

## Article

# Fungi Classification in Various Growth Stages Using Shortwave Infrared (SWIR) Spectroscopy and Machine Learning

Zhuo Liu \* and Yanjie Li 

Research Institute of Subtropical Forestry, Chinese Academy of Forestry, Hangzhou 311400, China

\* Correspondence: lzhao0824@163.com

**Abstract:** Dark septate endophytes (DSEs) fungi are beneficial to host plants with regard to abiotic stress. Here, we examined the capability of SWIR spectroscopy to classify fungus types and detected the growth stages of DSEs fungi in a timely, non-destructive and time-saving manner. The SWIR spectral data of five DSEs fungi in six growth stages were collected, and three pre-processing methods and sensitivity analysis (SA) variable selection methods were performed using a machine learning model. The results showed that the De-trending + first Derivative (DET\_FST) processing spectra combined with the support vector machine (SVM) model yielded the best classification accuracy for fungi classification at different growth stages and growth stage detection on different fungus types. The mean accuracy of generic model for fungi classification and growth stage detection are 0.92 and 0.99 on the calibration set, respectively. Seven important bands, 1164, 1456, 2081, 2272, 2278, 2448 and 2481 nm, were found to be related to the SVM fungi classification. This study provides a rapid and efficient method for the classification of fungi in different growth stages and the detection of fungi growth stage of various types of fungi and could serve as a tool for fungi study.

**Keywords:** dark septate endophytes (DSEs); fungi identification; preprocessing; support vector machine (SVM); variable selection



**Citation:** Liu, Z.; Li, Y. Fungi Classification in Various Growth Stages Using Shortwave Infrared (SWIR) Spectroscopy and Machine Learning. *J. Fungi* **2022**, *8*, 978. <https://doi.org/10.3390/jof8090978>

Academic Editor: Gary A. Strobel

Received: 13 August 2022

Accepted: 17 September 2022

Published: 19 September 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Endophytes fungi, which mainly inhabit healthy plant tissues, have been widely noticed due to their diversity and ecological significance in recent years [1]. It is a crucial component in plant root mycobiome and could reveal masked, undiscovered fungal diversity [2]. Dark septate endophytes (DSEs) are able to alleviate the host plant abiotic stresses [3] by producing melanized hyphae and microsclerotia in host plant roots which are crucial for host salt tolerance [4]. However, different DSE species could yield various influences on their host plant for abiotic stresses [5]. As a result, it is of vital importance to study the growth characteristics of DSEs and identify various DSEs species, so as to understand the influence of fungal growth on host plants.

Numerous studies have been carried out on fungal identification and growth characteristics study using traditional techniques, including enzyme assays, time lapse cameras, cytochemical processing and photo microscopy [6], which are cost- and time-consuming and often lead to further biological or chemical contamination [7], and are not suitable for high-throughput and rapid measurement. Therefore, a non-destructive and highly efficient method is needed.

Near-infrared (NIR) spectroscopy, a vibrational spectroscopy ranged 800–2500 nm, mainly used in hydrogen-containing groups such as C–H, O–H, N–H and S–H stretching modes, has received extensive attention for its great potential in many fields [8–11]. In particular, NIR has been shown to have advantages in the classification of fungi species [12] and monitoring fungi growth stages [13]. To create a robust and promising classification model, spectral pre-processing and important variable selections need to be coupled with pattern recognition algorithms, such as support vector machine (SVM) [14], random forests

(RF) [15] and multilayer perceptron (MLP) [16]. SVM, a data-driven approach for pattern recognition tasks, has extended its usage from initially binary classification to multiclass classification, which allows for a broad range of classification approaches in spectroscopy and remote sensing [17,18]. RF is also a powerful algorithm that could be used for solving both classification and regression problems [19]. It was first introduced by Breiman [15] and has been used in a broad field, including ecology [20], agriculture [21], and fungi study [22,23]. MLP is a deep learning methodology [24], widely used for classification and regression in recent years [25,26]. It mainly contains three types of layers: the input layer, the output layer, and the hidden layer. The input layer will load the input data to be processed, the output layer needs to be set based on the task of prediction or classification, and the true computational engine of the MLP is the multiple hidden layers which locate between the input and output layer. Pattern classification, recognition, prediction, and approximation are the main uses of MLP [27].

Spectral pre-processing methods are mainly used to reduce the overlapping and noise influence from the raw NIR spectra which may reduce the model accuracy. The most common spectral pre-processing methods are block scaling (BS) [28], De-trending (DET) [28], and first (FST) and second (SED) derivatives [29]. Spectra not only contain the useful information that could highly contribute to the model accuracy, but also contain irrelevant variables which may reduce the model accuracy [30]. Therefore, it is crucial to use the most important variable that is highly correlated with the target traits instead of the full spectra band. There are many variable selection algorithms available including genetic algorithm (Ga) [31], Sub-window permutation analysis (SwPA) [32], backward variable elimination (BVE) [33], significant Multivariate Correlation (sMC) algorithm [34], and sensitivity analysis (SA) [35]. Among them, the Ga, SwPA, BE, and sMC methods are mostly applied on the partial least-squares regression (PLSR) algorithm [36–38]. The SA techniques have been studied for the selection of important variables in recent years [39–41]. The SA method usually considers the effects of both individual input variables and the interaction among input variables on the output, and it works well on the SVM model for classification [42].

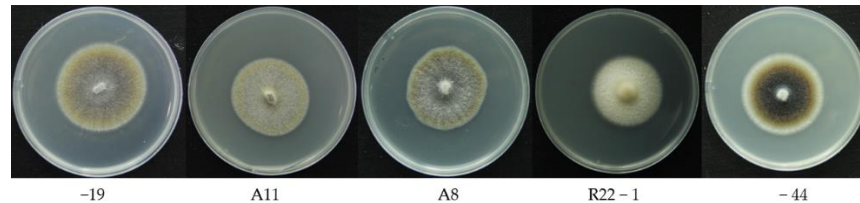
Although it has been proved that the NIR spectroscopy can be used for fungi detection, limited research outcome has been found on its use in the classification of different fungal species in a variety growth stages [12,43]. Lu, Wang, Huang, Ni, Chu, and Li [12] evaluated and classified five cereal fungi in different growth stages using Visible/Near-Infrared hyper-spectral imaging and machine learning algorithm ranged from 400 to 1000 nm and yielded a satisfactory result using SVM model coupled with the successive projection algorithm (SPA) wavelength selection method. However, to the best of our knowledge, there is no research that identifies endophytes fungi on the culture medium in different growth stages using the shortwave NIR spectroscopy (SWIR). SWIR ranges from 1100–2500 nm, which is relatively low compared to the visible and near-infrared region (VNIR, 400–1100 nm), mainly due to its lack of sufficient technology and high-cost silicon detectors [44]. However, compared to the VNIR region, SWIR has the advantages of easy penetrating the atmosphere, sensitive to the water content, high total transmittance in the atmospheric, and containing many unique absorption features that are not available in the VNIR [45].

To fill the gap, the study aims to use the SWIR technique combined with machine learning methodology to evaluate the possibility on the classification of different fungi in different growth stages, the specific objectives are: (1) comparing the advantages of different machine learning models to establish an optimal model for fungi species classification; (2) exploring the possibility of classifying different endophytes fungi in an early stage and determining the earliest classifiable time; (3) selecting the most relevant wavelengths correlated with the classification of different endophytes fungi and different growth stages on model building; (4) testing the possibility of using a mixed growth stage model with optimal informative variables for a rapid fungi classification.

