

## Article

# Non-Destructive Detection of Different Pesticide Residues on the Surface of Hami Melon Classification Based on tHBA-ELM Algorithm and SWIR Hyperspectral Imaging

Yating Hu <sup>1</sup>, Benxue Ma <sup>1,2,3,\*</sup>, Huting Wang <sup>1,2,3</sup>, Yujie Li <sup>1</sup> , Yuanjia Zhang <sup>1</sup> and Guowei Yu <sup>1</sup> 

- <sup>1</sup> College of Mechanical and Electrical Engineering, Shihezi University, Shihezi 832003, China; hyt\_0207@163.com (Y.H.); xgb@shzu.edu.cn (H.W.); lyj\_shzdx@163.com (Y.L.); zyj\_shz@163.com (Y.Z.); ygw986119533@163.com (G.Y.)
- <sup>2</sup> Key Laboratory of Northwest Agricultural Equipment, Ministry of Agriculture and Rural Affairs, Shihezi 832003, China
- <sup>3</sup> Engineering Research Center for Production Mechanization of Oasis Characteristic Cash Crop, Ministry of Education, Shihezi 832003, China
- \* Correspondence: mbx\_mac@shzu.edu.cn

**Abstract:** In the field of safety detection of fruits and vegetables, how to conduct non-destructive detection of pesticide residues is still a pressing problem to be solved. In response to the high cost and destructive nature of existing chemical detection methods, this study explored the potential of identifying different pesticide residues on Hami melon by short-wave infrared (SWIR) (spectral range of 1000–2500 nm) hyperspectral imaging (HSI) technology combined with machine learning. Firstly, the classification effects of classical classification models, namely extreme learning machine (ELM), support vector machine (SVM), and partial least squares discriminant analysis (PLS-DA) on pesticide residues on Hami melon were compared, ELM was selected as the benchmark model for subsequent optimization. Then, the effects of different preprocessing treatments on ELM were compared and analyzed to determine the most suitable spectral preprocessing treatment. The ELM model optimized by Honey Badger Algorithm (HBA) with adaptive t-distribution mutation strategy (tHBA-ELM) was proposed to improve the detection accuracy for the detection of pesticide residues on Hami melon. The primitive HBA algorithm was optimized by using adaptive t-distribution, which improved the structure of the population and increased the convergence speed. Compared the classification results of tHBA-ELM with HBA-ELM and ELM model optimized by genetic algorithm (GA-ELM), the tHBA-ELM model can accurately identify whether there were pesticide residues and different types of pesticides. The accuracy, precision, sensitivity, and F1-score of the test set was 93.50%, 93.73%, 93.50%, and 0.9355, respectively. Metaheuristic optimization algorithms can improve the classification performance of classical machine learning classification models. Among all the models, the performance of tHBA-ELM was satisfactory. The results indicated that SWIR-HSI coupled with tHBA-ELM can be used for the non-destructive detection of pesticide residues on Hami melon, which provided the theoretical basis and technical reference for the detection of pesticide residues in other fruits and vegetables.

**Keywords:** pesticide residues; SWIR hyperspectral imaging; metaheuristic optimization; machine learning; non-destructive detection



**Citation:** Hu, Y.; Ma, B.; Wang, H.; Li, Y.; Zhang, Y.; Yu, G. Non-Destructive Detection of Different Pesticide Residues on the Surface of Hami Melon Classification Based on tHBA-ELM Algorithm and SWIR Hyperspectral Imaging. *Foods* **2023**, *12*, 1773. <https://doi.org/10.3390/foods12091773>

Academic Editor: Soledad Cerutti

Received: 20 March 2023

Revised: 18 April 2023

Accepted: 21 April 2023

Published: 25 April 2023



**Copyright:** © 2023 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

Hami melon is a kind of thick-skinned melon with delicious taste and rich nutritional value, which is a national geographical indication product. Hami melon is susceptible to powdery mildew, leaf blight, downy mildew, aphids, and other diseases during planting [1]. In order to prevent the above diseases, Hami melon growers often use Acetamiprid, Difenconazole, Chlorpyrifos, and other pesticides for prevention and control [2]. However,

in agricultural production, pesticides are often overused intentionally or unintentionally, which not only seriously threatens human health, but also causes environmental damage. Therefore, it is necessary to accurately detect pesticide residues on Hami melon.

The Food and Agriculture Organization of the United Nations (FAO) defines pesticides as substances or mixtures used to prevent, eliminate, or control harmful organisms that harm agriculture, as well as to purposefully regulate, control, and affect the metabolism, growth, and other processes of plants and harmful organisms [3]. The growing population has put forward a stronger demand for food, and the pesticide market has also increased significantly. With the increasing awareness of consumer health, people are increasingly focusing on food safety, more and more researchers begin to pay attention to the detection of pesticide residues in fruits [4]. In previous studies, techniques such as supercritical fluid chromatography (SFC), immunoassay method, liquid chromatography (LC), and gas chromatography-mass spectrometry (GC-MS) have been widely applied in the field of pesticide residue detection [5–8]. Although these traditional methods had high detection accuracy and sensitivity, they were time-consuming, costly, complex in operation, and dependent on many chemical reagents, which were wasteful and polluting, and were not conducive to popularization. What's more, traditional chemical detection methods are destructive, and the detected Hami melon could not be sold. Therefore, it is essential and significant to find a rapid and non-destructive method to detect pesticide residues on Hami melon.

Scholars have conducted related research to search the capacity of different non-destructive detection technologies, such as machine vision, electronic nose, acoustic technology, and spectral technology, for food quality assessment [9–11]. Among them, the spectroscopic technology depending on simple and non-destructive methods have attracted great attention [12]. Spectroscopic techniques, including near infrared spectroscopy, fluorescence spectroscopy, Raman spectroscopy, and terahertz time-domain spectroscopy, have been proved to be useful for the measurement of pesticide residues in food [13–17]. However, one of the main disadvantages of these spectral analysis methods is that the spectral data are obtained from a single point (or regional area). Although there is no destruction or preprocessing, the point source sampling method limits the integrity of collected data and cannot ensure that the collected data can completely and accurately express the required information, which easily leads to unsatisfactory pesticide residue detection effect.

In contrast, hyperspectral imaging (HSI) is considered as a potential technology to solve the above-mentioned defects [18,19]. HSI is a rapidly developing non-destructive testing technology, which has been gradually applied to agricultural products testing in recent years [20,21]. It has the advantages of no pollution, no damage, automation, high efficiency and so on. The spatial distribution spectral information at each pixel of the object can be obtained. The spectral and image information of the sample can be obtained simultaneously [22]. At present, the research regarding HSI technology in agricultural product quality detection is relatively mature [23].

Some scholars have used HSI to discriminate and assess the quality and safety of diverse foods, and obtained satisfactory results, such as defect detection of citrus [24], classification of mildew, health, and damage of peanuts [25], prediction of potential pest infection of apples [26], and identification of Cucumber Green Mottle Mosaic Virus (CGMMV) infection in watermelon seeds [27]. HSI technology has also been extensively employed in the study of pesticide residues. Sun et al. studied chlorpyrifos residues in mulberry leaves at different concentrations and combined hyperspectral imaging technology with SVR prediction model, obtaining good results with root mean square error (RMSE) and determination coefficient ( $R^2$ ) of prediction set of 27.719 and 0.874, respectively [28]. Relevant studies showed that HSI technology can be applied to identify pesticide residues in lettuce. Two models, CARS-SPA-LSSVR and RF-RFE-SPA-LSSVR, were used to detect fenvalerate and dimethoate, respectively, and good results were obtained [29]. Some scholars also applied this technology to detect the contamination of chlorpyrifos and imidacloprid in

edible jujube fruits, and achieved good results by using the ES-AWLSGSD-RC-LWPLSR model constructed by using only eight feature wavelengths [30]. However, these studies mainly used HSI in the spectral range of 400–1000 nm or 900–1700 nm to assess food quality and safety. Any substance has energy and produces unique light waves. Therefore, the light waves reflected by different substances are not the same. In recent years, the use of short-wave infrared (SWIR) spectral range (1000–2500 nm) has been gradually favored by researchers, such as the classification of apple potential scratches [31], the identification of peanut kernel mildew [32], and the rapid prediction of the water content of single cucumber seed [33] and the soluble solid content of Hami dried jujube [21]. Some studies have proved that the discovery of this spectral region was affected by the vibration and overtones of the main structural components of organic molecules [34]. Most pesticides are organic compounds and contain more organic molecules. So, the SWIR spectral range may be more suitable for the discrimination of pesticide residues on fruits. However, there are few studies using HSI to detect pesticide residues on thick-skinned fruits in the SWIR spectral range. Therefore, one of the objectives of this study is to investigate the possibilities of using SWIR-HSI for the accurate detection of pesticide residues on the surface of Hami melon.

The collected original spectrum often contains the random noise of the instrument. In addition, the introduction of baseline drift, light stray scattering, and other information will also affect the corresponding relationship between the spectrum information and the indicators to be measured, thus affecting the stability of the model. Therefore, the preprocessing of spectral data is a key step before modeling and analysis [35]. Preprocessing treatments were commonly used in spectral analysis, such as Savitzky–Golay smoothing (SGS), multiple scatter correction (MSC), and standard normal variate transformation (SNV), etc. These treatments were widely used in hyperspectral imaging to determine the quality of agricultural products [36,37]. Extreme learning machine (ELM), support vector machine (SVM), and partial least squares discriminant analysis (PLS-DA), three classical machine learning classification models, have been widely used in HSI field because of their simple classification ideas, strong learning ability, and good classification effect [38–40]. Considering that the model accuracy of classical machine learning model is affected by model parameters, data volume, and other factors, resulting in problems such as poor model stability and generalization, it is very important to optimize the classical classification model. The metaheuristic algorithm has many advantages, such as faster optimization speed, more effective search for the global optimal solution of complex optimization problems, and stronger stability and adaptability. The most typical problem dealt with by the metaheuristic algorithm is optimization. Therefore, some researchers tried to use the metaheuristic algorithm to adjust the traditional model, and it has been successfully applied in bacterial food-borne pathogen classification, pesticide residues in lettuce, rice mold colony detection, and other fields [41–43].

