Deep Learning with Unsupervised Data Labeling for Weed Detection in Line Crops in UAV Images

M Dian Bah [1], Adel Hafiane [2],* and Raphael Canals [1]

[1]   University of Orleans, PRISME, EA 4229, F45072 Orleans, France; m-dian.bah@univ-orleans.fr (M.D.B.);
      raphael.canals@univ-orleans.fr (R.C.)
[2]   INSA Centre Val de Loire, PRISME, EA 4229, F18020 Bourges, France
*   Correspondence: adel.hafiane@insa-cvl.fr

check for
updates

Abstract: In recent years, weeds have been responsible for most agricultural yield losses. To deal with this threat, farmers resort to spraying the fields uniformly with herbicides. This method not only requires huge quantities of herbicides but impacts the environment and human health. One way to reduce the cost and environmental impact is to allocate the right doses of herbicide to the right place and at the right time (precision agriculture). Nowadays, unmanned aerial vehicles (UAVs) are becoming an interesting acquisition system for weed localization and management due to their ability to obtain images of the entire agricultural field with a very high spatial resolution and at a low cost. However, despite significant advances in UAV acquisition systems, the automatic detection of weeds remains a challenging problem because of their strong similarity to the crops. Recently, a deep learning approach has shown impressive results in different complex classification problems. However, this approach needs a certain amount of training data, and creating large agricultural datasets with pixel-level annotations by an expert is an extremely time-consuming task. In this paper, we propose a novel fully automatic learning method using convolutional neuronal networks (CNNs) with an unsupervised training dataset collection for weed detection from UAV images. The proposed method comprises three main phases. First, we automatically detect the crop rows and use them to identify the inter-row weeds. In the second phase, inter-row weeds are used to constitute the training dataset. Finally, we perform CNNs on this dataset to build a model able to detect the crop and the weeds in the images. The results obtained are comparable to those of traditional supervised training data labeling, with differences in accuracy of 1.5% in the spinach field and 6% in the bean field.

Keywords: weed detection; deep learning; unmanned aerial vehicle; image processing; precision agriculture; crop line detection

1. Introduction

Currently, losses due to pests, diseases, and weeds can reach 40% of global crop yields each year, and this percentage is expected to increase significantly in the coming years [1]. Common weed control practices consist of spraying the entire field with herbicides, a practice that involves significant waste and cost for farmers and that causes environmental pollution [2]. In order to reduce the volume of chemicals while continuing to increase productivity, the concept of precision agriculture was introduced [3,4]. Precision agriculture can be defined as the application of technology for the purpose of improving crop performance and environmental quality [3]. The main goal of precision agriculture is to select the right management practice in order to allocate the right doses of inputs, such as fertilizers, herbicides, seed, fuel, etc., to the right place and at the right time [5]. Weed detection and characterization represent one of the major challenges of precision agriculture, since, in current

farming practice, herbicides are typically applied uniformly across fields, despite the fact that weeds exhibit uneven spatial distributions.

In the literature, several methods of weed detection are proposed with different acquisition systems [6–8]. Compared to robot or satellite acquisitions, drones have been considered more efficient since they allow a fast acquisition of the field with very high spatial resolution and at a low cost [9,10]. Despite significant advances in unmanned aerial vehicle (UAV) acquisition systems, the automatic detection of weeds remains a challenging problem. In recent years, deep learning techniques have shown a dramatic improvement for many computer vision tasks, and recent developments have shown the importance of these techniques for weed detection [11,12]. They are still not widely used in agriculture, however, as the huge quantities of data required in the learning phase have accentuated the problem of the manual annotation of these datasets. The same problem arises in agriculture data, where labeling plants in a field image is very time consuming. So far, very little attention has been paid to the unsupervised annotation of data to train deep learning models, particularly for agriculture.

In this article, we propose a fully automatic method for weed detection in drone images. Our method is based on the unsupervised collection of a training data set and convolutional neuronal networks (CNNs). The proposed method is performed in three steps. First, we detect crop rows and exploit them to detect inter-row weeds. In the second step, these inter-row weeds are used to form the training dataset. Finally, the database created in the previous step is used to generate a model from deep learning. The advantage of this method is that it is adaptive and robust, which means that it is possible not only to use the generated model to detect weeds in a new field with the same crop type, but also to generate a new model by applying this method in a new field without any feature selection methods.

The paper is divided into five parts. In Section 2, we discuss related work. Section 3 presents the proposed method. In Section 4, we comment on and discuss the experimental results obtained. Section 5 concludes the paper.

2. Related Work

In the literature, several approaches have been used to detect weeds with different acquisition systems. The main approach for weed detection is to extract vegetation from the image using segmentation and then to discriminate crop and weeds. Common segmentation approaches use color and multispectral information to separate vegetation and background (soil and residues). Specific indices are calculated from this information to segment the vegetation effectively [13].

However, weeds and crop are hard to discriminate using spectral information because of their strong similarities. Regional approaches and a spatial arrangement of pixels are preferred in most cases. In [14], the Excess Green Vegetation Index (ExG) [15] and Otsu's thresholding [16] helped to remove background (soil, residues) before performing a double Hough transform [17] in order to identify the main crop lines in perspective images. Then, to discriminate crop and weeds in the segmented image, the authors applied a region-based segmentation method to develop a blob coloring analysis. Thus, any region with at least one pixel belonging to the detected rows is considered to be crop; otherwise, it is weeds. Unfortunately, this technique failed to handle weeds that were close to crop regions. In [18], an object-based image analysis (OBIA) procedure was developed based on a series of UAV images for the automatic discrimination of crop rows and weeds in a maize field. The UAV images were segmented into homogeneous multi-pixel objects using the multiscale algorithm [19]. The large scale highlighted the structures of crop rows and the small scale brought out objects that lay within the crop rows. The authors found that the process was strongly affected by the presence of weeds very close to or within the crop rows.

In [20], 2-D Gabor filters were applied to extract the features, and an artificial neural network (ANN) was used to classify broadleaf and grass weeds. Their results showed that joint space-frequency texture features have the potential for weed classification. In [21], the authors relied on morphological variation and used neural network analysis to separate weeds from a maize crop. Support vector machine (SVM) and shape features were proposed for the effective classification of crops and weeds

in digital images in [22]. In their experiment, a total of 14 features that characterized crops and weeds in the images were tested to find the optimal combination of features which provided the highest classification rate. Latha et al. [23] suggested that, in the image, edge frequencies and veins of the crop and the weeds have different density properties (strong and weak edges) that could be used to separate crop from weed. A semi-supervised method was proposed in [24] to discriminate weeds and crop. The Ostu thresholding was applied twice on ExG. In the first step, the authors used segmentation to remove the background and then created two classes considered to be crop and weeds. K-means clustering was used to select 100 samples in each class for the training. Classification was then performed with an SVM classifier using the geometric features, spatial features, and first- and second-order statistics extracted in the red, blue, green, and ExG bands. The method proved to be effective in the sunflower field, but less robust in the corn field because of the shade produced by the corn plants. In [25], the authors used texture features extracted from wavelet subimages to detect and characterize four types of weeds in a sugar beet field. Neural networks were applied as the classifier. The use of wavelets proved to be efficient for the detection of weeds, even at a stage of growth of beet greater than six leaves. Bakhshipour and Jafari [26] evaluated weed detection with support vector machine and artificial neural networks in four species of common weeds in sugar beet fields using shape features. In [27], a semiautomatic object-based image analysis (OBIA) procedure was developed with random forests (RF) combined with feature selection techniques to classify soil, weeds, and maize. A common point in all these studies is that the selected features change, in general, from one type of crop to another or from one type of data to another.

Recently, convolutional neural networks have emerged as a powerful approach to computer vision tasks. CNNs [28] have progressed mostly through its successful use as a method in the ImageNet Large-Scale Vision Recognition Challenge 2012 (ILSCVR12) and the creation of the AlexNet network in 2012, which showed that a large, deep convolutional neural network is capable of achieving record-breaking results on a highly challenging dataset using purely supervised training [29]. Nowadays, deep learning is applied to several domains to help solve many big data problems, such as computer vision, speech recognition, and natural language processing. In the agriculture domain, CNNs were applied to classify patches of water hyacinth, serrated tussock, and tropical soda apple in [30]. Mortensen et al. [31] used CNNs for semantic segmentation in the context of mixed crops from images of an oil radish plot trial with barley, grass, weed, stump, and soil. Milioto et al. [32] provided accurate weed classification in real sugar beet fields with mobile agricultural robots. Dos Santos Ferreira et al. [11] applied AlexNet to the detection of weeds in soybean crops. In [12], AlexNet was applied to weed detection in different crop fields, such as beet, spinach, and bean in UAV imagery.

The main common point between the supervised machine learning algorithms is the need for training data. For a good optimization of deep learning models, it is necessary to have a certain amount of labeled data. However, as mentioned previously, creating large agricultural datasets with pixel-level annotations is an extremely time-consuming task. Few attempts have been made to develop fully automatic systems for the training and identification of weeds in agricultural fields. In a recent study, Di Cicco et al. [33] suggested the use of synthetic training datasets. However, this technique requires precise modeling in terms of texture, 3D models, and light conditions. In [34], an automatic image processing method was developed to discriminate between crop and weed pixels by combining spatial and spectral information extracted from four-band multispectral images. Image data were captured at 3 m above ground with a camera mounted on a manually held pole. The spatial approach (Hough transform) was used to detect crop rows and to build a training dataset. SVM was applied to the spectral information to perform classification. This method assumes that weeds and crops have different spectral information, which is not always the case in agricultural fields. The success of this kind of method relies on better feature selection which involves human analysis of each particular field. To the best of our knowledge, no studies have been carried out on weed detection in UAV images using automatic labeling of training images and deep learning.

