

Article

Comparison and Analysis of Several Quantitative Identification Models of Pesticide Residues Based on Quick Detection Paperboard

Yao Zhang ^{1,2}, Qifu Zheng ^{2,*}, Xiaobin Chen ^{2,*} , Yingyi Guan ², Jingbo Dai ^{1,2}, Min Zhang ², Yunyuan Dong ² and Haodong Tang ¹

¹ College of Chemical Engineering, Zhejiang University of Technology, Hangzhou 310014, China

² College of Chemical and Material Engineering, Quzhou University, Quzhou 324000, China; yunyuandong@qzc.edu.cn (Y.D.)

* Correspondence: zqf@qzc.edu.cn (Q.Z.); xiaobinchen@qzc.edu.cn (X.C.)

Abstract: Pesticide residues have long been a significant aspect of food safety, which has always been a major social concern. This study presents research and analysis on the identification of pesticide residue fast detection cards based on the enzyme inhibition approach. In this study, image recognition technology is used to extract the color information RGB eigenvalues from the detection results of the quick detection card, and four regression models are established to quantitatively predict the pesticide residue concentration indicated by the quick detection card using RGB eigenvalues. The four regression models are linear regression model, quadratic polynomial regression model, exponential regression model and RBF neural network model. Through study and comparison, it has been shown that the exponential regression model is superior at predicting the pesticide residue concentration indicated by the rapid detection card. The correlation value is 0.900, and the root mean square error is 0.106. There will be no negative prediction value when the expected concentration is near to 0. This gives a novel concept and data support for the development of image recognition equipment for pesticide residue fast detection cards based on the enzyme inhibition approach.

Keywords: pesticide residue; image processing; RGB color model; prediction model; data averaging



Citation: Zhang, Y.; Zheng, Q.; Chen, X.; Guan, Y.; Dai, J.; Zhang, M.; Dong, Y.; Tang, H. Comparison and Analysis of Several Quantitative Identification Models of Pesticide Residues Based on Quick Detection Paperboard. *Processes* **2023**, *11*, 1854. <https://doi.org/10.3390/pr11061854>

Academic Editor: Jesús M. Marín-Benito

Received: 11 May 2023

Revised: 8 June 2023

Accepted: 17 June 2023

Published: 20 June 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

With the evolution of civilization, people's living standards have steadily risen, and food safety has become a growing everyday worry. Specifically, pesticide residues are extremely detrimental to human health. Utilizing pesticides to eliminate pests is crucial to ensuring that crop production meets expectations. However, excessive pesticide application will result in agricultural goods with high pesticide residues, that once on the market, will result in irreparable losses and compromise the security of human life and private property [1–5]. So it is very necessary to study the rapid detection technology of pesticide residues, which can help certain situations that require large-scale and rapid detection of pesticide residues, such as farms, orchards and supermarkets. Reduce testing costs and improve testing efficiency. Therefore, this article conducts in-depth research on the rapid identification technology of pesticide residue rapid detection cards, using machine vision and mathematical modeling to quickly and accurately identify pesticide residue rapid detection cards, improving detection efficiency and accuracy.

China is a huge agricultural nation; thus, pesticides are essential. However, the misuse of pesticides is also widespread. At present, the methods used to measure pesticide residues in vegetables and fruits include enzyme inhibition method [6–8], biosensor method [9–12], chromatographic detection method [13–17], etc.

Enzyme inhibition is the most commonly used rapid detection technique for pesticide residues. It is a technique that makes use of the phenomenon in which the functional group

of an enzyme is altered by a substance, causing the enzyme's activity to decrease or lose its effect. Its benefits include simple operation, quick speed, low cost, easy generalization, and widespread use, but its precision and reproducibility are marginally subpar.

It comprises of two components, a biosensor and a converter, and has a selective and reversible reaction to particular chemicals or bio-actives. Pesticide residues can be calculated based on signal changes such as pH and conductivity in biologically vulnerable areas. It has the benefits of rapid measurement, good selectivity, high anti interference ability, and high sensitivity, but it also has the disadvantages of a short service life and test results that are easily influenced by physical and chemical environmental factors, resulting in the deviation of final results.

Chromatographic detection is subdivided into gas chromatography and liquid chromatography. The basis of chromatographic detection is to utilize chromatograms to assess pesticide residues; however, the carrier employed varies between gas and liquid. Gas chromatography and liquid chromatography are the most commonly used techniques in non-rapid detection methods, which can detect pesticide content with very low detection limits and maintain high accuracy but require professional technology and expensive instruments. Gas chromatography is a gas-based method for detecting substances. It has advantages such as high sensitivity, low difficulty, good selectivity, etc., but it can only be utilized when the measured pesticides are easy to crack and volatilize. Consequently, the varieties measured have restrictions. After the development of gas chromatography, liquid chromatography emerged. Its carrier is liquid, which is easier to operate than gas, resulting in a less costly and more versatile detecting system. When there are many different types of pesticide residues, however, it is difficult to detect the various pesticide components and amounts with traditional liquid chromatography in a timely and precise manner.