Therefore, the objective of this study was to apply HSI technology combined with the ELM model optimized by the metaheuristic algorithm to identify a single class of pesticide residues (None, Acetamiprid, Malathion, Difenconazole and Beta-cypermethrin). The main objectives of this study were: (1) to explore the feasibility of HSI technology based on SWIR spectral region (1000–2500 nm) for pesticide residues on Hami melon; (2) to assess the effectiveness of different preprocessing treatments for spectral data analysis; (3) to develop an ELM optimized by the adaptive t-distribution honey badger algorithm (tHBA) for identifying pesticide residues; and (4) to compare the optimization effects of different metaheuristic algorithms and the classification effects of the established models.

## 2. Materials and Methods

### 2.1. Samples Preparation

Four standard pesticides were purchased from a local agricultural raw material market in Shihezi, Xinjiang, China. Acetamiprid (Active ingredient content 70%, water dispersible granule, Shandong Baixin Biotechnology Co., Ltd., Jinan, China), Malathion (Active ingredient 70%, emulsifiable oil, Ningbo Sanjiang Yinong Chemical Co., LTD., Ningbo, China),

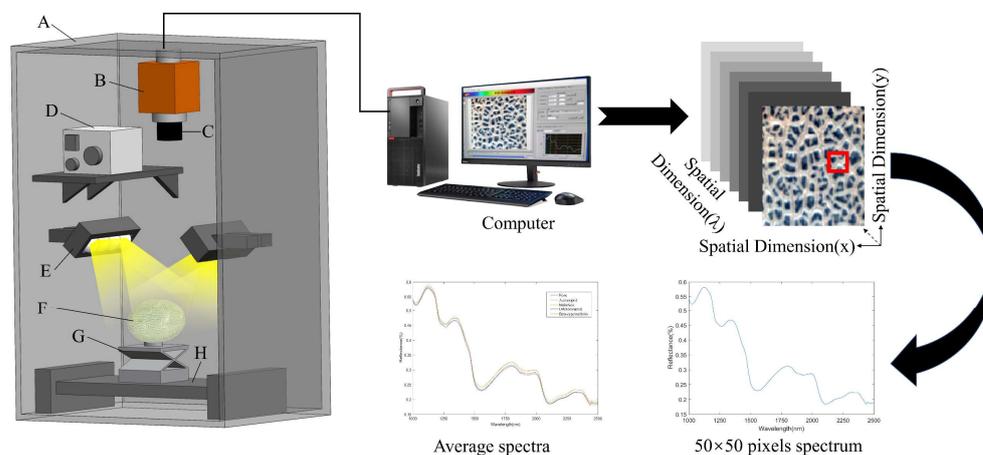
Difenoconazole (Active ingredient 20%, microemulsion, Chengdu Kelilong Biochemical Co., Ltd., Chengdu, China), and Beta-cypermethrin (Active ingredient 4.5%, emulsion, Jiangsu Yixing Xingnong Chemical Products Co., Ltd., Yixing, China) were used as research objects.

Hami melons (Xizhoumi No.25) were purchased from the local agricultural products trading center in Shihezi, Xinjiang, China, in August 2021. All samples were oval and weighed approximately 3–4 kg. Before collecting data, all Hami melons samples were numbered after wiping, and then reposed in a well-ventilated room at 22 °C and 40% relative humidity for 24 h to avoid being affected by the external environment. The 200 samples were arbitrarily selected and separated into five equal groups (40 samples per group). Then, Acetamiprid, Malathion, Difenoconazole, and Beta-cypermethrin were mixed with distilled water respectively to prepare 1:1000 pesticide solution, which was evenly sprayed on the surface of Hami melon, and recorded as group 1, group 2, group 3, and group 4. In addition, the remaining 40 Hami melon samples were used as the control group, and samples were homogeneously sprayed with distilled water, which was recorded as group 0. In the end, every sample was deposited indoors for 12 h.

## 2.2. Hyperspectral Imaging Acquisition and ROI Spectrum Acquisition

### 2.2.1. Hyperspectral Imaging System

The SWIR hyperspectral imaging system consists of an imaging spectrometer (ImSpector N25E23<sup>™</sup>, Specim, Oulu, Finland), a high-resolution camera (Zephyr-2.5–320, Photon Etc., Montreal, QC, Canada), two 150 W halogen surface light sources adjusted at an angle of about 45° to illuminate the camera's field of view, an electrically positioned sample stand operated by a stepper motor, a computer equipped with data acquisition software (Spectral Image System, Isuzu Optics Corp., Taiwan, China), and a box with black inner surface. Hyperspectral images of Hami melon samples containing pesticide residues were obtained in diffuse reflection mode, and the spectral range was 1000–2500 nm. The measurement was carried out in a dark room to avoid the interference of stray light. In addition, the SWIR hyperspectral imaging system is shown in Figure 1.



**Figure 1.** SWIR-HSI acquisition system and three-dimensional (3D) data cube. (A) Box with black inner surface; (B) Imaging Spectrograph; (C) Lens; (D) Light source regulator; (E) Halogen lamps; (F) Sample of Hami melon; (G) Lifting platform; (H) Electric moving stage.

### 2.2.2. Imaging Acquisition and Calibration

The instrument was preheated for 30 min before use, and the samples of Hami melon were collected on the electric moving stage. Proper adjustment of parameters was crucial for gaining perspicuous and distortion-free hyperspectral images. In this study, the correlation coefficient of HSI was set as follows. The exposure time of the camera was adjusted to 4.1 ms, and the speed of the electric mobile stage was adjusted to 53.4 mm/s. As the electric moving stage moved, the acquisition system simultaneously acquired the image and spectral information of the Hami melon, and repeated the operation until all the

sample information was obtained. In fact, the directly obtained hyperspectral images of Hami melon need to be black and white corrected first and cannot be immediately used in experiments. When collecting hyperspectral images, because the illumination direction of the light source was fixed, the light intensity on the sample surface changed greatly, and the existence of camera dark current can bring noise to the spectral data. Whiteboard data and blackboard data are reference data for reflectivity calibration objects. By the white and dark references, the gap of data with different light intensities can be decreased, so that the relevant spectra can be compared and analyzed under different environments and light intensities. The white and dark references principle is shown in the equation below:

$$R_c = \frac{R_r - R_d}{R_w - R_d} \quad (1)$$

where  $R_r$  is the raw spectral value,  $R_w$  is the all-white reference value,  $R_d$  is the all-black reference value, and  $R_c$  is the corrected value.

### 2.2.3. ROI Spectrum Acquisition

The image resolution of the hyperspectral camera is  $320 \times 256$  pixels, the spectral resolution is 6.20 nm. The spectral range is 982.38–2618.37 nm, which can collect hyperspectral images of 288 wavelengths. After edge band screening, the obtained band range was 1000–2500 nm, and the hyperspectral images of 233 wavelengths were retained. For each Hami melon, one hyperspectral image was collected every  $90^\circ$  rotation along the equatorial direction. So, 4 hyperspectral images can be collected from one Hami melon sample, and 800 hyperspectral images can be collected from 200 Hami melon samples. After the hyperspectral image information was collected, the hyperspectral image was imported into ENVI software, and pixel blocks of  $50 \times 50$  at the equatorial position of Hami melon were randomly designated as regions of interest (ROI). A total of 800 average spectra were collected, and these spectra were stored for the establishment of classification model.

### 2.3. Spectral Data Preprocessing

Preprocessing can remove as much extraneous noisy data as possible that has an interfering effect on the predicted results. Spectral information is often mixed with unwanted noise and background interference, such as stray light and electrical noise, which should be eliminated or reduced, which is the goal of spectral preprocessing [44]. In this study, we applied and compared three commonly used treatments, including MSC, NM and SNV, respectively. The purpose of MSC is to compensate the decrease of signal-to-noise ratio caused by light scattering, and the spectral data obtained after scattering correction can effectively eliminate the spectral differences caused by different scattering levels [45]. NM uses scaling, panning and other operations to convert the data into a specific interval, thereby eliminating the influence of data dimension and making the data indicators comparable [46]. SNV can eliminate the differences in individual samples due to factors such as particle size and can be used to correct spectral errors caused by scattering between samples [47].

### 2.4. Establishment of tHBA-ELM Model

#### 2.4.1. Honey Badger Algorithm

Currently, more and more researchers are trying to introduce various numerical optimization methods into their algorithms, among which metaheuristic optimization algorithms have been developed substantially. Enlightened by the foraging habits of honey badgers in nature, a more intelligent metaheuristic optimization algorithm, the honey badger algorithm (HBA) [48], has been proposed. This algorithm developed an effective search strategy for solving optimization problems in mathematics. In the HBA algorithm, HBA imitates the foraging behavior of badgers in nature and divides optimization into two modes: “digging” and “honey”. In the digging mode, the badger mainly depends on the odor intensity of the target, estimates the position of the target according to the odor, and

selects the appropriate position near the target for digging. In the honey mode, the badger mainly locates the target location directly according to the location of the honeyguide bird. HBA is effective in solving optimization problems with complex search space and has advantages in convergence speed and exploration-development balance. It has strong search ability and fast convergence speed.

