

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, Brian J. Steffenson ³ and Cory D. Hirsch 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

³ Department of Plant Pathology, University of Minnesota, Saint Paul, MN 55108, USA; bsteffen@umn.edu (B.J.S.); cdhirsch@umn.edu (C.D.H.)

* Correspondence: ceyang@umn.edu

Received: 11 September 2019; Accepted: 7 November 2019; Published: 13 November 2019



Abstract: Fusarium head blight (FHB) is a devastating disease of wheat worldwide. In addition to reducing the yield of the crop, the causal pathogens also produce mycotoxins that can contaminate the grain. The development of resistant wheat varieties is one of the best ways to reduce the impact of FHB. To develop such varieties, breeders must expose germplasm lines to the pathogen in the field and assess the disease reaction. Phenotyping breeding materials for resistance to FHB is time-consuming, labor-intensive, and expensive when using conventional protocols. To develop a reliable and cost-effective high throughput phenotyping system for assessing FHB in the field, we focused on developing a method for processing color images of wheat spikes to accurately detect diseased areas using deep learning and image processing techniques. Color images of wheat spikes at the milk stage were collected in a shadow condition 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 detected spikes very accurately in the images since the coefficient of determination for the number of spikes tallied by manual count and the model 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 for spike detection, a new color feature was applied to obtain the gray image of each spike and a modified region-growing algorithm was implemented to segment and detect the diseased areas of each spike. Results showed that the region growing algorithm performed better than the K-means and Otsu's method in segmenting diseased areas. We demonstrated that deep learning techniques enable accurate detection of FHB in wheat based on color image analysis, and the proposed method can effectively detect spikes and diseased areas, which improves the efficiency of the FHB assessment in the field.

Keywords: fusarium head blight disease; color imaging; deep neural network

1. Introduction

As a widely cultivated crop and staple food in the world, wheat (*Triticum aestivum* L.) is crucial for ensuring food security. Unfortunately, its production is limited by many abiotic and biotic stresses. Among biotic stresses, fungal diseases are the most important. Worldwide, one of the most devastating fungal diseases of wheat is Fusarium head blight (FHB) or scab, which is caused by a number of different *Fusarium* species, but most commonly *F. graminearum*. FHB attacks wheat spikes, which causes yield losses as high as 40% under severe epidemics [1]. In addition to causing yield losses, FHB pathogens

also produce mycotoxins that can contaminate the grain, often rendering it unusable. One of the most common mycotoxins produced by *F. graminearum* in grain is deoxynivalenol (DON), which is toxic to humans and animals, especially swine. When humans ingest grain contaminated with DON, the common symptoms include vomiting, dizziness, and nausea [2,3]. Due to the problems associated with this mycotoxin in food and feed chains, DON levels in grain are routinely assayed.

The development of resistant varieties is one of the most important means for ameliorating the impact of FHB. However, the tasks of manually phenotyping FHB severity on hundreds of lines in the field and selecting based on resistance is a costly, labor-intensive, and time-consuming process. Moreover, conventional phenotyping protocols suffer from inaccurate subjective assessments (raters often do not capture the true level of disease), rater-fatigue (people tire quickly from this work), and inter-rater variation (individual raters will score diseased kernels on spikes differently), which are all factors contributing to increased experimental error in the disease assessments [4]. It is clear that breeding programs and research requiring FHB field phenotyping would be well-served by developing a high throughput phenotyping system.

FHB pathogens infect spikelets (i.e., inflorescence unit comprised of two glumes surrounding one or more florets) within the wheat spike. In doing so, they destroy the tissue, altering the chlorophyll and water content as well as the structure of the organ [5]. Thus, the optical properties of infected florets are different from those of healthy florets within a spike. Innovations in sensors and data processing methods offer potential methods to detect FHB, and techniques such as hyperspectral and color imaging have been investigated.

Hyperspectral imaging can provide considerable spectral and spatial information. A hyperspectral image contains hundreds of images that correspond to each wavelength [6]. This technique has been used to detect several crop diseases [7–9]. With respect to FHB, Ropelewska et al. [10] designed a hyperspectral imaging system to capture images of infected and healthy wheat kernels and then used the MaZda software to extract approximately 200 textures at selected wavelengths. The researchers analyzed these features to develop classification models to distinguish infected wheat kernels, and the accuracy attained was 94%–100%. Similarly, Barbedo et al. [11] processed hyperspectral images of a wheat kernel and segmented them from the background using four selected bands. After completing the tests, they 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 complete spectral data for wheat spikes. Then, these researchers used partial least squares regression (PLSR) analyses to establish models that could assess the percentage of coverage by 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) and used them successfully to distinguish diseased and healthy wheat spikes. The previously mentioned studies were all conducted under laboratory conditions. The researchers also investigated the potential of applying hyperspectral imaging for detecting FHB in the field. Whetton et al. [14] continued their work of processing hyperspectral images collected in winter wheat fields, with a spectral range of 400–1000 nm. PLSR was used to predict FHB and assess disease severity, and the residual prediction deviation value was 2.27. Furthermore, Jin et al. [15] 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 reached as high as 74.3%. Most studies have selected several specific wavebands and have developed models or 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 [16]. Although hyperspectral imaging can be useful in detecting FHB, several problems limit its application. First, hyperspectral imaging is costly and, thus, difficult to widely promote. Second, some specific accessories (e.g., calibration board, stage, and objective lenses) are essential when collecting hyperspectral data, especially in the field. Lastly, the extraction of information from huge hyper-spectral data sets is challenging [17].

Alternatively, color imaging is inexpensive and is easy to handle. Color imaging can capture the color and texture information of an object, and has been implemented to detect and assess plant diseases [18–20]. For *Fusarium* or FHB detection, researchers have explored the usefulness of color imaging. Cambaza et al. [21] found that color imaging can be applied to monitor *Fusarium* based on the color differences between healthy and infected oat grains. The textural features from channels R (red), G (green), B (blue), and other color models were analyzed to develop classification models. Ruan et al. [22] used 55 color and texture features to develop a four-layer neural network to estimate the percentage of visual scabby wheat kernels. Color imaging was also investigated to recognize wheat spikes under field conditions, which provides the basis of identifying FHB infection. Bakhouché et al. [23] extracted seven features including mean, variance, energy, entropy, contrast, skewness, and kurtosis of intensities, from wheat color images, and then used the K-means algorithm to detect spikes. Li et al. [24] used the Laws' texture energy method and trained a classification model using a neural network. Zhu et al. [25] chose the RGB (red, green, blue), Lab (L for lightness and a and b for the color-opponent dimensions), and HSV (hue, saturation, value) color spaces. The researchers applied all color components to train a support vector machine (SVM) model that can detect wheat spikes. Zhou et al. [26] proposed an approach using a color feature (color coherence vectors), a texture feature (gray level co-occurrence matrix), and an image feature (edge histogram descriptor) to train a twin SVM model. Thus, this successfully extracted 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. Most recently, Hasan et al. [27] tested a Faster Region-based Convolutional Network (RCNN) model to detect spikes and output their bounding boxes. Pound et al. [28] trained a deep convolutional neural network (DCNN) model that localized wheat spikes in the laboratory. Generally, extracting features is not necessary, since accurate models can be obtained using a sufficiently large dataset. Deep learning (DL) is promising for detecting wheat spikes and FHB, but its application requires further exploration [29,30].

Assessment of FHB levels in the field is important for a number of different applications such as breeding for resistance to the disease, assessing the efficacy of various disease management practices (e.g., fungicide applications and crop residue treatments), and determining crop quality for various end uses. Therefore, accurate assessments of FHB on wheat plants in the field, particularly quantitative ones, are very important. To obtain enough sample data for accurate assessments of FHB, a population of spikes must be assessed. This requires not only distinguishing the boundaries of individual wheat spikes but also resolving which florets within the spikes are diseased or healthy. Previous studies have revealed the potential of color imaging for assessing FHB infection in wheat spikes. Moreover, DL techniques can improve the efficiency of color imaging.