The rapid test card is a technique of detection based on the enzyme inhibition method. Compared to other detection technologies, it is easy to use and does not need expert training or specialized equipment. Therefore, regular families may utilize it without difficulty. Although the pesticide detection card can be used to quickly determine the concentration of pesticide residues, there are many uncontrollable factors in the process of identifying the concentration indicated by the color of the detection card through human eyes, which affects the accuracy of the detection. Currently, mathematical modeling theory has reached its full maturity, and a number of academics have conducted research on the quantitative identification algorithm of pesticide residues based on detection cards. For example, Chen Zipeng, Ma Jianbo and others [18] extracted image information from the detection results of the pesticide residue quick detection card using image recognition technology. Through research, it was determined that the color saturation parameter in the test results of the quick test card has an approximately linear relationship with the pesticide residue degree. Consequently, a model was developed to determine the pesticide residue degree based on color saturation. Ranveer Soniya A, Harshitha CG and others [19] created a paper strip-based sensor for detecting pesticide residues in various dairy environmental samples, and the paper strip biosensor is capable of detecting pesticides at extremely low concentrations. Jia Binghui, Zhang Peiqiang and others [20] AA retrieved the gray value from the gray picture on the fast test card and qualitatively determined if the pesticide residue concentration exceeded the standard by establishing a gray value threshold. Malay Kishore Dutta, Namita Sengar and others [21] employed image processing technology to distinguish between pesticide-treated and untreated grapes based on the picture characteristics and a support vector machine (SVM) classifier to determine the pesticide concentration in the sample. Yan Lin, Lu Anxiang and others [22] analyzed the correlation between the RGB parameters of the detection card image and the concentration change and determined that the R parameter has the highest correlation with the concentration change; consequently, they constructed a linear regression model with the R parameter to predict pesticide residue concentration. The advantages and disadvantages of each identification method have not been compared and studied horizontally by the researchers.

This paper obtained a large number of RGB values from the detection card and the corresponding pesticide residue data through experiments in order to build identification models of pesticide residues, including a linear regression prediction model, an exponential regression prediction model, a quadratic polynomial regression prediction model, and an RBF neural network prediction model, and analyze the effect of each model on predicting pesticide residue concentrations. Moreover, taking into account the variances in RGB data at the same concentration gathered throughout the experiment. Errors caused by the quality of the detection card, test solutions that were unevenly smeared, equipment, environment, and human factors, which will decrease the data quality and the model fit of the prediction model, have been identified as the root cause of the poor quality of the data and the poor model fit. After eliminating outliers during data preparation, this work, therefore, presents the notion of averaging to handle the training set data. After averaging, the RGB value of the detection card color at the same concentration tends towards its average value, improving the data. The rapid detection technology of pesticide residues has a huge market demand, so research in this area is very meaningful. In the future, pesticide residue detection methods will develop in a convenient, rapid, and low-cost manner. People can detect pesticide residue concentrations with less cost and time, greatly reducing the risk of accidental consumption of pesticides when consuming vegetables and fruits.

2. Methodology

2.1. Research Ideas

This research compares and contrasts the advantages and drawbacks of four models for estimating the quantity of pesticide residue on a rapid test card and analyzes the effect of data averaging on modeling outcomes.

The research process is as follows: (As shown in Figure 1). First, create a collection of pesticide residue detection card samples, convert them to electronic images via photography, and input them into the computer. Second, improve the image using image processing technologies and extract RGB parameters from the optimized image. Perform data analysis and data processing for RGB parameters as the third step. Finally, the model was developed and its performance was assessed.

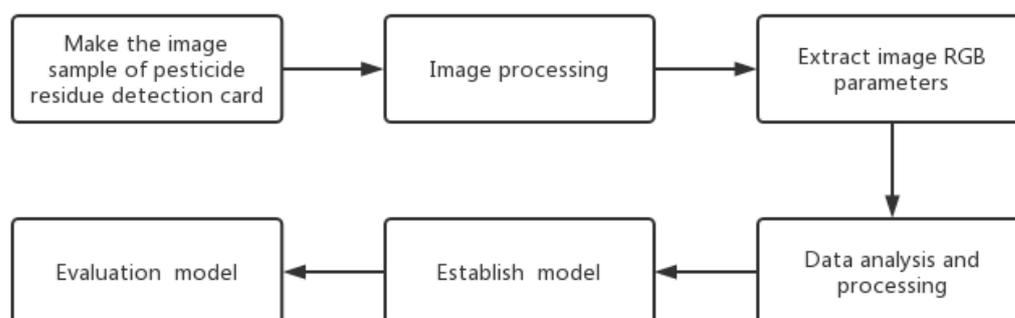


Figure 1. Research roadmap.

2.2. Image Sample Acquisition

2.2.1. Stores Reserve

Detection card: Chip of paper with cholinesterase and indophenol acetate reagent.