The HBA algorithm is simple in design and clear in thought. The algorithm mainly consists of steps such as population initialization, fitness assessment, and honey attraction calculation, providing good local search capability based on honey attraction. In addition, the density factor is set to ensure the global searching ability of the algorithm. Through testing, the algorithm showed great properties in solving most test functions.

Compared with the common metaheuristic algorithm, the algorithm has the advantages of simple mechanism, less parameters, fast convergence speed, and has advantages in convergence speed and exploration and development balance. Therefore, this study intended to use HBA to optimize the parameters and optimize the classification model to improve the classification accuracy.

#### 2.4.2. Honey Badger Algorithm Based on Adaptive t-Distribution Mutation Strategy

Aiming at the issue that the metaheuristic algorithm tends to become trapped in local convergence, mutation operation is a common method used to leap out of local convergence. Introducing mutation operator is beneficial to further improve the ability of the algorithm to find the optimal solution. Because the introduction of mutation operator can make the algorithm have a certain local random search ability, on the one hand, it accelerates the convergence to the optimal solution in the late stage of solution, and on the other hand, it also maintains the diversity of solutions. The iteration times was regarded as the degrees of freedom parameter of t-distribution, and the adaptive t-distribution mutation method was used to disturb the individual position to improve the algorithm's ability [49]. This method was called tHBA.

Gaussian distribution (GD) and Cauchy distribution (CD) have been proven to effectively lift the optimization capability of the algorithm. GD can enhance the searching performance of individuals near the optimal point and accelerate the convergence speed of the algorithm. CD can enhance the search ability of individuals in the solution space and increase the diversity of the population, while t-distribution has the advantages of both GD and CD. As the iteration of the algorithm begins, the t-distribution gradually changes from similar to CD to similar to GD as the number of iterations increases. Therefore, the adaptive t-distribution mutation method was used to disturb the individual position and improve the ability of the algorithm to jump out of the local optimum. In this study, the t-distribution mutation operator with the degree of freedom parameter of  $iter$  as t-distribution is used to disturb the position of the badger, so that the algorithm has better global development capability in the early iteration and has better local exploration ability in the later iteration, and enhances the convergence speed of the algorithm. Specific location updates are as follows:

$$X_i^{t+1} = X_i^t + X_i^t \cdot t(iter) \quad (2)$$

where,  $X_i^{t+1}$  is the position of the disturbed badger, and  $X_i^t$  is the position of the badger  $i$  at the  $t$  iteration. The proposed update equation adds a random interference term  $X_i^t \cdot t(iter)$  on the basis of  $X_i^t$ , which helps the algorithm leap out of the local optimum by introducing random disturbance information on the basis of the current position information. With the number of iterations  $iter$  increased, t-distribution gradually draws close to the Gaussian distribution, which is useful to enhance the convergence speed. The addition of adaptive t-distribution mutation operator effectively enhances the property optimization of the algorithm and avoids local optimization.

The algorithm flow of tHBA is as follows, the flow chart for which is shown in the tHBA Algorithm part in Figure 2:

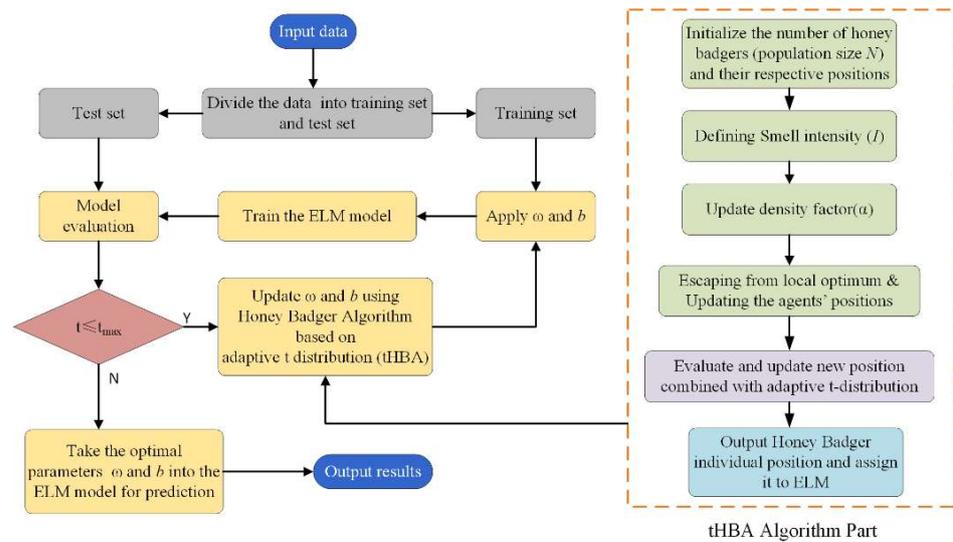


Figure 2. tHBA-ELM Algorithm Flow Chart.

- (1) Determine the tHBA algorithm parameters (population size  $N$ , mutation ratio, etc.), using Equation (3) to initialize the population.
- (2) Define smell intensity of prey  $I$ . Strength and prey concentration are related to the distance between it and the badger. If the smell intensity is strong, the exercise speed is higher, and the reverse is also true, as defined by Equation (4).
- (3) Update density factor ( $\alpha$ ). The value of the  $\alpha$  decays with the increase in the number of iterations to ensure the stability of the overall environment when the algorithm transitions from the exploration state to the exploitation state, as in Equation (5).
- (4) Escape from local optimum and update the agents' positions. HBA uses a sign  $F$  ( $F$ : is a sign to change the agent's search direction.) to change the search direction to take advantage of the high opportunity for agents to scan the search space strictly. The HBA location update process is separated into two modes: "digging" and "honey".
- (5) Generate the random number  $rand$  between the interval of 0 to 1. If  $rand < P$  ( $P$ : mutation probability) then according to Equation (2) adaptive  $t$ -distribution mutation, calculate the fitness value and update the badger position.
- (6) When the number of iterations exceeds  $t_{max}$ , the optimization process will terminate and output the global optimal position, namely the optimal values of  $\omega$  (connection weight between input layer and hidden layer) and  $b$  (neuron threshold of hidden layer) in ELM.

$$x_i = lb_i + r_1 \times (ub_i - lb_i) \tag{3}$$

where  $x_i$  is the coordinate of the  $i$ th badger individual,  $lb_i$  and  $ub_i$  are the upper and lower boundaries of the optimization space, and  $r_1$  is a random variable in the interval of 0 to 1.

$$\begin{cases} I_i = r_2 \times \frac{S}{4\pi d_i^2} \\ S = (x_i - x_{i+1})^2 \\ d_i = x_{Prey} - x_i \end{cases} \tag{4}$$

where  $r_2$  is a random variable in the interval of 0 to 1,  $S$  is the aggregation degree of honey, and  $d_i$  is the distance between the badger and honey.

$$\alpha = C \times \exp\left(\frac{-t}{t_{max}}\right) \tag{5}$$

where  $t_{max}$  is the maximum number of iterations,  $C$  is a constant  $\geq 1$  (default = 2).

### 2.4.3. ELM Classification Model Based on tHBA Optimization

ELM is a simple and reliable single hidden layer feedforward neural network learning algorithm. The improved neural network structure improves the learning speed and got a wide application. ELM is mainly composed of the input layer, output layer, and hidden layer in the middle. ELM only needs to set the appropriate number of hidden layer nodes, arbitrarily produce all the parameters required for the hidden layer, and finally use the least square method to determine the weight of the output layer. ELM possesses the advantages of low training complexity, high learning efficiency for targets, and high accuracy after generalization in comparison with traditional neural networks. In this study, the tHBA algorithm was used to search for optimal  $\omega$  and  $b$  in ELM. The flow chart is shown in Figure 2.

The procedure of the tHBA-ELM algorithm was as follows:

- (1) Input the preprocessed spectral data of Hami melon samples and divide the data into training set and test set according to the ratio of 3:1, with 600 spectra in training set and 200 spectra in test set. Then the training set was used as the training sample of tHBA-ELM model, and the test set was used as the model verification.
- (2) The position of the badger was updated by tHBA and assigned to  $\omega$  and  $b$  in ELM. The ELM model was trained by using the  $\omega$  and  $b$  obtained from each update.
- (3) Set the number of iterations to 300. When the number of iterations reached 300, tHBA output the best fitness value corresponding to the badger position, using the best  $\omega$  and  $b$  to establish the best ELM model, and output the model results.