3. Proposed Method

In modern agriculture, most crops are grown in regular rows, separated by a defined space that depends on the type of crop. Generally, plants that grow outside the rows are considered weeds, commonly referred to as inter-row weeds. Several studies have used this assumption to locate weeds using the geometric properties of the rows [35]. The main advantage of this technique is that it is unsupervised and does not depend on the training data. Based on this hypothesis, we first detected the crop rows, then inter-row vegetation was used to constitute our training database, with data categorized into two classes: crop and weed. Thereafter, we performed CNNs on this database to build a model able to detect the crop and weeds in the images. The flowchart in Figure 1 depicts the main steps of the proposed method. The following sections describe each step in detail.
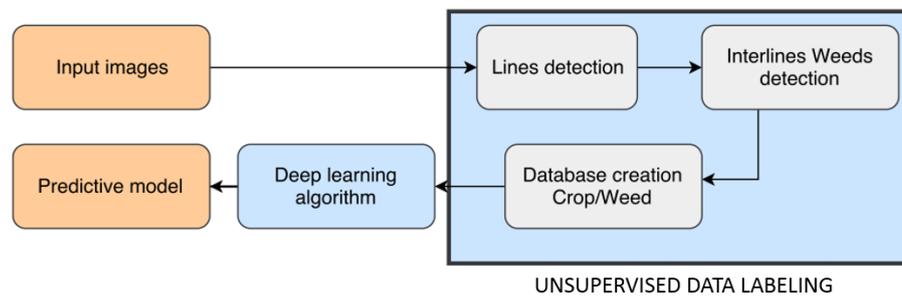

Figure 1. Flowchart of the proposed method.

3.1. Detection of Crop Lines

A crop row can be defined as a composition of several parallel lines. The aim of this step is to detect the main line of each crop row. For that purpose, we used the Hough transform to highlight alignments of the pixels. In Hough space, there is one cell per line, which means that cells are aggregated by crop row. The main lines in Hough space correspond to cells which contain the maximum number of votes on each aggregation. Before starting any line detection procedure, generally, preprocessing is required to remove undesirable perturbations, such as shadows, soil, or stones. Here, we used the ExG (Equation (1)) with Otsu adaptive thresholding to discriminate between vegetation and background.

$$ExG = 2g - r - b \tag{1}$$

where r, g, and b are the normalized RGB color values.

The Hough transform is one of the most widely used methods for line detection, and it is often integrated into tools to guide agricultural machines because of its robustness and ability to adjust discontinuous lines caused by missing crop plants in a row or by poor germination [36]. Usually, for crop line detection, the Hough transform is directly applied to the segmented image. This procedure is computationally expensive and depends on the density of vegetation in the crop rows. There is also a risk of line overdetection. We addressed this problem by using the skeleton of each row, which is an approach that showed better performance in our previous study [37]. We found that the Hough transform applied to the skeleton gave a good overall detection rate of crop lines, close to 100%, and a low overdetection rate even for images with high infestation rates. We also discovered that the skeleton provided a good overall representation of the field structure, namely, orientations and periodicity. The Hough transform $H(\theta, \rho)$ was computed on the skeleton with a θ resolution of 0.1°, letting θ take values in the range $[-90°; 90°]$, and a ρ resolution of 1. Thanks to a histogram of the skeletons' directions, the most frequently represented angle was chosen as the main orientation θ_{lines} of crop lines. $H(\theta, \rho)$ was normalized to $H_{norm}(\theta, \rho)$ in order to give the same weight to all the crop lines, especially the short ones close to the borders of the image [14]. $H_{norm}(\theta, \rho)$ is defined as the ratio between the accumulator of the vegetation image and the accumulator of a totally white image of the same size $H_{ones}(\theta, \rho)$. To disregard

the small lines created by the aggregation of weeds in the inter-row space, a threshold of 0.1 was applied to the normalized Hough transform. Moreover, in modern agriculture, crops are usually sown in parallel lines with the same inter-row distance, which means that the main peaks corresponding to the crop lines are aligned around an angle in the Hough space with the same gaps. Unfortunately, because of the realities in the agricultural field, lines are not perfectly parallel; thus, peaks in the Hough space have close but different angles, and the inter-row distance is not constant. In order to avoid skipping any crop line during the detection, the lines were kept if they had a peak in Hough space whose angle did not exceed $20°$ compared to the overall orientation (θ_{lines}) of the lines. Figure 2 presents the flowchart of the line detection method. However, to avoid detecting more than one peak in an aggregation (i.e., to reduce overdetection), whenever a peak of a crop row was spotted in $H_{norm}(\theta, \rho)$, we identified the corresponding skeleton, and then we deleted the votes of this skeleton in $H_{norm}(\theta, \rho)$ before continuing. All the steps are summarized in Algorithm 1.
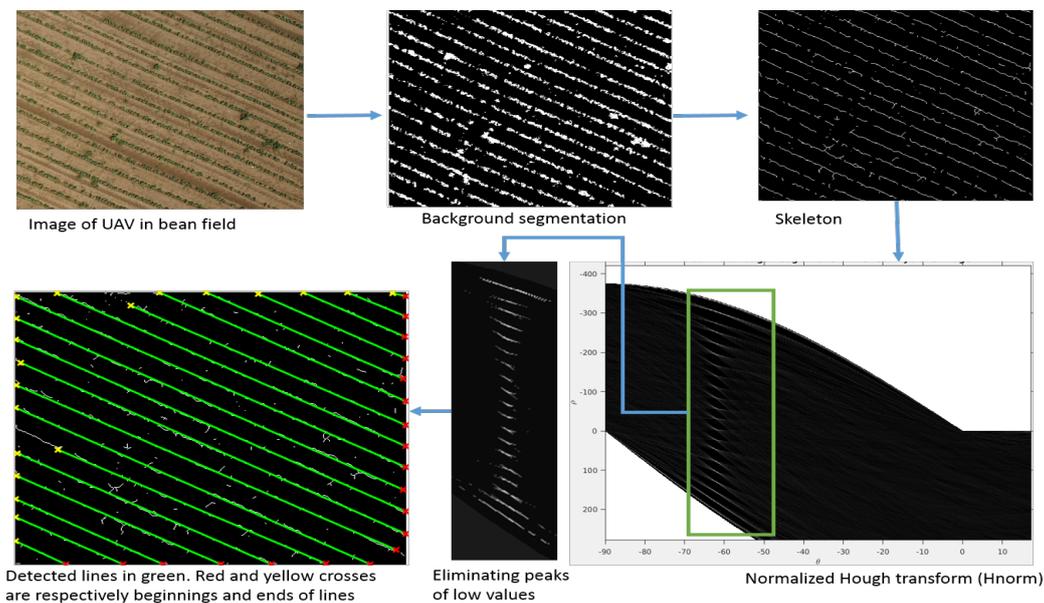

Image of UAV in bean field   Background segmentation   Skeleton

Detected lines in green. Red and yellow crosses are respectively beginnings and ends of lines   Eliminating peaks of low values   Normalized Hough transform (Hnorm)

Figure 2. Flowchart of crop line detection method.

Algorithm 1: Crop line detection.

---

    input :skeletons
    output:crop lines
1 Computation of the skeleton angle
2 Computation of the main orientation θ_{lines} of the crop lines
3 Hough transform of the skeletons $H(\theta, \rho)$
4 $H_{norm}(\theta, \rho)=H(\theta, \rho)/\text{Hones}(\theta, \rho)$
5 **while** *maximum of $H_{norm}(\theta, \rho) > 0.1$* **do**
6   Computation of the maximum of $H_{norm}(\theta, \rho)$ and the corresponding angle θ_m
7   Recovery of the line corresponding to the maximum ($Line_{skeleton}$)
8   Computation of the normalized Hough transform ($H_{temp}(\theta, \rho)$) of the $Line_{skeleton}$
9      $H_{norm}(\theta, \rho)=H_{norm}(\theta, \rho) -H_{temp}(\theta, \rho)$
10   **if** $\theta_m > \theta_{lines} - 20°$ *and* $\theta_m < \theta_{lines} + 20°$ **then**
11 | The detected line is a crop line
   **end**
   end

3.2. Unsupervised Training Data Labeling

The unsupervised training dataset annotation is based on the detected lines obtained following the procedure in the previous section. Assuming that the lines detected are mainly at the center of the crop rows (Figure 3), we applied a mask to delimit the crop rows. Hence, vegetation overlapped by the mask corresponds to the crop. This mask was obtained from the intersection of superpixels formed by the simple linear iterative clustering (SLIC) algorithm [38] and the detected lines. SLIC was chosen since it is simple and efficient in terms of the quality of results and the computation time. It is an adaptation of the k-means approach for superpixel generation, with a control for the size and compactness of the superpixels. SLIC creates a local grouping of pixels based on their spectral values, which are defined by the values of the CIELAB color space, and their spatial proximity [11,38]. A higher value of compactness makes superpixels more regularly shaped while a lower value makes superpixels adhere better to boundaries, making them irregularly shaped. Since here the goal was to create a mask around the detected crop lines that is able to delimit the crop rows, we chose a compactness of 20 because it was found that the process was less sensitive to variations of color caused by the effects of light and shadow. Figure 4 shows examples of images segmented with different sizes of superpixels.

Once the crop has been identified, the next step consists of detecting the inter-row weeds. An inter-row weed is defined as a plant growing between the crop lines. To detect weeds that lie in inter-rows, we applied a blob coloring algorithm. Hence, any region that does not intersect with the crop mask is regarded as a weed. Also, vegetation pixels which belong neither to the crop mask nor to the inter-row weeds are attributed to the potential weeds. Figure 5 presents the mask of crop, inter-row weeds, and potential weeds. To construct the training dataset, we extracted patches from the original images using positions of the detected inter-row weeds and crops. For weed samples, we applied bounding boxes to each segmented intra-row weed. For the crop samples, a sliding window was applied to the input image using positions relative to the segmented crop lines. Thus, for a given position of the window, if it intersects the binary mask and there are no inter-row weed pixels, it is attributed to the crop class. Generally, the crop class has many more samples than the weed one. In cases where there were few inter-row weed samples but a large number of potential weeds, as in Figure 5, we included the latter in the training dataset of weeds. Hence, the window which contained only potential weeds was labeled weeds. On the other hand, windows which contained crop and potential weeds, where we had more potential weeds than crop, were not retained.