Detection solution: Using chloropyrhion as a pesticide simulant, 10 distinct sets of pesticide residual solutions were prepared and their concentrations were $0 \text{ mg}\cdot\text{mL}^{-1}$, $0.0002 \text{ mg}\cdot\text{mL}^{-1}$, $0.01 \text{ mg}\cdot\text{mL}^{-1}$, $0.03 \text{ mg}\cdot\text{mL}^{-1}$, $0.05 \text{ mg}\cdot\text{mL}^{-1}$, $0.1 \text{ mg}\cdot\text{mL}^{-1}$, $0.3 \text{ mg}\cdot\text{mL}^{-1}$, $0.5 \text{ mg}\cdot\text{mL}^{-1}$, $0.7 \text{ mg}\cdot\text{mL}^{-1}$, $1 \text{ mg}\cdot\text{mL}^{-1}$, 10 samples per group.

2.2.2. Detection Card Assay

Using a detection card, add two or three drops of the test solution and set it in an incubator at 37 °C for 10 min to initiate a pre-reaction. During the pre-reaction period, the detection card's surface remained damp. After the pre-reaction, half-fold the detection card and maintain 37 °C at a steady temperature for three minutes to create the detection card-reacting solution.

2.2.3. Image Acquisition

Place 100 gathered detection cards on white paper, photograph them vertically with a cell phone under indoor natural light, and save the images.

2.3. Image Processing

2.3.1. Image Preprocessing

Python's OpenCV tool is responsible for image preparation. Initially, the resolution of the captured pictures is scaled to a consistent 256×256 pixels, which facilitates uniform processing in the next steps. During the process of picture acquisition, environmental or human variables might cause distortions and noise spots to appear in the obtained images; thus, preprocessing operations must be performed on the images to reduce the impacts of distortions and noise spots, hence enhancing image quality. The photos in this study are preprocessed using median filtering.

2.3.2. Image Cropping

The obtained photos contained background portions outside the chromogenic regions of the detection card; hence, photographs must be cropped to eliminate the background and retain only the chromogenic regions required for this experiment [23–29]. Image cropping was performed with the Python OpenCV tool.

The cutting method is as follows:

Step one: Convert the original RGB color model picture to the HSV color model image, defining thresholds by the hue and brightness of HSV color models. The HSV of the lower threshold is set to (0, 43, 46), and the higher HSV threshold is set to (180, 255, 255). After filtering the background information, obtain a preliminary mask image [30]. The masked picture is a binary image in which the only grayscale values are 0 and 255. The resultant mask picture has a central color rendering region with a gray value of 255 and a background area with a gray value of 0.

Step two: A little amount of background information remains unfiltered in the mask picture. Image etching is used to remove the remaining background information. To obtain the optimal mask picture, the kernel size of the etch operation is set to 3·3 and the number of iterations is set to 5.

Step three: do the "and" operation on the mask image and the original picture ($1 \& 1 = 1$, $1 \& 0 = 0$, $0 \& 1 = 0$, $0 \& 0 = 0$). Two pictures with corresponding pixel positions accomplish the "and" action. Following the "and" operation, the cropped image is obtained.

2.4. RGB Eigenvalue Extraction

Assign as the desired value, the RGB value with the greatest frequency in the color rendering section of the detection card. Using the Python-OpenCV tool, summarize and organize the cropped photos and retrieve histogram array information on RGB characteristic values. Histograms provide access to RGB values.

2.5. Eliminate Outliers

A preliminary study of RGB data indicates that the RGB values of certain samples (10 samples at each dosage) with the same quantity of pesticide residue vary somewhat. It is regarded as an error produced by sampling errors, hence it is important to preprocess the data in order to limit its influence on modeling.

In this paper, the Mahalanobis distance test [31–33] is adopted to eliminate outliers with large differences in the original data. The Mahalanobis distance is a more accurate tool for identifying multivariate outliers. It is independent of the unit selection of each component when compared to Euclidean distance since it does not take into consideration the dimensions of each characteristic parameter. Second, it considers the correlation between variables to more correctly recognize outliers, and it may also be used on data that does not exactly correspond to a normal distribution. Because the Mahalanobis distance roughly follows the Chi-square distribution, the Chi-square test is performed after each sample data's Mahalanobis distance is computed, and any value larger than the critical value is considered an outlier.

2.6. Average Processing

The original data are acquired after image processing and the extraction of RGB feature values. The Mahalanobis distance and the Chi-square test were employed to identify outliers with considerable differences. However, there are still legitimate variances, and this divergence cannot be removed. It is caused by the combined impact of the detection card's quality, unequal application of the liquid to be tested, the environment at the moment of taking the photo, equipment, and human factors. It causes the RGB values of samples with the same concentration (10 samples for each concentration) to deviate to varying degrees. The approach of averaging is employed to remove the effect of this difference on the creation of a prediction model.

Average processing is a way of correcting data that, in principle, should be the same but are impacted by environmental, equipment, human, and other variables, resulting in varying degrees of variation. Such deviations are not outliers; they are within the permitted range and take into account the actual conditions of the experiment.

The objective behind averaging is to average data that should be the same in theory, and the processing object is the data set with outliers eliminated. The data farther away from the mean moved faster, whereas the data closest to the mean moved slower. Formula 1 shows the formula for calculating movement distance (1). The value of penalty factor C can be used to influence the total movement specification.

