


## Article

# Goji Disease and Pest Monitoring Model Based on Unmanned Aerial Vehicle Hyperspectral Images

Ruixin Zhao <sup>1,2,3</sup>, Biyun Zhang <sup>4</sup>, Chunmin Zhang <sup>1,2,3,\*</sup>, Zeyu Chen <sup>1,2,3,5</sup>, Ning Chang <sup>1,2,3,5</sup>, Baoyu Zhou <sup>6</sup>, Ke Ke <sup>1,2,3</sup> and Feng Tang <sup>1,2,3</sup> 

<sup>1</sup> School of Physics, Xi'an Jiaotong University, Xi'an 710049, China; xjtu\_zrx@163.com (R.Z.); chenzeyu@opt.ac.cn (Z.C.); channingjean@163.com (N.C.); kexjtu@163.com (K.K.); fengtang@stu.xjtu.edu.cn (F.T.)

<sup>2</sup> The Institute of Space Optics, Xi'an Jiaotong University, Xi'an 710049, China

<sup>3</sup> Key Laboratory for Nonequilibrium Synthesis and Modulation of Condensed Matter, Xi'an Jiaotong University, Ministry of Education, Xi'an 710049, China

<sup>4</sup> BA Trading (Guangzhou) Co., Ltd., Guangzhou 510000, China; lauren\_zhang@live.cn

<sup>5</sup> Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences, Xi'an 710119, China

<sup>6</sup> Ningxia Bing He Technology Co., Ltd., Shizuishan 753099, China; xianzby@163.com

\* Correspondence: zcm@xjtu.edu.cn

**Abstract:** Combining near-earth remote sensing spectral imaging technology with unmanned aerial vehicle (UAV) remote sensing technology, we measured the Ningqi No. 10 goji variety under conditions of health, infestation by psyllids, and infestation by gall mites in Shizuishan City, Ningxia Hui Autonomous Region. The results indicate that the red and near-infrared spectral bands are particularly sensitive for detecting pest and disease conditions in goji. Using UAV-measured data, a remote sensing monitoring model for goji pest and disease was developed and validated using near-earth remote sensing hyperspectral data. A fully connected neural network achieved an accuracy of over 96.82% in classifying gall mite infestations, thereby enhancing the precision of pest and disease monitoring in goji. This demonstrates the reliability of UAV remote sensing. The pest and disease remote sensing monitoring model was used to visually present predictive results on hyperspectral images of goji, achieving data visualization.

**Keywords:** hyperspectral; diseases and pests; unmanned aerial vehicle (UAV); remote sensing monitoring



**Citation:** Zhao, R.; Zhang, B.; Zhang, C.; Chen, Z.; Chang, N.; Zhou, B.; Ke, K.; Tang, F. Goji Disease and Pest Monitoring Model Based on Unmanned Aerial Vehicle Hyperspectral Images. *Sensors* **2024**, *24*, 6739. <https://doi.org/10.3390/s24206739>

Academic Editor: Alfred Stein

Received: 2 September 2024

Revised: 3 October 2024

Accepted: 16 October 2024

Published: 20 October 2024



**Copyright:** © 2024 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

Ningxia is a major goji berry-producing area in China, and its goji berry industry has become one of the pillars of Ningxia's agriculture after many years of development. Goji berries have been cultivated in Ningxia for over 500 years. In recent years, as the main cash crop of the region, the planting area for goji berries has been expanding annually. The region now maintains a cultivation area of 380,000 mu (approximately 25,333 hectares), with a fresh fruit output of 300,000 tons and a total industry value exceeding 27 billion yuan. With the expansion of the cultivated area and the increase in years of cultivation, the occurrence of pests and diseases has become increasingly severe, causing significant economic losses. Common pests and diseases include the goji psyllids and the goji gall mites [1,2].

The spectral characteristics of ground objects form the basis for both theoretical and applied research in the field of remote sensing. As early as the 1980s, researchers have utilized portable spectrometers to study crop diseases and pest infestations [3]. Near-earth remote sensing spectral imaging technology involves the use of handheld or portable spectrometers to measure the spectral reflectance of crop canopies and leaves affected by diseases and pests in both laboratory and field settings. This technology is not only used

for analyzing the spectral characteristics of various diseases and pests and identifying sensitive spectral bands after damage, but it also allows for the calibration of target spectral data on the ground before satellite remote sensing applications. Near-earth remote sensing spectral imaging can also determine the spectral characteristics of plants to assess the health of vegetation. Moreover, the continuous monitoring of spectral data from near-earth remote sensing spectral imaging in fields allows for the real-time tracking of crop growth, providing critical data support for scientific agricultural management.

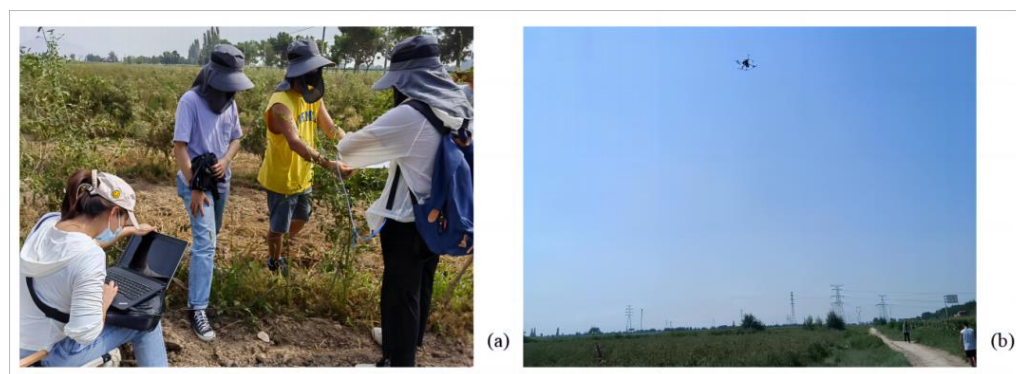
Drone remote sensing involves the use of sensors mounted on unmanned aerial vehicles (UAVs) to gather surface information. Drones can carry high-resolution optical and multispectral sensors to capture high-resolution images of farmland, providing detailed spatial information. They can quickly cover large areas of farmland and gather real-time data, offering timely support for agricultural decision-making. The combined use of near-surface spectral technology and drone remote sensing, such as equipping drones with near-earth remote sensing instruments, can more comprehensively capture spectral information of farmland, thereby enhancing the precision of crop and soil monitoring. This is of significant importance for the management of agricultural production, pest and disease control, and the optimization of resource utilization.

In recent years, deep learning has become a pivotal tool in modern intelligent agriculture, owing to its ability to extract complex features from large datasets [4,5]. As a result, deep learning models are increasingly used in agricultural research. For example, Guo, Q et al. employed a principal component analysis and the successive projections algorithm to reduce the dimensionality of hyperspectral data for wheat rust and powdery mildew, subsequently developing discriminative models. In 2017, Whetton, R.L et al. used partial least squares regression (PLSR) to create calibration models that predict the infection rates of wheat and barley by leaf rust and fusarium diseases [6]. In 2019, Wang, X.L et al. developed a pest and disease identification model based on hyperspectral data from cotton leaves infested by aphids and spider mites using support vector machine (SVM) algorithms [7]. Similarly, in 2021, Lakshmi V et al. extracted texture features from healthy and diseased cotton leaves using the gray-level co-occurrence matrix and constructed an SVM-based pest and disease identification model [8]. In 2022, Dilixiatl, Y et al. selected specific vegetation indices to build regression models for identifying and monitoring the occurrence of cotton aphids, cotton spider mites, and cotton bollworms [9]. In the case of goji berries, pest and disease monitoring currently relies on linear methods, such as multiple regression, while research into nonlinear approaches involving neural networks remains limited [10].