(**a**)　　　　　　　　　　　　　　　　　　　　　　　　　　(**b**)
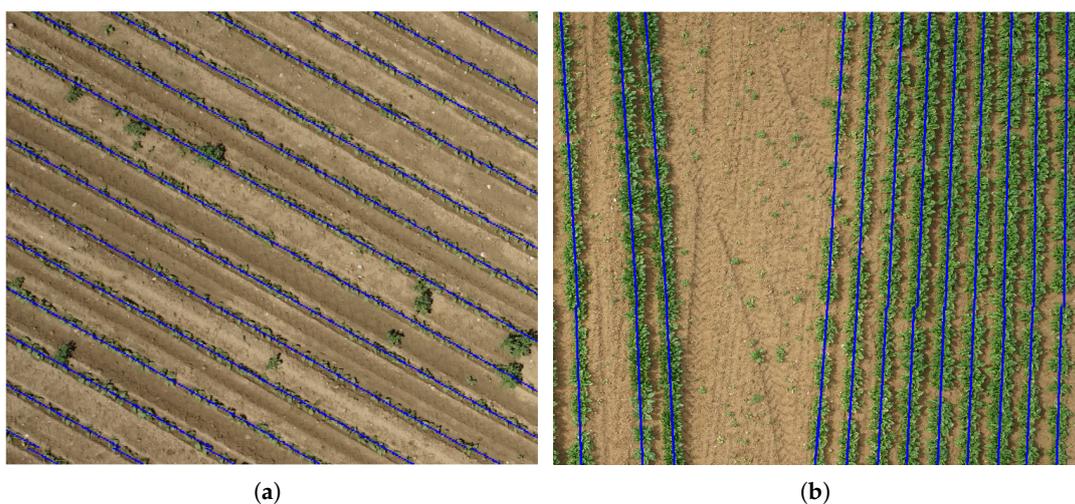
Figure 3. From left to right: line detection in bean (**a**) and spinach (**b**) fields. Detected lines are in blue. In the spinach field, inter-row distance and the crop row orientation are not regular. The detected lines are mainly located in the center of the crop rows.
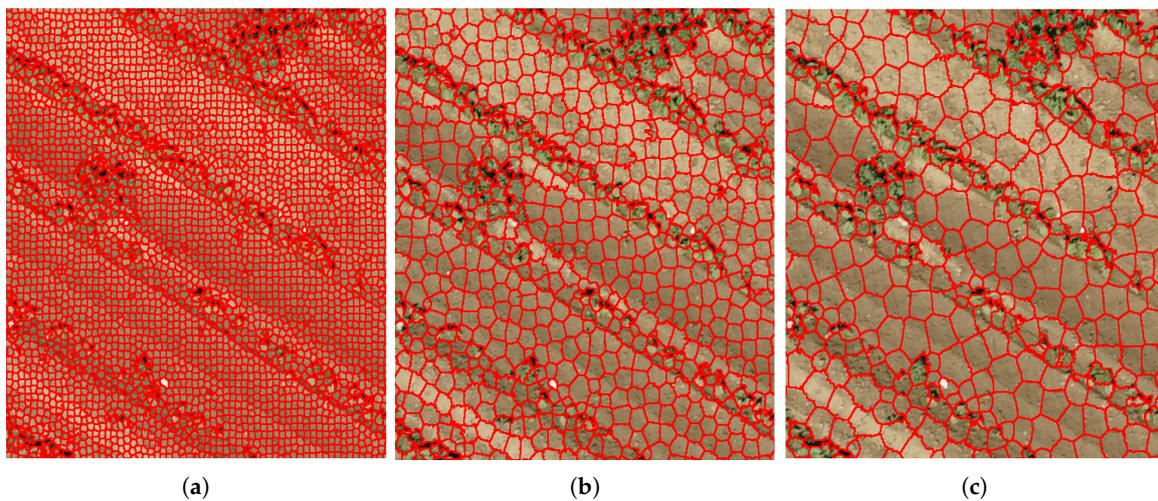
Figure 4. Examples of superpixels computed on images with the dimensions $N = 7360 \times 4912$ pixels. From left to right: the image is segmented with a number of superpixels equal to $0.5\% \times N$, $0.1\% \times N$, and $0.01\% \times N$, respectively.
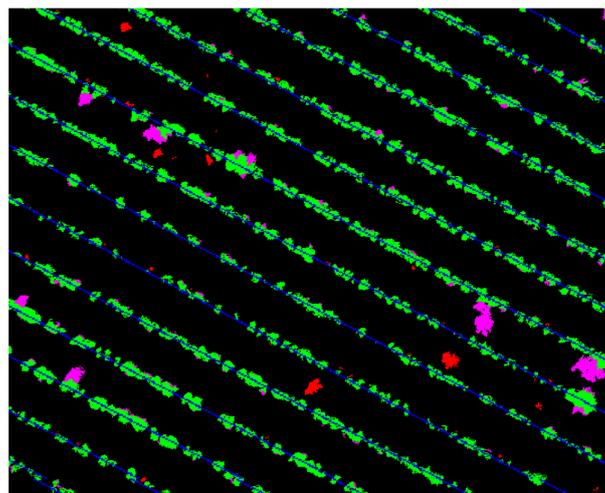
Figure 5. Detection of inter-row weeds (red) after line detection (blue) in a bean image. The crop mask is represented in green and the potential weeds in magenta.

3.3. Crop/Weed Classification Using Convolutional Neural Networks

CNNs are part of the deep learning approach and have shown an impressive performance in many computer vision tasks [39]. CNNs are made up of two types of layers: the convolutional layers which extract different characteristics from images, and the fully connected layers based on the multilayer perceptron to perform classification. The number of convolutional layers depends on the classification task and also the number and the size of the training data.

In this work, we used a Residual Network (ResNet). This network architecture was introduced in 2015 [40]. It won the ImageNet Large-Scale Vision Recognition Challenge 2015 with 152 layers. However, given the size of our data, we used the ResNet with 18 layers (ResNet18) described in [40] because it achieved a better result than AlexNet and VGG13 [41] in the ImageNet challenge.

However, given the number of parameters to be updated in ResNet18 and the data we had at our disposal, we decided to use transfer learning. Transfer learning aims to extract knowledge from one or more source tasks and applies this knowledge to a target task. In other words, transfer learning is

a machine learning method in which a model developed for one task is reused as the starting point for a model in a second task [42]. Transfer learning is the most popular approach in deep learning; models pretrained on a dataset such as ImageNet are used as the starting point to solve another problem in computer vision—weed and crop classification, in our case. Due to the large number of categories and images in ImageNet, some studies have shown that transfer learning of networks trained with the ImageNet database could be successfully used [31,43]. Thus, we performed a transfer learning technique called fine-tuning to train the networks with our data. Fine-tuning means that we started with the learned features from the ImageNet dataset, then we truncated the last layer (softmax layer) of the pretrained network and replaced it with new softmax layers that are relevant to our own problem. Here, the thousand categories of ImageNet were replaced by two categories (crop and weeds).

3.4. Feature Extraction

Although color indices make sense in distinguishing between vegetation and background, they become less effective when applied to classify plant species. Sometimes, the color of weeds and crop leaves look almost the same. Moreover, the result becomes unreliable under different lighting conditions. To solve this problem, several image features were analyzed. We computed a series of statistical features, shape features, and texture features which have been selected in other works [23–25,44]. A procedure for feature selection was then used to analyze the most suitable features.

3.4.1. Color Features

The color features are means and standard deviations of the three RGB image bands and of the ExG image. In order to make the color features consistent with different lighting levels, each color band was normalized by the sum of all three color bands.

3.4.2. Geometric Shape Features

Based on [22], three parameters, namely, Form Factor, Elongatedness, and Solidity, were computed as geometric features. We named the feature vector created by these three *Geo*3.

$$FormFactor = \frac{4 * \pi * area}{perimeter^2} \tag{2}$$

$$Elongatedness = \frac{area}{thickness^2} \tag{3}$$

$$Solidity = \frac{area}{convex_{area}} \tag{4}$$

Here, area is defined as the number of pixels with a value of '1' in the binary image. Perimeter is defined as the number of pixels with a value '1' for which at least one of the eight neighboring pixels has the value '0', implying that the perimeter is the number of border pixels. Convex area is the area of the smallest convex hull that covers all the plant pixels in an image.

3.4.3. Edge density

Edge detection is a method of image segmentation which uses the fact that the edge frequencies and veins of both crop and weeds have different density properties (strong and weak edges) to separate crop from weed [23]. In the remainder of this article, we denote edge density as *Edensity*. It is defined as:

$$Edensity = \frac{edge_{area}}{area} \tag{5}$$

Here, area is defined as the number of pixels with a value '1' in the binary image. The image edges were computed by the Sobel edge detection method. All the pixels marked as edge were summed, and their sum is called $edge_{area}$.

3.4.4. Histogram of Oriented Gradients (HOG)

Contour attributes generally correspond to the histogram of the gradient orientation. HOG counts the occurrences of gradient orientation in localized regions in an image. It is fast, compared to the SIFT algorithm (because no smoothing is computed); it is processed on a large number of cells uniformly spaced in the image and overlapping. Thanks to normalization of the local contrast, it is invariant to conditions of illumination. HOG was initially used for pedestrian detection [45], but it has proved its robustness for many other issues. In agriculture, it is used to identify plant leaves [46,47]. These experiments are inspiring and indicate that we can combine the features extracted by HOG methods for the classification of leaves. The principle of HOG is the division of the image into small regions called cells. For each cell, a histogram of the gradient is computed. Depending on the gradient orientation, each cell is discretized into angular bins. Finally, adjacent cells are merged into blocks and then normalized.

3.4.5. Haralick Texture

The co-occurrence matrix makes it possible to obtain the occurrence frequency of a pattern of two pixels separated by a distance d along a direction θ. In [48], the authors proposed 14 features that can be computed on this matrix. These features have the aim of highlighting some visual characteristics, statistics, the randomness of the gray level distribution, and the linear dependence of the gray levels on a neighborhood of pixels (homogeneity, coarseness, periodicity, smoothness, etc.). In 2012, the Haralick method was applied to extract texture features in a classification of plant species [49]. Here, we used six Haralick features, namely, autocorrelation, contrast, correlation, dissimilarity, energy, and entropy. More details of these features can be found in [48,50,51].

