The Detection of Kiwifruit Sunscald Using Spectral Reflectance Data Combined with Machine Learning and CNNs

Ke Wu 1, Zhicheng Jia 2*, Qifeng Duan 2

1 College of Science, Nanjing Forestry University, Nanjing 210037, China; william@njfu.edu.cn
2 Mechanical and Electronic Engineering College, Nanjing Forestry University, Nanjing 210037, China; sidneystrouth@njfu.edu.cn
* Correspondence: jcznfu@njfu.edu.cn

Abstract: Sunscald in kiwifruit, an environmental stress caused by solar radiation during the summer, reduces fruit quality and yields and causes economic losses. The efficient and timely detection of sunscald and similar diseases is a challenging task but helps to implement measures to control stress. This study provides high-precision detection models and relevant spectral information on kiwifruit physiology for similar statuses, including early-stage sunscald, late-stage sunscald, anthracnose, and healthy. Primarily, in the laboratory, 429 groups of spectral reflectance data for leaves of four statuses were collected and analyzed using a hyperspectral reflection acquisition system. Then, multiple modeling approaches, including combined preprocessing methods, feature extraction algorithms, and classification algorithms, were designed to extract bands and evaluate the performance of the models to detect the statuses of kiwifruit. Finally, the detection of different stages of kiwifruit sunscald under anthracnose interference was accomplished. As influential bands, 694–713 nm, 758–777 nm, 780–799 nm, and 1303–1322 nm were extracted. The overall accuracy, precision, recall, and F1-score values of the models reached 100%, demonstrating an ability to detect all statuses with 100% accuracy. It was concluded that the combined processing of moving average and standard normal variable transformations (MS) could significantly improve the data; the near-infrared support vector machine and visible convolutional neural network with MS (NIR-MS-SVM and VIS-MS-CNN) were established as high-precision detection techniques for the classification of similar kiwifruit statuses, demonstrating 25.58% higher accuracy than the single support vector machine. The VIS-MS-CNN model reached convergence with a stable cross-entropy loss of 0.75 in training and 0.77 in validation. The techniques developed in this study will improve orchard management efficiency and yields and increase researchers’ understanding of kiwifruit physiology.

Keywords: kiwifruit; sunscald; anthracnose; spectral reflectance; multilayer perceptron; random forests; support vector machine; convolutional neural network

1. Introduction

The kiwifruit (Actinidia chinensis Planch) belongs to the genus Kiwifruit in the family Rhododendron. According to FAO data for 2021 [1], China produces 2.381 million tons of kiwifruit; the world produces approximately 4.46 million tons. However, kiwifruit, an essential component of Chinese agriculture [2,3], is subjected to increasingly severe sunscald in its planted orchards [4,5]. Sunscald is an environmental stress caused by excessive solar radiation that frequently occurs in the summer and is common in a wide range of fruits, such as grapes, apples, and tomatoes [6–8]. Sunscald produces patches on the fruit and leaf surfaces [9,10] and reduces fruit yields and quality [11], negatively impacting farmers’ incomes. Due to large-scale planting, large areas of sunscald can occur in orchards under high-temperature conditions, resulting in yield reductions of more than 30% [12]. It is crucial to find methods to effectively reduce sunscald losses.
Field treatments for sunscald include evaporative cooling, shade nets, and inhibitors [13]. Spraying systems can reduce the ambient temperature and thus the potential for sunscald, but are accompanied by water and electricity consumption [14,15]. Shade nets can reduce the intensity of sunlight, lower plant canopy temperatures, and increase the relative humidity but increase the risk of fungal diseases [16]. As part of an integrated orchard management strategy, the timely detection of sunscald at different stages for early control strategies before large-scale occurrence is necessary for better resource utilization and disease protection. Problems such as high labor and time costs, inefficiency, and unsuitability for large areas exist for manual detection.

In summer, sunscald occurs after high-temperature exposure; the surface of kiwifruit begins to turn leathery and brown spots appear, and leaf water deficiency is accompanied by curling and drying. As the sunscald deepens, the fruit stops developing, becoming soft and ulcerating [17]. The symptoms of kiwifruit sunscald facilitate the extraction of information on the leaf status from the phenotype. Changes in the physiological status of early-stage sunscald on leaves and fruits are not noticeable compared with healthy fruits, posing a challenge for detection. In addition, studies summarizing early plant stress detection have concluded that it is necessary to consider the influence of other disease factors to determine the plant physiological status [18]. For kiwifruit, anthracnose and sunscald have similar symptoms and occur at overlapping times. Anthracnose is a fungal disease with rapid onset, forming irregularly shaped brown spots at the leaf margins or leaf tips that turn grayish-brown or grayish-white in late stages [19]. For the detection of sunscald, difficulties are caused by early detection and disease interference; practical field identification requires a long observation process and substantial labor costs.

Hyperspectral assays can capture rich and adequate plant information, which is convenient for the determination of the plant status. In agricultural research, it has been reported that hyperspectral methods can be used to detect sunscald and other physiological statuses. For example, spectral reflectance in the wavelength regions of 500–600 nm, 650–700 nm, and 800–850 nm was used to predict the sunscald grade of ‘Packham’s Triumph’ pears with accuracy of 94% [10]; a prediction model for apple sunscald used VIS-NIR reflectance data to determine the effect of predicting apple sunscald in advance [20]. To study the environmental stress of kiwifruit, Ge used spectral technology and SVM to detect chilling injuries in kiwifruit in 2023 and achieved 94.2% accuracy [21]. However, research on how to use hyperspectral data to detect kiwifruit sunscald is still scarce.

Traditionally, various techniques, such as autoregression (AR) [22,23], moving average (MA) [24,25], exponential smoothing (ES) [26], the hybrid method (HM) [27–29], and autoregressive integrated moving average (ARIMA) [30], have been used to construct detection models. In addition, some recent techniques, such as transfer learning and convolutional neural networks [31,32], have been used in the agricultural field for fruit image classification and annotation and disease classification [32,33]. Among these techniques, machine learning and CNN mostly outperform other techniques in terms of precision and accuracy [34,35]. In precision agriculture, classification algorithms are closely combined with spectral technology, generally divided into machine learning and deep learning. Researchers have considered multilayer perceptron, support vector machine, and random forest techniques as classic and popular machine learning methods. The multilayer perceptron is an artificial neural network algorithm [36] that relies on continuously adjusting the parameters in the network to improve the model. It has been used to detect avocado laurel wilt and achieved detection accuracy of 98% [37]. The support vector machine is a typical machine learning algorithm that has been used to distinguish between special and traditional coffee, achieving 96% detection accuracy [38]. Random forest is an algorithm based on ensemble learning that can run in the case of many variable inputs and has strong computational efficiency [39]. It has been used to detect diseases in apples, corn, potatoes, and other plants and has achieved overall accuracy of 96.1% [40]. The convolutional neural network is a popular deep learning algorithm that has been used to detect tomato diseases, including sunscald and anthracnose, with average accuracy of 99.64%
It has also been used to analyze the phenotypes of diseased plant leaves, demonstrating good predictive performance and generalization ability [42].

