Identification of *Fusarium* head blight in wheat ears using vertical angle-based reflectance spectroscopy

Linsheng Huang[1] · Hansu Zhang[1,2] · Wenjiang Huang[1,2,3] · Yingying Dong[2] · Huichun Ye[2,3] · Huiqin Ma[2,4] · Jinling Zhao[1]

Abstract

Fusarium head blight (FHB) is a major disease that negatively affects wheat yield in China. Given that conventional reflectance spectroscopy measurements are perpendicular to crop canopy, the identification of FHB in wheat ears with the spectral data from the vertical angle can provide the possibility for large-scale monitoring. In this study, multi—features were selected and constructed to realize the identification of FHB in wheat ears from the vertical angle, and the influence of leafy and leafless samples were discussed. Firstly, the multi-features, such as band features, spectral position features, and vegetation indices for the leafy and leafless samples, were used to evaluate the ability to identify FHB, and correlation analysis was performed to select the effective features. In order to further reduce redundancy and enhance the separation capability of features, these candidate features were categorized into different feature sets based on Fisher score values. Then, the support vector machine (SVM) algorithm was used to construct the FHB identification model based on different feature sets of leafy and leafless samples. The optimal multi-features and the best classification accuracy were finally determined. The results were showed in the following: (1) The overall accuracies and Kappa coefficients of leafy samples could reach up to 65% and 0.28, respectively, whereas the values for the leafless samples could reach 81% and 0.63 in this model; (2) the optimal multi-features had great potential in identifying FHB-infected wheat ears; and (3) the presence of leaves would reduce the model's identification capability and adversely affected the identification of FHB in wheat ears. These results provide realistic theoretical references for large-scale FHB monitoring, which are conducive to the selective harvest of wheat.

Keywords *Fusarium* head blight · Identification · Wheat ears · Vertical angle · Leaves · Multi-features

✉ Wenjiang Huang
huangwj@radi.ac.cn

1   National Engineering Research Center for Agro-Ecological Big Data Analysis & Application, Anhui University, Hefei 230601, China

2   Key Laboratory of Digital Earth Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China

3   Key Laboratory for Earth Observation of Hainan Province, Sanya 572029, China

4   Collaborative Innovation Center on Forecast and Evaluation of Meteorological Disasters, Nanjing University of Information Science & Technology, Nanjing 210044, China

Introduction

Fusarium head blight (FHB), which is caused by *Fusarium graminearum* Schwabe, is one of the primary diseases that negatively affect wheat production in China (Goswami and Kistler 2004). FHB occurs seriously in the Yangtze River and Huai River basin, China (McBeath et al. 2010). The disease has also been found in and around the Yellow River basin and is spreading northward gradually (Zhang et al. 2014). The yield and quality of FHB-infected wheat will be reduced, and a variety of mycotoxins, such as deoxynivalenol (DON) and zearalenone (ZEN), may be produced, which seriously harm the health of animals and humans and cause food safety problems (Li et al. 2018). Therefore, the accurate identification of FHB is important for disease prevention and control.

Various methods have been applied to identify wheat FHB. The incidence of wheat FHB in the field can be directly

observed through visual interpretation traditionally. However, this method is time-consuming and laborious. What is more, it is easily affected by subjective consciousness. The morphology and internal physiological structure of wheat will changed after fungus infection, which can be effectively reflected by the spectral reflectance of diseased plant tissues (Kuenzer and Knauer 2013). Therefore, many scholars used hyperspectral imaging technology to identify FHB in wheat ears. For example, Zhang et al. (2019a) proposed a special FHB classification index combining spectral bands of 417, 539, and 668 nm based on hyperspectral microscopic images of the side of wheat ears, and proved that this index could be successfully applied to the classification of wheat hyperspectral image data. Jin et al. (2018) applied the convolutional neural network classification algorithm to hyperspectral image pixels and successfully classified healthy and infected wheat ears. Mahlein et al. (2019) calculated six spectral vegetation indices based on hyperspectral images of the side of wheat ears to investigate the potential of specific bands in detecting FHB; the results showed that pigment-specific simple ratio (PSSRa and PSSRb) had the highest sensitivity for early detection of *Fusarium* infection. These studies have made progress in the use of hyperspectral imaging technology to detect wheat ears. However, hyperspectral imaging technology consumes large amounts of memory and transmission bandwidth, resulting in a significant increase in computing costs (Jin et al. 2018). Compared with hyperspectral imaging technology, non-imaging spectrometer can detect spectral information of crops with low cost and fast speed and can reduce the time of data pretreatment (Wang et al. 2011). Several scholars used non-imaging hyperspectral instruments for wheat FHB detection. Ma et al. (2020) used a ground surface spectrometer to measure the spectra from the side angle of wheat ears and carried out wavelet transform combined with Fisher linear analysis to establish a wheat FHB identification model with an overall accuracy higher than 88%. Based on the spectra measured by Analytical Spectral Devices (ASD) spectrometer, Huang et al. (2019a) extracted the derivative and absorption features and vegetation indices from the side angle of winter wheat ears, and then these features were used to construct effective identification models of disease severity under the combination of Fisher's linear discriminant analysis and support vector machine (SVM). These studies indicated that non-imaging hyperspectral technology had great potential in the identification of FHB in wheat ears.

In addition, we found that the spectral features used for FHB identification in the above studies were all derived from the combination or transformation of the spectral reflectance collected from the side angle of wheat ears. This spectral measurement method had been applied to the identification of wheat FHB in the field by relying on large machines or tool carriers, such as specific tractors (Whetton et al. 2018; Dammer et al. 2011). However, it caused damage to wheat.