3.4.6. Gabor Wavelets

This method performs joint space-frequency analysis. The short-time Fourier transform with a Gaussian window is called the Gabor transform. It is able to preserve both local and global information in the image and is particularly useful for analyzing texture images containing highly specific frequency or orientation characteristics [52,53]. In 2003, Tang et al. fixed the filter orientation at 90° for the classification of broad and narrow leaves [20]. By analyzing the separation between classes of each feature, they concluded that a filter bank with four frequency levels, from 4 to 7, was suitable for the classification task. Therefore, we chose from 4 to 7 as the frequencies and 0°, 45°, and 90° as the orientations. We generated 12 Gabor features.

3.5. SVM or Support Vector Machine

The ideal foundation of a good classification system is to have a fast classifier which avoids overfitting and is able to respond to multi-class problems, to separate classes with a large gap or margin, and to manage large feature vectors. In this study, we applied the SVM or support vector machine or large margin separator [54]. It is one of the most successful machine learning methods [55,56]. Its popularity comes not only from the fact that it provides class separation with a very large margin if provided with data in two classes but also because it is suitable for linear and nonlinear cases [57].

3.6. Random Forest (RF)

Random forest [58] is a meta-classifier, which combines several weak classifiers to form a strong one. RF easily handles multi-class problems, and it is robust to large features and has a very low risk of overfitting. It is used in several applications, such as point tracking in video surveillance, medical imaging, and games in Microsoft's Kinect. In addition, RF has been shown to be ideally suited for classifying high-resolution UAV data [59]. It is structured like a real forest with trees, where each tree has roots, branches, and leaves. Trees correspond to the different classifiers. The first node corresponds to the root of the tree (the point of entry of our data), each node is then separated into intermediate

nodes, and each leaf corresponds to a terminal node where the final decision is stored. The forest trees are built using bagging or bootstrap aggregating [60]. The principle of bagging is the construction of each tree by selecting a subset of n observations among the N learning data (n < N) obtained by random sampling with delivery. The objective is to get trees as different as possible, or, in other words, to obtain uncorrelated trees, because the more different the trees are, the more robust the forest is. The other advantage of bagging is that it makes it possible to estimate the prediction error of the forest by using "out-of-bag" (OOB) data or data not used during the construction of trees.

4. Results and Discussion

Experiments were conducted on two different fields of bean and spinach (Figure 6). Images were acquired by a DJI Phantom 3 Pro drone that embeds a 36-megapixel (MP) RGB camera at an altitude of 20 m. This acquisition system produces very high resolution images with a spatial resolution of approximately 0.35 cm.

To build the unsupervised training database, we selected two different parts of each field. The first one (Part1) was used to collect training data, and the other one (Part2) was used for test data collection.

To create the crop binary mask after line detection, the superpixels' compactness was set at 20 and the number of superpixels was set to $0.1\% \times N$, where $N = 7360 \times 4912$ pixels (Figure 4b). In this experiment, we used a 64×64 window to create the weed and crop training databases. This window size provides a good trade-off between plant type and overall information. A small window is not sufficient to capture the whole plant and can lead to confusing crops and weeds because, in some conditions, crop and weed leaves have the same visual characteristics. On the other hand, too large a size presents a risk of having crop and weeds in the same window.

(a)   (b)
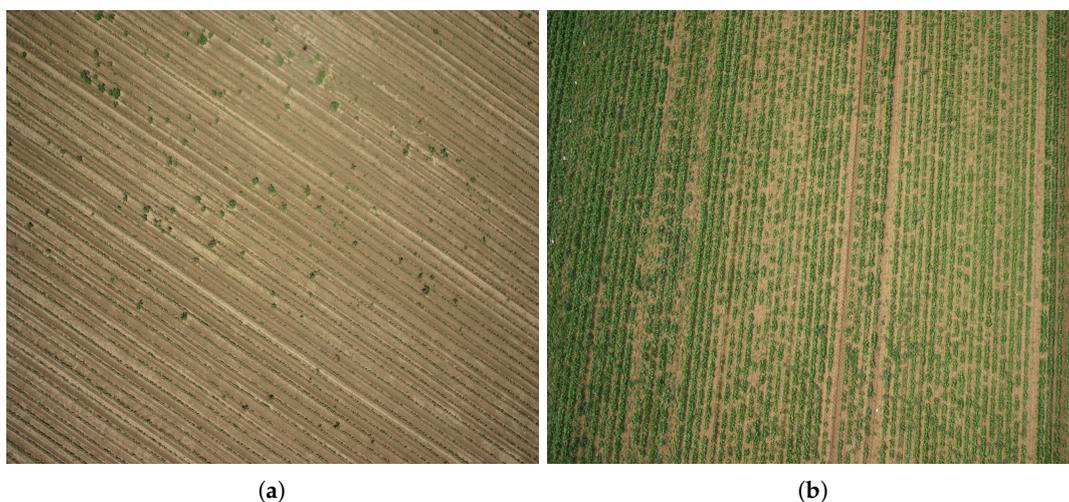
Figure 6. Example of images taken in bean (**a**) and spinach fields (**b**). The bean field has fewer inter-row weeds and is predominantly composed of potential weeds. The inter-row distance is stable and plants are sparse compared to the spinach field, which presents a dense vegetation in the crop rows and irregular inter-row distances. The spinach field has more inter-row weeds and has few potential weeds.

In the bean field, the weeds present are thistles and young potato sprouts from a previous sowing on the same field. This field has few inter-row weeds, so we decided to include potential weeds directly in the weed samples. After applying the unsupervised labeling method, the number of samples collected was 673 for weeds and 4861 for crop. Even with potential weeds included, the collected samples were unbalanced. To address this problem, we carried out data augmentation. Hence, we performed two contrast changes, smoothing with a Gaussian filter, and three rotations (90°, 180°, 270°). A strong heterogeneity in the fields can often be encountered from one part of the field to another one. This heterogeneity may correspond to

a difference in soil moisture, presence of straw, etc. In order to make our models robust to background, we mixed samples with and without background. Samples without background were obtained by applying ExG followed by Otsu thresholding on previously created samples (Figure 7). We evaluated the performance of our method by comparing models created by data labeled in supervised and unsupervised ways.

The supervised training datasets were labeled by human experts. A mask was applied manually to the pixels of weeds and crop. Figure 8 presents weeds delineated in red by an expert. The supervised data collected were also unbalanced, so we carried out the same data augmentation procedure performed on the unsupervised data. The total number of samples is shown in Table 1.

(a)　　　　　(b)　　　　　(c)　　　　　(d)

(e)　　　　　(f)　　　　　(g)　　　　　(h)

Figure 7. Example of crop and weed samples of size 64 × 64 pixels with and without background. Bean: samples of crop (**a**,**b**), samples of weed (**c**,**d**). Spinach: samples of crop (**e**,**f**) and samples of weed (**g**,**h**). Depending on the plant size and the window position, we obtain a plant or aggregation of plants per window.

(**a**)   (**b**)
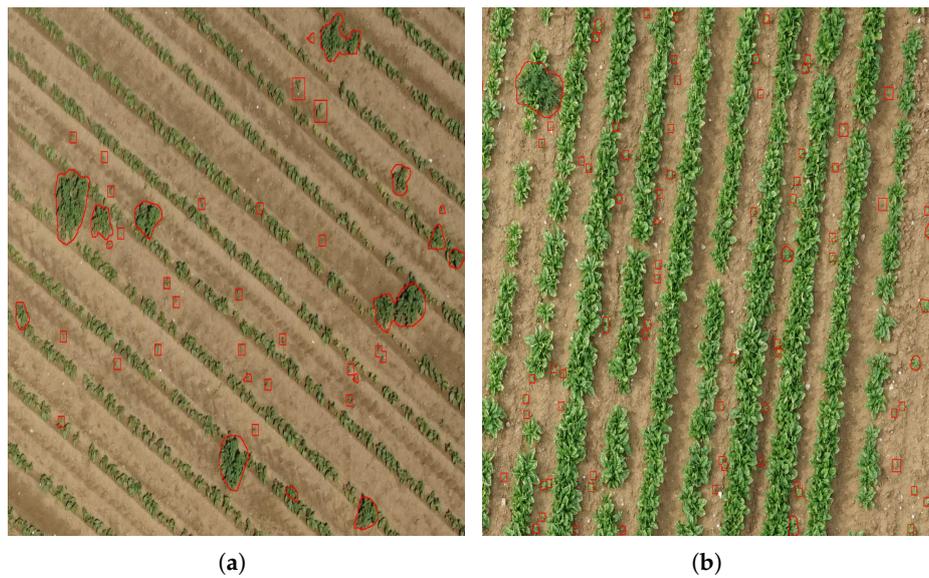
Figure 8. Parts of the bean field (**a**) and the spinach field (**b**) with the weeds labeled manually by an expert in red. The manual labeling took about 2 working days.

The spinach field is more invaded by weeds (mainly thistles) than the bean field. Altogether, 4303 samples of crop and 3626 samples of weed were labeled in an unsupervised way. Unlike for the bean field, we obtained less unbalanced data. Therefore, the only data augmentation applied was adding samples without background. The same processing was applied to the supervised data. Table 1 presents the number of samples.

Table 1. Training and validation data in the bean and spinach fields.

Bean Field				
Data	**Class**	**Training**	**Validation**	**Total**
Supervised	Crop	17,192	11,694	28,886
labeling	Weed	17,076	9060	16,136
Total		34,868	20,754	45,022
Unsupervised	Crop	7688	1928	9616
labeling	Weed	5935	1493	7428
Total		13,623	3421	17,044
Spinach field				
Data	**Class**	**Training**	**Validation**	**Total**
Supervised	Crop	11,350	2838	14,188
labeling	Weed	8234	2058	10,292
Total		19,584	4896	34,772
Unsupervised	Crop	6884	1722	8606
labeling	Weed	5800	1452	7252
Total		12,684	3174	15,858

4.1. Results and Discussion

After the creation of both weed and crop classes, 80% of the samples were selected randomly for the training, and the remaining ones were used for validation. Table 1 presents the training and validation data performed on each field.

