Article

Research on Vision-Based Navigation for Plant Protection UAV under the Near Color Background

Hehu Zhang ¹, Xiushan Wang ^{1,*}, Ying Chen ², Guoqiang Jiang ¹ and Shifeng Lin ³

¹ College of Mechanical & Electrical Engineering, Henan Agricultural University, Zhengzhou 450003, China; hehuzhang@hotmail.com (H.Z.); jgq19951013@gmail.com (G.J.)

² College of Humanity & Law, Henan Agricultural University, Zhengzhou 450003, China; henaucy@hotmail.com

³ Faculty of Engineering, University of New South Wales, Sydney 2052, Australia; tostarklin@gmail.com

* Correspondence: xswang@henau.edu.cn

Received: 26 February 2019; Accepted: 9 April 2019; Published: 12 April 2019

Abstract: GPS (Global Positioning System) navigation in agriculture is facing many challenges, such as weak signals in orchards and the high cost for small plots of farmland. With the reduction of camera cost and the emergence of excellent visual algorithms, visual navigation can solve the above problems. Visual navigation is a navigation technology that uses cameras to sense environmental information as the basis of an aircraft flight. It is mainly divided into five parts: Image acquisition, landmark recognition, route planning, flight control, and obstacle avoidance. Here, landmarks are plant canopy, buildings, mountains, and rivers, with unique geographical characteristics in a place. During visual navigation, landmark location and route tracking are key links. When there are significant color-differences (for example, the differences among red, green, and blue) between a landmark and the background, the landmark can be recognized based on classical visual algorithms. However, in the case of non-significant color-differences (for example, the differences between dark green and vivid green) between a landmark and the background, there are no robust and high-precision methods for landmark identification. In view of the above problem, visual navigation in a maize field is studied. First, the block recognition method based on fine-tuned Inception-V3 is developed; then, the maize canopy landmark is recognized based on the above method; finally, local navigation lines are extracted from the landmarks based on the maize canopy grayscale gradient law. The results show that the accuracy is 0.9501. When the block number is 256, the block recognition method achieves the best segmentation. The average segmentation quality is 0.87, and time is 0.251 s. This study suggests that stable visual semantic navigation can be achieved under the near color background. It will be an important reference for the navigation of plant protection UAV (Unmanned Aerial Vehicle).

Keywords: landmark location; route tracking; inception-V3; visual navigation; grayscale gradient law

1. Introduction

In recent years, the UAV (Unmanned Aerial Vehicle) has been widely used in agriculture, which has an important application value for crop growth data acquisition, pesticide spraying, pest detection, and so on. In particular, visual-based UAV navigation has attracted more and more attention. Visual navigation is a navigation technology that uses cameras to sense environmental information as the basis of an aircraft flight. It is mainly divided into five parts: Image acquisition, landmark recognition, route planning, flight control, and obstacle avoidance. Here, landmarks are plant canopy, buildings, mountains, and rivers, with unique geographical characteristics in a place. During UAV visual navigation, landmark recognition and route planning are the core links, and complex image algorithms are often used to extract and match colorful landmarks. Compared with navigations in other areas,

farmland mainly consists of a near color background. The classical segmentation methods are able to achieve good results under the background of significant color differences [1,2]. However, under the near color background, there is no mature, effective, and universal solution for plant canopy landmark segmentation. For example, a green leaves background makes it difficult to recognize green cucumbers, green peppers, and green apples when robots pick fruits or vegetables [3,4]. Similarly, the green duckweed and cyanobacteria background makes it hard to extract the navigation line when robots navigate a paddy field visually [5]. In order to effectively segment a plant canopy landmark under the near color background, it is necessary to develop specific segmentation algorithms.