## 2. Materials and Methods

### 2.1. Subsection

In the elite goji berry ecological demonstration garden located in the Dawukou District, Shizuishan City, Ningxia Hui Autonomous Region, detection was carried out using the Flame-T-VIS-NIR-ES spectrometer and the Nano-Hyperspec miniature airborne hyperspectral imager from Ocean Optics, located in Dunedin, Florida, USA. In July, when pest infestations are most common, Ningqi No. 10 goji berry cultivars in their peak fruiting phase were used as detection targets. The two experimental fields were selected under the guidance of Ms. Gao Min from the Ningxia Forest Pest Control and Quarantine Station. The fields are free of mixed pest infestations and only have single infections of either goji psyllids or goji gall mites. Each field included healthy goji plants as controls. The field detection experiment picture is shown in Figure 1. The technical specifications of the Flame-T-VIS- NIR-ES spectrometer and Nano-Hyperspec miniature airborne hyperspectral imager are detailed in Tables 1 and 2, respectively.



**Figure 1.** Field detection experiment diagram. (a) Near-ground hyperspectral detection; (b) UAV detection.

**Table 1.** Parameters for handheld spectrometer.

Parameters	Value
Model	FLAME-T-VIS-NIR-ES
Signal-to-noise ratio	250:1
Spectral resolution	1.33 $\mu\text{m}$
Dark noise	50 RMS counts
Integration time	1 ms~65 s

**Table 2.** Parameters for drone-mounted hyperspectral imager.

Parameters	Value
Unmanned aerial vehicle model	DJI M300
Imaging spectrometer model	Nano-Hyperspec
Altitude	15 m
Geospatial resolution	1 cm
Lateral overlap	30%
Flight speed	1.1 m/s
Spectral wavebands	400~1000 nm
Number of channels	270
Field of view	32°

## 2.2. Spectral Reflectance Measurement

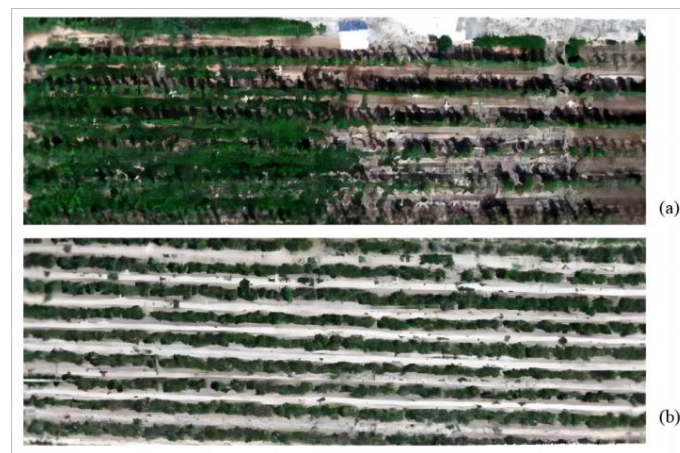
Under the guidance of Ms. Gao Min from the Ningxia Forest Pest Control and Quarantine Station, the detection was performed around 13:00 on clear, sunny days with wind speeds of less than 3 on the Beaufort scale.

**Ground Spectral Data:** Measurements with the handheld spectrometer were conducted with the probe positioned vertically downward and consistently about 1 m above the canopy. The field of view angle of the probe was fixed at 25°. Before each measurement, the instrument was calibrated against a standard whiteboard. To ensure the representativeness of the spectral data, branches were randomly sampled from five directions (north, south, east, west, and center) within each detection area, with 10 repeats per sample. The average value was used to represent the spectral reflectance of the goji canopy.

**Drone Spectral Data:** A single flight detection was conducted over the entire experimental area at a flight altitude of 15 m.

## 2.3. Data Analysis

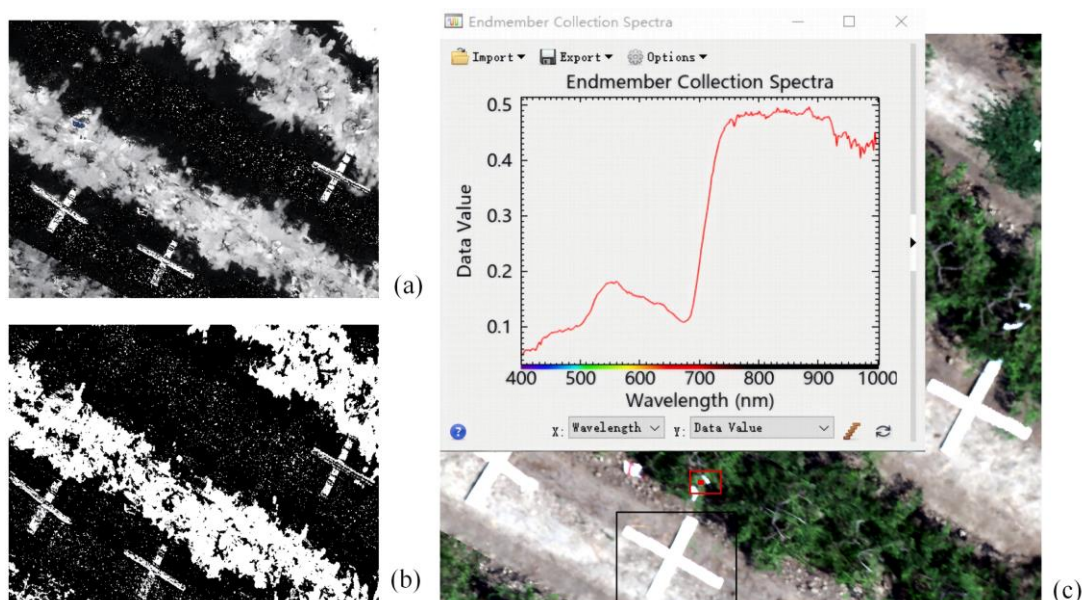
The images collected by the drone are shown in Figure 2. The experimental field infected with goji psyllids pests is shown in Figure 2a, while the field infected with goji gall mites pests is shown in Figure 2b. The collected data were processed and analyzed using ENVI 5.6, SPSS 17.0, Excel 2007, and Matlab R2022a.



**Figure 2.** Drone images collected in Dawukou District, Shizuishan City, Ningxia Hui Autonomous Region. (a) Experimental field infected with goji psyllids pests; (b) experimental field infected with goji gall mites.

### 2.3.1. Spectral Data Preprocessing

Hyperspectral image preprocessing involves a series of steps performed on images before data analysis to enhance data quality, reduce noise, and eliminate invalid information. To mitigate the effects of background factors, outlier data are removed from the collected spectral data, and the ground spectral data are subjected to first-order derivative processing [11,12]. For the stitched drone images, noise reduction and smoothing are first performed using ENVI 5.6 software. To reduce soil interference, a masking process is applied to the drone images to segment green vegetation from the image, excluding soil and other non-target areas to enhance data quality and analytical precision. The preprocessing steps for hyperspectral images acquired by the drone are illustrated in Figure 3.



