see Figure 11d), the extraction results were relatively accurate. Table 2 illustrates the proposed segmentation method's average P, R and F1 values for 100 sample images. The average P, R and F1 values are 93.2%, 93.5% and 93.4%, respectively, with an accuracy rate over 90%. On the other hand, the average P value decreases significantly in the test set when the canopy and weeds overlap heavily. Hence, the overlapping degree between the canopy and weeds will affect the segmentation accuracy. Meanwhile, the average R values drop in the test set under both overexposure and underexposure conditions. Thus, the time of the image shooting mainly impacts the segmentation completeness. These discoveries underpin the proposed method's effectiveness in tree-crown recognition under different overlapping conditions and at different day times of image shooting.

Table 2. Average statistics for 100 images under different overlapping conditions and at different times of the day.

Indices	Overlapping Degrees			Image Shooting Day Times			In Total
	Slightly	Partly	Heavily	Morning	Noon	Evening	
Average P/%	94.9	92.9	90.8	94.1	93.9	93.1	93.2
Average R/%	95.3	94.9	93.8	94.9	90.4	92.1	93.5
Average F1/%	95.1	93.9	92.3	94.5	92.1	92.6	93.4

4.3. Segmentation Results in Different Years and Seasons

This section analyzes the proposed method's effectiveness in different years and seasons, and the image segmentation results are plotted in Figure 12. Figure 12a shows images and their segmentation results in 2018 spring and summer. Figure 12b displays images taken in 2019 spring and summer. Figure 12c only contains images in 2020 autumn, as the COVID-19 pandemic interrupted the image acquisition [57]. The highlighted patches show the image acquisition data. The test results of the datasets show that the proposed method can satisfy the segmentation of images taken in different seasons and years. To carry out quantitative verification, this study has analyzed a total of 150 sets of tree images that were taken continuously in 2018, 2019 and 2020.
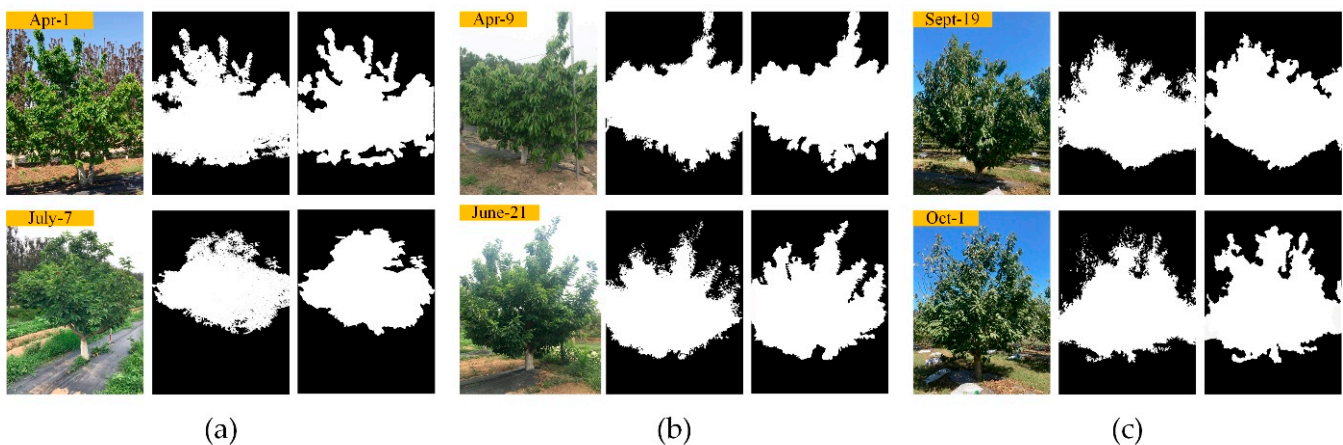

(a)   (b)   (c)

Figure 12. Segmentation results of the proposed method in different seasons and years: (**a**) samples in 2018 spring and summer; (**b**) samples in 2019 spring and summer; and (**c**) samples in 2020 autumn.