Visual navigation based on a plant canopy landmark is one of the main navigation modes in agriculture, which can be achieved by three steps. First, vegetation is segmented; then, plant canopy landmarks are segmented from the vegetation; finally, the navigation line is detected according to the landmark. For the first two steps, many algorithms have been developed for vegetation and crop canopy segmentation, such as color-index-based segmentation: CIVE (color index of vegetation), NGRDI (normalized green red difference index), VEG (vegetative index), COM1 (combined indices 1) [6–9]; threshold-based segmentation: EH (entropy of a histogram), AT (automatic threshold) [10,11]; and learning-based segmentation: FC (fuzzy clustering), SVM (support vector machines), DTSM (decision tree based segmentation model), PSO-MM (particle swarm Optimization clustering and morphology modelling) [12–15]. Color-index-based and threshold-based segmentation occupy less memory resources, but color-index-based segmentation is susceptible to light interference. As there is no significant difference from the green leaf base to the green leaf edge, it is difficult to achieve crop canopy segmentation based on a pixel feature threshold. Not only that, the green background is also one of the main reasons for poor segmentation quality. For some learn-based methods, such as FC and PSO-MM, poor real-time performance is a common drawback. About the last step, a variety of navigation line detection methods have been proposed in recent years. Generally, these methods are classified into several categories according to their detection principles, such as HT (Hough transform), LR (linear regression), SA (speckle analysis), SV (stereo vision), and HF (horizontal fringes) [16–20]. It is worth noting that the crop row is approximated as a straight line in the above methods. However, actual crop rows are discrete broken lines and a more specific route is often needed during the slow-speed and high-precision navigation process for weeding and fixed-point fertilization and so on. After analysis, it was found that piecewise linearization of the crop row is a good idea to solve the deficiencies of the existing research.

In order to stabilize navigation, plenty of colorful landmarks are often created for route tracking. As mentioned, there are still great challenges for the landmark location under the near color background. Fortunately, deep learning has developed rapidly in recent years [21,22]. It has successfully solved various recognition problems, such as speech recognition, image processing, and natural language processing. At the same time, good results have also been achieved in crop segmentation and recognition [23–27]. Based on the above facts, a good solution is to apply deep learning to landmark location under the near color background, which will improve the segmentation effect, and further ensure the robustness of visual navigation. In this study, the fine-tuned Inception-V3 deep convolution neural network framework is developed for landmark recognition [28]. First, the block recognition method based on fine-tuned Inception-V3 is developed; then, the maize canopy landmark is recognized based on the above method; finally, piecewise linearized local navigation lines are extracted from the landmarks based on the maize canopy gray gradient law.

2. Materials and Methods

2.1. Image Acquisition and Data Preparation

2.1.1. Maize Samples in Laboratory Environment

Cultivated maize seedlings were used as ground landmarks for visual navigation in a laboratory environment as shown in Figure 1. Up to now, "Zhengdan 958" has been the largest maize variety in

China. It has been popularized as evidenced by the nearly 30 million hectares in the whole country. In order to ensure the experiment was universal, "Zhengdan 958" seedlings were used as canopy recognition objects in the laboratory environment. According to the estimation of the canopy and root growth, maize seedlings were planted in white pots with a diameter of 18 cm. In addition, maize is adapted to grow in the hot summer and early autumn, and has high requirements for temperature and light intensity. Therefore, the temperature was adjusted to 30 degrees Celsius, and the light intensity was 10,000 to 60,000 lx.

Figure 1. Planting of maize seedlings.

2.1.2. Maize Samples in Farmland Environment

The research was carried out in the State Key Laboratory of Crop Science and the farmland environment samples were from Mao Village Crop Cultivation Base, Zhengzhou. In order to facilitate mechanized operation and intelligent recognition, maize plants were planted according to agronomic requirements that the row spacing and row spacing of maize plants were 60 cm and 40 cm, respectively [29]. A training set and verification set was created: First, 1000 maize plant images and 1000 vegetation background images were captured, respectively, in different scenes, time, weather, and light conditions; then, the captured images were divided into image blocks whose size could be recognized by the fine-tuned Inception-V3; finally, 2800 image blocks were randomly selected as a training set and 800 image blocks as a verification set.

