Wheat FHB resistance assessment using hyperspectral feature band image fusion and deep learning

Kun Liang[1,2*], Zhizhou Ren[1,2], Jinpeng Song[1,2], Rui Yuan[1,2], Qun Zhang[1,2]

(1. *College of Artificial Intelligence, Nanjing Agricultural University, Nanjing 210031, China*;
2. *Jiangsu Province Engineering Lab for Modern Facility Agriculture Technology & Equipment, Nanjing 210031, China*)

Abstract: The breeding and selection of resistant varieties is an effective way to minimize wheat *Fusarium* head blight (FHB) hazards, so it is important to identify and evaluate resistant varieties. The traditional resistance phenotype identification is still largely dependent on time-consuming manual methods. In this paper, the method for evaluating FHB resistance in wheat ears was optimized based on the fusion feature wavelength images of the hyperspectral imaging system and the Faster R-CNN algorithm. The spectral data from 400-1000 nm were preprocessed by the multiple scattering correction (MSC) algorithm. Three feature wavelengths (553 nm, 682 nm and 714 nm) were selected by analyzing the X-loading weights (XLW) according to the absolute value of the peaks and troughs in different principal component (PC) load coefficient curves. Then, the different fusion methods of the three feature wavelengths were explored with different weight coefficients. Faster R-CNN was trained on the fusion and RGB datasets with VGG16, AlexNet, ZFNet, and ResNet-50 networks separately. Then, the other detection models SSD, YOLOv5, YOLOv7, CenterNet, and RetinaNet were used to compare with the Faster R-CNN model. As a result, the Faster R-CNN with VGG16 was best with the mAP (mean Average Precision) ranged from 97.7% to 98.8%. The model showed the best performance for the Fusion Image-1 dataset. Moreover, the Faster R-CNN model with VGG16 achieved an average detection accuracy of 99.00%, which was 23.89%, 1.21%, 0.75%, 0.62%, and 8.46% higher than SSD, YOLOv5, YOLOv7, CenterNet, and RetinaNet models. Therefore, it was demonstrated that the Faster R-CNN model based on the hyperspectral feature image fusion dataset proposed in this paper was feasible for rapid evaluation of wheat FHB resistance. This study provided an important detection method for ensuring wheat food security.

Keywords: *Fusarium* head blight, resistance evaluation, hyperspectral feature band image fusion, deep learning, Faster R-CNN
DOI: 10.25165/j.ijabe.20241702.8269

Citation: Liang K, Ren Z Z, Song J P, Yuan R, Zhang Q. Wheat FHB resistance assessment using hyperspectral feature band image fusion and deep learning. Int J Agric & Biol Eng, 2024; 17(2): 240–249.

1 Introduction

Fusarium head blight (FHB), mainly caused by *Fusarium graminearum*, is a global disease[1,2]. The normal physiological functions of wheat are disturbed by fungi after infection, and the internal structure and external shape of wheat also change[3,4]. The disease can cause significant yield and quality reductions in the form of atrophy, weight reduction, and discoloration[5]. *Fusarium* produces secondary metabolites, especially deoxynivalenol (DON), which can interfere with normal eukaryotic cell function by inhibiting protein synthesis and can cause great harm to humans and animals[6-9]. It has been well documented that DON may cause toxemia, which may contribute to the emergence of cancer[10,11]. Recently, chemical, biological, and agricultural approaches have been implemented to control FHB in wheat[12]. Chemical control approaches primarily reduce the amount of *Fusarium* contamination through the application of chemicals, but issues of increased resistance to *Fusarium* and environmental contamination do exist[1]. Biological control approaches mainly control *Fusarium* contamination through microorganisms, but the effectiveness of microorganisms against the disease varies under different conditions. Agricultural control approaches change farming methods and select resistant varieties, as changing the farming methods could reduce the survival of *Fusarium* to some extent with respect to the environment. However, *Fusarium* outbreaks would be uncontrollable when the weather is hot or rainy[13]. Therefore, research on and development of wheat varieties with FHB resistance is essential[14]. Wheat FHB resistance is composed of five types: resistance to the original fungal infection (type I), resistance to fungal diffusion in the spike (type II), low accumulation of DON (type III), low *Fusarium*-damaged kernel (type IV), and other FHB resistance types (type V)[15]. However, only type II resistance is stable and can be easily used to evaluate and select FHB-resistant wheat varieties, which is conducive to the fundamental control of FHB[16]. Currently, manual methods are mainly adopted for resistance identification, which is slow and restricts the development cycle of resistant varieties. In addition, manual evaluation approaches are not objective or accurate. Therefore, proposing a new rapid identification technique is of great importance[17].

In recent years, some new technologies for phenotypic analysis have been rapidly and accurately developed to improve plant yield and quality[18,19]. Plant phenotype detection has developed rapidly, especially hyperspectral imaging (HSI) technology, which

Received date: 2023-03-27 **Accepted date:** 2024-04-22
Biographies: Zhizhou Ren, MS, research interest: non-destructive testing of agricultural products, Email: rzz2031686205@163.com; **Jinpeng Song**, MS, research interest: non-destructive testing of agricultural products, Email: songjinpeng1214@163.com; **Rui Yuan**, MS, research interest: non-destructive testing of agricultural products, Email: yuanmarco@163.com; **Qun Zhang**, UG, research interest: non-destructive testing of agricultural products, Email: m18646094302@163.com.
***Corresponding author: Kun Liang**, Professor, research interest: non-destructive testing of agricultural products. No.40, Dian Jiangtai Road, Pukou, Nanjing 210031, Jiangsu province, China. Tel: +86-18751882013. E-mail: kliang@njau.edu.cn.

integrates imaging technology and spectroscopy technology and provides methods for the detection of plants and the identification of plant diseases[20,21]. In the study of wheat ear FHB resistance based on HSI technology, Bauriegel et al.[5] identified FHB infection of wheat ears by a HSI system based on wavelength ranges of 500-533 nm, 560-675 nm, 682-733 nm and 927-931 nm, with an identification accuracy of 87%. Vincke et al.[22] developed a method for evaluating the degree of FHB infection in winter wheat spikes using near-infrared hyperspectral imaging (NIR-HSI) technology in the wavelength range of 900-1700 nm, with an accuracy of 99.4% in detecting the degree of wheat spike infection. Zhang et al.[23] collected the spectral and image information of wheat spikelets through a hyperspectral microscopic imaging system, obtained the FHB classification index (FCI) by combining a single wavelength and spectral index, and identified FHB-infected spikelets with a classification accuracy of 89.8%. The selection of the feature wavelength could effectively reduce the high-dimensional data of the hyperspectral images, and the extraction of the sensitive band could effectively improve the detection accuracy. In the research of hyperspectral feature wavelength fusion images, Zhang et al.[24] extracted central wavelengths of 660, 560 and 480 nm based on hyperspectral images, and three wavelengths were combined into the RGB image and converted to the YDbDr color space for extracting the color features to investigate the different severities of FHB in wheat ears.