Moreover, in practical large-scale satellite remote sensing-scale applications, collecting side spectral information of wheat is difficult. The sensors generally capture the wheat canopy vertically, which means that the collected spectra mainly reflect the information on the top of wheat ears. The features extracted from the side angle of wheat ears cannot provide real reference at the canopy, field, or large regional scales. Application to precision agriculture may lead to misclassification of healthy and diseased wheat, resulting in waste of pesticide spraying and environmental pollution. Summarizing and analyzing the proprietary spectral features of wheat ears from the vertical angle can provide a powerful basis for the large-scale identification of wheat FHB in the future. Wheat FHB mainly occurs in wheat ears, so scholars only study diseases occurring at the ear level of wheat. However, in addition to wheat ears, sensors can also detect the spectral information of wheat leaves. As an important part of wheat, the information on wheat leaves may interfere with FHB identification. Few included wheat leaves in previous studies on wheat ear scales. Therefore, the assessment of identification results of wheat samples with and without leaves in this model is necessary.
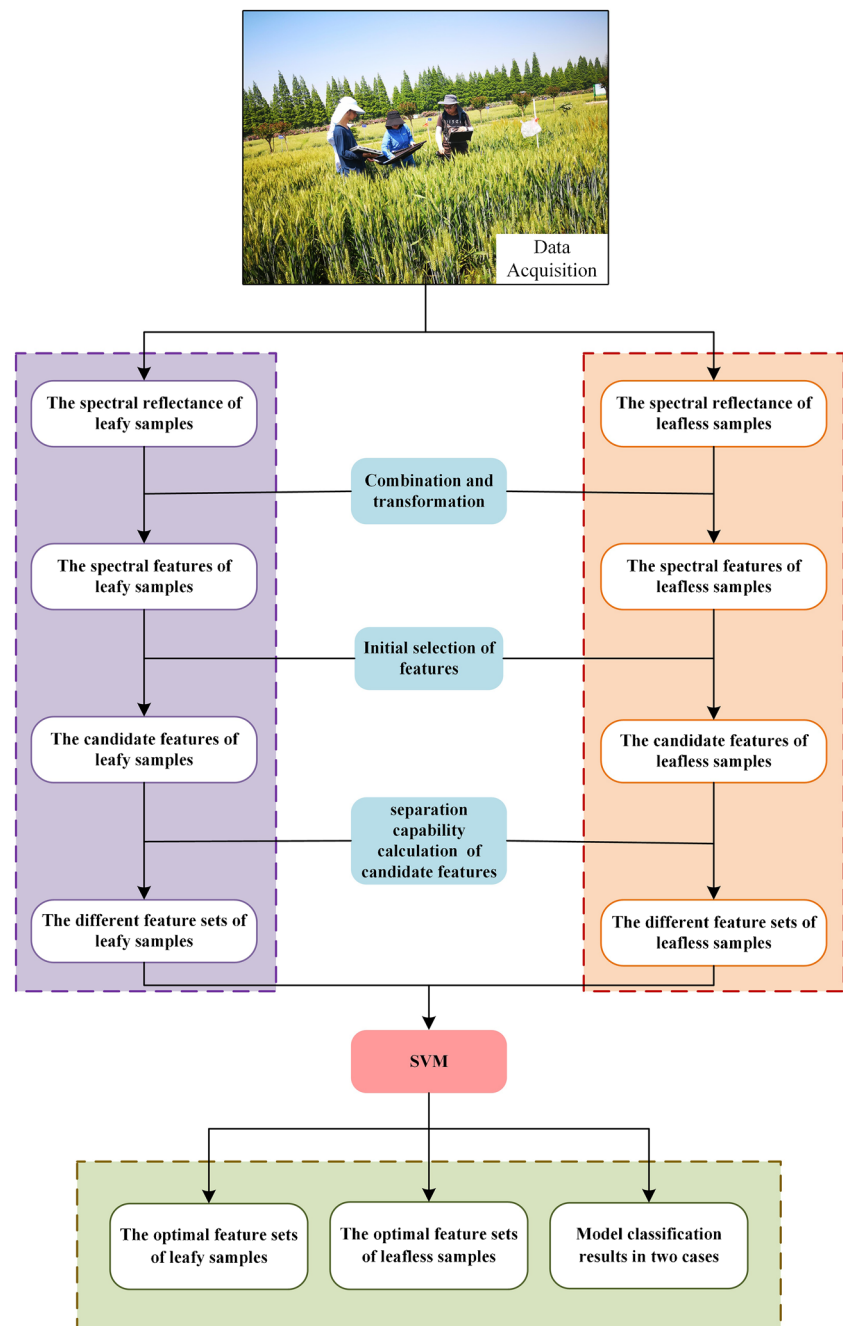
In our study, we attempted to establish a method for identifying FHB in wheat ears from the vertical angle with non-imaging spectral reflectance to provide references for large-scale vertical observation. The influence of wheat leaves presence on model accuracy was evaluated to provide detailed prior knowledge for the improvement of model accuracy in the later stage. The purposes of this paper was listed in the following: (1) to determine the optimal features for identification of wheat FHB at the vertical angle; (2) to evaluate the efficiencies of FHB identification model; and (3) to compare and analyze the influence of wheat leaves on the model. Figure 1 presents the workflow of the study.

Materials and methods

Study area and data acquisition

Experimental area

Wheat FHB is a typical climatic disease. The high temperature and humidity during the heading and flowering periods are conducive to disease outbreaks (Chen et al. 2017). According to the climate history data of Anhui Climate Center, the average temperature in Lujiang County in April 2019 was 14–23 °C, the highest temperature in early May was 28 °C, and the number of rainy days in April was 11 days. Sufficient moisture and suitable temperatures provide suitable climatic conditions for FHB occurrence. The present experiment was conducted in a test field in Guohe Town (31° 29′ N, 117° 13′ E), Lujiang County, Anhui Province (Fig. 2). The

Fig. 1 Workflow of this study

wheat was at the grain-filling period. Data were collected on May 3 and May 8, 2019. The abundant *Fusarium oxysporum* resulted in the serious occurrence of FHB in this test field.

Spectral measurement

The hyperspectral data of wheat ears was collected using the ASD FieldSpec Pro spectrometer (Analytical Spectral Devices, Inc., Boulder, CO, USA) on May 3 and 8, 2019. The spectral sampling interval was 1 nm, the spectral range was from 350 to 2500 nm for the spectrometer, and the spectral resolution was 3 nm in the 350–1000-nm range and 10 nm in the 1000–2500-nm range. In the experiment, a 1 m × 1 m black cloth was used to reduce the interference of other background factors such as soil. The field of view of the spectrometer was 25°, and all the wheat ears were measured at a height of 0.5 m above the black cloth. To explore the influence of leaves on the results of FHB identification, we considered the flag leaf of wheat as a representative and performed spectral measurements on leaf and leafless samples (Fig. 3).

Following the spectral measurement method described by Huang et al. (2019a), a small hole was cut at the center of the
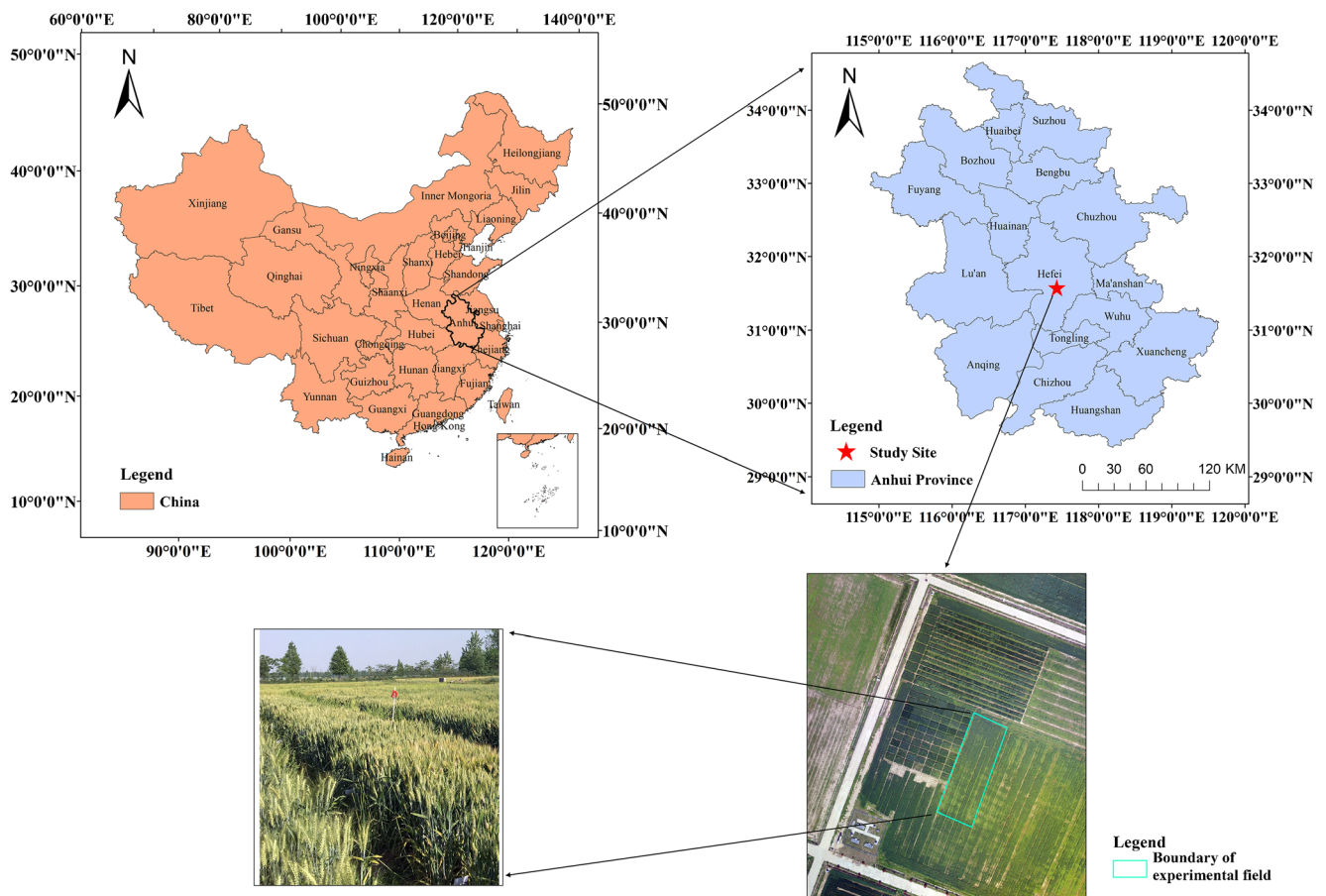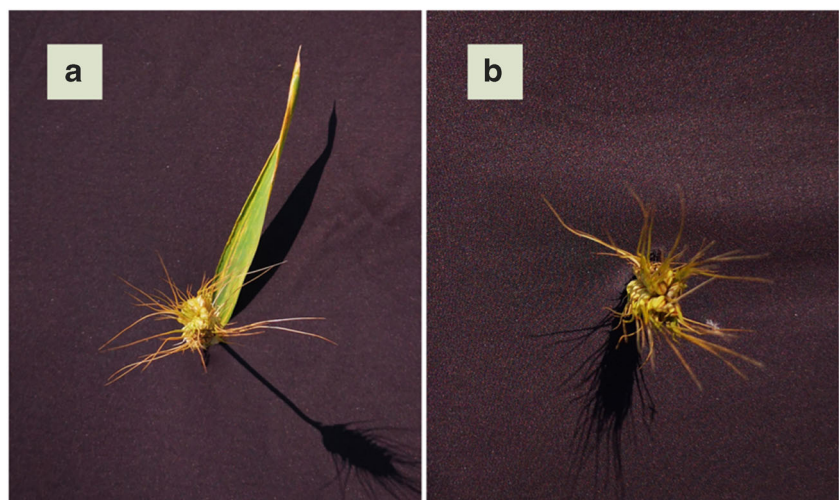
Fig. 2 Location of the study site