## 2. Materials and Methods

### 2.1. Sample Preparation

Five strains of *Laburnicola rhizohalophila* (Figure 1) [46] purchased from the China General Microbiological Culture Collection Center (CGMCC 3.19615) were propagated on the modified Melin-Norkrans liquid medium [47] at 28 °C for 7 days to obtain the pure fungi with high vigor and consistent activity. Then, the fungi were inoculated on potato dextrose agar (PDA) and placed in a mold incubator at 22 °C, with 22 identical replicates of each group. A total of 110 spectra were obtained for each growth stage of five fungi, and a total of 660 spectra were obtained for five species.



**Figure 1.** Five strains of *Laburnicola rhizohalophila* in the same growth stage.

### 2.2. SWIR Spectrum Acquisition

To monitor the growth stages of each fungi, the SWIR spectra were collected every two days with an interval of 1 h from the first day after inoculation. The NIR spectra data were taken from the top of the Petri dish with a lid using the field-based spectrometer (LF-2500, Spectral evolution, Haverhill, MA, USA) with a handheld fiber optic contact probe. Each spectrum was averaged 24 scans and saved in a 6 nm resolution with a range of 1100 to 2500 nm.

### 2.3. Model Calibration and Validation

SVM, RF, and MLP, three common machine learning methods, were generated to obtain the optimal classification model for five endophytes fungi. Three different spectral pre-processing methods and their combination (BS, DET, FST, DET\_BS, DET\_FST) were performed on the spectrum of each sample when building classification model to reduce the effects of overlapping or light noise at different wave numbers [48]. A variable selection method called sensitivity analysis (SA) [49] was applied to extract the most relevant spectra variables from the pre-processing or raw spectra that contribute highly to the accuracy of the classification model. Data were randomly split into three sets: calibration set (60%), validation set (20%), and test set (20%). To check the stability of overall data on the classification models, each model with different pre-processing methods and variable selection methods were repeated 100 times.

### 2.4. Model Evaluation

To check the quality of the calibration models, four parameters, i.e., overall accuracy (OA), precision (P), recall (R), and the confusion matrix of each model were generated. Similar to the study of Li, et al. [50], we used true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN) calculated from the confusion matrix of each model. Since the models in this study are for multiclass classification, the confusion matrix will have a  $N \times N$  dimension where  $N$  is the number of different class labels  $C_0, C_1, \dots, C_N$  (e.g.,  $N = 5$  for fungi in this study). The performance metrics for the OA and the specific class  $C_i$  in the model can be expressed as follows [51]:

$$OA = \frac{\sum_{i=1}^N TP(C_i)}{\sum_{i=1}^N \sum_{j=1}^N C_{i,j}} \quad (1)$$

$$R = \frac{TP(C_i)}{TP(C_i) + FN(C_i)} \tag{2}$$

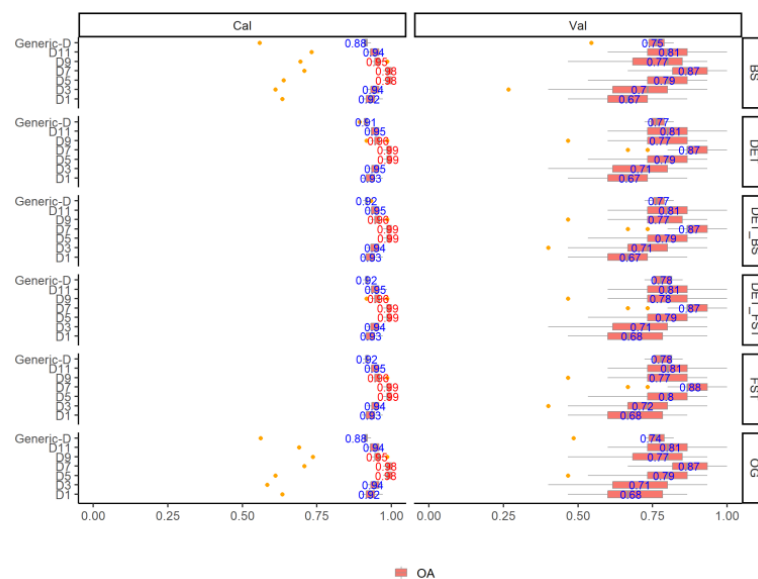
$$P = \frac{TP(C_i)}{TP(C_i) + FP(C_i)} \tag{3}$$

All of the analyses were conducted using R software [52] integrated with R studio [53]. The *caret* [54] packages have been used for SVM modeling, the *prospectr* [55] package for spectra pre-processing, the *rminer* package [56] for variable selection, and the *ggplot2* [57] for data visualization.

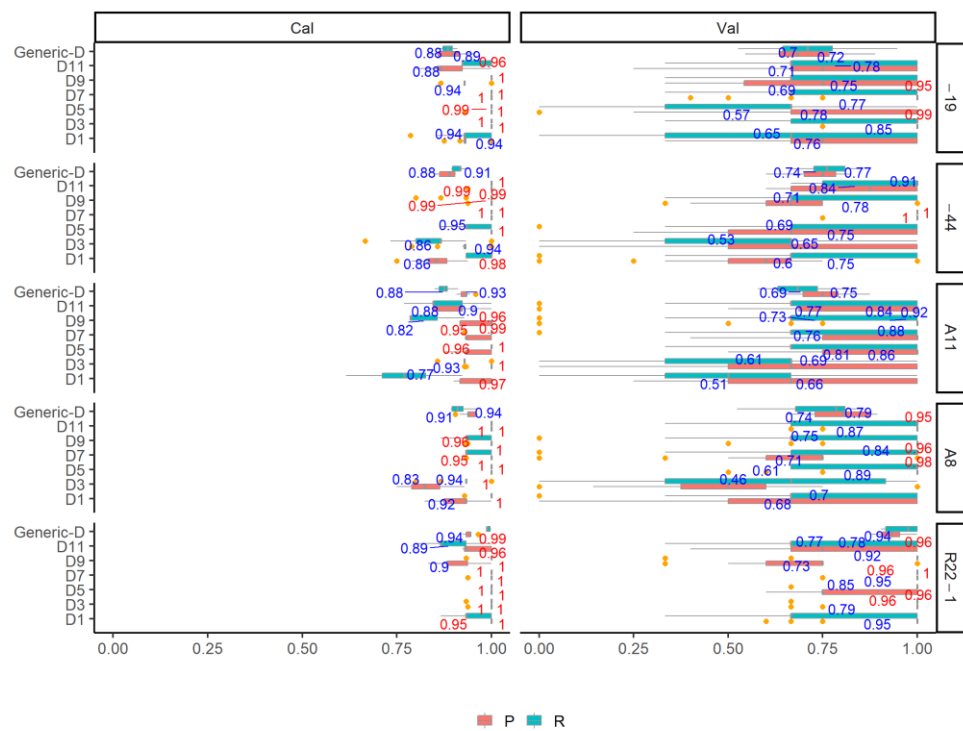
### 3. Results

#### 3.1. Fungi Classification in Different Growth Stages

Figure 2 shows the OA results of fungi classification using the SWIR spectra in six individual growth stages and generic model (generic-D). Regardless of the pre-processing methods, all the individual growth stage models and the generic model yield a promising and reliable classification accuracy with the mean  $OA_{Cal}$  higher than 0.9. The DET\_FST and FST pre-processing methods yield the highest mean  $OA_{Cal}$  with a range from 0.92 to 0.99 in all of the models, followed by the DET\_BS (range: 0.92–0.99) and DET (range: 0.91–0.99), the BS pre-processing and the original spectra yield the lowest mean  $OA_{Cal}$  (range: 0.92–0.98). When using the DET\_FST processing spectra for model classification, different growth stages influence the  $OA_{Cal}$ , with the highest mean  $OA_{Cal}$  of 0.99 in D5 and D7. D1 yields the lowest classification result compared to other individual growth stage models, but still high with the mean  $OA_{Cal}$  of 0.93. The generic-D model using of all growth stage data yield a satisfactory classification result with the highest mean  $OA_{Cal}$  of 0.92. The classification error is low in calibration and slightly high in validation from the 100 simulated models for both generic and individual growth stages. The P and R of the calibration and validation set using the DET\_FST processing spectra in all models are shown in Figure 3. All of the fungi show a stable  $P_{Cal}$  and  $R_{Cal}$  in all different growth stages, The highest mean  $P_{Cal}$  and  $R_{Cal}$  (range: 0.89–1) is found in R22-1 fungi, and A11 shows the lowest mean  $P_{Cal}$  and  $R_{Cal}$  (range: 0.77–1) compared to other fungi.



