

Article

Autonomous Detection of *Spodoptera frugiperda* by Feeding Symptoms Directly from UAV RGB Imagery

Jiedong Feng ¹, Yaqin Sun ^{1,2,*}, Kefei Zhang ^{1,2}, Yindi Zhao ^{1,2}, Yi Ren ¹, Yu Chen ^{1,2}, Huifu Zhuang ^{1,2} and Shuo Chen ¹

¹ School of Environment and Spatial Informatics, China University of Mining and Technology, Xuzhou 221116, China; stdfengjiedong@cumt.edu.cn (J.F.); profkzhang@cumt.edu.cn (K.Z.); zhaoyd@cumt.edu.cn (Y.Z.); 07182463@cumt.edu.cn (Y.R.); chenyu@cumt.edu.cn (Y.C.); huifuzhuang@cumt.edu.cn (H.Z.); stdchensh@cumt.edu.cn (S.C.)

² Key Laboratory of Land Environment and Disaster Monitoring MNR, China University of Mining and Technology, Xuzhou 221116, China

* Correspondence: syqin@cumt.edu.cn

Abstract: The use of digital technologies to detect, position, and quantify pests quickly and accurately is very important in precision agriculture. Imagery acquisition using air-borne drones in combination with the deep learning technique is a new and viable solution to replace human labor such as visual interpretation, which consumes a lot of time and effort. In this study, we developed a method for automatic detecting an important maize pest—*Spodoptera frugiperda*—by its gnawing holes on maize leaves based on convolution neural network. We validated the split-attention mechanism in the classical network structure ResNet50, which improves the accuracy and robustness, and verified the feasibility of two kinds of gnawing holes as the identification features of *Spodoptera frugiperda* invasion and the degree. In order to verify the robustness of this detection method against plant morphological changes, images at the jointing stage and heading stage were used for training and testing, respectively. The performance of the models trained with the jointing stage images has been achieved the validation accuracy of ResNeSt50, ResNet50, EfficientNet, and RegNet at 98.77%, 97.59%, 97.89%, and 98.07%, with a heading stage test accuracy of 89.39%, 81.88%, 86.21%, and 84.21%.

Keywords: *Spodoptera frugiperda*; deep learning; convolutional neural network; corn insect



Citation: Feng, J.; Sun, Y.; Zhang, K.; Zhao, Y.; Ren, Y.; Chen, Y.; Zhuang, H.; Chen, S. Autonomous Detection of *Spodoptera frugiperda* by Feeding Symptoms Directly from UAV RGB Imagery. *Appl. Sci.* **2022**, *12*, 2592. <https://doi.org/10.3390/app12052592>

Academic Editors: Pawel Kielbasa and Manuel Armada

Received: 5 January 2022

Accepted: 25 February 2022

Published: 2 March 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 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 (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Spodoptera frugiperda, originating from the American continent, has invaded Europe, Asia, and Africa [1]. As a migratory pest, *Spodoptera frugiperda* has a strong survival ability and a rapid reproduction rate, colonizing the above continents in a short time and causing great damage to corn, rice, and other main food crops [2–5].

At present, the main standard method to control this pest is pesticides, including (a) detecting the occurrence and status of pests by field sampling investigation, which relies on agronomists or trained surveyors [3], and (b) spraying pesticides evenly in the corresponding area [6,7]. It is simple and easy to indiscriminately spray, but the process of obtaining the information is time consuming and laborious, which depends on the subjectivity of surveyors. Uniform spraying would cause pesticide waste and environmental pollution [8,9]. In this context, there is an urgent need for a low-cost, high-efficiency, and high-precision method to quickly and effectively obtain field information, including the occurrence location, extent, and overall distribution of insect pests [10].

There have been several research studies focusing on the identification of pests and diseases affecting plant leaves. Most of the image data come from ground-based sensors such as mobile phones and digital cameras [11–13], and a small part of this is collected by unmanned aerial vehicles (UAV), which belong to remote sensing (RS) technology [14–16]. RS has been frequently adopted as a rapid, non-destructive, and cost-effective means for

plant disease and pest detection that can be adapted to different scenarios and different objects [17–19]. All of the abbreviations in the introduction are found in Table 1.

Table 1. Abbreviations.

CNN	Convolutional Neural Networks	UAV	Unmanned Aerial Vehicle
RS	Remote Sensing	RGB	R (red), G (green), B (blue)

Compared to satellite remote sensing and aerial remote sensing, UAVs have great advantages in terms of cost, operation, carrying, etc. [19], and they have been widely used in crop classification, growth monitoring, yield estimation, and other aspects, especially for large fields [20].

On the other hand, deep learning—originating from machine learning—has gradually gained popularity because of its ability to automatically extract representative features from a large number of input images [21,22]. Konstantinos et al. developed CNN models to perform plant disease detection and diagnosis using simple leaves images of healthy and diseased plants through deep learning methodologies [23]. Chen et al. used the UNet-based BLSNet to automatic identify and segment the diseased region of Rice bacterial leaf streak from the camera photos [24]. The appearance of the attention mechanism also further improves the performance of the network [22,25].

The following are studies based on UAV imagery combined with machine learning or deep learning: Tetila et al. detected soybean foliar diseases subjected to biological stress based on the simple linear iterative clustering segmentation method through foliar physical properties using RGB imagery captured by the low-cost unmanned aerial vehicle model DJI Phantom 3 [26]. Harvey et al. used an unmanned aerial vehicle (UAV) to acquire high-resolution images in the field, and they built an automated, high-throughput system based on a convolutional neural network (CNN) for the detection of northern leaf blight of maize plants [27]. Jin et al. proposed a computerized system based on CNN to process images captured by UAVs at low altitudes, which can detect *Fusarium* wilt of radish with high accuracy [28]. Ryo et al. used CNN to implement a detection method of virus-infected plants in a potato seed production field, with UAV RGB images being captured at an altitude of 5–10 m from the ground [7].

Compared with disease studies, insect pests are more flexible. There are two primary approaches to insect identification [29]: (i) direct, focusing on the ontology of the insects, and (ii) indirect, which focus on the damage caused by the insects [6]. For example, Liu et al. used a field insect light trap to obtain images and combined the CNN and attention mechanism to construct a direct classification model for insect identification [30]. Zhang focused on the significant change in the plant's leaf area index caused by *Spodoptera frugiperda* to indirectly monitor the infestation [31]. On the other hand, using the camera to closely capture pest images is also a widely used method, such as Li et al. integrating Convolutional Neural Network (CNN) and non-maximum inhibition for positioning and counting aphids in rice images obtained by a close view camera, achieving 0.93 accuracy and 0.885 mAP by optimizing key parameters and feature extraction network [13].

The above methods may have defects in accuracy or cannot be applied to large area practice. Thanks to the development of UAV technology, the pest identification based on UAV images is worth further research [32]. Ana et al. carry small aircraft RGB camera drones to obtain the vineyard plant image, and the application of geometric vision and computer vision technology, combined with landform factors on the influence of pests on the vineyard of the quantitative analysis for the farm digital management provides accurate low-cost information, which helps in the implementation and improvement of farm management and decision-making processes [33]. Farian et al. also used the corn leaves damaged by *Spodoptera frugiperda* and applied VGG16 and InceptionV3 to detect the infected corn leaves captured by the UAV (UAV) remote sensing technology while using

the angular detection method in computer vision to strengthen the feature representation and improve the detection accuracy [34].