Furthermore, the formula was verified through practical application in the paper. In the model validation section, it can be found that the model established using the averaging method is superior to the untreated model, and the point is that the data used for validation is raw data without any processing.

The training set data were averaged as follows:

$$Y = X - C \cdot |X - X_{\text{mean}}| \cdot (X - X_{\text{mean}}) / (X_{\text{max}} - X_{\text{min}}) \quad (1)$$

Y denotes the averaging result; X denotes the RGB value of each sample under the same pesticide residue concentration; X_{mean} denotes the mean of all sample RGB values under the same pesticide residue concentration; X_{max} denotes the maximum RGB value of all samples under the same pesticide residue concentration and X_{min} denotes the minimum RGB value of all samples under the same pesticide residue concentration. C is the penalty factor that governs the data movement specification.

Input the data into the data processing algorithm, then enter the penalty factor. C is set to 1, which ensures that each RGB data value is near its mean value.

2.7. Modeling and Evaluation

2.7.1. Linear Regression Model

A linear regression model describes the linear connection between independent and dependent variables. The goal of predicting dependent variables through independent factors is achieved by fitting the data of dependent variables and independent variables with a linear regression equation [34–37]. For model estimates, the modeling tool utilized SPSS 26.0 statistical analysis software. For model estimation, the RGB value (independent

variable) and matching pesticide residue concentration (dependent variable) were entered into the SPSS 26.0 program. The regression equation of the linear regression model:

$$Y = X_1 + X_2 \cdot R + X_3 \cdot G + X_4 \cdot B \quad (2)$$

In the equation, Y represents the estimated concentration of pesticide residue ($\text{mg} \cdot \text{mL}^{-1}$) and X_1 – X_4 represents the parameters to be determined. R, G, and B represent the color values R, G, and B on the color test card.

2.7.2. Nonlinear Regression Model

The concept of a nonlinear regression model resembles the approach of linear regression. The regression equation of a nonlinear regression model is frequently a complicated nonlinear function, and the regression law is visually represented by different-shaped curves; such models are referred to as nonlinear regression models [38–40]. For model estimates, the modeling tool utilized SPSS 26.0 statistical analysis software. The quadratic polynomial and exponential function model better reflect the link between the RGB value and the related pesticide residue concentration, as determined by repeated modeling.

Quadratic polynomial regression equation:

$$Y = X_1 + X_2 \cdot R + X_3 \cdot G + X_4 \cdot B + X_5 \cdot R^2 + X_6 \cdot G^2 + X_7 \cdot B^2 + X_8 \cdot R \cdot G + X_9 \cdot R \cdot B + X_{10} \cdot G \cdot B \quad (3)$$

In the equation, Y denotes the projected value of pesticide residue concentration ($\text{mg} \cdot \text{mL}^{-1}$). X_1 – X_{10} denotes the required parameter. R, G, B represents the color values R, G, B on the detection card.

Exponential function regression equation:

$$Y = X_1 + \exp(X_2 + X_3 \cdot R + X_4 \cdot G + X_5 \cdot B) \quad (4)$$

Y denotes the expected concentration of pesticide residues ($\text{mg} \cdot \text{mL}^{-1}$) in the formula. The parameter to be acquired is indicated by X_1 – X_5 . R, G, B represents the color values R, G, B of the detection card.

2.7.3. Neural Network Model

The neural network model is one of the most used models in machine learning. By emulating the operating mechanism of biological neural networks and connecting a large number of neuron nodes to construct a complex network that processes signals in parallel, it is possible to simulate biological neural networks. The RBF neural network of the feedforward neural network is employed in this paper [41–46]. It possesses the benefits of simple network topology, rapid training speed, robust local approximation ability, and high approximation precision.

The establishment of the RBF neural network model was carried out using Matlab R2016a software. The full name of Matlab R2016a is Matrixlaboratory, which is a powerful data analysis and processing software and enjoys a high reputation internationally. The core of the RBF neural network established in this paper is the Intrinsic function 'newrb' in Matlab R2016a. The hidden layer uses the Gaussian function. The parameter goal in newrb is set to 0.005, spread is set to 1, MN is set to 25, and DF is set to 1. When the maximum number of neurons increases to 25, the expected RBF neural network model can be obtained.

3. Results and Discussion

3.1. Image Processing Results

Figure 2 depicts an exhaustive collection of image processing procedures. Figure 2a shows the first representation. Figure 2b depicts the picture after median filtering, with the window size set at 5·5. The mask is seen in Figure 2c. By superimposing it on the original image, just the region of interest is preserved while the backdrop is filtered off. The

picture in RGB color space must be transformed to HSV color space in order to obtain the mask image. Then, by configuring the threshold, the color information of pixels within the threshold range is set to 255 and the color information of pixels beyond the threshold range is set to 0. Thus, it is possible to acquire the necessary mask picture.