### 2.5. Assessment Standard of Models

To evaluate and compare the effect of the model, four evaluation indexes (accuracy, precision, sensitivity, and F1-score) were selected to represent the performance of the model. Accuracy refers to the percentage of discriminated correct results in the total sample; precision refers to the probability of actually positive samples among all the discriminated positive samples; sensitivity refers to the probability of being discriminated as a positive sample among the actually positive samples. The F1-score combines the performance of precision and sensitivity. To find a balance between the two, an F1-score appears. These evaluation parameters are calculated as follows:

$$Accuracy = \frac{(TP + TN)}{(TP + FP + TN + FN)} \quad (6)$$

$$Precision = \frac{TP}{TP + FP} \quad (7)$$

$$Sensitivity = \frac{TP}{TP + FN} \quad (8)$$

$$F1\text{-score} = \frac{2 \times Precision \times Sensitivity}{Precision + Sensitivity} \quad (9)$$

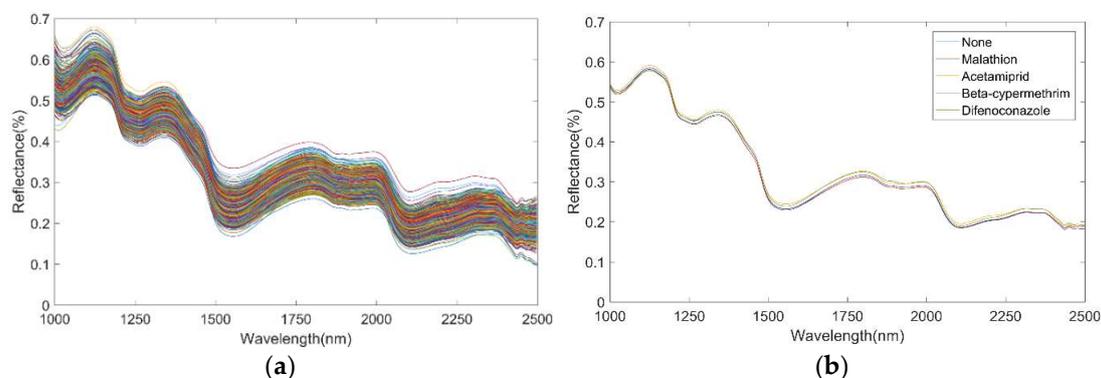
Among them,  $TP$ : true positive, positive samples are classified as positive samples.  $FP$ : false positive, negative samples are classified as positive samples.  $TN$ : true negative, negative samples are classified as negative samples.  $FN$ : false negative, positive samples are classified as negative samples.

## 3. Results

### 3.1. Spectral Characteristics of Pesticide Residues

The original and average diffuse reflectance spectra of different pesticide residues on Hami melon are shown in Figure 3. Figure 3a shows the original diffuse spectra and Figure 3b the average diffuse reflectance spectra. It can be observed from Figure 3 that the hyperspectral reflectance of various pesticide residues was varied, but the variation

tendency of spectral curves was parallel. Specifically, there were relatively obvious spectral differences between the samples with different kinds of pesticide residues and the samples without residue in the range of 1493–2038 nm. Among them, the reflectance of Hami melon samples without residues in the range of 1493–1663 nm was the lowest, the reflectance of Hami melon samples sprayed with pesticides was higher than that of no residue samples, and the reflectance of Hami melon samples sprayed with Malathion was the highest. There were characteristic peaks and valleys in the specific spectral range, six peaks (1125, 1340, 1801, 2000, 2313, and 2370 nm) and six valleys (1020, 1269, 1552, 1929, 2432, and 2466 nm) appeared at similar wavelengths. These absorption wavelengths were related to the periodic tensile vibrations of C-H, O-H, and N-H bonds, which were the most fundamental chemical bonds of organic compounds [50]. The 1020 nm was in the low reflectivity range with high absorbance, which corresponded to the second overtone of C-H tensile vibration [51]. Wavelengths around 1125 and 1269 nm can be designated as the second overtone of C-H stretching in carbohydrates [52]. The 1552 nm and 1340 nm were mainly related to the chlorophyll absorption band of Hami melon, and 1552 nm was also related to N-H first-order frequency doubling [53]. The absorption peaks of water were approximately 1801 nm and 1929 nm, which was also the first overtone and combination mode of O-H group stretching [54]. While the strongest reflection trough (1935–1952 nm) can be attributed to the combination of O-H stretching and O-H bending of water [55]. The peaks and valleys at 2000 nm, 2313 nm, 2370 nm, 2432 nm, and 2466 nm were related to the frequency synthesis of C-H and O-H groups [56]. However, the appearance of these peaks and valleys cannot accurately and directly represent the presence of specific components in the sample, as the reflectivity at each location represents a complex set of component information. Especially, there is overlap in the spectral information of organic pesticides. In this study, the implied relationship between spectral data and pesticide residues was explored and analyzed.



**Figure 3.** Raw spectra and average reflectance spectra. (a) raw spectra; (b) average reflectance spectra.

### 3.2. Establishment and Analysis of Classical Machine Learning Classification Model

SVM, PLS-DA, and ELM were commonly used machine learning classification models, which were widely used in hyperspectral. SVM model is a supervised learning method based on statistical learning theory. After mapping the input data to high-dimensional space, linear SVM is applied to obtain nonlinear SVM, which was widely used in pattern recognition. The PLS-DA method is a discriminative classification strategy based on PLS, which can decompose the spectrum and the category matrix concomitantly, enhance the function of category information in spectral decomposition, and maximizes the extraction of the differences between different categories of spectra. ELM is a simple and reliable single hidden layer feedforward neural network learning algorithm, which has the advantages of fast training speed and high efficiency during learning, and high accuracy of generalization model. It is widely applied in prediction, classification, discrimination, and other issues. In this section, three classical machine learning classification models were compared and

analyzed, and the better model was selected. The experimental results are shown in the Table 1.

**Table 1.** Classification performance of three classical machine learning models for pesticide residues.

Models	Class	Accuracy (%)	Precision (%)	Sensitivity (%)	F1-Score
ELM	All	79.50	80.29	79.50	0.7962
	None	75.00	78.95	75.00	0.7692
	Acetamiprid	82.50	89.19	82.50	0.8571
	Malathion	82.50	73.33	82.50	0.7765
	Difenoconazole	82.50	71.74	82.50	0.7674
	Beta-cypermethrin	75.00	88.24	75.00	0.8108
SVM	All	77.50	78.03	77.50	0.7760
	None	75.00	78.95	75.00	0.7692
	Acetamiprid	77.50	86.11	77.50	0.8158
	Malathion	75.00	78.95	75.00	0.7692
	Difenoconazole	80.00	68.09	80.00	0.7356
	Beta-cypermethrin	80.00	78.05	80.00	0.7901
PLS-DA	All	75.50	76.42	75.50	0.7555
	None	75.00	76.92	75.00	0.7595
	Acetamiprid	67.50	87.10	67.50	0.7606
	Malathion	82.50	71.74	82.50	0.7674
	Difenoconazole	77.50	67.39	77.50	0.7209
	Beta-cypermethrin	75.00	78.95	75.00	0.7692

In the training of the model, RBF was selected as the kernel function of SVM, and the number of latent variables (LVs) with the smallest cross-validation error was selected to construct the PLS-DA model. The number of hidden layers of the ELM model was set to 135, and the kernel function was set as sig. The results showed that the three classical classification models can complete the classification task. Scholars in related fields also used machine learning model to complete classification tasks and obtain good results [23]. The accuracy of test sets of three classical classification models was 79.50%, 77.50%, and 75.50%, respectively.

Among these models, the performance of ELM classifier was significantly better than the other two. The precision value of ELM model was the highest, which was 80.29%. The precision values of SVM and PLS-DA were similar. The sensitivity of ELM model was 79.50%, which was 2.00% and 4.00% higher than that of SVM and PLS-DA models, respectively. There was sometimes a trade-off between precision and sensitivity. In order to achieve a balance between these two indicators, the F1-score indicator was selected, which was the harmonic average of the above two. Among them, the ELM model got the highest F1-score of 0.7962, the second was 0.7760 of the SVM model, and the lowest was 0.7555 of the PLS-DA model. The closer the F1-score value is to 1, the better the classification effect of the model is.

The performance test results of three different classical machine learning classification models showed that ELM had more advantages than SVM and PLS-DA, so ELM was selected to complete the subsequent model optimization.

### 3.3. Comparative Analysis of Spectral Preprocessing Treatments

Three types of preprocessing treatments (NM, MSC, and SNV) were used in this study to correct the original spectra. After preprocessing, the spectral data retained the absorption characteristics of 233 original spectra, and then ELM was used to model and analyze the spectral data after different preprocessing, which improved the classification ability of the model to a certain extent.

ELM classification models were established based on NM, MSC, SNV, and the intact spectrum without preprocessing. All three preprocessing treatments can enhance the classification accuracy of the model. Table 2 shows the model performance of the test set in terms of accuracy, precision, sensitivity, and F1-score. Compared with the untreated original spectrum, the ELM classification model developed based on the preprocessing spectral data

showed better overall discrimination performance. Proper spectral preprocessing improved the recognition accuracy of the model. The classification effects of the MSC-ELM and SNV-ELM were similar, and the difference between them was less than 1.00% under various evaluation criteria. In contrast, when the NM-ELM model was used to classify different pesticide residues on the surface of Hami melon, in single classification, the accuracy of Acetamiprid was 90.00%, the precision was 92.31%, the sensitivity was 90.00%, and the F1-score was 0.9114. Compared with the ELM classification model constructed after MSC and SNV preprocessing, the classification effect of the NM-ELM model was better, showed higher accuracy (82.00%), and its precision, sensitivity, and F1-score were 82.07%, 82.00%, and 0.8201, respectively. So, the NM-ELM model was selected for further optimization.