black cloth, and a single wheat sample was randomly selected at each experimental point in the study area, and leaves of the wheat sample were retained. We inserted the wheat sample vertically into the center of the black cloth and ensured that the ear and leaves were above the black cloth, and then put our hands under the black cloth to fix the wheat sample. A probe was vertically placed on top of the ear for spectral measurement. After the measurement of the leaf sample, we removed the sample leaves and placed the probe on the top of the wheat ear for spectral measurement. Thus, two sets of spectral data were collected for the same wheat sample. All spectral measurements were obtained at 10:00 a.m.–14:00 p.m. (local time) under clear and cloudless weather conditions. Each wheat sample with or without leaves was

Fig. 3 Leafy (**a**) and leafless samples (**b**) for spectral measurement

measured 10 times, and the average value was considered the final spectrum. A 40 cm × 40 cm BaSO$_4$ calibration panel was used for spectral correction before every measurement. The spectral reflectance was calculated using the following formula (Huang et al., 2019b):

$$R_m = \frac{\text{DN}_m}{\text{DN}_b} \times R_b \tag{1}$$

where R_m is the target reflectance, DN_m is the gray value of the target sample measured using a spectrometer, DN_b is the gray value of the calibration panel, and R_b is the standard reflectance of the calibration panel.

Determination of disease severity

According to the Rules for Monitoring and Forecast of Wheat Head Blight (*Fusarium graminearum* Schw./*Gibberella zeae* (Schw.) Petch) published in 2011 (GB/T 15796-2011), the disease severity of each wheat ear is determined by the infection ratio of diseased kernels to the total number of kernels (Ma et al. 2020). Herein, the disease infestation conditions of each wheat sample were reorganized into two classes for identification analysis: healthy (infection ratio is 0) and FHB infected (0 < infection ratio ≤ 100%). A total of 130 samples were selected, and the number of diseased kernels of each wheat sample was recorded based on visual interpretation. This study included 47 healthy and 83 infected samples.

Selection and processing of spectral features

Hyperspectral remote sensing can generate hundreds or thousands of narrow, continuous spectral bands, which may provide important additional information. The analysis of a single feature is often insufficient to explore such a wealth of information. It is significant to prevent and control FHB by systematic summarizing and analyzing the performance of different spectral features of hyperspectral remote sensing data in FHB recognition. However, hyperspectral data is considerably complicated than multispectral data, and most adjacent bands are redundant and often highly correlated (Thenkabail et al. 2004). Therefore, the best spectral features obtained from hyperspectral data must be determined to accurately identify FHB. The spectral features employed in this research mainly included band features, spectral position features, and vegetation indices. Based on the spectrum or mechanism changes caused by wheat under disease stress, the features commonly used in the identification of crop diseases were summarized, and the candidate spectral features were obtained through correlation analysis. We obtained the final feature sets for model construction by further calculation of the candidate spectral features.

Selection of candidate spectral features

First-derivative spectra were studied to select the waveband features in our study. Spectral derivative is one of the most important techniques in analyzing and processing hyperspectral remote sensing data. Compared with the original band, the spectra after differential transformation eliminated the background effect in a certain extent and highlighted the required information. After sorting the wavebands in a descending order based on the correlation between the first-derivative spectral reflectance and disease severity, the top 10 wavebands in leafy and leafless conditions were selected as the candidate band features. The first-derivative spectra were calculated using the following equation (Huang et al. 2019b):

$$R'(\lambda_i) = \frac{R(\lambda_{i+1}) - R(\lambda_{i-1})}{2\Delta\lambda} \tag{2}$$

where $R'(\lambda_i)$ is the first-derivative reflectance of band i; $R(\lambda_{i+1})$ and $R(\lambda_{i-1})$ are the spectral reflectance of wavebands $i-1$ and $i+1$, respectively; and $\Delta\lambda$ is the wavelength of adjacent wavebands.

After wheat is infected with FHB, the shrinkage of ear tissue and the reduction of chlorophyll and water content cause changes in the absorption position and spectral reflectance of the red band (Jing et al. 2010). Continuum removal is an effective method for extracting feature information on the absorption valley of hyperspectral data. This method can enhance the spectral features in the red-absorption valley and highlight the differences among spectral curves (Fu et al. 2013). Herein, continuum removal was performed on the spectrum within 550–780 nm and used for the selection of absorption feature parameters. As shown in Fig. 4, the continuum-removal reflectance can be obtained by dividing the spectral reflectance of each waveband in the absorption position by the continuum line value at the corresponding wavelength. Three parameters, namely, depth, area, and normalized depth (ND), were proposed as the position features.
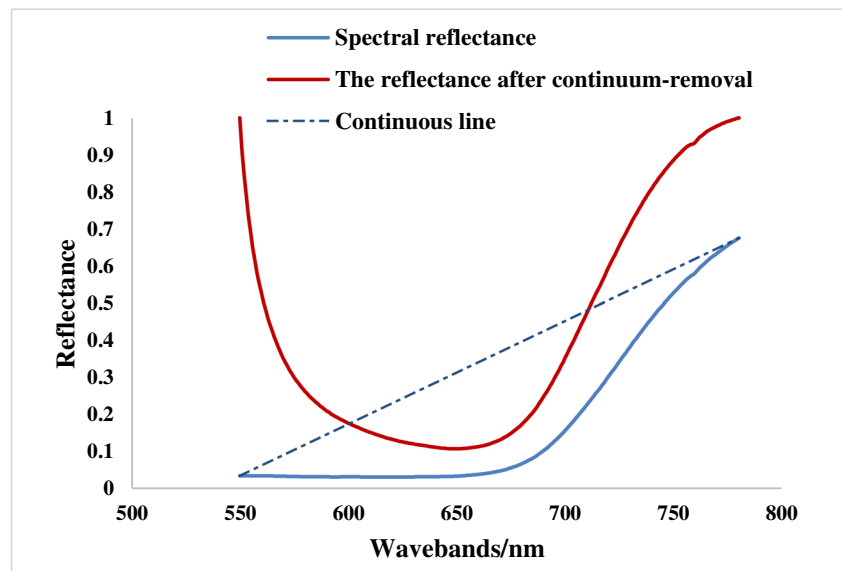
The depth feature was computed as follows:

$$\text{depth} = 1 - R'_i(\lambda_{\min}) = 1 - \frac{R_i(\lambda_{\min})}{R_c(\lambda_{\min})} \tag{3}$$

where $R_i(\lambda_{\min})$ is the minimum reflectance value within 550–780 nm, $R_c(\lambda_{\min})$ is the continuum line value of the band with minimum reflectance value, and $R_i'(\lambda_{\min})$ represents the corresponding continuum-removal reflectance value.