For fine-tuning, we tested different values of the learning rate. The initial learning rate was set to 0.01 and updated every 200 epochs. The update was done by dividing the learning rate by a factor of 10. Figure 9 shows the evolution of the loss function during training for supervised and unsupervised datasets for spinach and bean fields. From these figures, it can be seen that the validation loss curves decrease during about the first 80 epochs before increasing and converging (behavior close to overfitting). Overfitting was less pronounced in the supervised labeled bean data. The best models were obtained during the first learning phase with a learning rate of 0.01.

The performance of the models was assessed on test ground truth data collected in Part2 in a supervised way on each field; Table 2 presents the samples. The performance of the classification results is illustrated with receiver operating characteristic (ROC) curves.

The ROC curves (Figure 10) show that the AUCs (area under the curve) are close to or greater than 90% and that both types of learning data provide good results that are comparable. For both fields, a false positive rate of 20% provides a true positive rate greater than 80%. The differences in performance between supervised and unsupervised data labeling are about 6% in the bean field and about 1.5% in the spinach field. The performance gap in the bean field can be explained by the sparsity of inter-row weeds.

Both fields are infested mainly by thistles; we tested the robustness of our models by exchanging the samples of weeds from the bean field with those of the spinach field.

In Figure 11, the results obtained show that, despite the small number of samples harvested in the bean field, those data are suitable for the spinach field, and the model created with unsupervised labeling in the spinach field is most sensitive to the presence of young potato sprouts among bean weed samples.
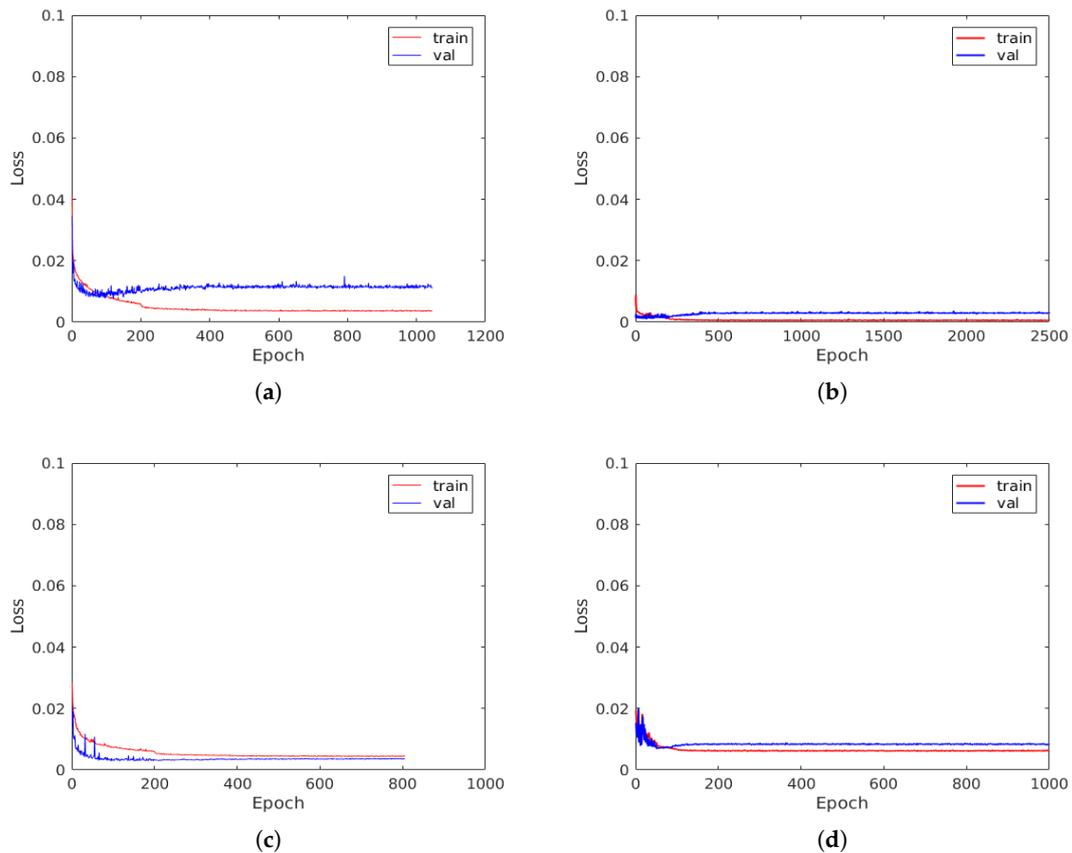
Figure 9. Evolution of the loss during training for supervised and unsupervised data in the spinach and bean fields. The validation loss curves decrease during about the first 80 epochs before increasing and converging. The top two figures represent the spinach field, and the bottom two correspond to the bean field. Figures on the left are from training on the supervised data, and those on the right are from training on the unsupervised data.
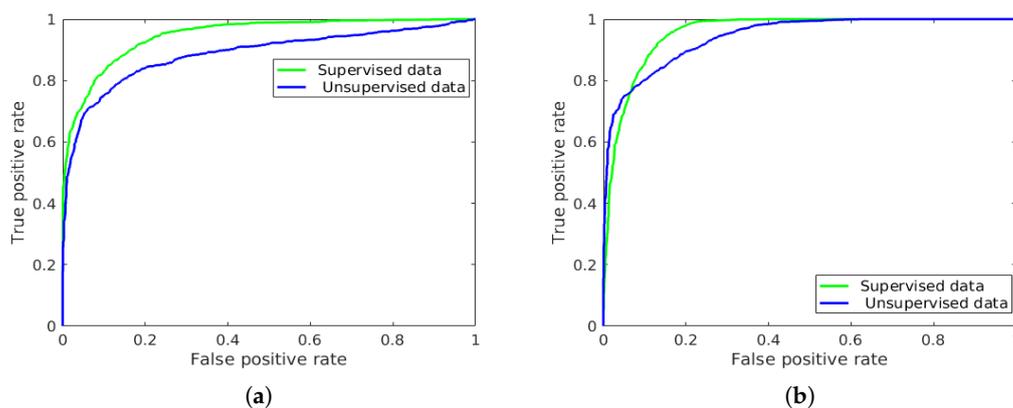
Figure 10. Receiver operating characteristic (ROC) curves of the test data with unsupervised and supervised data labeling. From left to right, the ROC curves computed on the bean (**a**) and spinach (**b**) test data. In the bean field, the areas under the curve (AUCs) are 88.73% for unsupervised data and 94.84% for supervised data. In the spinach field, the AUCs are 94.34% for unsupervised data and 95.70% for supervised data. Supervised and unsupervised data mean, respectively, data labeled in supervised and unsupervised ways.
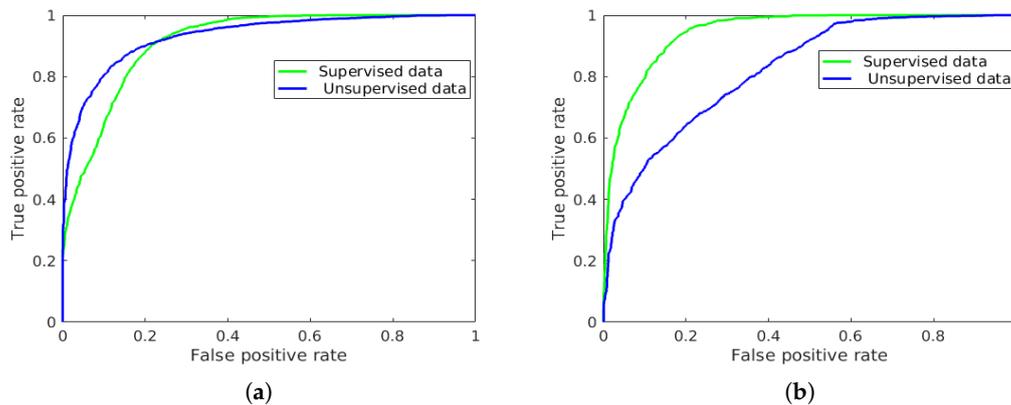
Figure 11. ROC curves of test data with weed data from the bean field exchanged with those of the spinach field. From left to right: the ROC curves computed on the bean (**a**) and spinach (**b**) test data. In the bean field, the areas under the curve (AUCs) are 91.37% for unsupervised data and 93.25% for supervised data. In the spinach field, the areas under the curve (AUC) are 82.70% for unsupervised data and 94.34% for supervised data. Supervised and unsupervised data mean, respectively, data labeled in supervised and unsupervised ways.

Table 2. Number of test samples used for each field.

Field	Crop Samples	Weed Samples
Bean	2139	1852
Spinach	1523	1825

4.2. ResNet vs. Feature Extraction with SVM and RF

SVM and RF were applied to features extracted from the datasets (Table 1). RF was performed with 200 trees. As for Resnet18, models were created based on data labeled in supervised and unsupervised ways. In order to assess the effectiveness of the selected features, we applied them separately, and to select the set of features that gives the optimal classification result, we combined them. Figures 12 and 13 show that the color, Haralick, and geometric features give the best results. In the spinach field, the abundance of thistles with a different color of leaf from that of spinach at a certain level of growth explains the effectiveness of the color features. In the bean field, the color features were less effective than the texture features (Haralick) in both datasets since we have young potato shoots from the previous sowing among the weeds, and their color is almost the same as that of the bean plants.

By using SVM, when the features are combined, the improvement is less than 2% for the data labeled in a supervised manner and about 10% for unsupervised data in the spinach field. In the bean field, the same remark applies to the data collected in a supervised manner; for the unsupervised data collected, no improvement was found. Another remark that can be made is that from one type of data to another, the best features are not the same. We also noticed that the selected features are not suitable to detect the weeds present in the bean field. With the RF, the feature selection procedure only increased performance by about 1% for both spinach datasets. In the bean field, an improvement of about 1% was observed with the data labeled in an unsupervised way and about 5% for the data labeled in a supervised way. Tables 3 and 4 present the results of SVM and RF with the best selected features.

Table 3. Results of test data collected in the bean field with ResNet18, support vector machine (SVM), and random forest (RF). For the SVM and RF, only the results of the best selected features are presented.