Figure 13 shows that season has a greater impact on the segmentation effect than the growth year. The average F1 value in spring is higher than in other seasons because the images taken in spring have brighter colors and fewer weeds. Images taken in autumn have the lowest segmentation accuracy and segmentation completeness because canopy characteristics have changed in autumn, with changed leaf color and scattered crown. Overall, the average P, R and F1 values in 2018, 2019 and 2020 are consistently above 87.7%. The results indicate that the proposed method is robust in canopy recognition in different years and seasons.
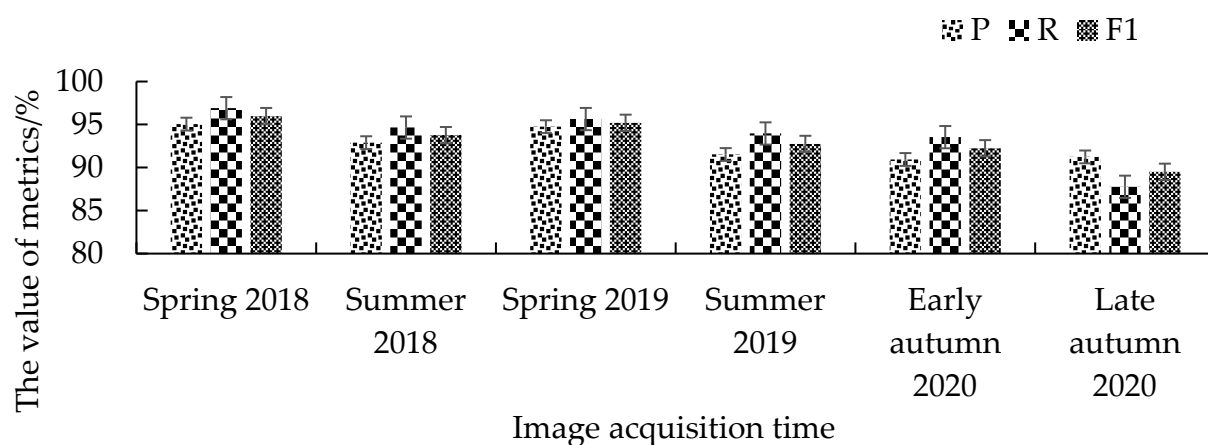
Figure 13. Average statistics of the proposed method for 200 sample images.

5. Conclusions

This study contrived to extract cherry-tree crowns from the complex background properly. The proposed method takes three stages: to compute Mahalanobis distance from the image's color, brightness and location features to acquire an initial classification of the canopy and background; to create a conditional random field, using the Mahalanobis distance calculations as unary potential energy and the Gaussian kernel function based on the image color and pixels distance as binary potential energy; and, finally, to complete the image segmentation, using mean-field approximation. In comparison with other methods, the proposed method has the highest average P, R and F1-score values, i.e., 92.1%, 94.5% and 93.3%, respectively, which were 34%, 14.8% and 24.2% higher than that of K-means traditional supervised algorithms. Compared with Grabcut interactive segmentation algorithms and DeepLabV3+ deep learning algorithm, the proposed method has lower image annotation and model training costs. The study also verified the feasibility and validity of the proposed method under different overlapping conditions, at different times of image acquisition, and in different years and seasons, and their results indicate that the overlapping conditions mainly affect the accuracy of the algorithm, but the image acquisition time affects the completeness of segmentation. The result also demonstrates that the season has a greater impact on the segmentation effect than the growth year. In a nutshell, the proposed method can outstand different environmental conditions, with the overall average P, R and F1 values higher than 87.7%. This study has exemplified that computer vision technology has great potential in crop identification. Future work will test the proposed method's application on other orchard tree crops and study new techniques that do not require data labeling.