However, it took considerable time to process large image datasets, such as hyperspectral images[25]. Therefore, it was necessary to investigate faster methods to improve data processing efficiency. Deep learning methods are capable of extracting more valuable information rapidly from a large number of large image datasets[26]. With the development of deep learning, Faster R-CNN[27] has been widely used in plant phenotypes due to its great performance. Quan et al.[28] used the Faster R-CNN model based on the pretrained network VGG19 to quickly and accurately detect maize seedlings in different field environments and different growth stages, and the model accuracy was over 97.71%. Ozguven and Adem[29] proposed a Faster R-CNN model that automatically detected leaf spot in sugar beet leaves. The overall correct classification rate of the model reached 95.48%. Wu et al.[30] used the transfer learning method combined with the Faster R-CNN algorithm to count wheat berries and calculate the relevant traits in various situations, with an average accuracy of 91%. Simultaneously, multiple recent studies have applied deep learning methods in plant phenotype analysis based on hyperspectral images. Several researchers have focused on high-dimension hyperspectral feature extraction with deep learning networks. Yang et al.[31] proposed a method combining hyperspectral imaging and convolutional neural networks based on the Spectral and Spatial Attention Mechanism (Spl-Spal-At) module, aimed at identifying different types of maize kernels. Li et al.[32] proposed a technique based on convolutional neural networks (CNN) and hyperspectral imaging (HSI) for rapid, non-destructive, and accurate identification of copper (Cu) pollution levels in apple rootstocks. The results showed that the accuracy of the CNN model was 99.6%, which was better than traditional linear and nonlinear models. Moreover, recent studies of various deep learning algorithms have emphasized the feasibility of extracting deep spectral features from hyperspectral images to develop deep learning models for plant phenotype prediction[33]. Yu et al.[34] acquired hyperspectral images of oilseed rape leaves and used SAE to extract deep spectral features from pixel spectral data for nonlinear data dimension reduction. A deep learning regression model composed of SAE and a fully connected neural network (FNN) was established to predict the nitrogen content in rape leaves with a prediction accuracy of 90.3%. Jin et al.[35] used deep learning models based on hyperspectral image pixels to identify FHB-infected regions in wheat ears. The high-dimensional hyperspectral data were reduced by principal component analysis (PCA) and input into the deep convolutional recurrent neural network (DCRNN) for extracting features that achieved high accuracy in classifying healthy and diseased wheat ears. Based on the mean spectral reflectance and 1-D CNN, Rehman et al.[36] implemented an end-to-end deep learning model to predict the relative water content (RWC) of plants with a prediction accuracy of 95.6%.

Therefore, in order to explore the effect of feature wavelength fusion methods, the evaluation method of wheat ear FHB resistance was optimized based on hyperspectral feature band image fusion and Faster R-CNN. In this study, the spectral data from 400-1000 nm were preprocessed by the multiple scattering correction (MSC) algorithm. Three feature wavelengths were selected by analyzing the X-loading weights (XLW) according to the absolute value of the peaks and troughs in different principal component (PC) load coefficient curves. Then, the different fusion methods of the three feature wavelengths were explored with different weight coefficients. Faster R-CNN was trained on the fusion and RGB datasets with VGG16, AlexNet, ZFNet, and ResNet-50 networks separately. Then, the other detection models SSD, YOLOv5, YOLOv7, CenterNet, and RetinaNet were used to compare with the Faster R-CNN model. The best deep learning model was selected to evaluate the FHB resistance in wheat ears. This work provides a primer for future studies on the genetic improvement of wheat with high FHB resistance.

2 Materials and methods

2.1 Sample preparation

The experimental materials were winter wheat harvested from the test rotation fields in Yun Sheng Road, Liu He District, Nanjing, China (32°42.33 ′N, 118°64.02 ′E). The moderate sensitivity of wheat samples to the *Fusarium* variety (cv. “YangMai 23”) were provided by Jiangsu Academy of Agricultural Sciences[37]. The wheat ears were infected by single-floret inoculation to evaluate resistance to fungal diffusion in the spike (type II). The spores of *Fusarium* suspensions for inoculation were provided by Jiangsu Academy of Agricultural Sciences. The wheat ovary was the site for spore suspension injection[38]. Generally, the disease levels of the wheat ears fell within a certain range after inoculation with the same resistant variety. Thus, to increase the number of samples with different disease levels and improve the robustness of the models, the inoculated wheat ears were divided into three groups and inoculated separately with three solutions. In the first group of 50 wheat ears, each wheat ear was injected once with 0.02 mL of spore suspension (1×10^6 CFU/mL). In the second and third groups of 50 wheat ears, each wheat ear was injected twice and three times, respectively (Figure 1). Another fifty healthy samples without injection were selected as controls. After injection, the glume and bag were closed to maintain moisture, and the bag was removed after 3 d to allow natural disease onset. After being injected for 21 d, wheat ears with different injection levels were cut according to the label and brought back to the laboratory. In total, 133 wheat ears were selected with the same proportion of infected spikelets on both sides to facilitate and simplify the evaluation of the resistance of a single wheat ear.

Figure 1 Diagram of the three wheat ear inoculation solutions

2.2 Resistance identification by an artificial method

Wheat FHB resistance (fungal diffusion in the spike (type II)) has traditionally been evaluated by artificial methods[37,39]. The process of resistance evaluation by the artificial method is as follows. First, the percentage of diseased spikelets is determined by the ratio of the number of diseased spikelets to the number of spikelets in the whole wheat ear. The number of diseased spikelets and the number of spikelets in the whole wheat ear were manually counted, and the severity level of each wheat ear was determined according to Table 1[40]. Second, FHB resistance (type II) was evaluated according to the average severity (AS) of each wheat ear in Table 2[40]. The AS is defined in Equation (1).

$$AS = \frac{\sum_{i=0}^{4} G_i N_i}{\sum_{i=0}^{4} N_i} \quad (1)$$

where, G_i is the ith severity level, and N_i is the sample number of the ith severity level.

Table 1 Severe wheat FHB classification under different wheat ear inoculation conditions

Severity level	Symptom description
0	Inoculated spikelet with no visible disease symptom (0%)
1	Diseased spikelet under a quarter of total spikelet (0-25%)
2	Diseased spikelet accounts for 1/4-1/2 of total spiketlet (25-50%)
3	Diseased spikelet accounts for 1/2-3/4 of total spiketlet (50-75%)
4	Diseased spikelet more than 3/4 of total spikelet (75%-100%)

Table 2 Evaluation of resistance to wheat FHB under different wheat ear inoculation conditions

Average severity (AS)	Resistance evaluation
0.00<AS≤0.25	High resistance (HR)
0.25<AS≤0.50	Moderately resistance (MR)
0.50<AS≤0.75	Moderately susceptible (MS)
0.75<AS≤1.00	High susceptible (HS)

2.3 Hyperspectral imaging system

As shown in Figure 2, hyperspectral images of wheat ears were obtained using a line-scan HSI system that primarily consisted of the following sections: a spectral imaging unit, a lighting unit, a sliding unit and a controlling unit. The spectral unit mainly includes an imaging spectrograph (ImSpectorV10E-1621, Spectral Imaging Ltd., Finland) with a spectral range covering 616 bands from 358-1021 nm, a 1555×1200 pixel resolution CCD camera (GEV-B1621M-TC000, Imperx, USA), and a standard C-mount 2.4/23 mm zoom lens (Schneider Optics Inc., Germany). The lighting unit consisted of a 21 V/150 W quartz tungsten-halogen light source (Illumination, USA), and the samples were illuminated by 2 separated linear light guides. The sliding unit included a sample holder and a horizontal stepping-motor-driven slider. The HSI system was controlled by computer software.