This study aimed to improve the detection effect in four aspects: data source, preprocessing, feature extraction algorithms, and classification algorithms. The research hypothesis of this study was that the detection model could achieve higher accuracy and better detection of kiwifruit sunscald under anthracnose interference than the widely used plant status detection techniques through the specialized design of the model content described above. To determine the physiological statuses of plants, in addition to building an extensive sample database [43], informative bands should be selected from hyperspectral data [37,44]. The full spectrum (400 nm to 2400 nm), visible spectrum (400 nm to 760 nm), and near-infrared spectrum (761 nm to 2400 nm) were input into PLS-DA to detect ice plants (Aizoaceae), achieving good detection results. The Kappa value reached 0.9 [45]; preprocessing tools such as MA, SNV, and airPLS were used to enhance the noise resistance of the detection models [46–48], which helped to extract information from the spectral data; and spectral feature extraction methods, including PCA and RFE, have been involved in the task of detecting potato chlorophyll content [49].

The contributions of this study are as follows.

1. Improving kiwifruit yield and quality: by detecting sunscald under the interference of anthracnose, farmers could identify potentially problematic plants early and cool or irrigate the plants to avoid further damage, effectively improving kiwifruit yields and quality in the orchard.

2. Optimizing resource utilization: by accurately detecting sunscald in kiwifruit, farmers could avoid wasting resources such as water, nutrients, and pesticides on healthy plants, instead providing them to those already affected by sunscald. This will reduce the wastage of resources, lower costs, and contribute to the development of sustainable agriculture and the protection of the ecological environment.

3. Understanding plant physiological processes: the hyperspectral reflectance technique was used to provide information on the spectral responses of plant leaves in different bands and to extract the relevant bands. Analyzing and interpreting this information helps researchers to understand plants’ growth processes, metabolic activities, and response mechanisms and further advance the research and development of plant biology beyond visual methods.

4. Developing smart agriculture: this study emphasizes the implementation of plant status detection for early-stage sunscald and similar diseases, which can be implemented in combination with other modern agricultural technologies such as IoT, drones, and data analytics for smart agriculture applications.

5. Comparison and selection of models: comparing and evaluating the detection effects of various models based on four parts, including the data source, preprocessing, feature extraction algorithms, and classification algorithms, can provide researchers and decision makers with a basis for selecting the best model. This helps to identify the most suitable model to solve a particular problem and can identify directions for the optimization of the model to improve the algorithm further and enhance the model’s performance.

To the best of our knowledge, this study is the first in the field of precision agriculture to describe the spectral characteristics of kiwifruit sunscald based on hyperspectral technology and develop high-precision detection technology.

2. Materials and Methods

2.1. Study Area and Plant Material

The study area was located in an orchard with both management and experiments in Liuhe District, Nanjing City, Jiangsu Province, China, as shown in Figure 1. Nanjing is situated between 118°22'–119°14' E and 31°14'–32°37' N, with distinct climate changes, abundant light, and an extensive annual range of temperatures; there are 1955.5 total
sunshine hours in a year, the extreme yearly temperature is 39.7 °C, and the mean annual precipitation is 1106 mm. Kiwifruits were planted over approximately 1500 acres in the orchard—the main variety was Yuhuang—and their growth was observed and recorded for an extended period. In summer, the orchard often suffers from sunscald stress, the physiological status of fruit trees is destroyed, and production is restricted.

From July to September 2022, sunscald appeared on the kiwifruits in the orchard, as shown in Figure 1a,b. One hundred mature kiwifruit trees growing in an independent area within a radius of 2 m were selected, and it was ensured that all fruit trees grew healthily for more than one year. These principles ensured the consistency of the trees’ physiological status and structure. Due to the maintenance of orchards by fruit growers, kiwifruit trees are not exposed to other stresses or disease infections. After the visual observation of trees by experienced local farmers and botanists and PCR testing to assess the severity of kiwifruit sunscald and other diseases, a number of independent samples of healthy, early-stage sunscald, late-stage sunscald, and anthracnose kiwifruits were obtained.

Leaves were collected from 12:00 to 14:00 on a sunny day at the end of the month. The collection method was as follows: randomly collecting leaves from four places of the same tree and picking at least five leaves of similar size in each tree. The influence of the leaf growth position and size on the experimental results had to be avoided as much as possible. The collected leaves were placed in sealed bags to reduce the possibility of leaf deterioration and water loss, and then frozen in a sealed container at −20 °C, a standard method for storing plant leaf samples [50]. The time from picking leaf samples in the orchard to using the instrument to collect spectral data in the laboratory was controlled to within four hours, which ensured that the physiological status of the kiwifruit leaves was consistent with that of field leaves. The types and numbers of sample leaves are shown in Figure 2 and Table 1.

Figure 1. Sunscald kiwifruit orchard in Liuhe District, Nanjing, Jiangsu Province, China, showing (a) early-stage sunscald and (b) late-stage sunscald.
Figure 2. Statuses of kiwifruit leaves: (a) healthy; (b) early-stage sunscald; (c) late-stage sunscald; and (d) anthracnose.

Table 1. Sample numbers of different leaf statuses.