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

References

1. Cisternas, I.; Velásquez, I.; Caro, A.; Rodríguez, A. Systematic literature review of implementations of precision agriculture. *Comput. Electron. Agric.* **2020**, *176*, 105626. [CrossRef]
2. Miles, C. The combine will tell the truth: On precision agriculture and algorithmic rationality. *Big Data Soc.* **2019**, *6*, 1–12. [CrossRef]
3. Shafi, U.; Mumtaz, R.; García-Nieto, J.; Hassan, S.A.; Zaidi, S.A.R.; Iqbal, N. Precision agriculture techniques and practices: From considerations to applications. *Sensors* **2019**, *19*, 3796. [CrossRef] [PubMed]
4. Xiao, K.; Ma, Y.; Gao, G. An intelligent precision orchard pesticide spray technique based on the depth-of-field extraction algorithm. *Comput. Electron. Agric.* **2017**, *133*, 30–36. [CrossRef]
5. Solanelles, F.; Escolà, A.; Planas, S.; Rosell, J.R.; Camp, F.; Gràcia, F. An electronic control system for pesticide application proportional to the canopy width of tree crops. *Biosyst. Eng.* **2006**, *95*, 473–481. [CrossRef]
6. Miranda-Fuentes, A.; Llorens, J.; Rodríguez-Lizana, A.; Cuenca, A.; Gil, E.; Blanco-Roldán, G.L.; Gil-Ribes, J.A. Assessing the optimal liquid volume to be sprayed on isolated olive trees according to their canopy volumes. *Sci. Total Environ.* **2016**, *568*, 269–305. [CrossRef] [PubMed]
7. Tona, E.; Calcante, A.; Oberti, R. The profitability of precision spraying on specialty crops: A technical-economic analysis of protection equipment at increasing technological levels. *Precis. Agric.* **2018**, *19*, 606–629. [CrossRef]
8. Zhang, Z.; Wang, X.; Lai, Q.; Zhang, Z. Review of variable-rate sprayer applications based on real-time sensor technologies. In *Automation in Agriculture—Securing Food Supplies for Future Generations*; Hussmann, S., Ed.; London, UK, 2018; pp. 53–79. [CrossRef]
9. Pallottino, F.; Antonucci, F.; Costa, C.; Bisaglia, C.; Figorilli, S.; Menesatti, P. Optoelectronic proximal sensing vehicle-mounted technologies in precision agriculture: A review. *Comput. Electron. Agric.* **2019**, *162*, 859–873. [CrossRef]
10. Virlet, N.; Gomez-Candon, D.; Lebourgeois, V.; Martinez, S.; Jolivot, A.; Lauri, P.E.; Costes, E.; Labbe, S.; Regnard, J.L. Contribution of high-resolution remotely sensed thermal-infrared imagery to high-throughput field phenotyping of an apple progeny submitted to water constraints. *Acta Hortic.* **2016**, *1127*, 243–250. [CrossRef]
11. Jurado, J.M.; Ortega, L.; Cubillas, J.J.; Feito, F.R. Multispectral mapping on 3D models and multi-temporal monitoring for individual characterization of olive trees. *Remote Sens.* **2020**, *12*, 1106. [CrossRef]
12. Ma, X.; Meng, Q.; Zhang, L.; Liu, G.; Zhou, W. Image mosaics reconstruction of canopy organ morphology of apple trees. *Nongye Gongcheng Xuebao Trans. Chin. Soc. Agric. Eng.* **2014**, *30*, 154–162.
13. Dong, W.; Isler, V. Tree morphology for phenotyping from semantics-based mapping in orchard environments. *arXiv* **2018**, arXiv:1804.05905.
14. Xu, H.; Ying, Y. Detecting citrus in a tree canopy using infrared thermal imaging. In *Monitoring Food Safety, Agriculture, and Plant Health*; International Society for Optics and Photonics: San Diego, CA, USA, 2004; Volume 5271, pp. 321–327. [CrossRef]
15. Coupel-Ledru, A.; Pallas, B.; Delalande, M.; Boudon, F.; Carrié, E.; Martinez, S.; Regnard, J.L.; Costes, E. Multi-scale high-throughput phenotyping of apple architectural and functional traits in orchard reveals genotypic variability under contrasted watering regimes. *Hortic. Res.* **2019**, *6*, 52. [CrossRef]
16. Pusdá-Chulde, M.R.; Salazar-Fierro, F.A.; Sandoval-Pillajo, L.; Herrera-Granda, E.P.; García-Santillán, I.D.; de Giusti, A. Image analysis based on heterogeneous architectures for precision agriculture: A systematic literature review. *Adv. Intell. Syst. Comput.* **2020**, *1078*, 51–70. [CrossRef]
17. Moreno, W.F.; Tangarife, H.I.; Escobar Díaz, A. Image analysis aplications in precision agriculture. *Visión Electrónica* **2018**, *11*, 200–210. [CrossRef]
18. Delgado-Vera, C.; Mite-Baidal, K.; Gomez-Chabla, R.; Solís-Avilés, E.; Merchán-Benavides, S.; Rodríguez, A. Use of technologies of image recognition in agriculture: Systematic review of literature. *Commun. Comput. Inform. Sci.* **2018**, *883*, 15–29. [CrossRef]
19. Hernández-Hernández, J.L.; García-Mateos, G.; González-Esquiva, J.M.; Escarabajal-Henarejos, D.; Ruiz-Canales, A.; Molina-Martínez, J.M. Optimal color space selection method for plant/soil segmentation in agriculture. *Comput. Electron. Agric.* **2016**, *122*, 124–132. [CrossRef]
20. Hamuda, E.; Glavin, M.; Jones, E. A survey of image processing techniques for plant extraction and segmentation in the field. *Comput. Electron. Agric.* **2016**, *125*, 184–199. [CrossRef]
21. Xue, J.; Su, B. Significant remote sensing vegetation indices: A review of developments and applications. *J. Sens.* **2017**, *1*, 1–17. [CrossRef]
22. Woebbecke, D.M.; Meyer, G.E.; von Bargen, K.; Mortensen, D.A. Color indices for weed identification under various soil, residue, and lighting conditions. *Trans. Am. Soc. Agric. Eng.* **1995**, *38*, 259–269. [CrossRef]
23. Meyer, G.E.; Neto, J.C.; Jones, D.D.; Hindman, T.W. Intensified fuzzy clusters for classifying plant, soil, and residue regions of interest from color images. *Comput. Electron. Agric.* **2004**, *42*, 161–180. [CrossRef]
24. Weier, J.; Herring, D. *Measuring Vegetation (NDVI & EVI)*; Normalized Difference Vegetation Index (NDVI); Nasa Earth Observatory: Washington, DC, USA, 2011.
25. Tucker, C.J. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens. Environ.* **1979**, *8*, 127–150. [CrossRef]