2.1.3. UAV Image Acquisition Device

As shown in Figure 2, the 's70w' UAV with a size of 40 cm * 40 cm * 16 cm was used for maize canopy image acquisition, carrying a nine-axis high precision gyroscope and GPS sensors. Its main functions include a high pressure setting; 2.4 GHz remote control; real-time picture transmission; one-button take-off, return, and landing; pointing follow, which greatly facilitated the work of image acquisition. The navigation time was about 15 minutes, and the picture quality was 1080 P, which ensured the canopy image quality.

Figure 2. Unmanned aerial vehicle and image acquisition device.

2.2. Machine Vision Processing Platform

2.2.1. Deep Learning Workstation

The parameters of the workstation are shown in Table 1.

Table 1. Parameters of the workstation.

GPU	GPU Memory	CUDA-Core Number	CPU	DDR4
GTX1080Ti	12 GB	7168	E5-2620	128 GB

2.2.2. Deep Learning Framework and Programming Language

Operating System: Ubuntu16.04;

Deep Learning Framework: Tensor-Flow;

Programing Languages: Python and C++.

2.3. Image Preprocessing

The data enhancement method can greatly increase the sample size of a training data set and the generalization ability of the network model. Most deep learning frameworks have some basic functions, which can directly implement common data conversion. In order to build a specific image recognition system, our task is to determine a meaningful transformation method for existing data sets. Common conversion methods include: Pixel color jitter, rotation, cutting, random clipping, horizontal flip, lens stretching, and lens correction, etc.

In a farmland environment, the spatial orientation of objects depends on the shooting position. It is difficult to photograph maize leaf images from every angle to meet all possibilities. Image rotation is an effective preprocessing method to solve the above problems. The pixel position can be represented by the process of rotation transformation as show in Equation (1):

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} * \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} \quad (1)$$

where x_0 , y_0 are the original pixel coordinates, θ is the rotation angle, and x and y are the coordinates after rotation.

As shown in Figure 3, the image is rotated and generates four images, in which the angle of rotation is 90° , 180° , and 270° , and there is mirror symmetry.

Figure 3. Image rotation: (A) Initial; (B) 90° ; (C) 180° ; (D) 270° ; and (E) mirror symmetry.

Furthermore, the change of illumination causes great interference to target recognition under the near color background. To improve the generalization ability of the CNN (Convolutional Neural Network), it is essential to train the images under different illumination conditions. The corresponding image was generated by adjusting the brightness, contrast, and sharpness as shown in Figure 4.

Figure 4. Illumination disturbance: (A) Initial; (B) 70% brightness; (C) 130% brightness; (D) 30% sharpness; (E) 75% sharpness; (F) 70% contrast; (G) 130% contrast.

2.4. Fine-Tuned Inception-v3 and Network Layer Feature Representation

In order to classify image blocks, the fine-tuned Inception-V3 classifier was designed as shown in Figure 5. First, the block features were extracted by Inception-V3, which includes the convolution layer, pooling layer, and mixed layer. Figure 4 shows the feature representation at different layers. Then, the pool_2 layer was connected to the feature classifier that consists of a FC layer (full connection layer) and a soft-max layer. Finally, the probability distribution results were output.

Figure 5. Fine-tuned Inception-V3

Among them, the filter was the basic function unit of the above network layer as shown in Figure 6. The former convolution layer, X_i^{L-1} , data, namely an image or a feature map, was processed by filter k_{ij}^L to generate the latter convolution layer, X_j^L , data, namely a new feature map, which can be expressed as Equation (2). The lower sampling layer is also called the pooling layer. Its specific operation is basically the same as that of the convolution base layer, except that the convolution core of the lower sampling only takes the maximum and average values of the corresponding positions (maximum pooling and average pooling), and does not undergo the modification of reverse propagation, as shown in Equation (3):

$$X_j^L = f \left(\sum_{i \in M_j} X_i^{L-1} * k_{ij}^L + b_j^L \right) \quad (2)$$

$$X_j^L = f(\beta_j^L \text{down}(X_j^{L-1}) + b_j^L) \quad (3)$$

where '*' is the convolution operation, b_j^L is the bias term, $f(\cdot)$ is the activation function, β_j^L is the weight term, and $\text{down}(\cdot)$ is the down-sampling function.