This paper presents a CNN-based deep learning system for the automatic detection of maize leaves infected by *Spodoptera frugiperda*; from RGB UAV remote sensing images at high spatial resolution. UAV remote sensing images have excellent potential for agricultural data acquisition, while deep learning has agricultural data processing potential. Through the combination, this study is based on a multi-stage pest detection classification model applied to actual maize production environmental characteristics based on the ResNest model. The model has the following capabilities: (1) Collecting corn images from the actual field agricultural production conditions for the automatic detection of leaves infected by *Spodoptera frugiperda*; (2) According to the feeding characteristics of corn grass, Accurately and quickly determining the pest stage of the infected leaves, providing a reliable reference for the formulation and implementation of prevention and control measures; (3) The potential and generalization ability of indirect pest detection based on UAV remote sensing images are verified. This provides a reference for the automated detection of pest invasion status in the field. The remainder of this paper is organized as follows: In Section 2, we describe the study area, data collection, and methods. Section 3 presents results, and Section 4 provides a discussion. Finally, Section 5 summarizes this work and highlights future works.

2. Materials and Methods

2.1. Study Area

The UAV RGB imagery of the maize pest *Spodoptera frugiperda* was captured at the Experimental Station (117.552616, 34.309942) of the China University of Mining and Technology, Xuzhou city, Jiangsu Province, China. Image are shown in Figure 1. The experimental site was invaded by the grass moth because of a later maize planting cycle than the surrounding fields. When we started the data collection, the larvae in the field were in a transition phase from low to medium age.

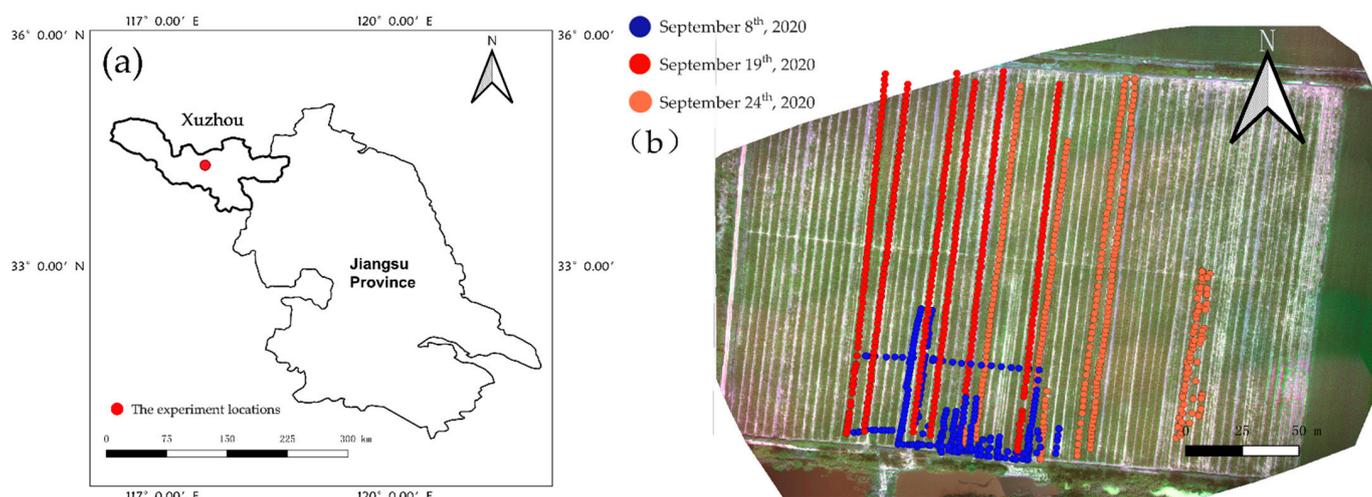


Figure 1. (a) The study area located in Xuzhou and the experimental field; (b) images of the experimental field obtained on 8, 9, and 24 September 2020 using a low-altitude UAV equipped with an RGB sensor. Each of the colors represents a different experiment date. Each drone image has its own coordinate.

2.2. UAV Image Collection

The image acquisition device was a DJI Mavic air2 equipped with a half-inch CMOS sensor, which can achieve an effective pixel rate of 48 million. What is more, it is an ultra-small drone, weighing just 570 g, which is capable of capturing high-resolution images (standard red–green–blue or RGB photos) of corn at ultra-low altitudes.

The data were collected three times in September 2020 during the critical growth period of maize across the jointing stage to the heading stage, and the specific time and resolution are shown in Table 2. In the jointing stage, corn was collected for the first time, and in the heading stage, it was collected for the second and third time. Between these two periods, maize grows rapidly, and its morphology changes greatly, especially the appearance of stamen changes the overall morphology of the maize plant to a great extent, which can be applied to the generalized type test of model. That is the reason we set the interval. The specific differences between the two stages are shown in Figure 2Part A below, and maize leaf categories are shown in the Figure 2Part B.

Table 2. Date and quantity of data collection.

Date	#Images	Stage	Image Resolution
8 September 2020	295	Jointing	4000 × 3000
19 September 2020	249	Heading	6000 × 8000
24 September 2020	84	Heading	6000 × 8000
	272		4000 × 3000

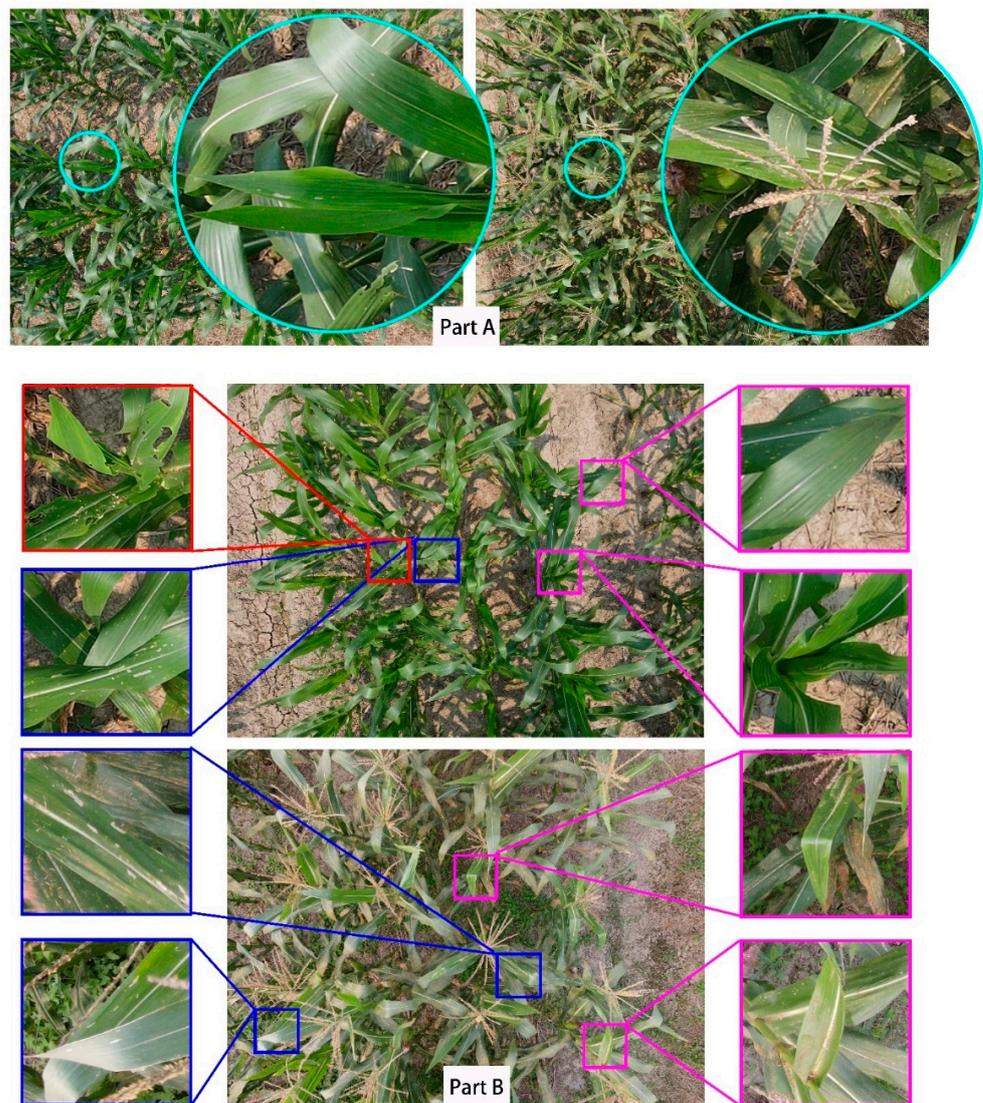


Figure 2. (Part A) The picture on the left shows corn in the jointing stage, the right picture shows corn in the heading stage; (Part B) The red box shows severely infected corn leaves, the blue box shows slightly infected corn leaves, and the purple boxes show corn leaves in healthy condition.