**Figure 3.** Hyperspectral image preprocessing process: (a) calculation of vegetation indices; (b) mask processing; (c) extraction of hyperspectral curves from the region of interest. the red box indicates the selected area of interest, the black box shows the representative goji plants marked for infection with pests and diseases.



### 2.3.2. Feature Parameter Selection

The spectral index method is an approach that utilizes spectral information from various bands in remote sensing imagery to compute specific index values, thereby reflecting the characteristics and information of the target object [13,14]. Vegetation reflectance characteristics are influenced by vegetation structure, leaf surface features, and vegetation coverage. By analyzing the reflectance across different bands, information about the growth state of the vegetation can be obtained [15,16]. The commonly used hyperspectral feature parameters in the field of remote sensing are shown in Table 3.

**Table 3.** Hyperspectral characteristic parameters and their descriptions.

Characteristic Parameter	Definition
$SD_b$	Sum of the first-order derivative values within the blue edge wavelength range (492~530 nm)
$SD_g$	Sum of the first-order derivative values within the green edge wavelength range (505~553 nm)
$SD_y$	Sum of the first-order derivative values within the yellow edge wavelength range (555~571 nm)
$SD_r$	Sum of the first-order derivative values within the red edge wavelength range (680~760 nm)
$SD_r/SD_b$	Ratio of the total sum of first-order derivatives within the red edge to that within the blue edge
$SD_r/SD_g$	Ratio of the total sum of first-order derivatives within the red edge to that within the green edge
$(SD_r - SD_b)/(SD_r + SD_b)$	The normalized sum of first-order derivatives within the red edge compared to that within the blue edge.
$(SD_r - SD_g)/(SD_r + SD_g)$	The normalized sum of first-order derivatives within the red edge compared to that within the green edge.
$PSSR_a$	$R_{800}/R_{680}$
$PSSR_b$	$R_{800}/R_{635}$
$PSSR_c$	$R_{800}/R_{470}$

Vegetation indices are calculated by analyzing the light reflected from vegetation at different wavelengths to assess vegetation health, growth status, and other biological characteristics. The seven representative vegetation indices selected are shown in Table 4.

**Table 4.** Representative vegetation indices and their definitions.

Vegetation indices	Formula	Reference
Normalized Difference Vegetation Index	$NDVI = (NIR - Red)/(NIR + Red)$	[17]
Green Normalized Difference Vegetation Index	$GNDVI = (NIR - Green)/(NIR + Green)$	[18]
Modified Triangular Vegetation Index	$MTVI = 1.2 \times (1.2 \times (NIR - Green) - 2.5 \times (Red - Green))$	[19]
Modified NDVI 705	$mNDVI_{705} = (NIR - RedEdge)/(NIR + RedEdge)$	[20]
Modified Simple Ratio 705	$mSR_{705} = NIR/RedEdge$	[21]
Red-Green Ratio Index	$RGRI = Red/Green$	[22]
Triangular Vegetation Index	$TVI = 0.5 \times (120 \times (NIR - Green) - 200 \times (Red - Green))$	[23]

### 2.3.3. Method Selection

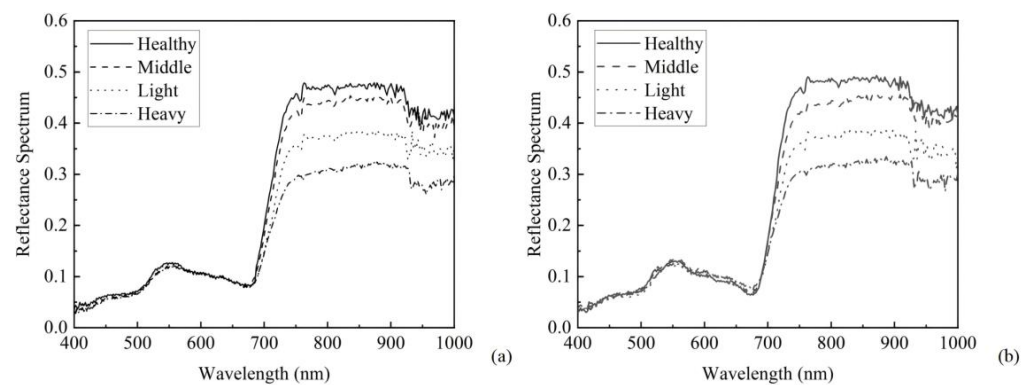
From the goji psyllid data, 6481 pixels were selected, including 98 healthy, 2064 light, 3480 middle, and 839 heavy instances. From the goji gall mite data, 2722 pixels were selected, comprising 723 healthy, 1036 light, 505 middle, and 458 heavy instances. Three methods were employed to construct remote sensing monitoring models for goji berry pests and diseases, including monadic regression (MR), multiple linear regression (MLR), and fully connected neural network (FCN), with a comparative analysis conducted thereafter. The models were validated using the coefficient of determination ( $R^2$ ) and the root mean square error (RMSE), where an  $R^2$  closer to 1 and a smaller RMSE indicate a better predictive performance of the model [24].

## 3. Discussion

### 3.1. Spectral Reflectance Detection Results

According to the Ningxia Hui Autonomous Region Local Standard DB64/T 852-2023, plants infected with psyllids and gall mites are classified into three levels of damage: light,

moderate, and severe. The canopy reflectance of goji trees infected with goji psyllids and goji gall mites are shown in Figures 4a and 4b, respectively [25].



**Figure 4.** Spectral reflectance measurement results of goji canopies (a) for goji psyllids and (b) for goji gall mites.

In Figure 4, it is evident that goji exhibits typical vegetation spectral characteristics, with a small reflection peak around 550 nm in the green light band. Due to chlorophyll absorption, there is strong absorption of red light, initially leading to low reflectance around 680 nm, which then increases, reaching a peak near 800 nm and reflecting strong near-infrared radiation intensely.

The canopy spectrum of psyllid-infected goji in the blue, green, and red light bands between 600 and 700 nm shows no significant difference in reflectance; however, in the red light band from 700 to 760 nm and in the near-infrared band, as the infestation level increases, the spectral reflectance decreases. This is because the nymphs and adults of psyllids adhere to the goji leaves, damaging the internal structure of the leaves, thereby reducing the reflectance in the near-infrared band.

For goji infected with gall mites, the canopy spectrum in the blue and green light bands between 500 and 550 nm shows no significant difference in reflectance. However, in the green light band from 550 to 600 nm and in the red light band from 600 to 700 nm, as the infestation level increases, the spectral reflectance rises; in the red light band from 700 to 760 nm and in the near-infrared band, as the infestation level increases, the spectral reflectance decreases. This occurs because the mites damage the leaves, turning the affected areas from green to reddish-brown and, in severe cases, to purple or black. In the visible light band, the destruction of chlorophyll in the leaves causes an increase in spectral reflectance, while in the near-infrared band, the mites cause the leaves to form patchy galls, similarly damaging the internal structure, thus reducing reflectance in this band [26].

