

Article

# Hyperspectral Reflectance Imaging Combined with Multivariate Analysis for Diagnosis of *Sclerotinia* Stem Rot on *Arabidopsis Thaliana* Leaves

Jing Liang <sup>1</sup>, Xiaoli Li <sup>2</sup>, Panpan Zhu <sup>1</sup>, Ning Xu <sup>1,\*</sup>  and Yong He <sup>2</sup> 

<sup>1</sup> Institute of Drug Development & Chemical Biology, College of Pharmaceutical Science, Zhejiang University of Technology, Hangzhou 311122, Zhejiang, China; 2111607031@zjut.edu.cn (J.L.); 2111607040@zjut.edu.cn (P.Z.)

<sup>2</sup> College of Biosystems Engineering and Food Science, Zhejiang University, 866 Yuhangtang Road, Hangzhou 310058, China; Xiaolili@zju.edu.cn (X.L.); yhe@zju.edu.cn (Y.H.)

\* Correspondence: xuning@zjut.edu.cn

Received: 28 February 2019; Accepted: 14 May 2019; Published: 21 May 2019



**Featured Application:** Hyperspectral imaging for *Sclerotinia* stem rot identification.

**Abstract:** *Sclerotinia* stem rot (SSR) is one of the most destructive diseases in the world caused by *Sclerotinia sclerotiorum* (*S. sclerotiorum*), resulting in significant yield loss. Early and high-throughput detection would be critical to prevent SSR from spreading. This study aimed to propose a feasible method for SSR detection based on the hyperspectral imaging coupled with multivariate analysis. The performance of different detecting algorithms were compared by combining the extreme learning machine (ELM), K-nearest neighbor algorithm (KNN), linear discriminant analysis (LDA), naïve Bayes classifier (NB) and the support vector machine (SVM) with the random frog (RF), successive projection algorithm (SPA) and sequential forward selection (SFS). The similarity of selected optimal wavelengths by three different feature selection methods indicated a high correlation between selected wavelengths and SSR. Compared with KNN, LDA, NB, and SVM, three wavelengths (455, 671 and 747 nm) selected by SFS-CA combined with ELM could achieve relatively better results with the overall accuracy of 93.7% and the lowest false negative rate of 2.4%. These results demonstrated the potential of the presented method using hyperspectral reflectance imaging combined with multivariate analysis for SSR diagnosis.

**Keywords:** *Arabidopsis thaliana*; *Sclerotinia sclerotiorum*; *Sclerotinia* stem rot (SSR); hyperspectral reflectance imaging

## 1. Introduction

*Sclerotinia* stem rot (SSR) is considered as an important disease of major crops [1] and some medicinal plants worldwide. It is caused by a typical necrotrophic phytopathogenic fungus, *Sclerotinia sclerotiorum* (*S. sclerotiorum*), which could induce significant crop yield loss in oilseed rape, peanut, sunflower [2], safflower [3] and perilla frutescens [4]. The traditional detection methods for SSR are mainly based on artificial identification and laboratory biochemical methods including flow cytometry, polymerase chain reaction (PCR), enzyme-linked immune sorbent assay (ELISA) and immunofluorescence (IF) [5]. Although these methods could gain relatively accurate results, they are often time-consuming and labor-intensive. Therefore, a reliable and rapid diagnose method for SSR is urgently needed.

Currently, optical techniques, especially hyperspectral imaging (HSI) has been widely used for plant diseases detection [6–8] due to the advantage of providing spectral and spatial information

simultaneously which contain information of chemical compositions and physical structures of the plant. Several studies have reported that HSI was feasible for plant disease detection, such as *Plasmopara viticola* in grape [9], green diseases in citrus [10], *Magnaporthe oryzae* in rice [11], gray mold in tomato [12,13]. As for SSR detection, one study conducted by Zhang et al. on SSR detection of oilseed rape leaves using mid-infrared spectrum provided an over 80% detection accuracy [14]. However, point to point or one dimensional (1D) measurement could cause high uncertainty in disease detection due to that the disease would cause large variability and heterogeneity from a single leaf to the entire plant. Considering the large heterogeneity of disease distributing within a leaf, two- or three-dimensional optical imaging techniques containing spectral and spatial information have been employed. Kong et al. carried out the SSR identification in oilseed rape stems using hyperspectral imaging with an overall accuracy of over 90% [15]. Although a reasonable result was achieved, it still needed 13 selected wavelengths to train the discriminant model, which can limit laboratory-based studies to a low-cost application. To our knowledge, there is little information about changes in light absorption and scattering in leaf tissue caused by *S. sclerotiorum* infection. Therefore, the model plant, *Arabidopsis thaliana* (*A. thaliana*), was introduced in this study for the investigation of diffuse reflectance heterogeneity within a whole leaf infected by *S. sclerotiorum*. Moreover, the findings would enable the effective detection of SSR based on HSI.

In the study, HSI was proposed for the investigation of SSR infected leaves, as well as detection of SSR by establishing a robust discriminant model based on hyperspectral imaging and multivariate analysis. To achieve this goal, different optimal wavelength selection methods and classification algorithms were evaluated in order to select the most optimal model that was feasible for SSR detection. The specific objectives were: (1) Select the optimal wavelengths for SSR detection; (2) establish a suitable classification model for SSR diagnosis based on selected features.

## 2. Materials and Methods

### 2.1. *A. thaliana* Cultivation and Pathogen Inoculation

The wild type (WT) *A. thaliana* seeds were provided by the Plant Environmental Sensing Laboratory (Hangzhou Normal University). The seeds were pretreated using 2.5% plant preservative mixture (PPM) and then stored in the refrigerator at 4 °C for 3 days. *A. thaliana* seeds were cultured on murashige and skoog (MS) medium in Petri dishes within a chamber under 80  $\mu\text{mol}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$  with 16-photoperiod at 22 °C, 65% humidity. Two-week-old *A. thaliana* plants were then transplanted to pots, and then the pots were covered by plastic wrap for three days in the same incubator. *A. thaliana* plants were watered every four days before the start of the experiment. *S. sclerotiorum* was cultured on a potato dextrose agar. In this study, a 5 mm diameter hyphal block was inoculated on the center of the upper surface of the *A. thaliana* leaf. After 3 h, plants were considered successfully infected when at least one leaf presented typical symptoms, while some were still asymptomatic but with watery lesions in the back of mesophyll. Finally, 285 samples were collected including 138 healthy, 72 infected asymptomatic and 75 infected symptomatic leaves, respectively.