The flight speed was controlled at 1.5 m/s and the flight altitude was from 2 to 5 m from the ground, which was very close to the corn canopy. Moreover, the shooting angle of the flight path along the ridges of the field was 90° , with more efficient harvesting of corn canopy information. The specific information of the images is shown in Table 2, including the date, number, and image resolution.

2.3. Image Preprocessing

The data processing included two main steps: Cropping and Classing.

Cropping: Due to the size of the images, we tailored them from the sizes in Table 2 to 200×200 to speed up training and to reduce the pressure on the graphics memory. We used the OpenCV-Python tool in Python language to read and crop images in batches. For visual effect, Figure 3 shows the conversion process of an image from $1 \times 3000 \times 4000$ to $25 \times 200 \times 200$.

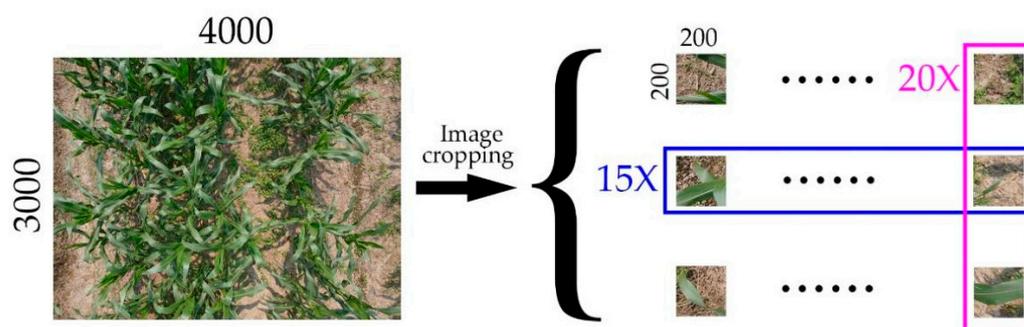


Figure 3. An image cropped from $1 \times 3000 \times 4000$ to $300 \times 200 \times 200$.

Classing: After cropping, the main body of the image is basically composed of corn leaves, and part of the image is land and weeds. By combining the edge detection tool in Opencv2-Python library and RGB channel calculation, the image containing only land and weeds is removed. According to the habits of *Spodoptera frugiperda* and specific representation on the image, we divided the rest of the images into 3 categories by visual interpretation, as shown in Figure 4 below.

(Condition 1): Healthy leaves: Green healthy leaves, complete without damage.

(Condition 2): Translucent silver window (TSW), the first instar and part of the second instar of *Spodoptera frugiperda* only feed on one side of corn leaves, creating a translucent silver windowpane.

(Condition 3): Irregular wormhole (IW), the rest of the instars of *Spodoptera frugiperda* cause significant irregular pores, and some of the leaves infected in the undeveloped period show symmetry in the holes.

(Condition 4): Other objects, the picture contains only land and weeds (drop).

For pest control, the earlier the intervention, the fewer losses and pesticides used, so the first translucent silver windowpane is the most important object. However, in the image, most of the leaves were presented as healthy. In order to balance the number of positive and negative samples, we selected the number of healthy leaves in condition 1 according to the number of infected leaves in conditions 2 and 3, and we removed the leaves in condition 4 at the same time.

After the above processing, we finally obtained more than 5000 maize images in the joint stage, including 2043 healthy leaves, 1866 condition 1 images, and 1430 condition 2 images. The quantities are shown in Table 3. The above data are used for training and validation of the models. During the training, the ratio of training to verification was 9:1.

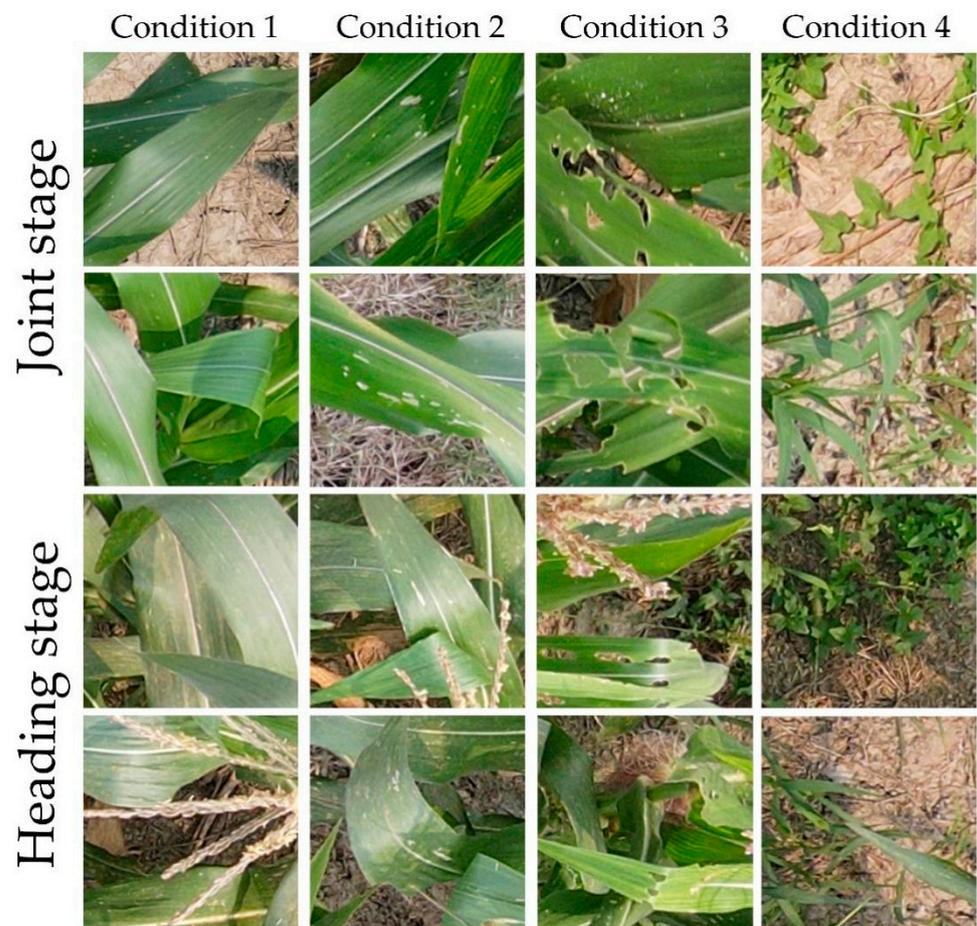


Figure 4. Classing: Four categories, two stages.

Table 3. Number of processed images.

Stage	Health	TSW	IW	Sum
Jointing	2043	1904	1744	5691
Heading	532	417	596	1545

To test the robustness of the model detection ability, 1545 images in the heading stage were used, including 532 in condition 1 and 417 in condition 2. This part of the image does not participate in training and verification at all, and it was used as an independent test set for testing the model after training.