The absorption area is defined as the space enclosed between the reflectance curve of bands and the continuum line. This feature was calculated as follows:

$$\text{area} = \int_{\lambda_1}^{\lambda_2} (R_c(\lambda) - R_i(\lambda)) d\lambda \tag{4}$$

Fig. 4 Continuum removal curve of spectral reflectance after transformation

where $R_c(\lambda)$ and $R_i(\lambda)$ are the continuum line and spectral reflectance values of band λ, respectively. Variables λ_1 and λ_2 are the start and end positions of the wavelength, respectively. Herein, λ_1 was 550 nm, and λ_2 was 780 nm.

ND can be expressed as the ratio of maximum depth index to absorption area index. ND was computed as follows:

$$\text{ND} = \frac{\text{depth}}{\text{area}} \tag{5}$$

Wheat suffered from "green loss" under FHB stress, and the position of the green peak shifted to the red light direction ("red shift"). The three edges of vegetation spectra ("blue edge," "yellow edge," and "red edge") exhibited varying degrees of changes from the viewpoint of absorption and reflection characteristics, all of which can reflect vegetation growth (Huang et al. 2018). Aside from depth, area, and ND, 11 derivative features (Table 1) obtained by combining and transforming the spectral derivative values of the blue (490–530 nm), yellow (550–582 nm), and red edges (670–737 nm) and green peak (510–560 nm) were also considered as position features. By calculating the correlation between these 14 position features and disease severity, the features with correlation coefficient greater than 0.6 ($|R| > 0.6$) and significantly (P-value < 0.001) correlated with disease severity were selected as the candidate position features.

Different vegetation indices have been reported in the literature and are widely used in hyperspectral remote sensing monitoring of crop diseases (Huang et al. 2018; Yang et al. 2007; Bravo et al. 2003). Herein, 18 vegetation indices commonly used for crop disease monitoring were selected (Table 2). By calculating the correlation between these 18 vegetation indices and disease severity, the features with correlation coefficient greater than 0.6 ($|R| > 0.6$) and significantly (P-value < 0.001) correlated with disease severity were selected as the candidate vegetation indices.

Table 1 Definition of the 11 spectral position features

Feature	Definition
db	Maximum first-derivative value of blue edge
dy	Maximum first-derivative value of yellow edge
Dr	Maximum first-derivative value of red edge
sdb	Sum of the first-derivative values of blue edge
sdy	Sum of the first-derivative values of yellow edge
sdr	Sum of the first-derivative values of red edge
sdg	Sum of the first-derivative values of green peak
sdg/sdb	Ratio of sdg to sdb
sdy/sdb	Ratio of sdy to sdb
(sdr − sdy)/(sdr + sdy)	Normalized ratio of sdr to sdy
(sdg − sdb)/(sdg + sdb)	Normalized ratio of sdg to sdb

Table 2 Preliminary selection of vegetation indices

Feature	Definition	Formula	Reference
SIPI	Structure Insensitive Pigment Index	$(R_{800} - R_{445})/(R_{800} - R_{680})$	Devadas et al. 2009
PRI	Photochemical Reflectance Index	$(R_{570} + R_{531})/(R_{570} - R_{531})$	Gamon et al. 1997
TCARI	Transformed Chlorophyll Absorption and Reflectance Index	$3 \times [(R_{700} - R_{670}) - 0.2 \times (R_{700} - R_{550}) \times (R_{700}/R_{670})]$	Haboudane et al. 2002
NDVI	Normalized Difference Vegetation Index	$(R_{830} - R_{675})/(R_{830} + R_{675})$	Roy et al. 2016
NPCI	Normalized Pigment Chlorophyll Index	$(R_{680} - R_{430})/(R_{680} + R_{430})$	Peñuelas et al. 1994
PSRI	Plant Senescence Reflectance Index	$(R_{680} - R_{500})/R_{750}$	Merzlyak et al. 1999
PHRI	Physiological Reflectance Index	$(R_{550} - R_{531})/(R_{550} + R_{531})$	Gamon et al. 1992
ARI	Anthocyanin Reflectance Index	$(R_{550})^{-1} - (R_{700})^{-1}$	Gitelson et al. 2001
MSR	Modified Simple Ratio	$(R_{800}/R_{670} - 1)/\mathrm{sqrt}(R_{800}/R_{670} + 1)$	Haboudane et al. 2004
MCARI	Modified Chlorophyll Absorption Reflectance Index	$[(R_{700} - R_{670}) - 0.2(R_{700} - R_{550})] \times (R_{700}/R_{670})$	Daughtry et al. 2000
TVI	Triangular Vegetation Index	$0.5[120(R_{750} - R_{550}) - 200(R_{670} - R_{550})]$	Broge and Mortensen 2002
NRI	Nitrogen Reflectance Index	$(R_{570} - R_{670})/(R_{570} + R_{670})$	Filella et al. 1995
RVSI	Ratio Vegetation Stress Index	$(R_{714} + R_{752})/2 - R_{733}$	Davoud et al. 2014
GI	Greenness Index	R_{554}/R_{677}	Zarco-Tejada et al. 2005
GNDVI	Greenness Normalized Difference Vegetation Index	$(R_{747} - R_{537})/(R_{747} + R_{537})$	Thenkabail et al. 2000
NBNDVI	Narrow-Band Normalized Difference Vegetation Index	$(R_{850} - R_{680})/(R_{850} + R_{680})$	Filella et al. 1995
NDVI705	Normalized Difference Vegetation Index of Red Edge	$(R_{750} - R_{705})/(R_{750} + R_{705})$	Yang et al. 2016
MSR705	Modified Simple Ratio of Red Edge	$(R_{750} - R_{445})/(R_{705} + R_{445})$	Wang et al. 2014

Further processing of candidate spectral features

The correlation coefficients between spectral features and disease severity can connect the features with the disease infestation of wheat ears, and the threshold setting can reduce the redundancy among these features. However, in addition to correlation, the capabilities of features to classify and identify categories are important factors that affect model accuracy. The Fisher score (F-score) of features is an excellent indicator that can represent their classification capability based on the ratio of the sum of interclass distances of features to the sum of intraclass distances (Lai and He 2017). The intraclass distance of features decreased, whereas the interclass distance increased as the proportion increased, thereby enhancing the class separation capability of features, that is, the larger the F-score, the better the effect on class recognition. The F-score values of the selected candidate features were further calculated. F-score can be defined as follows:

$$\text{F-score} = \frac{\left(x_i^{(1)} - \bar{x}_i\right)^2 + \left(x_i^{(2)} - \bar{x}_i\right)^2}{\frac{1}{n_1 - 1}\sum_{k=1}^{n_1}\left(x_{k,i}^{(1)} - \bar{x}_i^{(1)}\right)^2 + \frac{1}{n_2 - 1}\sum_{k=1}^{n_2}\left(x_{k,i}^{(2)} - \bar{x}_i^{(2)}\right)^2} \tag{6}$$

where $\bar{x}_i$ is the average value of feature i, $\bar{x}_i^{(1)}$ and $\bar{x}_i^{(2)}$ represent the average of feature i in the class 1 and 2 datasets, respectively, n_1 and n_2 correspond to the number of samples corresponding to classes 1 and 2, respectively, and $x_{k,i}^{(1)}$ and $x_{k,i}^{(2)}$ denote the value of feature i at the kth sample point in classes 1 and 2, respectively.

Model construction

SVM was used to identify FHB in our study. As a traditional supervision model, SVM aims to identify a hyperplane that can correctly divide the feature space of data, and the support vector refers to the training sample point at the edge of the interval region (Ghaddar and Naoum-Sawaya 2018). The input vector can be mapped to a high-dimensional eigenvector space by kernel function, and the optimal classification surface can be constructed in this eigenspace, so as to maximize the sum of the distance between heterogeneous samples and hyperplane, and finally achieve the goal of accurate classification. Herein, the most common radial basis function (RBF) was used as the kernel function for model construction, and cross-validation was applied to identify the best parameters. The RBF is defined as follows:

$$K(x_i, x) = \exp\left\{-\frac{\|x - x_i\|}{2\sigma^2}\right\} \tag{7}$$

where σ is the kernel parameter.