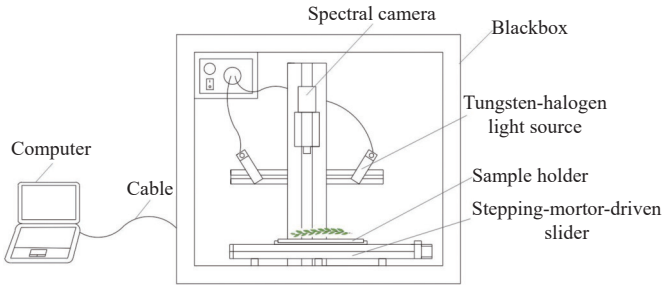


Figure 2 Hyperspectral imaging system

To maintain the integrity and accuracy of the hyperspectral images, it was necessary to set the system parameters at fixed values before scanning the samples. The HSI system was preheated for 15 minutes before acquiring the images. The regulated height between the wheat surface and lens was 330 mm, and the slider moving speed was 3.5 mm/s. The exposure time was 53 s, which could ensure hyperspectral image clearance and accuracy. To reduce the influence of illumination and different physical configurations of cameras and imaging systems, the images were immune to dark-current effects, and the original hyperspectral images (I_s) were corrected[41], as shown in Equation (2).

$$I_c = \frac{I_s - I_d}{I_w - I_d} \quad (2)$$

where, the I_c is the calibrated hyperspectral image; I_s is the hyperspectral wheat sample image, I_d is the dark reference image with 0 reflectance, and I_w is the white reference image with 1 reflectance. The calibrated hyperspectral images formed the following spectral collection. After the exclusion of the wheat ears with different severity levels over both surfaces, 133 wheat ears were selected and scanned by the HSI system over both surfaces. The number of wheat ear hyperspectral images at different severity levels are listed in Table 3. Therefore, the number of samples for each severity level was half of the number of collected hyperspectral images.

Table 3 Accuracy test of the technical specifications for FHB in wheat ears

Severity level	Level 0	Level 1	Level 2	Level 3	Level 4
Diseased spikelet rate	0%	0%-25%	25%-50%	50%-75%	75%-100%
Number of wheat ears images	38	52	58	64	54

2.4 Spectral data analysis

2.4.1 Spectral data extraction and preprocessing

The region of interest (ROI) from calibrated hyperspectral images was selected for further spectral processing by Environment for Visualizing Images (ENVI) software (version 5.1, Research Systems Inc., Boulder, CO, USA). The ROI of each wheat ear was manually selected from the calibrated hyperspectral images. The average spectrum was obtained through the calculated spectral data of every extracted pixel, and the average spectrum was considered the spectral information. Because of the noise disturbance at the beginning and end of the spectral range, spectrum data analysis was performed in the 400-1000 nm wavelength range with 555 bands[42,43]. The spectral data over 400-1000 nm were preprocessed by the MSC algorithm to remove noise and enhance the spectral

features in MATLAB R2014a (MathWorks Inc., Natick, MA)[44].

2.4.2 Feature wavelengths acquisition

XLW analysis is an effective feature variable selection method. The principle is to select the most relevant primary components of different variables that contain their distinctive features and then extract the eigenvector with the largest eigenvalue as the weight and bias parameters to derive the relationship between the extracted primary components and the variables and then screen out the feature variables[45]. XLW analysis was carried out on the spectral information of the feature wavelengths. First, the spectral data of the diseased and healthy areas in the hyperspectral images of wheat spikelets were selected. Second, two kinds of spectral data were input to the PCA algorithm, PCA was used to select the appropriate number of PCs, and the individual and cumulative explained variances of the PCs were calculated in Figure 3. The first three PCs explained 97.78% of the cumulative variance and were selected as the PCs. Under each principal component variable, the weight coefficient corresponding to each wavelength could be calculated. The larger the absolute value of the weight coefficient is, the greater the influence of the wavelength[46]. Therefore, the wavelength with the larger absolute value of the weight coefficient in each PC was selected as the feature wavelength (Figure 4). The selected feature wavelengths were 682 nm, 714 nm and 553 nm.

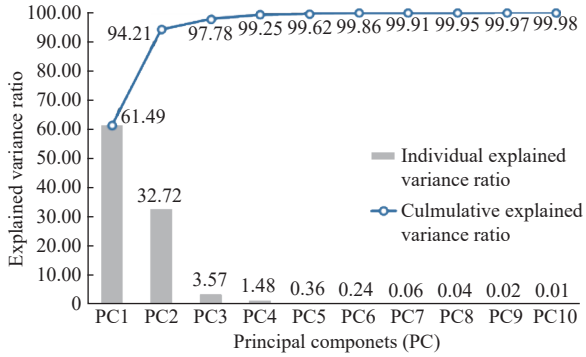


Figure 3 Individual and cumulative explained variance ratios for the first ten PCs

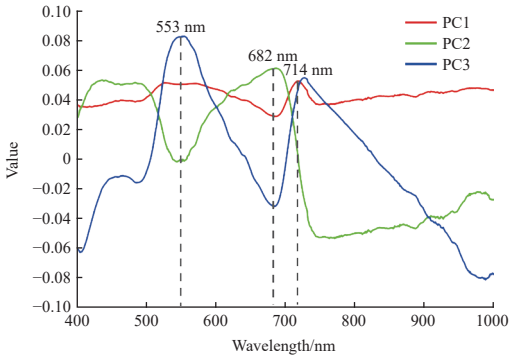


Figure 4 Load coefficient curves of the first three PCs

2.5 Deep learning methods

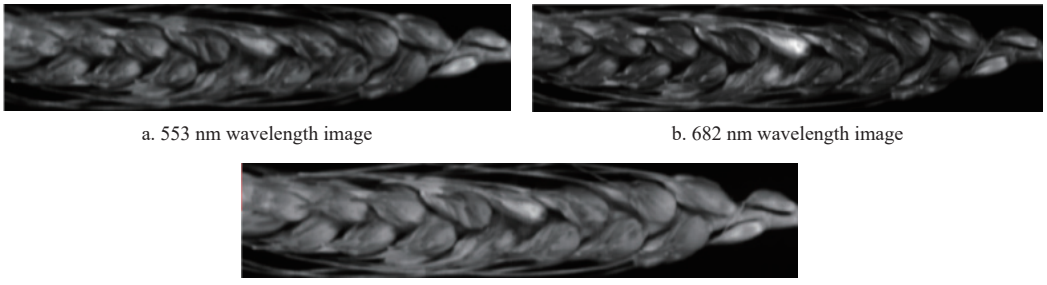
2.5.1 Algorithm operating environment

The graphics processing unit (GPU) was a GTX2080Ti NVIDA (11 GB graphic memory), while the central processing unit (CPU) was an Intel Core i7-9700 (16 GB memory), the Ubuntu operating system (version 16.04) and Python language (version 2.7). Based on the Caffe frame, computed unified device architecture (CUDA) version 9.0.176 and version 7.3.1 were used to improve the use rate of the graphic card. The Caffe framework was widely used, and the overall code was simple to write. It had comprehensive functions and was widely used in target detection.

2.5.2 Datasets preprocessing

After the XLW analysis, three feature wavelengths of the spectral images were screened, and different fusion methods of the three feature wavelengths were investigated. Three wavelengths of 663 nm, 555 nm and 450 nm were selected as the Fusion Image-RGB dataset, and its performance was compared with the different fusion datasets of the three feature wavelengths. Figure 5 shows the three feature wavelengths selected by XLW analysis. The red head feature of FHB wheat ears was distinguished at 682 nm wavelength, which was the red absorption region of visible light. Therefore, according to naked eye observations, the images at a wavelength of 682 nm had the richest wheat ear information and the most significant difference between diseased and normal wheat spikelets. Moreover, different weight coefficients for each feature wavelength image were added as a new approach to explore the effect of fusion image datasets. The element-level weighted fusion of hyperspectral images at three selected feature wavelengths was performed to obtain a feature-fused image. Three fusion images (Fusion Image-1, Fusion Image-2 and Fusion Image-3) with different feature wavelengths and weight coefficients are listed in Table 4. For Fusion Image-2 and Fusion Image-3, the weight coefficient of the feature wavelength at 682 nm was increased to retain more information clearly distinguishing between diseased and normal wheat ears by naked eye observation. The information from all three feature bands was presented in Fusion Image-1, which included more comprehensive information. The Fusion Image-2 highlighted the 682-nanometer red band while filtering the 714-nanometer red band to highlight the most prominent features of the FHB spike. Fusion Image-3 also highlighted the 682-nanometer red band, eliminating the 553-nanometer green band, which focused entirely on the red band about the FHB feature.