2.5. Training and Verification of the Inception-V3 Classifier

2.5.1. Loss Function

Before the emergence of CNN, the RBF function was often used as an activation function for a single hidden layer neural network because of its excellent local approximate feature. However, the RBF function does not have good probability characteristics compared with the soft-max function. Therefore, the soft-max function is generally adopted as the activation function of the output layer of CNN. For the soft-max function, the probability that the image block, x , belongs to n is expressed as follows:

$$P(y_j = n) = \frac{e^{z(x_j)}}{\sum_{i=1}^n e^{z(x_i)}} \quad (4)$$

An extreme example is given to illustrate the probabilistic properties of soft-max. Equation (4) shows that if one of $z(x_j)$ is far greater than the others, the largest $z(x_j)$ mapping values gradually approximate to 1 while the rest are approximated to 0. Next, the cross-entropy loss function was used to measure the prediction performance for the N class samples of the M group, as shown in Equation (5):

$$C = -\frac{1}{M} \left[\sum_{m=1}^M \sum_{i=n}^N 1\{y_i = n\} \log P(y_j = n) \right] \quad (5)$$

where $1\{.\}$ represents an indicative function, and its value was equal to 1 only if the i th image of the test was correct.

Figure 6. Feature representation at different network layers.

2.5.2. Weight Optimization

In order to prevent the activation value from being too large or too small, $w = 1/n$ was set, where the regularized Gauss distribution was used to initialize the weights of the full connection layer, as shown in Equation (6):

$$W^L = \begin{bmatrix} G(W_{11}) & G(W_{12}) & \dots \\ G(W_{21}) & G(W_{22}) & \dots \\ \dots & \dots & \dots \end{bmatrix} * \sqrt{1/n^{L-1}} \quad (6)$$

where W^L is the weight matrix of layer L , and $G(\cdot)$ is the regularized Gauss distribution function.

For the Inception-V3 classifier, the activation value is denoted as X^{L-1} and the forward propagation can be expressed as Equations (7) and (8):

$$Z^L = W^L X^{L-1} + b^L \quad (7)$$

$$Y^L = f(Z^L) \quad (8)$$

where W^L is the weight of the full connection layer, b^L is the bias term, and Y^L is the activation value calculated by the $f(\cdot)$ activation function.

In addition, the error, ε^L , produced by soft-max back propagation can be expressed as Equation (9):

$$\varepsilon^L = \nabla_{Y^L} C \odot \frac{\Delta Y^L}{\Delta X^{L-1}} \quad (9)$$

where ∇ is the gradient operator, $\odot$ is the Hadamard product operator used for point-to-point multiplication between matrices or vectors, and $\Delta(\cdot)$ is the difference operator.

Finally, the SGD training parameters can be expressed as Equations (10) and (11):

$$W^L = W^L - \frac{\eta}{M} \sum \varepsilon^L (X^{L-1})^T \quad (10)$$

$$b^L = b^L - \frac{\eta}{M} \sum \varepsilon^L \quad (11)$$

where η is the learning rate.

2.5.3. Other Parameter Settings

Table 2 shows the specific parameter settings. In general, if η is too small, the model converges too slowly. On the contrary, the cost function will oscillate and the optimal solution will not be obtained. The batch size determines the direction of optimization. The bigger the batch size is, the more accurate the descent direction is, but an excessive batch size leads to high memory occupancy. Weight decay is multiplied before the regular term and controls the weight of the regularization term in the loss function. The size of epoch selection varies with different data sets. When epoch is small, under-fitting will occur. On the contrary, there will be over-fitting.

Table 2. Other parameter settings.