### 3.2. Correlation Analysis of Disease Index

From Figure 4, it is evident that goji diseases and pests are sensitive to the red light band from 700 to 760 nm and the near-infrared band. Therefore, we calculate the spectral reflectance parameters for the sensitive bands at 700, 760, 850, 955, and 975 nm.

Using SPSS 17.0 software and a Pearson two-tailed test, a correlation analysis was conducted on the hyperspectral characteristic parameters in Table 3, the vegetation indices in Table 4, and the selected sensitive bands with respect to the disease index. The results are shown in Table 5.

The results from Table 5 indicate that whether infected with psyllids or gall mites, the disease index of goji pests and diseases correlates strongly with the sum of the first-order derivative values within the blue edge, yellow edge wavelength range, or normalized value of the sum of the first-order derivatives within the red edge relative to that within the green edge, GNDVI and first-order derivative of  $R_{700}$ .

**Table 5.** Correlation analysis results between hyperspectral feature parameters and sensitive bands and the disease index of goji psyllids and goji gall mites.

Characteristic Parameter and Sensitive Bands	DI			
	Goji Psyllids		Goji Gall Mites	
	Pearson Correlation	Sig.	Pearson Correlation	Sig.
SD <sub>b</sub>	0.985 **	0.002	0.996 **	0
SD <sub>g</sub>	−0.921 *	0.026	−0.94 *	0.018
SD <sub>y</sub>	0.965 **	0.008	0.962 **	0.009
SD <sub>r</sub>	0.893 *	0.041	0.951 *	0.013
SD <sub>r</sub> /SD <sub>b</sub>	0.405	0.499	−0.149	0.811
SD <sub>r</sub> /SD <sub>g</sub>	0.753	0.142	−0.22	0.722
(SD <sub>r</sub> − SD <sub>b</sub> )/(SD <sub>r</sub> + SD <sub>b</sub> )	0.443	0.455	−0.051	0.935
(SD <sub>r</sub> − SD <sub>g</sub> )/(SD <sub>r</sub> + SD <sub>g</sub> )	0.977 **	0.004	0.975 **	0.005
PSSR <sub>a</sub>	0.483	0.41	0.501	0.39
PSSR <sub>b</sub>	0.059	0.925	0.789	0.113
PSSR <sub>c</sub>	0.725	0.166	0.448	0.449
GNDVI	−0.983 **	0.003	−0.99 **	0.001
first-order derivative of R <sub>700</sub>	−0.993 **	0.001	−0.99 **	0.001
first-order derivative of R <sub>760</sub>	−0.767	0.13	−0.561	0.325
R <sub>850</sub>	−0.945 *	0.015	0.62	0.264
R <sub>975</sub> /R <sub>955</sub>	0.89 *	0.043	0.144	0.817
NDVI	0.853	0.066	0.863	0.06
GNDVI	−0.983 **	0.003	−0.99 **	0.001
MTVI	0.508	0.382	0.56	0.326
mNDVI <sub>705</sub>	0.654	0.232	0.623	0.262
mSR <sub>705</sub>	−0.676	0.21	−0.777	0.122
RGRI	−0.533	0.355	−0.227	0.713
TVI	−0.472	0.422	−0.561	0.325

Note: “\*\*” indicates significant correlation, while “\*\*\*” indicates highly significant correlation.

### 3.3. Univariate Linear Regression Model

The construction of regression equations holds significant importance in statistical analysis. By using regression equations, one can predict the dependent variable using known values of independent variables, which is very useful for estimating and predicting future events or phenomena.

The univariate linear regression model is the simplest and most basic model in regression equation formulation [27,28]. Based on the highly significant disease index parameters in Table 5, univariate linear regression models are constructed for goji psyllids and goji gall mites. The results are shown in Table 6.

**Table 6.** Univariate linear regression models of goji psyllids and goji gall mites with respect to disease index.

Types of Diseases and Pests	Simulation Equation	R <sup>2</sup>	Sig.
goji psyllids	DI = 10.119 − 1264.591 × first-order derivative of R <sub>700</sub>	0.986	0.001
	DI = 13.587 − 520.533 × SD <sub>b</sub>	0.967	0.003
	DI = 8.264 + 1496.754 × SD <sub>g</sub>	0.827	0.016
goji gall mites	DI = 5.732 − 474.858 × first-order derivative of R <sub>700</sub>	0.983	0.001
	DI = 5.959 − 173.345 × SD <sub>b</sub>	0.98	0.001
	DI = 6.072 − 155.721 × SD <sub>g</sub>	0.972	0.002

A higher R<sup>2</sup> coefficient, closer to 1, indicates better model fit; a smaller standard error (Sig.) value indicates better regression performance. According to Table 5, the regression models using the first-order derivative of R<sub>700</sub> as variables proved to be the best for both goji psyllids and goji gall mites.

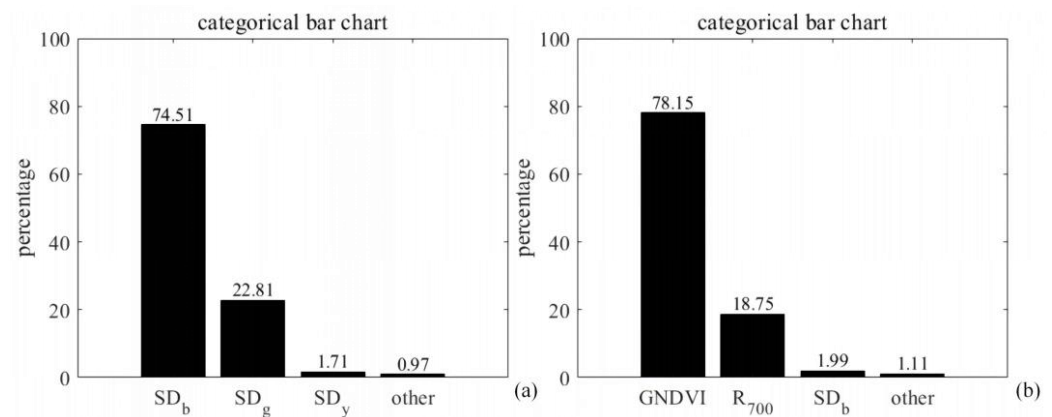
### 3.4. Multiple Linear Regression

When multiple independent variables are involved, some factors may not significantly impact the dependent variable, and there might be interactions among the independent variables, preventing them from independently influencing the dependent variable. In such cases, a stepwise regression analysis is used to filter the independent variables, identifying a subset of variables that significantly and independently affect the dependent variable [29]. Based on the highly significant disease index parameters in Table 5, multivariate stepwise regression models are constructed for goji psyllids and goji gall mites. The results are shown in Table 7.