Best Features	SVM (AUC%)		RF (AUC%)		ResNet18 (AUC%)	
	Sup Labeling	Unsup Labeling	Sup Labeling	Unsup Labeling	Sup Labeling	Unsup Labeling
ALL features	**60.60**	44.76	**70.16**	63.95	-	-
Geo3	40.80	**59.51**	48.91	44.86	-	-
Haralick, Color	59.78	40.46	68.15	**65.40**	-	-
-	-	-	-	-	94.84	88.73

Table 4. Results of test data collected in the spinach field with ResNet18, SVM, and random forest. For the SVM and RF, only the results of the best selected features are presented. Sup and Unsup mean, respectively, supervised and unsupervised.

Best Features	SVM (AUC%)		RF (AUC%)		ResNet18 (AUC%)	
	Sup Labeling	Unsup Labeling	Sup Labeling	Unsup Labeling	Sup Labeling	Unsup Labeling
Color, HOG, Gabor	**95.94**	87.38	93.50	95.131	-	-
Haralick, Color, HOG, Gabor	93.93	**90.77**	95.464	**96.177**	-	-
All features	93.352	90.70	**96.99**	95.162	-	-
-	-	-	-	-	95.70	94.34
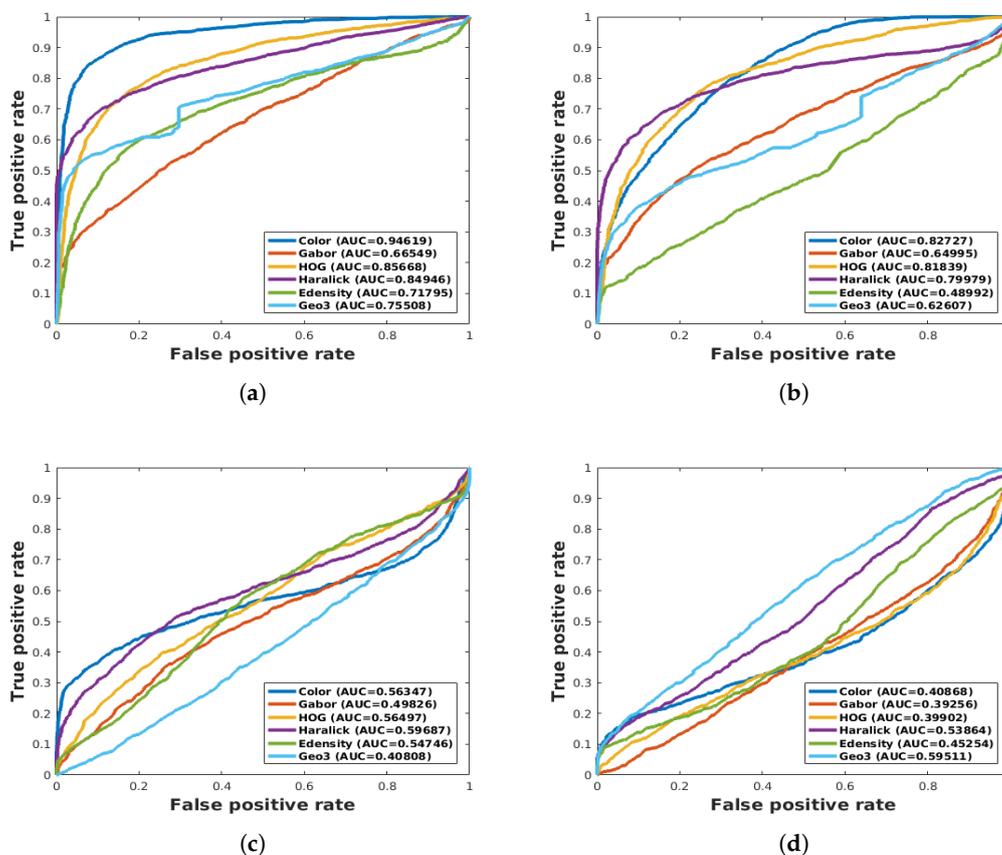
Figure 12. ROC curves of the SVM models created by each feature for each field. The first line represents the spinach field and the second one is the bean field. The first and second columns are the results of the models trained on the supervised and unsupervised data, respectively.
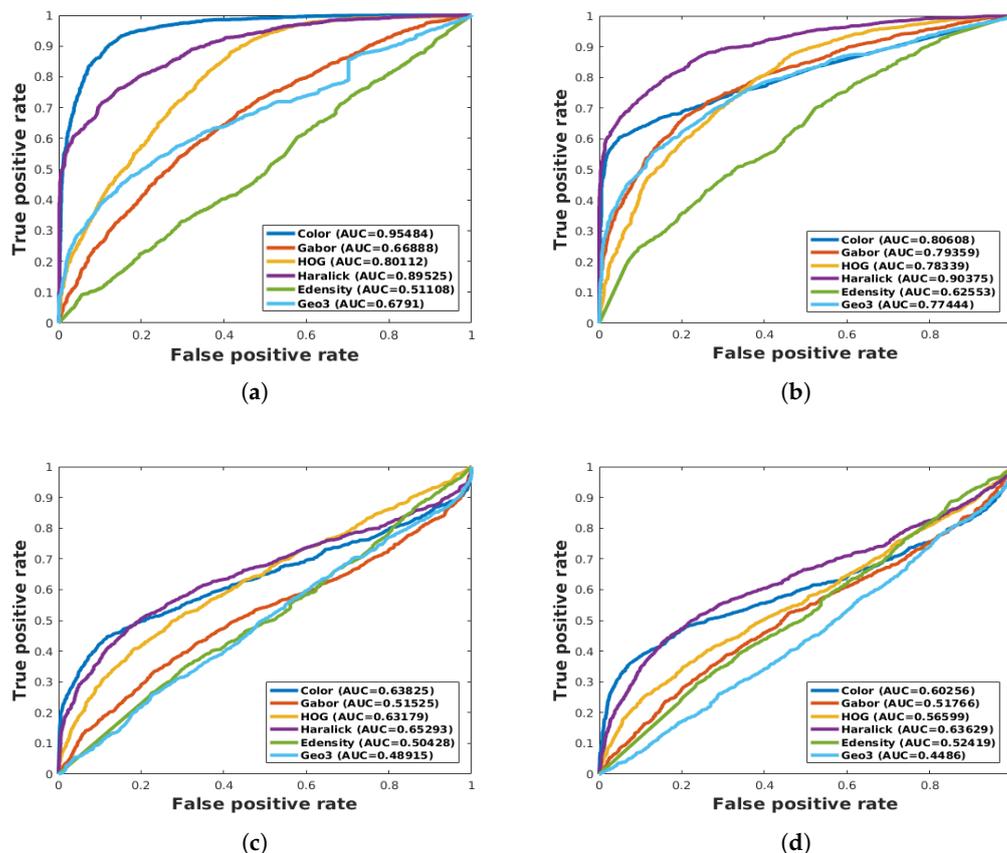
Figure 13. ROC curves of the RF models created by each feature for each field. The first line represents the spinach field and the second one is the bean field. The first and second columns are the results of the models trained on the supervised and unsupervised data, respectively.

As reported in Table 4, ResNet18 provides much better results than SVM and RF in the bean field, with a performance difference greater than 20%. However, in the spinach field, the results obtained are comparable and sometimes the results of ResNet18 are lower than those of SVM and RF (Table 4). This performance of ResNet18 can be explained by the small amount of data used for training in the spinach field. For deep learning algorithms, the more data we have, the better the algorithm learns. We can also note that the performance of the models formed by the two types of data collected is comparable for the three classification methods. The maximum difference is about 6% in both fields.

Based on the results, it can be concluded that even if we manage to select the most suitable features to identify weeds in a field, these features may not be adapted to another field with a different type of crop. They also show that the features considered better by a classifier may not necessarily be the best if you change the classifier. However, in the fields from one year to another, new types of weeds may appear and the level of growth of the plants can sometimes cause confusion between weeds and crops, which leads to a new collection of weed/crop data and a new selection of features. Thus, for an efficient classification, it would be interesting to use a tool capable of automatically generating relevant samples and features to detect weeds, hence the interest in using deep learning with unsupervised data labeling.

4.3. Weed Detection

In order to detect weeds in an entire UAV image, we applied an overlapping window for weed detection. For each position of the window, the CNN models provide the probability of the plants

being weeds or crops. Thus, the center of the extracted image is marked by a colored dot according to the probabilities. Blue, red, and white dots mean, respectively, that the extracted image is identified as crop, weed, and an uncertain decision (Figure 14a,c). Uncertain decision means that the two probabilities are very close to 0.5. Thereafter, we used crop line information and the previously created superpixels to classify all the pixels of the image. On each superpixel, we identify which dot color is dominant. A superpixel is classed as crop or weed if the majority of dots are blue or red, respectively. For superpixels where the majority of dots are white, we used crop line information. Hence, superpixels which are in the crop lines are regarded as crop and the others are weeds. The superpixels created in the background are removed. Figure 14b,d present the classification results in parts of the spinach and bean fields. It can be seen that inter-row and intra-row weeds are slightly overdetected. Overdetections are mainly found at the edges of the crop rows where the window cannot overlap the whole plant. Some weed pixels are not entirely in red because, after applying the threshold to the ExG, the parts of these plants which are less green are considered soil.