The flow chart of detection by Faster R-CNN models based on HSI data is shown in Figure 6. The MSC method was used for spectral pretreatment, and feature wavelengths were selected by XLW algorithm. After the feature wavelength fusion datasets were created, the performance of three feature wavelength fusion datasets and the RGB dataset were compared in different Faster R-CNN models along with four different pretrained networks, AlexNet,



a. 553 nm wavelength image

b. 682 nm wavelength image

c. 714 nm wavelength image

Figure 5 Feature wavelength images of wheat ears

VGG16, ZFNet, ResNet-50[47-50] models. To reduce overfitting and improve precision in the Faster R-CNN model, four datasets were expanded by brightness, color, contrast and rotation enhancement[28], and then the number of images in each dataset was increased to 1330 images. The labeled images in each dataset were divided into training (931 images, 70%) and testing (399 images, 30%) sets, and the images of each level by manual method in each dataset were randomly placed into the training and test sets at a ratio of 7:3.

2.5.3 Faster R-CNN model

As a classic object detection model, the Faster R-CNN adopted a two-stage detection approach. This approach integrated the Region Proposal Network (RPN) to effectively generate a series of object candidate regions. Subsequently, the targets in the candidate regions were fed into the detection network for further classification and localization. In contrast, one-stage networks like SSD and YOLO directly completed both region proposal and target detection in a single network, which had faster detection speed but may lost detection accuracy. In this study, the Faster R-CNN was used as the detection model to obtain higher FHB wheat ear detection accuracy[28].

Table 4 Four fusion methods of the image datasets

Datasets	Input image1		Input image2		Input image3	
	Feature wavelength	Weight coefficient	Feature wavelength	Weight coefficient	Feature wavelength	Weight coefficient
Fusion Image-1	682nm	1/3	714nm	1/3	553nm	1/3
Fusion Image-2	682nm	2/3	-	-	553nm	1/3
Fusion Image-3	682nm	2/3	-	-	714nm	1/3
Fusion Image-RGB	663nm	1/3	555nm	1/3	450nm	1/3

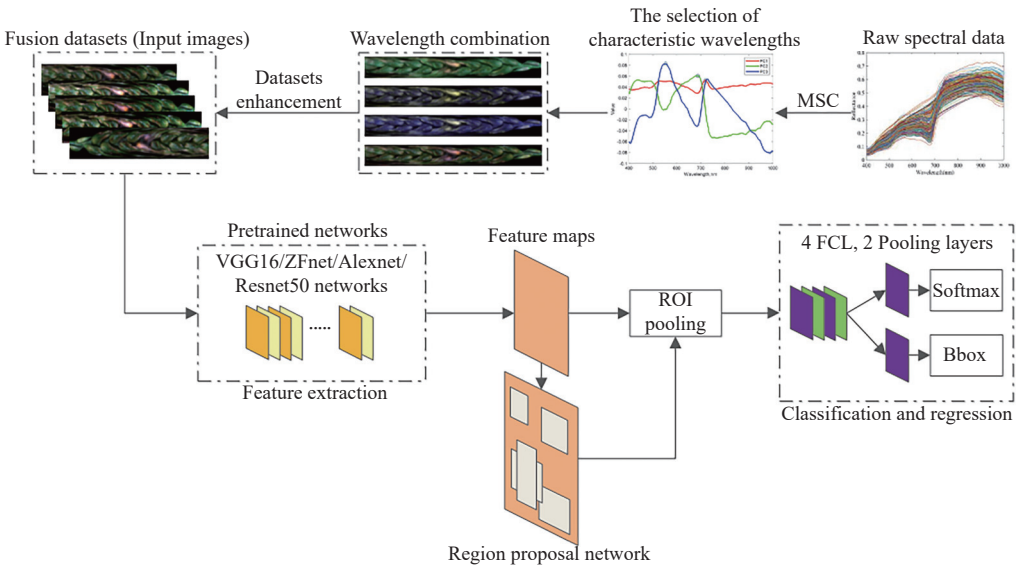


Figure 6 Flow chart of detection by Faster R-CNN models based on HSI data

The Faster R-CNN object detection algorithm was combined with AlexNet, VGG16, ZFNet and ResNet-50 to form four models. There were three parameters adjusted in the experiment, namely, learning rate, iterations and batch size. The learning rate directly controlled the magnitude of the network gradient update during training, and the learning rate affected the effective tolerance capability of the model. If the learning rate is set too small, the convergence process will become extremely slow. If the learning rate is set too large, the gradient will oscillate back and forth near the minimum value and may even fail to converge. Iterations influenced the restraint of the loss curve and model accuracy. If the iterations are set too small, the loss curve will not be constrained, and the network will not select complete features. Moreover, when the number of iterations reached a certain value, the increase in iterations had no optimization effect on the model[51-53]. To share features between RPN and Faster R-CNN, stochastic gradient descent (SGD) was used as a training optimization method in Faster R-CNN, and batch size influenced the performance of SGD. In the process of training, choosing too few training samples leads to a longer runtime, and the loss function oscillates and fails to converge. A large number of sample calculations made too many weight update changes, which caused the model to fall into a local optimal point, and the high memory occupancy rate caused the training to fail[4,54]. In this study, the loss function oscillated and did not converge when the initial learning rate was set to 0.1. After adjusting the learning rate and the number of iterations many times, the loss function stabilized when the learning rate was set to 0.001 and the number of iterations was set to 20,000. According to the limitation of the graphics card and the size of the data sets, the batch size was set to 64, 128, 256, and 512. The performance was compared with other classical target detection models, such as SSD, YOLOv5, YOLOv7, CenterNet, and RetinaNet. Then, the optimized model was used to evaluate the wheat ear FHB resistance.

2.6 Model evaluation

The inference time, precision (P), recall (R), average precision (AP) and mAP were used to precisely evaluate the performance of the model. The inference time showed the time spent on model inference. The precision-recall (PR) curve showed variations in the precision and recall values as the sample size changed[55]. A good classifier retains a high precision as the recall increases, and a poor classifier loses much precision as the recall increases. The P and R were defined in Equation (3) and Equation (4), respectively.

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

$$R = \frac{TP}{TP + FN} \quad (4)$$

where, TP is the number of correctly detected diseased objects (true positives), FP is the number of falsely detected diseased objects (false positives), and FN is the number of diseased samples incorrectly classified as normal.

AP was defined in Equation (5) as the area under the PR curve and indicated the detection accuracy of diseased and normal wheat samples, and mAP was defined in Equation (6) as the average of all the AP values.

$$\mathrm{AP} = \int_0^1 P(R)dR \tag{5}$$

$$\mathrm{mAP} = \sum_{n=1}^{K} \frac{\mathrm{AP}(n)}{K} \tag{6}$$

2.7 Resistance evaluation

For each wheat ear in a band fusion image, diseased spikelets and healthy spikelets were labeled ‘diseased’ and ‘normal’ in the outputs of the Faster R-CNN models, respectively. The prediction result of each wheat ear is presented in a visualization image. The total numbers of ‘diseased’ and ‘normal’ labels were calculated by traversing the visualization image. The diseased spikelet rate was defined in Equation (7).

$$L = \frac{S}{S+M} \tag{7}$$

where, S is the total number of spikelets labeled ‘diseased’, M is the total number of spikelets labeled ‘normal’, and L is the diseased spikelet rate of a single wheat ear image. Then, the resistance evaluation process based on deep learning is summarized in the algorithm given below:

Step 1：Determine the severity level of the image.