**Table 7.** Multivariate stepwise regression models for the disease index of goji psyllids and goji gall mites.

Types of Diseases and Pests	Simulation Equation	R <sup>2</sup>	Sig.
goji psyllids	$DI = 13.587 - 520.533 \times SD_b$	0.976	0.003
	$DI = 12.267 - 614.427 \times SD_b - 176.924 \times SD_y$	0.985	0.001
goji gall mites	$DI = 5.732 - 474.858 \times GNDVI$	0.972	0.003
	$DI = 6.223 - 544.749 \times \text{first-order derivative of } R_{700} - 209.454 \times GNDVI$	0.988	0.001

To verify whether multiple stepwise regression correctly eliminates the interaction between variables, a principal component analysis (PCA) was employed. PCA aims to transform data from a high-dimensional space to a lower-dimensional space through linear transformations while preserving as much variability as possible. This method allows the dataset in the lower-dimensional space to exhibit maximum variance, with the direction of maximum variance representing the linear direction of the majority of the data. The chart in Figure 5 illustrates the proportion of each eigenvalue's contribution to the total eigenvalues (variance contribution) for the pests and diseases of goji berries related to the psyllids and gall mites.



**Figure 5.** Contribution rate of a single feature parameter to the total feature parameters. (a) for goji psyllids and (b) for goji gall mites.

In Figure 5, it can be observed that the cumulative contribution rates of the first principal component (F1) and the second principal component (F2) exceed 90%. This indicates that a significant amount of the original variables' limited information has been retained, thereby allowing the construction of multiple regression models for the pest and disease of goji psyllids and goji gall mites using the first and second principal components, respectively.



### 3.5. Fully Connected Neural Network Model

Fully connected neural networks are a fundamental type of neural network architecture, in which each neuron is connected to every neuron in the previous layer. They learn patterns and relationships from input data for accurate predictions by continuously adjusting connection weights and biases and utilizing activation functions. Input signals are passed to the output layer through forward propagation, and parameters are adjusted via backward propagation to minimize prediction errors [30,31].

Fully connected neural networks (also known as dense neural networks) are a basic architecture of neural networks, in which each neuron in one layer is connected to every neuron in the next layer. Here, the fundamental mathematical principles of fully connected neural networks will be discussed through the following key components: forward propagation, loss functions, and backpropagation.

#### 3.5.1. Forward Propagation

Consider a simple fully connected neural network with one hidden layer. The network structure is as follows: input layer with  $n$  inputs:  $x_1, x_2, \dots, x_n$ ; hidden layer with  $h$  neurons; and output layer with  $m$  outputs.

1. Hidden layer: Each neuron  $j$  in the hidden layer receives a weighted sum of all inputs from the input layer plus a bias, as follows:

$$z_j = \sum_{i=1}^n w_{ij}x_i + b_j \quad (1)$$

where  $w_{ij}$  is the weight from input neuron  $i$  to hidden neuron  $j$ , and  $b_j$  is the bias for hidden neuron  $j$ .

2. Apply an activation function  $f$  (such as ReLU) to each input of the hidden neurons to obtain the following output:

$$a_j = f(z_j) \quad (2)$$

3. Output layer: Each neuron  $k$  in the output layer similarly receives a weighted sum of all outputs from the hidden layer plus a bias, as follows:

$$y_k = \sum_{j=1}^h v_{jk}a_j + c_k \quad (3)$$

where  $v_{jk}$  is the weight from hidden neuron  $j$  to output neuron  $k$ , and  $c_k$  is the bias for output neuron  $k$ .

#### 3.5.2. Loss Functions

Commonly used loss functions include mean squared error (for regression problems) and cross-entropy loss (for classification problems).

1. The mean squared error is as follows:

$$L = \frac{1}{2} \sum_{k=1}^m (y_k - t_k)^2 \quad (4)$$

where  $t_k$  is the target output.

2. The cross-entropy loss is as follows:

$$L = - \sum_{k=1}^m [t_k \log(y_k) + (1 - t_k) \log(1 - y_k)] \quad (5)$$

### 3.5.3. Backpropagation

Backpropagation is used to compute the gradient of the loss function with respect to each weight, allowing for the use of gradient descent methods to update the weights.

1. Compute the error term for the output layer as follows:

$$\delta_k = (y_k - t_k) \cdot f'(y_k) \quad (6)$$

2. Propagate the error term to the hidden layer as follows:

$$\delta_j = \left( \sum_{k=1}^m v_{jk} \delta_k \right) \cdot f'(z_j) \quad (7)$$

3. Update weights and biases as follows:

- (1) Update weights in the output layer as follows:

$$v_{jk} = v_{jk} - \eta \cdot a_j \cdot \delta_k \quad (8)$$

- (2) Update weights in the hidden layer as follows:

$$w_{ij} = w_{ij} - \eta \cdot x_i \cdot \delta_j \quad (9)$$

where  $\eta$  is the learning rate.

### 3.5.4. Application of Fully Connected Neural Networks in the Detection of Diseases and Pests in Goji Berries

The fully connected neural network consists of several key components tailored for processing goji berry hyperspectral data. The network architecture starts with a feature input layer designed to receive 13 input features, reflecting the dimensionality of the input data. This is followed by a series of fully connected layers, each interleaved with ReLU activation layers to introduce non-linearity and aid in preventing vanishing gradients. The network culminates in a fully connected output layer with four neurons corresponding to the classification categories, topped off with a softmax layer to convert the output into a probability distribution over the predicted categories. The final piece is a classification layer, which interprets the softmax output to produce a categorical prediction.

The hidden layers of a fully connected neural network are one of the key factors influencing the performance of the network. The number of hidden layers and the number of neurons within them directly affect the model's training and inference times. A three-layer architecture is typically a reasonable choice, as it provides sufficient flexibility and expressive capability to capture nonlinear relationships within the data. This structure can maintain high performance while avoiding excessive computational costs, thereby reducing network complexity and the risk of overfitting.

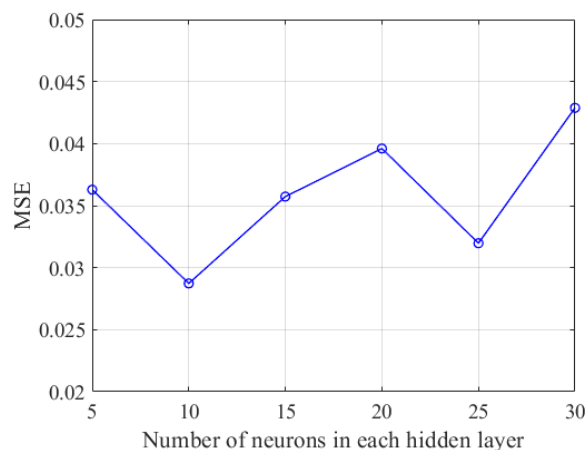
The number of neurons in the hidden layers directly determines the network's learning capacity and performance. A higher number of neurons increases the model's capacity to learn and express more complex patterns and features. By augmenting the number of neurons, the network can better fit the training data in a higher-dimensional space, thus enhancing model performance. However, for small datasets, an excessive number of neurons may lead to overfitting, in which the model performs well on the training set but poorly on the test set. Conversely, too few neurons may prevent the model from adequately learning the features of the training data, resulting in underfitting. Therefore, the careful selection of the number of neurons is essential to achieve good generalization ability [32,33].

To address this issue, the mean squared error (MSE) can be calculated to determine the optimal number of neurons in the hidden layers. A smaller MSE indicates better predictive performance. If a particular configuration yields the lowest MSE on the validation set

without signs of overfitting, it is generally considered the optimal configuration. The formula for mean squared error is as follows:

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^N (E_i - P_i)^2 \quad (10)$$

where  $E_i$  is the actual value,  $P_i$  is the predicted value, and  $N$  is the total number of samples. As shown in Figure 6, when the number of neurons is set to 15, the MSE is minimized, indicating that the neural network architecture exhibits the best performance in this configuration.