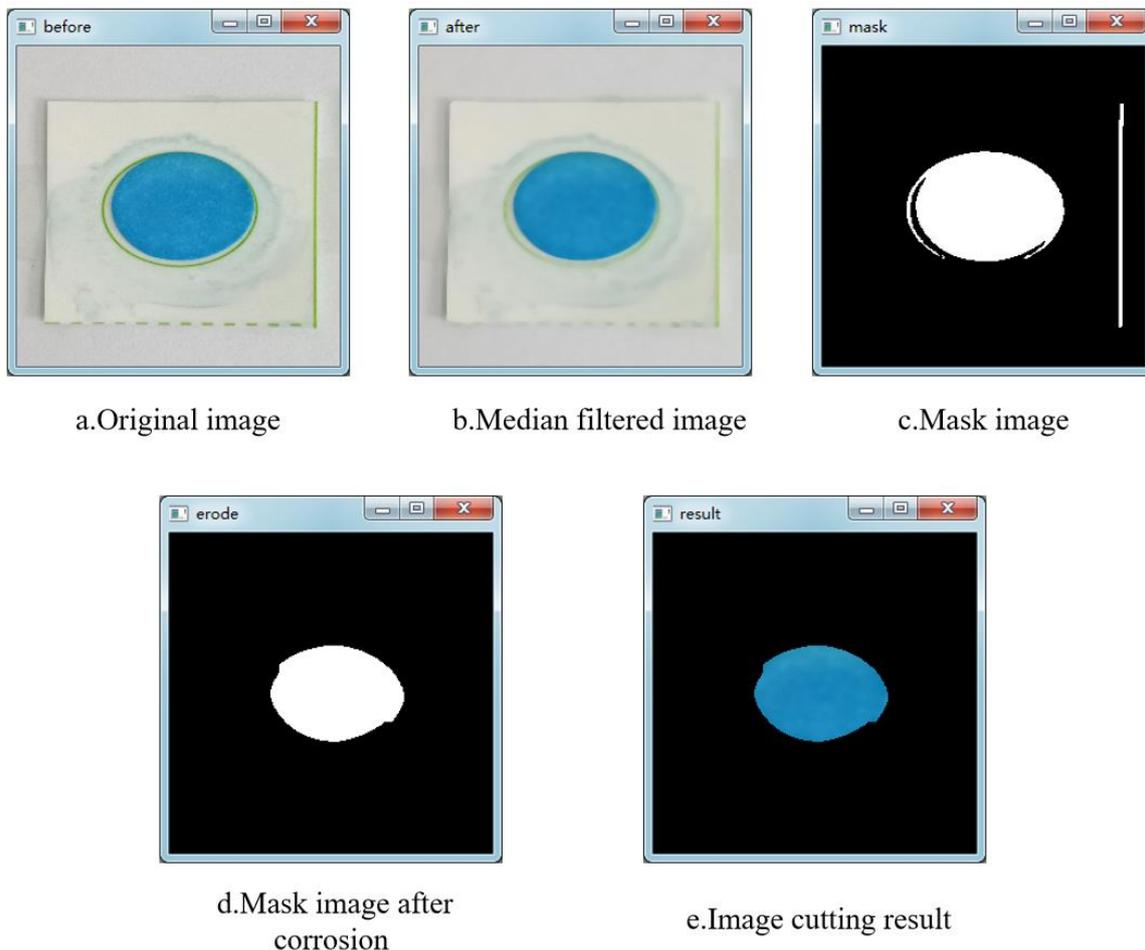


Figure 2. Image processing process.

The lowest limit of the HSV threshold is $[0, 43, 46]$, while the top limit is $[180, 255, 255]$. Figure 2d depicts the picture following the corrosion of the mask image. The etching process can fine-tune the mask picture in order to filter out leftover background information. In general, the edge of a pixel whose color information is 255 is diminished (the white part in the figure). The size of the etching operation's kernel and the number of iterations affect the degree of corrosion. The number of iterations is set to 5 and the kernel size of the corrosion operation is set to 3-3. Figure 2e depicts the outcome of cutting the picture.

3.2. Extraction Results of RGB Feature Values

The RGB histogram of the cropped image is calculated using the Python-OpenCV program. The histogram information returned is saved in an array. Extract the data from the array and generate the histogram based on this information. The histogram of the frequency of RGB feature values is depicted in Figure 3. By scanning the array data of each RGB histogram, the RGB value with the highest frequency is eliminated. After classification, induction, and sorting, the original data are acquired, and the results are depicted in Figure 4.