By defining the kernel function, the decision function can be rewritten as follows:

$$f(x) = \text{Sgn}\left\{\sum_{i=1}^{l} y_i \alpha_i K(x_i, x) + b\right\} \qquad (8)$$

Among the 130 samples, 87 of them were used as training samples for model construction, and the others were used as test samples for model validation. Based on SVM algorithm, our study was conducted as follows:

- The candidate features were ranked in descending order based on their F-score values. First, the first feature was used as input variable for model construction. Then, the first two features were used as input variables for model construction. Next, the first three features were used as input variables for model construction until all features were inputted into the model. The model with the highest accuracy was the optimal model.
- The above steps were repeated from two aspects—leafy and leafless—to analyze the influence of leaves on the model.
- The producer's accuracy, user's accuracy, overall accuracy, and Kappa coefficient were used as indicators to evaluate the identification capability of the model.

Results

Spectral reflectance of wheat ears

Given that the spectrum after 1000 nm is greatly affected by external noise, this study mainly analyzes the spectrum within the band range of 350–1000 nm. As shown in Fig. 5a, the spectral reflectance of healthy and infected wheat is significantly different. When wheat is infected by FHB, the spectral reflectance in the visible region increases, while the reflectance in the near-infrared region decreases. The reason for this phenomenon is that the internal pigment and water content of the wheat are affected (Zhang et al. 2012). In Fig. 5b, we can see that the spectral difference between leafy and leafless wheat is more obvious in the near-infrared region, indicating that the reflectance in the near-infrared region is greatly affected by the leaves.

Candidate feature selection by correlation analysis

Band features

The spectral reflectance data of the wheat ears and their first-derivative spectra correlated with disease severity (Fig. 6). Figure 6 a showed the correlation results between the

waveband reflectance and disease severity, whereas Fig. 6b displayed the correlation results between first-derivative reflectance and disease severity. The correlation value in Fig. 6a was low, and the maximum value did not exceed 0.7. After the first derivative of reflectance, the maximum correlation coefficients reached 0.79 and 0.77, respectively, under leafy and leafless conditions, indicating that the first derivative could improve the correlation between spectral reflectance and disease severity. Table 3 showed the first 10 bands with an extremely significantly correlation (P-value < 0.001) between the first-derivative reflectance and disease severity for leafy and leafless samples. These 20 bands were considered candidate band features.

Spectral position features

The position features primarily included 3 continuum and 11 derivative features. Correlation analysis was performed on the 14 features and disease severity (Table 4). The features that were significantly correlated with disease severity (P-value < 0.001) for the leafy samples primarily included two continuous features, namely, depth and DN, and three derivative features, including sdy, sdr, and (sdr − sdy)/(sdr + sdy). However, only the correlation coefficient between depth and disease severity was greater than 0.6 ($|R| > 0.6$). Seven position features for leafless samples, namely, depth, DN, dy, sdy, sdr, sdy/sdb, and (sdr − sdy)/(sdr + sdy), were significantly correlated (P-value < 0.001) with disease severity, and the correlation coefficients of depth, sdy, sdy/sdb, and (sdr − sdy)/(sdr + sdy) were greater than 0.6 ($|R| > 0.6$). The depth for leafy samples and depth, sdy, sdy/sdb, and (sdr − sdy)/(sdr + sdy) for leafless samples were considered candidate position features.

Vegetation indices

Table 5 listed the correlation between the 18 vegetation indices and the severity of FHB-infected wheat. SIPI, TCARI, NDVI, NPCI, PSRI, MSR, NRI, GI, NBNDVI, NDVI705, and MSR705 significantly correlated (P-value < 0.001) with disease severity, where SIPI, TCARI, NPCI, and PSRI showed positive correlation coefficients, whereas the rest showed negative correlations. The sensitive vegetation indices ($|R| > 0.6$) for leafy samples were SIPI, NDVI, PSRI, MSR, NBNDVI, and NDVI705, whereas those for the leafless samples were SIPI, NDVI, NPCI, PSRI, MSR, and NDVI705; these sensitive vegetation indices were considered candidate features.

F-score values of candidate features

According to the "Candidate feature selection by correlation analysis" section, the candidate features for leafy samples were 590, 594, 597, 598, 599, 604, 625, 626, 629, and
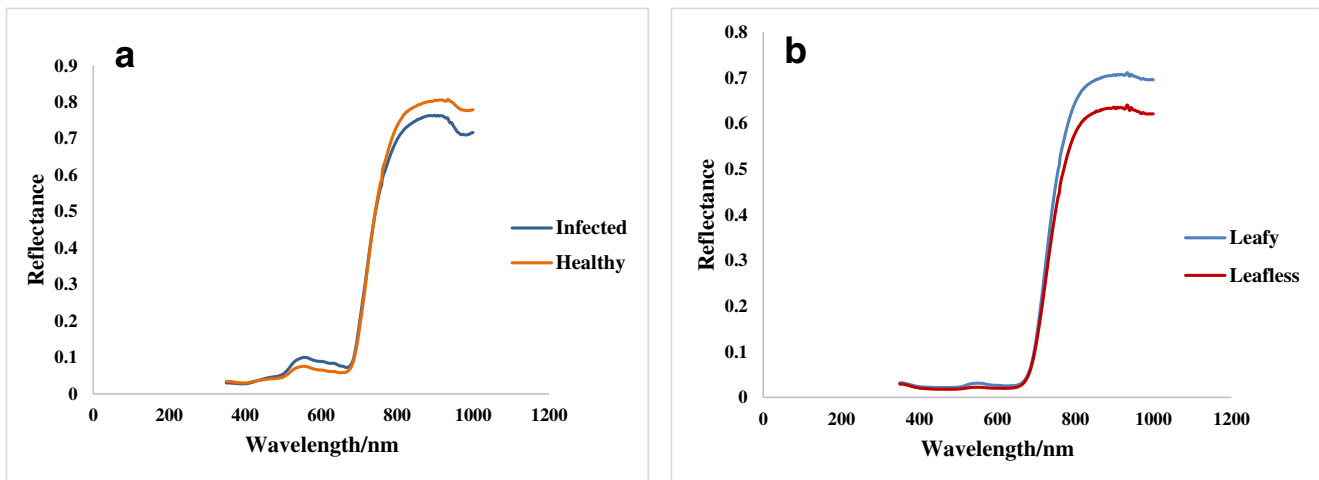
Fig. 5 Comparison of hyperspectral curves between (**a**) infected and healthy ears and (**b**) leafy and leafless ears

630 nm; depth; and SIPI, NDVI, PSRI, MSR, NBNDVI, and NDVI705. The candidate features for leafy samples were 561, 562, 563, 581, 582, 585, 590, 597, 598, and 599 nm; depth, sdy, sdy/sdb, and (sdr − sdy)/(sdr + sdy); and SIPI, NDVI, NPCI, PSRI, MSR, NBNDVI, and NDVI705. The F-score values of these 38 candidate features were calculated, and the results are shown in Fig. 7.

Best feature sets for model classification

The optimal feature sets for FHB identification of wheat ears with vertical angle were determined. Given the F-score calculation results in Fig. 7, the leafy and leafless features were arranged in descending order, and they were individually superimposed to form different input variable sets for model construction. We summarized the identification results of the optimal feature sets for FHB in wheat ears (Table 6). For leafless samples, when the input variables were the top nine features, namely, 561, 562, 581, 585, 597, and 598 nm; depth; PSRI; and (sdr − sdy)/(sdr + sdy), the constructed model

