Peer-reviewed version available at Remote Sens. 2019, 11, 2658; doi:10.3390/rs11222658

Article

Detection of *Fusarium* Head Blight in Wheat Using a Deep Neural Network and Color Imaging

Ruicheng Qiu [1], Ce Yang [2,*], Ali Moghimi [2], Man Zhang [1], and Brian Steffenson [3]

[1] Key Laboratory of Modern Precision Agriculture System Integration Research-Ministry of Education, China Agricultural University; Beijing 100083, China; qrc@cau.edu.cn (R.Q.); cauzm@cau.edu.cn (M.Z.)
[2] Department of Bioproducts and Biosystems Engineering, University of Minnesota; Saint Paul, MN 55108, USA; moghi005@umn.edu (A.M.);
[3] Department of Plant Pathology, University of Minnesota; Saint Paul, MN 55108, USA; bsteffen@umn.edu (B.S.);
* Correspondence: ceyang@umn.edu

Abstract: *Fusarium* head blight (FHB) disease is extensively distributed worldwide. This disease damages grain quality and reduces yield. The detection of this disease in a high throughput way is crucial to planters and breeders. Our study focused on developing a method for processing wheat color images and accurately detecting disease areas using deep learning and image processing techniques. The color images of wheat at the milky stage were collected and processed to construct datasets, which were used to retrain a deep convolutional neural network model using transfer learning. Testing results showed that the model can detect spikes, and the coefficient of determination of the number of spikes between the manual count and the detection was 0.80. The model was assessed, and the mean average precision for the testing dataset was 0.9201. On the basis of the results of spike detection, a new color feature was applied to obtain the gray image of each spike. Then, a modified region growing algorithm was implemented to segment and detect the diseased areas of each spike. Results show that the region growing algorithm performs better than K-means and Otsu's method in segmenting the FHB disease. Overall, this study demonstrates that deep learning techniques enable the accurate detection of FHB in wheat using color images, and the proposed method can effectively detect spikes and diseased areas, thereby improving the efficiency of FHB detection.

Keywords: *Fusarium* head blight disease; color imaging; deep neural network

1. Introduction

As a widely cultivated crop and staple food in the world, wheat is crucial to ensuring food security. Wheat is susceptible to numerous factors, such as ambient light and moisture, soil fertility, human management, and germ infection. In terms of diseases, *Fusarium* head blight (FHB) or scab from a *Fusarium* fungus is frequently observed in several humid areas; wheat spikes present characteristic symptoms once the fungus occupies them [1]. *Fusarium* can not only destroy the structure of wheat spikes but also generate toxins in wheat kernels, thereby decreasing grain production and quality. A major *Fusarium* toxic metabolite is deoxynivalenol, which can cause vomiting, dizziness, and nausea in humans when ingested through infected grain products [2, 3]. Thus, FHB should be monitored and assessed in time to reduce grain loss. Furthermore, breeding resistant varieties is deemed an effective method for solving FHB. The quantitative monitoring and assessment of FHB are vital during breeding. However, distinct wheat varieties have different resistances to *Fusarium* [4]; therefore, no universal method of detecting FHB is available. Currently, FHB can be manually monitored on the basis of its typical symptoms, but experience and prior knowledge are required. The monitoring is subject to human errors. Furthermore, attempts have been made to select specific gene sequences to detect *Fusarium*, which is time-consuming and

labor-intensive. Monitoring FHB in a high throughput manner is urgent and promising, and special attention is given to resolve the issues.

Fusarium can invade and destroy wheat tissues and change chlorophyll contents, water contents, and structure [5]; thus, the optical properties of infected spike regions are different from those of healthy spikes. Innovations in sensors and data processing methods offer potentials in detecting FHB, and techniques, such as hyperspectral and color imaging, have been investigated.

Hyperspectral imaging can provide considerable spectral and spatial information in an image format. A hyperspectral image contains numerous images that correspond to each wavelength. Every pixel of the hyperspectral image has full wavelengths [6]. This technique has been evaluated to detect crop diseases [7–9]. With respect to FHB, Ropelewska et al. [10] designed a hyperspectral imaging system to capture images of infected and healthy kernels and then used MaZda software to extract approximately 200 textures at selected wavelengths. These researchers analyzed these features to develop classification models to distinguish infected wheat kernels, and the accuracy reached 94%–100%. Similarly, Barbedo et al. [11] processed hyperspectral images of a wheat kernel and segmented the kernels from the background using four selected bands. After their tests, these authors selected a wavelength of 1411 nm to define a *Fusarium* index, which could indicate the likelihood of the kernel being infected by FHB. Whetton et al. [12] scanned a wheat canopy using a hyperspectral imager and obtained the spectrum information of wheat spikes. Then, these researchers used partial least squares regression (PLSR) analyses to establish models that can assess the percentage coverage of FHB. Bauriegel et al. [13] applied principal component analysis to identify four wavelength ranges (i.e., 500–533, 560–675, 682–733, and 927–931 nm) of diseased and healthy spikes and classified them successfully. The aforementioned studies were all conducted under laboratory conditions. The researchers also investigated the potential of applying hyperspectral imaging in detecting FHB in the fields. West et al. [14] reported that air sampling can be integrated with hyperspectral imaging to detect FHB by analyzing the spectrum differences at specific wavelengths. Whetton et al. [15] continued their work of processing hyperspectral images collected in four fields, with a spectral range of 400–1000 nm. PLSR was used to predict FHB and assess the diseased coverage, and the residual prediction's deviation value was 2.27. Furthermore, Jin et al. [16] designed a system for collecting hyperspectral images of wheat spikes in the field. A deep neural network (DNN) was used to classify diseased and healthy pixels, and the accuracy could reach 0.743. Most studies have selected several specific wavebands and have developed models or have proposed indexes using machine learning to detect FHB. In recent years, DNN has been extensively implemented in processing hyperspectral images and has exhibited excellent performance [17]. Although hyperspectral imaging can be useful in detecting FHB, several shortages that limit the application of hyperspectral imaging remain. First, the cost of hyperspectral imaging is high and thus difficult to promote. Second, some accessories are essential when collecting hyperspectral data, especially in the field. Finally, the spectral dataset of hyperspectral images is huge, and extracting significant and removing redundant information are the major challenges [18].

Alternatively, color imaging is inexpensive and easy to process and has a simple operation. Color imaging can capture the color and texture information of an object, and many researchers have implemented color imaging to detect or assess plant diseases [19–21]. For FHB detection, researchers have explored the usefulness of color imaging. Cambaza et al. [22] revealed that color imaging can be applied to monitor *Fusarium* on the color differences between healthy and infected spikes. The textural features from channels R (red), G (green), B (blue), and other color models were analyzed to develop classification models. Ruan et al. [23] used 55 color and texture features to develop a four-layer neural network to detect *Fusarium* in wheat kernels. Color imaging has also been investigated to recognize wheat spikes under field conditions, thereby providing the basis of identifying FHB. Bakhouche et al. [24] extracted seven features, namely, mean, variance, energy, entropy, contrast, skewness, and kurtosis, from wheat color images, and then used the K-means algorithm to detect spikes. Li et al. [25] selected laws of texture energy for texture analysis and trained a classification model using a neural network. Zhu et al. [26] chose the RGB, Lab, and HSV color spaces and applied all color components to train a support vector machine (SVM) model that

Peer-reviewed version available at Remote Sens. 2019, 11, 2658; doi:10.3390/rs11222658

can detect spikes. Zhou et al. [27] proposed an approach using a color feature (color coherence vectors), texture features (gray level co-occurrence matrix), and an image feature (edge histogram descriptor) to train a twin SVM model, thus successfully extracting each spike pixel from the background. In summary, specific features were extracted or calculated from color images and used as input parameters of classification models. The models were trained to detect FHB or spikes. In the process, suitable features and classification models were crucial. Most recently, Hasan et al. [28] tested a Faster RCNN model to detect spikes and output their bounding boxes. Pound et al. [29] trained a deep convolutional neural network (DCNN) model that can localize wheat spikes indoor. Generally, extracting features is not necessary, accurate models can be obtained using a sufficient dataset. Deep learning (DL) is promising in detecting wheat spikes and FHB, but its application requires further exploration [30, 31].

The detection of FHB under laboratory conditions must sample spikes and kernels, which will interrupt wheat growth and cannot realize long-term monitoring. Therefore, FHB are detected and assessed in a field is significant. Although hyperspectral imaging can detect FHB, its high cost is a major issue. Furthermore, to cultivate a new variety, breeders are eager to assess the severity of FHB, not merely detect it. As a prerequisite, every spike must be recognized, and their boundaries must be obtained. Previous studies have revealed the potential of color imaging for determining these issues. Moreover, DL techniques improve the efficiency of color imaging.