2.4. Augmentation

The training of a deep learning model requires a very large amount of data, so we used data augmentation to amplify the data [35]. Data enhancement technology can enhance images and reduce over-fitting by flipping, mirroring, and contrast transformation without changing the original form of an image [36]. In the classification task, the image geometric transformation, color space enhancement, random erasure, and feature space enhancement operation can change the image status without changing the image category so as to improve the quantity and quality of data, play the effect of reducing the distance between the training data and the test data, and reduce the overfitting in the model training process [37,38].

For example, the contrast change can change the brightness of the image to a certain extent and enhance the sensitivity of the image to the illumination change. In this study, we used a variety of image enhancement methods, as shown in Figure 5 below.

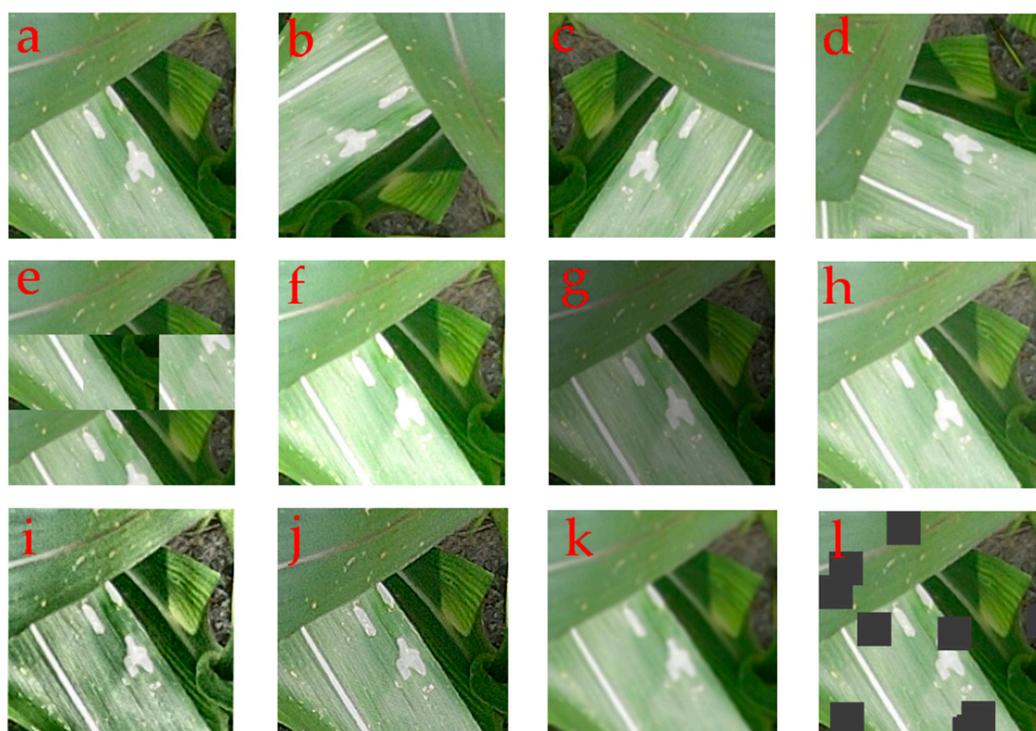


Figure 5. Augmentation: (a) The original image; (b,c) Horizontal flips and rotations of 90° , respectively; (d) Shift scale rotate; (e) Random grid transformation; (f,g) Contrast transformation; (h) Random Contrast transformation; (i) IAASharp; (j) IAAEmb; (k) CLAHE; (l) Shift cutout.

2.5. Convolutional Neural Network

After a long period of development since AlexNet [39], convolutional neural networks, composed of a convolution layer, a pooling layer, and a fully connected layer, have evolved into series of models that can automate the extraction of features through training iterations [40,41]. ResNet [42], through the application of residual blocks, has solved the problem of network degradation and parameter disappearance with a continuous increase in neural network layers, making an indelible contribution to the progress of deep learning, called Deep Convolutional Neural Network (DCNN). DCNN can automatically extract the features of convolutional check images with different specifications to obtain higher data classification accuracy, and it has become the most common identification method [43].

Based on ResNet, the integration of different methods leads to the development of various network structures, such as grouping convolution [44], self-attention mechanism [45], and selective attention mechanism [46]. Therefore, in this study, the feasibility of the feeding symptoms method based on maize *Spodoptera frugiperda* was verified by using several kinds of ResNet related networks, including ResNet, ResNeS [47], SE-Net, and SK-Net. Although the residual structure has been widely applied with its simple structure and convenient modular design, its performance is not satisfactory in downstream applications, which is affected by factors such as receptive field size and channel interaction. Recently, the successful application of the channel and attention mechanism has introduced new possibilities for its improvement. ResNext first introduced the idea of grouping convolution. The SE-Net introduces a channel-attention mechanism for feature construction by adaptively recalibrating the channel feature response. SK-Net extracted the channel information from the feature map through the construction of the grouped channels. Therefore, according to the idea of taking the channel as the operation unit and dividing the input data into more fine-grained weighted subgroups or subchannels based on the global context, it is able to build a channel-based split attention structure. During training, each subgroup is able to perform different mapping abstractions on the input channel data of its own part so as to build different feature representations. In the model, the module is named a distraction

block. Thanks to the simple and modular structure, the distraction blocks can perform multiple reuse and stacking and then construct the universal structure bodies similar to the same residue model. Therefore, the block can be simply described as replacing the original residual part of the attention operations with channels as units and thus giving the corresponding weight to the identity.

At the same time, to increase the comparison, other classic network models such as EfficientNet [48] and RegNet [49] are also selected. The networks used in this study were consistent with the architecture of ResNet50, and the original block was replaced with a split-attention block, as shown in the Figure 6 below.