### 2.2. Hyperspectral Images Acquisition and Processing

Hyperspectral images of *A. thaliana* leaves were collected by a push-broom hyperspectral imaging system. The system briefly was consisted of a charge coupled device (CCD) camera with 672 × 512 pixels resolution (C8484-05, Hamamatsu Photonics, Hamamatsu City, Japan), a spectrograph with 2.8 nm resolution in the range of 379–1023 nm (ImSpector V10E, Spectral Imaging Ltd. Oulu, Finland), two-line lights (Fiber-Lite DC950, Dolan Jenner Industries Inc. Boxborough, MA, USA), a mobile platform (IRCP0076, Isuzu Optics Crop, Taiwan, China), a dark box and a computer installed with V10E software of four different spectral resolutions (1.25, 2.51, 5.03 and 10.08 nm) using the interpolation algorithm. Before the final hyperspectral images acquisition, some key parameters including spectral resolutions, working distance, exposure time, light intensity and scanning speed were optimized with

the aim of achieving the best output performance. Two halogen line lights were installed on two sides of the lens above the samples with an angle of 45° to reduce the shadow. The spectral resolutions were set as 1.25 nm, and the working distance was 28 cm, and the exposure time of the charge coupled device (CCD) camera was adjusted to 40 ms. The *A. thaliana* leaves were scanned at a speed of 3 mm/s on the sampling platform with a dark background to simplify the image processing procedure. All components were fixed in a dark box to avoid any impact on the hyperspectral images collection.

Prior to further images processing, the raw hyperspectral images were calibrated by the following equation to reduce the bias caused by the uneven distribution of light.

$$R = I_{raw} - I_{dark} / (I_{white} - I_{dark}) \tag{1}$$

where  $R$ ,  $I_{raw}$ ,  $I_{dark}$ , and  $I_{white}$  were the calibrated images, original images, dark current and reference images, respectively. The calibrated spectral images of first 60 wavelengths were removed due to high noises. As shown in Figure 1, the background of each spectral images was removed firstly. The Otsu’s model was then introduced to calculate a global threshold that can segment abnormal tissues (regions of interest, ROIs) from healthy ones from the spectral image [16]. As the spectral images at 750 nm of the infected tissues (Figure 2a) presented a great contrast to healthy ones, it was determined for image segmentation. For healthy samples, the whole leaf area was considered as the ROIs. The mean spectrum of all pixels within each ROI in the range of 450–1023 nm was calculated. In order to further remove the undesirable noise that could impact the extraction of spectral features, and eventually improve the discrimination accuracy, the Savitsky-Golay (SG) smoothing algorithm was also implemented as previously described [10].

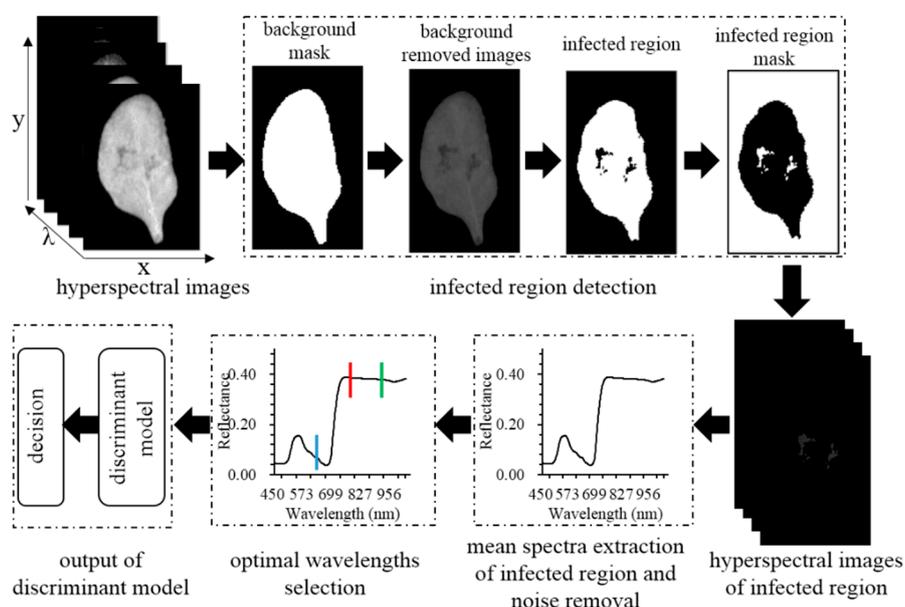


Figure 1. The pipeline of *Sclerotinia* stem rot (SSR) detection based on hyperspectral imaging.

### 2.3. Spectral and Textural Features Selection for Disease Detection

Although hyperspectral images contain rich information spatially and spectrally that was associated with the structural and biochemical properties of plant leaves, they also included redundant information. Therefore, it is generally preferred to select the most important wavelengths to remove irrelevant information so that they could be applied online for disease diagnosis with less expensive hardware setup. In this study, three commonly used optimal selection methods, including the random frog (RF), successive projection algorithm (SPA) and sequential forward selection (SFS) were used to select the most sensitive wavelengths for SSR. RF was performed to select optimal wavelengths for discriminating different leaf health conditions based on the invertible jump Markov Chain Monte

Carlo technology [17,18]. It outputs the chosen probability of each selected variable. A higher probability indicates a more important waveband. RF is an effective variable selection algorithm from high-dimension that can gain the most relevant variables for classification. SPA was also performed to select the sensitive wavelengths for SSR detection in *A. thaliana* leaves through the projection operation in a vector space. The candidate variable was projected to the maximum value on the orthogonal subspace with the minimum colinearity. The redundant wavelengths were eliminated without significant loss of the predicting ability once the root mean square error (RMSE) of multiple linear regression (MLR) model was achieved [19]. The SFS algorithm is a bottom-up process that starts with an empty subset and repeatedly adds the most important features selected by the objective function. Once the attribute is preserved, it cannot be discarded at a later stage. Fisher criteria are often used as objective functions [20].