Therefore, our study focuses on the usefulness of color imaging, and the overall goal of this study is to apply DL techniques to extract spikes and detect FHB disease. This study aims to (1) train a DCNN model that can extract wheat spikes from one image without the influence of a complex background and (2) propose a novel approach to detecting diseased spikelets and FHB from each spike.

The remainder of this paper is organized as follows. Section 2 briefly describes the data collection, DCNN model training, and the methods of FHB detection. Section 3 provides the results of model training and FHB detection in detail. Section 4 discusses the analysis results and the shortcomings of this study, and proposes some improvements. Finally, the conclusion is provided in Section 5.

2. Materials and Methods

The pipeline of our system is depicted in Figure 1. The system involves four main steps, that is, data collection, dataset construction, DCNN model training, and FHB detection. Each step is described in detail in the following sections. A desktop with a 3.4 GHz AMD R5-2600 CPU, NVIDIA GTX 1070 (8 GB memory), and 8 GB RAM was utilized for offline processing.

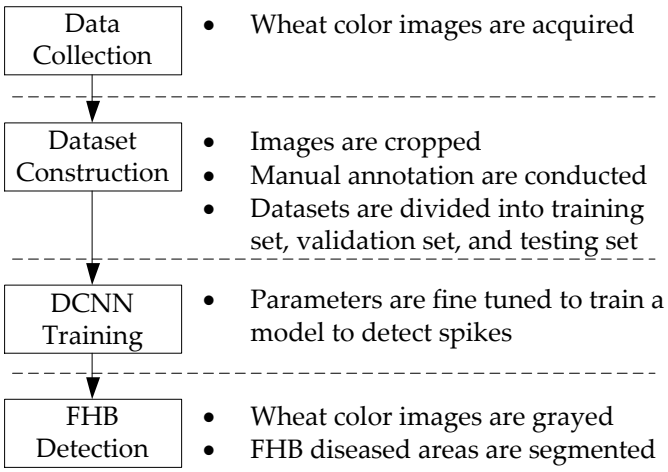


Figure 1. Workflow diagram of the proposed FHB detection.

2.1 Data collection

Five wheat lines were cultivated on Saint Paul campus at the University of Minnesota, United States, and each wheat plot is approximately 1.6 m × 0.7 m (length × width). The wheat lines are Wheaton, Roblin, Bac Up, MN00269, and Alsen, respectively. All wheat spikes were passively infected with *Fusarium* manually. The experiment was conducted between 11:00 and 14:00 on July 16, 2018, during which most of the wheat was at the milky stage. A phenotype platform, which is approximately 2.7 m × 2.5 m (length × width), was used to collect data. The platform was manually pulled and stopped to capture wheat images. A camera (Canon EOS Rebel T6i) with a large field of view was installed on the platform, as illustrated in Figure 2 (a) so that the whole wheat plot was covered in each image. In addition, a shelter was designed to protect the camera from sunlight (Figure 2 (b)). All images were saved in JPEG compressed format, and a computer was used to control the camera and collect wheat images.

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

Figure 2. Color image collection in the field: (**a**) Position of the color camera; (**b**) Platform used in this study.

2.2 Data augmentation

The captured images have a 6000 × 4000 pixel resolution, and only 20 images that correspond to 20 plots were obtained. These images were insufficient to train the DL model. To increase the sample size, two common steps were performed. First, the raw images were divided into small images, as demonstrated in Figure 3. The sub-images are 700 × 700 pixels in dimension, and 100 overlapping pixels existed between adjacent images. Then, additional datasets were generated by mirroring the sub-images horizontally and vertically. Subsequently, 2829 images were generated after the augmentation.

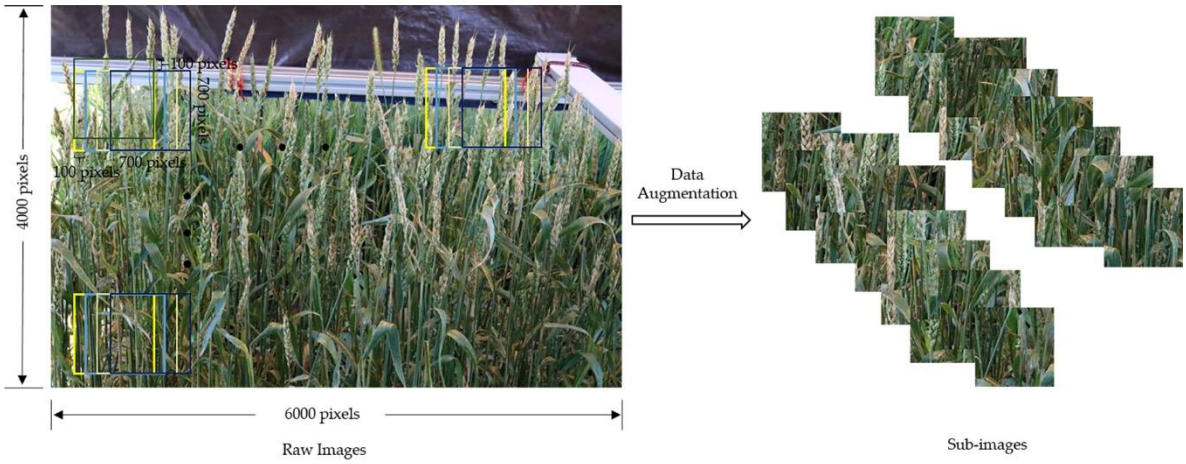


Figure 3. Sketch of data augmentation. Each rectangle box is a sub-image.

2.3 DCNN model

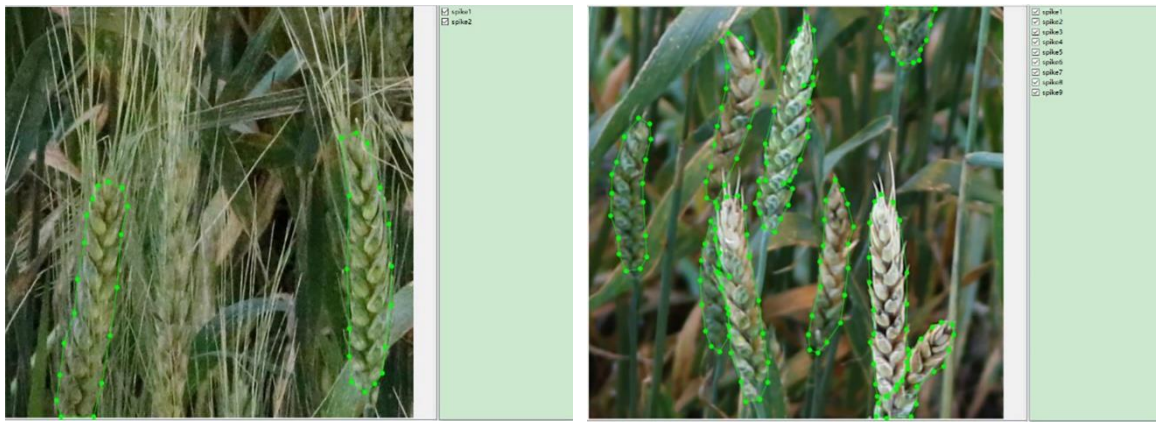
The task of this study is to detect diseased areas of spikes, and object detection techniques are the key to accomplishing this task. Primarily, detecting the boundaries of spikes is essential to remove the background influence. In 2017, based on the DL technique, He et al. [32] proposed the Mask RCNN model, which can not only classify objects but also output their outlines at the pixel level. It extends Faster RCNN model; combines classification, bounding boxes, and mask generation; and enables instance segmentation. Specifically, Mask RCNN has two main steps. First, it scans an image and generates numerous probable anchor regions using the region proposal network to extract features. Second, the regions are evaluated and classified, thereby generating bounding boxes and masks. In addition, the "RoIAlign" layer was proposed to improve the segmentation accuracy, and loss function (L) was defined as function 1.

$$L=L_{cls}+L_{box}+L_{mask} \quad (1)$$

where L_{cls}, L_{box}, and L_{mask} represent the classification, bounding box, and mask losses, respectively.

In the present study, Mask RCNN was implemented using Tensorflow API. The model was pre-trained using the COCO dataset, and only the last few layers were trained and fine-tuned using transfer learning. ResNet 50 and ResNet 101 are two backbone network options. The former was used here given its accuracy and small training dataset. Batch normalization was applied to prevent overfitting. The batch size was set to 2 because it requires a small memory. Five scales (i.e., 32, 64, 128, 256, and 512) and three ratios (i.e., 0.5, 1, and 2) were set for the anchors. If the values of intersection over union ratio with the ground-truth bounding boxes were higher than 0.5, the anchors were considered available. The other parameters were in the default setting.

Before training the model, a training dataset and a validation dataset must be annotated. An open-source software, namely, Lableme [33], was used to label the datasets. The spikes were selected and labeled "spike" class by drawing polygons manually, as exhibited in Figure 4. Furthermore, the software generated JSON files that record the boundaries of the spikes.