**Figure 2.** The overall accuracy (OA) Distribution (95% confidence intervals) of calibration and validation statistics from 100 simulations of models classifying the fungi in different growth stages with different spectra pre-processing methods using the SVM model.



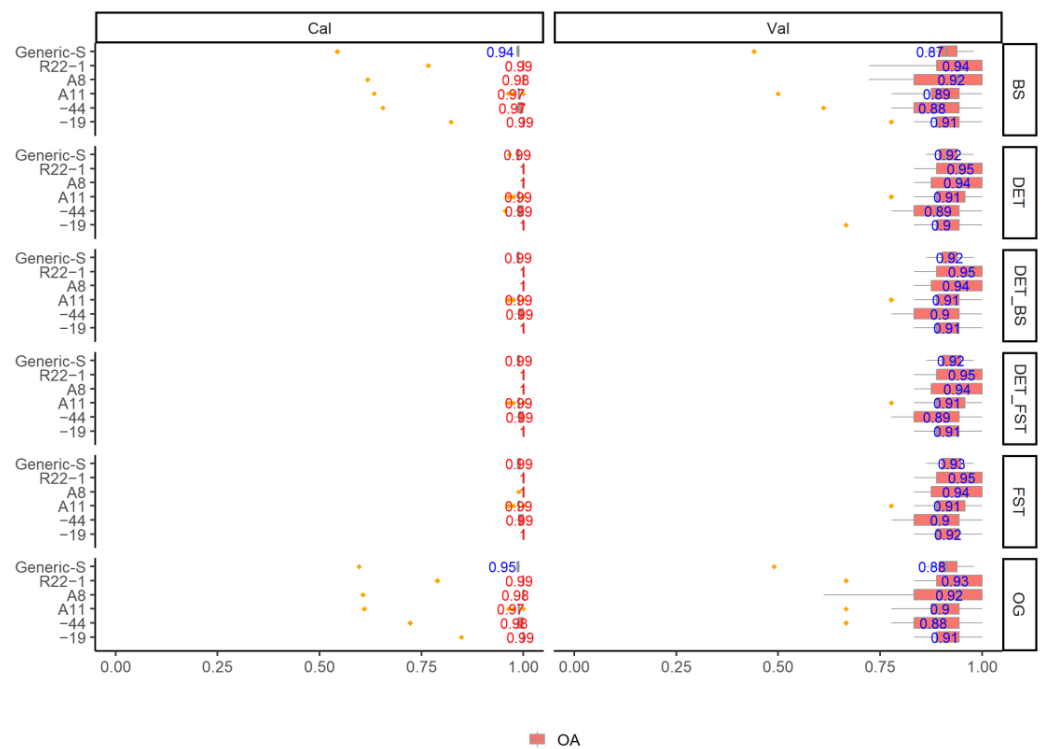
**Figure 3.** The performance of P and R in different fungi using the optimal DET\_FST pre-processing methods and SVM model in different growth stages. Values more than 0.95 are shown in red, otherwise in grey.

### 3.2. Early Detection of Fungi in Different Species

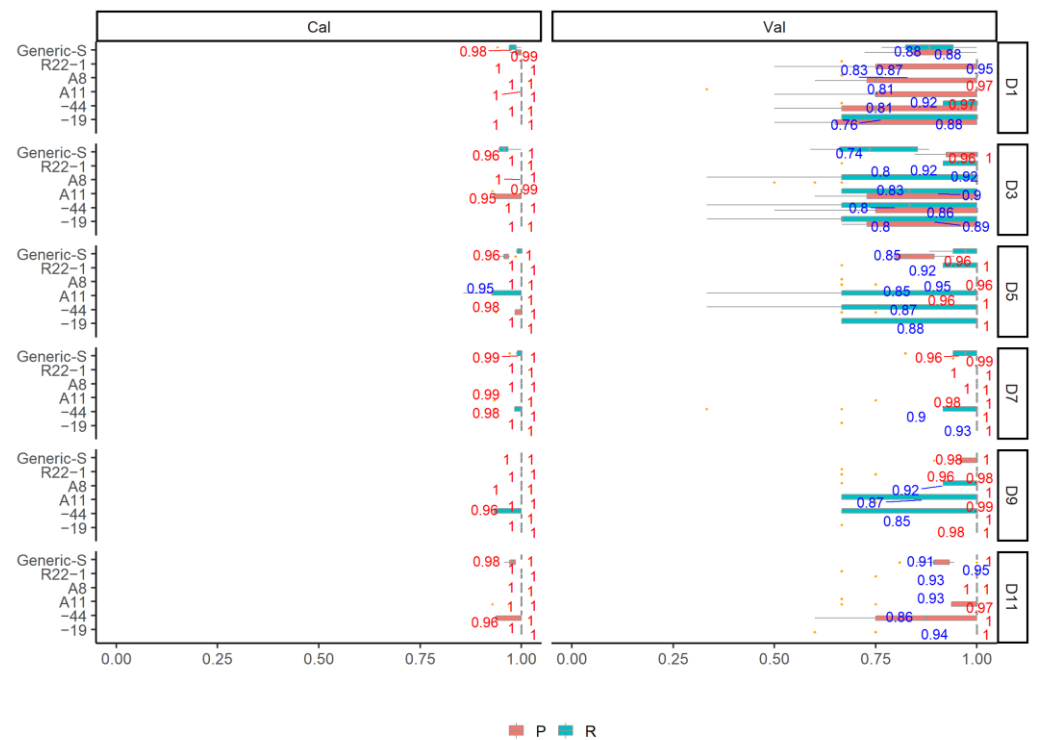
The results of classification on different growth stages using five fungi are shown in Figures 4 and 5. It can be clearly seen that the growth stages of different fungi have been successfully detected using various pre-processing methods. All of the models yield a high mean  $OA_{Cal}$  value ranging from 0.97 to 1. Pre-processing methods have slightly improved the model accuracy. DET, DET\_BS, DET\_FST, and FST yield similar mean  $OA_{Cal}$  (range: 0.99–1) in all individual sample model and generic model. Models using BS processing spectra and original spectra without processing yield promising classification  $OA_{Cal}$ , but with some  $OA_{Cal}$  outliers from the 100 simulated model. Similarly, when choosing the DET\_FST processing spectra to generate models, the P and R are high in both calibration and validation set in different types of fungi. All of the  $P_{Cal}$  and  $R_{Cal}$  in six growth stages are higher than 0.95. In addition, the generic model (Generic-S) for detecting the different growth stages using all types of fungi yield a high and promising  $P_{Cal}$  and  $R_{Cal}$  result for all growth stages, with  $P_{Cal}$  and  $R_{Cal}$  mean value ranging from 0.96 to 1. The data show a stable result with small error in OA,  $P_{Cal}$  and  $R_{Cal}$  for 100 simulated calibration models.