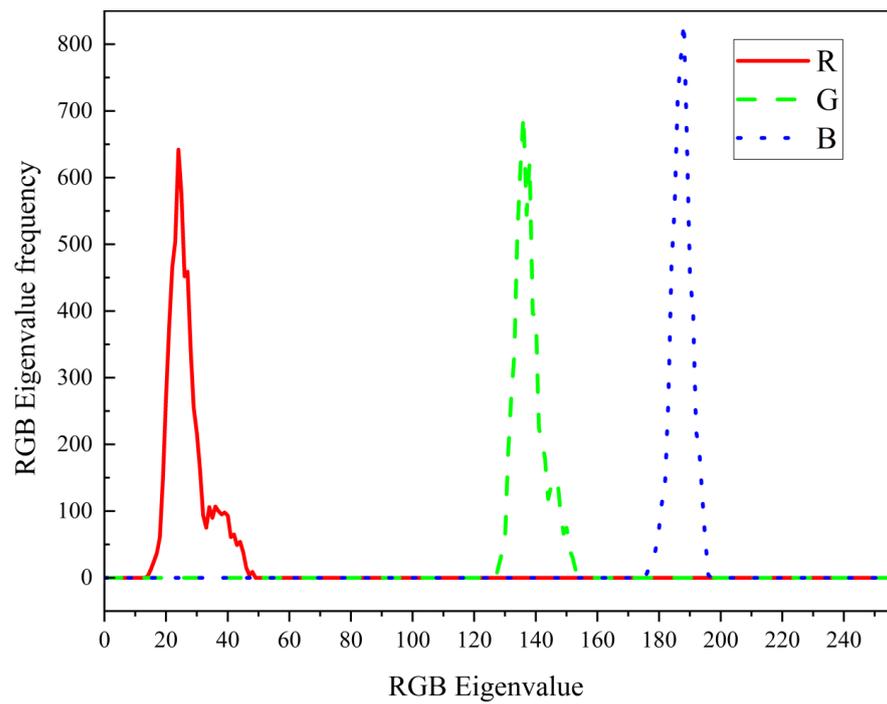


Figure 3. Frequency histogram of RGB eigenvalues.

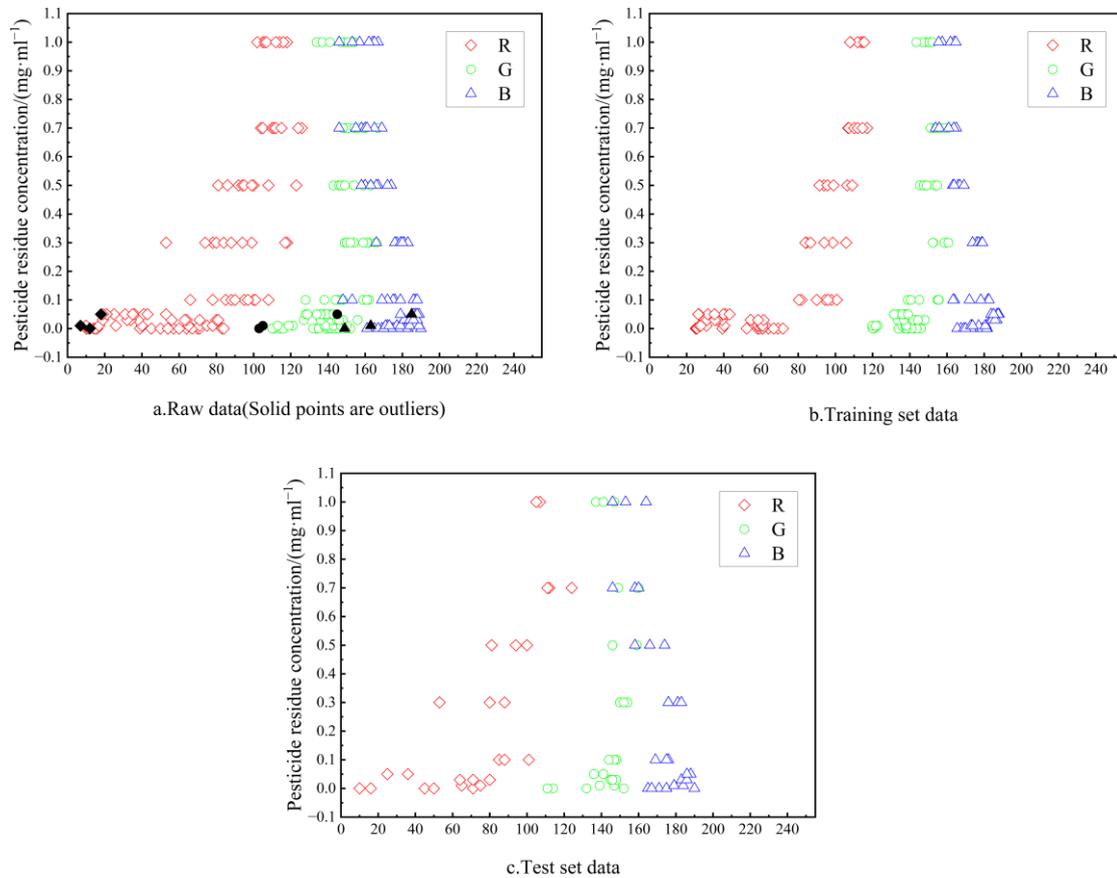


Figure 4. Data set.

3.3. Outlier Processing Result

The SPSS 26.0 statistical program—was used to compute the Mahalanobis distance. Using the Mahalanobis distance and Chi-square test, the original data were compared to exclude outliers. The Mahalanobis distance roughly follows the Chi-square distribution. Consequently, after computing the Mahalanobis distance of each sample data, the Chi-square test is performed. If the value exceeds the critical threshold, it is considered to be an outlier. The level of significance was fixed at 0.05, and the degree of freedom at three. Check the table of critical values for the Chi-square test; the critical value is 7.815. In conclusion, samples 12, 21, and 45 are identified as outliers based on the comparison of their Mahalanobis distances, which were 12.00773, 7.87728, and 19.60156, respectively. Delete these three samples, as detailed in Table 1 below.