achieved the strongest identification capability. For the leafy samples, the overall accuracy of the model based on 590, 597, 598, 599, and 626 nm and depth was highest. These features with strong separation capability could reduce the operation time of the model, decreased the redundancy among features, and improved the model accuracy. As can be seen from Table 6, the optimal feature sets of the leafy and leafless samples produced overall accuracies of 65% and 81%, and the Kappa coefficients of 0.28 and 0.63, respectively, indicating that the selected features performed well in identifying healthy and FHB-infected wheat ears from the vertical angle. In addition, Table 6 revealed that the overall accuracy and Kappa coefficient of the leafy model were decreased by 16% and 0.35 than those of the leafless model, respectively. The producer's and user's accuracies for the leafless samples reached 94% and 95%, respectively, whereas those of the leafy model only reached 67% (producer's accuracy) and 75% (user's accuracy). Tables 3, 4, and 5 showed that the correlation between most features of the leafless samples and disease severity were higher than that of leafy samples. The F-score values
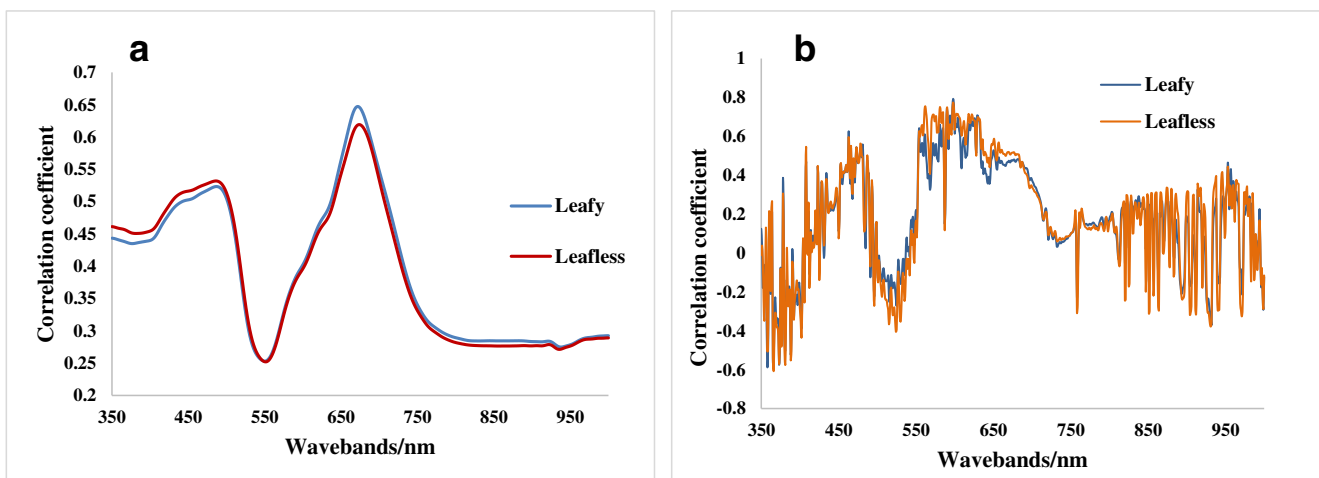

Fig. 6 Response of different spectra and disease severity. **a** Original spectra. **b** First-derivative spectra

Table 3 Correlation between the first-derivative reflectance of the first 20 bands and disease severity

Leafy			Leafless		
Wavebands	R	R^2	Wavebands	R	R^2
590***	0.748	0.559	561***	0.738	0.544
594***	0.707	0.500	562***	0.755	0.570
597***	0.750	0.562	563***	0.722	0.521
598***	0.793	0.628	581***	0.748	0.559
599***	0.759	0.576	582***	0.720	0.519
604***	0.698	0.487	585***	0.729	0.532
625***	0.684	0.468	590***	0.737	0.543
626***	0.690	0.477	597***	0.743	0.552
629***	0.710	0.504	598***	0.766	0.587
630***	0.710	0.504	599***	0.734	0.538

*** indicates the significant difference at 0.999 confidence level

of the features of leafless samples in Fig. 7 were also higher than those of the features of leafy samples on the whole. The above results illustrated that features of leafless samples were more remarkable in response to FHB identification than those of leafy samples, proving that the presence of leaves will reduce the model's capability to identify FHB.

Discussion

The selection of multi-features

The band features, spectral position features, and vegetation indices were selected to evaluate the severity of wheat FHB in our study. In recent decades, scholars have successively proposed and improved many methods to extract the spectral features of spectral information and indicated that spectral features could be basically divided into three categories. (i) Waveband features: The band of a specific part in the spectral data can be used as the basis for judging vegetation suffering from disease stress. Kobayashi et al. (2001) found that the spectral reflectance in the ranges of 430–530, 580–680, and 1480–2000 nm is more sensitive to rice panicle blast at the canopy or leaf scale. Graeff et al. (2006) analyzed the spectra of wheat leaves infected with powdery mildew and take-all disease and found that the 490, 510, 516, 540, 780, and 1300-nm bands would have a strong spectral response to the occurrence of the diseases. (ii) Spectral position features: The absorption and reflectance features that characterize hyperspectral data are also related to the specific physical and chemical characteristics of vegetation (Li et al. 2014). Huang et al. (2004) observed that the absorption depth and area could be used to establish remarkable correlation with the severity of wheat stripe rust after continuum removal within 540–740 nm. The "three edges" features are often used as key indicators to explore the spectral changes of crops under disease stress (Wang et al. 2011). (iii) Vegetation indices: Vegetation indices are special expressions that mathematically combine different wavebands, and they can enhance vegetation information and minimize non-vegetation information simultaneously (Li et al. 2014). Therefore, band features, spectral position features, and vegetation indices were selected as our research objects in our study.

The determination of model input variables

Figure 6 clearly showed the candidate features and F-score calculation results obtained after correlation analysis. The

Table 4 Correlation between position features and disease severity

Features	Leafy			Leafless		
	R	R^2	P-value	R	R^2	P-value
depth	− 0.688	0.473	0.000	− 0.734	0.539	0.000
area	0.137	0.019	0.060	0.130	0.017	0.070
DN	−0.438	0.192	0.000	−0.478	0.229	0.000
db	−0.117	0.014	0.093	−0.219	0.048	0.006
dy	0.269	0.072	0.001	0.337	0.114	0.000
dr	0.222	0.049	0.006	0.202	0.041	0.010
sdb	−0.125	0.016	0.078	−0.215	0.046	0.007
sdy	0.566	0.320	0.000	0.724	0.524	0.000
sdr	0.295	0.087	0.000	0.288	0.083	0.000
sdg	−0.055	0.003	0.267	−0.160	0.026	0.034
sdg/sdb	0.027	0.001	0.381	−0.063	0.004	0.240
sdy/sdb	0.136	0.019	0.061	0.657	0.432	0.000
(sdr−sdy)/(sdr+sdy)	−0.578	0.334	0.000	− 0.725	0.526	0.000
(sdg−sdb)/(sdg+sdb)	0.006	0.000	0.474	−0.063	0.004	0.238

Table 5 Correlation between vegetation indices and disease severity

Features	Leafy			Leafless		
	R	R^2	P-value	R	R^2	P-value
SIPI	0.718	0.516	0.000	0.720	0.518	0.000
PRI	−0.097	0.009	0.136	−0.147	0.022	0.048
TCARI	0.382	0.146	0.000	0.416	0.173	0.000
NDVI	−0.651	0.423	0.000	−0.655	0.429	0.000
NPCI	0.558	0.311	0.000	0.612	0.375	0.000
PSRI	0.694	0.482	0.000	0.751	0.564	0.000
PHRI	−0.148	0.022	0.047	−0.247	0.061	0.002
API	−0.126	0.016	0.076	−0.147	0.022	0.047
MSR	−0.601	0.361	0.000	−0.636	0.405	0.000
MCARI	0.185	0.034	0.018	0.175	0.031	0.023
TVI	0.227	0.051	0.005	0.219	0.048	0.006
NRI	−0.346	0.120	0.000	−0.396	0.157	0.000
RVSI	−0.255	0.065	0.002	−0.281	0.079	0.001
GI	−0.323	0.105	0.000	−0.371	0.138	0.000
GNDVI	−0.042	0.002	0.318	−0.039	0.002	0.330
NBNDVI	−0.676	0.457	0.000	−0.424	0.180	0.000
NDVI705	−0.634	0.402	0.000	−0.684	0.468	0.000
MSR705	−0.295	0.087	0.000	−0.363	0.132	0.000

candidate features of bands in the regions of 560–600 and 590–630 nm and the depth are located near the "green peak" and "red valley" in the visible region. The spectral reflectance in the visible region is mainly affected by pigments, especially chlorophyll content. The content of plant pigments will changes as crops are invaded by diseases, making the absorption valley in the visible region inconspicuous, and "green peak" decreases and flattens. This change provides a strong theoretical basis for identifying wheat FHB (Zhang et al. 2002). The candidate features, namely, sdy, sdy/sdb, and (sdr − sdy)/(sdr + sdy), are the parameters of "red edge," "yellow edge," and

"blue edge," which can well reflect the spectral characteristics of green vegetation and are sensitive to changes in chlorophyll, water content, and other factors (Jiang et al. 2007). The vegetation indices of SIPI, NDVI, NPCI, PSRI, MSR, NBNDVI, and NDVI705 also produce strong correlation. The selected candidate features can well represent the changes in wheat after FHB infection, but the number of features is still large. Furthermore, because the band features being spaced close to each other, redundancy may still exist. By further screening the features based on the feature separation capability, the model achieved the highest classification accuracy with the least features and its performance was improved.