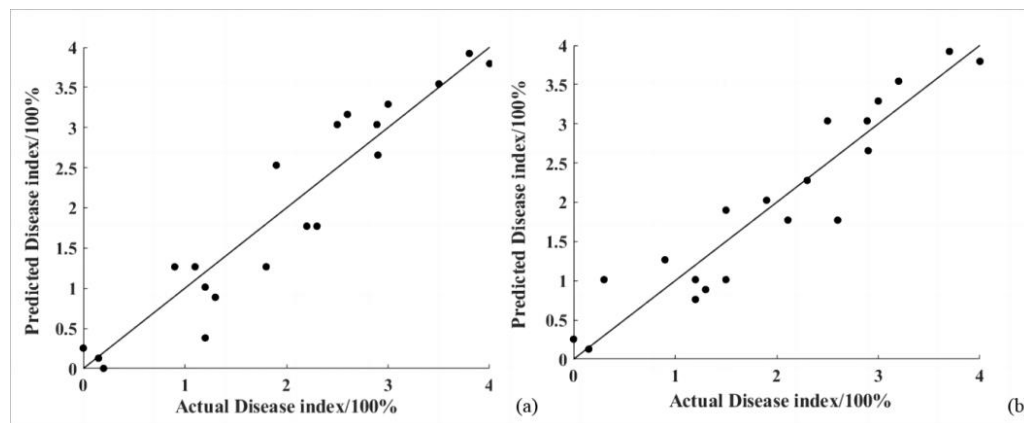


**Figure 6.** Line chart of MSE variations with changes in the number of neurons in the hidden layer.

## 4. Results

### 4.1. Univariate Linear Regression Model Validation

Near-ground hyperspectral data were collected from a total of 544 samples composed of 192 samples of goji psyllids, 272 samples of goji gall mites, and 80 healthy goji samples. Twenty randomly selected near-ground hyperspectral measurements were used as a validation set, and the disease index predictions for these 20 data points were calculated using the estimated univariate linear regression models with parameter  $R_{700}'$ 's first derivative as the variable. A 1:1 scatter plot of the predicted versus actual values was created. Finally, the model's performance was evaluated comprehensively using the coefficient of determination and root mean square error as metrics. The results are shown in Figure 7.



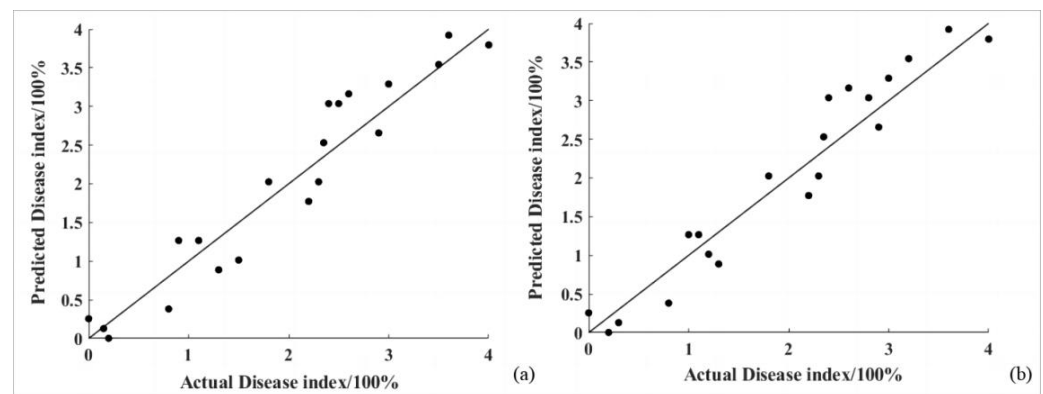
**Figure 7.** The 1:1 line plot of predicted disease indices with measured values under univariate linear regression model (a) for goji psyllids and (b) for goji gall mites.

For the goji psyllid pest disease, the univariate linear regression model with parameter  $R_{700}$ 's first derivative as the independent variable had a coefficient of determination of 0.8860 and a root mean square error of 0.3946.

For the goji gall mite pest disease, the univariate linear regression model with parameter  $R_{700}$ 's first derivative as the independent variable had a coefficient of determination of 0.8827 and a root mean square error of 0.3860.

#### 4.2. Multivariate Linear Regression Model Validation

The validation set was composed of 20 actual disease index values randomly selected from the near-ground remote sensing data. The predicted values for these 20 disease index data points were calculated using the binary linear regression model obtained in Section 4.2. A scatter plot was created to illustrate the predicted values against the actual values in a 1:1 ratio. The performance of the model was comprehensively evaluated using the coefficient of determination and the root mean square error as metrics. The results are shown in Figure 8.



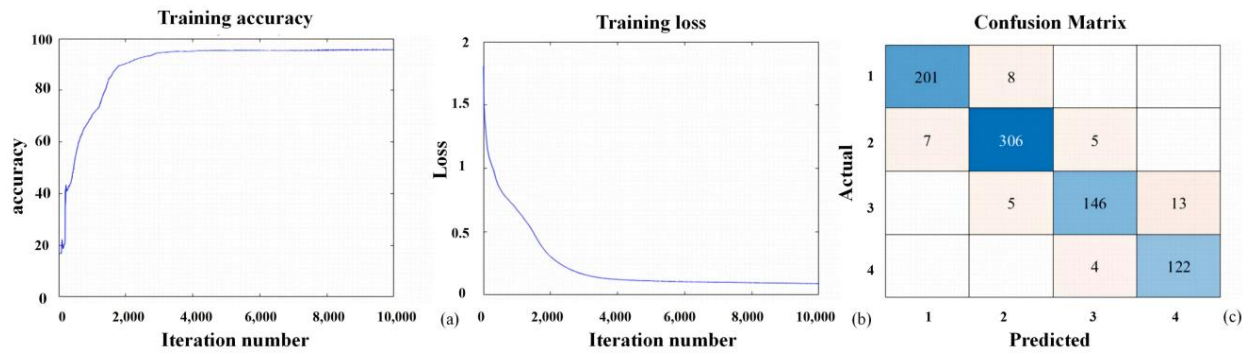
**Figure 8.** The 1:1 line plot of predicted disease indices with measured values under multiple linear regression model (a) for goji psyllids and (b) for goji gall mites.

The goji psyllid pest disease binary linear regression model, with  $SD_b$  and  $SD_y$  as the independent variables, had a coefficient of determination of 0.9059 and a root mean square error of 0.3531.