26. Hassaan, O.; 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]

27. Wang, A.; Zhang, W.; Wei, X. A review on weed detection using ground-based machine vision and image processing techniques. *Comput. Electron. Agric.* **2019**, *158*, 226–240. [CrossRef]

28. Mu, Y.; Fujii, Y.; Takata, D.; Zheng, B.; Noshita, K.; Honda, K.; Ninomiya, S.; Guo, W. Characterization of peach tree crown by using high-resolution images from an unmanned aerial vehicle. *Hortic. Res.* **2018**, *5*, 74. [CrossRef]

29. Dong, X.; Zhang, Z.; Yu, R.; Tian, Q.; Zhu, X. Extraction of information about individual trees from high-spatial-resolution UAV-acquired images of an orchard. *Remote Sens.* **2020**, *12*, 133. [CrossRef]

30. Chen, Y.R.; Chao, K.; Kim, M.S. Machine vision technology for agricultural applications. *Comput. Electron. Agric.* **2002**, *36*, 173–191. [CrossRef]

31. Patrício, D.I.; Rieder, R. Computer vision and artificial intelligence in precision agriculture for grain crops: A systematic review. *Comput. Electron. Agric.* **2018**, *153*, 69–81. [CrossRef]

32. Sridharan, M.; Gowda, P. Application of statistical machine learning algorithms in precision agriculture. In Proceedings of the 7th Asian-Australasian Conference on Precision Agriculture, Hamilton, New Zealand, 16–18 October 2017. [CrossRef]

33. Liu, W.; Zhong, T.; Song, Y. Prediction of trees diameter at breast height based on unmanned aerial vehicle image analysis. *Nongye Gongcheng Xuebao Trans. Chin. Soc. Agric. Eng.* **2017**, *33*, 99–104.

34. Qi, L.; Cheng, Y.; Cheng, Z.; Yang, Z.; Wu, Y.; Ge, L. Estimation of upper and lower canopy volume ratio of fruit trees based on M-K clustering. *Nongye Jixie Xuebao Trans. Chin. Soc. Agric. Mach.* **2018**, *49*, 57–64.

35. Abdalla, A.; Cen, H.; El-Manawy, A.; He, Y. Infield oilseed rape images segmentation via improved unsupervised learning models combined with supreme color features. *Comput. Electron. Agric.* **2019**, *162*, 1057–1068. [CrossRef]

36. Rehman, T.U.; Mahmud, M.S.; Chang, Y.K.; Jin, J.; Shin, J. Current and future applications of statistical machine learning algorithms for agricultural machine vision systems. *Comput. Electron. Agric.* **2019**, *156*, 585–605. [CrossRef]

37. Elavarasan, D.; Vincent, D.R.; Sharma, V.; Zomaya, A.Y.; Srinivasan, K. Forecasting yield by integrating agrarian factors and machine learning models: A survey. *Comput. Electron. Agric.* **2018**, *155*, 257–282. [CrossRef]

38. Valente, J.; Doldersum, M.; Roers, C.; Kooistra, L. Detecting rumex obtusifolius weed plants in grasslands from UAV RGB imagery using deep learning. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2019**, *4*, 179–185. [CrossRef]

39. Chen, Y.; Hou, C.; Tang, Y.; Zhuang, J.; Lin, J.; He, Y.; Guo, Q.; Zhong, Z.; Lei, H.; Luo, S. Citrus tree segmentation from UAV images based on monocular machine vision in a natural orchard environment. *Sensors* **2019**, *19*, 5558. [CrossRef] [PubMed]

40. 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. *Drones* **2018**, *11*, 1–7.