### 3.3. Model Comparison

After selecting the best pre-processing method using SVM model, other machine learning methodologies were also applied and compared to the SVM model using DET\_FST processing spectra. The OA results show that other machine learning models also provide a stable and promising result in both fungi classification and early detection (Table 1). The OA value of  $RF_{Cal}$  and  $MLP_{Cal}$  are 0.91 and 0.91 in fungi classification and 0.99 and 0.98 in early detection, respectively.



**Figure 4.** The overall accuracy (OA) Distribution (95% confidence intervals) of calibration and validation statistics from 100 simulations of models classifying the fungi growth stages in different types of fungi with different spectra pre-processing methods using the SVM model.



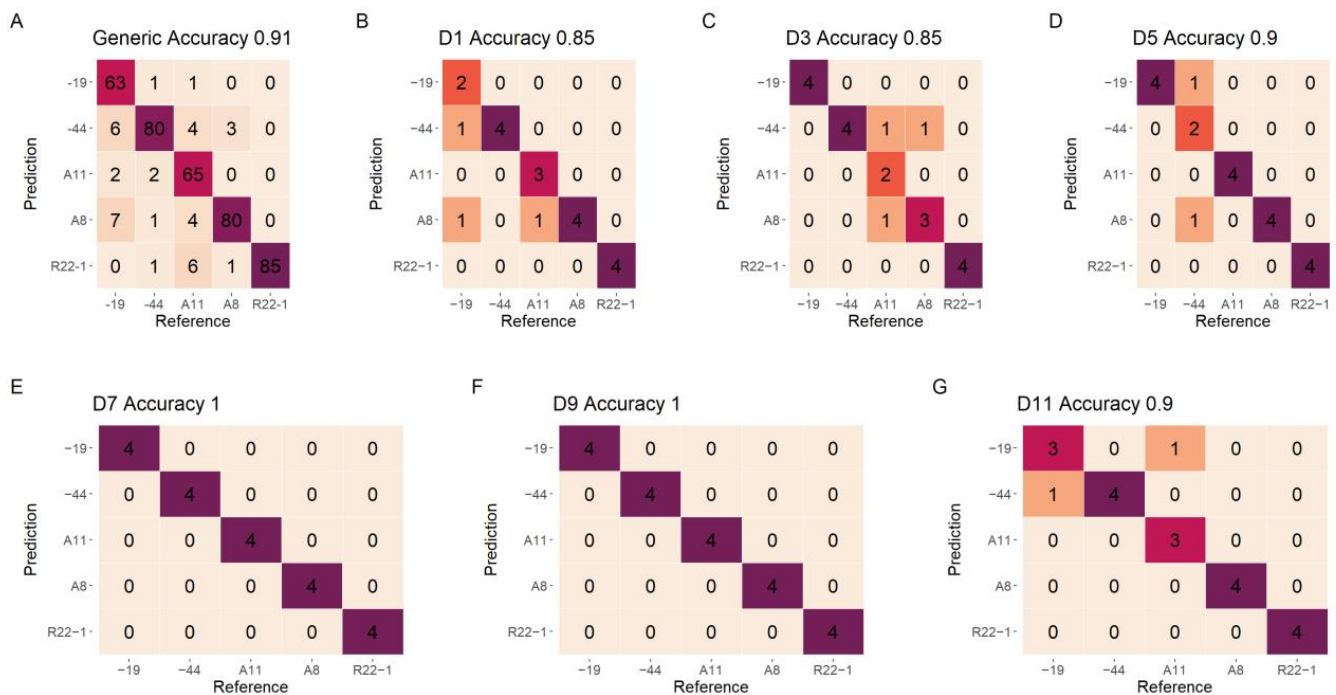
**Figure 5.** The performance of P and R in different growth stage using the optimal DET\_FST pre-processing methods and SVM model on different types of fungi. Values more than 0.95 are shown in red, otherwise in grey.

**Table 1.** The mean OA accuracy of 100 simulated generic models for the fungi classification and early detection using SVM, RF, and MLP methodology and the optimal DET\_FST processing spectra.

Model Types	Generic Model	OA	
		Cal	Val
Fungi classification	SVM	0.92	0.76
	RF	0.91	0.75
	MLP	0.91	0.77
Early detection	SVM	0.99	0.92
	RF	0.99	0.92
	MLP	0.98	0.91

3.4. Model Evaluation for Fungi Classification

The confusion matrixes for fungi classification of various growth stages and the generic model on the test data set are plotted in Figure 6. The model using the D1 and D3 data shows the lowest result on test data, with only 0.85 accuracy. As the fungi grow, the classification accuracy shows a growing trend, with 1 classification accuracy on the D7 and D9, respectively. The generic model using all growth stages data show slightly low classification accuracy on the test set (0.91). The most unstable fungi for classification is -19, which has misclassified six samples to -44, two samples to A11, and seven samples to A8.