#### 2.4. Discriminant Analysis Method for Disease Detection

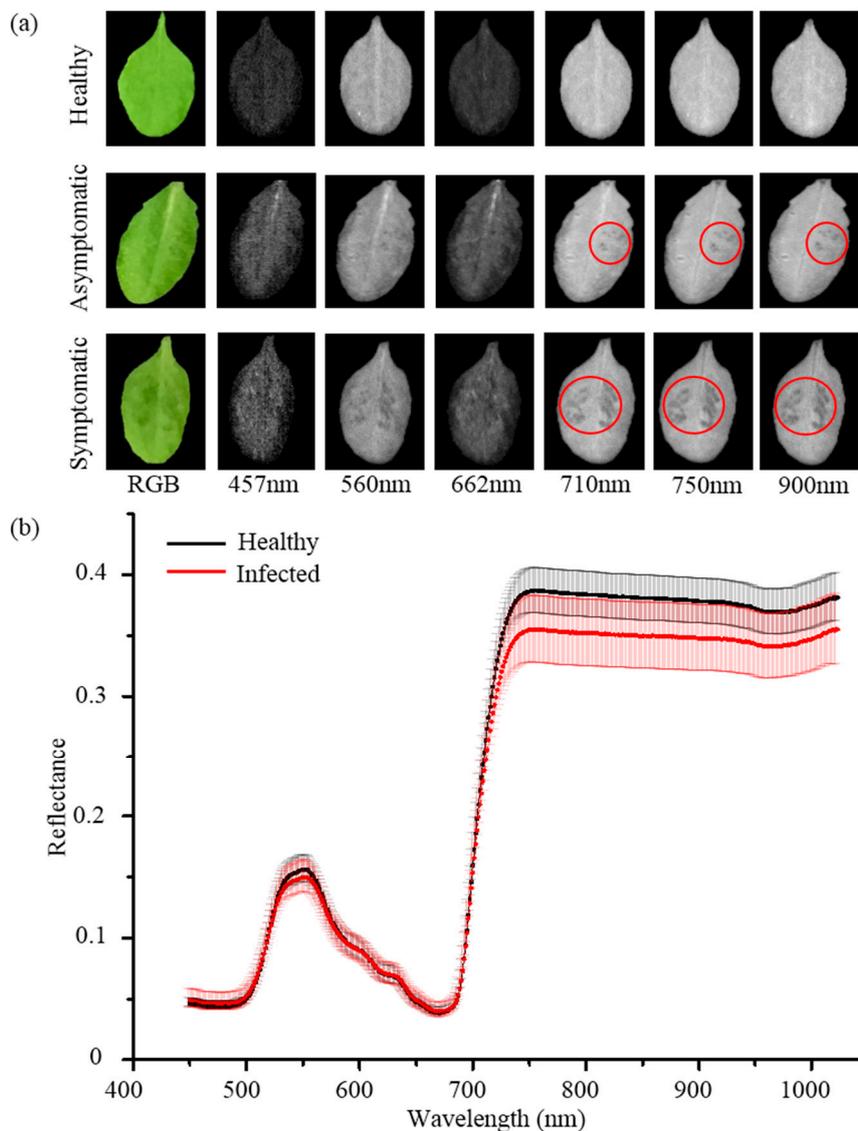
Machine learning methods have been proven useful for the detection of diseased or damaged plant materials based on hyperspectral imaging [21–25]. However, different kinds of stress would induce different reflectance patterns implying that the most feasible model should be constructed for a specific disease. In this study, four supervised classifiers including the extreme learning machine (ELM), K-nearest neighbor algorithm (KNN), linear discriminant analysis (LDA), naïve Bayes classifier (NB) and the support vector machine (SVM) were introduced. ELM is a feedforward network with a single hidden layer carrying the performance of fast learning speed and good generalization and was widely used for pattern recognition [26,27]. The classification ability largely depends on the number of hidden nodes. Therefore, a range of 1–60 hidden node number was tested in the study according to their detection performance. KNN is a non-parametric classification model by training  $k$  neighbors from the training set using a distance function, and then classifying the features into a class based on the largest voting rules [28]. In this case, four neighbors and Manhattan distance were determined to establish the KNN model. LDA is an effective subspace technique because it optimizes the Fisher score and has no need to adjust the free parameters [29,30]. NB classifier is a probabilistic classifier based on Bayes' theory with independent assumptions between features [31,32], which was also used for citrus disease recognition [33]. The SVM classifier is developed on the basis of the statistical learning theory to give the largest distance between the margins of the training data set with a hyperplane, which can be achieved by solving a convex quadratic programming problem using a kernel function [34]. The SSR detection performance of four classifiers was evaluated using the confusion matrix from which the overall classification accuracy, true positive (TP), true negative (TN), false negative (FN), and false positive (FP) were derived. The kappa coefficient is often used to measure the consistency of a classification model deriving from the confusion matrix. In general, the kappa coefficient higher than 0.8 was regarded as a good consistency. In the discrimination process, all samples were assigned to healthy and infected groups, labeled as "1" and "2" respectively. Two-thirds of the whole sample set was used as a training subset and one third as a validation subset via Kennard-Stone (KS) algorithm [35]. Data analysis in this study was carried out using MATLAB R2014a (MathWorks, Inc. Natick, MA, USA).