Therefore, our study focuses on the usefulness of color imaging collected in the shadow condition and the application of DL techniques for assessing FHB infection in wheat. The specific aims of this study were to: (1) train a DCNN model that can extract individual wheat spikes from one image with a complex background and (2) propose a novel approach to detecting FHB-infected areas 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, with proposals for improvements. Lastly, the conclusions are provided in Section 5.

2. Materials and Methods

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

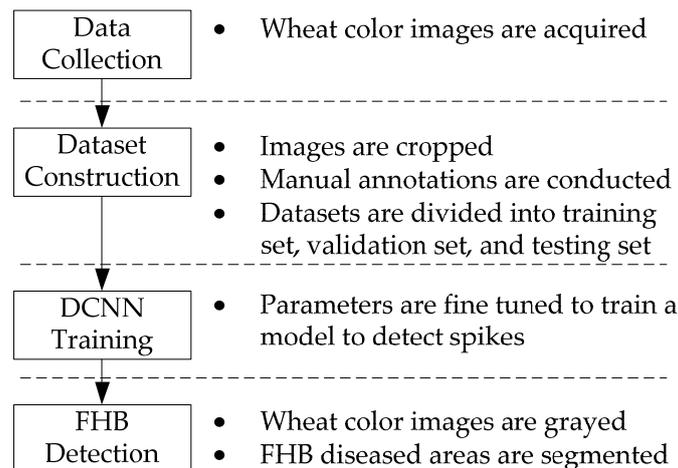


Figure 1. Workflow diagram of the proposed FHB detection.

2.1. Data Collection

Three wheat lines were cultivated in plots of approximately 1.6 m × 0.7 m (length × width) on the St. Paul campus at the University of Minnesota (Minnesota, USA). The wheat lines selected had different levels of susceptibility to FHB and included Wheaton (susceptible with awns (i.e. bristle-like appendages extending from the glumes, i.e. outside bracts of the spikelet)), Roblin (susceptible, awnless), and MN00269 (late maturing susceptible, awned). To initiate FHB infection in the plots, all lines were inoculated using the grain spawn method [31]. Collection of images was performed with a phenotyping platform (approximately 2.7 m × 2.5 m, length × width) between 11:00 and 14:00 on 16 July, 2018, during which most of the wheat lines were at the milk stage of development. The phenotyping platform was manually advanced and stopped to capture the wheat images. A camera (Canon EOS Rebel T6i) with a large field of view was installed on the platform, as illustrated in Figure 2a so that the entire wheat plot was captured in each image. In addition, a shelter was designed to achieve uniform light on the subject matter (Figure 2b). The camera was controlled through an onboard computer to collect the wheat images. All images were saved in a JPEG (Joint Photographic Experts Group) compressed format for later analysis. The captured images have a 6000 × 4000-pixel resolution. In total, 12 images were collected including four images of each of the three wheat lines.

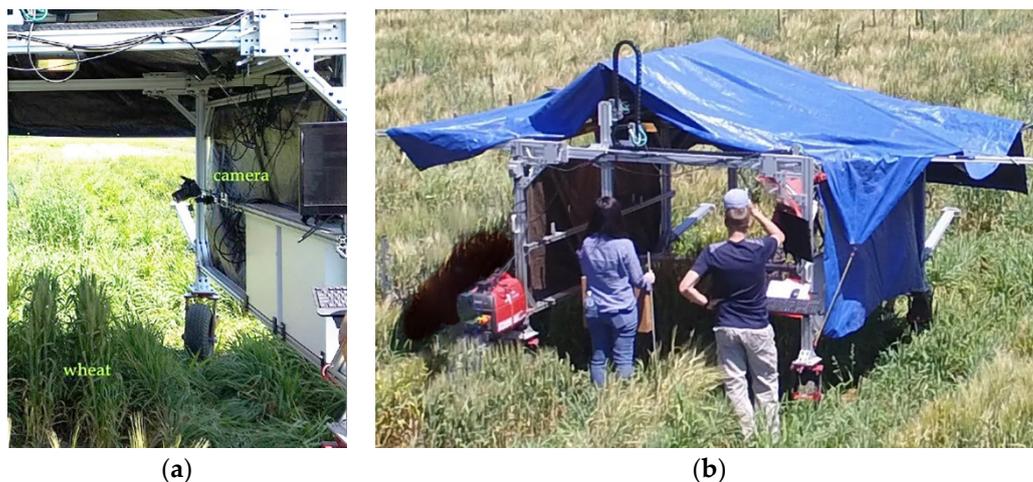


Figure 2. Phenotyping platform for collecting color images of wheat plots in the field: (a) Position of the color camera within the phenotyping platform. (b) Overall view of the phenotyping platform.

2.2. Data Augmentation

This small set of images (12 images) was insufficient to train the DL model. To increase the sample size, two common steps were performed. First, the raw images were divided into smaller sub-images, as demonstrated in Figure 3. The sub-images were 700×700 pixels in resolution with 100 overlapping pixels between adjacent images. Then, additional datasets were generated by mirroring the sub-images horizontally and vertically. In the end, 2829 images were generated after the augmentation.

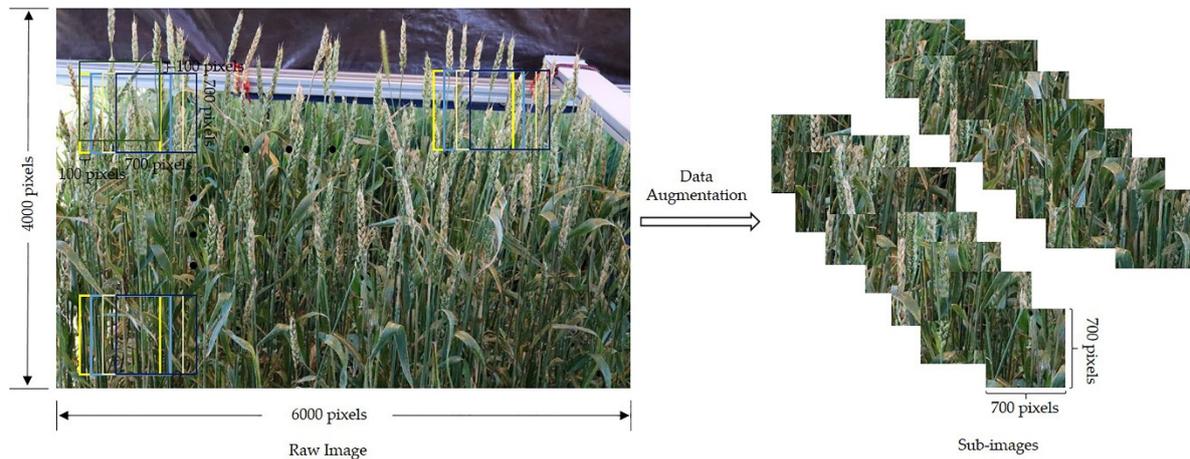


Figure 3. Process used to augment the number of data images for analysis. Each rectangular box to the right is a sub-image.

2.3. DCNN Model

The ultimate goal of this study was to reliably detect diseased areas of individual spikes, which is a task where object detection techniques are key. The first essential step in this process is detecting the boundaries of the spikes and removing the background. Based on the DL technique, He et al. [32] proposed the Mask RCNN model, which can not only classify objects, but can also output their outlines at the pixel level. It extends the Faster RCNN model by combining classification, bounding boxes, and mask generation, which 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, which generates 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 [33], and only the last few layers were trained and fine-tuned using transfer learning. ResNet 50 and ResNet 101 [34] are two backbone network options. The former was used in this case given its accuracy and small training dataset. Batch normalization was applied to prevent overfitting. The batch size was set to 2 because it requires only a small amount of 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 the union ratio with the ground-truth bounding boxes were higher than 0.5, the anchors were considered to be available. The other parameters were configured to the default setting.

Before training the model, a training dataset and a validation dataset were annotated. An open-source software, Lableme [35], was used to label the datasets. Then, the spikes were selected and labeled the “spike” class by drawing polygons manually, as exhibited in Figure 4.