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

Figure 4. Annotation of spike images: (**a**) Annotated result of an awned spike image; (**b**) Annotated result of an awnless spike image.

The model was retrained on a desktop with CUDA 9.0. A dataset containing 1959 images was obtained. Among these images, 989 and 970 are awned and awnless wheat images, correspondingly. A total of 420 wheat images were randomly selected to generate the validation dataset, with 220 and 200 awned and awnless wheat images, respectively. The remaining images were used as the testing dataset. Every epoch contains 100 iterations, and the maximum number of training epochs was set to 770 with a momentum of 0.9. In addition, the weight decay was fixed to 0.01. Learning rate (LR) was initially set to 0.002. After 161, 332, and 508 epochs, the LR was decreased to 0.001, 0.0005, and 0.0002, separately.

Peer-reviewed version available at Remote Sens. 2019, 11, 2658; doi:10.3390/rs11222658

After training, the model was evaluated using parameters, namely, recall (r) and precision (p). These parameters were calculated using the following functions:

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

$$p = \frac{TP}{TP + FP} \tag{3}$$

where TP is the predicted mask pixels that correspond to the labeled mask pixels, FN represents the labeled mask pixels that are inexistent in the predicted mask pixels, and FP is the predicted mask pixels that are inexistent in the labeled mask pixels. Normally, a high TP, a low FN, and a low FP are expected to obtain improved performance. However, an individual parameter could not quantify the model adequately. Therefore, the average precision (AP) that is the area under the precision–recall curve was applied as an indicator to estimate the model.

Furthermore, several statistics parameters were used to assess the spike detection performance. These parameters are root mean squared error ($RMSE$), relative RMSE ($rRMSE$), and the coefficient of determination (R^2), as described in the following functions.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{n}(t_i - d_i)^2} \tag{4}$$

$$rRMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{n}\left(\frac{t_i - d_i}{t_i}\right)^2} \tag{5}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(t_i - d_i)^2}{\sum_{i=1}^{n}(t_i - \bar{t}_i)^2} \tag{6}$$

where t_i is the manually counted number of spikes, d_i is the automatically counted number of spikes, and $\bar{t}_i$ is the mean value of t_i.

2.4 FHB detection

In contrast to healthy spikes, the colors of spikes are white or pink if they are infected with *Fusarium*. However, in the entire image, the symptom areas are small and scattered. Consequently, the diseased areas are difficult to detect directly. The DCNN can segment the spikes from the background and output their boundaries, thus aiding the detection of diseased areas. Nonetheless, a spike consists of spikelets, rachis, rachilla, and spaces are in-between spikelets. Therefore, further steps must be performed to process the segmented spikes and extract the diseased areas accurately.

Spikelets, rachis, and rachillas have different colors. Thus, spikelets could be extracted using color features. In addition, healthy spikelets are generally green, and diseased spikelets are white or pink. In accordance with tests [34], a new color feature (GB) was proposed using the green and blue components to highlight the diseased areas of each spike; this new color feature is calculated using function 7.

$$GB = \alpha \times B - G \tag{7}$$

where α is a coefficient, B is the blue component value, and G is the green component value.

The GB components of the segmented spikes in the color images were calculated to obtain their gray images. Although the diseased areas of a spike are scattered, each area shows a similar appearance. In accordance with the characteristic, every pixel could be classified on the basis of the correlations between such a pixel and adjacent pixels. Region growing is an effective segmentation algorithm and widely used in machine learning. In general, a pixel is initially selected as the seed point, and then its surrounding pixels are compared and evaluated with it. If their similarities are consistent with the set rules, then the pixels are classified into the same class. The selection of the

Peer-reviewed version available at Remote Sens. 2019, 11, 2658; doi:10.3390/rs11222658

initial seed point and setting of the growing rules are crucial to the region growing algorithm. However, in the present study, a single initial seed point setting is insufficient to recognize all unconcentrated diseased areas. Therefore, a modified region growing algorithm is applied to process the gray image [35]. The flowchart of this algorithm is displayed in Figure 5. First, the pixels of the gray spike images were scanned to find an unmarked pixel (x_0, y_0), which is tagged as the Kth class and saved into the "stack." Second, a pixel was selected from the "stack" as the seed point, and the pixels around the seed point (eight neighbors) were processed. When the gray difference between the seed point and the processing pixel was in the range, the pixel was regarded as the same class and saved into the "stack." Third, the second operation was repeated until the "stack" was empty. Finally, Steps 1–3 were repeated until every pixel was marked. Many small areas are inevitably generated because of noises, especially at the junctions of spikelets, rachis, and rachillas. Thus, every region's area was calculated, and the diseased potential regions whose areas were smaller than the threshold value set were removed. At this point, the region growing algorithm was implemented completely.

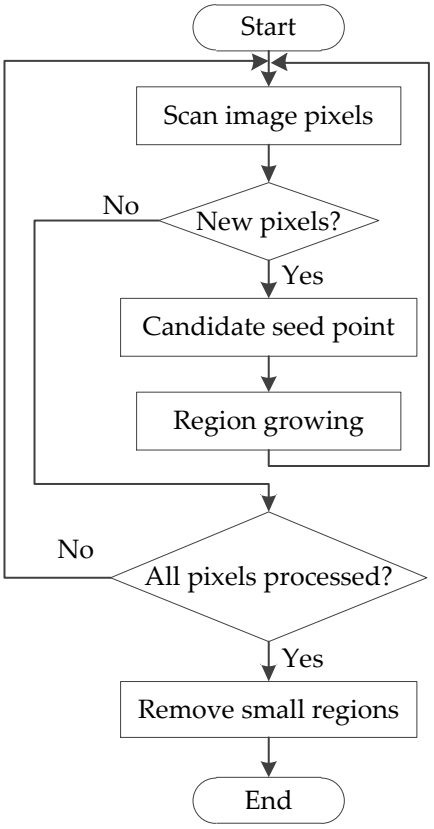


Figure 5. Flowchart of the region growing algorithm.

Subsequently, the close operation was performed on the results. Breeders typically classify the FHB severities into four classes [36], as listed in Table 1. Finally, the ratio of FHB diseased areas to the entire spike area is calculated in accordance with the standard [37]. The approach was implemented using MATLAB r2016a software.

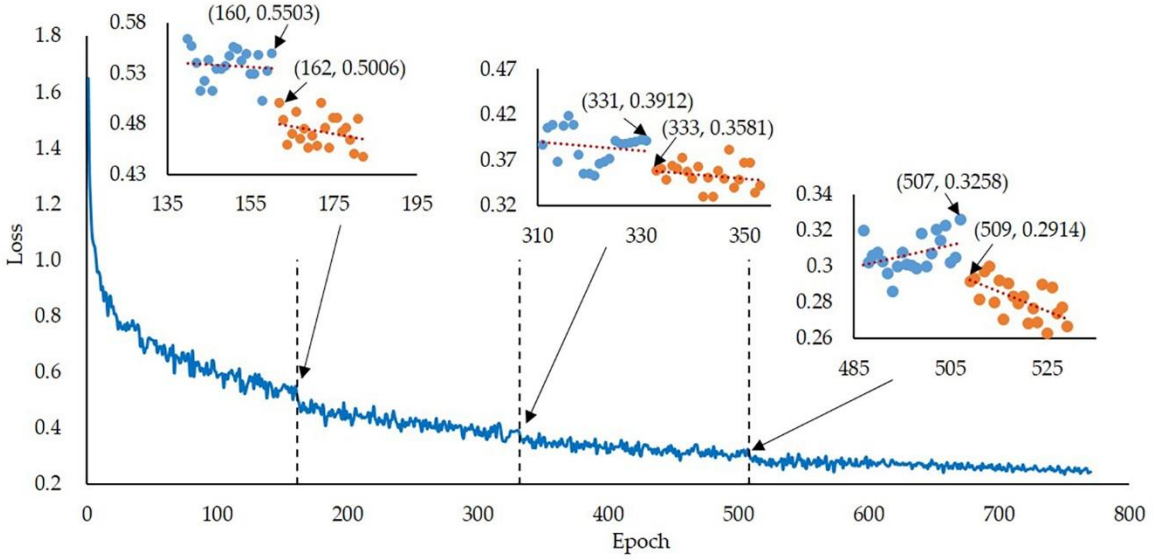
Table 1. Standard for assessing FHB severity.