### 3. Results and Discussion

#### 3.1. Analysis of Hyperspectral Reflectance

With the aim to compare the differences between healthy and infected leaves, the six representative wavelengths distributing in the red light, green light, blue light, red edge, and near-infrared region, respectively, were selected for analysis. Figure 2a displays the representative color and spectral images of healthy and infected (asymptomatic and symptomatic) *A. thaliana* leaves. No significant difference can be observed from color images between healthy and asymptomatic leaves, implying the challenges of SSR identification only using color images. With the development of *S. sclerotiorum* infection, it could be found that some typical spots appear in symptomatic leaves. Considering

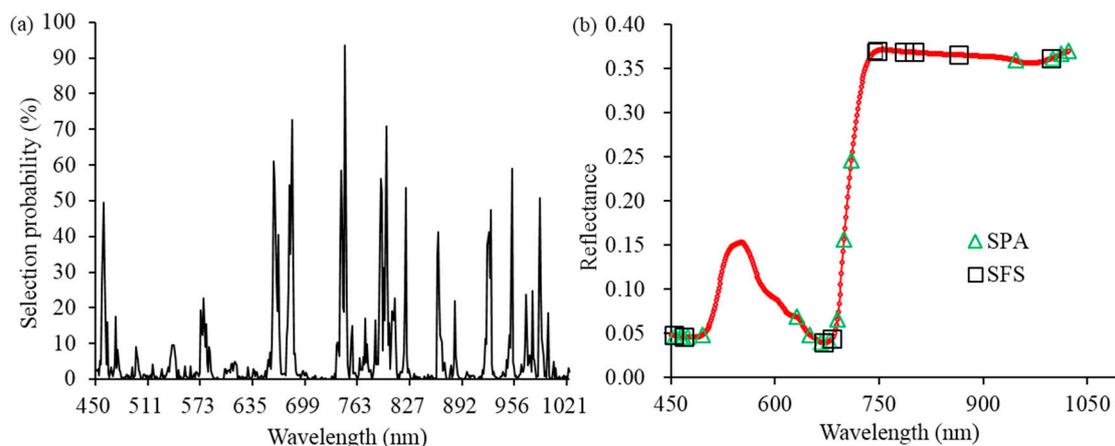
spectral images at different wavelengths, the images in the near-infrared range showed a better performance in comparison with those in the visible range and that the infected tissues (asymptomatic and symptomatic leaves) presented a great contrast to uninfected ones. Figure 2b shows the mean spectra of healthy and infected leaves. A general reflectance pattern of green leaves in the range of 450–1023 nm can be observed. The reflectance from 450 to 700 nm that carried carotenoids, surface color, and the chlorophyll contents information, while reflectance from 700 to 1023 nm was associated with the internal characteristics of *A. thaliana* leaves. The reflectance of infected leaves in the invisible range presented some overlaps to that of healthy ones, except for a slight difference in the blue and green light region, which indicated the difficulty of detecting the infected leaves of *A. thaliana* with sole color images. That is consistent with the spectral images shown in Figure 2a. The difference of reflectance in the near-infrared region between healthy and infected leaves is significant. The lower reflectance of infected leaves in the near-infrared region might result from tissue collapse and cell structure breakdown due to *S. sclerotiorum* infection, which suggests the potential of HSI for discriminating infected leaves from healthy ones.



**Figure 2.** (a) The color images and six representative spectral images of the healthy and infected samples (asymptomatic and symptomatic). (b) Mean spectra and standard deviations of the healthy ( $n = 138$ ) and infected ( $n = 147$ ) samples in the range of 450–1023 nm.

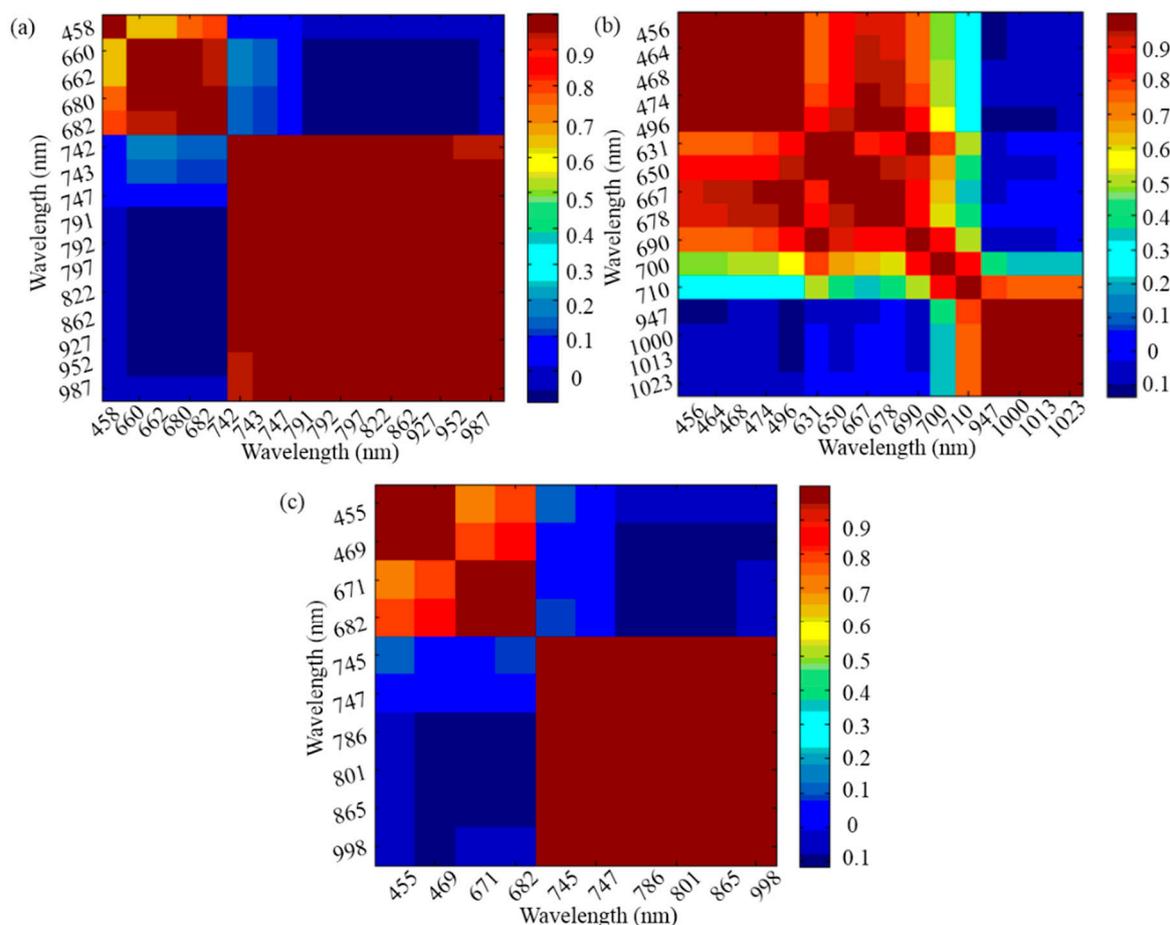
### 3.2. Optimal Wavelengths Selection for Disease Detection of *A. thaliana* Leaves