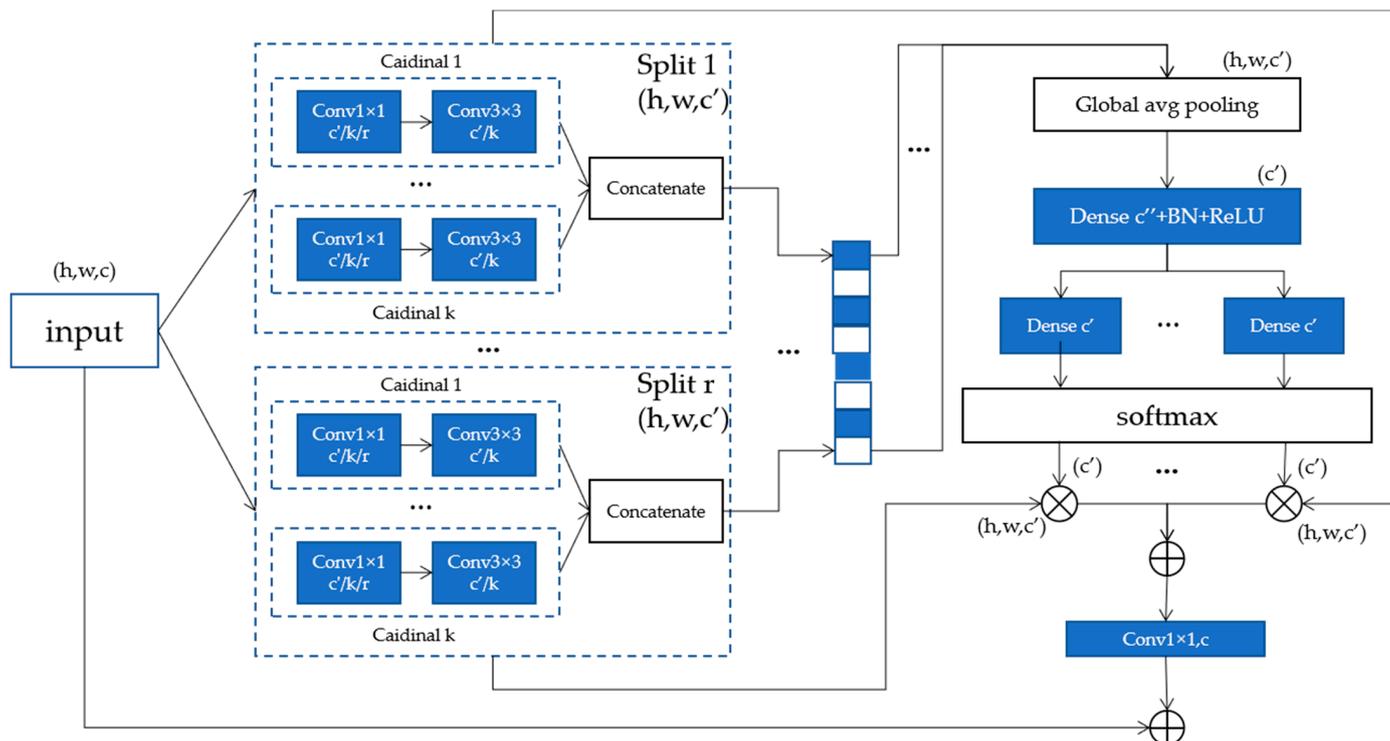


Figure 6. ResNest block.

2.6. Transfer Learning

Transfer learning can transplants the weights obtained through the pre-training of large data sets to the network. Fine tuning based on these weights can accelerate the network training speed and reduce the amount of data required for training [50,51]. In this paper, the ImageNet data set [52] was used as the source data to pre-train the model.

3. Results

3.1. Experimental Setup

In this experiment, we used Pytorch 1.4 as the framework, which is an open-source package in deep learning based on the python programming language. The selected optimizer was Stochastic Gradient Descent (SGD) with the momentum of 0.9, the initial learning rate was 0.002, which decreases with loss, the batch size was 32, and the loss function was CrossEntropy. The training was performed on a machine with the graphics processor of NVIDIA GTX2080s and 32 GB of memory. We trained and tested the models with the data set consisting of the corn leaf in the jointing stage, and to valid the robust ability, we tested the models with the heading stage data set. Figure 7 illustrates the processes involved in obtaining the images used for the experiments.

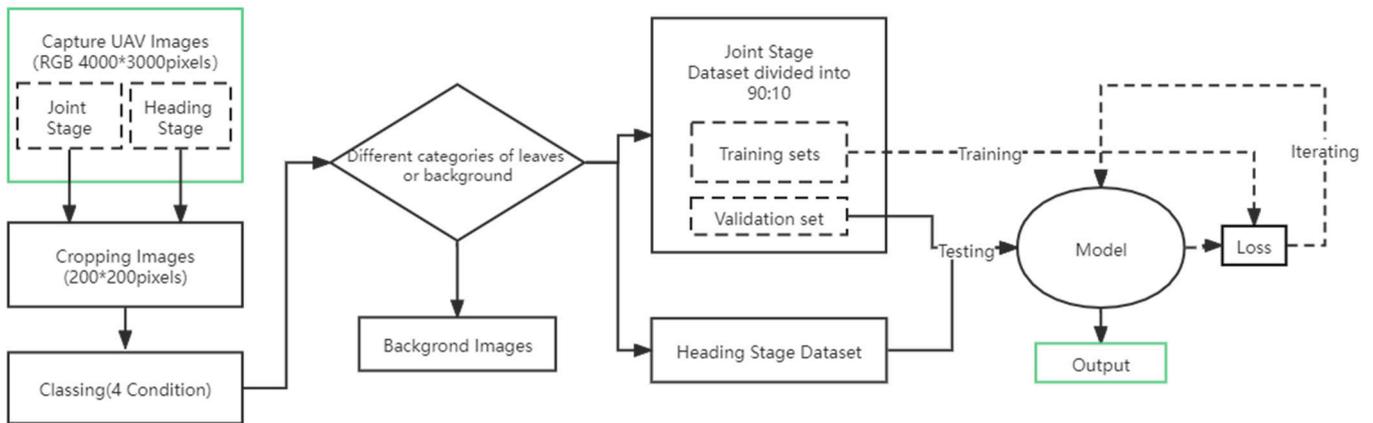


Figure 7. Illustration of the process used in this study.

3.2. Evaluation Parameters

Model performance was assessed using six parameters: Accuracy, Sensitivity, Specificity, Precision, F1 Score, and Kappa.

Accuracy is used as the main method to calculate the accuracy of a model, and the classification ability of the model is represented by the proportion of the correct number of samples in the total number of samples. The specific formula is as follows, and in the formula, T (True) and F (False) represent whether the prediction is correct, and P (Positive) and N (Negative) represent the category of the model prediction; TP and TN is the sum of True Predictions, and FP and FN are the opposite:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

Sensitivity represents the model’s recognition ability for positive samples, consisting of TP and FN:

$$Sensitivity = \frac{TP}{TP + FN} \tag{2}$$

Specificity is defined to show true negative assessment ability, consisting of TN and FP:

$$Specificity = \frac{TN}{TN + FP} \tag{3}$$

Precision shows the accuracy of all model-identified positive samples, consisting of TN and FP:

$$Precision = \frac{TP}{TP + FP} \tag{4}$$

F1 Score is an aggregative indicator based on the harmonic mean of precision and recall.

$$F1\ score = \frac{2 * Precision * Sensitivity}{Precision + Sensitivity} \tag{5}$$

Kappa coefficient is a consistency test confusion matrix-based indicator with values between −1 and 1; closer to 1 indicates the overall effect of the classification. In the formula, a_i and b_i represent the true number and the predicted number of the i category, respectively, and the sum means the number of all data.

$$Kappa = \frac{p_0 - p_1}{1 - p_1} \tag{6}$$

$$p_0 = Accuracy$$

$$p_1 = \frac{a_1 \times b_1 + a_3 \times b_3 + a_3 \times b_3}{sum^2}$$

3.3. Experimental Results

In this study, we compared the performance of four models in the data set, and the results of each model are shown in Figure 8 below. The accuracy for ResNeSt50, ResNet50, EfficientNet, and RegNet is 98.77%, 97.59%, 97.89%, and 98.07%. It can be seen from the data that all the four network structures can obtain high reconnaissance accuracy in this classification problem. Among them, the accuracy of ResNet50 with split attention is the highest. At the same time, with the addition of transfer learning, all networks basically reach the steady state in about 20 epochs, which means in the production environment, we can complete the training and validation of the model in a relatively short time.