(**a**)　　　　　　　　　　　(**b**)

(**c**)　　　　　　　　　　　(**d**)
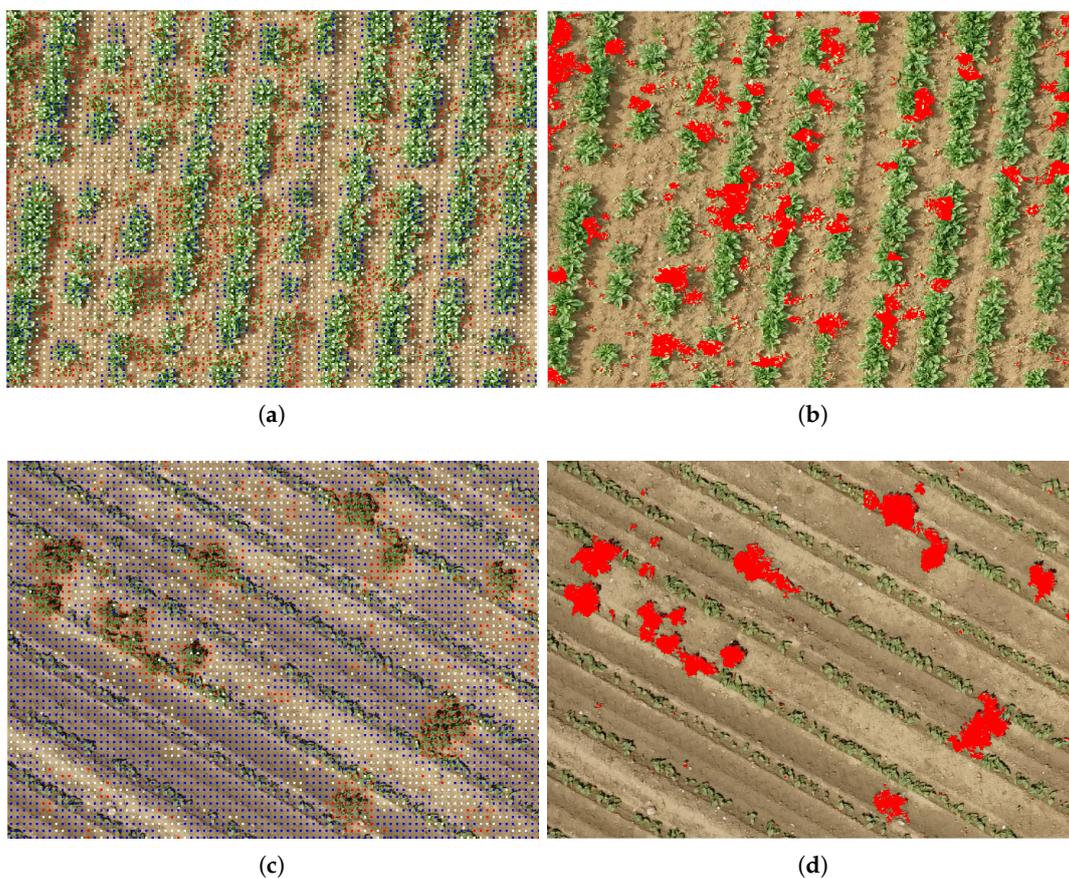
Figure 14. Examples of unmanned aerial vehicle (UAV) image classification with models created by unsupervised data in two different fields. The top two figures show samples from the spinach field and the bottom two samples are from the bean field. On the left are the samples obtained after using a sliding window, without crop line and background information. Blue, red, and white dots mean that the plants are identified as crop, weed, and an uncertain decision, respectively. On the right in red are the weeds detected after crop line and background information has been applied.

However, the unsupervised data collection method strongly depends on the efficiency of the crop line detection method and also on the presence of weeds in the inter-row. The line detection approach used here has already shown its effectiveness in beet and corn fields in our previous work [37]. With the bean field, we found that even if a field does not have a lot of samples of weeds in the

inter-row, it is possible to create a robust model with data augmentation. By using a deep learning architecture such as ResNet18, robust models can be created for the classification of weeds in bean or spinach fields with supervised or unsupervised data labeling. This work can be compared to recent studies which also aim to develop unsupervised detection approaches. A semi-automatic object-based image analysis (OBIA) procedure was developed with random forest combined with feature selection techniques to classify soil, weeds, and maize in [27]. An overall accuracy of 0.945 and a Kappa value of 0.912 were obtained. This method was applied to only one field, but we have found that the feature selection approach, even with random forest, is not robust when the field or crop type changes. In [34], an automatic image processing method was developed to discriminate between crop and weed pixels on images acquired by a camera mounted on a manually held pole. The authors combined spatial and spectral information extracted from four-band multispectral images. SVM was applied to the spectral information to perform the classification. On all images, the mean value of the weed detection rate was 89% for their spatial and spectral combination method. This method assumes that weeds and crops have different spectral information, which is not always the case in farm fields. Di Cicco et al. [33] used synthetic training datasets. However, this technique requires a precise modeling in terms of texture, 3D models, and light conditions. Overall, this illustrates that the main advantage of our method compared to the ones that use unsupervised labeling is that it is fully automatic and that no feature selection is required.

Currently, our method has only been evaluated on images acquired in the visible spectrum. As the line detection approach depends on the background segmentation, we intend to adapt the proposed method to multispectral images in future work. Then, we can implement a robust background segmentation algorithm using the Normalized Difference Vegetation Index (NDVI) [61]. Beyond segmentation, the multispectral bands could also provide additional information to distinguish crops from certain weed species.

5. Conclusions

In this paper, we propose a novel fully automatic learning method using convolutional neuronal networks with unsupervised training dataset collection for weed detection in UAV images acquired from bean and spinach fields. The results obtained show a performance close to that of supervised data labeling. The area under curve (AUC) differences are 1.5% in the spinach field and 6% in the bean field. Supervised labeling is an expensive task for human experts, and given the differences in accuracy between supervised and unsupervised labeling, our method can be a better choice in the detection of weeds, especially when crop rows are spaced well apart. The proposed method is interesting in terms of flexibility and adaptivity, since a model can be easily trained on a new dataset. We also found that the ResNet18 architecture can extract useful features for weed classification in bean or spinach fields with data labeled in a supervised or unsupervised manner. In addition, the developed method could be a key technique for online weed detection with UAV.

As future work, we plan to use multispectral images because, in some conditions, multispectral bands such as red edge or near infrared could help to distinguish plants, even if they have similarities in the visible spectrum and leaf shape. With multispectral information, we also expect to improve the background segmentation. To enhance the simplicity of use and rapidity of the method, we intend to implement an application with a graphical interface that will automatically chain the different methods used in the processing flowchart and generate a weed infestation map. This map can then be integrated into a robot or tractor for selective herbicide spraying, thereby helping farmers to save money while applying the right amount of herbicide where it is needed.

Author Contributions: M.D.B., A.H. and R.C. conceived and designed the method; M.D.B implemented the method and performed the experiments; M.D.B., A.H. and R.C. wrote the paper, discussed the results and revised the manuscript. All authors have read and approved the manuscript.

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

References

1. European Crop Protection. *With or Without Pesticides?* ECPA: Brussels, Belgium, 2017.
2. Oerke, E.C. Crop losses to pests. *J. Agric. Sci.* **2006**, *144*, 31. [CrossRef]
3. Pierce, F.J.; Nowak, P. Aspects of Precision Agriculture. *Adv. Agron.* **1999**, *67*, 1–85. [CrossRef]
4. McBratney, A.; Whelan, B.; Ancev, T.; Bouma, J. Future Directions of Precision Agriculture. *Precis. Agric.* **2005**, *6*, 7–23. [CrossRef]
5. Mulla, D.J. Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps. *Biosyst. Eng.* **2013**, *114*, 358–371. [CrossRef]
6. Weis, M.; Gutjahr, C.; Ayala, V.R.; Gerhards, R.; Ritter, C.; Schölderle, F. Precision farming for weed management: Techniques. *Gesunde Pflanzen* **2008**, *60*, 171–181. [CrossRef]
7. Herrmann, I.; Shapira, U.; Kinast, S.; Karnieli, A.; Bonfil, D.J. Ground-level hyperspectral imagery for detecting weeds in wheat fields. *Precis. Agric.* **2013**, *14*, 637–659. [CrossRef]
8. Alchanatis, V.; Ridel, L.; Hetzroni, A.; Yaroslavsky, L. Weed detection in multi-spectral images of cotton fields. In *Computers and Electronics in Agriculture*; Elsevier: New York, NY, USA, 2005; Volume 47, pp. 243–260. [CrossRef]
9. Torres-Sánchez, J.; López-Granados, F.; Peña, J.M. An automatic object-based method for optimal thresholding in UAV images: Application for vegetation detection in herbaceous crops. *Comput. Electron. Agric.* **2015**, *114*, 43–52. [CrossRef]
10. Zhang, C.; Kovacs, J.M. The application of small unmanned aerial systems for precision agriculture: A review. *Precis. Agric.* **2012**, *13*, 693–712. [CrossRef]
11. Dos Santos Ferreira, A.; Matte Freitas, D.; Gonçalves da Silva, G.; Pistori, H.; Theophilo Folhes, M. Weed detection in soybean crops using ConvNets. *Comput. Electron. Agric.* **2017**, *143*, 314–324. [CrossRef]
12. Bah, M.D.; Dericquebourg, E.; Hafiane, A.; Canals, R. Deep learning based classification system for identifying weeds using high-resolution UAV imagery. In Proceedings of the 2018 Computing Conference, London, UK, July 2018.
13. 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]
14. Gée, C.; Bossu, J.; Jones, G.; Truchetet, F. Crop/weed discrimination in perspective agronomic images. *Comput. Electron. Agric.* **2008**, *60*, 49–59. [CrossRef]
15. Woebbecke, D.; Meyer, G.; Von Bargen, K.; Mortensen, D. Color indices for weed identification under various soil, residue, and lighting conditions. *Trans. ASAE* **1995**, *38*, 259–269. [CrossRef]
16. Otsu, N. A threshold selection method from gray-level histograms. *IEEE Trans. Syst. Man Cybern.* **1979**, *9*, 62–66. [CrossRef]
17. Hough, P.V.C. Method and Means for Recognizing Complex Patterns. US Patent 3,069,654, 18 December 2018. [CrossRef]
18. Peña, J.M.; Torres-Sánchez, J.; Isabel De Castro, A.; Kelly, M.; López-Granados, F. Weed Mapping in Early-Season Maize Fields Using Object-Based Analysis of Unmanned Aerial Vehicle (UAV) Images. *PLoS ONE* **2013**, *8*. [CrossRef] [PubMed]
19. Blaschke, T. Object based image analysis for remote sensing. *ISPRS J. Photogramm. Remote Sens.* **2010**, *65*, 2–16. [CrossRef]
20. Tang, L.; Tian, L.; Steward, B.L. Classification of Broadleaf and Grass Weeds Using Gabor Wavelets and an Artificial Neural Network. *Trans. Am. Soc. Agric. Eng.* **2003**, *46*, 1247–1254. [CrossRef]
21. Jeon, H.Y.; Tian, L.F.; Zhu, H. Robust crop and weed segmentation under uncontrolled outdoor illumination. *Sensors* **2011**, *11*, 6270–6283. [CrossRef] [PubMed]
22. Ahmed, F.; Al-Mamun, H.A.; Bari, A.S.M.H.; Hossain, E.; Kwan, P. Classification of crops and weeds from digital images: A support vector machine approach. *Crop Protect.* **2012**, *40*, 98–104. [CrossRef]
23. Latha, M.; Poojith, A.; Reddy, A.; Kumar, V. Image Processing in Agriculture. *Int. J. Innov. Res. Electr. Electron. Instrum. Control Eng.* **2014**, *2*, ISSN 2321-2004.