Analysis of classification results

As can be seen from Table 6, the optimal identification accuracy of FHB was 81% from the vertical angle in this study. Under the same field conditions, Huang et al. (2019a) had identified wheat FHB from multiple angles, and the highest precision of the model under the vertical angle was only 68.8%, which was nearly 12% lower than our maximum result. These results also proved the effectiveness of our research method. In addition, Huang et al. (2019) also proved that the identification accuracy of FHB from the side was the highest at 88.6% and he concluded that the side of wheat ears was the best angle to identify FHB, because it contained more useful spectral information. However, this method of identifying wheat FHB from the side could not be applied to practice accurately. Our study simply analyzed the spectral changes of wheat after disease from the vertical angle and found effective features which was suitable for identification of FHB and provided guidance for larger scale FHB research. Actually, the field environment is relatively complicated with the effect of multiple factors such as temperature, wind, and multiple heads. Future studies need to take these factors into consideration.
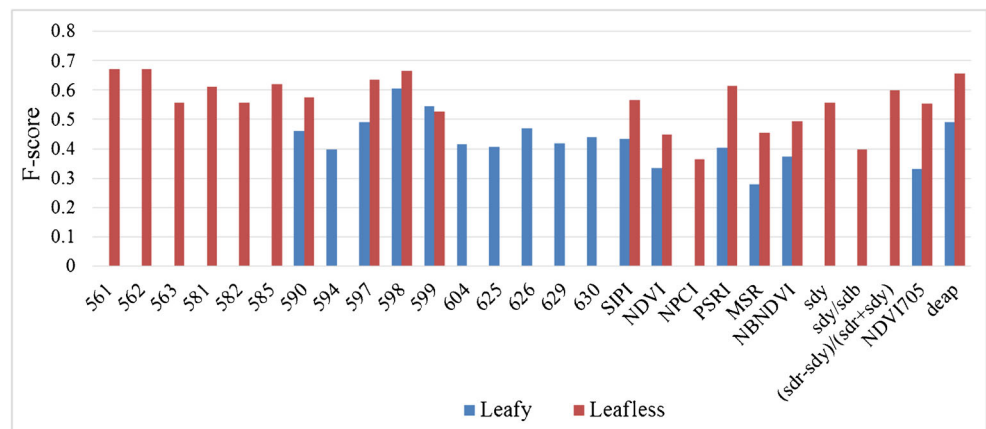
Fig. 7 F-score distribution of candidate features

Table 6 The classification accuracies of the optimal feature sets as input variables

		Healthy	Diseased	Sum	UA%	OA%	Kappa coefficient
Leafy	Healthy	10	9	19	53	65	0.28
	Diseased	6	18	24	75		
	Sum	16	27	43			
	PA/%	63	67				
Leafless	Healthy	15	7	22	68	81	0.63
	Diseased	1	20	21	95		
	Sum	16	27	43			
	PA%	94	74				

PA producer's accuracy, *UA* user's accuracy, *OA* overall accuracy

Outlook and shortcomings

Our study can directly provide guidance for canopy-scale FHB identification, and it is the first step to identify FHB on a large regional scale by applying remote sensing technology. The selected features can be further used in the identification of FHB at the canopy scale, which has high practicability in actual precision agriculture management. Image segmentation had been used to extract single wheat plant in the field to accurately identify the severity of FHB (Zhang et al., 2019b). In the future, wheat ears and leaves can be separated to further explore the best identification method for FHB applied on a large regional scale based on the canopy images. Considering that most of the information detected by the sensor was the leaf information on the top of wheat, this study only selected the flag leaf as the research object. In fact, due to different growth heights of wheat in the field, multiple leaves on the upper part of wheat can be further considered. In addition, cultivar is also an influencing factor for disease spectra detection (Cao et al. 2013). Alisaac et al. (2018) confirmed that the spectral reflectance was related to the susceptibility of wheat varieties. This result directly indicated that different wheat varieties had different sensitive wavelengths and, thus, varying spectral identification results. Given the complexity of wheat varieties in the experimental field, we disregarded the influence of different varieties on the identification of FHB in this study. The influence of wheat varieties on the performance of FHB identification model should be discussed in the future.

Conclusions

This study explored the possibility of identifying FHB in wheat ears from a vertical angle and discussed the influence of leaves on model accuracy to provide novel insights into the further study of FHB identification. Based on two sets of different hyperspectral data collected vertically, correlation analysis was performed to select the candidate spectral features, including band features, position features, and vegetation indices. The F-score values of the candidate features, which were divided into different spectral sets, were calculated. The optimal spectral feature sets for leafless samples were 561, 562, 581, 585, 597, and 598 nm; depth; PSRI; (sdr − sdy)/(sdr + sdy) and 590, 597, 598, 599, and 626 nm; and depth for the leafy samples. SVM algorithm was used to build the FHB identification model. The overall accuracy of the leafy model was 65% and that of the leafless model reached 81%. The results showed that the selected features had great potential in vertical-angle FHB identification. Compared with the leafy samples, leafless samples were more beneficial to the identification of FHB. The above results provide detailed prior knowledge for the identification of FHB at the canopy or regional scale.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

Alisaac E, Behmann J, Kuska MT, Dehne HW, Mahlein AK (2018) Hyperspectral quantification of wheat resistance to Fusarium head blight: comparison of two Fusarium species. Eur J Plant Pathol 152(4):869–884

Bravo C, Moshou D, West J, McCartney A, Ramon H (2003) Early disease detection in wheat fields using spectral reflectance. Biosyst Eng 84(2):137–145

Broge NH, Mortensen JV (2002) Deriving green crop area index and canopy chlorophyll density of winter wheat from spectral reflectance data. Remote Sens Environ 81(1):45–57

Cao XR, Luo Y, Zhou YL, Duan XY, Cheng DF (2013) Detection of powdery mildew in two winter wheat cultivars using canopy hyperspectral reflectance. Crop Prot 45:124–131

Chen Y, Wang JQ, Yang RM, Ma ZH (2017) Current situation and management strategies of Fusarium head blight in China. Plant Prot 5:11–17

Dammer KH, Möller B, Rodemann B, Heppner D (2011) Detection of head blight (Fusarium ssp.) in winter wheat by color and multispectral image analyses. Crop Prot 30(4):420–428

Daughtry CST, Walthall CL, Kim MS, Brown De Colstoun E, McMurtrey JE III (2000) Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance. Remote Sens Environ 74(2):229–239

Davoud A, Mohammad M, Alfredo H (2014) Evaluating the effect of different wheat rust disease symptoms on vegetation indices using hyperspectral measurements. Remote Sens 6(6):5107–5123

Devadas R, Lamb DW, Simpfendorfer S, Backhouse D (2009) Evaluating ten spectral vegetation indices for identifying rust infection in individual wheat leaves. Precis Agric 10(6):459–470

Filella I, Serrano L, Serra J, Penuelas J (1995) Evaluating wheat nitrogen status with canopy reflectance indices and discriminant analysis. Crop Sci 35(5):1400–1405

Fu YY, Wang JH, Yang GJ, Song XY, Xu XG, Feng HK (2013) Band depth analysis and partial least square regression based winter wheat biomass estimation using hyperspectral measurements. Spectrosc Spectr Anal 33(5):1315–1319

Gamon JA, Peñuelas J, Field CB (1992) A narrow-waveband spectral index that tracks diurnal changes in photosynthetic efficiency. Remote Sens Environ 41(1):35–44