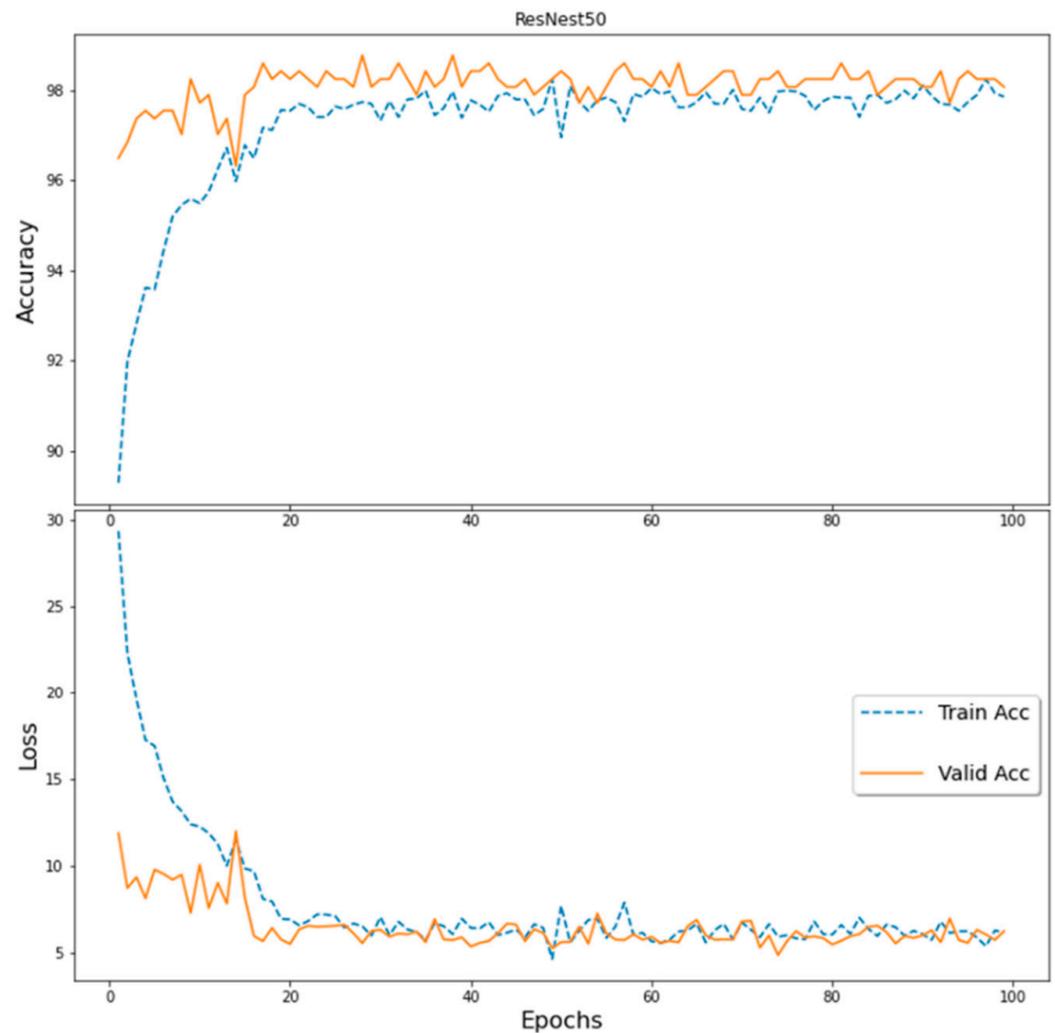


Figure 8. Accuracy and loss of the ResNest50 during training and validation.

An image was randomly selected from the test data set (clipping completed), and the model operation was carried out. The infected image blocks in the calculation results were given different colors according to their severity, and then, they were spliced together for display. A blue box represents a slight silver window, while a red box represents an irregular wormhole (see Figure 9).



Figure 9. Visualization of model’s outputs. The upper part of the figure is the original image, the lower part is the image after adding the infected color block, and the left and right yellow image cells are the enlarged image of the infected area.

4. Discussion

In this study, we proposed the deep learning model to detect the invasion of *Spodoptera frugiperda* by the features of the damaged leaves. At the same time, four different neural network structures are used to verify the feasibility of the proposed method. In addition, in order to verify the ability of this feature against maize morphological change, the four models were trained on the images at the jointing stage and tested on the images at the heading stage. The appearance of stamen and the fall of stamen in corn leaves images at the heading stage has a certain influence on the overall structure and color of the image. However, the neural network based on features of the damaged leaves still has good accuracy. Accuracy, Sensitivity, Specificity, Precision, F1 score, and Kappa were used to demonstrate the recognition ability of maize leaves with holes (see Table 4) (TSW: Translucent silver window, IW: Irregular wormhole).

Table 4. Test results—based on images at the heading stage.

Model	Class	Health	TSW	IW	Accuracy	Sensitivity	Specificity	Precision	F1 Score	Kappa
ResNeSt50	Health	473	47	12	89.39%	0.950.86	0.97	0.94	0.91	0.84
	TSW	7	398	12						
	IW	22	64	510						
ResNet50	Health	390	37	105	81.88%	0.84	0.91	0.79	0.81	0.72
	TSW	14	349	54						
	IW	13	57	526						
EfficientNet	Health	456	39	37	86.21%	0.86	0.95	0.91	0.88	0.79
	TSW	18	377	22						
	IW	25	72	499						
RegNet	Health	419	30	83	84.21%	0.79	0.97	0.94	0.86	0.76
	TSW	15	342	60						
	IW	10	46	540						

It can be seen from the table that the models based on four different network structures all have a good ability to identify the infected leaves from the corn images at the heading stage. However, compared with the original valid accuracy, the current accuracy has a degree of decline, respectively 89.39%, 81.88%, 86.21%, and 84.21%. The split-attention

models outperformed the origin ResNet50 structures and the classical network model on the performance in terms of Accuracy, Precision, etc. ReNest50 also achieves the best results on the Kappa coefficients and F1 Score. What is more, we can see the differences between these networks more clearly in the CAM (Class Activation Map based on the average gradient) in Figure 10 below. Compared with the results of other network structures, the split-attention network can identify the target more accurately and closely.

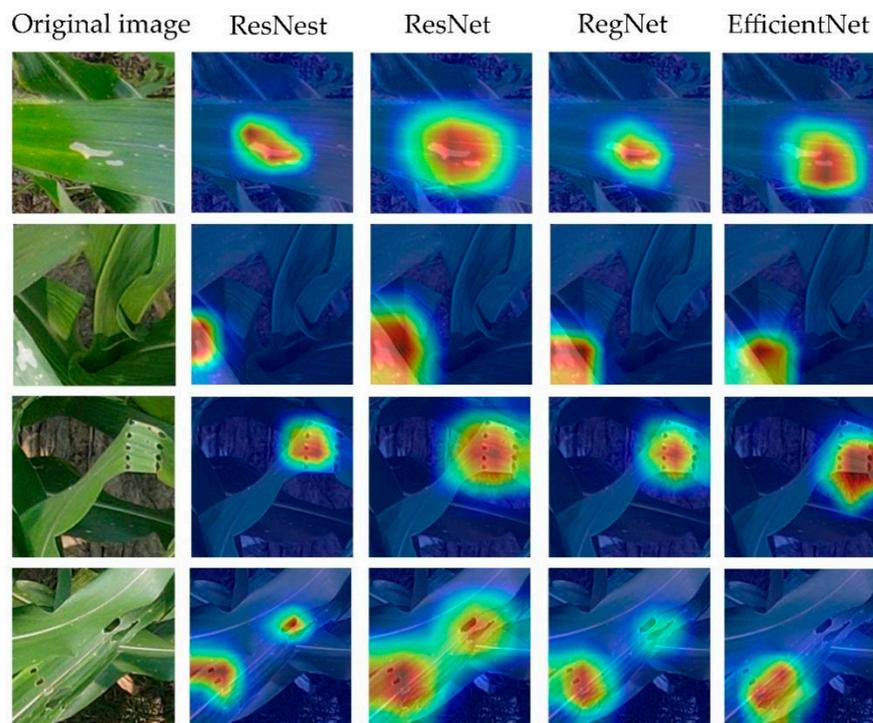


Figure 10. Attention map base on grad CAM.

5. Conclusions and Future Directions

This study aimed to detect maize images that included leaves infected by *Spodoptera frugiperda* in the early stages. Four different models including ResNeSt50, ResNet50, EfficientNet, and RegNet were used to verify the feasibility of using the above features for recognition and explore the split-attention mechanism to improve the accuracy and robustness of the model. The ResNeSt50 network achieved a high accuracy of 98.77% in the validation data set based on the jointing stage and of 89.39% in the test data set based on the heading stage. The model demonstrated its ability to identify infected maize leaves at various stages and allowed to classify them according to the degree of infection. In the process of model construction, methods such as data enhancement and transfer learning are adopted to speed up model construction, reduce overfitting, and improve robustness. Accurate treatment can carry out according to an image's coordinates the grade and distribution of infected leaves, which can significantly reduce the use of pesticides and assist in the implementation of biological control.