Table 1. Summary of outliers.

Number	Pesticide Residue Concentration/mg·mL ⁻¹	R	G	B	Mahalanobis Distance	Chi-Square Test Critical Value
12	0.0002	12	103	149	12.00774	7.815
21	0.01	7	105	163	7.87728	7.815
45	0.05	18	145	185	19.60156	7.815

3.4. Results of Averaging

Figure 4a depicts the remaining 97 sample data after the Mahalanobis distance and Chi-square test were used to exclude the original data. Next, 70 sample data are picked at random, and an average processing operation is performed on the 70 sample data. As indicated in Figure 4b, the collected outcomes were utilized as training set data. Figure 4c demonstrates that the remaining 27 samples are assigned to the test set, and the original condition of the test set is maintained.

3.5. Comparison and Analysis of the Prediction Effect of Identification Model

Using the training set data after averaging and the training set data without averaging, respectively, four types of models are developed in this paper: the linear regression model, the exponential regression model, the quadratic polynomial regression model, and the RBF neural network model.

Linear regression model:

$$Y = 1.928932 + 0.004150 \cdot R + 0.002906 \cdot G - 0.013762 \cdot B \text{ (No Averaging)}$$

$$Y = 2.287768 + 0.004221 \cdot R + 0.003953 \cdot G - 0.016686 \cdot B \text{ (Averaging)}$$

Exponential regression model:

$$Y = -0.007099 + \exp(0.830657 + 0.049982 \cdot R - 0.031633 \cdot G - 0.012129 \cdot B) \text{ (No Averaging)}$$

$$Y = -0.003991 + \exp(0.293173 + 0.055380 \cdot R - 0.028222 \cdot G - 0.014946 \cdot B) \text{ (Averaging)}$$

Quadratic polynomial regression model:

$$Y = -12.689855 + 0.114999 \cdot R - 0.136337 \cdot G + 0.216906 \cdot B + 0.000368 \cdot R^2 + 0.001305 \cdot G^2 - 0.000307 \cdot B^2 - 0.001225 \cdot R \cdot G + 0.000080 \cdot R \cdot B - 0.000860 \cdot G \cdot B \text{ (No Averaging)}$$

$$Y = -9.942219 + 0.138477 \cdot R - 0.197519 \cdot G + 0.226283 \cdot B - 0.000005 \cdot R^2 - 0.000696 \cdot G^2 - 0.001275 \cdot B^2 + 0.000569 \cdot R \cdot G - 0.001229 \cdot R \cdot B + 0.002064 \cdot G \cdot B \text{ (Averaging)}$$

For the RBF neural network model, we use Matlab R2016a for modeling, which is a powerful data processing and analysis software. The Intrinsic function 'newrb' in Matlab R2016a is used for RBF neural network modeling. The parameter goal in newrb is set to 0.005, spread is set to 1, MN is set to 25, and DF is set to 1. The hidden layer function uses the Gaussian function for model approximation, and the model reaches the expected state after adding 25 neurons during training. Because RBF is a black box model, we will analyze the model through the data in the following text.

The results of a comparison of several prediction models are displayed in Table 2 and Figures 5–8. After data averaging, the degree of model fit for the four established models improved to varying degrees, as shown by the table data. The neural network model

demonstrates the minimum gain in model fit, with the coefficient of determination improving from 0.959 to 0.965 and the improvement value is 0.06. The linear regression model had the smallest improvement effect on model fit, with the coefficient of determination increasing from 0.655 to 0.738 and an improvement value of 0.083. The improved impact of the other models fell between that of the first two. The coefficient of determination of the exponential regression model rose from 0.787 to 0.899, and the improvement value was 0.112. The coefficient of determination of the quadratic polynomial regression model rose from 0.823 to 0.908, and the increased value was 0.085.

Table 2. Influence of data averaging on goodness of fit of model.

Training Set Data	Linear Regression Model	Quadratic Polynomial Regression Model	Exponential Regression Model	RBF Neural Network Model
Coefficient of determination (R ²)	0.655 ¹ /0.738 ²	0.823/0.908	0.787/0.899	0.959/0.965
Root mean square error	0.197/0.171	0.141/0.102	0.155/0.106	0.067/0.061

¹ Data are not averaged, ² Data are averaged.