Furthermore, the software generated JSON (JavaScript Object Notation) files that record the boundaries of the spikes.



Figure 4. Annotation of wheat 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 computer with CUDA 9.0. A dataset containing 1959 images (Wheaton: 595 images, Roblin: 970 images, MN00269: 394 images) that were randomly selected from 2829 images was obtained as the training dataset. Among these images, 989 and 970 represent awned and awnless wheat spikes, respectively. A total of 420 (Wheaton: 101 images, Roblin: 200 images, MN00269: 119 images) wheat images were randomly selected to generate the validation dataset, with 220 and 200 images corresponding to awned and awnless wheat spikes, respectively. The remaining images (450 images, Wheaton: 116 images, Roblin: 222 images, MN00269: 112 images) were used as the testing dataset, with images 228 and 222 corresponding to awned and awnless wheat spikes, respectively. 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 and the 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.

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

$$r = \frac{TP}{TP + FN} \quad (2)$$

$$p = \frac{TP}{TP + FP} \quad (3)$$

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 indicative of improved performance. However, an individual parameter could not quantify the model adequately. Therefore, the average precision (AP), known as the area under the precision–recall curve, was applied as an indicator to estimate the model.

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

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

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

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

where n is the number of testing samples, 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 the green hues of chlorophyll in healthy spikes, FHB-infected spikes are tan or straw-colored, similar to senescing tissue. For this reason, it is best to assess FHB at a stage prior to the initiation of senescence in the spike. Occasionally, infected spikelets may exhibit a pinkish or salmon-colored hue. This is due to the mycelium and conidia of the *Fusarium* fungus growing on the surface of the tissue. Within an individual spike, the extent of symptomatic spikelets can range from 0% to 100%. The DCNN was able to segment the spikes from the background and output their boundaries, which aided the detection of diseased areas. The spike is a very complex structure consisting of the spikelets, rachis (the axis of the spike to which the spikelets are attached), and awns. Additionally, image analysis of individual spikes must also consider the spaces in between these structures. Therefore, further steps were performed to process the segmented spikes and extract the diseased areas accurately.

The stark color differences of healthy vs. diseased spikelets were exploited for assessing the percentage of FHB infection in individual spikes. In accordance with previous tests [36], a new color feature (GB) was proposed by using the green and blue components to highlight the diseased areas of each spike. This new color feature was calculated using function (7).

$$GB = \alpha \times B - G \quad (7)$$

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 can sometimes be scattered in wheat, they do show a similar appearance. In accordance with this 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, the pixels are classified into the same class. The selection of the 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 was insufficient to recognize all unconcentrated diseased areas. Therefore, a modified region growing algorithm was applied to process the gray image [37]. 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 was tagged as the K th 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 setting range, the pixel was regarded as being in the same class and saved into the “stack.” Third, the second operation was repeated until the “stack” was empty. Lastly, Steps 1–3 were repeated until every pixel was marked. Many small areas are inevitably generated because of background noise, especially at the junctions of the spikelets and rachis. Thus, the area of each region was calculated, and small diseased regions were removed. At this point, the region-growing algorithm was completely implemented.

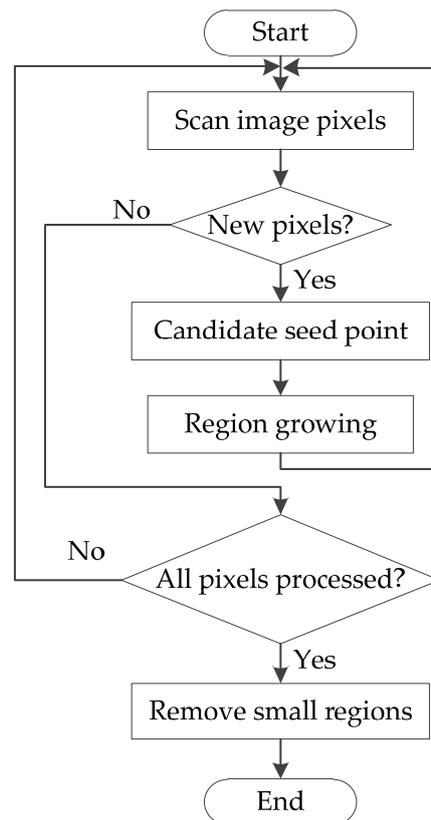


Figure 5. Flowchart of the region growing algorithm used to determine FHB-infected areas of wheat spikes.

Subsequently, the close operation was performed on the results. Lastly, the ratio of FHB-infected areas to the entire spike area was calculated in accordance with the standard [38,39]. The methods for classifying the reaction of wheat to FHB can vary depending on the intended purpose. For the general selection of resistant wheat lines in breeding programs, FHB severities are classified into four classes [2]. However, for other purposes (e.g., inheritance and molecular mapping studies), FHB assessments are made to the nearest integer on a 0% to 100% severity scale. The latter assessment was applied and the severity scoring was calculated using MATLAB r2016a software.

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 the Tensorboard [40], 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.

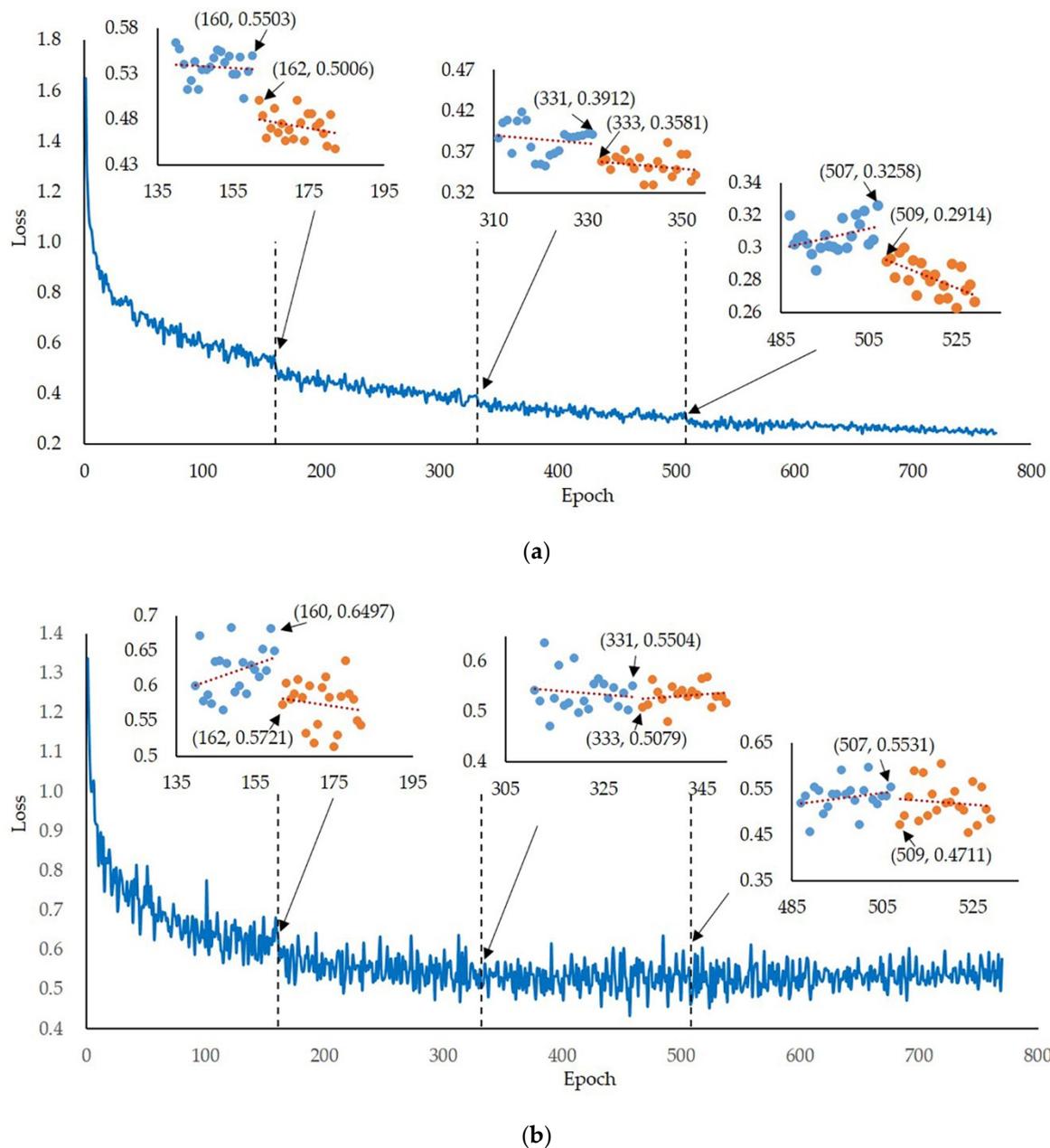


Figure 6. Losses of the Mask RCNN for training and validation: (a) Training loss. (b) Validation loss.