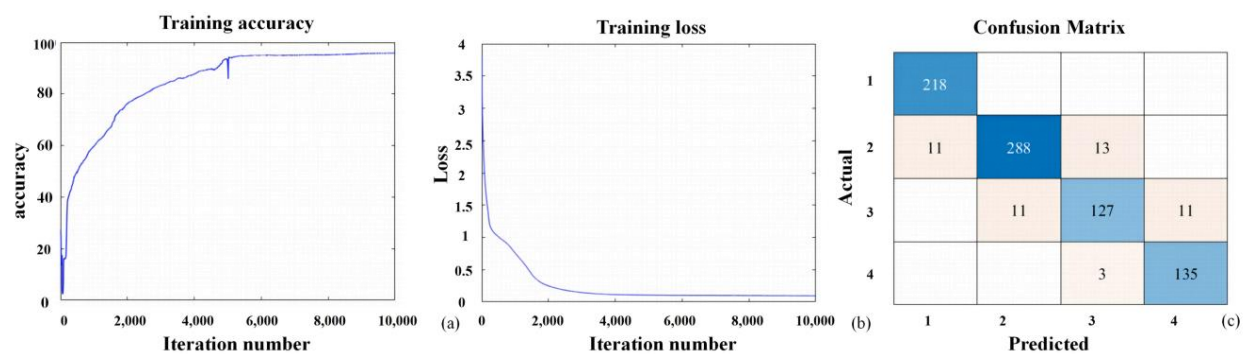
The goji gall mite pest disease binary linear regression model, with GNDVI and the first derivative of  $R_{700}$  as the independent variables, had a coefficient of determination of 0.9168 and a root mean square error of 0.3271.

#### 4.3. Fully Connected Neural Network Model Validation

The hyperspectral image data and near-ground remote sensing data acquired by the drone were annotated based on the corresponding disease index data recorded during the experiment. The pixels associated with the tagged plants were categorized into four levels: healthy, mild, moderate, and severe, represented by the numbers 1, 2, 3, and 4, respectively. Thirteen feature parameters with the highest correlations were selected as the input layer, while the disease index (DI) served as the output layer. Three hidden layers were chosen, each containing 15 neurons. To reduce the training time of the neural network, the dataset underwent random sampling. The resulting dataset was split into training and validation sets at a ratio of 7:3, with a maximum training iteration set to 10,000 rounds to establish the neural network model. The results indicate that the validation accuracy of the fully connected neural network classification model for the goji psyllids is 94.86%, while that for the goji gall mites is 96.82%. The training accuracy, loss, and confusion matrix of the fully connected neural network model constructed for the goji psyllids and goji gall mites are shown in Figures 9 and 10, respectively.

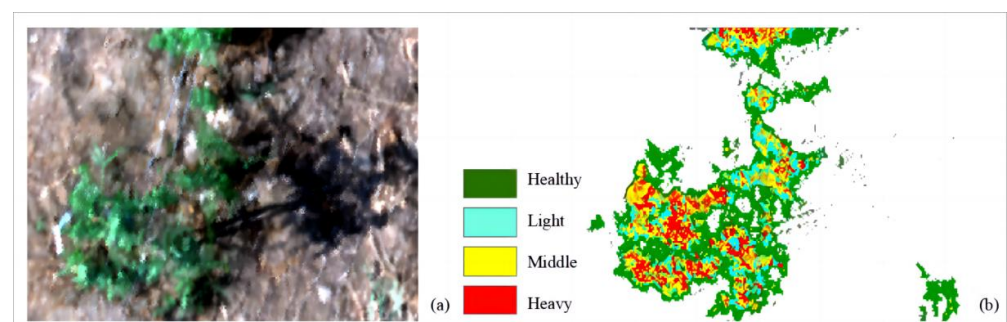


**Figure 9.** Model of goji psyllid pests and diseases constructed under a fully connected neural network: (a) accuracy; (b) loss value; (c) confusion matrix. The highest accuracy in the confusion matrix is represented in dark blue, while the rest are shown in light blue.



**Figure 10.** Model of goji gall mite pests and diseases constructed under a fully connected neural network: (a) accuracy; (b) loss value; (c) confusion matrix. The highest accuracy in the confusion matrix is represented in dark blue, while the rest are shown in light blue.

Display the classification results of a randomly selected high-spectral image of goji infected with psyllid pests using fully connected neural networks. The hyperspectral image is shown in Figure 11a. Extract spectral data from each pixel of the hyperspectral image and compute its feature parameter values, preserving spatial information of the pixels. Use these feature parameters as input to the fully connected neural network, which outputs the classification results. Different colors annotate the severity of psyllid infestation on goji, visually presenting the predicted classification results onto the hyperspectral image. The classification results of the fully connected neural network are depicted in Figure 11b.



**Figure 11.** Classification results diagram of goji psyllids using a fully connected neural network. (a) Hyperspectral image of goji psyllids; (b) classification results.

## 5. Conclusions

In the current research on monitoring diseases and pests affecting goji berries, most studies are still based on fundamental multiple linear regression models, primarily utilizing



near-ground spectral remote sensing data. This paper employs a fully connected neural network to enhance the accuracy of the goji berry disease and pest prediction model, combining near-ground spectral remote sensing with drone data for the first time.

By exploring the relationships between hyperspectral feature parameters, sensitive band variables, vegetation indices, and the disease indices of the goji psyllids and goji gall mites through drone remote sensing data, highly significant correlated variables were used to construct univariate linear regression models, multiple linear regression models, and fully connected neural network models. The predictive performance of these three models was then validated using near-ground remote sensing data. The consistency between drone remote sensing and near-ground spectral remote sensing was confirmed. Due to the rapid and non-destructive detection capabilities of UAV hyperspectral technology, it holds significant promise for future agricultural crop monitoring applications.

The model based on the fully connected neural network exhibited optimal performance, achieving a prediction accuracy of 94.86% for goji psyllids and 96.82% for goji gall mites, thereby enhancing the accuracy of the goji berry disease and pest prediction model. Finally, by employing pixel-based classification, the predicted severity of goji berry diseases was directly visualized on hyperspectral images, providing a theoretical foundation and reference standard for the establishment of rapid monitoring models for goji berry diseases and pests in the field.

**Author Contributions:** Investigation, B.Z. (Baoyu Zhou), B.Z. (Biyun Zhang), Z.C., K.K. and F.T.; data curation, N.C. and R.Z.; writing—original draft preparation, R.Z. and N.C.; writing—review and editing, R.Z., B.Z. (Biyun Zhang), C.Z. and N.C.; supervision, C.Z.; project administration, C.Z. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the Major International (Regional) Joint Research Project of National Natural Science Foundation of China (42020104008), the Key Program of National Natural Science Foundation of China (41530422), the Shaanxi Fundamental Science Research Project for Mathematics and Physics (Grant No. 22JSZ007), the National High Technology Research and Development Program of China (863 Program) (Grant No. 2012AA121101), the National Key R&D Program of China (Grant No. 2023YFB3906000), and the Key Research and Development Plan of Ningxia Hui Autonomous Region (2022BBF03015).

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

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data will be available upon request.

**Acknowledgments:** We sincerely thank Xuejian Zhang for his theoretical guidance throughout the project, Yong Ma for their active communication and assistance with drone detection and goji berry experimental field selection during the Ningxia fieldwork, and Min Gao for her guidance in selecting specific goji berry plants infected with pests and diseases. We extend gracious thanks to the farmers who permitted access to their fields for data collection.