The procedure of reducing the data size through optimal wavelengths selection is generally considered important in HSI data analysis for a more robust classification model with a high computing efficiency [27,36]. Figure 3 shows the optimal wavelengths selected by RF, SPA and SFS. The selection probability of each wavelength according to RF was displayed in Figure 3a. It can be found that the most of wavelengths with relatively higher selection probabilities were distributed near-infrared range, except for some in the red light region. The first selected wavelength was 747 nm with the highest selection probability of 93.6%. The selected wavelengths distributed in the near infrared range were related to the changes in tissue structure and internal chemical composition. In this study, diseased spots could be found in spectral images of near-infrared range as shown in Figure 2a, which was consistent with the results of the RF algorithm. Additionally, some wavelengths in the red light range could be due to the damage of chlorophyll after the infection of *S. sclerotiorum*. 16 optimal wavelengths selected by SPA for discriminating diseased leaves from healthy ones were shown in Figure 3b. The wavelengths in the blue and red light region associated with carotenoid and chlorophyll absorption, including 456, 464, 468, 474, 496, 631, 650, 667, 678, and 690 nm, were also selected by SPA. 710, 947, 1000, 1013, and 1023 nm were also considered as the sensitive wavelength due to the damage of mesophyll cell structure for SSR detection. Ten optimal wavelengths including those distributed in blue and red light (i.e. 455 nm, 469 nm, 667 nm, and 678 nm) were selected by SFS, which was similar with optimal wavelengths selected by SPA (Figure 3b). Compared with 452 full wavelengths, the optimal wavelengths chosen via RF, SFS, and SPA were decreased by 93.1%, 93.1% and 95.7%, respectively.



**Figure 3.** (a) The selection probability of each wavelength for SSR detection by random frog (RF). (b) The selected wavelengths for SSR detection by successive projections algorithm (SPA) and sequential forward selection (SFS), respectively.

Based on Figure 3, it could be observed that some selected wavelengths were close, indicating a high correlation between them. Therefore, it was necessary to further remove some selected wavelengths based on correlation analysis (CA). The Pearson's correlation coefficients between two sensitive wavelengths selected by RF, SPA, and SFS were calculated, and heat maps are shown in Figure 4. In this study, the correlation coefficient higher than 0.9 was regarded as a strong relationship between two wavelengths, one of which was removed based on the following criterion. The one with higher correlation coefficients with the rest wavelengths was discarded. Based on these criteria, three optimal wavelengths (458 nm, 660 nm, and 742 nm) from RF, four optimal wavelengths (456 nm, 678 nm, 710 nm, and 947 nm) from SPA and three optimal wavelengths (455 nm, 671 nm, and 747 nm) from SFS were finally selected respectively for establishing classification models for SSR detection (Table 1). The number of optimal wavelengths less than five indicated the feasibility to be applied for a rapid and low-cost SSR detection system in comparison with full wavelengths.



**Figure 4.** Correlation analysis (CA) among optimal wavelengths selected by (a) random frog (RF), (b) successive projections algorithm (SPA) and (c) sequential forward selection (SFS), respectively.

**Table 1.** Optimal wavelengths selected by RF, SFS and SPA.

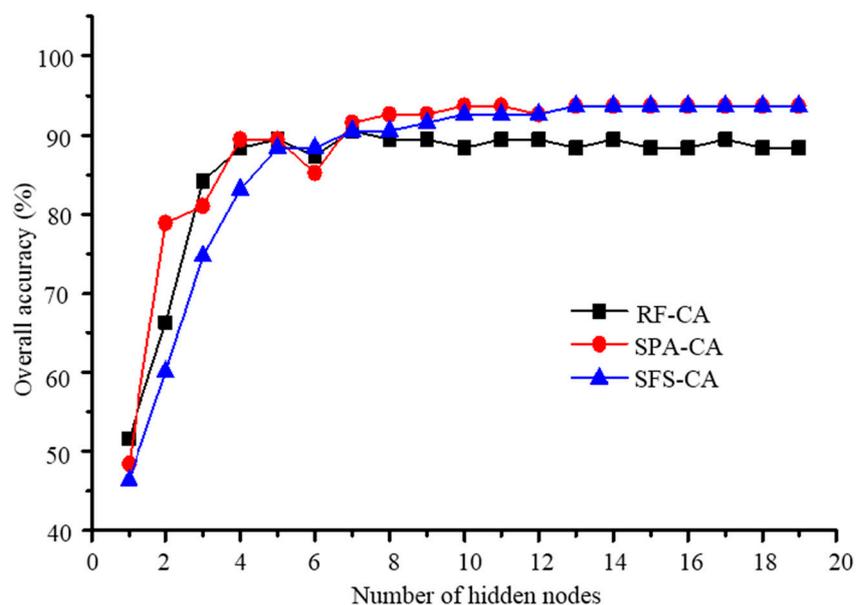
Feature Selection	Number	Wavelength (nm)
RF	16	458, 660, 662, 680, 682, 742, 743, 747, 791, 792, 797, 822, 862, 927, 952, 987
SPA	16	456, 464, 468, 474, 496, 631, 650, 667, 678, 690, 700, 710, 947, 1000, 1013, 1023
SFS	10	455, 469, 671, 682, 745, 747, 786, 801, 865, 998
RF-CA	3	458, 660, 742
SPA-CA	4	456, 678, 710, 947
SFS-CA	3	455, 671, 747

Note: NB, RF, the random frog algorithm; SPA, successive projection algorithm; SFS, sequential forward selection; CA, correlation analysis.