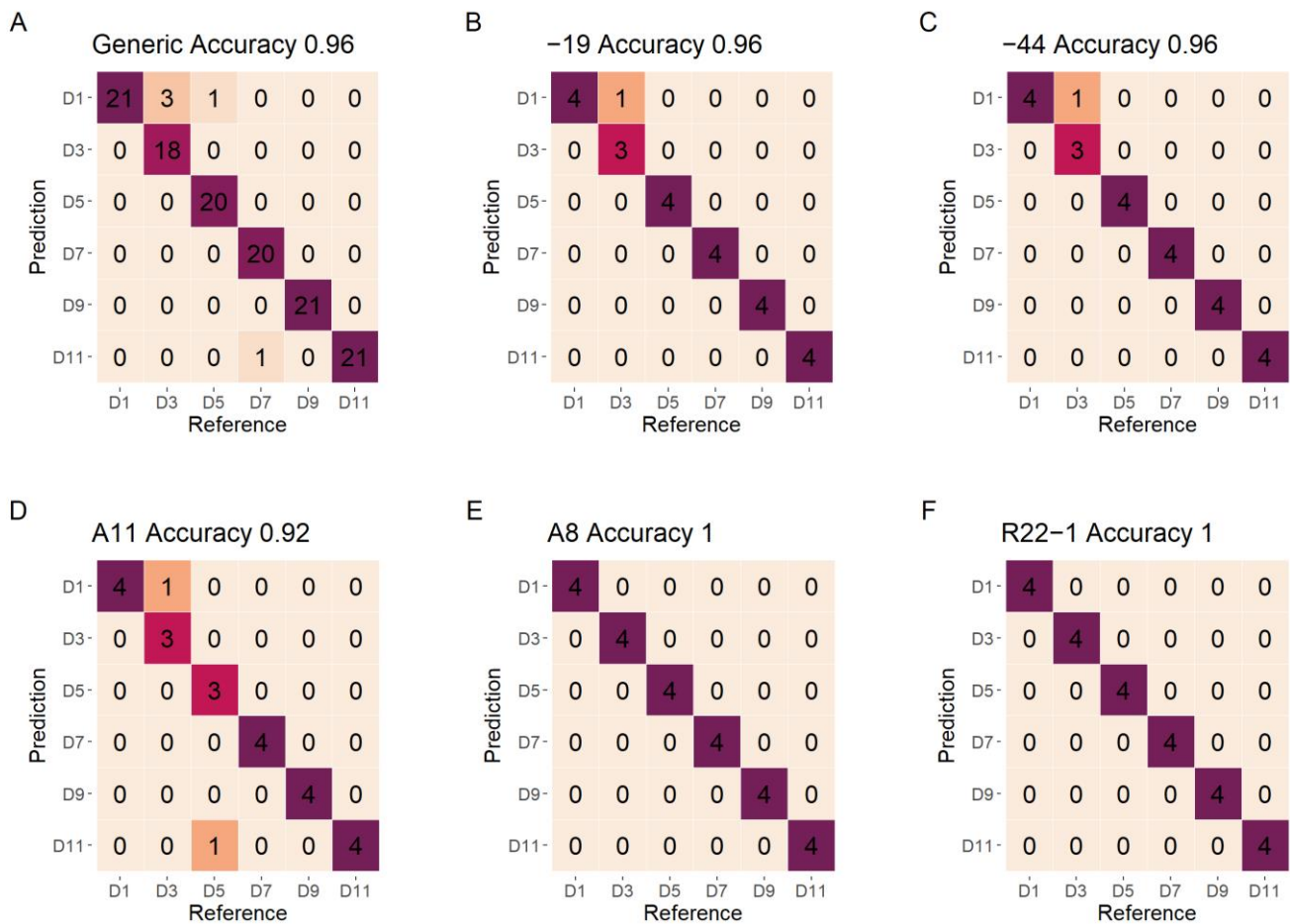


**Figure 6.** The confusion matrixes of fungi classification for test data of individual growth stages and the generic growth stages model using DET\_FST processing spectra. (A) The generic model; (B) D1 model; (C) D3 model; (D) D5 model; (E) D7 model; (F) D9 model; (G) D11 model. The color changed from yellow to fuchsia represent different classified numbers.

3.5. Model Evaluation for Early-Stage Detection

The confusion matrixes for the classification of different growth stages using various fungi and generic fungi models on the test set are shown in Figure 7. All of the models show a high classification result with accuracy higher than 0.92. Different types of fungi show a slight difference on the detection of growth stages, The R22-1 and A8 fungi show an accuracy of 1, followed by the -44 and -19 fungi (accuracy of 0.96). The A11 fungi

show the lowest classification accuracy of 0.92 on test data. The generic model for different growth stages shows a high and reliable accuracy on the test data (0.96).

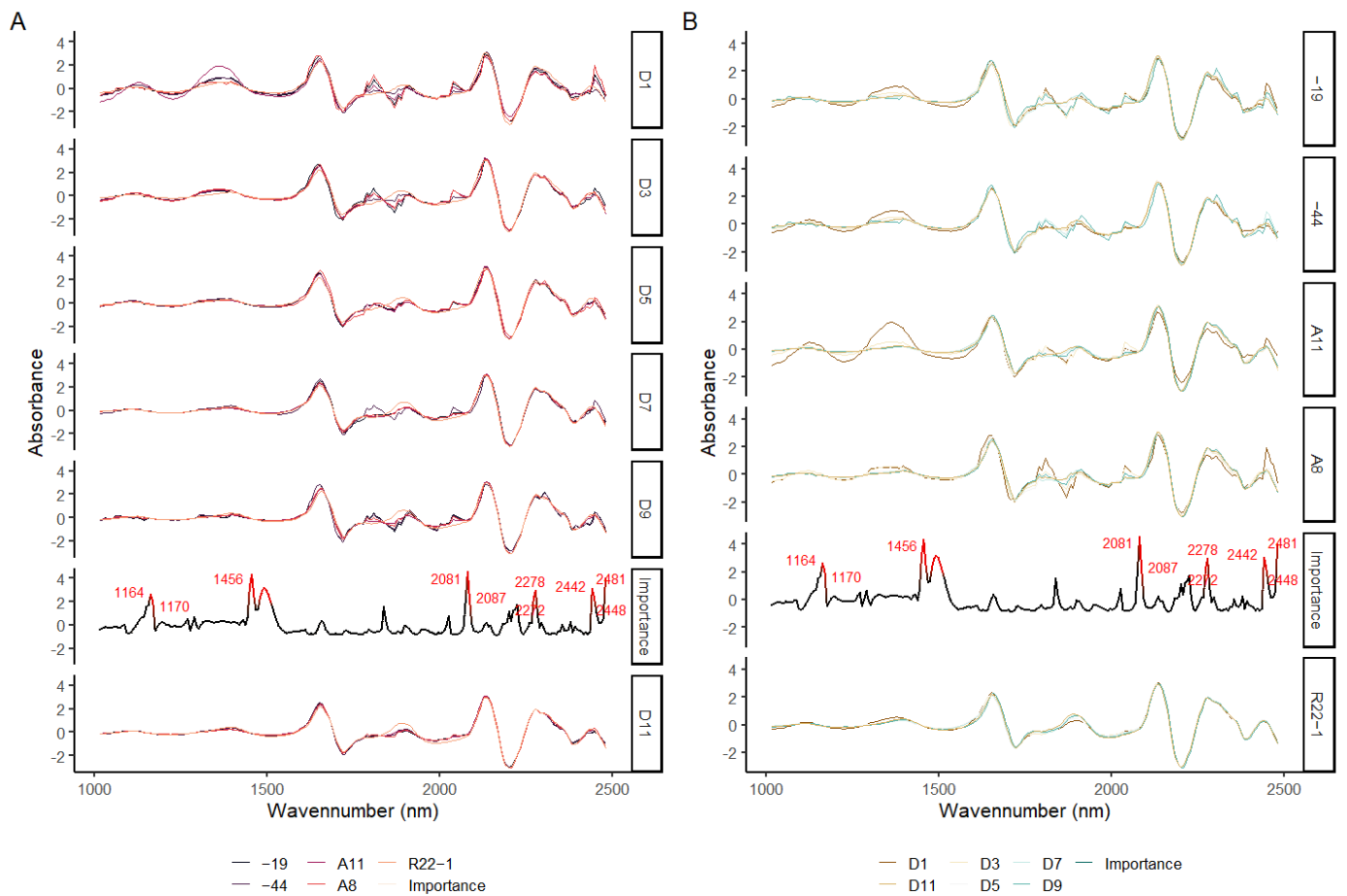


**Figure 7.** The confusion matrixes of growth stages detection for test data of individual fungi and the generic fungi model using DET\_FST processing spectra. (A) The generic model; (B) -19 model; (C) -44 model; (D) A11 model; (E) A8 model; (F) R22-1 model. The color changed from yellow to fuchsia represent the different classified numbers.

### 3.6. Importance Variable Selection

The average of DET\_FST processing spectra for five fungi and six growth stages and the important variables for these two kinds of SVM models are shown in Figure 8. For the averaged spectra of fungi samples, the region around 1800 to 2000 nm and 1300 nm are found to have various changes for different fungi types with growth (Figure 8A), and similar results are also found in the growth date of different fungi (Figure 8B). The top 10% important spectral variables are labeled in red. Similar important variables are found for both fungi classification model and growth stage model, which are 1164, 1456, 2081, 2272, 2278, 2442, 2448, and 2481 nm, respectively. The important variables found by the SA from SVM model are located slightly differently to the change area of DET\_FST processing spectra.





**Figure 8.** The DET\_FST processing spectra of different fungus types (A) and growth stage (B); the importance values of wavelengths for fungi classification (A) and growth stage (B) detection using SVM model and DET\_FST processing spectra.