Leaning Rate: η	Batch Size	Weight Decay	Epoch Number
0.01	128	0.005	4000

2.6. Inception-V3 Network Training

Figure 7A shows that the average recognition accuracy is about 0.95 after 500, which shows that the Inception-V3 classifier has good recognition ability for maize image blocks. At the same time, it can

be seen that the validation curve and training curve have good consistency, and the distribution range of the residuals is narrow, which indicates that the network has strong generalization ability under different backgrounds and weather. Figure 7B shows that the loss reaches the requirements quickly after 250, which shows that the SGD optimization algorithm can make the weight converge quickly.

Figure 7. Training and verification of the Inception-V3 classifier. (A) Accuracy rate; (B) loss function.

2.7. Design of the Mask Extractor

After the implementation of the Inception-V3 classifier, the block mask extractor needs to be designed. First, the Inception-V3 classifier was integrated into the classifier-array, which converts the patch into the class-matrix; then, the class-matrix was transformed into the block classification label for maize regions; finally, the mask was generated according to the label as shown in Figure 8.

2.8. Maize Canopy Landmark Location and Local Route Tracking

Previous studies have found that the canopy of maize plants has a gray gradient distribution law in the radial direction (Figure 9B) (Wang et al., 2018). The canopy landmarks in the maize region were identified according to the above law (Figure 9C). As a result, local route tracking was realized based on maize canopy landmark information and the canopy gray gradient distribution (Figure 9D).

Figure 8. Mask extractor.

Figure 9. Maize canopy landmark location and local route tracking.

3. Results and Discussion

3.1. The Effect of Neural Technologies for Fine-Tuned Inception-V3

In order to verify the influence of image preprocessing, a contrast experiment was performed under the same experimental conditions. The experiment showed that the Inception-V3 network with image preprocessing obtains an accuracy of 95.01%, and the proposed model without image preprocessing only obtains an accuracy of 93.15%. The experimental result shows that the recognition accuracy is improved by about 1.86%. It shows that image enhancement can improve the generalization ability and robustness of the model by increasing the amount of training set data. In addition, batch regularization technology and dropout technology bring about obvious improvements, and if these

model enhancement technologies are used simultaneously, the accuracy of the verification set will be raised from about 90.15% to about 95.01%, and the improvement effect is very obvious.

In the process of model training, it was found that not all optimization algorithms can converge quickly, for example, a comparison of the convergence using the RMS-prop and SDG algorithms.

As shown in Figure 10, the training accuracy of RMSProp is still below 0.5 under the same iteration times, while the SGD algorithm quickly reaches above 0.8. Compared with the RMSProp optimization algorithm, the SGD algorithm can converge quickly. Generally, the SGD optimization algorithm has a stable convergence rate and it can optimize the network with the minimum number of training times and prevent over-fitting.

Figure 10. Comparison of optimization algorithms.

3.2. Performance of Fine-Tuned Inception-V3 on the Training and Test Set

Figure 11 shows that the accuracy of BP-Net is the lowest, averaging 0.64. SVM is an excellent algorithm in small sample data classification. It can converge after less iteration, but the recognition accuracy is also relatively low, averaging at 0.68. As Inception-V3 is able to extract general features, only the back-end network needs to be fine-tuned. Compared with the former two, it achieves amazing classification accuracy on the training set, averaging 0.95. Then, taking VGG16 and the original Inception-V3 as benchmarks, the fine-tuned Inception-V3 was evaluated on the test set.

Where, MAP (mean average precision) is the solving of the limitation of the single value of the P, R, F-measure, taking into account the ranking of the retrieval results. Table 3 shows that the recognition accuracy and speed of the fine-tuned Inception-V3 on the test set were not significantly decreased.

Table 3. Comparison of the recognition accuracy of the classifier on the test set.

Framework	Accuracy	MAP	Time (s)
VGG16	0.8925	0.8157	0.231
Inception-V3	0.965	0.8025	0.172
Fine-tuned	0.9501	0.8688	0.179

Figure 11. Contrast of training characteristics of classifiers.

3.3. Recognition Results of Maize Region Under the Background of Non-Significant Color Differences

Figure 12 shows that the green vegetation in the background is highly similar to the green maize plant, and some green background, such as weeds, will be identified as part of the maize region. However, the misidentified blocks are in the "island" state, which can be easily filtered out by post processing [30]. In general, the algorithm realizes the recognition of the maize region under the background of the non-significant color difference. Next, the accuracy and effect of the block recognition were evaluated and compared with the classical green vegetation segmentation method, as shown in Table 4.