24. Pérez-Ortiz, M.; Peña, J.M.; Gutiérrez, P.A.; Torres-Sánchez, J.; Hervás-Martínez, C.; López-Granados, F. Selecting patterns and features for between- and within-crop-row weed mapping using UAV-imagery. *Expert Syst. Appl.* **2015**, *47*, 85–94. [CrossRef]

25. Bakhshipour, A.; Jafari, A.; Nassiri, S.M.; Zare, D. Weed segmentation using texture features extracted from wavelet sub-images. *Biosyst. Eng.* **2017**, *157*, 1–12. [CrossRef]

26. Bakhshipour, A.; Jafari, A. Evaluation of support vector machine and artificial neural networks in weed detection using shape features. *Comput. Electron. Agric.* **2018**, *145*, 153–160. [CrossRef]

27. Gao, J.; Liao, W.; Nuyttens, D.; Lootens, P.; Vangeyte, J.; Pižurica, A.; He, Y.; Pieters, J.G. Fusion of pixel and object-based features for weed mapping using unmanned aerial vehicle imagery. *Int. J. Appl. Earth Observ. Geoinf.* **2018**, *67*, 43–53. [CrossRef]

28. LeCun, Y.; Bottou, L.; Bengio, Y.; Haffner, P. Gradient-based learning applied to document recognition. *Proc. IEEE* **1998**, *86*, 2278–2323. [CrossRef]

29. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet Classification with Deep Convolutional Neural Networks. In *Advances in Neural Information Processing Systems*; MIT Press Ltd.: Cambridge, MA, USA, 2012; pp. 1–9. [CrossRef]

30. Hung, C.; Xu, Z.; Sukkarieh, S. Feature Learning Based Approach for Weed Classification Using High Resolution Aerial Images from a Digital Camera Mounted on a UAV. *Remote Sens.* **2014**, *6*, 12037–12054. [CrossRef]

31. Mortensen, A.K.; Dyrmann, M.; Karstoft, H.; Nyholm Jørgensen, R.; Gislum, R. Semantic Segmentation of Mixed Crops using Deep Convolutional Neural Network. In Proceedings of the International Conference on Agricultural Engineering, Aarhus, Denmark, 26–29 June 2016.

32. Milioto, A.; Lottes, P.; Stachniss, C. Real-time blob-wise sugar beets vs. weeds classification for monitoring fields using convolutional neural networks. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2017**, *IV-2/W3*, 41–48. [CrossRef]

33. Di Cicco, M.; Potena, C.; Grisetti, G.; Pretto, A. Automatic model based dataset generation for fast and accurate crop and weeds detection. In Proceedings of the 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Vancouver, BC, Canada, 24–28 September 2017; pp. 5188–5195. [CrossRef]

34. Louargant, M.; Jones, G.; Faroux, R.; Paoli, J.N.; Maillot, T.; Gée, C.; Villette, S. Unsupervised Classification Algorithm for Early Weed Detection in Row-Crops by Combining Spatial and Spectral Information. *Remote Sens.* **2018**, *10*, 761. [CrossRef]

35. Jones, G.; Gée, C.; Truchetet, F. Modelling agronomic images for weed detection and comparison of crop/weed discrimination algorithm performance. *Precis. Agric.* **2009**, *10*, 1–15. [CrossRef]

36. Montalvo, M.; Pajares, G.; Guerrero, J.M.; Romeo, J.; Guijarro, M.; Ribeiro, A.; Ruz, J.J.; Cruz, J.M. Automatic detection of crop rows in maize fields with high weeds pressure. *Expert Syst. Appl.* **2012**, *39*, 11889–11897. [CrossRef]

37. Bah, M.D.; Hafiane, A.; Canals, R. Weeds detection in UAV imagery using SLIC and the hough transform. In Proceedings of the 2017 Seventh International Conference on Image Processing Theory, Tools and Applications (IPTA), Montreal, QC, Canada, 28 November–1 December 2017; pp. 1–6. [CrossRef]

38. Achanta, R.; Shaji, A.; Smith, K.; Lucchi, A.; Fua, P.; Süsstrunk, S. SLIC Superpixels Compared to State-of-the-Art Superpixel Methods. *IEEE Trans. Pattern Anal. Mach. Intell.* **2011**, *34*, 2274–2282. [CrossRef] [PubMed]

39. Kamilaris, A.; Prenafeta-Boldú, F.X. Deep learning in agriculture: A survey. *Comput. Electron. Agric.* **2018**, *147*, 70–90. [CrossRef]

40. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 26 June–1 July 2016; pp. 770–778. [CrossRef]

41. Simonyan, K.; Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv* **2015**, arXiv:1409.1556. [CrossRef]

42. Pan, S.J.; Yang, Q. A Survey on Transfer Learning. *IEEE Trans. Knowl. Data Eng.* **2010**, *22*, 1345–1359. [CrossRef]

43. Huang, Z.; Pan, Z.; Lei, B. Transfer learning with deep convolutional neural network for SAR target classification with limited labeled data. *Remote Sens.* **2017**, *9*, 907. [CrossRef]

44. Lottes, P.; Khanna, R.; Pfeifer, J.; Siegwart, R.; Stachniss, C. UAV-based crop and weed classification for smart farming. In Proceedings of the 2017 IEEE International Conference on Robotics and Automation (ICRA), Sands Expo, NV, USA, 29 May–3 June 2017; pp. 3024–3031. [CrossRef]

45. Dalal, N.; Triggs, W.; Triggs, B. Histograms of Oriented Gradients for Human Detection. In Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), San Diego, CA, USA, 20–26 June 2005. [CrossRef]

46. Xiao, X.Y.; Hu, R.; Zhang, S.W.; Wang, X.F. HOG-based approach for leaf classification. Advanced Intelligent Computing Theories and Applications. In Proceedings of the 6th International Conference on Intelligent Computing, ICIC 2010, Changsha, China, 18–21 August 2010; Springer: Berlin/Heidelberg, Germany, 2010; Volume 6216, pp. 149–155. [CrossRef]

47. Ibrahim, Z.; Sabri, N.; Abu Mangshor, N.N. Leaf Recognition using Texture Features for Herbal Plant Identification. *Indones. J. Electr. Eng. Comput. Sci.* **2018**, *9*, 152–156. [CrossRef]

48. Haralick, R.M.; Shanmugam, K. Textural Features for Image Classification. *Syst. Man Cybern.* **1973**, *3*, 610–621. [CrossRef]

49. Agrawal, K.N.; Singh, K.; Bora, G.C.; Lin, D. Weed Recognition Using Image-Processing Technique Based on Leaf Parameters. *J. Agric. Sci. Technol. B J. Agric. Sci. Technol.* **2012**, *2*, 899.

50. Soh, L.K.; Tsatsoulis, C. Texture analysis of sar sea ice imagery using gray level co-occurrence matrices. *IEEE Trans. Geosci. Remote Sens.* **1999**, *37*, 780–795. [CrossRef]

51. Brynolfsson, P.; Nilsson, D.; Torheim, T.; Asklund, T.; Karlsson, C.T.; Trygg, J.; Nyholm, T.; Garpebring, A. Haralick texture features from apparent diffusion coefficient (ADC) MRI images depend on imaging and pre-processing parameters. *Sci. Rep.* **2017**, *7*, 4041. [CrossRef] [PubMed]

52. Bovik, A.C.; Clark, M.; Geisler, W.S. Multichannel Texture Analysis Using Localized Spatial Filters. *IEEE Trans. Pattern Anal. Mach. Intell.* **1990**, *12*, 55–73. [CrossRef]

53. Naghdy, G.A. Texture analysis using Gabor wavelets. *Proc. SPIE* **1996**, *2657*, 74–85. [CrossRef]

54. Cortes, C.; Vapnik, V. Support-vector networks. *Mach. Learn.* **1995**, *20*, 273–297. [CrossRef]

55. Kazmi, W.; Garcia-Ruiz, F.; Nielsen, J.; Rasmussen, J.; Andersen, H.J. Exploiting affine invariant regions and leaf edge shapes for weed detection. *Comput. Electron. Agric.* **2015**, *118*, 290–299. [CrossRef]

56. Yang, Y.; Rushmeier, H. Special object extraction from medieval books using superpixels and bag-of-features. *J. Electron. Imaging* **2016**, *26*, 011008. [CrossRef]

57. Cristianini, N.; Shawe-Taylor, J. *An Introduction to Support Vector Machines and Other Kernel Based Learning Methods*; Cambridge University Press: Cambridge, UK, 2000; Volume 22, p. 190,

58. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]

59. de Castro, A.I.; Torres-Sánchez, J.; Peña, J.M.; Jiménez-Brenes, F.M.; Csillik, O.; López-Granados, F. An automatic random forest-OBIA algorithm for early weed mapping between and within crop rows using UAV imagery. *Remote Sens.* **2018**, *10*, 285. [CrossRef]

60. Geurts, P.; Ernst, D.; Wehenkel, L. Extremely randomized trees. *Mach. Learn.* **2006**, *63*, 3–42. [CrossRef]

61. Tucker, C.J. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens. Environ.* **1979**, *8*, 127–150. [CrossRef]