L is the diseased spikelet rate of a single wheat ear image, as defined above. D is assumed to be the severity level of each wheat ear image, as shown in Equation (8)[40].

$$D = \begin{cases} 0, & L = 0\% \\ 1, & 0\% < L \leqslant 25\% \\ 2, & 25\% < L \leqslant 50\% \\ 3, & 50\% < L \leqslant 75\% \\ 4, & 75\% < L \leqslant 100\% \end{cases} \tag{8}$$

Step 2: Calculate the number of wheat ear images at each level.

Step 3: Determine the AS of the wheat ear samples.

P is assumed to be the AS of each wheat ear and n_i is the number of wheat ear images at each severity level from 0 to 4, as shown in Equation (9)[40].

$$P = \frac{\sum_{i=0}^{4} \frac{n_i}{2} \times D_i}{\sum_{i=0}^{4} \frac{n_i}{2}} \tag{9}$$

Step 4: Evaluate the resistance

The resistance of the samples that are predicted by the deep learning method according to Table 2 is evaluated. The sample resistance was defined in Equation (10)[40].

$$\text{Resistance} = \begin{cases} \text{HR}, & 0 < P < 2 \\ \text{MR}, & 2 \leqslant P < 3 \\ \text{MS}, & 3 \leqslant P < 3.5 \\ \text{HS}, & P \geqslant 3.5 \end{cases} \tag{10}$$

3 Results and discussion

3.1 Raw spectral reflectance analysis

After selecting the ROI from the calibrated hyperspectral images, the raw spectral range from 358-1021 nm was used to obtain the average wheat ear spectrum. Spectral data ranging from 400 to 1000 nm were selected for spectral preprocessing with the MSC algorithm. The algorithm completed spectral data preprocessing by eliminating the spectral differences caused by different scattering levels and enhancing the spectral absorption information related to component content[44]. The preprocessed spectral data were used for subsequent feature wavelength selection.

3.2 Hyperspectral feature band image fusion

The three fusion images with feature wavelengths and Fusion Image-RGB were shown in Figure 7. The 553 nm, 682 nm, and 714 nm wavelength fused images（Fusion Image-1, Fusion Image-2, and Fusion Image-3） were all exhibited more distinctive color and texture characteristics compared to RGB images. The characteristic of spike red head of FHB wheat was more obvious in the fusion images. The fusion images in Figure 7b and Figure 7c highlighted the red light wavelength information at 682 nm, making the red head feature of the FHB ear, but it was also accompanied by the loss of some other key band information. However, Figure 7a contained hyperspectral image information in three bands of 553 nm, 682 nm, and 714 nm, and this fused image exhibited a more distinct FHB spike signature, at the same time, the characteristics of normal ears become more distinct. This fusion method effectively improved the ability to distinguish between diseased and normal ears.

a. Fusion Image-1

b. Fusion Image-2

c. Fusion Image-3

d. Fusion Image-RGB

Figure 7 Fusion images of wheat ear

3.3 Faster R-CNN model

3.3.1 Detection results of the Faster R-CNN model based on the pretraining network

Faster R-CNN models were trained with pretrained AlexNet, VGG16, ZFNet, and ResNet-50 networks. The trained datasets were Fusion Image-1, Fusion Image-2, Fusion Image-3, and Fusion Image-RGB,respectively. The overall trend in the mAP of the four Faster R-CNN models as the number of iterations increased is shown in Figure 8. The mAP results of the four networks at 20 000 iterations are listed in Table 5. The overall results of Faster R-CNN with VGG16 and ResNet-50 for the four fusion image datasets were better than those of the Faster R-CNN models with AlexNet and

ZFNet; the highest mAP of the VGG16 and ResNet-50 models trained on Fusion Image-1 was 98.8%, and the lowest mAP of the model trained on Fusion Image-RGB was 97.7%. The reason might be because the Faster R-CNN model with VGG16 and ResNet-50 had better feature extraction networks than AlexNet and ZFNet. Moreover, the parameters of the convolution in the feature extraction network affect the memory, speed and performance of the detector[56]. The runtimes of the training and prediction phases for the four networks are shown in Figure 9. Considering the complexity of the network structure, Faster R-CNN with ZFNet had the shortest training time, which was approximately 52 min. The training time for Faster R-CNN with AlexNet was approximately 1.6 h, which was longer than that of Faster R-CNN with ZFNet. Faster R-CNN with ResNet-50 had the longest training time of approximately 6 h. Compared to ResNet-50, the Faster R-CNN with VGG16 has a shorter training time of approximately 3 h. In the prediction time of pretrained Faster R-CNN networks, Faster R-CNN with AlexNet and ZFNet had shorter prediction times of 0.014 s per image and 0.016 s per image, respectively. Faster R-CNN with VGG16 had a relatively short prediction time of 0.044 s per image. Faster R-CNN with ResNet-50 had the longest prediction time of 0.182 s per image. The better performance of VGG-16 feature extraction may be due to the relatively deep network structure and efficient design. Compared to AlexNet and ZFNet, VGG-16 had a deeper network that better captured the complex features and abstract representations in images. VGG-16 used a smaller 3×3 convolution kernel to capture these local subtle features well, and enable parameter sharing, thus reducing model complexity and computational cost. Overall, Faster R-CNN with VGG16 performed best, and the parameters of this network was optimized and trained with different datasets to obtain the optimized deep learning model.

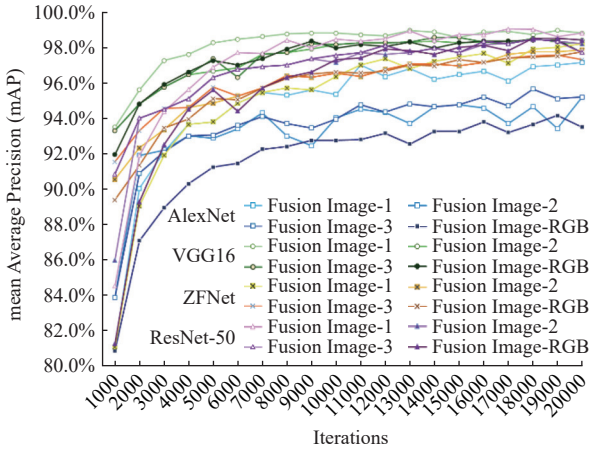


Figure 8 Detection results of the four pretrained Faster R-CNN models with increasing numbers of iterations

Table 5 The mAP of four Faster R-CNN networks at 20,000 iterations

Faster R-CNN network	Datasets			
	Fusion Image-1	Fusion Image-2	Fusion Image-3	Fusion Image-RGB
Faster R-CNN with VGG16	98.80%	98.40%	98.20%	97.70%
Faster R-CNN with AlexNet	97.15%	94.50%	95.20%	93.50%
Faster R-CNN with ZFNet	98.20%	97.90%	97.30%	97.80%
Faster R-CNN with ResNet-50	98.80%	98.20%	97.70%	98.40%

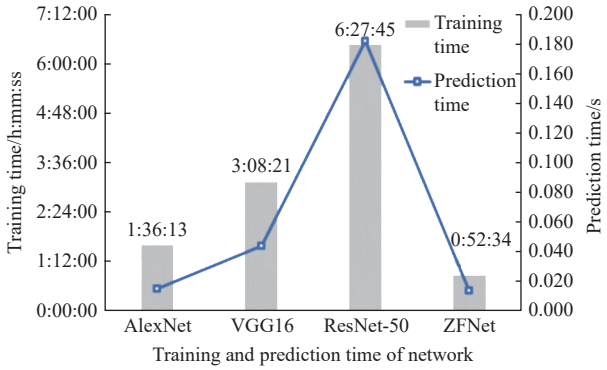


Figure 9 Training and prediction time of the four pretrained Faster R-CNN models

In Faster R-CNNs, the batch size determines the direction of gradient descent. The memory utilization rate and parallelization efficiency of large matrix multiplication can be improved by reasonably increasing the batch size to achieve the optimal convergence accuracy. The number of iterations required to run an epoch was decreased to accelerate the processing of the same amount of data[4]. The initial value of the batch size confirmed in the experiment was 64, and the dereferencing of the batch size was 64, 128, 256 and 512.