Severity	Percentage of FHB diseased areas
I class	< 25%
II class	25%–50%
III class	50%–75%
IV class	> 75%

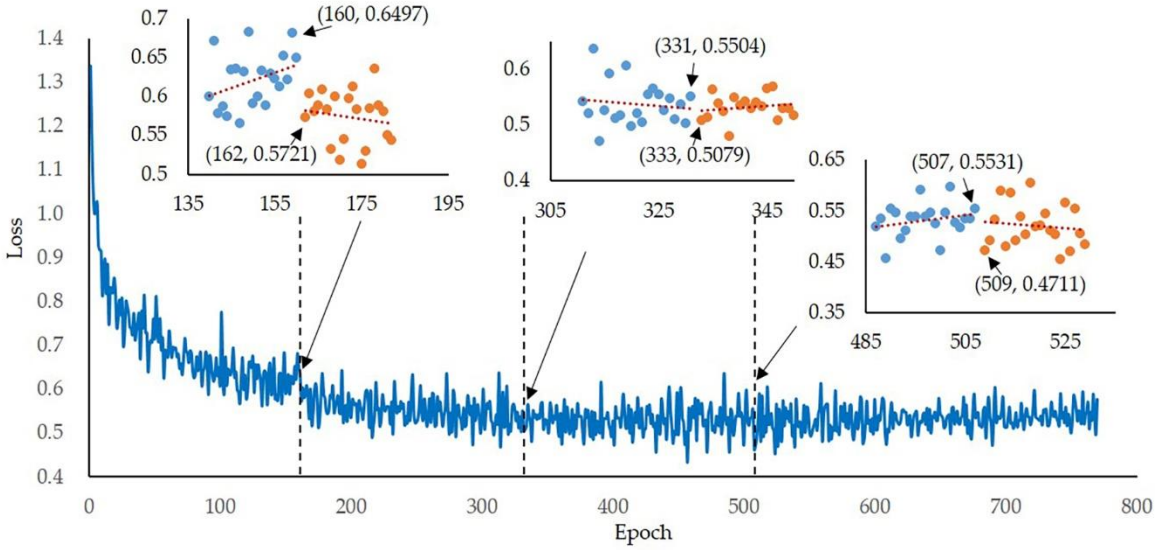
3. Results

3.1. Model training for spike detection

The training and validation datasets for spike detection were used to retrain the Mask RCNN model and find the hyperparameters of the DCNN. Losses were calculated after each epoch and visualized using Tensorboard, which helped monitor the training process and adapted the hyperparameters. After 770 epochs of training, the losses of the model for the training and validation datasets are summarized in Figure 6. Clearly, the loss of the initial epoch would become smaller than the last epoch when the LR is reset at epochs 161, 332, and 508, especially for training loss. This result indicated that decreasing the loss by lowering the LR during the training process is effective. Figure 6 displays that the losses change very slowly after 508 epochs because the LR becomes 0.0002 at this point. To ensure a convincing result, additional training epochs were performed to monitor the variation in losses and prevent the model from overfitting. The final losses (the average loss of the last five epochs) of the model of the training and validation datasets were approximately 0.2479 and 0.5456, respectively.



(**a**)



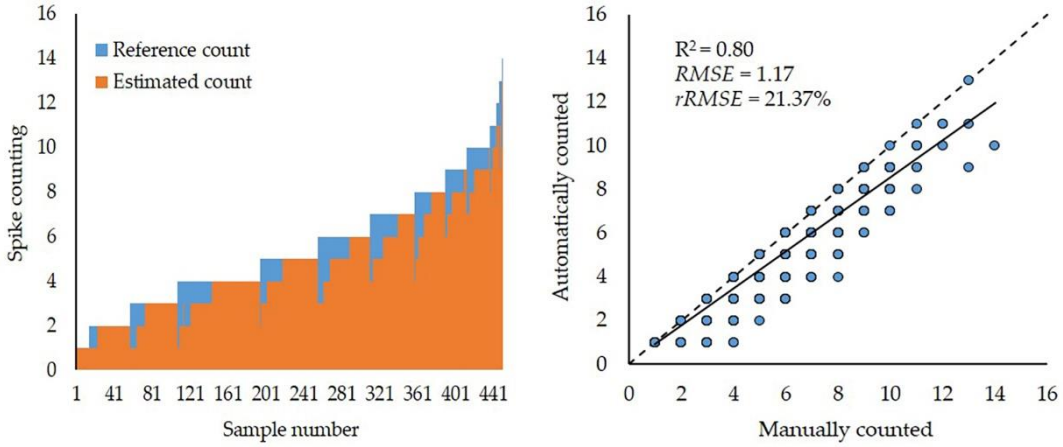
(**b**)

Figure 6. Losses of the Mask RCNN for training and validation: (**a**) Training loss; (**b**) Validation loss.

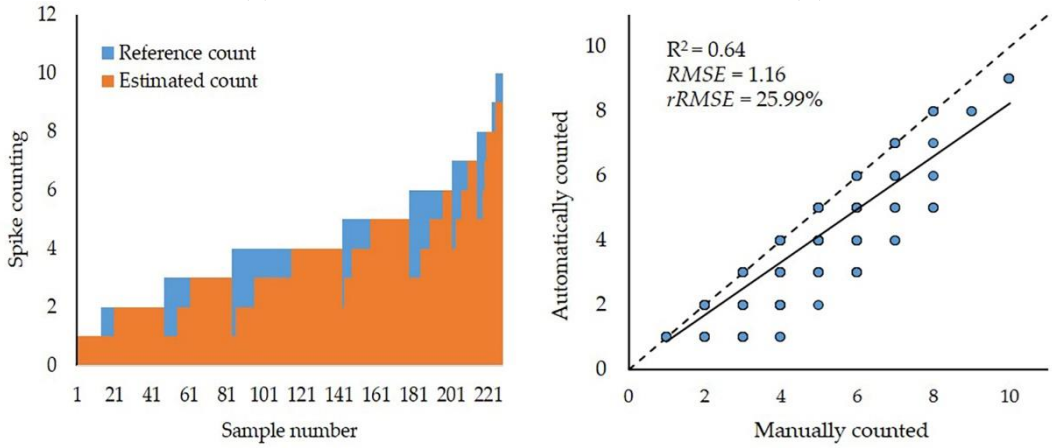
3.2. Model testing for spike detection

Mask RCNN model was retrained to accomplish spike detection. To evaluate its performance, the testing dataset, which contains 450 images, including 228 awned and 222 awnless wheat samples, was used to assess the model. In this study, the APs of the testing dataset were calculated, and the mean AP (mAP) was 0.9201, thereby indicating that the retrained model can detect wheat spikes with different shapes and features.

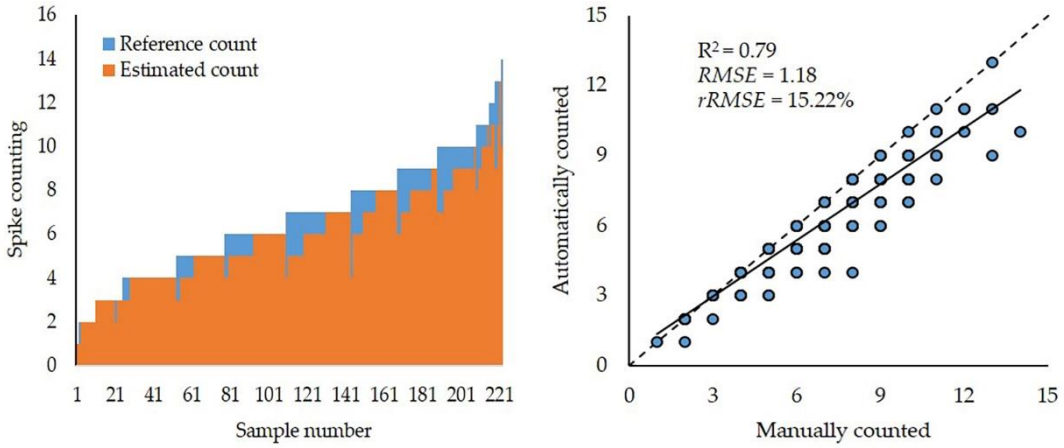
The numbers of spikes in every image were counted using the retrained model and compared with the true number obtained manually. The testing dataset was utilized, and the statistics are plotted in Figure 7. The R^2 between the automatically and manually counted spikes for the testing dataset (450 samples), the awned wheat spikes (228 samples), and the awnless wheat spikes (222 samples) were 0.80, 0.64, and 0.79, correspondingly; the *RMSEs* were 1.17, 1.16, and 1.18, respectively, and the *rRMSEs* were 21.37%, 25.99%, and 15.22%, separately.