Gamon JA, Serrano L, Surfus JS (1997) The photochemical reflectance index: an optical indicator of photosynthetic radiation use efficiency across species, functional types and nutrient levels. Oecologia 112(4):492–501

Ghaddar B, Naoum-Sawaya J (2018) High dimensional data classification and feature selection using support vector machines. Eur J Oper Res 265(3):993–1004

Gitelson AA, Merzlyak MN, Chivkunova OB (2001) Optical properties and nondestructive estimation of anthocyanin content in plant leaves. J Photochem Photobiol 74(1):38–45

Goswami RS, Kistler HC (2004) Heading for disaster: Fusarium graminearum on cereal crops. Mol Plant Pathol 5(6):515–525

Graeff S, Link J, Claupein W (2006) Identification of powdery mildew (Erysiphe graminis sp. tritici) and take-all disease (Gaeumannomyces graminis sp. tritici) in wheat (Triticum aestivum L.) by means of leaf reflectance measurements. Cent Eur J Biol 1(2):275–288

Haboudane D, Miller JR, Pattey E, Zarco-Tejada PJ, Strachan IB (2004) Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: modeling and validation in the context of precision agriculture. Remote Sens Environ 90(3):337–352

Haboudane D, Miller JR, Tremblay N, Zarco-Tejada PJ, Dextraze L (2002) Integrated narrow-band vegetation indices for predicting of crop chlorophyll content for application to precision agriculture. Remote Sens Environ 81(2–3):416–426

Huang LS, Wu ZC, Huang WJ, Ma HQ, Zhao JL (2019a) Identification of Fusarium head blight in winter wheat ears based on Fisher's linear discriminant analysis and a support vector machine. Appl Sci 9(18):3894

Huang LS, Zhang HS, Ding WJ, Huang WJ, Hu TG, Zhao JL (2019b) Monitoring of wheat scab using the specific spectral index from ASD hyperspectral dataset. J Spectrosc 1:1–9

Huang MY, Huang WJ, Liu LY, Huang YD, Wang JH, Wan AM (2004) Spectral reflectance feature of winter wheat single leaf infected with stripe rust and severity level inversion. Trans Chin Soc Agric Eng 20(1):176–180

Huang WJ, Zhang JC, Shi Y, Dong YY, Liu LY (2018) Progress in monitoring and forecasting of crop pests and diseases by remote sensing. J Nanjing Univ Inf Sci Technol 10(1):30–43

Jiang JB, Chen YH, Huang WJ (2007) Using hyperspectral derivative index to monitor winter wheat disease. Spectrosc Spectr Anal 27(12):2475

Jin X, Jie L, Wang S, Qi HJ, Li SW (2018) Classifying wheat hyperspectral pixels of healthy heads and Fusarium head blight disease using a deep neural network in the wild field. Remote Sens 10(3):395

Jing X, Wang JH, Song XY, Xu XG, Chen B, Huang WJ (2010) Continuum removal method for cotton verticillium wilt severity monitoring with hyperspectral data. Trans CSAE 26(1):193–198

Kobayashi T, Kanda E, Kitada K (2001) Detection of rice panicle blast with multispectral radiometer and the potential of using airborne multispectral scanners. Am Phytopathol Soc 91(3):316–323

Kuenzer C, Knauer K (2013) Remote sensing of rice crop areas. Remote Sens 34(6):2101–2139

Lai XF, He XS (2017) Method based on minimum redundancy and maximum separability for feature selection. Comput Eng Appl 53(12):70–75

Li WG, Chen H, Jin ZT, Zhang ZZ, Ge GX, Ji FJ (2018) Remote sensing monitoring of winter wheat scab based on suitable scale selection. J Triticeae Crops 38(11):1374–1380

Li XC, Zhang YJ, Bao YS, Luo JH, Jin XL, Xu XG, Song XY, Yang GJ (2014) Exploring the best hyperspectral features for LAI estimation using partial least squares regression. Remote Sens 6(7):6221–6241

Ma HQ, Huang WJ, Jing YS, Pignatti S, Laneve G, Dong YY, Ye HC, Liu LY, Guo AT, Jiang J (2020) Identification of Fusarium head blight in winter wheat ears using continuous wavelet analysis. Sensors 20(1):20

Mahlein AK, Alisaac E, Al Masri A, Behmann J, Dehne HW, Oerke EC (2019) Comparison and combination of thermal, fluorescence, and hyperspectral imaging for monitoring Fusarium head blight of wheat on spikelet scale. Sensors 19(10):2281

McBeath JH, McBeath J (2010) Plant diseases, pests and food security. Springer Neth 35:117–156

Merzlyak MN, Gitelson AA, Chivkunova OB, Rakitin VY (1999) Nondestructive optical detection of pigment changes during leaf senescence and fruit ripening. Physiol Plant 106(1):135–141

Peñuelas J, Gamon JA, Fredeen AL, Merino J, Field CB (1994) Reflectance indices associated with physiological changes in introgen and water-limited sunflower leaves. Remote Sens Environ 48(2):136–146

Roy DP, Kovalskyy V, Zhang HK, Vermote EF, Yan L, Kumar SS, Egorov A (2016) Characterization of Landsat-7 to Landsat-8 reflective wavelength and normalized difference vegetation index continuity. Remote Sens Environ 185:57–70

Thenkabail PS, Enclona EA, Ashton MS, Van Der Meer B (2004) Accuracy assessments of hyperspectral waveband performance for vegetation analysis applications. Remote Sens Environ 91(3–4):354–376

Thenkabail PS, Smith RB, Pauw ED (2000) Hyperspectral vegetation indices and their relationships with agricultural crop characteristics. Remote Sens Environ 71(2):158–182

Wang DC, Zhang DY, Zhao JL, Li CJ, Zhu DZ, Huang WJ, Li YF, Yang XD (2011) Using extraction of red edge position to validate consistency of hyperspectral imaging and non-imaging data. Spectrosc Spectr Anal 31(9):2450–2454

Wang RH, Song XY, Li ZH, Yang GJ, Guo WS, Tan CW, Chen LP (2014) Estimation of winter wheat nitrogen nutrition index using

hyperspectral remote sensing. Trans Chin Soc Agric Eng 30(19): 191–198

Whetton RL, Waine TW, Mouazen AM (2018) Hyperspectral measurements of yellow rust and Fusarium head blight in cereal crops: part 2: on-line field measurement. Biosyst Eng 167:144–158

Yang CM, Cheng CH, Chen RK (2007) Changes in spectral characteristics of rice canopy infested with brown lanthopper and leaffolder. Crop Sci 47(1):329–335

Yang FQ, Feng HK, Li ZH, Gao L, Yang GJ, Dai HY (2016) Hyperspectral estimation of leaf area index for winter wheat based on Akaike's information criterion. Trans Chin Soc Agric Eng 32(3): 163–168

Zarco-Tejada PJ, Berjón A, López-Lozano R, Miller JR, Martín P, Cachorro V, González MR, De Frutos A (2005) Assessing vineyard condition with hyperspectral indices: leaf and canopy reflectance simulation in a row-structured discontinuous canopy. Remote Sens Environ 99(3):271–287

Zhang DY, Wang DY, Gu CY, Jin N, Zhao HT, Chen G, Liang HY, Liang D (2019b) Using neural network to identify the severity of wheat Fusarium head blight in the field environment. Remote Sens 11(20):2375

Zhang J, Pu R, Huang W, Yuan L, Luo J, Wang J (2012) Using in-situ hyperspectral data for detecting and discriminating yellow rust disease from nutrient stresses. Field Crop Res 134:165–174

Zhang J, Yi YJ, Wang JS, Chen SH, Ling GL (2014) Research progress of control techniques on wheat scab. China Plant Prot 34(1):24–28

Zhang M, Liu X, O'neill M (2002) Spectral discrimination of Phytophthora infestans infection on tomatoes based on principal component and cluster analyses. Int J Remote Sens 23(6):1095–1107

Zhang N, Pan YC, Feng HK, Zhao XQ, Yang XD, Ding CL, Yang GJ (2019a) Development of Fusarium head blight classification index using hyperspectral microscopy images of winter wheat spikelets. Biosyst Eng 186:83–99