<table>
<thead>
<tr>
<th>Statuses of Leaves</th>
<th>Number of Leaves</th>
</tr>
</thead>
<tbody>
<tr>
<td>Healthy</td>
<td>105</td>
</tr>
<tr>
<td>Early-stage sunscald</td>
<td>90</td>
</tr>
<tr>
<td>Late-stage sunscald</td>
<td>130</td>
</tr>
<tr>
<td>Anthracnose</td>
<td>104</td>
</tr>
<tr>
<td>Total</td>
<td>429</td>
</tr>
</tbody>
</table>

2.2. Experiment Apparatus and Data Acquisition

In this study, a hyperspectral data acquisition system was used to collect the spectral reflection data of kiwifruit leaves. The system consisted of a spectrometer, lamp, scanning table, whiteboard, power supply, laptop computer, and control software, as shown in Figure 3a. The spectrometer was an ASD FieldSpec 3 Spectroradiometer (ASD, Inc., Falls Church, VA, USA) with spectral coverage in the wavelength range of 350–2500 nm, 2151 spectral acquisition points, and a spectral resolution of 1.0 nm. The system provides high-resolution data in the visible and infrared bands. The light source power was 75 W, and the vertical incidence angle was 15 degrees. The calibrated reflectance panel was used for light intensity correction, which helped to reduce noise in the spectral data, and for sensitivity optimization, which helped to reduce the effect of noise in the circuit itself on the
results. The scanning table was a clean, solid black, velvet-covered bench, and the ViewSpecPro 6.20 control software operated the spectrometer’s shooting process.

The laboratory in which the measurements were taken was a dark environment. The spectrometer and computer were fully charged before each collection of reflectance data. Before collecting spectral data, the spectrometer was first powered on, and the instrument was allowed to warm up for 15 min; subsequently, the lens was aligned to the calibration panel under illumination conditions, and the calibration panel covered the field of view of the lens to complete light intensity correction and spectrometer sensitivity optimization. Each kiwifruit leaf sample was fixed in the center of the scanning table, with the probe located 30 cm above the leaf surface. This ensured that the collected spectra were the reflectance spectra of the leaf’s central part and a fixed area. The obtained spectral reflectance data are shown in Figure 3b.

![Diagram](image)

**Figure 3.** Laboratory data acquisition: (a) hyperspectral data acquisition system with all components, including a light, calibrated reflectance panel, scanning table, spectrometer, and laptop computer; (b) spectral reflectance data.

### 2.3. Data Analysis

In this study, a method was designed to determine the different statuses of kiwifruit based on the hyperspectral reflectance data of the leaves. The model-building process used a combination of band segmentation, preprocessing, feature extraction algorithms, and classification algorithms to establish multiple detection models. The data analysis process consisted of the following eight steps, as shown in the flowchart in Figure 4: (1) reflectance curve analysis to explore the spectral characteristics of different physiological statuses by calculating the average reflectance and sensitivity; (2) band segmentation to take 400–780 nm in the range of 350–2500 nm as the visible band, VIS, and 780–2500 nm as the near-infrared band, NIR, and to compare the input data in different bands on the model detection effects; (3) preprocessing, using a combination of MA, SNV, and airPLS to obtain MS and MAS methods to process the raw data, comparing the raw data (unprocessed) and the effect of different preprocessing procedures on the model’s detection effects; (4) feature extraction, for which three schemes were designed, namely no feature extraction (unprocessed), PCA, and RFE, which were used to select the essential variables and improve the model’s prediction performance; (5) machine learning classification, based on MLP, RF, and SVM; (6) deep learning classification, to build a convolutional neural network (CNN); (7) model evaluation, in which the OA, recall, precision, and F1-score were selected to evaluate the accuracy of detection, thus favoring the detection
models; and (8) variable analysis, which analyzed the variables obtained from the model preferences and explained the reasons for the improved detection performance of the models from the perspectives of data distribution and significance analysis. The data analysis process was implemented through Python programming.

**Figure 4.** Flowchart.

### 3. Theoretical Foundations

#### 3.1. Reflection Curve Analysis

In this study, parameters were utilized to explore the spectral differences of leaves in different statuses: (1) average reflectance, which was the average of the reflectance of all samples with the same status; and (2) sensitivity, which was calculated by dividing the average reflectance of a stressed or diseased leaf by the average reflectance of healthy leaves.

#### 3.2. Preprocessing

The main preprocessing methods for spectra include the moving average (MA), standard normal variable transformation (SNV), and adaptive iterative weighted penalized least squares (airPLS) [46–48]. In this study, different main preprocessing methods were combined in anticipation of enhancing the model’s predictive performance; the two combined preprocessing methods were MA-SNV (MS) and MA-airPLS-SNV (MAS).

MA is a digital signal processing method for smoothing spectral curves. Within 400–2500 nm, the reflectivity curve may fluctuate in parts of the spectrum due to signal noise. For the wavelength \( t \), the value \( x^* \) after MA processing can be calculated as shown in Equation (1). Here, \( \gamma \) is a factor, and \( \gamma \) is taken as one third at a sliding period of 3.

\[
    x^* = \gamma^1 x^*_{t-1} + \gamma^2 x^*_t + \gamma^3 x^*_{t+1}
\]

SNV is a preprocessing method to remove the variance in the spectral signal by standardizing the spectral data to correct the spectral errors due to scattering. The value of \( x^\text{SNV}_{mp} \) after SNV was calculated as follows: the corrected value was obtained by subtracting the mean value of the reflectance of the band and dividing it by its standard deviation. When the reflectance at wavelength \( t \) was \( x_t \), the reflectance of all samples constituted the column vector \( X = (x_1, x_2, \ldots, x_n)^T \), where \( n \) is the total number of samples.

\[
    x^\text{SNV}_t = \frac{x_t - \text{mean}(X)}{\sqrt{\sum_{i=1}^{n}(x_i - \text{mean}(X))^2 / (n-1)}}
\]

airPLS is based on least squares estimation to achieve the removal of the noise effect of baseline drift on spectral data without any a priori information and improve the
accuracy of the detection performance of spectral signal peaks. It uses a least squares algorithm to fit the baseline of the spectral data, sets a penalty function to evaluate the smoothness of the fitted curve, and uses an adaptive iterative weighting strategy to adjust the parameters in the penalty function.

Theoretically, the combined preprocessing used in this study can correct the sensitivity of the spectral data to wavelength deviation and remove the noise generated during data acquisition, transmission, and storage to obtain more accurate and reliable data, and it reduces the impact of excessive noise on subsequent analyses.

3.3. Feature Extraction

Principal component analysis (PCA) is a commonly used feature extraction technique to extract critical data features by finding the significant variances in the data [38], simplifying the dataset, and reducing the effect of data noise. We analyzed the principal components in the raw data to identify the spectral bands with higher information content and were able to understand the data features better, thus enabling the efficient processing and simplification of the hyperspectral data. Principal components with a cumulative contribution of 95% or more were retained to build the detection models and provide input variables for the kiwifruit sunscald detection models.