As shown in Figure 10 and Table 6, after training Faster R-CNN with VGG16 with different parameters using the Fusion Image-1 dataset, the mAP of the model was higher than that of the models trained with the other datasets. When training the model with Fusion Image-1, the mAP of the model reached 99.0% when the batch size was 128 or 512. When training the models with the Fusion Image-2, Fusion Image-3 and Fusion Image-RGB datasets, a certain increase in the batch size was effective in improving the mAP of the models. The Fusion Image-2 and Fusion Image-3 datasets yielded the highest mAPs of 98.6% and 98.5%, respectively, when the batch size was 256. The Fusion Image-RGB

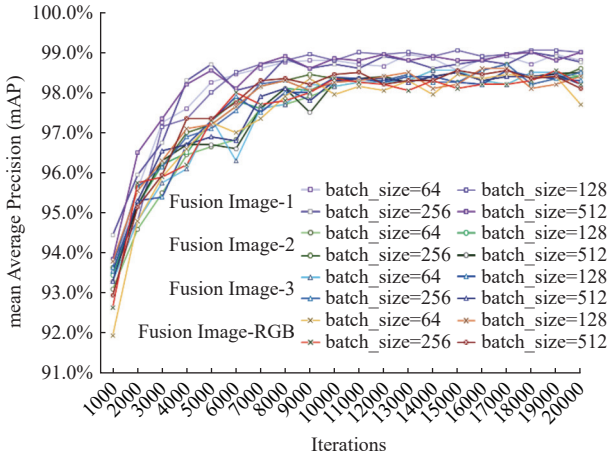


Figure 10 Detection results of Faster R-CNN with VGG16 under the different batch sizes and datasets

Table 6 The mAP of Faster R-CNN with VGG16 under the different batch sizes and datasets

Datasets	Batch size			
	64	128	256	512
Fusion Image-1	98.80%	99.00%	98.80%	99.00%
Fusion Image-2	98.40%	98.40%	98.60%	98.50%
Fusion Image-3	98.20%	98.30%	98.50%	98.20%
Fusion Image-RGB	97.70%	98.40%	98.20%	98.10%

dataset yielded the highest mAP of 98.4% when the batch size was 128. The mAP of the models trained on the Fusion Image-1, Fusion Image-2 and Fusion Image-3 datasets was better than that trained on the Fusion Image-RGB dataset, showing that the selection of the feature wavelengths could effectively facilitate the extraction of the spectral information of wheat ears while reducing the dimensionality of the hyperspectral images. Although there was relatively similar accuracy in the detection of terminal FHB wheat ears in the hyperspectral feature wavelength fusion image dataset and RGB image dataset models in this study, the experiment was able to provide a more effective method for the early detection of FHB wheat ears. In addition, the Fusion Image-1 dataset, with more comprehensive information on feature wavelengths, yielded higher mAP than the models trained by the Fusion Image-2 and Fusion Image-3 datasets. Therefore, we selected Faster R-CNN with VGG16, set the batch size to 128, and trained the best deep learning model with the Fusion Image-1 dataset for subsequent experiment.

3.3.2 Performance comparison

To compare the advantages of Faster R-CNN model in FHB ear detection, the average detection accuracy of SSD, Yolov5, Yolov7, CenterNet and RetinaNet algorithms were compared for the Fusion Image-1 dataset, and the results were listed in Table 7. The mAP of Faster R-CNN model was 99.00%, which was 23.89% higher than the mAP of the SSD model. In the process of feature extraction, SSD algorithm may gradually pay attention to the whole semantic information of the image, but neglect the local details, especially the small grain size. In contrast, the Faster Faster R-CNN model has higher detection accuracy because image features were extracted and recognized in the RPN region of Faster Faster R-CNN, which generated candidate boxes focused on ear size. However, this also led to a large increase in the speed of model reasoning. The Yolo, SSD and RetinaNet models saved the run time in ears of wheat detection, but these algorithms are part of a one-stage network and are less accurate than the two-stage network Faster R-CNN. CenterNet algorithm had the second higher recognition accuracy, because it could predict the center point of the object directly on the input image, effectively solved the problem of dense object detection, and was also suitable for wheat ear disease detection. Thus, the advantage of the Faster R-CNN model in detection effect was the most accurate detection effect, but the recognition speed was lower, which would be further improved in future research.

Table 7 Different algorithms detect results of the Fusion Image-1 dataset

Models	mAP/%	Inference time/ms
SSD	75.11	7.85
YOLOv5	97.79	11.67
YOLOv7	98.25	17.29
CenterNet	98.38	9.07
RetinaNet	90.54	23.94
Faster R-CNN with VGG16	99.00	81.44

To verify the detection effect of the model, the detection results of the six models with the same data are shown in Figure 11. As can be seen from the figure, SSD had the lowest recognition accuracy, while the other five models could correctly detect and locate FHB wheat ears. The Faster R-CNN model had the highest confidence in detecting the FHB spike in the graph, and the confidence level of four FHB ears was higher than 0.95. We can also clearly see that the healthy ear is in the upper right corner, which only the Faster R-CNN model can detect, but none of the other models could. The experiment further verified that the model has higher detection accuracy, and the method can accurately identify FHB wheat ears and healthy wheat ears.

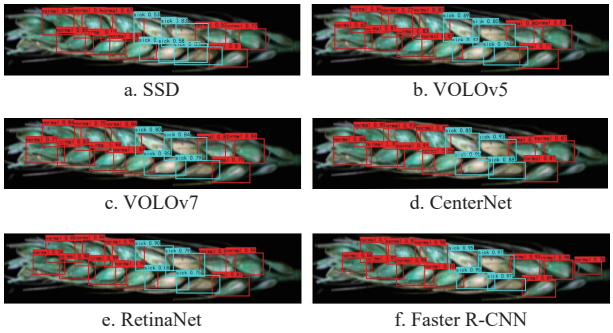


Figure 11 Wheat ear detection results with different models

3.3.3 Detection results of the Faster R-CNN model

The accuracy of the deep learning method was identified by comparing the disease classification results of the Faster R-CNN model with manual classification. Artificial results of manual classification was according to Table 1 and Table 2[40]. The accuracy rates of the deep learning method to predict severity levels of 0, 1, 2, 3 and 4 were 100%, 100%, 93.5%, 93.8% and 100%, respectively (Table 8). In Table 8, the manual classification and algorithm classification results showed the same performance for all levels except at Level 2 and Level 3. It indicated that the algorithm was able to accurately identify the health of the wheat ears in most cases, which was the further evidence of the effectiveness of the algorithm. However, differences in the performance of the two methods emerged when the wheat ears were in a moderately diseased state (levels 2 and 3). This may be because, in these cases, the phenotypic characteristics of the diseased ear are not obvious enough. The evaluation results of wheat FHB by the manual and deep learning methods are listed in Table 9. Both methods ultimately determined the sample resistance of FHB as moderately resistant (MR), and the relative error in the AS was 0.9%. Therefore, it was evident that the deep learning method was a viable alternative to manual resistance assessment, and the assessment of wheat FHB resistance (type II) could be optimized for rapid, objective and accurate pattern evaluation.