#### 4. Discussion

The SWIR technology, in our study, has been successfully applied to the DSEs fungi classification and the early growth stage detection. It shows a promising and reliable result using machine learning methodology. To the best of our knowledge, it is the first report on the DSEs fungi classification using SWIR spectra. Similar study on fungi classification and growth stage detection using NIR spectra was made by Lu, Wang, Huang, Ni, Chu, and Li [12], who reported that the cereal fungi can be successfully classified in all growth stages using the SVM model and successive projection algorithm variable selection methods combined with the Visible/Near-Infrared (Vis/NIR) hyperspectral imaging, with an overall accuracy of more than 95.87% on all growth stages and 98.89% on the first growth stage, which is close to our study (with a mean accuracy of 0.92 on all growth stages and 0.99 on the first growth day). In addition, SWIR hyperspectral imaging has also shown a promising ability on the *Fusarium* spp. Fungi detection using the partial least squares discriminant analysis (PLS-DA) models yielded a reasonable accuracy ranging from 78% to 100% on the train and test data sets [58]. Our results showed that different growth stages influence the fungi classification, with an OA value ranging from 0.92 to 0.99, supported by the study of Sun, Gu, Wang, Huang, Wei, Zhang, Tu, and Pan [13], who found that the best accuracy in the classification of three spoilage fungi is 97.5% for the test dataset at 36 h using the 400 to 1000 nm NIR hyperspectral imaging and PLS-DA model.

Three spectra pre-processing methods (BS, DET, FST) and their combination (DET\_BS, DET\_FST) were applied in our study to check the influence of processing spectra on the SVM classification accuracy. Our results found that BS showed no improvement on the model accuracy compared to the original raw spectra with more outliers when simulated

100 times, and that DET, DET\_BS, DET\_FST showed high improvement on both fungi classification and early detection models, with the highest OA reaching 1. Our results are different to what were reported in the study by Wu, et al. [59] who found the highest classification accuracy of waxy wheats was gained by using the raw non-processing SWIR spectra data combined with the SVM model, reaching 98.51%. Different applications may lead to different results. Xu, et al. [60] found the second-order derivatives and standard normal variate (SNV) processing spectra can largely enhance the classification of regular and aged shiitake using SVM model. Yao, et al. [61] found that the combination of different pre-processing methods (second derivative (2D), Savitzky–Golay (SG) filter and SNV) could improve the SVM model accuracy on the discrimination of five Boletaceae mushroom species, which is similar to our study where the combination of DET and first derivative yield the optimal classification results on the SVM model. Therefore, the spectral preprocessing methods may need to be performed based on the needs of practical applications [62].

Different types of machine learning methodologies including SVM, RF, and MLP have been compared in our study to find the most suitable model for fungi classification and growth stage detection. It has shown that all of the machine learning methodologies can achieve high classification aspects. All of the classification accuracy on the calibration set for the two kinds of models are above 0.9. Though all these three machine learning models can be applied to the fungi classification and growth stage detection, since the SVM model performed slightly better than the other two methods, we selected the SVM as the final model in our study, which is supported by Castro, et al. [63] who compared three classifiers, decision tree (DT), SVM, and K-nearest neighbor, and found that the SVM model yielded the best performance on the classification of different growth stages of coffee rust infection, using the Vis/NIR (400–1000 nm) hyperspectral imaging. In addition, the detection and classification of corn (*Zea mays* L.) seeds also yielded a high classification accuracy of 96.46% when SVM was applied using SWIR hyperspectral imaging [64].

The region of SWIR (1100–2500 nm) has been well studied for whether it is highly sensitive to the target water, lignin, cellulose, and proteins [13,65]. Seven prominent bands around 1164, 1456, 2081, 2272, 2278, 2448, and 2481 nm were found by SA variable selection method in our study which has highly contributed to the SVM classification model. These bands found to be related to the model classification are not exactly equal to the findings of other studies, maybe because the different spectral processing varied the number of bands in the spectral range [11]. However, similar bands can be found by other studies. The bands around 1164, 1456, and 2081 nm have been reported to be highly related to the carbohydrates [6,66], the bands at 2272, 2278 are mostly correlated to the stretching or deformation vibration of C-H or C-C, and the bands at 2448 and 2481 nm have been reported to be related to protein contents [11,67]. All of these band regions were responsible for the differences between fungi species and growth stages.

## 5. Conclusions

In our study, the growth stages of five different DSEs fungi inoculated on the PDA were classified using the SWIR spectroscopy based on machine learning models. It has been shown that the SWIR spectroscopy can be successfully used for the fungi classification in different growth stages and the growth stage detection in different fungi. Generic model for the two approaches also provides a reliable and promising classification accuracy. Furthermore, the hyperspectral imaging with SWIR regions should be studied in future to yield a better result for visualization. Our classification of fungi and detection of fungi growth stages using the SWIR spectroscopy is fast and accurate which could enhance the fungi study in the future.

**Author Contributions:** Z.L.: Conceptualization, Formal analysis, methodology, writing—original draft, Y.L.: Supervision, Methodology, Writing—review and editing. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data mentioned in this paper are available on request from the corresponding author.

**Conflicts of Interest:** The authors declare that they have not known competing financial interest or personal relationships that could have appeared to influence the work reported in this paper.