Recursive feature elimination (RFE) is a method to eliminate non-important features to select an optimal feature subset for feature extraction [51–53]. This paper used RFE to rapidly downscale the data to improve the recognition accuracy. RFE removes a specific number of non-important features per generation. When the number of features reaches a predetermined value, the algorithm stops and determines the optimal subset of features for the elimination process based on the test results. In this study, two features were eliminated per generation for visible spectral data until twenty features remained. Ten features were eliminated per generation for near-infrared spectral data until twenty features remained. The subset of features selected by RFE with the best prediction was used to build the final detection models for the detection of the status of kiwifruits.

3.4. Classification Algorithm

In this study, three supervised learning algorithms were applied.

(1) Multilayer perceptron (MLP), a representative artificial neural network [54]. The MLP was used for analysis and training, and its structure includes input, hidden, and output layers. Neurons within each layer calculate the output, \( y \), using the following equation:

\[
W^* = \sum_{i=1}^{n} w_i \cdot x_i + b
\]

\[
y = f(W^*)
\]

where \( w_i \) denotes the weight of the input variable \( x_i \) connected to the neuron, and \( b \) denotes the bias term after multiplying the input variable by the weight. \( W^* \) denotes the linear combination of the input and its corresponding weight plus the bias term. Then, the activation function, \( f(W^*) \), is applied to introduce nonlinearity to obtain the neuron’s output, \( y \). The gradient descent method was used to optimize network weights in the training process.

(2) Random forest (RF) is a classification algorithm based on the concept of ensemble learning [55], which makes a final prediction by constructing multiple decision trees and integrating the prediction results of each decision tree. The Gini coefficient method was used to divide the decision tree nodes and select a more appropriate way to classify the data. The equation for calculating the Gini coefficient is as follows:
where \( p_i \) is the probability of each class label and \( c \) denotes the number of classes.

(3) Support vector machine (SVM) performs the classification of the data by constructing an optimal hyperplane [56]. The SVM algorithm was used for the processing and classification of the data. The SVM algorithm separated different types of data by constructing an optimal hyperplane and using a nonlinear function to map the data into a higher-dimensional space to classify the hyperspectral datasets. The SVM was also computed using Equations (3) and (4). However, unlike MLP, which uses an implicit layer structure, SVM finds decision boundaries by solving optimization problems to determine support vectors.

The optimal combinations of hyperparameters were obtained using a network search optimization method, i.e., traversing a predefined hyperparameter space, thus improving the accuracy and reliability of the classification. Cross-validation was used to evaluate the classifier’s performance and the optimization results after each training round utilizing the validation data.

Convolutional neural networks (CNNs) are a popular deep learning algorithm whose structures consist of convolutional layers, pooling layers, activation functions, and fully connected layers; they perform well in feature extraction and classification [57]. For the one-dimensional reflection curve data, we set up a one-dimensional convolutional neural network whose structure consisted of three convolutional layers with a sliding window size of 3, three pooling layers with a maximum pooling method, two fully connected layers, and RELU activation functions to achieve automated feature extraction and feature mapping, as shown in Figure 5. To produce comparable results, all CNNs were trained with the same hyperparameter settings: the number of epochs was set to 250, the batch size was set to 32, the initial learning rate was set to 0.001, and the model parameters were optimized using the Adam optimizer during the training process. Due to the difference in algorithmic construction, there is a difference in feature extraction capability between CNN, MLP, and SVM.

![Figure 5. Structure of the convolutional neural networks. (a) Reflectance data in the visible range were used as inputs to the CNNs; (b) reflectance data in the near-infrared range were used as inputs to the CNNs.](image)

**3.5. Model Evaluation**

Based on the confusion matrix, the OA, precision, recall, and F1-score were used to evaluate the performance of the models. Overall accuracy (OA) was used to assess the overall performance of the models; precision evaluated the proportion of samples predicted by the models to be in a class that were actually in this class; recall evaluated the
ratio of actual samples in a class that were determined to be in this class; and the F1-score is the summed average of precision and recall and was used to evaluate the model’s balanced performance [58,59]. The equations for the OA, precision, recall, and F1-score are shown below for multiclassification.

\[
OA = \frac{1}{N} \sum_{i=1}^{N} \frac{TP_i + TN_i}{TP_i + TN_i + FP_i + FN_i} \times 100\%
\]

(6)

\[
Precision = \frac{1}{N} \sum_{i=1}^{N} \frac{TP_i}{TP_i + FP_i} \times 100\%
\]

(7)

\[
Recall = \frac{1}{N} \sum_{i=1}^{N} \frac{TP_i}{TP_i + FN_i} \times 100\%
\]

(8)

\[
F1 \text{- score} = \frac{2 \times Precision \times Recall}{Precision + Recall} \times 100\%
\]

(9)

4. Tests and Results

4.1. Spectral Reflection Curve Analysis

The average reflectance and sensitivity curves for different statuses of leaves were obtained from the spectral curve acquisition system, as shown in Figure 6. For the visible spectrum, the troughs of the average reflectance curves for healthy, early-stage sunscald, late-stage sunscald, and anthracnose status were located at 410 nm, 421 nm, 403 nm, and 400 nm, respectively. The peaks were concentrated at 760 nm. For the near-infrared spectrum, the troughs were located at 1928 nm, 1925 nm, 1932 nm, and 2491 nm, respectively, and the peaks were located at 1126 nm, 1119 nm, 1127 nm, and 1297 nm, respectively. Within the green band (495–570 nm), the maximum reflectance of the healthy leaves was 8.96%, the early-stage reflectance was 6.44%, the late-stage reflectance was 7.81%, and the reflectance was 8.89%. The average reflectance curves of the four plant statuses exhibited similar trends. In the near-infrared spectrum, the peak position and reflectance of anthracnose (1297 nm, 64.84%) showed significant positional separation and numerical differences from healthy (1126 nm, 57.12%), early-stage sunscald (1119 nm, 48.71%), and late-stage sunscald (1127 nm, 46.39%). The average reflectance curves of anthracnose were overall higher than those of other statuses.