3.2. Model Testing for Spike Detection

The Mask RCNN model was retrained to accomplish spike detection. To evaluate its performance, the testing dataset, which contained 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. This value indicated that the retrained model could reliably 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 s, $RMSE$ s, and $rRMSE$ s 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; 1.17, 1.16, and 1.18; and 21.37%, 25.99%, and 15.22%, respectively.

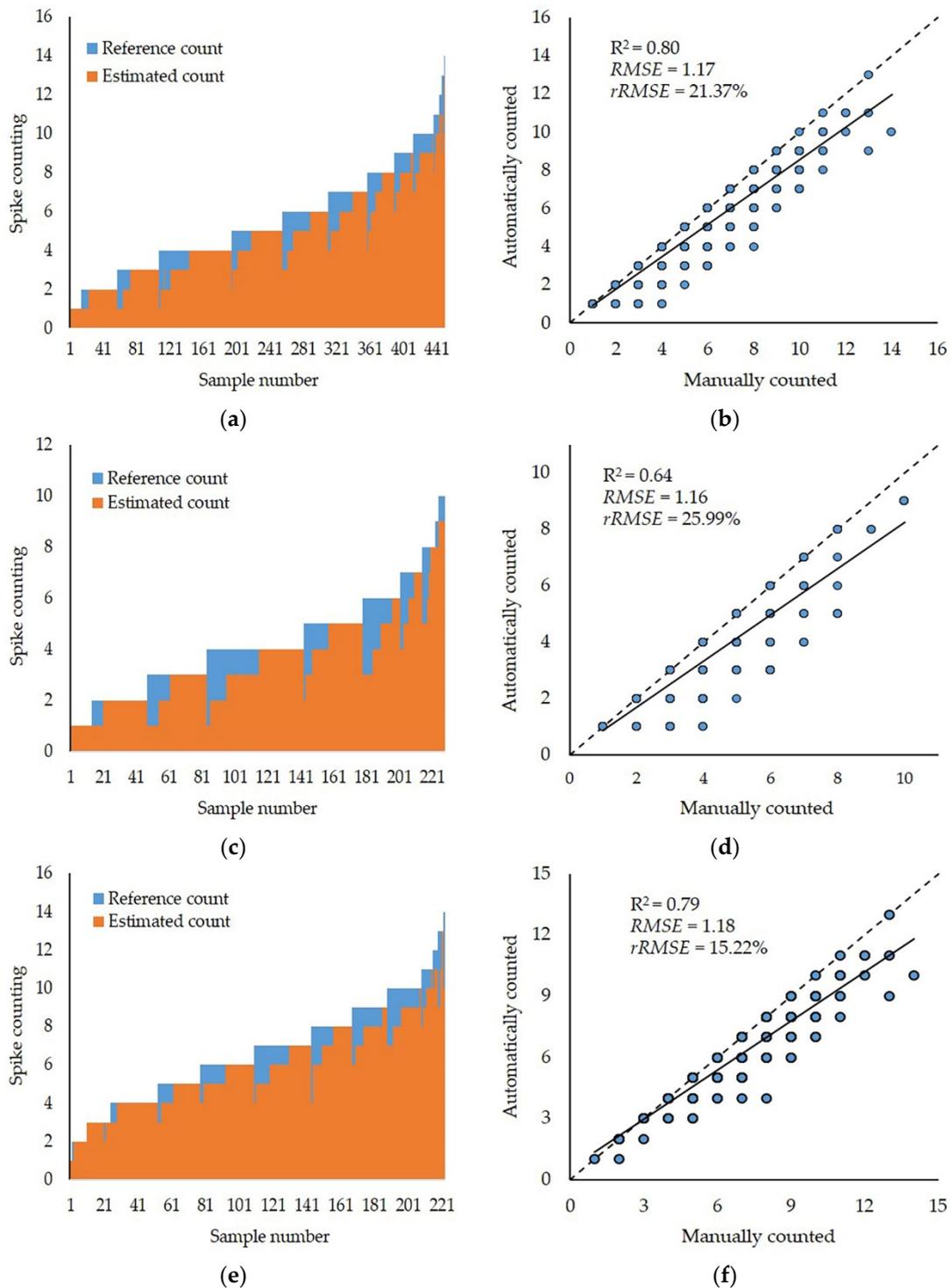


Figure 7. Comparison between the automatically and manually counted spike numbers: (a,b) Comparison and correlation of automatically and manually counted numbers for all testing spikes. (c,d) Comparison and correlation of automatically and manually counted numbers for the awned spikes. (e,f) Comparison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

The majority of the identified spikes could be assigned as either healthy or diseased. In Figures 8 and 9, the spike density in one image has no influence on the detection and is adaptable to both awned wheat and awnless wheat varieties. As shown in Figure 10e, some portions of spikes obscured by leaves can also be detected. 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 classified in the background because the camera could focus only on limited field of view and depth of field, as illustrated in Figure 10a,b. Furthermore, many spikes were severely infected by FHB or became prematurely senescent. Consequently, the tissue of the spikelets and rachis was severely compromised, and their features (e.g., shapes, spaces between spikelets) were not evident, which makes it hard for the model to detect spikes. Several samples are depicted in Figure 10c,d. Moreover, only a small area of a spike was visible in the image. Thus, extracting features and classifying pixels, such as the yellow rectangle portions displayed in Figure 10e,f, was difficult for the model. Lastly, some spikes were adjacent or intersecting, as demonstrated in Figure 10g,h, which, thereby, causes all or parts of them being missed.

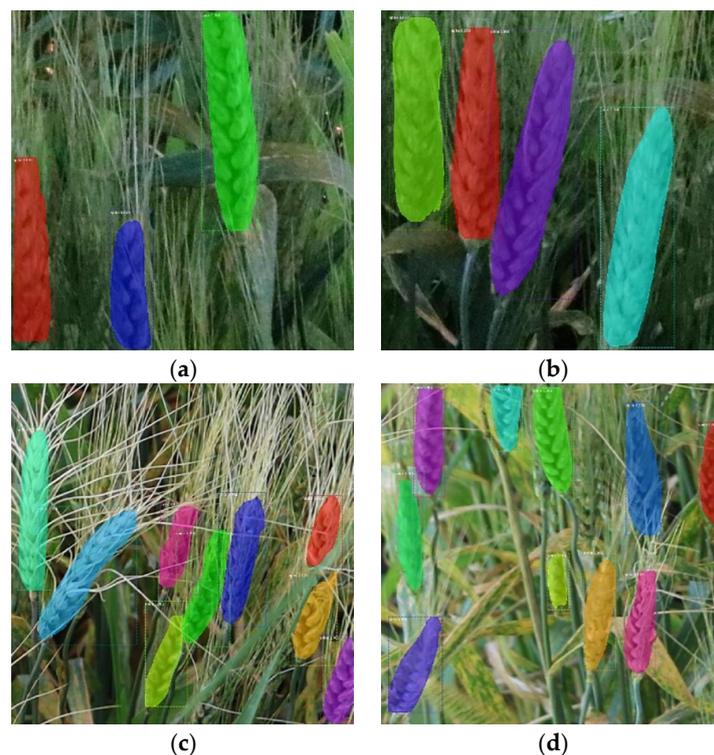


Figure 8. Detection results of awned spikes in different densities: (a,b) Detection results of awned spikes in low density. (c,d) Detection results of awned spikes in high density.