### 3.3. SSR Detection Based on Optimal Features

With the aim of detecting SSR, the optimal wavelengths selected by RF-CA, SPA-CA, and SFS-CA, respectively, were used as inputs of classifiers to build discriminant models. Figure 5 shows the classification performance of the ELM based on three different feature selection methods. In general, it presented a similar changing pattern using these three feature selection methods. However, the parameter of the ELM model, hidden nodes, would significantly impact the detecting performance, which indicated the importance of parameter optimization. It showed that the increasing tendency of classification accuracies with increasing hidden nodes in the range of 1 to 5, and gradually reached a steady pattern. The best detecting accuracies of ELM-SPA-CA and ELM-SFS-CA were both 93.7% with the hidden nodes of 13 and 10, respectively, while ELM-RF-CA achieved the best overall classification

accuracy of 90.5% with hidden nodes of 7. The results demonstrated that the discriminating performance using features from SPA-CA and SFS-CA was relatively better than those from RF-CA. The analysis above denoted that the three optimal wavelengths selected by SPA-CA and SFS-CA might involve relatively more important spectral signatures of SSR.



**Figure 5.** The discriminant results of the extreme learning machine (ELM) model based on optimal wavelengths selected by RF-CA, SFS-CA and SPA-CA under different hidden nodes.

The classification discriminant results could vary from different feature selection methods and model parameters. Therefore, it was also necessary to investigate the differences of discriminant ability among different classifiers, including ELM, KNN, LDA, NB and SVM. Table 2 presents detecting results from the different combination of classifiers and feature selection methods. The results from KNN, LDA, and NB were comparable and were all lower than 90% through different selected wavelengths were used. However, the ELM and SVM models could achieve outperformed detecting accuracies up to 93.7% and 94.7%, respectively, implying that the classifiers could impact the detecting result to a more extent than that of feature selection methods in this study. The similar detecting accuracy (95%) could also be found in SSR detection in oilseed rape stems with 16 optimal wavelengths as the model input [15]. In this study, only three wavelengths were used for SSR detection, which could be more suitable for high-speed online implementation due to the advantages of analyzing low-scale spectral images.

**Table 2.** The classification results of models based on the optimal wavelengths.

Feature Selection	Classifiers				
	NB	KNN	LDA	ELM	SVM
RF-CA	86.3	89.5	88.4	90.5	89.5
SPA-CA	86.3	86.3	89.5	93.7	94.7
SFS-CA	88.4	85.3	89.5	93.7	94.7

Note: NB, naïve Bayes classifier; KNN, K-nearest neighbor algorithm; LDA, linear discriminant analysis; ELM, the extreme learning machine; SVM, the support vector machine.

It was not conclusive to compare the performance of classification models based on just the overall classification accuracy. Therefore, step-by-step discrimination results were presented as the confusion matrix of the healthy and infected leaves from different combinations of classifier and feature selection

methods, respectively (Table 3). Besides the overall accuracy, a lower false positive rate (the infected sample was misjudged as the healthy one) is very critical to find all infected samples and prevent further infection. Although the overall classification accuracies of KNN, LDA, and NB were comparable (Table 2), the false negative rate and the false positive rate for healthy and infected samples were different. For KNN, the healthy samples could be all recognized with 100% accuracy for all three feature selection methods, while the discriminant ability for infected samples was quite poor with accuracies less than 80%. In comparison, the misclassification errors of LDA models were more balanced for both healthy and infected leaves. The number of incorrect/correct detections in healthy and infected leaves obtained from the NB model was similar to those from KNN. SVM models can obtain relatively better overall accuracies, however, the ELM-SFS-CA achieved the lowest false negative rate of 2.4% among 15 different combinations. In addition, the kappa value of SFS-CA-ELM was 0.873, which also indicated a good classification performance. Overall, the results demonstrated that hyperspectral imaging combined with SFS-CA-ELM has the potential for SSR detection in *A. thaliana* leaves.

**Table 3.** The classification results of models based on the optimal wavelengths.

Prediction	KNN					
	RF-CA		SPA-CA		SFS-CA	
	Infected	Healthy	Infected	Healthy	Infected	Healthy
Infected	39 (79.6%)	10 (20.4%)	32 (69.6%)	14 (30.4%)	32 (69.6%)	14 (30.4%)
Healthy	0	46 (100%)	0	49 (100%)	0	49 (100%)
Kappa value	0.791		0.702		0.702	
Prediction	LDA					
	RF-CA		SPA-CA		SFS-CA	
	Infected	Healthy	Infected	Healthy	Infected	Healthy
Infected	44 (89.8%)	5 (10.2%)	39 (88.6%)	5 (11.4%)	41 (89.1%)	5 (10.9%)
Healthy	6 (13%)	40 (87%)	5 (9.8%)	46 (90.2%)	5 (10.2%)	44 (89.8%)
Kappa value	0.768		0.788		0.789	
Prediction	NB					
	RF-CA		SPA-CA		SFS-CA	
	Infected	Healthy	Infected	Healthy	Infected	Healthy
Infected	38 (77.6%)	11 (22.4%)	32 (72.7%)	12 (27.3%)	35 (76.1%)	11 (23.9%)
Healthy	2 (4.3%)	44 (95.7%)	1 (2%)	50 (98%)	0	49 (100%)
Kappa value	0.728		0.720		0.766	
Prediction	ELM					
	RF-CA		SPA-CA		SFS-CA	
	Infected	Healthy	Infected	Healthy	Infected	Healthy
Infected	44 (91.7%)	4 (8.3%)	39 (97.5%)	1 (2.5%)	41 (97.6%)	1 (2.4%)
Healthy	5 (10.6%)	42 (89.4%)	5 (9.1%)	50 (90.9%)	5 (9.4%)	48 (90.6%)
Kappa value	0.810		0.872		0.873	
Prediction	SVM					
	RF-CA		SPA-CA		SFS-CA	
	Infected	Healthy	Infected	Healthy	Infected	Healthy
Infected	48 (88.9%)	6 (11.1%)	49 (96.1%)	2 (3.9%)	53 (94.6%)	3 (5.4%)
Healthy	4 (9.8%)	37 (90.2%)	3 (6.8%)	41 (93.2%)	2 (5.4%)	37 (94.6%)
Kappa value	0.787		0.894		0.892	