There was a small range of fluctuation in the sensitivity values for the three unhealthy statuses within 400–530 nm. Early-stage and late-stage sunscald sensitivity ranges were 0.67–1.83 and 0.25–1.04, respectively. Anthrax sensitivity values ranged from 0.73 to 2.17, greater than 1.00 overall, and showed multiple peaks in the near-infrared range. The sensitivity of anthrax exhibited an increasing trend with increasing wavelength. It significantly differed from the sensitivity values of sunscald, indicating a more significant spectral difference between anthrax and health.
Figure 6. Two parameter curves of kiwifruit leaf samples: (a) average reflectance; (b) sensitivity.

4.2. Feature Extraction

The results of PCA processing are shown in Table 2. In the visible spectrum, 96.46% of the data variation was explained by PC1 (89.76%) and PC2 (6.70%), and the 758–777 nm and 694–713 nm bands were identified as influential. In the near-infrared spectrum, 97.15% of the data variation was explained by PC1 (80.98%) and PC2 (16.47%), and the 1303–1322 nm and 780–799 nm bands were identified as influential. The scatter distribution of PC1 and PC2 obtained by extraction is shown in Figure 7. The confidence ellipse overlap region accounted for many of the leaf statuses that could not be separated effectively. In particular, the confidence ellipse of anthracnose exhibited a minor overlap with other statuses, and there was clear position separation.
Table 2. PCA-selected bands.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Principal Component</th>
<th>Contribution Rate</th>
<th>Wavelengths (nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>VIS</td>
<td>PC1</td>
<td>89.76%</td>
<td>758 759 760 761 762 763 764 765 766 767</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>768 769 770 771 772 773 774 775 776 777</td>
</tr>
<tr>
<td></td>
<td>PC2</td>
<td>6.70%</td>
<td>694 695 696 697 698 699 700 701 702 703</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>704 705 706 707 708 709 710 711 712 713</td>
</tr>
<tr>
<td>NIR</td>
<td>PC1</td>
<td>80.98%</td>
<td>1303 1304 1305 1306 1307 1308 1309</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1310 1311 1312 1313 1314 1315 1316</td>
</tr>
<tr>
<td></td>
<td>PC2</td>
<td>16.47%</td>
<td>780 781 782 783 784 785 786 787 788 789</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>790 791 792 793 794 795 796 797 798 799</td>
</tr>
</tbody>
</table>

Figure 7. Distributions of principal components and 95% confidence ellipses are shown: (a) principal components of the visible spectrum; (b) principal components of the near-infrared spectrum.

The accuracy changes of MLP, SVM, and RF during the iterations of RFE are shown in Figure 8. With the RFE iteration, the number of input variables decreased, and the algorithm’s accuracy showed an improving trend. In all preprocessing procedures, the accuracy of SVM and RF tended to be flat after increasing, while the accuracy of MLP showed strong fluctuation after increasing. RFE identified different best features, including VIS-SVM (44), VIS-RF (162), VIS-MLP (28), VIS-MS-SVM (76), VIS-MS-RF (26), VIS-MS-MLP (34), VIS-MAS-SVM (80), VIS-MAS-RF (352), VIS-MAS-MLP (26), NIR-SVM (880), NIR-RF (1700), NIR-MLP (70), NIR-MS-SVM (580), NIR-MS-RF (1590), NIR-MS-MLP (370), NIR-MAS-SVM (1740), NIR-MAS-RF (1620), and NIR-MAS-MLP (80).

As shown in Figure 7, the unsupervised learning PCA method could not achieve good discrimination between different leaf statuses. PCA and RFE were modeled as feature extraction in combination with machine learning algorithms RF, SVM, and MLP. The ability of machine learning algorithms to detect kiwifruit statuses was determined by evaluating all methods using the untreated detection models as a benchmark.
Figure 8. Accuracy variation under different preprocessing procedures and the number of iterations corresponding to the highest accuracy: (a) unprocessed visible spectrum; (b) MS-processed visible spectrum; (c) MAS-processed visible spectrum; (d) unprocessed near-infrared spectrum; (e) MS-processed near-infrared spectrum; and (f) MAS-processed near-infrared spectrum.

4.3. Machine Learning

Preprocessing, feature extraction, and machine learning algorithms build multiclassification models for visible and near-infrared spectrum data source analysis. The multiclassification detection models under the unprocessed, MS, and MAS methods are as follows: MLP, RF, SVM, PCA-MLP, PCA-RF, PCA-SVM, RFE-MLP, RFE-RF, and RFE-SVM. Model performance results for the test data are shown in Tables 3 and 4.

Table 3 shows the detection effectiveness metrics of the models in the visible spectrum (400–760 nm). The MS and MAS led to a significant improvement in the metrics of all models. In most cases, the SVM-based models were the optimal and suboptimal models in each group. RF (OA: 72.09%, precision: 72.97%, recall: 72.08%, and F1-score: 71.85%) appeared to be the only suboptimal model. The best model was VIS-MS-RFE-SVM (OA: 97.67%, precision: 97.87%, recall: 97.72%, and F1-score: 97.77%). Comparing VIS-MS-SVM and VIS-MAS-SVM, the airPLS increased the OA by 0.77%.

Table 4 shows the metrics of the models in the near-infrared spectrum (780–2500 nm). Again, MS and MAS showed an overall improvement in model performance, especially for SVM and MLP. The SVM-based models were generally optimal and suboptimal, with NIR-RFE-RF (OA: 74.42%, precision: 72.64%, recall: 72.59%, and F1-score: 72.57%) appearing to be the only RF-based suboptimal model. NIR-MS-SVM was the best predictor (OA: 100%, precision: 100%, recall: 100%, and F1-score: 100%). airPLS processing resulted in a 3.88% decrease in OA for NIR-MAS-SVM relative to NIR-MS-SVM.
Table 3. Performance of machine learning model detection using the visible spectrum. The best models are shown in bold; the second-best models are underlined.