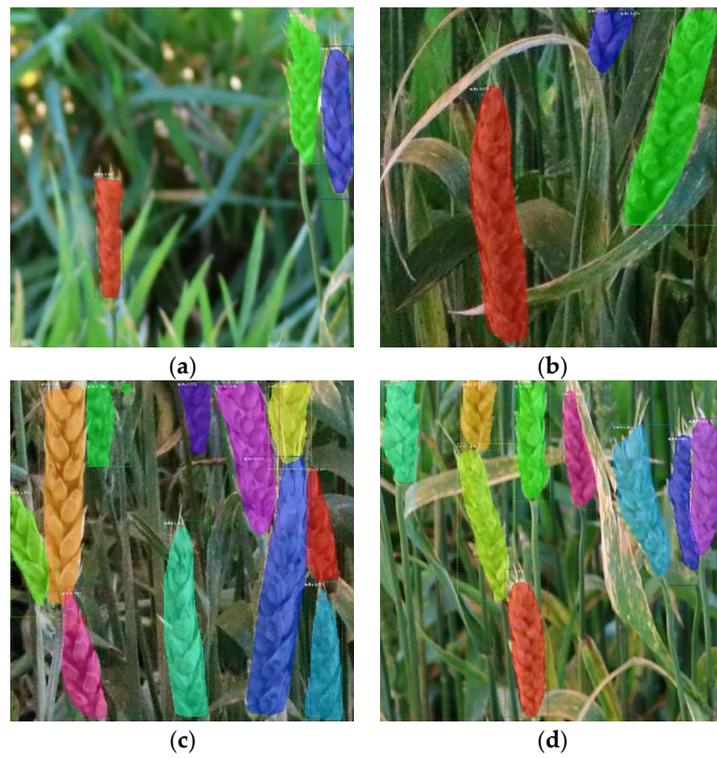


Figure 9. Detection results of awnless spikes in different densities: (a,b) Detection results of awnless spikes in low density. (c,d) Detection results of awnless spikes in high density.

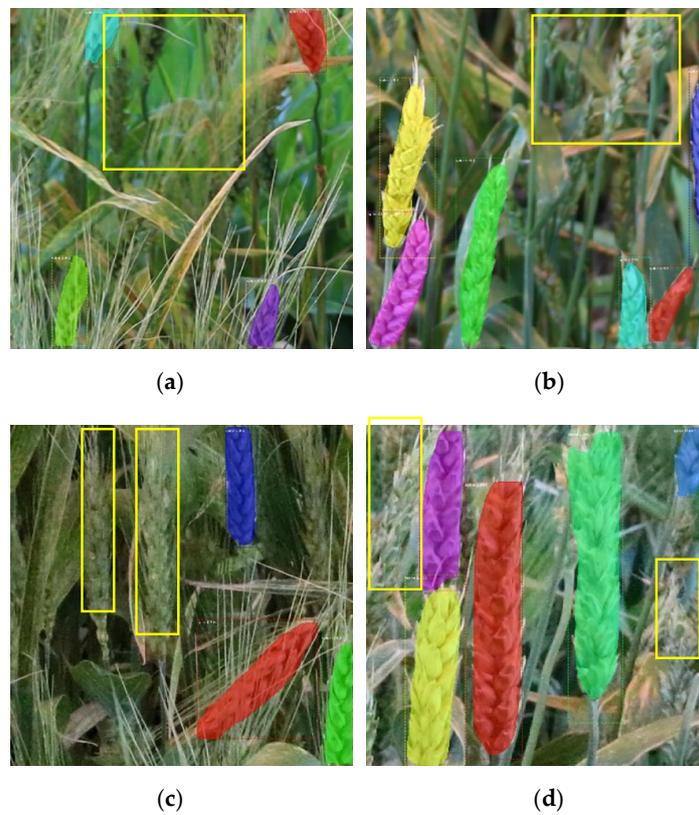


Figure 10. Cont.

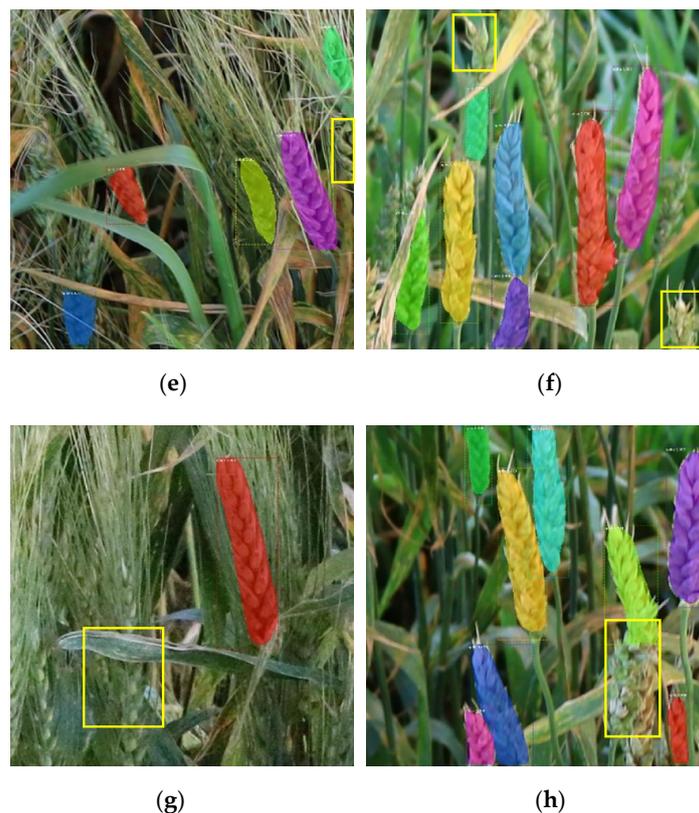


Figure 10. Examples of several spikes that failed to be detected by the model. They are pointed out using the yellow rectangles: (a, b) Spikes are taken as the background. (c, d) Spikes are severely destroyed. (e, f) Spikes are not entire. (g, h) Spikes are adjacent to each other.

3.3. FHB Detection

The Mask RCNN model can partially detect spikes and output their mask pixels, which enables 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 and healthy spikelets were able to be distinguished. The gray values were high if the pixels belong to the diseased spikelet class. Otherwise, the values were low. In addition, α was given a set of values, and the corresponding gray images were evaluated to obtain optimum performance. Tests showed that when α was set to 1.4, the diseased areas were brighter than the healthy areas in the gray images and more likely to be detected. In Figure 11, the images of three wheat varieties (Wheaton, Roblin, and MN00269) were processed using GB. Two spike samples of each wheat variety were selected to show the performance of the GB feature. The results indicated that the GB feature was suitable for detecting the FHB-infected areas of multiple wheat varieties.

The gray value differences between diseased and healthy pixels enable the detection of the FHB-infected areas of each spike using conventional segmentation algorithms of machine learning. In this scenario, the modified region-growing algorithm was used to segment the diseased areas from the entire spike. The K-means and Otsu's 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 background noise. The final results are exhibited in Figures 12 and 13. In the present study, K was assigned to 4 considering the spatial resolution of the images. Thus, the pixels in the gray images were categorized into four classes, which include the background, healthy tissue, diseased tissue, and connected areas between the spikelets. The segmentation results were unfavorable because the connected areas were typically classified by mistake (Figure 13b,e,h). In Figure 13c,f,i, the segmentation accuracy of Otsu's algorithm was the lowest, and some non-diseased areas were improperly considered

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

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 14. For most wheat spikes, the ratios of FHB-infected areas to the entire spike area were low, which indicates that the disease levels of the spikes were not high.