#### 4. Conclusions

In this study, we presented an investigation on SSR detection using HSI combined with multivariate analysis. To get the most feasible classification method for SSR detection in *A. thaliana* leaves, the classification performances obtained from different combinations of five classifiers and three feature selection algorithms were compared. It was found that changes of reflectance at three optimal wavelengths (455, 671 and 747 nm) selected by SFS-CA presented a good correlation with SSR. Among different combinations, SFS-CA-ELM could gain the best discriminant performance with the classification accuracy of 93.7% and the false negative rate of 2.4%. In overall, the results demonstrated that the method constructed in this study was suitable for diagnosis of SSR in *A. thaliana* leaves, and has the potential to be applied for high-throughput screening of fungal disease in plants.

**Author Contributions:** The work presented here was carried out in collaboration among all authors. J.L. conceived the idea and wrote the original draft of the paper. X.L., P.Z., N.X. and Y.H. contributed to revising the subsequent versions. We also appreciated Haiyong Weng providing valuable suggestions about the experiment design and manuscript writing.

**Funding:** This research was funded by the National Natural Science Foundation of China, grant number 31471417.

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

#### References

- Duan, Y.; Li, T.; Xiao, X.; Wu, J.; Li, S.; Wang, J.; Zhou, M. Pharmacological characteristics of the novel fungicide pyrisoxazole against *Sclerotinia sclerotiorum*. *Pestic. Biochem. Physiol.* **2018**, *149*, 61–66. [[CrossRef](#)] [[PubMed](#)]
- Melo, I.; Moretini, A.; Cassiolato, A.; Faull, J. Development of mutants of coniothyrium minitans with improved efficiency for control of *Sclerotinia sclerotiorum*. *J. Plant Protect. Res.* **2011**, *51*, 179–183. [[CrossRef](#)]
- Bardin, S.D.; Huang, H.C. Research on biology and control of *Sclerotinia* diseases in Canada. *Can. J. Plant Pathol.* **2001**, *23*, 88–98. [[CrossRef](#)]
- Kwon, J.H.; Choi, S.L.; Lee, H.S.; Shim, H.S. Stem rot of perilla caused by sclerotium rolfsii in Korea. *Korean J. Mycol.* **2012**, *40*, 177–178. [[CrossRef](#)]
- Lopez, M.M.; Bertolini, E.; Olmos, A.; Caruso, P.; Gorris, M.T.; Llop, P.; Penyalver, R.; Cambra, M. Innovative tools for detection of plant pathogenic viruses and bacteria. *Int. Microbiol.* **2003**, *6*, 233–243. [[CrossRef](#)]
- Sannakki, S.S.; Rajpurohit, V.S.; Nargund, V.B.; Arun, K.R. Leaf disease grading by machine vision and fuzzy logic. *Int. J. Comput. Technol. Appl.* **2011**, *2*, 1709–1716.
- Chaerle, L.; Hagenbeek, D.; Bruyne, E.D.; Straeten, D.V.D. Chlorophyll fluorescence imaging for disease-resistance screening of sugar beet. *Plant Cell Tissue Organ Cult.* **2007**, *91*, 97–106. [[CrossRef](#)]
- Xie, C.; Shao, Y.; Li, X.; He, Y. Detection of early blight and late blight diseases on tomato leaves using hyperspectral imaging. *Sci. Rep.* **2015**, *5*, 16564. [[CrossRef](#)]
- Oerke, E.C.; Herzog, K.; Toepfer, R. Hyperspectral phenotyping of the reaction of grapevine genotypes to *Plasmopara viticola*. *J. Exp. Bot.* **2016**, *67*, 5529–5543. [[CrossRef](#)]
- Weng, H.Y.; Lv, J.W.; Cen, H.Y.; He, M.B.; Zeng, Y.B.; Hua, S.J.; Li, H.Y.; Meng, Y.Q.; Fang, H.; He, Y. Hyperspectral reflectance imaging combined with carbohydrate metabolism analysis for diagnosis of citrus Huanglongbing in different seasons and cultivars. *Sens. Actuators B* **2018**, *275*, 50–60. [[CrossRef](#)]
- Zhou, R.Q.; Jin, J.J.; Li, Q.M.; Su, Z.Z.; Yu, X.J.; Tang, Y.; Luo, S.M.; He, Y.; Li, X.L. Early detection of *Magnaporthe oryzae*-infected Barley leaves and lesion visualization based on hyperspectral imaging. *Front. Plant Sci.* **2019**, *9*, 13. [[CrossRef](#)]
- Kong, W.W.; Yu, J.J.; Liu, F.; He, Y.; Bao, Y.D. Early diagnosis of gray mold on tomato stalks based on hyperspectral data. *Spectrosc. Spectr. Anal.* **2013**, *33*, 733–736.
- Xie, C.Q.; He, Y.; Li, X.L.; Liu, F.; Du, P.P.; Feng, L. Study of detection of SPAD value in tomato leaves stressed by grey mold based on hyperspectral technique. *Spectrosc. Spectr. Anal.* **2012**, *32*, 3324–3328.
- Zhang, C.; Feng, X.; Wang, J.; Liu, F.; He, Y.; Zhou, W. Mid-infrared spectroscopy combined with chemometrics to detect *Sclerotinia* stem rot on oilseed rape (*Brassica napus* L.) leaves. *Plant Methods* **2017**, *13*, 9. [[CrossRef](#)]
- Kong, W.; Zhang, C.; Huang, W.; Liu, F.; He, Y. Application of hyperspectral imaging to detect *Sclerotinia sclerotiorum* on oilseed rape stems. *Sensors* **2018**, *18*, 123. [[CrossRef](#)]