<table>
<thead>
<tr>
<th>Data</th>
<th>Model</th>
<th>OA (%)</th>
<th>Precision (%)</th>
<th>Recall (%)</th>
<th>F1-score (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MLP</td>
<td>49.61</td>
<td>67.08</td>
<td>50.14</td>
<td>46.29</td>
</tr>
<tr>
<td></td>
<td>RF</td>
<td>72.09</td>
<td>72.97</td>
<td>72.08</td>
<td>71.85</td>
</tr>
<tr>
<td></td>
<td>SVM</td>
<td>74.42</td>
<td>74.08</td>
<td>74.12</td>
<td>73.76</td>
</tr>
<tr>
<td></td>
<td>PCA-MLP</td>
<td>37.98</td>
<td>47.96</td>
<td>39.43</td>
<td>34.11</td>
</tr>
<tr>
<td></td>
<td>PCA-RF</td>
<td>51.16</td>
<td>50.61</td>
<td>50.04</td>
<td>50.13</td>
</tr>
<tr>
<td></td>
<td>PCA-SVM</td>
<td>56.59</td>
<td>59.42</td>
<td>55.22</td>
<td>55.26</td>
</tr>
<tr>
<td></td>
<td>RFE-MLP</td>
<td>72.09</td>
<td>70.61</td>
<td>70.17</td>
<td>70.27</td>
</tr>
<tr>
<td></td>
<td>RFE-RF</td>
<td>71.32</td>
<td>70.43</td>
<td>70.68</td>
<td>70.34</td>
</tr>
<tr>
<td></td>
<td>RFE-SVM</td>
<td>72.87</td>
<td>70.09</td>
<td>68.61</td>
<td>68.70</td>
</tr>
</tbody>
</table>

Table 4. Performance of machine learning model detection using near-infrared spectrum. The best models are shown in bold; the second-best models are underlined.

<table>
<thead>
<tr>
<th>Data</th>
<th>Model</th>
<th>OA (%)</th>
<th>Precision (%)</th>
<th>Recall (%)</th>
<th>F1-score (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MLP</td>
<td>58.59</td>
<td>67.90</td>
<td>55.73</td>
<td>55.81</td>
</tr>
<tr>
<td></td>
<td>RF</td>
<td>70.54</td>
<td>71.28</td>
<td>71.43</td>
<td>70.28</td>
</tr>
<tr>
<td></td>
<td>SVM</td>
<td>56.59</td>
<td>61.86</td>
<td>54.49</td>
<td>53.17</td>
</tr>
<tr>
<td></td>
<td>PCA-MLP</td>
<td>43.41</td>
<td>54.53</td>
<td>44.15</td>
<td>41.94</td>
</tr>
<tr>
<td></td>
<td>PCA-RF</td>
<td>87.60</td>
<td>87.99</td>
<td>87.20</td>
<td>87.26</td>
</tr>
<tr>
<td></td>
<td>PCA-SVM</td>
<td>82.17</td>
<td>80.63</td>
<td>81.67</td>
<td>80.73</td>
</tr>
<tr>
<td></td>
<td>RFE-MLP</td>
<td>89.92</td>
<td>90.55</td>
<td>89.11</td>
<td>89.46</td>
</tr>
<tr>
<td></td>
<td>RFE-RF</td>
<td>88.37</td>
<td>87.79</td>
<td>87.35</td>
<td>87.49</td>
</tr>
<tr>
<td></td>
<td>RFE-SVM</td>
<td>93.80</td>
<td>93.71</td>
<td>92.08</td>
<td>92.45</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Data</th>
<th>Model</th>
<th>OA (%)</th>
<th>Precision (%)</th>
<th>Recall (%)</th>
<th>F1-score (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MLP</td>
<td>86.82</td>
<td>87.94</td>
<td>85.47</td>
<td>86.53</td>
</tr>
<tr>
<td></td>
<td>RF</td>
<td>86.82</td>
<td>86.12</td>
<td>85.74</td>
<td>85.80</td>
</tr>
</tbody>
</table>
### Table 1: Performance Comparison of Different Models

<table>
<thead>
<tr>
<th>Model Type</th>
<th>OA (%)</th>
<th>Precision (%)</th>
<th>Recall (%)</th>
<th>F1-score (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVM</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td>PCA-MLP</td>
<td>68.22</td>
<td>69.78</td>
<td>67.43</td>
<td>67.62</td>
</tr>
<tr>
<td>PCA-RF</td>
<td>84.50</td>
<td>84.16</td>
<td>83.71</td>
<td>83.84</td>
</tr>
<tr>
<td>PCA-SVM</td>
<td>78.29</td>
<td>78.01</td>
<td>78.12</td>
<td>77.83</td>
</tr>
<tr>
<td>RFE-MLP</td>
<td>84.50</td>
<td>83.26</td>
<td>84.16</td>
<td>83.41</td>
</tr>
<tr>
<td>RFE-RF</td>
<td>94.57</td>
<td>94.06</td>
<td>94.59</td>
<td>94.30</td>
</tr>
<tr>
<td>RFE-SVM</td>
<td>96.12</td>
<td>96.32</td>
<td>95.88</td>
<td>95.98</td>
</tr>
</tbody>
</table>

MAS-processed

<table>
<thead>
<tr>
<th>Model Type</th>
<th>OA (%)</th>
<th>Precision (%)</th>
<th>Recall (%)</th>
<th>F1-score (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLP</td>
<td>94.57</td>
<td>95.83</td>
<td>94.66</td>
<td>95.21</td>
</tr>
<tr>
<td>RF</td>
<td>90.70</td>
<td>90.32</td>
<td>91.06</td>
<td>90.30</td>
</tr>
<tr>
<td>SVM</td>
<td>96.12</td>
<td>95.88</td>
<td>96.15</td>
<td>95.86</td>
</tr>
<tr>
<td>PCA-MLP</td>
<td>71.32</td>
<td>75.53</td>
<td>71.91</td>
<td>72.51</td>
</tr>
<tr>
<td>PCA-RF</td>
<td>87.60</td>
<td>90.02</td>
<td>86.53</td>
<td>87.25</td>
</tr>
<tr>
<td>PCA-SVM</td>
<td>94.57</td>
<td>94.13</td>
<td>95.40</td>
<td>94.62</td>
</tr>
<tr>
<td>RFE-MLP</td>
<td>96.12</td>
<td>96.23</td>
<td>96.54</td>
<td>96.31</td>
</tr>
<tr>
<td>RFE-RF</td>
<td>93.80</td>
<td>93.70</td>
<td>94.23</td>
<td>93.81</td>
</tr>
<tr>
<td>RFE-SVM</td>
<td>98.45</td>
<td>98.48</td>
<td>98.32</td>
<td>98.37</td>
</tr>
</tbody>
</table>