**Conflicts of Interest:** Author Biyun Zhang was employed by the company BA trading Co., Ltd. Author Baoyu Zhou was employed by the company Ningxia Bing He Technology Co., Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## References

1. Xue, Q.F.; Zou, G.; Xie, M.T. Occurrence and control measures of main pests and diseases of Goji berry. *Plant Health Med.* **2011**, *24*, 29–31.
2. Zhang, G.L. Pest and disease control techniques for Goji berry. *Contemp. Hortic.* **2013**, *24*, 69.
3. Nilsson, H.E. Remote Sensing of Oil Seed Rape Infected by Sclerotinia Stem Rot and Verticillium Wilt. *Sver. Lantbruksuniv* **1985**, *33*, 231998.
4. Anand, T.; Sinha, S.; Mandal, M. AgriSegNet: Deep Aerial Semantic Segmentation Framework for IoT-Assisted Precision Agriculture. *IEEE Sens. J.* **2021**, *21*, 17581–17590. [[CrossRef](#)]
5. Hu, L.B.; Lan, Y.B.; Zhang, S.L. Research Progress on Remote Sensing Monitoring of Pests and Diseases in Cotton. *Shandong Agric. Sci.* **2024**, *56*, 164–171.

6. Whetton, R.L.; Hassall, K.L.; Waive, T.W. Hyperspectral measurements of yellow rust and fusarium head blight in cereal crops: Part 1: Laboratory study. *Biosyst. Eng.* **2018**, *166*, 101–115. [[CrossRef](#)]
7. Wang, X.L.; Deng, J.Z.; Huang, H.S. Identification of pests in cotton field based on hyperspectral data. *J. South China Agric. Univ.* **2019**, *40*, 97–103.
8. Lakshmi, V.; Ramesh, A.; Teja, Y.S. An efficient frame-work for disease detection and classification in cotton plants. *J. Eng. Sci.* **2021**, *12*, 520–525.
9. Dilixiatl, Y.; Zhou, J.P.; Xu, Y. Cotton pest monitoring based on Logistic algorithm and remote sensing image. *China Agric. Univ.* **2022**, *43*, 87–95.
10. Wu, W.L.; Liu, G.; Ma, Q. Adaptability Analysis of *Aceria macrodonis* Keifer in Ningxia Based on Maxent Model. *Biol. J. Mt. Agric. Biol.* **2021**, *41*, 66–70.
11. Ge, G.X. The Study of Monitoring Winter Wheat Growth and *Rhizoctonia Solani* Based on Remote Sensing. Master's Thesis, Nanjing University of Information Science & Technology, Nanjing, China, 2015.
12. Chu, X.L.; Yuan, H.F.; Lu, W.Z. Progress and application of spectral data pretreatment and wavelength selection methods in NIR analytical technique. *Prog. Chem.* **2004**, *16*, 528–542.
13. Tsai, F.; Philpot, W. Derivative analysis of hyperspectral data. *Remote Sens. Environ.* **1998**, *66*, 41–51. [[CrossRef](#)]
14. Xie, Y.P.; Chen, F.N.; Zhang, J.C.; Zhou, P.; Wang, H.J.; Wu, K.H. Study on monitoring of common diseases of crops based on hyperspectral technology. *Spectrosc. Spect. Anal.* **2018**, *38*, 2233–2240.
15. Shen, W.Y. Inversion of Winter Wheat Powdery Mildew Based on Hyperspectral Remote Sensing. Master's Thesis, Nanjing University of Information Science & Technology, Nanjing, China, 2016.
16. Pu, R.L.; Gong, P. *Hyperspectral Remote Sensing and Its Applications*; Higher Education Press: Beijing, China, 2000.
17. Luo, J.; Huang, W.; Zhao, J.; Zhang, J.; Zhao, C. Detecting aphid density of winter wheat leaf using hyperspectral measurements. *Earth Obs. Remote Sens.* **2013**, *6*, 690–698. [[CrossRef](#)]
18. Wang, S.Q.; Li, W.D.; Li, J.; Liu, X.S. Prediction of soil texture using FT-NIR spectroscopy and PXRF spectrometry with data fusion. *Soil. Sci.* **2013**, *178*, 626–638. [[CrossRef](#)]
19. Gumma, M.K. Mapping rice areas of south asia using MODIS multitemporal data. *Remote Sens.* **2011**, *5*, 053547. [[CrossRef](#)]
20. Candiago, S.; Remondino, F.; De Giglio, M.; Dubbini, M.; Gattelli, M. Evaluating multispectral images and vegetation indices for precision farming applications from UAV images. *Remote Sens.* **2015**, *7*, 4026–4047. [[CrossRef](#)]
21. Haboudane, D. Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modeling and validation in the context of precision agriculture. *Remote Sens. Environ.* **2004**, *90*, 337–352. [[CrossRef](#)]
22. Guo, R.; Zhu, X.; Liu, T. Automatic detection of crop lodging from multitemporal satellite data based on the isolation forest algorithm. *Comput. Electron. Agric.* **2023**, *215*, 108415. [[CrossRef](#)]
23. Castro, K.; Sanchez-Azofeifa, G. Changes in spectral properties, chlorophyll content and internal mesophyll structure of senescing populus balsamifera and populus tremuloides leaves. *Sensors* **2008**, *8*, 51–69. [[CrossRef](#)]
24. Liang, W.; Li, Y.J.; Cen, H.Y.; Zhu, J.P.; Yin, W.X.; Wu, W.K.; Zhu, H.Y.; Sun, D.W.; Zhou, W.J.; He, Y. Combining UAV-based vegetation indices and image classification to estimate flower number in oilseed rape. *Remote Sens.* **2018**, *10*, 1484. [[CrossRef](#)]
25. Fang, W.; Rong, Z.; Wei, S. *Technical Regulation for Monitoring and Forecasting of Goji Insect Pests and Diseases*; Ningxia Academy of Agricultural and Forestry Sciences: Yinchuan, China, 2023.
26. Ekwe, M.C.; Adeluyi, O.; Verrelst, J.; Kross, A.; Odiji, C. Estimating rainfed groundnut's leaf area index using sentinel-2 based on machine learning regression algorithms and empirical models. *Precis. Agric.* **2024**, *25*, 1404–1428. [[CrossRef](#)]
27. Sims, D.A.; Gamon, J.A. Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages. *Remote Sens. Environ.* **2002**, *81*, 337–354. [[CrossRef](#)]
28. Shi, Y.J.; Chen, P.F. Maize above-ground biomass retrieval using unmanned aerial vehicle (UAV) hyperspectral remote sensing imagery. *Chin. Agric. Sci. Bull.* **2019**, *35*, 117–123.
29. You, S.B.; Yan, Y. Stepwise regression analysis and its application. *Stat. Decis.* **2017**, *14*, 31–35.
30. Hornik, K.; Stinchcombe, M.; White, H. Multilayer feedforward networks are universal approximators. *Neural Netw.* **1989**, *2*, 359–366. [[CrossRef](#)]
31. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. In Proceedings of the 3rd International Conference on Learning Representations ICLR 2015, San Diego, CA, USA, 7–9 May 2015.
32. Tian, X.J.; Gong, H.; Tu, Y. Die Casting Quality Prediction Algorithm Based on Fully Connected Neural Network. *J. Netshape Form. Eng.* **2024**, *16*, 159–164.
33. He, T.; Zhou, N.; Wu, X.Y. Thickness Prediction of Reservoir Effective Sand Body by Deep Fully Connected Neural Network. *J. Jilin Univ. (Earth Sci. Ed.)* **2023**, *53*, 1262–1274.

**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.