Table 8 Accuracy rates of the algorithm incidence judgments

Severity level	Level 0	Level 1	Level 2	Level 3	Level 4
Diseased Kernel Rate/%	0%	0%-25%	25%-50%	50%-75%	75%-100%
Artificial results	38	52	58	64	54
Algorithm results	38	52	62	60	54
Accuracy/%	100%	100%	93.50%	93.80%	100%

Table 9 Evaluation results of wheat FHB compared between deep learning and artificial methods

Method type	Average severity	Resistance evaluation
Artificial	2.17	MR
Deep leaning	2.15	MR

4 Conclusions

In this study, the evaluation method of wheat ear FHB resistance was optimized based on the hyperspectral feature band image fusion and Faster R-CNN. Three feature wavelengths were selected by the XLW algorithm and different fusion methods for the three feature wavelengths with different weight coefficients were

explored. The fusion datasets were trained by Faster R-CNN model. We compared different feature extraction networks, including VGG 16, AlexNet, ZFNet and ResNet-50. We compared the detection performance of this model with models such as SSD, YOLOv5, YOLOv7, CenterNet, and RetinaNet. The optimized Faster R-CNN (VGG16) model was used to assess FHB resistance and compared with manual resistance assessment results, and finally, the following conclusions could be drawn:

1) The Fusion Image-1, Fusion Image-2, and Fusion Image-3 data were formed using different weights of three feature wavelengths with the XLW method, and they were compared with the RGB dataset. The Faster R-CNN model with different networks was trained on the four datasets. It was found that the model with the VGG16 network performed best, and the mAP ranged from 97.7% to 98.8%. The model showed the best performance for the Fusion Image-1 dataset.

2) The Faster R-CNN model with VGG16 achieved an average detection accuracy of 99.00%, which was 23.89%, 1.21%, 0.75%, 0.62%, and 8.46% higher than SSD, YOLOv5, YOLOv7, CenterNet, and RetinaNet models, respectively. In addition, the Faster R-CNN model was able to detect FHB wheat ears with high confidence and performed well for healthy wheat ears.

3) Compared with the FHB resistance results of the manual method, the accuracy of the optimized Faster R-CNN with the VGG16 model was 100% for severity levels 0, 1, and 4 and 93.5% and 93.8% for severity levels 2 and 3, respectively.

Acknowledgements

This study was supported by the Natural Science Foundation of Jiangsu Province (Grant No. BK20221518), and the Jiangsu Agriculture Science and Technology Innovation Fund (Grant No. CX (23)1002). The authors declare that there are no conflicts of interest.

[References]

[1] Wegulo S N, Baenziger P S, Nopsa J H, Bockus W W, Hallen-Adams H. Management of Fusarium head blight of wheat and barley. Crop Protection, 2015; 73: 100–107.

[2] Zhu Z W, Hao Y F, Mergoum M, Bai G H, Humphreys G, Cloutier S, et al. Breeding wheat for resistance to Fusarium head blight in the Global North: China, USA, and Canada. The Crop Journal, 2019; 7(6): 730–738.

[3] Li T, Zhang H J, Huang Y W, Su Z Q, Deng Y, Liu H W, et al. Effects of the *Fhb1* gene on Fusarium head blight resistance and agronomic traits of winter wheat. The Crop Journal, 2019; 7(6): 799–808.

[4] Giancaspro A, Lionetti V, Giove S L, Zito D, Fabri E, Reem N, et al. Cell wall features transferred from common into durum wheat to improve Fusarium Head Blight resistance. Plant Science, 2018; 274: 121–128.

[5] 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(2): 304–312.

[6] Landschoot S, Waegeman W, Audenaert K, Van Damme P, Vandepitte J, De Baets B, et al. A field-specific web tool for the prediction of Fusarium head blight and deoxynivalenol content in Belgium. Computers and Electronics in Agriculture, 2013; 93: 140–148.

[7] Palacios S A, Erazo J G, Ciasca B, Lattanzio V M T, Reynoso M M, Farnochi M C, et al. Occurrence of deoxynivalenol and deoxynivalenol-3-glucoside in durum wheat from Argentina. Food Chemistry, 2017; 230: 728–734.

[8] Peiris K H S, Dong Y H, Davis M A, Bockus W W, Dowell F E. Estimation of the deoxynivalenol and moisture contents of bulk wheat grain samples by FT-NIR spectroscopy. Cereal Chemistry, 2017; 94(4): 677–682.

[9] Visconti A, Pascale M. An overview on Fusarium mycotoxins in the durum wheat pasta production chain. Cereal Chemistry, 2010; 87(1): 21–27.

[10] Pestka J J, Smolinski A T. Deoxynivalenol: toxicology and potential effects on humans. Journal of Toxicology and Environmental Health Part B, 2005; 8(1): 39–69.

[11] Ma Y Y, Guo H W. Mini-review of studies on the carcinogenicity of deoxynivalenol. Environmental Toxicology and Pharmacology, 2008; 25(1): 1–9.

[12] Steiner B, Buerstmayr M, Michel S, Schweiger W, Lemmens M, Buerstmayr H. Breeding strategies and advances in line selection for Fusarium head blight resistance in wheat. Tropical Plant Pathology, 2017; 42: 165–174.

[13] Jennings P, Köhl J, Gosman N. Control of mycotoxins: raw material production. In: Magan N, Olsen M. (Eds.), Mycotoxins in Food. Sawston: Woodhead Publishing. 2004; pp.443–460.

[14] Su Z Q, Jin S J, Zhang D D, Bai G H. Development and validation of diagnostic markers for *Fhb1* region, a major QTL for Fusarium head blight resistance in wheat. Theoretical and Applied Genetics, 2018; 131: 2371–2380.

[15] Mesterházy A. Types and components of resistance to Fusarium head blight of wheat. Plant Breeding, 1995; 114(5): 377–386.

[16] Manavalan, R. Automatic identification of diseases in grains crops through computational approaches: A review. Computers and Electronics in Agriculture, 2020; 178: 105802.

[17] Bai G H, Shaner G. Management and resistance in wheat and barley to Fusarium head blight. Annual Review of Phytopathology, 2004; 42: 135–161.

[18] Shete S, Srinivasan S, Gonsalves T A. TasselGAN: An application of the generative adversarial model for creating field-based maize tassel data. Plant Phenomics, 2020; 2020: 8309605.

[19] David E, Madec S, Sadeghi-Tehran P, Aasen H, Zheng B Y, Liu S Y, et al. Global wheat head detection (GWHD) dataset: A large and diverse dataset of high-resolution RGB-Labelled images to develop and benchmark wheat head detection methods. Plant Phenomics, 2020; 2020: 3521852.

[20] Almoujahed M B, Rangarajan A K, Whetton R L, Vincke D, Eylenbosch D, Vermeulen P, et al. Detection of fusarium head blight in wheat under field conditions using a hyperspectral camera and machine learning. Computers and Electronics in Agriculture, 2022; 203: 107456.

[21] Zhao T X, Chen M, Jiang X S, Shen F, He X M, Fang Y, et al. Integration of spectra and image features of Vis/NIR hyperspectral imaging for prediction of deoxynivalenol contamination in whole wheat flour. Infrared Physics and Technology, 2020; 109: 103426.

[22] Vincke D, Eylenbosch D, Jacquemin G, Chandelier A, Pierna J A F, Stevens F, et al. Near infrared hyperspectral imaging method to assess Fusarium Head Blight infection on winter wheat ears. Microchemical Journal, 2023; 191: 108812.