### 4.4. Deep Learning

Figure 9 shows the convergence curves of all the deep learning models. The validation accuracy of VIS-MS-CNN was 63.33% at epoch one and gradually increased and stabilized to 100%, and the validation loss decreased and stabilized to 0.74. The validation accuracy of NIR-MS-CNN was 56.66% at epoch one and gradually increased and stabilized to 96.64%, and the validation loss decreased and stabilized to 0.79. After 200 epochs, VIS-MS-CNN and NIR-MS-CNN converged on the training and validation data without overfitting. The test results detailed in Table 5 show that the VIS-MS-CNN model achieved optimal performance in plant status detection (OA: 100%, precision: 100%, recall: 100%, and F1-score: 100%). Although MAS is a more sophisticated preprocessing procedure, the predictive metrics of the MAS-processed models were lower overall than those of the MS-processed models. In the deep learning algorithm, the visible spectrum obtained by band segmentation was enhanced in the prediction performance.
Figure 9. Model convergence curves: (a) accuracy of VIS-CNN; (b) accuracy of VIS-MS-CNN; (c) accuracy of VIS-MAS-CNN; (d) loss of VIS-CNN; (e) loss of VIS-MS-CNN; (f) loss of VIS-MAS-CNN; (g) accuracy of NIR-CNN; (h) accuracy of NIR-MS-CNN; (i) accuracy of NIR-MAS-CNN; (j) loss of NIR-CNN; (k) loss of NIR-MS-CNN; and (l) loss of NIR-MAS-CNN.

Table 5. Detection performance of the deep learning models on the test data. The best models are shown in bold.

<table>
<thead>
<tr>
<th>Data</th>
<th>Model</th>
<th>Model Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>OA (%)</td>
</tr>
<tr>
<td>VIS</td>
<td>CNN</td>
<td>77.52</td>
</tr>
<tr>
<td></td>
<td>MS-CNN</td>
<td>100.00</td>
</tr>
<tr>
<td></td>
<td>MAS-CNN</td>
<td>98.45</td>
</tr>
<tr>
<td></td>
<td>CNN</td>
<td>51.16</td>
</tr>
<tr>
<td>NIR</td>
<td>MS-CNN</td>
<td>96.90</td>
</tr>
<tr>
<td></td>
<td>MAS-CNN</td>
<td>91.47</td>
</tr>
</tbody>
</table>

4.5. Model Comparison and Analysis

A confusion matrix of the predicted results is shown in Figure 10. The best machine learning models for each preprocessing and the deep learning models for both data sources were obtained by screening, as shown in Figure 11. As analyzed above, the introduction of MS and MAS improved the model prediction metrics. The number of models with extracted and unextracted RFE features in machine learning accounted for one half of the total. PCA was unsuitable for modeling kiwifruit sunscald detection, and the validity of the preferred subset of variables for RFE and RFE was demonstrated. SVM performed well among all machine learning algorithms. The prediction metrics of the deep learning models were higher than 95% for both spectral ranges. Among all models, NIR-MS-SVM and VIS-MS-CNN achieved the highest detection metrics.
Figure 10. The confusion matrices: (a) VIS-SVM; (b) VIS-MS-RFE-SVM; (c) VIS-MAS-SVM; (d) NIR-RFE-SVM; (e) NIR-MS-SVM; (f) NIR-MAS-RFE-SVM; (g) VIS-MS-SVM; and (h) NIR-MS-SVM. *Early: Early-stage sunscald. *Late: Late-stage sunscald.
To evaluate the effectiveness of the models in detecting various plant statuses, the detection accuracy was calculated based on the confusion matrix, as shown in Figure 12. NIR-MS-SVM and VIS-MS-CNN could detect all leaf statuses well (accuracy: 100%). In particular, the NIR-MS-CNN model achieved 100% accuracy in detecting early-stage sunscald, and NIR-MAS-RFE-SVM and VIS-MS-RFE achieved 100% accuracy in detecting anthracnose. However, some models were not adapted to the detection problem; the accuracy of NIR-RFE-SVM and VIS-SVM in detecting early-stage sunscald was below 70%.

The distribution of the reflectance values in the significant bands in the different kiwi statuses’ (healthy, early-stage sunscald, late-stage sunscald, and anthracnose) detection was compared, taking 695 nm as an example, as shown in Figure 13. The most significant difference between groups was found for the MS-processed data, with an F-value of 248.89. The significance levels of all data were <0.05, which was statistically significant.
Figure 13. Distribution of reflectance data at 695 nm in different statuses under different preprocessing procedures and their F and p statistics (obtained by ANOVA): (a) unprocessed; (b) MS-processed; and (c) MAS-processed.

5. Discussion and Future Work

The trends and characteristics of the spectral changes in kiwifruit leaves in each status determined in the laboratory are shown in Figure 6. As the sunscald developed, the pigment content of the kiwifruit leaves decreased, increasing the red light reflectance. Typically, healthy plant leaves absorb red and reflect near-infrared light. Comparing the visible spectrum, the difference in the average reflectance of kiwifruit leaves in the near-infrared band was greater, possibly due to the altered physiological status of the visible leaves caused by sunscald not visible to the human eye. In addition, the similarity of symptoms between early-stage and late-stage sunscald resulted in close reflectance values, which posed a challenge in correctly differentiating between early-stage and late-stage sunscald. The most dramatic changes in reflectance were observed in anthracnose due to the onset of large, black, disease-causing spots along the veins on the plant leaves, which affected the appearance and moisture content. As the water content of anthracnose leaves decreased, their reflectance in the near-infrared spectrum band increased significantly; Penuelas obtained results consistent with ours [60]. Early-stage and late-stage sunscald produced significant differences in sensitivity values in the visible spectrum band. The sensitivity values were overall higher for anthracnose than for sunscald. Differences in sensitivity regarding disease status were also obtained during the detection of tomato diseases by Jaffar using hyperspectral reflectance data [54].

PCA can identify features with more valuable data by calculating the entropy information [61]. As shown in Table 2, PCA focuses on a specific band range among the bands extracted from the raw data. The visible spectrum consists of bars within the red spectrum (630–760 nm), reflecting the ability of kiwifruit leaves to reflect red light. The near-infrared spectrum group consists of short wavelengths of near-infrared range (780–1400 nm), reflecting the moisture content of kiwifruit leaves; Balasundram reported that the spectral region between 500 and 800 nm showed the greatest discriminatory power, with the upper limit of effective wavelengths up to 1100 nm for the detection of citrus peel ulcers [62]. Different physiological processes reflect information on the leaf spectra, resulting in deviations from the anthracnose confidence ellipse. RFE effectively removes feature redundancy and is a feature extraction method that facilitates improvements in model prediction performance.