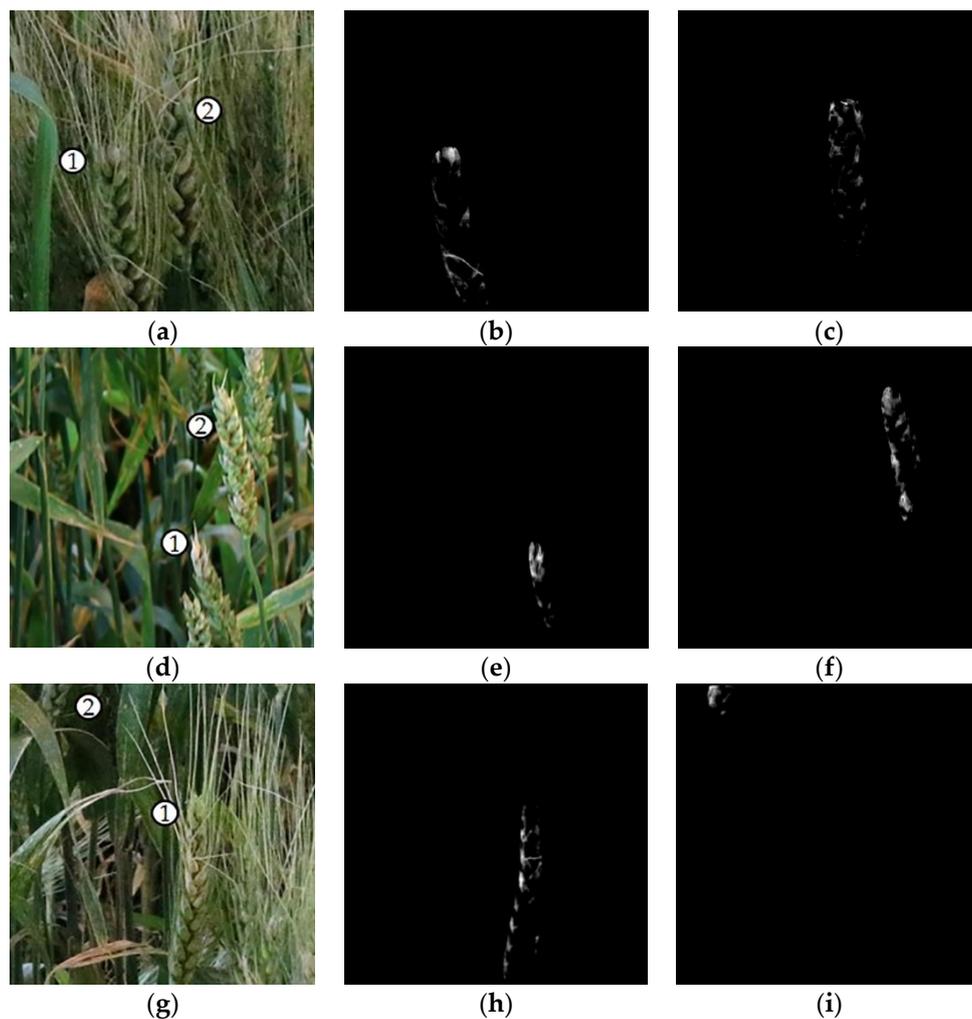


Figure 11. Examples of healthy and diseased spikes with three varieties (Wheaton, Roblin, and MN00269) and shapes in the gray images: (a, d, g) Color images of awned and awnless wheat varieties. (b, e, h) Gray images of number spike 1 in image a, d, and g. (c, f, i) Gray images of spike 2 in image a, d, and g.

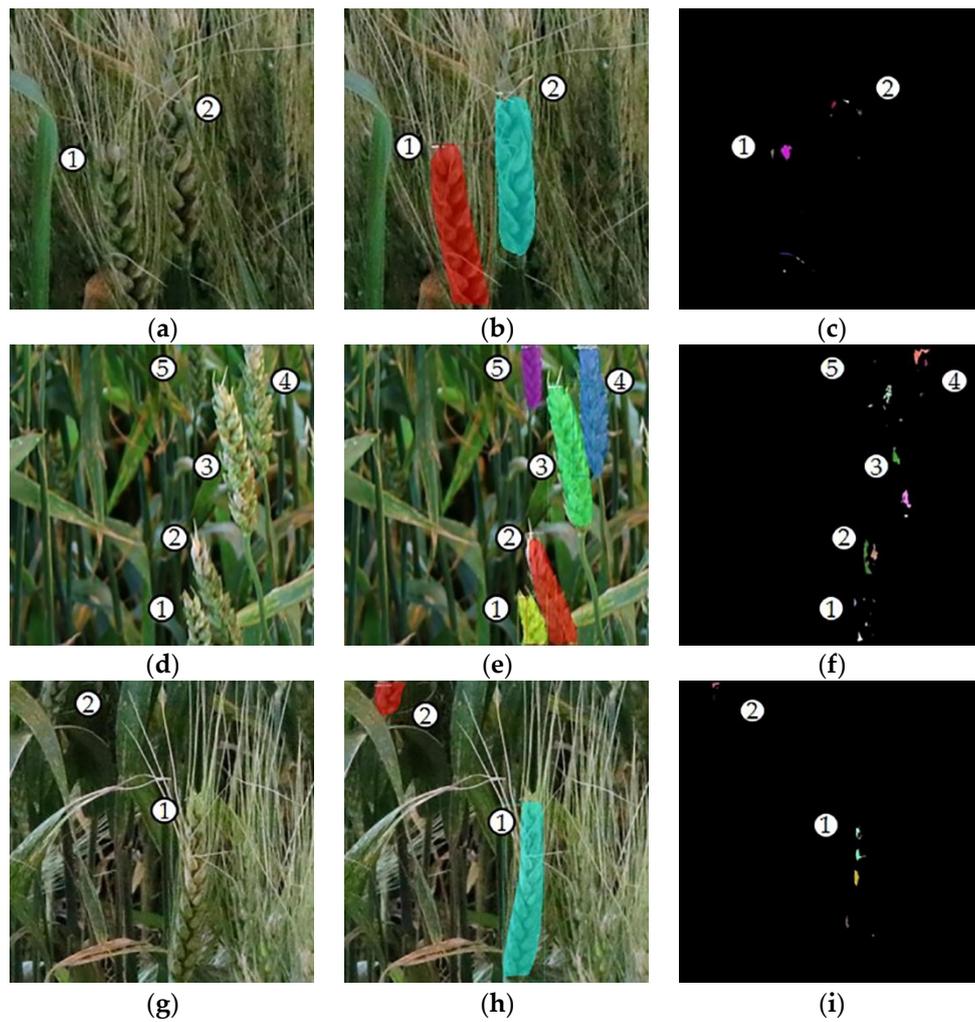


Figure 12. FHB detection results of three wheat varieties (Wheaton, Roblin, and MN00269) using the region growing algorithm: (a, d, g) Color images of awned and awnless wheat varieties. (b, e, h) Spike detection results of awned and awnless wheat varieties. (c, f, i) FHB detection results of awned and awnless wheat varieties.