**Table 2.** Influence of different preprocessing treatments on ELM model performance.

Models	Class	Accuracy (%)	Precision (%)	Sensitivity (%)	F1-Score
NM-ELM	All	82.00	82.07	82.00	0.8201
	None	85.00	79.07	85.00	0.8193
	Acetamiprid	90.00	92.31	90.00	0.9114
	Malathion	80.00	82.05	80.00	0.8101
	Difenoconazole	80.00	80.00	80.00	0.8000
	Beta-cypermethrin	75.00	76.92	75.00	0.7595
MSC-ELM	All	80.50	80.87	80.50	0.8062
	None	82.50	84.62	82.50	0.8354
	Acetamiprid	85.00	91.89	85.00	0.8831
	Malathion	82.50	78.57	82.50	0.8049
	Difenoconazole	75.00	69.77	75.00	0.7229
	Beta-cypermethrin	77.50	79.49	77.50	0.7848
SNV-ELM	All	80.00	80.08	80.00	0.8001
	None	82.50	84.62	82.50	0.8354
	Acetamiprid	82.50	84.62	82.50	0.8354
	Malathion	75.00	76.92	75.00	0.7595
	Difenoconazole	82.50	76.74	82.50	0.7952
	Beta-cypermethrin	77.50	77.50	77.50	0.7750

### 3.4. Establishment and Analysis of tHBA-ELM Models

Since the parameters of ELM directly affect its classification accuracy, how to choose the optimal parameters is a problem that must be considered to improve the ELM classification model. The classification performance of ELM is mainly affected by the connection weight  $\omega$  between the input layer and the hidden layer and the neuron threshold  $b$  of the hidden layer. The metaheuristic optimization algorithm is widely used in model parameter optimization [39,41,57]. In order to overcome this problem, the optimization technology of HBA was used to solve the selection and optimization of  $\omega$  and  $b$ .

Two optimization strategies, GA and HBA, are used to optimize ELM. In order to compare the classification performance of the two optimization models, experiments were carried out under the same data set, and the results were shown in the Table 3. The number of hidden layers of GA-ELM and HBA-ELM models was set to 120, the population size was set to 50, and the number of iterations was set to 300. The initial input weight and hidden layer bias of the ELM model were optimized by GA and HBA algorithms. The discrimination accuracy of different pesticide residues in Hami melon in GA-ELM and HBA-ELM models were all superior to that in ELM model, which indicated that the classification accuracy of different pesticide residues on the surface of Hami melon in ELM model could be effectively improved by optimizing the parameters of ELM model through Metaheuristic algorithm. Other scholars have also used the combination of GA and machine learning model to achieve good results [58,59]. The accuracy of ELM classification model optimized by GA and HBA increased by 6.50% and 8.50%, precision by 6.65% and 8.59%, sensitivity by 6.50% and 8.50%, and F1-score by 0.0651 and 0.0850, respectively. It can be seen from the evaluation criteria of individual categories that the classification accuracy of HBA-ELM was higher than GA-ELM, and the F1-score of Beta-cypermethrin

increased the most, indicating that the HBA-ELM model was more stable and accurate than the GA-ELM model.

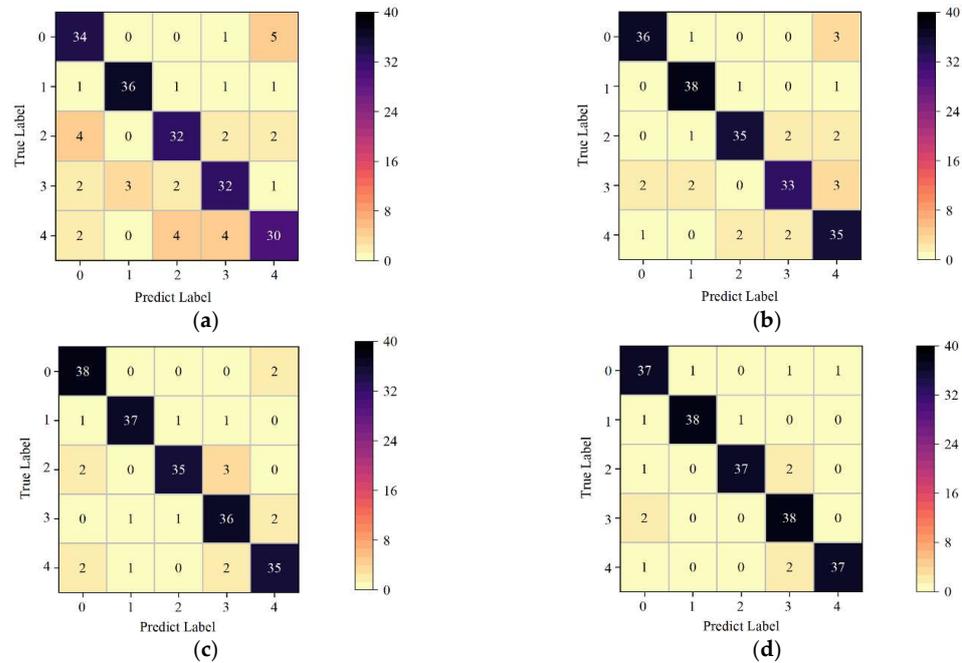
**Table 3.** Classification results of pesticide residues in Hami melon based on ELM model with different optimization algorithms.

Models	Class	Accuracy (%)	Precision (%)	Sensitivity (%)	F1-Score
NM-GA-ELM	All	88.50	88.72	88.50	0.8852
	None	90.00	92.31	90.00	0.9114
	Acetamiprid	95.00	90.48	95.00	0.9268
	Malathion	87.50	92.11	87.50	0.8974
	Difenoconazole	82.50	89.19	82.50	0.8571
	Beta-cypermethrin	87.50	79.55	87.50	0.8333
NM-HBA-ELM	All	90.50	90.66	90.50	0.9051
	None	95.00	88.37	95.00	0.9157
	Acetamiprid	92.50	94.87	92.50	0.9367
	Malathion	87.50	94.59	87.50	0.9091
	Difenoconazole	90.00	85.71	90.00	0.8780
	Beta-cypermethrin	87.50	89.74	87.50	0.8861
NM-tHBA-ELM	All	93.50	93.73	93.50	0.9355
	None	92.50	88.10	92.50	0.9024
	Acetamiprid	95.00	97.44	95.00	0.9620
	Malathion	92.50	97.37	92.50	0.9487
	Difenoconazole	95.00	88.37	95.00	0.9157
	Beta-cypermethrin	92.50	97.37	92.50	0.9487

For the sake of further improving the discrimination accuracy, the HBA-ELM model was optimized using adaptive t-distribution variation based on the above model. After many experiments, the coefficient of the mutation probability  $P$  was selected as 0.8. Compared with the HBA-ELM classification model, the evaluation criteria of the ELM model optimized by the tHBA were generally improved, with the overall accuracy of 93.50%, precision of 93.73%, sensitivity of 93.50%, and F1-score of 0.9355. Compared with the HBA-ELM, the discrimination accuracy of Acetamiprid and Difenoconazole in the individual categories reached 95.00%. The other three evaluation criteria of Acetamiprid were the highest among the five categories, precision was 97.44%, sensitivity was 95.00%, and F1-score was 0.9620.

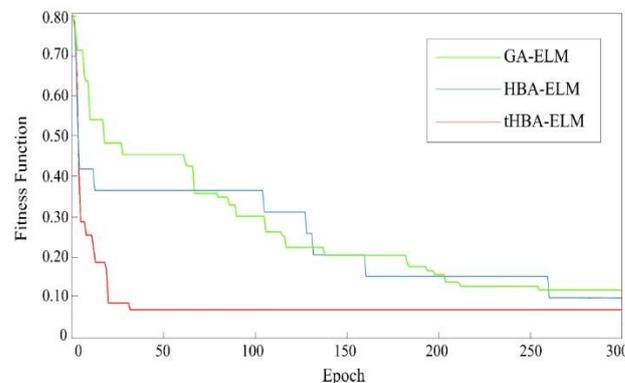
In order to further explore the detailed classification performance of each level, Figure 4 shows the confusion matrix of the classification results of the four classification models, i.e., NM-ELM, NM-GA-ELM, NM-HBA-ELM, and NM-tHBA-ELM, established by 200 different spectra of the test set. In the Figure 4, 0 in the horizontal and vertical coordinates represented none, 1 represented Acetamiprid, 2 represented Malathion, 3 represented Difenoconazole, and 4 represented Beta-cypermethrin. In terms of overall recognition ability, the NM-tHBA-ELM model only misjudged 13 samples, which was the least misjudged among the four models, as shown in Figure 4d. Moreover, the number of misjudged types was basically the same. In the NM-ELM model, as shown in Figure 4a, the misjudgment degree of Beta-cypermethrin in NM-ELM model was high, with 10 samples being misjudged (misjudgment rate: 20.00%). The misjudgment degree of Acetamiprid was the lowest in the five categories, and the misjudgment rate of the other four categories was 2.50%. In Figure 4b, the identification effect of Acetamiprid samples in the NM-GA-ELM model was better, and only two samples were misjudged, followed by the misjudgment rate of no residual samples of 3.00%, and most of them were misjudged as Beta-cypermethrin. In the NM-HBA-ELM model, as shown in Figure 4c, the misjudgment rate of Malathion and Beta-cypermethrin was 12.50%, but the overall discriminant ability was higher than that of the original ELM classification model. The test results showed that the classification accuracy of tHBA-ELM was higher than ELM, GA-ELM and HBA-ELM models, and it had good classification accuracy for different types of pesticide residues, which can be used for the identification of pesticide residues on Hami melon surface. In contrast, when SWIR was used to detect pesticide residues, the classification effect in this study was slightly lower

than that of garlic chive (*Allium tuberosum*) leaves. This may be because the proportion of pesticide solution in this study was low, and the types of pesticides were different. In addition, the samples used in this study were thick-skinned melons, which were quite different from the physicochemical properties of garlic chive (*Allium tuberosum*) leaves, which may also be the reason for the inconsistent results [51].