Figure 12. Recognition results of the maize region under the background of non-significant color differences.

Table 4. Comparison of the segmentation results with different segmentation methods.

Algorithm	Mean	Variance	Time (s)
COM1	0.65	0.017	0.011
AT	0.57	0.015	0.014
PSO-MM	0.71	0.012	0.215
Mask Extractor	0.87	0.005	0.251

As there is a color cross between the green vegetation and green maize plants, the color-based method and threshold-based method have poor segmentation, such as COM1 and AT, in Table 4. The clustering method is based on the principle that feature similarity is divided into the same super-pixel, so the green vegetation and green maize plants will be recognized as the same super-pixel. Moreover, the clustering method is time-consuming, for example, the time of PSO-MM is 0.215 s, as shown in Table 4. Compared with the above method, in the same time-consuming situation, good segmentation results can be obtained by the mask extractor based on the fine-tuning

3.4. Inception-V3 Performance of the Mask Extractor under Different Block Numbers

It is necessary to quantitatively evaluate the impact of block numbers on the recognition performance. Accuracy and recall were used as the main criteria to evaluate the segmentation effect. Typically, there are four kinds of classification results: TP (true positives, positive classes judged as positive classes), FN (false negatives, positive classes judged as negative classes), FP (false positives, negative classes judged as positive classes), and TN (true negatives, negative classes judged as negative classes). As a result, the equations of precision and recall can be expressed as Equations (12) and (13):

$$P = TP / (TP + FP) \quad (12)$$

$$R = TP / (TP + FN) \quad (13)$$

These two comprehensive parameters can be used to evaluate the segmentation performance comprehensively. The evaluations are shown in Table 5.

Table 5. The recognition results of the evaluation of the mask extractor under different block numbers.

Number	8*8	12*12	16*16	18*18	22*22	26*26
P-mean	0.92	0.90	0.91	0.92	0.91	0.85
P-variance	0.022	0.019	0.015	0.009	0.013	0.020
R-mean	0.35	0.61	0.82	0.93	0.91	0.87
R-variance	0.023	0.020	0.012	0.011	0.015	0.024

Table 5 shows that when the number of blocks is in the range (64, 144), the P-mean is high, but the R-mean is very low, which indicates that the integrity of the mask cannot be guaranteed when the number of blocks is small. The performance of the mask extractor is the best when the number of blocks is 256. P-variance and R-variance variances are large when the number of blocks is over 256, which indicates the Inception-V3 classifier performance decreases when the input image block is small.

3.5. Comparison of the Ideal Route and Visual Route

To verify the feasibility and accuracy of route planning in the field environment, the experiment was carried out in the maize plant plot. Figure 13 shows that the performance of the visual route tracking, taking the route planning in the laboratory environment as the benchmark. It shows that except for individual outliers, the visual path fluctuates around the ideal path in a small range, and has a stable positive and negative standard deviation.

Figure 13. Comparison of the ideal route and visual route.

4. Conclusions

This study set out to research the process of maize canopy landmark location under a background with non-significant color differences and realized visual route planning for agricultural intelligent equipment. This research has shown that there is a robust and effective block recognition method to identify landmarks based the fine-tuned inception-V3 classifier. The findings of this study suggest that stable visual semantic navigation can be achieved under the background of non-significant color differences. This research will serve as a base for future studies and visual navigation based on crop landmarks. A limitation of this study was that the optimal block size was bound by the Inception-V3 network structure. More research is required to eliminate the above limitation and promote the practical application of this method.

Author Contributions: Conceptualization, H.Z. and X.W.; Formal analysis, G.J.; Funding acquisition, X.W.; Investigation, Y.C.; Methodology, H.Z.; Project administration, Y.C.; Software, G.J. and S.L.; Visualization, S.L.; Writing—original draft, H.Z.