## References

- Rodríguez, R.; White, J., Jr.; Arnold, A.E.; Redman, A.R.A. Fungal endophytes: Diversity and functional roles. *New Phytol.* **2009**, *182*, 314–330. [[CrossRef](#)] [[PubMed](#)]
- Peay, K.G.; Kennedy, P.G.; Talbot, J.M. Dimensions of biodiversity in the Earth mycobiome. *Nat. Rev. Microbiol.* **2016**, *14*, 434–447. [[CrossRef](#)] [[PubMed](#)]
- Qin, Y.; Pan, X.; Kubicek, C.; Druzhinina, I.; Chenthamara, K.; Labbé, J.; Yuan, Z. Diverse plant-associated pleosporalean fungi from saline areas: Ecological tolerance and nitrogen-status dependent effects on plant growth. *Front. Microbiol.* **2017**, *8*, 158. [[CrossRef](#)] [[PubMed](#)]
- Gostinčar, C.; Grube, M.; De Hoog, S.; Zalar, P.; Gunde-Cimerman, N. Extremotolerance in fungi: Evolution on the edge. *FEMS Microbiol. Ecol.* **2009**, *71*, 2–11. [[CrossRef](#)] [[PubMed](#)]
- Yuan, Z.; Druzhinina, I.S.; Labbé, J.; Redman, R.; Qin, Y.; Rodríguez, R.; Zhang, C.; Tuskan, G.A.; Lin, F. Specialized microbiome of a halophyte and its role in helping non-host plants to withstand salinity. *Sci. Rep.* **2016**, *6*, 32467. [[CrossRef](#)]
- Williams, P.J.; Geladi, P.; Britz, T.J.; Manley, M. Growth characteristics of three *Fusarium* species evaluated by near-infrared hyperspectral imaging and multivariate image analysis. *Appl. Microbiol. Biotechnol.* **2012**, *96*, 803–813. [[CrossRef](#)]
- Yao, H.; Hruska, Z.; Kincaid, R.; Brown, R.L.; Cleveland, T.E. Differentiation of toxigenic fungi using hyperspectral imagery. *Sens. Instrum. Food Qual. Saf.* **2008**, *2*, 215–224. [[CrossRef](#)]
- Mohd Hilmi Tan, M.I.S.; Jamlos, M.F.; Omar, A.F.; Dzaharudin, F.; Chalermwisutkul, S.; Akkaraekthalin, P. Ganoderma boninense disease detection by near-infrared spectroscopy classification: A review. *Sensors* **2021**, *21*, 3052. [[CrossRef](#)]
- Miao, X.; Miao, Y.; Tao, S.; Liu, D.; Chen, Z.; Wang, J.; Huang, W.; Yu, Y. Classification of rice based on storage time by using near infrared spectroscopy and chemometric methods. *Microchem. J.* **2021**, *171*, 106841. [[CrossRef](#)]
- Zhang, Y.; Luan, Q.; Jiang, J.; Li, Y. Prediction and utilization of malondialdehyde in exotic pine under drought stress using near-infrared spectroscopy. *Front. Plant Sci.* **2021**, *12*, 735275. [[CrossRef](#)]
- Li, Y.; Sun, H.; Tomasetto, F.; Jiang, J.; Luan, Q. Spectrometric Prediction of Nitrogen Content in Different Tissues of Slash Pine Trees. *Plant Phenomics* **2022**, *2022*, 9892728. [[CrossRef](#)] [[PubMed](#)]
- Lu, Y.; Wang, W.; Huang, M.; Ni, X.; Chu, X.; Li, C. Evaluation and classification of five cereal fungi on culture medium using Visible/Near-Infrared (Vis/NIR) hyperspectral imaging. *Infrared Phys. Technol.* **2020**, *105*, 103206. [[CrossRef](#)]
- Sun, Y.; Gu, X.; Wang, Z.; Huang, Y.; Wei, Y.; Zhang, M.; Tu, K.; Pan, L. Growth simulation and discrimination of *Botrytis Cinerea*, *Rhizopus Stolonifer* and *Colletotrichum Acutatum* using hyperspectral reflectance imaging. *PLoS ONE* **2015**, *10*, e0143400. [[CrossRef](#)]
- Vapnik, V. *The Nature of statistical Learning Theory*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 1999.
- Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
- Atkinson, P.M.; Tatnall, A.R. Introduction neural networks in remote sensing. *Int. J. Remote Sens.* **1997**, *18*, 699–709. [[CrossRef](#)]
- Pérez-Cruz, F.; Artés-Rodríguez, A. Puncturing multi-class support vector machines. In Proceedings of the International Conference on Artificial Neural Networks, Madrid, Spain, 28–30 August 2002; pp. 751–756.
- Raghavendra, N.S.; Deka, P.C. Support vector machine applications in the field of hydrology: A review. *Appl. Soft Comput.* **2014**, *19*, 372–386. [[CrossRef](#)]
- Biau, G. Analysis of a random forests model. *J. Mach. Learn. Res.* **2012**, *13*, 1063–1095.
- Cutler, D.R.; Edwards, T.C., Jr.; Beard, K.H.; Cutler, A.; Hess, K.T.; Gibson, J.; Lawler, J.J. Random forests for classification in ecology. *Ecology* **2007**, *88*, 2783–2792. [[CrossRef](#)]
- Ok, A.O.; Akar, O.; Gungor, O. Evaluation of random forest method for agricultural crop classification. *Eur. J. Remote Sens.* **2012**, *45*, 421–432. [[CrossRef](#)]
- Wang, C.; Zhang, Y.; Han, S. Its2vec: Fungal species identification using sequence embedding and random forest classification. *BioMed Res. Int.* **2020**, *2020*, 2468789. [[CrossRef](#)]
- Chauhan, M.D. Detection of maize disease using random forest classification algorithm. *Turk. J. Comput. Math. Educ. (TURCO-MAT)* **2021**, *12*, 715–720.
- Delashmit, W.H.; Manry, M.T. Recent developments in multilayer perceptron neural networks. In Proceedings of the Seventh Annual Memphis Area Engineering and Science Conference, Memphis, TN, USA, 11 May 2005.
- Heidari, A.A.; Faris, H.; Mirjalili, S.; Aljarah, I.; Mafarja, M. *Ant Lion Optimizer: Theory, Literature Review, and Application in Multi-layer Perceptron Neural Networks*; Springer: Cham, Switzerland, 2020; pp. 23–46.