**Figure 4.** ELM and the improved ELM model confusion matrix. (a) NM-ELM; (b) NM-GA-ELM; (c) NM-HBA-ELM; (d) NM-tHBA-ELM.

The fitness curves of GA-ELM, HBA-ELM, and tHBA-ELM classification models are shown in Figure 5. It can be observed from Figure 5 that tHBA-ELM had the fastest convergence speed and was completely convergent when the number of iterations was 30 generations. However, GA-ELM and HBA-ELM were still not completely convergent when the number of iterations was 250 generations. In addition, the fitness value of tHBA-ELM classification model (the fitness value of this model was the error rate of the test set) was lower than 0.1, while the fitness value of GA-ELM and HBA-ELM classification models were both higher than 0.1. In view of the possibility that HBA-ELM may fall into local convergence, an adaptive t-distribution was added to the HBA-ELM model, so that it could jump out of the local optimum and converge as soon as possible. It can be seen from the Figure 5 that tHBA-ELM had a fast convergence speed and the lowest fitness value.



**Figure 5.** Optimization process of GA-ELM, HBA-ELM, and tHBA-ELM.

#### 4. Discussion

HSI is an effective method to realize the rapid and non-destructive detection of pesticide residues on Hami melon surface. In this study, SWIR-HSI was used to identify the pesticide residues on the surface of Hami melon. The high detection accuracy using the SWIR band may be due to the fact that the peaks and valleys of the spectral profile of this band are very distinct, or that the band has a higher reflectance for organic matter. At the same time, some scholars have used the band to detect pesticide residues in fruits and vegetables with good results [51].

When PLS-DA is used in multivariate modeling, if the dependent variables put into the model have little or no influence, the model structure will be unstable and even the discrimination results of the model will be poor. In addition, PLS-DA is a linear modeling method, SVM and ELM are both nonlinear modeling methods, and the spectral information of pesticide residue samples on the surface of Hami melon is nonlinear, which may also lead to poor classification effect of PLS-DA. In essence, both SVM and ELM map problems to a high-dimensional space, and then conduct classification research in the high-dimensional space. The classification effect is affected by the mapping method. If the mapping method is good, it will show obvious differences in high-dimensional space, and the classification effect will be good. If the mapping method is not shown in high-dimensional space, the classification effect will be poor. ELM is superior to SVM because it can be projected into high dimensions in infinite ways, and its training speed is very fast. We can adjust ELM many times and choose a better part from it. While with SVM, once the kernel function is determined, the mapping mode is uniquely determined, and its training speed is slower than ELM, so its performance is slightly worse than ELM. Similarly, some scholars have chosen spectroscopic techniques combined with ELM models for non-destructive detecting of the quality of agricultural products [60].

The acquisition of original hyperspectral data will be affected by the detection environment, operation level, instrument status, and other factors, so there will be more noise and interference in the original hyperspectral data. In order to minimize the impact of noise on data analysis and further explore and improve the classification performance of ELM, the original spectral data needs to be pretreated. The greatest improvement in model accuracy was achieved using the NM preprocessing method, and the model performance was also improved using MSC and SNV, and both improvements were similar. This may be because the purpose of MSC and SNV is basically the same, mainly to eliminate the scattering effect caused by uneven particle distribution and particle size. There is a linear correlation between MSC and SNV, which may lead to similar results of the two treatments. NM can unify the range of data, eliminate the influence of dimensions, and make the absolute value of numerical values become some kinds of relative value relationship, which is an effective way to simplify calculation and reduce numerical values. In similar studies, it can also be proved that NM has good spectral preprocessing effect [14]. This meant that the NM preprocessing treatment was suitable for the identification of pesticide residues in Hami melon.

In order to verify the classification effect of the tHBA-ELM model, different optimization strategies (GA, HBA) were used to optimize the ELM model and experiments were performed under the same dataset. After NM preprocessing, three ELM classification models based on different metaheuristic optimization algorithms have improved the classification effect and accuracy. Among them, the classification accuracy of NM-tHBA-ELM model was the highest, followed by the NM-HBA-ELM effect, and NM-GA-ELM effect was the worst. This may be because the introduction of adaptive t-distribution variation enhanced the diversity of population, which was beneficial to the algorithm to jump out of local optimization and improve the search speed.

Among single-class discrimination, Acetamiprid and Difenconazole have the highest discrimination accuracy. A possible reason for this is that these two pesticides are stable, non-volatile, and the reflection ability for the spectral is strong [7,61]. The classification accuracy of Beta-cypermethrin was basically lower than that of the other four categories

in the classification experiment of pesticide residues on the surface of Hami melon. A possible reason for this is that Beta-cypermethrin belongs to an agricultural pesticide, and its main component is cypermethrin, which is stable and not volatile in low temperature environments. However, when collecting hyperspectral images, the temperature in the dark box of HIS system will be slightly higher than the room temperature due to the energy excitation of halogen lamps. Therefore, when collecting Hami melon samples with Beta-cypermethrin on the surface, a small amount of surface pesticides may be heated and volatilized, resulting in poor classification effect compared with other pesticides.

Suitable preprocessing methods and meta-heuristic optimizers can effectively optimize the ELM model and thus improve the classification performance. A comparison of results from classification experiments on the identification of pesticide residue species on the surface of Hami melon based on different models reveals the superiority of NM-tHBA-ELM. Moreover, the fast convergence speed and good adaptation values further confirmed the effectiveness of our improvements on HBA. The results of the present study will help in the non-destructive detection of pesticide residues on the surfaces of other fruits and vegetables, thus effectively addressing the issue of food safety.

## 5. Conclusions

In this study, the tHBA-ELM model was successfully combined with SWIR-HSI in the spectral wavelength range of 1000–2500 nm, and a novel method for pesticide residue detection on Hami melon surface was proposed. The SWIR hyperspectral image was collected, and the spectral information of the sample was obtained with  $50 \times 50$  pixel blocks as the region of interest. The classification effect of classical machine learning classification models (ELM, SVM, and PLS-DA) on Hami melon surface pesticide residues was compared, and ELM was selected as the benchmark model. Different preprocessing treatments were adopted to measure their impact on the ELM classification model, and NM preprocessing treatment was proven to have the best performance. In order to improve the accuracy of NM-ELM model, the tHBA optimization algorithm was proposed, and it was proven that the tHBA optimization algorithm had faster convergence speed and better optimization effect than HBA and GA. The overall accuracy, precision, sensitivity, and F1-score of the NM-tHBA-ELM model was 93.50%, 93.73%, 93.50%, and 0.9355, which were higher than other classification models. Overall, the results showed that tHBA-ELM combined with SWIR hyperspectral imaging can be used as a non-destructive and efficient method for detecting pesticide residues on the surface of Hami melon. In addition, this method can provide a reference for the detection of pesticide residues in other fruits and vegetables. Furthermore, the types of pesticide residues on the surface of Hami melon and the varieties of Hami melon are not the same. Some parameters of the model should be fine-tuned with changes of pesticides and Hami melon varieties. Future work will optimize the generalizability of the proposed method by spraying more kinds of pesticides and adding different varieties of Hami melon samples, as well as exploring the use of deep learning and HSI technology to complete the non-destructive detection of pesticide residues on the surface of Hami melon [62,63].

**Author Contributions:** Conceptualization, Y.H. and Y.Z.; data curation, Y.H. and Y.L.; software, Y.H., Y.Z., and Y.L.; funding acquisition, B.M.; project administration, Y.H. and B.M.; methodology, Y.H., B.M., and H.W.; writing—original draft preparation, Y.H.; writing—review and editing, Y.H. and G.Y. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the National Natural Science Foundation of China, grant number 32060411.

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

**Acknowledgments:** The authors are also grateful for the experimental conditions provided by Shihezi University, Key Laboratory of Northwest Agricultural Equipment, Ministry of Agriculture and Rural

Affairs and Engineering Research Center for Production Mechanization of Oasis Characteristic Cash Crop, Ministry of Education.

**Conflicts of Interest:** The authors declare that the research was conducted without any commercial or financial relationships that could be construed as a potential conflict of interest.