Funding: This research was funded by the following projects: Henan science and technology tackling key project (Grant: 182102110249); Key research projects of universities in Henan (Grant: 18A416002); Henan province innovation and entrepreneurship training platform for University Students (Grant: s201810466023).

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

References

1. Abbasgholipour, M.; Omid, M.; Keyhani, A.; Mohtasebi, S.S. Color image segmentation with genetic algorithm in a raisin sorting system based on machine vision in variable conditions. *Expert Syst. Appl.* **2011**, *38*, 3671–3678. [[CrossRef](#)]
2. Xiong, J.; Lin, R.; Liu, Z.; He, Z. The recognition of litchi clusters and the calculation of picking point in a nocturnal natural environment. *Biosyst. Eng.* **2018**, *166*, 44–57. [[CrossRef](#)]
3. Eizentals, P.; Oka, K.; Harada, A. Fruit Pose Estimation and Stem Touch Detection for Green Pepper Automatic Harvesting. In Proceedings of the International Symposium on Experimental Robotics, Tokyo, Japan, 3–6 October 2016; Springer: Cham, Switzerland, 2016.

4. Sun, S.; Wu, Q.; Jiao, L.; Long, Y. Recognition of Green Apples Based on Fuzzy Set Theory and Manifold Ranking Algorithm. *Optik* **2018**, *165*, 395–407. [[CrossRef](#)]
5. Zhang, Q.; Chen, M.E.S.; Li, B. A visual navigation algorithm for paddy field weeding robot based on image understanding. *Comput. Electron. Agric.* **2017**, *143*, 66–78. [[CrossRef](#)]
6. Kataoka, T.; Kaneko, T.; Okamoto, H.; Hata, S. Crop Growth Estimation System Using Machine Vision. In Proceedings of the IEEE/ASME International Conference on Advanced Intelligent Mechatronics, Kobe, Japan, 20–24 July 2003; IEEE: Piscataway, NJ, USA, 2003.
7. Hunt, E.R.; Cavigelli, M.; Daughtry, C.S.T.; McMurtrey, J.E.; Walthall, C.L. Evaluation of Digital Photography from Model Aircraft for Remote Sensing of Crop Biomass and Nitrogen Status. *Precis. Agric.* **2005**, *6*, 359–378. [[CrossRef](#)]
8. Hague, T.; Tillett, N.D.; Wheeler, H. Automated Crop and Weed Monitoring in Widely Spaced Cereals. *Precis. Agric.* **2006**, *7*, 21–32. [[CrossRef](#)]
9. Guijarro, M.; Pajares, G.; Riomoros, I.; Herrera, P.J.; Burgos-Artizzu, X.P.; Ribeiro, A. Automatic segmentation of relevant textures in agricultural images. *Comput. Electron. Agric.* **2011**, *75*, 75–83. [[CrossRef](#)]
10. Tellaeche, A.; Burgosartizzu, X.P.; Pajares, G.; Ribeiro, A. A vision-based method for weeds identification through the Bayesian decision theory. *Pattern Recognit.* **2008**, *41*, 521–530. [[CrossRef](#)]
11. Jeon, H.Y.; Tian, L.F.; Zhu, H. Robust Crop and Weed Segmentation under Uncontrolled Outdoor Illumination. *Sensors* **2011**, *11*, 6270–6283. [[CrossRef](#)] [[PubMed](#)]
12. 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](#)]
13. Guerrero, J.M.; Pajares, G.; Montalvo, M.; Romeo, J.; Guijarro, M. Support Vector Machines for crop/weeds identification in maize fields. *Expert Syst. Appl.* **2012**, *39*, 11149–11155. [[CrossRef](#)]
14. Guo, W.; Rage, U.K.; Ninomiya, S. Illumination invariant segmentation of vegetation for time series wheat images based on decision tree model. *Comput. Electron. Agric.* **2013**, *96*, 58–66. [[CrossRef](#)]