The detection accuracy of the SVM, RF, and MLP base models varied widely under different preprocessing types, where MAS processing yielded stable and uniform prediction results. Using MAS-enhanced data, it was easier to obtain relevant information [63,64]. The MLP accuracy results fluctuated drastically, probably due to the inability of the MLP algorithm to converge within a limited number of iterations under limited data conditions to obtain a model with stable prediction results [65,66]. The extraction of wavebands promotes the use of inexpensive equipment to discriminate plant statuses, such as multispectral cameras or drones.
In machine learning, the MS-SVM with full-band inputs of the NIR spectrum had the highest prediction index and achieved 100% detection for all plant statuses. Better detection metrics were obtained for the models without feature extraction; other researchers have observed similar findings. Sankaran used full-band data (350–2500 nm) as input features to the detection models and obtained 98% detection accuracy for citrus yellow dragon disease [67].

Studies have shown that SVM outperforms RF and MLP in detection. The advantage of SVM is that it is suitable for training on small sample datasets and can perform well with many features [68]. RF has good resistance to overfitting; however, it requires larger samples to meet the training requirements [55]. Winston Pinheiro collected 1048 coffee bean samples for special and conventional coffee beans and obtained 97% and 88% prediction accuracy for SVM and RF, respectively [38]. With the large number of parameters computed during the training of the MLP [69], the study demonstrated low accuracy under small sample conditions. On the other hand, the computational efficiency advantage of SVM will promote its application in large survey area applications.

Detecting early-stage and late-stage sunscald is challenging, making misclassification by models with poor VIS-SV and NIR-RFE-SVM detection capabilities possible. In field management, early and late treatments are not timely, causing tree destruction and economic losses; the misclassification of different statuses of plants affects the protection of tree health, or even kills them [70,71]. In 2022, Zhao used hyperspectral and continuous wavelet analysis (CWA) to distinguish oil tea sunscald from similar diseases; the accuracy of sunscald detection was 82.50% to 83.91%, and the anthracnose detection accuracy was 94.12% to 94.28% [72]. Additionally, Samah Alhazmi obtained a recall value of 93.94% using a convolutional neural network to detect diseases, including anthracnose [73]. However, after considering the effects of early-stage sunscald and disease interference, the MS-SVM and VIS-MS-CNN models presented in this study eventually achieved 100% detection of all leaf statuses, with greater robustness. In addition, the accuracy of NIR-MS-SVM and NIR-MS-CNN improved by 25.58% compared with the best model in traditional machine learning (VIS-SVM). After MS and MAS processing, the data distributions of different classes of spectra were improved, and the significance of their differences was increased by MS, which facilitated the models in finding better information for leaf status detection.

In the last few years, many breakthroughs have been made in sensors for plant trait analysis [18], including multispectral sensing and hyperspectral instruments, which provide basic conditions for the development and extension of acquisition systems. In kiwifruit disease detection, the combination of preprocessing, feature extraction, machine learning, and deep algorithms has the advantages of high accuracy, diverse methods, and scalability [74]. Machine learning detection is suitable for small- and medium-sized growers. Although deep learning has higher data volume requirements, preprocessing is vital.

Furthermore, collecting large datasets that include more kiwi varieties, more similar diseases, and more plant species would provide opportunities to improve the detection performance of machine learning and CNN models. To further reduce the number of epochs for convergence and the possibility of overfitting, it is beneficial to compute the parameters using a strategic heuristic optimization algorithm. At the same time, this is a future research direction for the more efficient extraction of spectral band information for similar statuses.

With global warming, sunscald is more likely to occur and needs more attention as a worldwide environmental stress [75,76]. Despite its wide distribution and abundant germplasm in China, kiwifruit’s quality is affected by stresses that prevent it from entering the global kiwifruit consumer market. For kiwifruit sunscald outbreaks, existing pesticide-spraying helicopters operating autonomously can be used for control [77]. Research on kiwifruit sunscald and related technologies is essential to improve the production and management practices of the kiwifruit industry in China and around the world.
6. Conclusions

Our results have demonstrated that NIR-MS-SVM and VIS-MS-CNN are the best detection models for the identification of similar plant statuses, such as healthy, early-stage sunscald, late-stage sunscald, and anthracnose. MS and MAS preprocessing can significantly improve the ability of machine learning and deep learning models’ ability to predict kiwifruit’s physiological status; MS is considered the best preprocessing method. Information in the visible and near-infrared spectrum bands enabled machine learning and deep learning to distinguish kiwifruit health, sunscald, and anthracnose compared with full-band inputs. PCA and RFE machine learning increased the model’s detection capability. In addition, we selected information-rich spectrum bands based on PCA; the 694–713 nm, 758–777 nm, 780–799 nm, and 1303–1322 nm bands provided more practical information for the nondestructive determination of the plant status. The results of this study will be used to develop agricultural machines that automatically identify sunscald and similar diseases, improving the efficiency of orchard management. At the same time, the information extracted will support the work of botanists and growers in disease management and plant protection to improve kiwifruit yields and quality.

Author Contributions: Conceptualization, K.W. and Z.J.; Data curation, Z.J.; Formal analysis, K.W.; Funding acquisition, Z.J.; Investigation, K.W. and Q.D.; Methodology, K.W., Z.J. and Q.D.; Project administration, Z.J.; Resources, K.W. and Z.J.; Software, K.W.; Supervision, K.W. and Z.J.; Validation, K.W.; Visualization, K.W.; Writing—original draft, K.W.; Writing—review and editing, K.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Modern Agricultural Machinery Equipment and Technology Demonstration and Promotion Project of Jiangsu Province, grant number NJ2022-12; and the Jiangsu Agricultural Science and Technology Innovation Fund, grant number CX (19)3075.

Data Availability Statement: The datasets generated for this study are available on request to the corresponding author.

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

References


47. Wang, X.; Xie, F.; Yang, Y.; Zhao, J.; Wu, G.; Wang, S. Rapid Diagnosis of Ductal Carcinoma In Situ and Breast Cancer Based on Raman Spectroscopy of Serum Combined with Convolutional Neural Network. Bioengineering 2023, 10, 65.


**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.