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



(**c**) (**d**)



(**e**) (**f**)

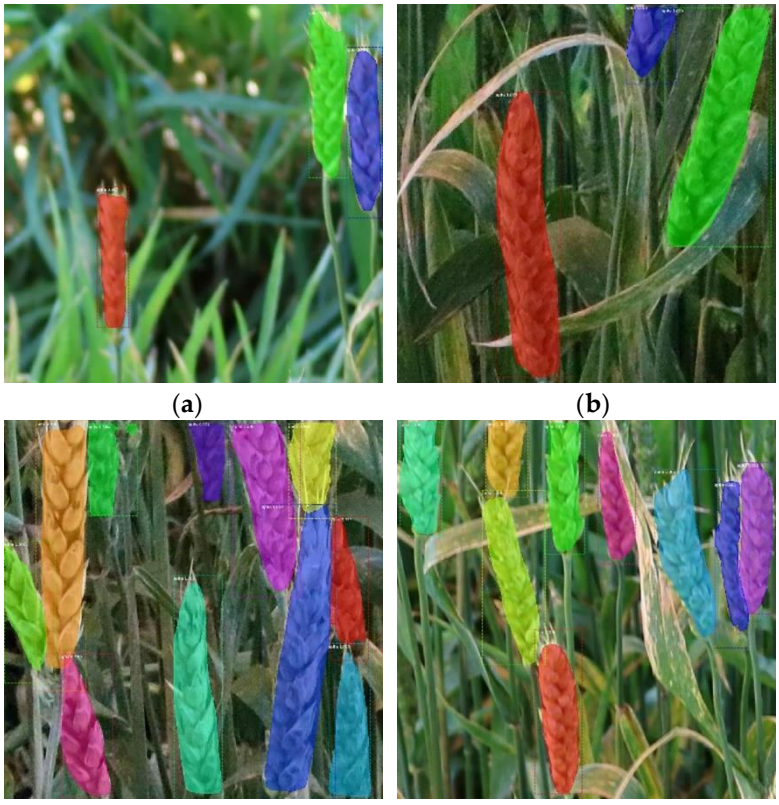
Figure 7. Comparison between the automatically and manually counted spike numbers: (**a**, **b**) Comparison and correlation of automatically and manually counted numbers for all testing spikes; (**c, d**) Comparison and correlation of automatically and manually counted numbers for the awned spikes; (**e**, **f**) Comparison and correlation of automatically and manually counted numbers for the awnless spikes. The dotted lines of (**b**, **d**, **f**) are 1:1 lines, and the straight lines are fitted lines.

Majority of the spikes could be detected as either healthy or diseased spikes. In Figures 8 and 9, the spike density in one image has no influence on the detection and is adaptable to awned wheat and awnless wheat varieties. Nonetheless, among the 450 testing samples, the spikes of approximately 218 samples were not entirely detected. The model's failure to detect some specific spikes is inevitable. Primarily, many spikes were blurred or taken as the background because the camera could merely focus on a limited space, as illustrated in Figures 10 (a) and (b). Furthermore, many spikes were severely infected by FHB or other diseases. Consequently, the spikelets, rachis, and rachillas were severely destroyed, their feathers are not evident. Several samples are depicted in Figures 10 (c) and (d). Moreover, only a small area of a spike is visible in the image. Thus, extracting features and classifying pixels, such as the yellow rectangle portions displayed in Figures 10 (e) and (f), were difficult for the model. Finally, some spikes were adjacent or intersecting, as demonstrated in Figures 10 (g) and (h), thereby causing all or parts of them being missed.

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

(c) (d)

Figure 8. Detection results of awned spikes in different densities.

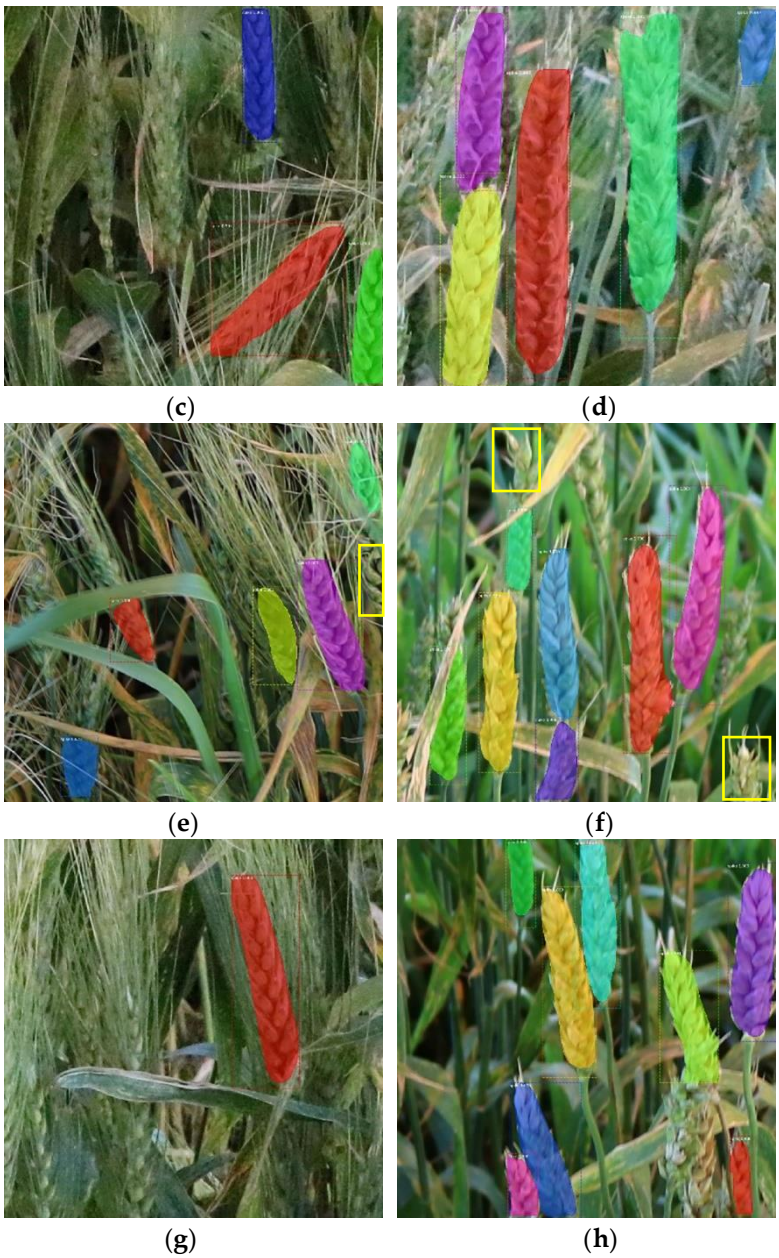


(a) (b)

(c) (d)

Figure 9. Detection results of awnless spikes in different densities.

(a) (b)



(**c**) (**d**)

(**e**) (**f**)

(**g**) (**h**)

Figure 10. Examples of several spikes that failed to be detected: (**a**, **b**) Spikes are taken as the background; (**c**, **d**) Spikes are severely destroyed; (**e**, **f**) Spikes are not entire; (**g**, **h**) Spikes are adjacent to each other.

3.3. FHB detection

The Mask RCNN model can detect spikes and output their mask pixels, thereby enabling us to detect the diseased regions of spikes. The new color feature GB was applied to each spike, and the gray values of the diseased spikelets, healthy spikelets, and others are different. The gray values are high if the pixels belong to the diseased spikelets; otherwise, the values are low. In addition, α was given a set of values, and the corresponding gray images were evaluated to acquire the optimum performance. Tests show that when α was set to 1.4, the diseased areas were much brighter than the healthy areas in the gray images and more likely to be detected. In Figure 11, the images of several wheat varieties were processed using GB to verify their robustness. The results indicated that the GB feature is suitable for detecting the FHB areas of multiple wheat varieties.