Although the model can accurately and quickly identify and judge the maize leaves present in the image for insect pests, the following problems still need to be further studied and explored in practical application: (1) under the condition of positive projections, so that some pest leaves may be ignored due to occlusion; (2) the image acquisition parameters such as height, angle, and resolution and actual field planting conditions; (3) the overall statistical analysis of field pest distribution and subsequent application should be further explored with agronomic knowledge. In future research, we will use the model of *Spodoptera frugiperda* based on more accurate network architecture for real-time field corn image recognition. In addition, according to the optimal resolution combination obtained, we

will conduct a new round of data collection in this year's maize planting period to further verify the method. At the same time, we will collect more data to build a model that can identify more pests and diseases faster and more accurately.

Author Contributions: Conceptualization, Y.S. and J.F.; methodology, J.F.; software, J.F.; validation, Y.Z., J.F. and Y.R.; formal analysis, J.F.; investigation, S.C. and J.F.; resources, S.C.; data curation, S.C. and J.F.; writing—original draft preparation, J.F.; writing—review and editing, Y.S., K.Z., Y.C., H.Z. and Y.Z.; visualization, J.F. and Y.R.; supervision, K.Z.; project administration, Y.S. and K.Z.; funding acquisition, K.Z. and Y.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Fundamental Research Funds for the Central Universities (Grant No. 2017XKQY019). This work was supported by a project founded by the Priority Academic Program Development (PAPD) of Jiangsu Higher Education Institutions.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available within the article.

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

References

1. Spark, A.N. A Review of the Biology of the Fall Armyworm. *Florida Entomol.* **1979**, *62*, 82–87. [CrossRef]
2. Food and Agriculture Organization of the United Nations. Map of the Worldwide Spread of Fall Armyworm since 2016. 2022. Available online: <http://www.fao.org/fall-armyworm/monitoring-tools/faw-map/en/> (accessed on 24 February 2022).
3. Sarkowi, F.N.; Mokhtar, A.S. The fall armyworm (Faw) spodoptera frugiperda: A review on biology, life history, invasion, dispersion and control. *Outlooks Pest Manag.* **2021**, *32*, 27–32. [CrossRef]
4. Poisot, A.; Hruska, A.; Fredrix, M. Integrated Management of the Fall Armyworm on Maize. 2018. Available online: <https://www.preventionweb.net/publication/integrated-management-fall-armyworm-maize-guide-farmer-field-schools-africa> (accessed on 24 February 2022).
5. Ministry of Agriculture and Rural Affairs. The General Office of the Ministry of Agriculture and Rural Affairs on Continuously Strengthening the Prevention and Control of Grass Moths. 2019. Available online: http://www.zzys.moa.gov.cn/tzgg/201907/t20190731_6321854.htm (accessed on 24 February 2022). (In Chinese)
6. Bieganski, A.; Dammer, K.; Siedliska, A.; Bzowska-Bakalarz, M.; Bereś, P.K.; Dąbrowska-Zielińska, K.; Pflanz, M.; Schirrmann, M.; Garz, A. Sensor-based outdoor monitoring of insects in arable crops for their precise control. *Pest Manag. Sci.* **2020**, *77*, 1109–1114. [CrossRef] [PubMed]
7. Sugiura, R.; Tsuda, S.; Tamiya, S.; Itoh, A.; Nishiwaki, K.; Murakami, N.; Shibuya, Y.; Hirafuji, M.; Nuske, S. Field phenotyping system for the assessment of potato late blight resistance using RGB imagery from an unmanned aerial vehicle. *Biosyst. Eng.* **2016**, *148*, 1–10. [CrossRef]
8. Dammer, K.; Adamek, R. Sensor-Based Insecticide Spraying to Control Cereal Aphids and Preserve Lady Beetles. *Agron. J.* **2012**, *104*, 1694–1701. [CrossRef]
9. Karimzadeh, R.; Hejazi, M.J.; Helali, H.; Iranipour, S.; Mohammadi, S.A. Assessing the impact of site-specific spraying on control of *Eurygaster integriceps* (Hemiptera: Scutelleridae) damage and natural enemies. *Precis. Agric.* **2011**, *12*, 576–593. [CrossRef]
10. Bock, C.H.; Poole, G.H.; Parker, P.E.; Gottwald, T.R. Plant Disease Severity Estimated Visually, by Digital Photography and Image Analysis, and by Hyperspectral Imaging. *Crit. Rev. Plant Sci.* **2010**, *29*, 59–107. [CrossRef]
11. Nanni, L.; Maguolo, G.; Pancino, F. Research on insect pest image detection and recognition based on bio-inspired methods. *arXiv* **2019**, arXiv:1910.00296.
12. Thenmozhi, K.; Reddy, U.S. Crop pest classification based on deep convolutional neural network and transfer learning. *Comput. Electron. Agric.* **2019**, *164*, 104906. [CrossRef]
13. Li, W.; Chen, P.; Wang, B.; Xie, C. Automatic Localization and Count of Agricultural Crop Pests Based on an Improved Deep Learning Pipeline. *Sci. Rep.* **2019**, *9*, 1–11. [CrossRef]
14. Zhang, X.; Yue, Q.; Meng, F.; Fan, C.; Zhang, M. Identification of maize leaf diseases using improved deep convolutional neural networks. *IEEE Access* **2018**, *6*, 30370–30377. [CrossRef]
15. Alvaro, F.; Sook, Y.; Sang, K.; Dong, P. A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition. *Sensors* **2017**, *17*, 2022. [CrossRef]
16. Selvaraj, M.G.; Vergara, A.; Ruiz, H.; Safari, N.; Blomme, G. AI-powered banana diseases and pest detection. *Plant Methods* **2019**, *15*, 1–11. [CrossRef]
17. Salami, E.; Barrado, C.; Pastor, E. UAV Flight Experiments Applied to the Remote Sensing of Vegetated Areas. *Remote Sens.* **2014**, *6*, 11051–11081. [CrossRef]