15. Bai, X.; Cao, Z.; Wang, Y.; Yu, Z.; Hu, Z.; Zhang, X.; Li, C. Vegetation segmentation robust to illumination variations based on clustering and morphology modelling. *Biosyst. Eng.* **2014**, *125*, 80–97. [[CrossRef](#)]
16. Vidovi, I.; Scitovski, R. Center-based clustering for line detection and application to crop rows detection. *Comput. Electron. Agric.* **2014**, *109*, 212–220. [[CrossRef](#)]
17. Choi, K.H.; Han, S.K.; Han, S.H.; Park, K.-H.; Kim, K.S.; Kim, S. Morphology-based guidance line extraction for an autonomous weeding robot in paddy fields. *Comput. Electron. Agric.* **2015**, *113*, 266–274. [[CrossRef](#)]
18. Jiang, G.; Wang, X.; Wang, Z.; Liu, H. Wheat rows detection at the early growth stage based on Hough transform and vanishing point. *Comput. Electron. Agric.* **2016**, *123*, 211–223. [[CrossRef](#)]
19. Javier, G.; Santiago, T.; Lucio, S.; Ricardo, C. Bounded memory probabilistic mapping of out-of-structure objects in fruit crops environments. *Comput. Electron. Agric.* **2018**, *151*, 11–20.
20. Lyu, H.K.; Park, C.H.; Han, D.H.; Seong, K. Orchard Free Space and Center Line Estimation Using Naive Bayesian Classifier for Unmanned Ground Self-Driving Vehicle. *Symmetry* **2018**, *10*, 355. [[CrossRef](#)]
21. Hinton, G.E.; Salakhutdinov, R.R. Reducing the Dimensionality of Data with Neural Networks. *Science* **2006**, *313*, 504–507. [[CrossRef](#)] [[PubMed](#)]
22. Krizhevsky, A.; Sutskever, I.; Hinton, G. Imagenet Classification with Deep Convolutional Neural Networks. In Proceedings of the NIPS, Lake Tahoe, CA, USA, 3–8 December 2012; Curran Associates Inc.: New York, NY, USA, 2012.
23. Mohanty, S.P.; Hughes, D.P.; Salathé, M. Using Deep Learning for Image-Based Plant Disease Detection. *Front. Plant Sci.* **2016**, *7*, 1419. [[CrossRef](#)]
24. Dyrmann, M.; Karstoft, H.; Midtby, H.S. Plant species classification using deep convolutional neural network. *Biosyst. Eng.* **2016**, *151*, 72–80. [[CrossRef](#)]
25. Alessandro, D.S.F.; Matte Freitas, D.; Gercina, G.D.S.; Hemerson, P.; Marcelo, T.F. Weed detection in soybean crops using ConvNets. *Comput. Electron. Agric.* **2017**, *143*, 314–324.
26. Inkyu, S.; Zongyuan, G.; Feras, D.; Ben, U.; Tristan, P.; Chris, M.C. DeepFruits: A Fruit Detection System Using Deep Neural Networks. *Sensors* **2016**, *16*, 1222. [[CrossRef](#)]
27. Park, K.; Hong, Y.K.; Kim, G.H.; Lee, J. Classification of apple leaf conditions in hyper-spectral images for diagnosis of Marssonina, blotch using mRMR and deep neural network. *Comput. Electron. Agric.* **2018**, *148*, 179–187. [[CrossRef](#)]

28. Szegedy, C.; Vanhoucke, V.; Shlens, J.; Wojna, Z. Rethinking the Inception Architecture for Computer Vision. *arXiv*, 2014; arXiv:1512.00567.
29. Xiu, Y.; Lin, H.; Wang, R.; Li, Q.; Yi, C. Study on the adaptability of corn machinery and agronomic requirements in the transition to precision agriculture. In Proceedings of the 2010 World Automation Congress, Kobe, Japan, 19–23 September 2010; pp. 483–488.
30. Wang, X.; Zhang, H.; Chen, Y. Research on maize canopy center recognition based on nonsignificant color difference segmentation. *PloS ONE* **2018**, *13*. [[CrossRef](#)]

© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).