26. Salehi, M.; Farhadi, S.; Moieni, A.; Safaie, N.; Ahmadi, H. Mathematical modeling of growth and paclitaxel biosynthesis in *Corylus avellana* cell culture responding to fungal elicitors using multilayer perceptron-genetic algorithm. *Front. Plant Sci.* **2020**, *11*, 1148. [CrossRef] [PubMed]
27. Ramchoun, H.; Ghanou, Y.; Ettaouil, M.; Janati Idrissi, M.A. Multilayer Perceptron: Architecture Optimization and Training. 2016. Available online: <https://reunir.unir.net/handle/123456789/11569> (accessed on 12 August 2022).
28. Barnes, R.; Dhanoa, M.S.; Lister, S.J. Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra. *Appl. Spectrosc.* **1989**, *43*, 772–777. [CrossRef]
29. Cameron, D.G.; Moffatt, D.J. A generalized approach to derivative spectroscopy. *Appl. Spectrosc.* **1987**, *41*, 539–544. [CrossRef]
30. Yang, S.-Y.; Han, Y.; Chang, Y.-S.; Park, J.-H.; Park, Y.; Chung, H.; Yeo, H. Classification of the hot air heat treatment degree of larch wood using a multivariate analysis of near-infrared spectroscopy. *J. Wood Sci.* **2018**, *64*, 220–225. [CrossRef]
31. Babatunde, O.H.; Armstrong, L.; Leng, J.; Diepeveen, D. A Genetic Algorithm-Based Feature Selection. *Br. J. Math. Comput. Sci.* **2014**, *4*, 889–905.
32. Li, H.-D.; Zeng, M.-M.; Tan, B.-B.; Liang, Y.-Z.; Xu, Q.-S.; Cao, D.-S. Recipe for revealing informative metabolites based on model population analysis. *Metabolomics* **2010**, *6*, 353–361. [CrossRef]
33. Frank, I.E. Intermediate least squares regression method. *Chemom. Intell. Lab. Syst.* **1987**, *1*, 233–242. [CrossRef]
34. Tran, T.N.; Afanador, N.L.; Buydens, L.M.; Blanchet, L. Interpretation of variable importance in partial least squares with significance multivariate correlation (sMC). *Chemom. Intell. Lab. Syst.* **2014**, *138*, 153–160. [CrossRef]
35. Pannell, D.J. Sensitivity analysis: Strategies, methods, concepts, examples. *Agric. Econ.* **1997**, *16*, 139–152. [CrossRef]
36. Mehmood, T.; Liland, K.H.; Snipen, L.; Sæbø, S. A review of variable selection methods in partial least squares regression. *Chemom. Intell. Lab. Syst.* **2012**, *118*, 62–69. [CrossRef]
37. Guzmán, E.; Baeten, V.; Pierna, J.A.F.; García-Mesa, J.A. Application of low-resolution Raman spectroscopy for the analysis of oxidized olive oil. *Food Control.* **2011**, *22*, 2036–2040. [CrossRef]
38. Li, Y.; Sharma, M.; Altaner, C.; Cookson, L.J. An approach to quantify natural durability of *Eucalyptus bosistoana* by near infrared spectroscopy for genetic selection. *Ind. Crops Prod.* **2020**, *154*, 112676. [CrossRef]
39. Plischke, E. An effective algorithm for computing global sensitivity indices (EASI). *Reliab. Eng. Syst. Saf.* **2010**, *95*, 354–360. [CrossRef]
40. Zouhri, W.; Homri, L.; Dantan, J.-Y. Handling the impact of feature uncertainties on SVM: A robust approach based on Sobol sensitivity analysis. *Expert Syst. Appl.* **2022**, *189*, 115691. [CrossRef]
41. Şandru, E.-D.; David, E.; Kovacs, I.; Buzo, A.; Burileanu, C.; Pelz, G. Modeling the Dependency of Analog Circuit Performance Parameters on Manufacturing Process Variations With Applications in Sensitivity Analysis and Yield Prediction. *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.* **2021**, *41*, 129–142. [CrossRef]
42. Jiang, S.-F.; Fu, D.-B.; Wu, S.-Y. Structural Reliability Assessment by Integrating Sensitivity Analysis and Support Vector Machine. *Math. Probl. Eng.* **2014**, *2014*, 586191. [CrossRef]
43. Kimuli, D.; Wang, W.; Jiang, H.; Zhao, X.; Chu, X.; Yang, Y.; Jia, B. Evaluation of growth characteristics of a mixed culture of toxigenic fungi by Visible/Near-infrared hyperspectral imaging. In Proceedings of the 2018 ASABE Annual International Meeting, Detroit, MI, USA, 29 July–1 August 2018; p. 1.
44. Herrmann, I.; Karnieli, A.; Bonfil, D.; Cohen, Y.; Alchanatis, V. SWIR-based spectral indices for assessing nitrogen content in potato fields. *Int. J. Remote Sens.* **2010**, *31*, 5127–5143. [CrossRef]
45. Hansen, M.P.; Malchow, D.S. Overview of SWIR detectors, cameras, and applications. In Proceedings of the Thermosense Xxx, Orlando, FL, USA, 17 March 2008; pp. 94–104.
46. Yuan, Z.; Druzhinina, I.S.; Wang, X.; Zhang, X.; Peng, L.; Labbé, J. Insight into a highly polymorphic endophyte isolated from the roots of the halophytic seepweed *Suaeda salsa*: *Laburnicola rhizohalophila* sp. nov. (Didymosphaeriaceae, Pleosporales). *Fungal Biol.* **2020**, *124*, 327–337. [CrossRef]
47. Marx, D.H. The Influence of ectotrophic mycorrhizal fungi on the resistance of pine roots to pathogenic infections. V. Resistance of mycorrhizae to infection by vegetative mycelium of *Phytophthora cinnamomi*. *Phytopathology* **1970**, *60*, 1472–1473. [CrossRef]
48. Alchanatis, V.; Schmilovitch, Z.; Meron, M.J.P.A. In-field assessment of single leaf nitrogen status by spectral reflectance measurements. *Precis. Agric.* **2005**, *6*, 25–39. [CrossRef]
49. Cortez, P.; Embrechts, M.J. Using sensitivity analysis and visualization techniques to open black box data mining models. *Inf. Sci.* **2013**, *225*, 1–17. [CrossRef]
50. Li, Y.; Al-Sarayreh, M.; Irie, K.; Hackell, D.; Bourdot, G.; Reis, M.M.; Ghamkhar, K. Identification of weeds based on hyperspectral imaging and machine learning. *Front. Plant Sci.* **2021**, *11*, 611622. [CrossRef]
51. Markoulidakis, I.; Rallis, I.; Georgoulas, I.; Kopsiaftis, G.; Doulamis, A.; Doulamis, N. Multiclass Confusion Matrix Reduction Method and Its Application on Net Promoter Score Classification Problem. *Technologies* **2021**, *9*, 81. [CrossRef]
52. Team, R.C. R: A Language and Environment for Statistical Computing. Available online: <https://www.r-project.org> (accessed on 1 May 2022).
53. Allaire, J. *Integrated Development Environment*; RStudio: Boston, MA, USA, 2012; Volume 770, p. 394.
54. Kuhn, M.; Caret: Classification and Regression Training. Astrophysics Source Code Library. Available online: <https://ui.adsabs.harvard.edu/abs/2015ascl.soft05003K> (accessed on 7 May 2022).

55. Stevens, A.; Ramirez-Lopez, L. An Introduction to the Prospectr Package. 2014. Available online: <https://cran.r-project.org/web/packages/prospectr/vignettes/prospectr-intro.pdf> (accessed on 1 May 2022).
56. Cortez, P.; Rminer: Data Mining Classification and Regression Methods. R Package Version 1.4.2. Available online: <https://cran.r-project.org/web/packages/rminer/rminer.pdf> (accessed on 7 May 2022).
57. Wickham, H. *Ggplot2: Elegant Graphics for Data Analysis*; Springer: Berlin/Heidelberg, Germany, 2016.
58. Williams, P.J.; Geladi, P.; Britz, T.J.; Manley, M. Near-infrared (NIR) hyperspectral imaging and multivariate image analysis to study growth characteristics and differences between species and strains of members of the genus *Fusarium*. *Anal. Bioanal. Chem.* **2012**, *404*, 1759–1769. [[CrossRef](#)] [[PubMed](#)]
59. Wu, Y.; Yun, Y.; Chen, J.; Liu, D. Discrimination of Waxy Wheats Using Near-Infrared Hyperspectral Spectroscopy. *Food Anal. Methods* **2021**, *14*, 1704–1713. [[CrossRef](#)]
60. Xu, L.; Fu, X.-S.; Cai, C.-B.; She, Y.-B. The Feasibility of Using Near Infrared Spectroscopy for Rapid Discrimination of Aged Shiitake Mushroom (*Lentinula edodes*) after Long-Term Storage. *J. Chem.* **2015**, *2015*, 692983. [[CrossRef](#)]
61. Yao, S.; Li, J.; Li, T.; Liu, H.; Wang, Y. Discrimination of Boletaceae mushrooms based on data fusion of FT-IR and ICP–AES combined with SVM. *Int. J. Food Prop.* **2018**, *21*, 255–266. [[CrossRef](#)]
62. Wu, H.; Song, Z.; Niu, X.; Liu, J.; Jiang, J.; Li, Y. Classification of Toona sinensis Young Leaves Using Machine Learning and UAV-Borne Hyperspectral Imagery. *Front. Plant Sci.* **2022**, *13*, 940327. [[CrossRef](#)]
63. Castro, W.; Oblitas, J.; Maicelo, J.; Avila-George, H. Evaluation of expert systems techniques for classifying different stages of coffee rust infection in hyperspectral images. *Int. J. Comput. Intell. Syst.* **2018**, *11*, 86. [[CrossRef](#)]
64. Wakholi, C.; Kandpal, L.M.; Lee, H.; Bae, H.; Park, E.; Kim, M.S.; Mo, C.; Lee, W.-H.; Cho, B.-K. Rapid assessment of corn seed viability using short wave infrared line-scan hyperspectral imaging and chemometrics. *Sens. Actuators B Chem.* **2018**, *255*, 498–507. [[CrossRef](#)]
65. Kumar, L.; Schmidt, K.; Dury, S.; Skidmore, A. Imaging Spectrometry and Vegetation Science. In *Imaging Spectrometry: Basic Principles and Prospective Applications*; Meer, F.D.V.D., Jong, S.M.D., Eds.; Springer: Dordrecht, The Netherlands, 2001; pp. 111–155.
66. Schwanninger, M.; Rodrigues, J.C.; Fackler, K. A Review of Band Assignments in near Infrared Spectra of Wood and Wood Components. *J. Near Infrared Spectrosc.* **2011**, *19*, 287–308. [[CrossRef](#)]
67. Curran, P.J. Remote sensing of foliar chemistry. *Remote Sens. Environ.* **1989**, *30*, 271–278. [[CrossRef](#)]