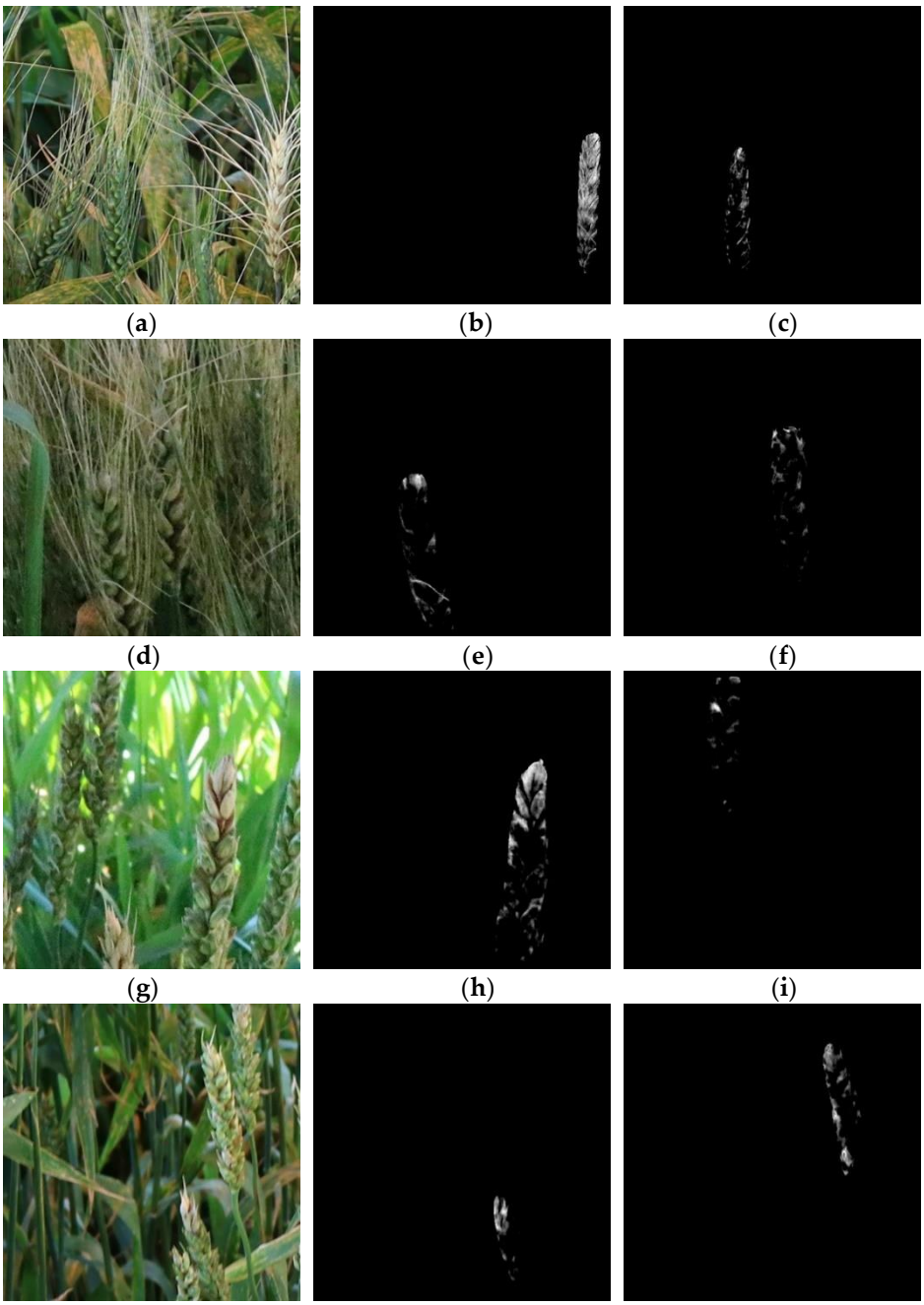


Figure 11. Examples of healthy and diseased spikes with different varieties and shapes in the gray images.

The gray differences between diseased and healthy pixels enable the detection of the FHB areas of each spike using conventional segmentation algorithms of machine learning. Here, the modified region growing algorithm was used to segment the diseased areas from the entire spike. The K-means and OTSU (the maximum between-class variance) algorithms were also implemented to process the gray spike images used for the comparative tests. After segmentation, small potential areas were removed, and close operation was applied to eliminate noises. The final results are exhibited in Figures 12, 13, and 14. In the present study, K was assigned with 4 considering the spatial resolution of the images. Thus, the pixels in the gray images were categorized into four classes, namely, the background, the healthy, the diseased, and the connected areas between the spikelets and the others. The segmentation results were unfavorable because the connected areas were typically classified by mistake (Figure 13). In Figure 14, the segmentation accuracy of OTSU is the lowest, and some non-diseased areas were improperly considered diseased areas. The region

growing algorithm can handle the abovementioned problems and demonstrated the optimum performance in segmenting the diseased areas. Although the masks may contain some non-spike pixels, the region growing algorithm can segment diseased areas accurately. Furthermore, the diseased areas were divided into several small regions, rather than merely a group, as displayed in Figures 12 (c), (f), (i), and (l) in different colors.

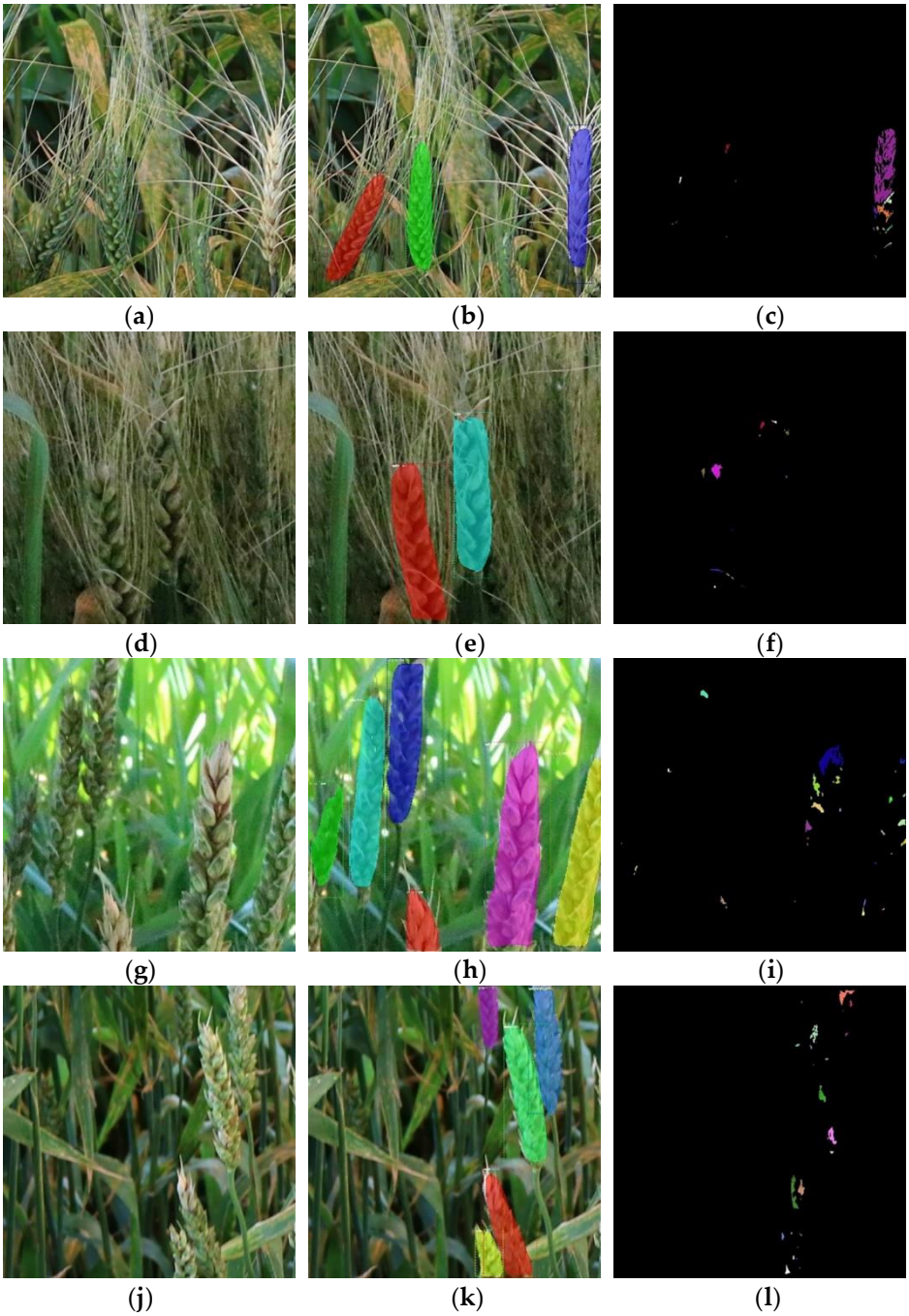
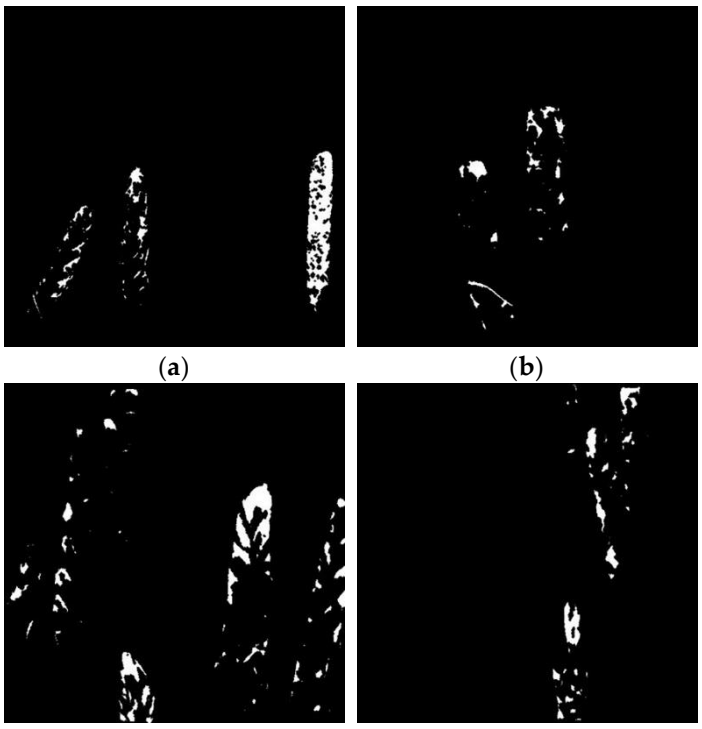