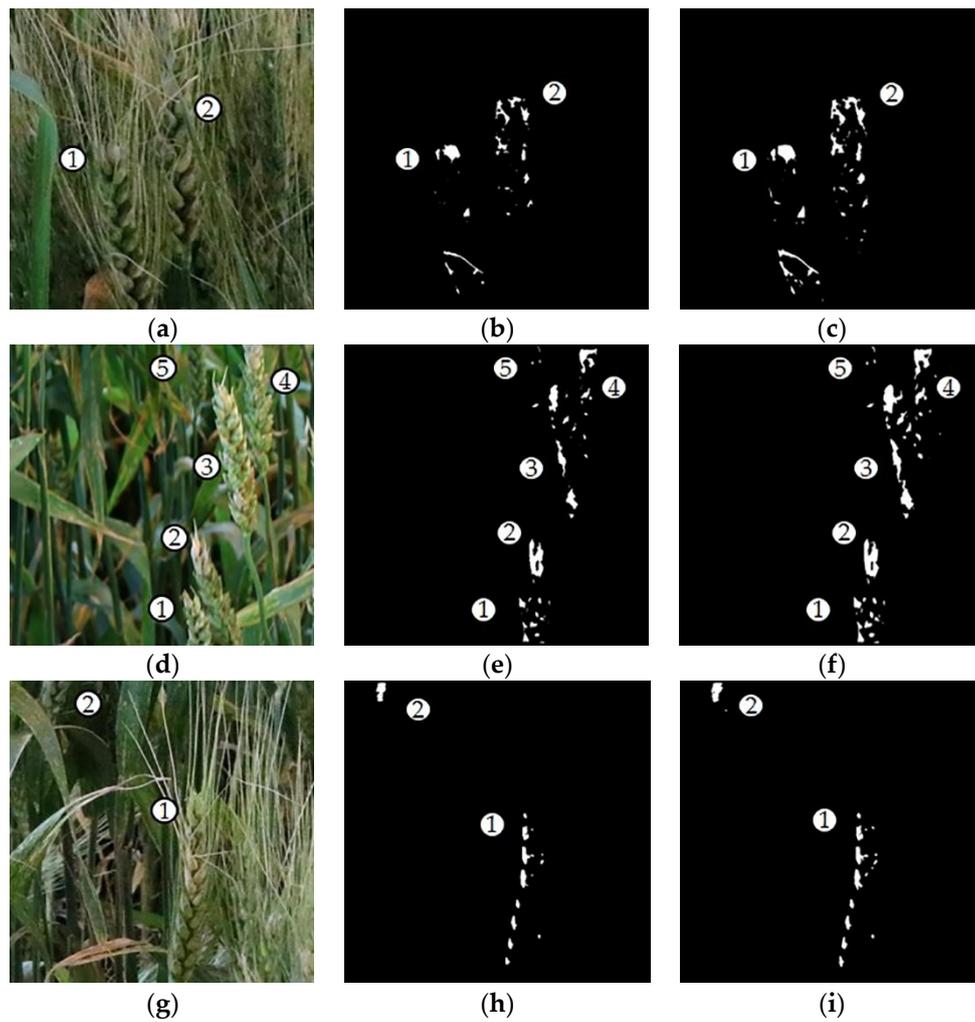


Figure 13. Segmentation results of the K-means algorithm and Otsu’s algorithm: (a, d, g) Color images of awned and awnless wheat varieties. (b, e, h) FHB detection results of awned and awnless wheat varieties using the K-means algorithm. (c, f, i) FHB detection results of awned and awnless wheat varieties using Otsu’s algorithm.

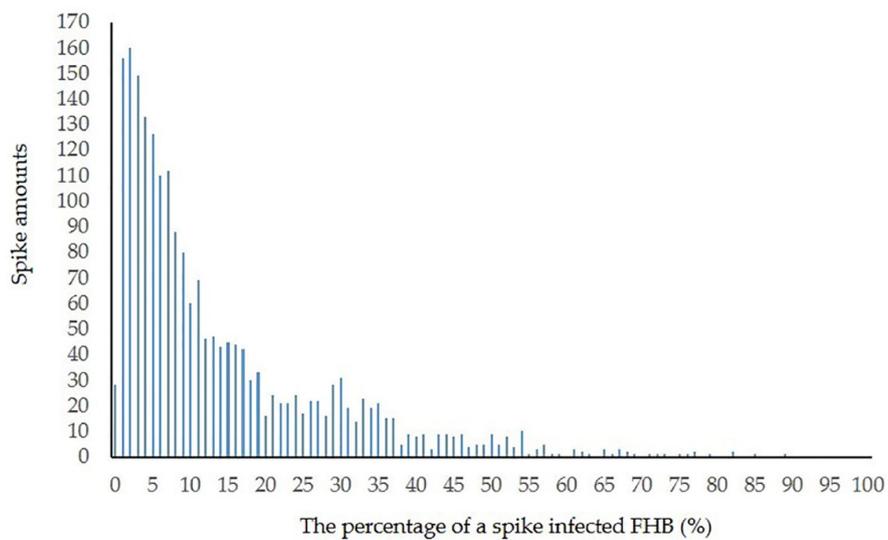


Figure 14. Number of spikes in each diseased percentage.

4. Discussion

4.1. Analysis of Mask RCNN Model Performance

A Mask 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. The retrained model demonstrated favorable performance in spike detection (mAP is 0.9201). However, some questions should be addressed and analyzed further. First, the spikes represent various orientations in the color images. Some neighboring spikes overlapped or their edges touched, which made detecting each spike difficult. This difficulty is also the shortcoming of the Mask RCNN model. Second, the color images contain not only entire spikes, but also many partial spikes. The tests showed that the retrained model could detect some partial spikes in the middle of the color images. However, it was less effective at recognizing spikes located at the edges of an image or those not in the depth of field focus. Lastly, the setting of anchor parameters will affect the precision of spike detection. The spike sizes in color images need to be considered. Thus, the anchor parameters should be consistent with the spike sizes. If these parameters are not established, small spikes will be missed and only a portion of large spikes will be detected.

In addition, the tests showed that the awns of wheat could slightly influence spike detection. The DCNN model was retrained using the labeled spikes. The labeled spikes contain spikelets, and no leaves for both awned and awnless wheat lines, so other leaf diseases (e.g., leaf rust, leaf spot) can not affect spike detection. For awned spikes, some labeled regions may contain many awns, but the spikelets are visible. If the spikes are not completely covered by surrounding awns, the model is able to detect spikelets and recognize spikes.

4.2. Analysis of FHB Detection

A new color feature, GB, was proposed in this study. The results showed that this color feature adequately used the differences between healthy and diseased spikelets to recognize and segment the FHB-infected areas of spikes. However, some factors negatively affecting FHB detection remain. First, the timing for image capture is critical. The best time to assess FHB in wheat is when maximum disease severity has been reached, but before senescence due to maturity starts to occur in the healthy tissue. The reason for this is due to the color of senescent tissue closely resembling tissue infected by FHB. The greater the contrast between the colors of healthy vs. diseased tissue, the better the model will perform for segmenting FHB-infected areas. Since different wheat varieties have different maturation dates, assessments for FHB must be made several times during the season to capture the optimal timing for each. Another potential problem is that many diseased areas are too small to be preserved when the close operation was conducted, which results in an underestimation of FHB infection. This is particularly true of diseased areas on the edges of spikelets, which are likely to be ignored.

Furthermore, the effect of wheat awns on detecting diseased areas should be considered. This effect is likely due to the colors of wheat awn being similar to the diseased spikes being similar to the diseased spikes as they begin to senesce.

4.3. Next Steps

The implementation of a DL technique increases the feasibility of FHB detection in wheat under field conditions, but further work needs to be conducted in the future to improve the precision of FHB detection. First, images should be taken across different developmental stages to obtain the optimal timing for capturing maximum disease progression before the beginning of senescence. Second, additional images of wheat spikes from diverse lines (i.e., with different shapes, sizes, and hues), with different views, should be captured and tested, to enlarge the training dataset and improve the robustness and applicability of the trained model and FHB detection. Third, in this paper, spikes were first detected and then FHB-infected areas of each spike were segmented. A new thought can be

verified to train a model that can detect FHB-infected areas in wheat directly. In this case, FHB-diseased samples must be labeled. Fourth, for applications related to genetic inheritance and mapping studies, greater precision is required for severity. With data for pixels distinguishing healthy and diseased areas or spikelets of a sample population of spikes, one could derive a mean FHB severity for individual wheat varieties at a tremendous savings in time, labor, and cost over manual scoring methods that are currently utilized.

Eventually, based on the results of this study, special attention can be given to processing 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 with wheat plants infected in the field. To detect and distinguish 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 different wheat varieties, even those with different morphologies and sizes. R^2 , $RMSE$, and $rRMSE$ for 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 was 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 performed favorably in detecting the FHB-infected 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, which decreases the cost and improves 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.J.S., and M.Z. Supervision, C.Y. Writing-original draft, R.Q. Writing-review & editing, M.Z., C.Y., B.J.S., and C.D.H.