16. Vala, H.J.; Baxi, A. A review on Otsu image segmentation algorithm. *Int. J. Adv. Res. Comput.* **2017**, *2*, 387–389.
17. Li, H.D.; Xu, Q.S.; Liang, Y.Z. Random frog: An efficient reversible jump Markov Chain Monte Carlo-like approach for variable selection with applications to gene selection and disease classification. *Anal. Chim. Acta* **2012**, *740*, 20–26. [[CrossRef](#)] [[PubMed](#)]
18. Yun, Y.H.; Li, H.D.; Wood, L.R.E.; Fan, W.; Wang, J.J.; Cao, D.S.; Xu, Q.S.; Liang, Y.Z. An efficient method of wavelength interval selection based on random frog for multivariate spectral calibration. *Spectrochim. Acta Part A* **2013**, *111*, 31–36. [[CrossRef](#)]
19. Liu, D.; Sun, D.W.; Zeng, X.A. Recent advances in wavelength selection techniques for hyperspectral image processing in the food industry. *Food Bioprocess Technol.* **2013**, *7*, 307–323. [[CrossRef](#)]
20. Sun, K.; Geng, X.R.; Ji, L.Y. An efficient unsupervised band selection method based on an autocorrelation matrix for a hyperspectral image. *Int. J. Remote Sens.* **2014**, *35*, 7458–7476. [[CrossRef](#)]
21. Kong, W.W.; Zhang, C.; Cao, F.; Liu, F.; Luo, S.M.; Tang, Y.; He, Y. Detection of Sclerotinia Stem Rot on Oilseed Rape (*Brassica napus* L.) Leaves Using Hyperspectral Imaging. *Sensors* **2018**, *18*, 1764. [[CrossRef](#)]
22. Ravikanth, L.; Chelladurai, V.; Jayas, D.S.; White, N.D.G. Detection of Broken Kernels Content in Bulk Wheat Samples Using Near-Infrared Hyperspectral Imaging. *Agric. Res.* **2016**, *5*, 285–292. [[CrossRef](#)]
23. Cen, H.; Lu, R.; Zhu, Q.; Mendoza, F. Nondestructive detection of chilling injury in cucumber fruit using hyperspectral imaging with feature selection and supervised classification. *Postharvest Biol. Technol.* **2016**, *111*, 352–361. [[CrossRef](#)]
24. Zheng, Z.X.; Qi, L.; Ma, X.; Zhu, X.Y.; Wang, W.J. Grading method of rice leaf blast using hyperspectral imaging technology. *Trans. Chin. Soc. Agric. Eng.* **2013**, *29*, 138–144.
25. Zhao, Y.R.; Yu, K.Q.; Li, X.L.; He, Y. Detection of fungus infection on petals of rapeseed (*Brassica napus* L.) using NIR hyperspectral imaging. *Sci. Rep.* **2016**, *6*, 38879. [[CrossRef](#)]
26. López-Fandiño, J.; Quesada-Barriuso, P.; Heras, D.B.; Argüello, F. Efficient ELM-based techniques for the classification of hyperspectral remote sensing images on commodity GPUs. *IEEE J. Sel. Top Appl.* **2017**, *8*, 2884–2893. [[CrossRef](#)]
27. Jiang, J.L.; Cen, H.Y.; Zhang, C.; Lyu, X.H.; Weng, H.Y.; Xu, H.X.; He, Y. Nondestructive quality assessment of chili peppers using near-infrared hyperspectral imaging combined with multivariate analysis. *Postharvest Biol. Technol.* **2018**, *146*, 147–154. [[CrossRef](#)]
28. Bo, C.; Lu, H.; Wang, D. Spectral-spatial K-Nearest Neighbor approach for hyperspectral image classification. *Multimed. Tools Appl.* **2017**, *77*, 10419–10436. [[CrossRef](#)]
29. Bandos, T.V.; Bruzzone, L.; Camps-Valls, G. Classification of hyperspectral images with regularized linear discriminant analysis. *IEEE Trans. Geosci. Remote Sens.* **2009**, *47*, 862–873. [[CrossRef](#)]
30. Wei, X.; Liu, F.; Qiu, Z.; Shao, Y.; He, Y. Ripeness classification of astringent persimmon using hyperspectral imaging technique. *Food Bioprocess Technol.* **2014**, *7*, 1371–1380. [[CrossRef](#)]
31. Pant, B.; Pant, K.; Pardasani, K.R. Naïve Bayes classifier for classification of plant and animal miRNA. *Int. J. Comput. Theor.* **2010**, *2*, 1793–8201. [[CrossRef](#)]
32. Kumar, D.; Singh, R.; Kumar, A.; Sharma, N. An adaptive method of PCA for minimization of classification error using Naïve Bayes classifier. *Procedia Comput. Sci.* **2015**, *70*, 9–15. [[CrossRef](#)]
33. Sankaran, S.; Ehsani, R. Detection of Huanglongbing disease in citrus using fluorescence spectroscopy. *Trans. ASABE* **2012**, *55*, 313–320. [[CrossRef](#)]
34. Cen, H.; Weng, H.; Yao, J.; He, M.; Lv, J.; Hua, S.; Li, H.; He, Y. Chlorophyll fluorescence imaging uncovers photosynthetic fingerprint of citrus Huanglongbing. *Front. Plant Sci.* **2017**, *8*, 1509. [[CrossRef](#)]
35. Macho, S.; Rius, A.; Callao, M.P.; Larrechi, M.S. Monitoring ethylene content in heterophasic copolymers by near-infrared spectroscopy: Standardisation of the calibration model. *Anal. Chim. Acta* **2001**, *445*, 213–220. [[CrossRef](#)]
36. Feng, Y.Z.; Sun, D.W. Near-infrared hyperspectral imaging in tandem with partial least squares regression and genetic algorithm for non-destructive determination and visualization of *Pseudomonas* loads in chicken fillets. *Talanta* **2013**, *109*, 74–83. [[CrossRef](#)]