[23] Zhang N, Pan Y C, Feng H K, Zhao X Q, Yang X D, Ding C L, et al. Development of *Fusarium* head blight classification index using hyperspectral microscopy images of winter wheat spikelets. Biosystems Engineering, 2019; 186: 83–99.

[24] Zhang D-Y, Chen G, Yin X, Hu R J, Gu C-Y, Pan Z G, et al. Integrating spectral and image data to detect Fusarium head blight of wheat. Computers and Electronics in Agriculture, 2020; 175: 105588.

[25] Gao C F, Guo W, Yang C H, Gong Z, Yue J B, Fu Y Y, et al. A fast and lightweight detection model for wheat fusarium head blight spikes in natural environments. Computers and Electronics in Agriculture, 2024; 216: 108484.

[26] Moghimi A, Yang C, Anderson J A. Aerial hyperspectral imagery and deep neural networks for high-throughput yield phenotyping in wheat. Computers and Electronics in Agriculture, 2020; 172: 105299.

[27] Ren S Q, He K M, Girshick R, Sun J. Faster R-CNN: Towards real-time object detection with region proposal networks. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017; 39(6): 1137–1149.

[28] Quan L Z, Feng H Q, Lv Y J, Wang Q, Zhang C B, Liu J G, et al. Maize seedling detection under different growth stages and complex field environments based on an improved Faster R–CNN. Biosystems Engineering, 2019; 184: 1–23.

[29] Ozguven M M, Adem K. Automatic detection and classification of leaf spot disease in sugar beet using deep learning algorithms. Physica A: Statistical Mechanics and its Applications, 2019; 535: 122537.

[30] Wu W, Yang T-l, Li R, Chen C, Liu T, Zhou K, et al. Detection and enumeration of wheat grains based on a deep learning method under various scenarios and scales. Journal of Integrative Agriculture, 2020; 19(8): 1998–2008.

[31] Yang D, Zhou Y X, Jie Y, Li Q Q, Shi T Y. Non-destructive detection of

defective maize kernels using hyperspectral imaging and convolutional neural network with attention module. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 2024; 313: 124166.

[32] Li J M, Yang Z H, Zhao Y R, Yu K Q. HSI combined with CNN model detection of heavy metal Cu stress levels in apple rootstocks. Microchemical Journal, 2023; 194: 109306.

[33] Zhang D Y, Luo H S, Cheng T, Li W F, Zhou X G, Gu C Y, et al. Enhancing wheat Fusarium head blight detection using rotation Yolo wheat detection network and simple spatial attention network. Computers and Electronics in Agriculture, 2023; 211: 107968.

[34] Yu X J, Lu H D, Liu Q Y. Deep-learning-based regression model and hyperspectral imaging for rapid detection of nitrogen concentration in oilseed rape (*Brassica napus* L.) leaf. Chemometrics and Intelligent Laboratory Systems, 2018; 172: 188–193.

[35] Jin X, Jie L, Wang S, Qi H J, Li S W. 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(3): 395.

[36] Rehman T U, Ma D D, Wang L J, Zhang L B, Jin J. Predictive spectral analysis using an end-to-end deep model from hyperspectral images for high-throughput plant phenotyping. Computers and Electronics in Agriculture, 2020; 177: 105713.

[37] Zhang P P, Guo C J, Liu Z, Bernardo A, Ma H X, Jiang P, et al. Quantitative trait loci for Fusarium head blight resistance in wheat cultivars Yangmai 38 and Zhengmai 9023. The Crop Journal, 2020; 9(1): 143–153.

[38] Guo J, Zhang X L, Hou Y L, Cai J J, Shen X R, Zhou T T, et al. High-density mapping of the major FHB resistance gene Fhb7 derived from Thinopyrum ponticum and its pyramiding with Fhb1 by marker-assisted selection. Theoretical and Applied Genetics, 2015; 128: 2301–2316.

[39] Isebaert S, Saeger S D, Devreese R, Verhoeven R, Maene P, Heremans B, Haesaert G. Mycotoxin-producing fusarium species occurring in winter wheat in belgium (Flanders) during 2002-2005. Journal of Phytopathology, 2009; 157(2): 108–116.

[40] GB/T 15796—2011. Rules for monitoring and forecast of the wheat head blight (*Fusarium graminearum* Schw./*Gibberella zeae* (Schw.) Petch). Standardization Administration of China, 2011. (in Chinese)

[41] Xie A G, Sun D-W, Zhu Z W, Pu H B. Nondestructive measurements of freezing parameters of frozen porcine meat by NIR hyperspectral imaging. Food and Bioprocess Technology, 2016; 9: 1444–1454.

[42] Alisaac E, Behmann J, Kuska M T, Dehne HW, Mahlein A K. Hyperspectral quantification of wheat resistance to Fusarium head blight: Comparison of two Fusarium species. European Journal of Plant Pathology, 2018; 152: 869–884.

[43] Zhang D Y, Xu L, Liang D, Xu C, Jin X L, Weng S Z. Fast prediction of sugar content in Dangshan pear (*Pyrus* spp.) using hyperspectral imagery data. Food Analytical Methods, 2018; 11: 2336–2345.

[44] Ma J, Cheng J-H, Sun D-W, Liu D. Mapping changes in sarcoplasmatic and myofibrillar proteins in boiled pork using hyperspectral imaging with spectral processing methods. LWT, 2019; 110: 338–345.

[45] Liang K, Zhang X X, Ding J, Xu J H, Han D S, Shen M X. Discrimination of wheat scab infection level by Fourier mid-infrared technology combined with sparse representation based on classified method. Spectroscopy and Spectral Analysis, 2019; 39: 3251–3255. (in Chinese)

[46] Zhu M Y, Yang H B, Li Z W. Early detection and identification of rice sheath blight disease based on hyperspectral image and chlorophyll content. Spectroscopy and Spectral Analysis, 2019; 39(6): 1898–1904. (in Chinese)

[47] Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks. Advances in Neural Information Processing Systems, 2012; 25(2): 3065386.

[48] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. Computer Science, 2014.

[49] Zeiler M D, Masum R. Visualizing and understanding convolutional networks. In: ECCV 2014, Zürich, Switzerland: Springer Science, Business Media, 2014; pp.818–833.

[50] He K M, Zhang X Y, Ren S Q, Sun J. Deep residual learning for image recognition. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA: IEEE, 2016; pp.770–778.

[51] Wilson D R, Martinez T R. The need for small learning rates on large problems. In: IEEE International Joint Conference on Neural Networks, Washington, DC, USA: IEEE, 2002; pp.115–119.

[52] Abbas Q, Bangyal W H, Ahmad J. Analysis of learning rate using BP algorithm for hand written digit recognition application. In 2010 International Conference on Information and Emerging Technologies, Karachi, Pakistan: IEEE, 2010; pp.1–5.

[53] Smith S L, Kindermans P-J, Le Q V. Don't decay the learning rate, increase the batch size. ICLR 2018, 2018;

[54] Ioffe S, Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. Journal of Machine Learning Research, 2015; 2: 03167.

[55] Gao F F, Fu L S, Zhang X, Majeed Y, Li R, Karkee M, et al. Multi-class fruit-on-plant detection for apple in SNAP system using Faster R-CNN. Computers and Electronics in Agriculture, 2020; 176: 105634.

[56] Peng H X, Huang B, Shao Y Y, Li Z S, Zhang C W, Chen Y, et al. General improved SSD model for picking object recognition of multiple fruits in natural environment. Transactions of the CSAE, 2018; 34(16): 155–162. (in Chinese)