41. Wu, J.; Yang, G.; Yang, H.; Zhu, Y.; Li, Z.; Lei, L.; Zhao, C. Extracting apple tree crown information from remote imagery using deep learning. *Comput. Electron. Agric.* **2020**, *174*, 105504. [CrossRef]

42. Lu, Y.; Young, S. A survey of public datasets for computer vision tasks in precision agriculture. *Comput. Electron. Agric.* **2020**, *178*, 105760. [CrossRef]

43. Cheng, Z.; Qi, L.; Cheng, Y.; Wu, Y.; Zhang, H. Interlacing orchard canopy separation and assessment using UAV images. *Remote Sens.* **2001**, *34*, 2259–2281. [CrossRef]

44. Cheng, H.D.; Jiang, X.H.; Sun, Y.; Wang, J. Color image segmentation: Advances and prospects. *Pattern Recognit.* **2001**, *34*, 2259–2281. [CrossRef]

45. Hamuda, E.; Ginley, B.M.; Glavin, M.; Jones, E. Automatic crop detection under field conditions using the HSV colour space and morphological operations. *Comput. Electron. Agric.* **2017**, *133*, 97–107. [CrossRef]

46. Kozlowski, T.T.; Pallardy, S.G. Environmental regulation of vegetative growth. In *Growth Control in Woody Plants*; Academic Press: Cambridge, MA, USA, 1997; pp. 195–322.

47. García-Santillán, I.D.; Pajares, G. On-line crop/weed discrimination through the Mahalanobis distance from images in maize fields. *Biosyst. Eng.* **2018**, *166*, 28–43. [CrossRef]

48. Lafferty, J.; Andrew, M.; Fernando, C.N.P. Conditional random fields: Probabilistic models for segmenting and labeling sequence data. In Proceedings of the 18th International Conference on Machine Learning, Williamstown, MA, USA, 28 June–1 July 2001; pp. 282–289. [CrossRef]

49. Krähenbühl, P.; Koltun, V. Efficient inference in fully connected crfs with Gaussian edge potentials. In Proceedings of the Advances in Neural Information Processing Systems 24: 25th Annual Conference on Neural Information Processing Systems, Granada, Spain, 12–14 December 2011; pp. 109–117.

50. Liu, T.; Huang, X.; Ma, J. Conditional random fields for image labeling. *Math. Probl. Eng.* **2016**, *2016*, 1–15. [CrossRef]

51. Cheng, Z.; Qi, L.; Cheng, Y.; Wu, Y.; Zhang, H.; Xiao, Y. Fruit tree canopy image segmentation method based on M-LP features weighted clustering. *Nongye Jixie Xuebao Trans. Chin. Soc. Agric. Mach.* **2020**, *51*, 191–198.

52. Liu, H.; Zhu, S.; Shen, Y.; Tang, J. Fast segmentation algorithm of tree trunks based on multi-feature fusion. *Nongye Jixie Xuebao Trans. Chin. Soc. Agric. Mach.* **2020**, *51*, 221–229.

53. Ferreira, M.P.; de Almeida, D.R.A.; de Almeida, D.P.; Minervino, J.B.S.; Veras, H.F.P.; Formighieri, A.; Santos, C.A.N.; Ferreira, M.A.D.; Figueiredo, E.O.; Ferreira, E.J.L. Individual tree detection and species classification of Amazonian palms using UAV images and deep learning. *For. Ecol. Manag.* **2020**, *475*, 118397. [CrossRef]

54. Selim, S.Z.; Ismail, M.A. K-means-type algorithms: A generalized convergence theorem and characterization of local optimality. *IEEE Trans. Pattern Anal. Mach. Intell.* **1984**, *6*, 81–87. [CrossRef] [PubMed]

55. Chen, L.C.; Zhu, Y.; Papandreou, G.; Schroff, F.; Adam, H. Encoder-decoder with atrous separable convolution for semantic image segmentation. *Lect. Notes Comput. Sci.* **2018**, 833–8521. [CrossRef]

56. Rother, C.; Kolmogorov, V.; Blake, A. "GrabCut"—Interactive foreground extraction using iterated graph cuts. *ACM Trans. Graph.* **2004**, *23*, 309–314. [CrossRef]

57. WHO. *Statement on the Second Meeting of the International Health Regulations (2005) Emergency Committee Regarding the Outbreak of Novel Coronavirus (2019-nCoV)*; WHO: Geneva, Switzerland, 2020.