## References

1. Wei, K.; Zhou, X.; Yan, Q.; Tao, Y.; Wang, C. Risk Assessment of Pesticide Residues in Muskmelon in Xinjiang. *Food Mach.* **2019**, *35*, 90–95. [[CrossRef](#)]
2. Dalcin, M.S.; Tschoeke, P.H.; de Souza Carlos Mourao, D.; Arguelles Osorio, P.R.; de Sousa Aguiar, R.W.; Cangussu, A.S.R.; dos Santos, G.R. Disease Progression and Productivity of Melon Plants with Different Cultural Management Techniques. *J. Plant Physiol. Pathol.* **2018**, *5*, 1–6. [[CrossRef](#)]
3. Food and Agriculture Organization of the United Nations. *International Code of Conduct on the Distribution and Use of Pesticides*; FAO: Rome, Italy, 2003.
4. Yazici, A.; Tiryaki, G.Y.; Ayvaz, H. Determination of Pesticide Residual Levels in Strawberry (*Fragaria*) by near-Infrared Spectroscopy. *J. Sci. Food Agric.* **2020**, *100*, 1980–1989. [[CrossRef](#)]
5. Tao, Y.; Dong, F.; Xu, J.; Liu, X.; Cheng, Y.; Liu, N.; Chen, Z.; Zheng, Y. Green and Sensitive Supercritical Fluid Chromatographic-Tandem Mass Spectrometric Method for the Separation and Determination of Flutriafol Enantiomers in Vegetables, Fruits, and Soil. *J. Agric. Food Chem.* **2014**, *62*, 11457–11464. [[CrossRef](#)] [[PubMed](#)]
6. Wang, X.; Tang, Q.; Wang, Q.; Qiao, X.; Xu, Z. Study of a Molecularly Imprinted Solid-Phase Extraction Coupled with High-Performance Liquid Chromatography for Simultaneous Determination of Trace Trichlorfon and Monocrotophos Residues in Vegetables. *J. Sci. Food Agric.* **2014**, *94*, 1409–1415. [[CrossRef](#)]
7. Garcia-Febrero, R.; Salvador, J.P.; Sanchez-Baeza, F.; Marco, M.P. Rapid Method Based on Immunoassay for Determination of Paraquat Residues in Wheat, Barley and Potato. *Food Control* **2014**, *41*, 193–201. [[CrossRef](#)]
8. Campillo, N.; Viñas, P.; Férrez-Melgarejo, G.; Hernández-Córdoba, M. Liquid Chromatography with Diode Array Detection and Tandem Mass Spectrometry for the Determination of Neonicotinoid Insecticides in Honey Samples Using Dispersive Liquid-Liquid Microextraction. *J. Agric. Food Chem.* **2013**, *61*, 4799–4805. [[CrossRef](#)] [[PubMed](#)]
9. Soltani Firouz, M.; Sardari, H. Defect Detection in Fruit and Vegetables by Using Machine Vision Systems and Image Processing. *Food Eng. Rev.* **2022**, *14*, 353–379. [[CrossRef](#)]
10. Sanaeifar, A.; Li, X.; He, Y.; Huang, Z.; Zhan, Z. A Data Fusion Approach on Confocal Raman Microspectroscopy and Electronic Nose for Quantitative Evaluation of Pesticide Residue in Tea. *Biosyst. Eng.* **2021**, *210*, 206–222. [[CrossRef](#)]
11. Zhao, K.; Zha, Z.; Li, H.; Wu, J. Early Detection of Moldy Apple Core Using Symmetrized Dot Pattern Images of Vibro-Acoustic Signals. *Trans. Chin. Soc. Agric. Eng.* **2021**, *37*, 290–298. [[CrossRef](#)]
12. Chen, S.; Zhao, Q.; Dong, D. Application of near Infrared Spectroscopy Combined with Comparative Principal Component Analysis for Pesticide Residue Detection in Fruit. *Spectrosc. Spectr. Anal.* **2020**, *40*, 917–921.
13. Xia, J.; Xu, W.; Li, Y.; Li, G.; Wang, X.; Xiong, Y.; Min, S. Quality Control of Pesticide Using Infrared Spectroscopic Coupled with Fingerprint Analysis. *Infrared Phys. Technol.* **2022**, *122*, 2–8. [[CrossRef](#)]
14. Yu, G.; Ma, B.; Chen, J.; Li, X.; Li, Y.; Li, C. Nondestructive Identification of Pesticide Residues on the Hami Melon Surface Using Deep Feature Fusion by Vis/NIR Spectroscopy and 1D-CNN. *J. Food Process Eng.* **2021**, *44*, e13602. [[CrossRef](#)]
15. Ma, S.; Ji, R.; Wang, X.; Yu, C.; Yu, Y.; Yang, X. Fluorescence Detection of Boscalid Pesticide Residues in Grape Juice. *Optik* **2019**, *180*, 236–239. [[CrossRef](#)]
16. Hu, J.; Zou, Y.; Sun, B.; Yu, X.; Shang, Z.; Huang, J.; Jin, S.; Liang, P. Raman Spectrum Classification Based on Transfer Learning by a Convolutional Neural Network: Application to Pesticide Detection. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **2022**, *265*, 120366. [[CrossRef](#)] [[PubMed](#)]
17. Ma, Q.; Teng, Y.; Li, C.; Jiang, L. Simultaneous Quantitative Determination of Low-Concentration Ternary Pesticide Mixtures in Wheat Flour Based on Terahertz Spectroscopy and BPNN. *Food Chem.* **2022**, *377*, 132030. [[CrossRef](#)]
18. López-Maestresalas, A.; Keresztes, J.C.; Goodarzi, M.; Arazuri, S.; Jarén, C.; Saeys, W. Non-Destructive Detection of Blackspot in Potatoes by Vis-NIR and SWIR Hyperspectral Imaging. *Food Control* **2016**, *70*, 229–241. [[CrossRef](#)]
19. Sellami, A.; Tabbone, S. Deep Neural Networks-Based Relevant Latent Representation Learning for Hyperspectral Image Classification. *Pattern Recognit.* **2022**, *121*, 108224. [[CrossRef](#)]
20. Wang, Z.; Fan, S.; Wu, J.; Zhang, C.; Xu, F.; Yang, X.; Li, J. Application of Long-Wave near Infrared Hyperspectral Imaging for Determination of Moisture Content of Single Maize Seed. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **2021**, *254*, 119666. [[CrossRef](#)]
21. Li, Y.; Ma, B.; Li, C.; Yu, G. Accurate Prediction of Soluble Solid Content in Dried Hami Jujube Using SWIR Hyperspectral Imaging with Comparative Analysis of Models. *Comput. Electron. Agric.* **2022**, *193*, 106655. [[CrossRef](#)]
22. Song, H.; Yang, W.; Yuan, H.; Bufford, H. Deep 3d-Multiscale Densenet for Hyperspectral Image Classification Based on Spatial-Spectral Information. *Intell. Autom. Soft Comput.* **2020**, *26*, 1441–1458. [[CrossRef](#)]