Funding: The USDA-ARS U.S. Wheat and Barley Scab Initiative (Funding No. 58-5062-8-018) supported this research. 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 National Natural Science Foundation of China (Grant No. 31971786) and the Industrial Partnership Grants from MnDRIVE Robotics, Sensors and Advanced Manufacturing Initiative also partially funded the project.

Acknowledgments: The authors are grateful to Youbing Wang from the University of Minnesota for his instructions in deep learning. We thank Tamas Szinyei, Ryan Johnson and Jacob Botkin for excellent technical assistance. We also appreciate the financial support from the China Scholarship Council.

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

References

1. Bottalico, A.; Perrone, G. Toxigenic *Fusarium* species and mycotoxins associated with head blight in small-grain cereals in Europe. *Eur. J. Plant Pathol.* **2002**, *108*, 611–624. [[CrossRef](#)]
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 Sci.* **2014**, *54*, 2077–2083. [[CrossRef](#)]
3. Bai, G.; Shaner, G. Management and resistance in wheat and barley to Fusarium head blight. *Annu. Rev. Phytopathol.* **2004**, *42*, 135–161. [[CrossRef](#)] [[PubMed](#)]
4. Steffenson, B.J. Chapter 10—Fusarium head blight of barley: Impact, epidemics, management, and strategies for identifying and utilizing genetic resistance. In *Fusarium Head Blight of Wheat and Barley*; Leonard, K.J., Bushnell, W.R., Eds.; The American Phytopathological Society: Saint Paul, MN, USA, 2003; pp. 241–295.
5. Kang, Z.; Buchenauer, H. Cytology and ultrastructure of the infection of wheat spikes by *Fusarium culmorum*. *Mycol. Res.* **2000**, *104*, 1083–1093. [[CrossRef](#)]

6. Xie, C.; Shao, Y.; Li, X.; He, Y. Detection of early blight and late blight diseases on tomato leaves using hyperspectral imaging. *Sci. Rep.* **2015**, *5*, 16564. [CrossRef]
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. *J. Plant Dis. Prot.* **2018**, *125*, 5–20. [CrossRef]
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. [CrossRef]
9. Yao, Z.; Lei, Y.; He, D. Early visual detection of wheat stripe rust using visible/near-infrared hyperspectral imaging. *Sensors* **2019**, *19*, 952. [CrossRef]
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. *Eur. Food Res. Technol.* **2018**, *244*, 1453–1462. [CrossRef]
11. Barbedo, J.G.A.; Tibola, C.S.; Fernandes, J.M.C. Detecting *Fusarium* head blight in wheat kernels using hyperspectral imaging. *Biosyst. Eng.* **2015**, *131*, 65–76. [CrossRef]
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. *Biosyst. Eng.* **2018**, *166*, 101–115. [CrossRef]
13. Bauriegel, E.; Giebel, A.; Geyer, M.; Schmidt, U.; Herppich, W.B. Early detection of *Fusarium* infection in wheat using hyper-spectral imaging. *Comput. Electron. Agric.* **2011**, *75*, 304–312. [CrossRef]
14. 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. *Biosyst. Eng.* **2018**, *167*, 144–158. [CrossRef]
15. 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 Sens.* **2018**, *10*, 395. [CrossRef]
16. Signoroni, A.; Savardi, M.; Baronio, A.; Benini, S. Deep learning meets hyperspectral image analysis: A multidisciplinary review. *J. Imaging* **2019**, *5*, 52. [CrossRef]
17. Qiu, R.; Wei, S.; Zhang, M.; Li, H.; Sun, H.; Liu, G.; Li, M. Sensors for measuring plant phenotyping: A review. *Int. J. Agric. Biol. Eng.* **2018**, *11*, 1–17. [CrossRef]
18. Sun, Y.; Jiang, Z.; Zhang, L.; Dong, W.; Rao, Y. SLIC_SVM based leaf diseases saliency map extraction of tea plant. *Comput. Electron. Agric.* **2019**, *157*, 102–109. [CrossRef]
19. Zhu, J.; Wu, A.; Wang, X.; Zhang, H. Identification of Grape Diseases Using Image Analysis and BP Neural Networks. *Multimed. Tools Appl.* **2019**. Available online: <https://link.springer.com/content/pdf/10.1007%2Fs11042-018-7092-0.pdf> (accessed on 10 September 2019). [CrossRef]
20. 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. *Int. J. Adv. Sci. Eng. Inf. Technol.* **2019**, *9*, 194–203. [CrossRef]
21. Cambaza, E.; Koseki, S.; Kawamura, S. Why RGB imaging should be used to analyze *Fusarium graminearum* growth and estimate deoxynivalenol contamination. *Methods Protoc.* **2019**, *2*, 25. [CrossRef]
22. 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 Chem. J.* **1998**, *75*, 455–459. [CrossRef]
23. Bakhouch, M.; Cointault, F.; Gouton, P. Texture analysis with statistical methods for wheat ear extraction. In Proceedings of the SPIE 6356, Eighth International Conference on Quality Control by Artificial Vision, Le Creusot, France, 29 May 2007.
24. 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.
25. Zhu, Y.; Cao, Z.; Lu, H.; Li, Y.; Xiao, Y. In-field automatic observation of wheat heading stage using computer vision. *Biosyst. Eng.* **2016**, *143*, 28–41. [CrossRef]
26. 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. *Front. Plant Sci.* **2018**, *9*, 1024. [CrossRef]
27. 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. [CrossRef]
28. 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. [CrossRef]

29. 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. *Agric. For. Meteorol.* **2019**, *264*, 225–234. [[CrossRef](#)]
30. Mohanty, S.P.; Hughes, D.P.; Salathé, M. Using deep learning for image-based plant disease detection. *Front. Plant Sci.* **2016**, *7*, 1419. [[CrossRef](#)] [[PubMed](#)]
31. Buerstmayr, H.; Ban, T.; Anderson, J.A. QTL mapping and marker-assisted selection for Fusarium head blight resistance in wheat: A review. *Plant Breed.* **2009**, *128*, 1–26. [[CrossRef](#)]
32. He, K.; Gkioxari, G.; Dollár, P.; Girshick, R. Mask R-CNN. *arXiv* **2017**, arXiv:1703.06870.
33. Lin, T.Y.; Maire, M.; Belongie, S.; Hays, J.; Perona, P.; Ramanan, D.; Dollár, P.; Zitnick, C.L. Microsoft COCO: Common objects in context. In *Computer Vision—ECCV 2014*; Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T., Eds.; Springer International Publishing: Cham, Switzerland, 2014; pp. 740–755.
34. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the European Conference on Computer Vision, Amsterdam, The Netherlands, 8–16 October 2016; pp. 770–778.
35. Russell, B.C.; Torralba, A.; Murphy, K.P.; Freeman, W.T. LabelMe: A database and web-based tool for image annotation. *Int. J. Comput. Vis.* **2008**, *77*, 157–173. [[CrossRef](#)]
36. Qiu, R.; Yang, C.; Moghimi, A.; Anderson, J.; Steffenson, B.J.; Marchetto, P. *Detection of Fusarium Head Blight in Small Grains Using Hyperspectral Imaging*; National Fusarium Head Blight Forum: Saint Louis, MO, USA, 2018; pp. 32–37.
37. Github. Available online: <https://github.com/ZQPei/Seeded-Region-Growing-Algorithm> (accessed on 10 September 2019).
38. Stack, R.W.; McMullen, M.P. *A Visual Scale to Estimate Severity of Fusarium Head Blight in Wheat*; North Dakota State University Extension Service: Fargo, ND, USA, 1995; p. 1095.
39. Alisaac, E.; Behmann, J.; Kuska, M.T.; Dehne, H.W. Hyperspectral quantification of wheat resistance to Fusarium head blight: Comparison of two *Fusarium* species. *Eur. J. Plant Pathol.* **2018**, *152*, 869–884. [[CrossRef](#)]
40. Rampasek, L.; Goldenberg, A. TensorFlow: Biology’s gateway to deep learning? *Cell Syst.* **2016**, *2*, 12–14. [[CrossRef](#)] [[PubMed](#)]



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