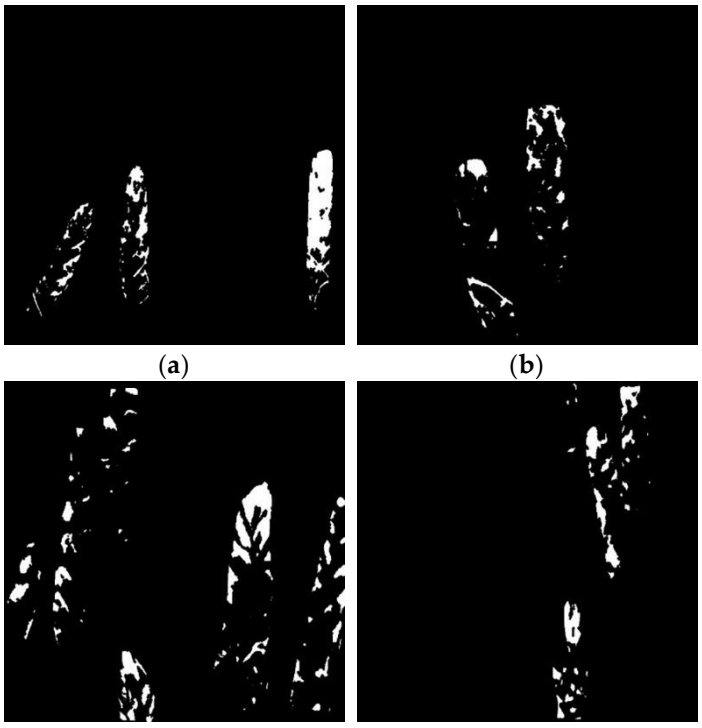
Figure 12. FHB detection results of different wheat varieties: (**a**, **d**, **g**, **j**) Color images of awned and awnless wheat varieties; (**b**, **e**, **h**, **k**) Spike detection results of awned and awnless wheat varieties; (**c**, **f**, **i**, **l**) FHB detection results of awned and awnless wheat varieties.

Peer-reviewed version available at Remote Sens. 2019, 11, 2658; doi:10.3390/rs11222658



(a) (b) (c) (d)

Figure 13. Segmentation results of the K-means algorithm.



(a) (b) (c) (d)

Figure 14. Segmentation results of the OTSU algorithm.

The testing dataset (450 samples) was processed through the method developed in this study to detect FHB. A total of 2127 spikes were successfully detected and extracted using the retrained Mask RCNN model. Furthermore, the ratios of FHB pixels to overall spike pixels were calculated to evaluate the FHB severities, as presented in Figure 15. For most wheat spikes, the ratios of the infected FHB areas to the entire spike are lower than 25%, thereby indicating that the disease severities of the spikes were I class.

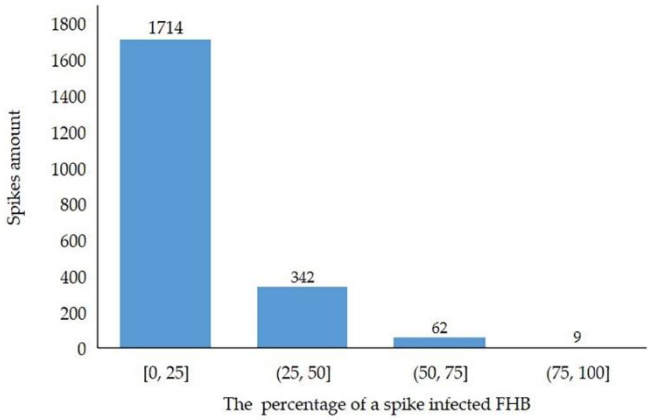


Figure 15. Statistical result of infected spikes.

4. Discussion

4.1. Analysis of Mask RCNN model performance

A Mash RCNN model was retrained and used to recognize the spikes in color images and generate masks for every spike, which is the basis for FHB detection. Therefore, the performance of the Mask RCNN model is essential. In this study, recall, precision, and mAP were implemented to assess the model. According to Section 3.2, the retrained model demonstrates favorable performance in spike detection (mAP is 0.9201). However, some questions should be addressed and analyzed further. First, the spikes represent various postures in the color images. Some neighbored spikes were overlapping or adhering, thereby making the detection of each spike in high precision difficult. This difficulty is also the shortcoming of Mask RCNN model. Second, the color images contain not only entire spikes but also many partial spikes. The tests showed that the retrained model can detect some incomplete spikes in the middle of the color images. However, recognizing the spikes located in image edges or with a deep depth of field is insufficient. Finally, the setting of anchor parameters will affect the precision of spike detection. The spike sizes in color images need to be considered, and the anchor parameters should be consistent with the spike sizes. Otherwise, small spikes will be missed, and only a portion of large spikes can be detected. In addition, the tests show that wheat awn slightly influences spike detection provided that the spikes are not completely covered.

4.2. Analysis of FHB detection

A new color feature, GB, was proposed in this study. The results showed that the color feature adequately uses the differences between healthy and diseased spikes to recognize and segment the FHB diseased areas of spikes. Moreover, some factors that may negatively affect FHB detection remain. Although FHB was effectively detected, FHB diseased areas are detected and classified on the basis of their severity is considerable. Specially, the representations of FHB diseased areas at the late stage are different from those at the early stage. At the late stage of FHB disease, *Fusarium* has invaded the inside of spikelets and destroyed their tissues and structures. Some diseased areas of different spikelets have merged into a large one. Therefore, rachis and rachillas may disappear at times, and the boundaries between healthy and diseased spikelets are unclear, thus possibly resulting in misclassification. At the early stage of FHB disease, many diseased areas have been overly small to be preserved when the close operation was conducted. Particularly, the diseased areas located on the edge of the spikelets are likely to be ignored. Furthermore, the effect of wheat awns on detecting diseased areas should be considered. This effect probably be due to the colors of wheat awn are similar to the diseased spikes after being dried.

4.3. Next Steps

The implementation of a DL technique increases the feasibility of FHB detection in wheat under field conditions. Further work needs to be conducted in the future to improve the precision of FHB detection. First, images at only one growth stage were processed in this study. Additional wheat images in different shapes, varieties, growth stages, and views should be captured to enlarge the training dataset and improve the robustness and applicability of the trained model. Second, a new thought can be verified to train a model that can detect FHB in wheat directly. In this case, FHB diseased samples must be labeled. In addition, the model can classify FHB based on their severities. However, a new criterion should be proposed to assess the disease resistance of wheat considering the lack of spike masks. Eventually, based on the results of this study, special attention can be given to process the color images collected in the natural ambient illumination environment.

5. Conclusions

In this study, the potential of FHB detection using color images was explored. To detect spikes in images, numerous annotated images were divided into training and validation datasets to retrain a Mask RCNN model through the transfer learning technique. Only the last few layers were trained, and the tests showed that the model can detect most spikes of several varieties in different shapes, R^2, *RMSE*, and *rRMSE* of the automatic detection and manual counts were 0.80, 1.17, and 21.37%, respectively. Moreover, the model can output the pixels of every spike, including healthy and diseased areas. The testing dataset used to assess the model and the mAP is 0.9201. A novel color feature, GB, was proposed to highlight the diseased portions of the spikes in the gray images. The modified region growing algorithm exhibited a favorable performance in detecting the FHB diseased areas of each spike. A total of 450 images were tested, and the results indicated the feasibility of FHB detection in wheat using color images. The method proposed in this study can effectively extract spikes and detect FHB, thereby decreasing the cost and improving the efficiency of FHB detection. In future work, a broad range of datasets should be developed and applied to our method to improve the accuracy of FHB detection further.

Author Contributions: Data collection, R.Q. and A.M.; Methodology, R.Q.; Funding acquisition, C.Y., B.S. and M.Z.; Supervision, C.Y.; Writing-original draft, R.Q.; Writing-review & editing, M.Z. and C.Y.

Funding: This material is based upon work supported by the USDA-ARS Funding No. 58-5062-8-018. This is a cooperative project with the U.S. Wheat & Barley Scab Initiative (USWBSI). Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the authors and do not necessarily reflect the view of the U.S. Department of Agriculture. The project is also partially funded by the National Natural Science Foundation of China (Grant No. 31971786) and the Industrial Partnership Grants from MnDRIVE Robotics, Sensors and Advanced Manufacturing Initiative.

Acknowledgments: The authors are grateful to Youbing Wang from the University of Minnesota for his instructions in deep learning. We also appreciate the financial support from the China Scholarship Council.