23. Alimohammadi, F.; Rasekh, M.; Sayyah, A.H.A.; Abbaspour-Gilandeh, Y.; Karami, H.; Sharabiani, V.R.; Fioravanti, A.; Gancarz, M.; Findura, P.; Kwaśniewski, D. Hyperspectral Imaging Coupled with Multivariate Analysis and Artificial Intelligence to the Classification of Maize Kernels. *Int. Agrophys.* **2022**, *36*, 83–91. [[CrossRef](#)]
24. Zhang, H.; Zhang, S.; Dong, W.; Luo, W.; Huang, Y.; Zhan, B.; Liu, X. Detection of Common Defects on Mandarins by Using Visible and near Infrared Hyperspectral Imaging. *Infrared Phys. Technol.* **2020**, *108*, 103341. [[CrossRef](#)]
25. Liu, Z.; Jiang, J.; Du, Y.; Xu, Z. A Band Influence Algorithm for Hyperspectral Band Selection to Classify Moldy Peanuts. *IEEE Access* **2021**, *9*, 147527–147536. [[CrossRef](#)]
26. Using, A.; Hyperspectral, P.N.I.R.; Ekramirad, N.; Khaled, A.Y.; Doyle, L.E.; Loeb, J.R.; Donohue, K.D.; Villanueva, R.T.; Adedeji, A.A. Nondestructive Detection of Codling Moth Infestation in Machine Learning and Feature Selection. *Foods* **2022**, *11*, 8.
27. Lee, H.; Kim, M.S.; Lim, H.S.; Park, E.; Lee, W.H.; Cho, B.K. Detection of Cucumber Green Mottle Mosaic Virus-Infected Watermelon Seeds Using a near-Infrared (NIR) Hyperspectral Imaging System: Application to Seeds of the “Sambok Honey” Cultivar. *Biosyst. Eng.* **2016**, *148*, 138–147. [[CrossRef](#)]
28. Jun, S.; Shuying, J.; Meixia, Z.; Hanping, M.; Xiaohong, W.; Qinglin, L. Detection of Pesticide Residues in Mulberry Leaves Using Vis-Nir Hyperspectral Imaging Technology. *J. Residuals Sci. Technol.* **2016**, *13*, S125–S131. [[CrossRef](#)]
29. Sun, J.; Cong, S.; Mao, H.; Wu, X.; Yang, N. Quantitative Detection of Mixed Pesticide Residue of Lettuce Leaves Based on Hyperspectral Technique. *J. Food Process Eng.* **2018**, *41*, e12654. [[CrossRef](#)]
30. Su, W.H.; Sun, D.W.; He, J.G.; Zhang, L.B. Variation Analysis in Spectral Indices of Volatile Chlorpyrifos and Non-Volatile Imidacloprid in Jujube (*Ziziphus jujuba* Mill.) Using near-Infrared Hyperspectral Imaging (NIR-HSI) and Gas Chromatograph-Mass Spectrometry (GC-MS). *Comput. Electron. Agric.* **2017**, *139*, 41–55. [[CrossRef](#)]
31. Nturambirwe, J.F.I.; Perold, W.J.; Opara, U.L. Classification Learning of Latent Bruise Damage to Apples Using Shortwave Infrared Hyperspectral Imaging. *Sensors* **2021**, *21*, 4990. [[CrossRef](#)]
32. Qi, X.; Jiang, J.; Cui, X.; Yuan, D. Moldy Peanut Kernel Identification Using Wavelet Spectral Features Extracted from Hyperspectral Images. *Food Anal. Methods* **2020**, *13*, 445–456. [[CrossRef](#)]
33. Xu, Y.; Zhang, H.; Zhang, C.; Wu, P.; Li, J.; Xia, Y.; Fan, S. Rapid Prediction and Visualization of Moisture Content in Single Cucumber (*Cucumis sativus* L.) Seed Using Hyperspectral Imaging Technology. *Infrared Phys. Technol.* **2019**, *102*, 103034. [[CrossRef](#)]
34. Li, B.; Cobo-Medina, M.; Lecourt, J.; Harrison, N.B.; Harrison, R.J.; Cross, J.V. Application of Hyperspectral Imaging for Nondestructive Measurement of Plum Quality Attributes. *Postharvest Biol. Technol.* **2018**, *141*, 8–15. [[CrossRef](#)]
35. Cen, H.; He, Y. Theory and Application of near Infrared Reflectance Spectroscopy in Determination of Food Quality. *Trends Food Sci. Technol.* **2007**, *18*, 72–83. [[CrossRef](#)]
36. Malvandi, A.; Kapoor, R.; Feng, H.; Kamruzzaman, M. Non-Destructive Measurement and Real-Time Monitoring of Apple Hardness during Ultrasonic Contact Drying via Portable NIR Spectroscopy and Machine Learning. *Infrared Phys. Technol.* **2022**, *122*, 104077. [[CrossRef](#)]
37. Yang, Y.; Zhao, C.; Huang, W.; Tian, X.; Fan, S.; Wang, Q.; Li, J. Optimization and Compensation of Models on Tomato Soluble Solids Content Assessment with Online Vis/NIRS Diffuse Transmission System. *Infrared Phys. Technol.* **2022**, *121*, 104050. [[CrossRef](#)]
38. Wang, W.; Ma, B.; Luo, X.; Li, X.; Lei, S.; Li, Y.; Sun, J. Study on the Moisture Content of Dried Hami Big Jujubes by Near-Infrared Spectroscopy Combined with Variable Preferred and GA-ELM Model. *Spectrosc. Spectr. Anal.* **2020**, *40*, 543–549.
39. Cong, S.; Liu, C.; Zhu, Z.; Hu, A. Study on Identification of Multiple Pesticide Residues in Lettuce Leaves Based on Hyperspectral Technology. *Commun. Comput. Inf. Sci.* **2021**, *1424*, 537–550. [[CrossRef](#)]
40. Pu, H.; Liu, D.; Wang, L.; Sun, D.W. Soluble Solids Content and PH Prediction and Maturity Discrimination of Lychee Fruits Using Visible and near Infrared Hyperspectral Imaging. *Food Anal. Methods* **2016**, *9*, 235–244. [[CrossRef](#)]
41. Sunli, C.; Jun, S.; Hanping, M.; Xiaohong, W.; Pei, W.; Xiaodong, Z. Non-Destructive Detection for Mold Colonies in Rice Based on Hyperspectra and GWO-SVR. *J. Sci. Food Agric.* **2018**, *98*, 1453–1459. [[CrossRef](#)]
42. Bonah, E.; Huang, X.; Yi, R.; Aheto, J.H.; Yu, S. Vis-NIR Hyperspectral Imaging for the Classification of Bacterial Foodborne Pathogens Based on Pixel-Wise Analysis and a Novel CARS-PSO-SVM Model. *Infrared Phys. Technol.* **2020**, *105*, 103220. [[CrossRef](#)]
43. Sun, J.; Ge, X.; Wu, X.; Dai, C.; Yang, N. Identification of Pesticide Residues in Lettuce Leaves Based on near Infrared Transmission Spectroscopy. *J. Food Process Eng.* **2018**, *41*, e12816. [[CrossRef](#)]
44. Xu, L.; Zhou, Y.P.; Tang, L.J.; Wu, H.L.; Jiang, J.H.; Shen, G.L.; Yu, R.Q. Ensemble Preprocessing of Near-Infrared (NIR) Spectra for Multivariate Calibration. *Anal. Chim. Acta* **2008**, *616*, 138–143. [[CrossRef](#)] [[PubMed](#)]
45. Isaksson, T.; Naes, T. Effect of Multiplicative Scatter Correction (MSC) and Linearity Improvement in NIR Spectroscopy. *Appl. Spectrosc.* **1988**, *42*, 1273–1284. [[CrossRef](#)]
46. Stordrange, L.; Libnau, F.O.; Malthe-Sørensen, D.; Kvalheim, O.M. Feasibility Study of NIR for Surveillance of a Pharmaceutical Process, Including a Study of Different Preprocessing Techniques. *J. Chemom.* **2002**, *16*, 529–541. [[CrossRef](#)]
47. Rinnan, Å.; van den Berg, F.; Engelsen, S.B. Review of the Most Common Pre-Processing Techniques for near-Infrared Spectra. *Trends Anal. Chem.* **2009**, *28*, 1201–1222. [[CrossRef](#)]
48. Hashim, F.A.; Houssein, E.H.; Hussain, K.; Mabrouk, M.S.; Al-Atabany, W. Honey Badger Algorithm: New Metaheuristic Algorithm for Solving Optimization Problems. *Math. Comput. Simul.* **2022**, *192*, 84–110. [[CrossRef](#)]
49. Yin, S.; Luo, Q.; Du, Y.; Zhou, Y. DTSM: Dominant Swarm with Adaptive T-Distribution Mutation-Based Slime Mould Algorithm. *Math. Biosci. Eng.* **2022**, *19*, 2240–2285. [[CrossRef](#)]

50. Li, J.; Tian, X.; Huang, W.; Zhang, B.; Fan, S. Application of Long-Wave near Infrared Hyperspectral Imaging for Measurement of Soluble Solid Content (SSC) in Pear. *Food Anal. Methods* **2016**, *9*, 3087–3098. [[CrossRef](#)]
51. He, W.; He, H.; Wang, F.; Wang, S.; Lyu, R. Non-Destructive Detection and Recognition of Pesticide Residues on Garlic Chive (*Allium tuberosum*) Leaves Based on Short Wave Infrared Hyperspectral Imaging and One-Dimensional Convolutional Neural Network. *J. Food Meas. Charact.* **2021**, *15*, 4497–4507. [[CrossRef](#)]
52. Dong, J.; Guo, W. Nondestructive Determination of Apple Internal Qualities Using Near-Infrared Hyperspectral Reflectance Imaging. *Food Anal. Methods* **2015**, *8*, 2635–2646. [[CrossRef](#)]
53. Sun, J.; Zhou, X.; Mao, H.P.; Wu, X.H.; Zhang, X.D.; Gao, H.Y. Identification of Pesticide Residue Level in Lettuce Based on Hyperspectra and Chlorophyll Fluorescence Spectra. *Int. J. Agric. Biol. Eng.* **2016**, *9*, 231–239. [[CrossRef](#)]
54. Siedliska, A.; Baranowski, P.; Zubik, M.; Mazurek, W.; Sosnowska, B. Detection of Fungal Infections in Strawberry Fruit by VNIR/SWIR Hyperspectral Imaging. *Postharvest Biol. Technol.* **2018**, *139*, 115–126. [[CrossRef](#)]
55. Manley, M.; du Toit, G.; Geladi, P. Tracking Diffusion of Conditioning Water in Single Wheat Kernels of Different Hardnesses by near Infrared Hyperspectral Imaging. *Anal. Chim. Acta* **2011**, *686*, 64–75. [[CrossRef](#)] [[PubMed](#)]
56. Sun, J.; Zhou, X.; Mao, H.; Wu, X.; Yang, N.; Zhang, X. Detection of Pesticide Residues on Lettuce Leaves Based on Piece-Wise Discrete Wavelet Transform and Hyperspectral Data. *Nongye Jixie Xuebao/Trans. Chin. Soc. Agric. Mach.* **2016**, *47*, 323–329. [[CrossRef](#)]
57. Li, C.; Ge, X.; Jin, Y.; Sun, J. Detection of Pesticide Residues in Vegetables Based on Near-Infrared Spectroscopy. *Agric. Eng.* **2019**, *9*, 33–39. [[CrossRef](#)]
58. Song, D.; Gao, D.; Sun, H.; Qiao, L.; Zhao, R.; Tang, W.; Li, M. Chlorophyll Content Estimation Based on Cascade Spectral Optimizations of Interval and Wavelength Characteristics. *Comput. Electron. Agric.* **2021**, *189*, 106413. [[CrossRef](#)]
59. Nagasubramanian, K.; Jones, S.; Sarkar, S.; Singh, A.K.; Singh, A.; Ganapathysubramanian, B. Hyperspectral Band Selection Using Genetic Algorithm and Support Vector Machines for Early Identification of Charcoal Rot Disease in Soybean Stems. *Plant Methods* **2018**, *14*, 86. [[CrossRef](#)]
60. Qiao, Z.; Zhan, C. Prediction of Sugar Content in Apple Based on Near-Infrared Spectroscopy of SSA-ELM. *Food Mach.* **2021**, *37*, 121–126. [[CrossRef](#)]
61. Zhou, Q.; Liu, Y.; Kou, K.; Zhao, G. Study on Spectral Properties of Difenoconazole by Terahertz Time-Domain Spectroscopy. In *2015 International Conference on Optical Instruments and Technology: Terahertz Technologies and Applications*; SPIE: Bellingham, WA, USA, 2015; Volume 9625, p. 962504. [[CrossRef](#)]
62. Lei, Z.; Zeng, Y.; Liu, P.; Su, X. Active Deep Learning for Hyperspectral Image Classification with Uncertainty Learning. *IEEE Geosci. Remote Sens. Lett.* **2022**, *19*, 1–5. [[CrossRef](#)]
63. Yuan, Y.; Wang, C.; Jiang, Z. Proxy-Based Deep Learning Framework for Spectral-Spatial Hyperspectral Image Classification: Efficient and Robust. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 1–15. [[CrossRef](#)]

**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.