18. Zhang, C.; Kovacs, J.M. The application of small unmanned aerial systems for precision agriculture: A review. *Precis. Agric.* **2012**, *13*, 693–712. [[CrossRef](#)]
19. Matese, A.; Toscano, P.; di Gennaro, S.F.; Genesio, L.; Vaccari, F.P.; Primicerio, J.; Belli, C.; Zaldei, A.; Bianconi, R.; Gioli, B. Intercomparison of UAV, Aircraft and Satellite Remote Sensing Platforms for Precision Viticulture. *Remote Sens.* **2015**, *7*, 2971–2990. [[CrossRef](#)]
20. Zhang, X.; Zhang, K.; Sun, Y.; Zhao, Y.; Zhuang, H.; Ban, W.; Chen, Y.; Fu, E.; Chen, S.; Liu, J.; et al. Combining Spectral and Texture Features of UAS-Based Multispectral Images for Maize Leaf Area Index Estimation. *Remote Sens.* **2022**, *14*, 331. [[CrossRef](#)]
21. Toda, Y.; Okura, F. How Convolutional Neural Networks Diagnose Plant Disease. *Plant Phenomics* **2019**, *2019*, 9237136. [[CrossRef](#)]
22. Zhao, S.; Peng, Y.; Liu, J.; Wu, S. Tomato Leaf Disease Diagnosis Based on Improved Convolution Neural Network by Attention Module. *Agriculture* **2021**, *11*, 651. [[CrossRef](#)]
23. Ferentinos, K.P. Deep learning models for plant disease detection and diagnosis. *Comput. Electron. Agric.* **2017**, *145*, 311–318. [[CrossRef](#)]
24. Chen, S.; Zhang, K.; Zhao, Y.; Sun, Y.; Ban, W.; Chen, Y.; Zhuang, H.; Zhang, X.; Liu, J.; Yang, T. An Approach for Rice Bacterial Leaf Streak Disease Segmentation and Disease Severity Estimation. *Agriculture* **2021**, *11*, 420. [[CrossRef](#)]
25. Liu, J.; Zhang, K.; Wu, S.; Shi, H.; Zhao, Y.; Sun, Y.; Zhuang, H.; Fu, E. An Investigation of a Multidimensional CNN Combined with an Attention Mechanism Model to Resolve Small-Sample Problems in Hyperspectral Image Classification. *Remote Sens.* **2022**, *14*, 785. [[CrossRef](#)]
26. Tetila, E.C.; Machado, B.B.; Belete, N.; Guimaraes, D.A.; Pistori, H. Identification of Soybean Foliar Diseases Using Unmanned Aerial Vehicle Images. *IEEE Geosci. Remote Sens. Lett.* **2017**, *1*, 5. [[CrossRef](#)]
27. Wu, H.; Wiesner-Hanks, T.; Stewart, E.L.; DeChant, C.; Kaczmar, N.; Gore, M.A.; Nelson, R.J.; Lipson, H. Autonomous Detection of Plant Disease Symptoms Directly from Aerial Imagery. *Plant Phenome J.* **2019**, *2*, 1–9. [[CrossRef](#)]
28. Ha, J.G.; Moon, H.; Kwak, J.T.; Hassan, S.I.; Dang, M.; Lee, O.N.; Park, H.Y. Deep convolutional neural network for classifying Fusarium wilt of radish from unmanned aerial vehicles. *J. Appl. Remote Sens.* **2017**, *11*, 1. [[CrossRef](#)]
29. Lima, M.C.F.; Leandro, M.E.D.D.A.; Valero, C.; Coronel, L.C.P.; Bazzo, C.O.G. Automatic Detection and Monitoring of Insect Pests—A Review. *Agriculture* **2020**, *10*, 161. [[CrossRef](#)]
30. Liu, L.; Wang, R.; Xie, C.; Yang, P.; Wang, F.; Sudirman, S.; Liu, W. PestNet: An End-to-End Deep Learning Approach for Large-Scale Multi-Class Pest Detection and Classification. *IEEE Access* **2019**, *7*, 45301–45312. [[CrossRef](#)]
31. Zhang, J.; Huang, Y.; Yuan, L.; Yang, G.; Jingcheng, Z.; Zhao, C. Using satellite multispectral imagery for damage mapping of armyworm (*Spodoptera frugiperda*) in maize at a regional scale. *Pest Manag. Sci.* **2016**, *72*, 335–348. [[CrossRef](#)]
32. Suwa, K.; Cap, Q.H.; Kotani, R.; Uga, H.; Kagiwada, S.; Iyatomi, H. A comparable study: Intrinsic difficulties of practical plant diagnosis from wide-angle images. In Proceedings of the IEEE International Conference on Big Data, Los Angeles, CA, USA, 9–12 December 2019; pp. 5195–5201. [[CrossRef](#)]
33. Del-Campo-Sanchez, A.; Ballesteros, R.; Hernandez-Lopez, D.; Ortega, J.F.; Moreno, M.A.; on behalf of Agroforestry and Cartography Precision Research Group. Quantifying the effect of *Jacobiasca lybica* pest on vineyards with UAVs by combining geometric and computer vision techniques. *PLoS ONE* **2019**, *14*, e0215521. [[CrossRef](#)]
34. Ishengoma, F.S.; Rai, I.A.; Said, R.N. Identification of maize leaves infected by fall armyworms using UAV-based imagery and convolutional neural networks. *Comput. Electron. Agric.* **2021**, *184*, 106124. [[CrossRef](#)]
35. Blok, P.M.; van Evert, F.K.; Tielen, A.P.M.; van Henten, E.J.; Kootstra, G. The effect of data augmentation and network simplification on the image-based detection of broccoli heads with Mask R-CNN. *J. Field Robot.* **2021**, *38*, 85–104. [[CrossRef](#)]
36. Kuznichov, D.; Zvirin, A.; Honen, Y.; Kimmel, R. Data Augmentation for Leaf Segmentation and Counting Tasks in Rosette Plants. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, Long Beach, CA, USA, 16–17 June 2019; pp. 2580–2589. [[CrossRef](#)]
37. El-Kenawy, E.-S.M.; Ibrahim, A.; Mirjalili, S.; Eid, M.M.; Hussein, S.E. Novel Feature Selection and Voting Classifier Algorithms for COVID-19 Classification in CT Images. *IEEE Access* **2020**, *8*, 179317–179335. [[CrossRef](#)]
38. Shorten, C.; Khoshgoftaar, T.M. A survey on Image Data Augmentation for Deep Learning. *J. Big Data* **2019**, *6*, 60. [[CrossRef](#)]
39. Gonzalez, T.F. *Handbook of Approximation Algorithms and Metaheuristics*; Chapman and Hall/CRC: New York, NY, USA, 2007; 1432p. [[CrossRef](#)]
40. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [[CrossRef](#)] [[PubMed](#)]
41. Girshick, R.; Donahue, J.; Darrell, T.; Malik, J. Rich feature hierarchies for accurate object detection and semantic segmentation. In Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23–28 June 2014; pp. 580–587. [[CrossRef](#)]
42. 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, 27–30 June 2016.
43. Zhang, X.; Han, L.; Dong, Y.; Shi, Y.; Huang, W.; Han, L.; González-Moreno, P.; Ma, H.; Ye, H.; Sobeih, T. A Deep Learning-Based Approach for Automated Yellow Rust Disease Detection from High-Resolution Hyperspectral UAV Images. *Remote Sens.* **2019**, *11*, 1554. [[CrossRef](#)]
44. Xie, S.; Girshick, R.; Dollar, P.; Tu, Z.; He, K. Aggregated residual transformations for deep neural networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 5987–5995. [[CrossRef](#)]

45. Hu, J.; Shen, L.; Albanie, S.; Sun, G.; Wu, E. Squeeze-and-Excitation Networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2020**, *42*, 2011–2023. [[CrossRef](#)]
46. Li, X.; Wang, W.; Hu, X.; Yang, J. Selective kernel networks. In Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 15–20 June 2019; pp. 510–519.
47. Zhang, H.L.H.; Wu, C.; Zhang, Z.; Zhu, Y. ResNeSt: Split-Attention Networks. *arXiv* **2020**, arXiv:2004.08955.
48. Tan, M.; Le, Q.V. EfficientNet: Rethinking model scaling for convolutional neural networks. In Proceedings of the Thirty-sixth International Conference on Machine Learning, ICML, Long Beach, CA, USA, 9–15 June 2019; pp. 10691–10700.
49. Radosavovic, I.; Kosaraju, R.P.; Girshick, R.; He, K.; Dollar, P. Designing network design spaces. In Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 13–19 June 2020; pp. 10425–10433.
50. Afifi, A.; Alhumam, A.; Abdelwahab, A. Convolutional Neural Network for Automatic Identification of Plant Diseases with Limited Data. *Plants* **2020**, *10*, 28. [[CrossRef](#)]
51. Hu, R.; Zhang, S.; Wang, P.; Xu, G.; Wang, D.; Qian, Y. The identification of corn leaf diseases based on transfer learning and data augmentation. *ACM Int. Conf. Proceeding Ser.* **2020**, 58–65. [[CrossRef](#)]
52. Deng, J.; Dong, W.; Socher, R.; Li, L.-J.; Li, K.; Fei-Fei, L. ImageNet: A large-scale hierarchical image database. In Proceedings of the 2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 20–25 June 2009; pp. 248–255.