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

References

1. Leplat, J.; Mangin, P.; Falchetto, L.; Heraud, C.; Gautheron, E; Steinberg, C. Visual assessment and computer–assisted image analysis of *Fusarium* head blight in the field to predict mycotoxin accumulation in wheat grains. *European Journal of Plant Pathology* **2018**, 150, 1065-1081.
2. Maloney, P.V.; Petersen, S.; Navarro, R.A.; Marshall, D.; Mckendry, A.L.; Costa, J.M.; Murphy, J.P. Digital image analysis method for estimation of *Fusarium*-damaged kernels in wheat. *Crop Science* **2014**, 54, 2077-2083.
3. Bai, G.; Shaner, G. Management and resistance in wheat and barley to *Fusarium* head blight. *Annual Review of Phytopathology* **2004**, 42, 135-161.

4. Popovski, S.; Kos, K.; Jakovac Strajn, B.; Celar, F.A. *Fusarium* spp. incidence and DON contamination in different wheat varieties correlated with the environmental factors. *Cereal Research Communications* **2017**, 45, 114-123.
5. Kang, Z.; Buchenauer, H. Cytology and ultrastructure of the infection of wheat spikes by *Fusarium* culmorum. *Mycological Research* **2000**, 104, 1083-1093.
6. Xie, C.; Shao, Y.; Li, X.; He Y. Detection of early blight and late blight diseases on tomato leaves using hyperspectral imaging. *Scientific Reports* **2015**, 5, 16564.
7. Thomas, S.; Kuska, M. T.; Bohnenkamp, D.; Brugger, A.; Alisaac, E.; Wahabzada, M.; Behmann, J.; Mahlein, A. Benefits of hyperspectral imaging for plant disease detection and plant protection: a technical perspective. *Journal of Plant Diseases and Protection* **2018**, 125, 5-20.
8. Thomas, S.; Behmann, J.; Steier, A.; Kraska, T.; Muller, O.; Rascher, U.; Mahlein, A. Quantitative assessment of disease severity and rating of barley cultivars based on hyperspectral imaging in a non-invasive, automated phenotyping platform. *Plant Methods* **2018**, 14, 45.
9. Yao, Z.; Lei, Y.; He, D. Early visual detection of wheat stripe rust using visible/near-infrared hyperspectral imaging. *Sensors* **2019**, 19, 952.
10. Ropelewska, E.; Zapotoczny, P. Classification of *Fusarium*-infected and healthy wheat kernels based on features from hyperspectral images and flatbed scanner images: a comparative analysis. *European Food Research and Technology* **2018**, 244, 1453-1462.
11. Barbedo, J.G.A.; Tibola, C.S.; Fernandes, J.M.C. Detecting *Fusarium* head blight in wheat kernels using hyperspectral imaging. *Biosystems Engineering* **2015**, 131, 65-76.
12. Whetton, R.L.; Hassall, K.L.; Waine, T.W.; Mouazen, A.M. Hyperspectral measurements of yellow rust and *Fusarium* head blight in cereal crops: part 1: laboratory study. *Biosystems Engineering* **2018**, 166, 101-115.
13. Bauriegel, E.; Giebel, A.; Geyer, M.; Schmidt, U.; Herppich, W.B. Early detection of *Fusarium* infection in wheat using hyper-spectral imaging. *Computers and Electronics in Agriculture* **2011**, 75, 304-312.
14. West, J.S.; Canning, G.G.M.; Perryman, S.A.; King, K. Novel technologies for the detection of *Fusarium* head blight disease and airborne inoculum. *Tropical Plant Pathology* **2017**, 42, 203-209.
15. Whetton, R.L.; Waine, T.W.; Mouazen, A.M. Hyperspectral measurements of yellow rust and *Fusarium* head blight in cereal crops: part 2: on-line field measurement. *Biosystems Engineering* **2018**, 167, 144-158.
16. Jin, X.; Jie, L.; Wang, S.; Qi, H.; Li, S. Classifying wheat hyperspectral pixels of healthy heads and *Fusarium* head blight disease using a deep neural network in the wild field. *Remote Sensing* **2018**, 10, 395.
17. Signoroni, A.; Savardi, M.; Baronio, A.; Benini, S. Deep learning meets hyperspectral image analysis: a multidisciplinary review. *Journal of Imaging* **2019**, 5, 52.
18. Qiu, R.; Wei, S.; Zhang, M.; Li, H.; Sun, H.; Liu, G.; Li, M. Sensors for measuring plant phenotyping: a review. *International Journal of Agricultural and Biological Engineering* **2018**, 11, 1-17.
19. Sun, Y.; Jiang, Z.; Zhang, L.; Dong, W.; Rao, Y. SLIC_SVM based leaf diseases saliency map extraction of tea plant. *Computers and Electronics in Agriculture* **2019**, 157, 102-109.
20. Zhu, J.; Wu, A.; Wang, X.; Zhang, H. Identification of grape diseases using image analysis and BP neural networks. *Multimedia Tools and Applications* 2019. Available online: https://link.springer.com/content/pdf/10.1007%2Fs11042-018-7092-0.pdf (accessed on 10 September 2019).
21. Salazar, I.F.; Huaman, S.G.; Kemper, G.; Telles, J.; Diaz, D. An algorithm for plant disease visual symptom detection in digital images based on superpixels. *International Journal on Advanced Science Engineering Information Technology* **2019**, 9, 194-203.

22. Cambaza, E.; Koseki, S.; Kawamura, S. Why RGB imaging should be used to analyze *Fusarium* graminearum growth and estimate deoxynivalenol contamination. *Methods and Protocols* **2019**, 2, 25.
23. Ruan, R.; Ning, S.; Song, A.; Ning, A.; Jones, R.; Chen, P. Estimation of *Fusarium* scab in wheat using machine vision and a neural network. *Cereal Chemistry Journal* **1998**, 75, 455-459.
24. Bakhouche, M.; Cointault, F.; Gouton, P. Texture analysis with statistical methods for wheat ear extraction. Proc. SPIE 6356, Eighth International Conference on Quality Control by Artificial Vision, Le Creusot, France, 29 May 2007.
25. Li, Q.; Cai, J.; Berger, B.; Okamoto, M.; Miklavcic, S.J. Detecting spikes of wheat plants using neural networks with laws texture energy. *Plant Methods* **2017**, 13, 83.
26. Zhu, Y.; Cao, Z.; Lu, H.; Li, Y.; Xiao, Y. In-field automatic observation of wheat heading stage using computer vision. *Biosystems Engineering* **2016**, 143, 28-41.
27. Zhou, C.; Liang, D.; Yang, X.; Yang, H.; Yue, J.; Yang, G. Wheat ears counting in field conditions based on multi-feature optimization and TWSVM. *Frontiers in Plant Science* **2018**, 9, 1024.
28. Hasan, M.M.; Chopin, J.P.; Laga, H.; Miklavcic, S.J. Detection and analysis of wheat spikes using convolutional neural networks. *Plant Methods* **2018**, 14, 100.
29. Pound, M.P.; Atkinson, J.A.; Wells, D.M.; Pridmore, T.P.; French, A.P. Deep learning for multi-task plant phenotyping. *bioRxiv* **2017**, 2055-2063.
30. Madec, S.; Jin, X.; Lu, H.; De Solan, B.; Liu, S.; Duyme, F.; Heritier, E.; Baret, F. Ear density estimation from high resolution RGB imagery using deep learning technique. *Agricultural and Forest Meteorology* **2019**, 264, 225-234.
31. Mohanty, S.P.; Hughes, D.P.; Salathé, M. Using deep learning for image-based plant disease detection. *Frontiers in Plant Science* **2016**, 7, 1419.
32. He, K.; Gkioxari, G.; Dollar, P.; Girshick, R. Mask R-CNN. *arXiv* **2017**, arXiv: 170306870
33. Russell, B.C.; Torralba, A.; Murphy, K.P.; Freeman, W.T. LabelMe: a database and web-based tool for image annotation. *International Journal of Computer Vision* **2008**, 77, 157-173.
34. Qiu, R.; Yang, C.; Moghimi, A.; Anderson, J.; Steffenson, B.; Marchetto, P. Detection of *Fusarium* head blight in small grains using hyperspectral imaging. National Fusarium Head Blight Forum, St. Louis, Missouri, USA, December 2018; pp. 32-37.
35. Github. Available online: https://github.com/ZQPei/Seeded-Region-Growing-Algorithm (accessed on 10 September 2019).
36. Liu, S.; Hall, M.D.; Griffey, C.A.; Mckendry, A.L. Meta-analysis of QTL associated with *Fusarium* head blight resistance in wheat. *Crop Science* **2009**, 49, 1955-1968.
37. Stack, R.W.; McMullen, M. P. A visual scale to estimate severity of *Fusarium* head blight in wheat. North Dakota State University Extension Service **1995**, p 1095.