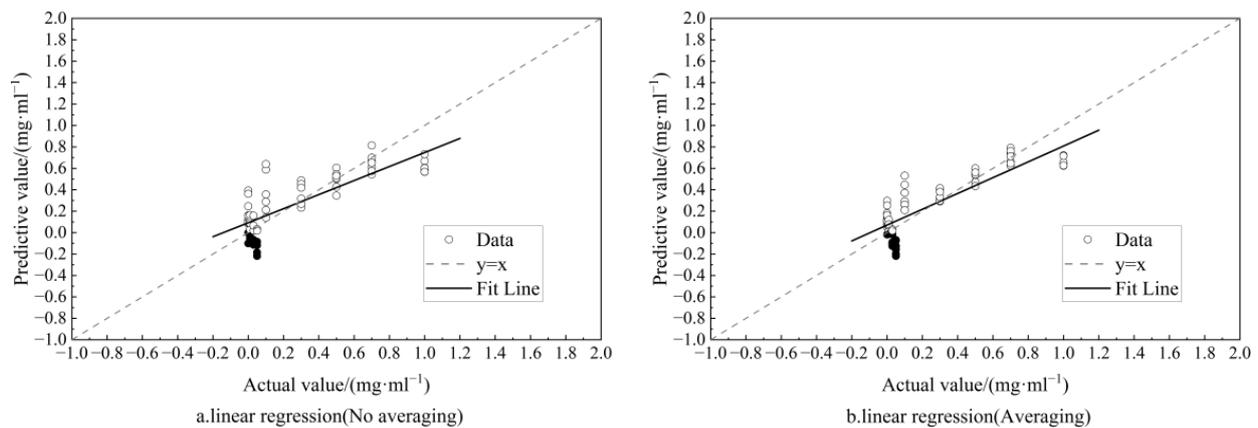


Figure 5. Comparison of prediction effects of linear regression models.

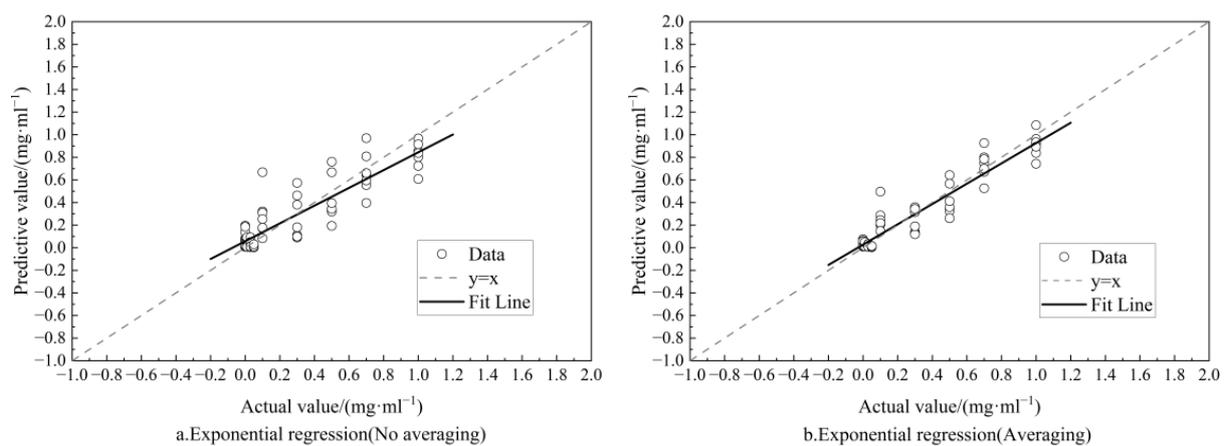


Figure 6. Comparison of prediction effect of exponential regression model.

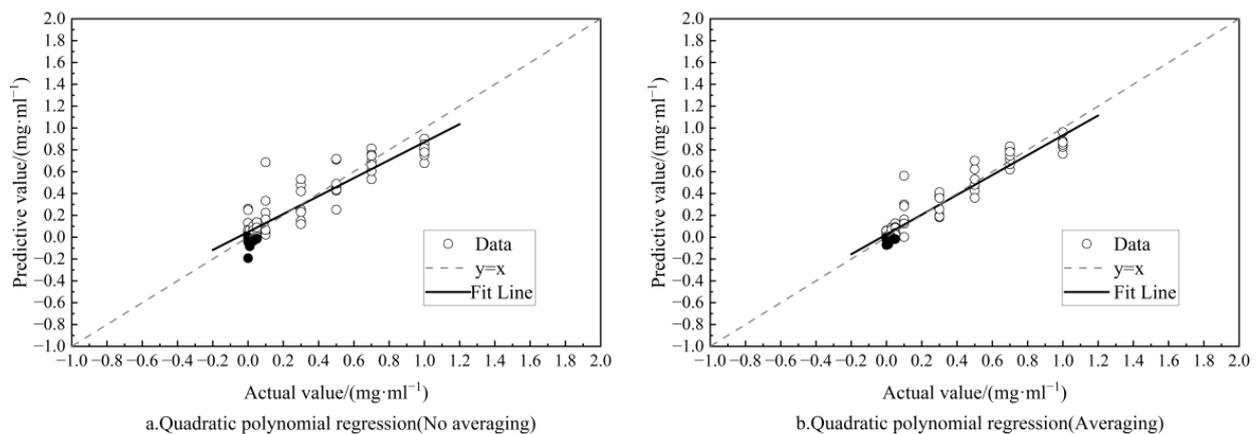


Figure 7. Comparison of prediction effect of quadratic polynomial regression model.

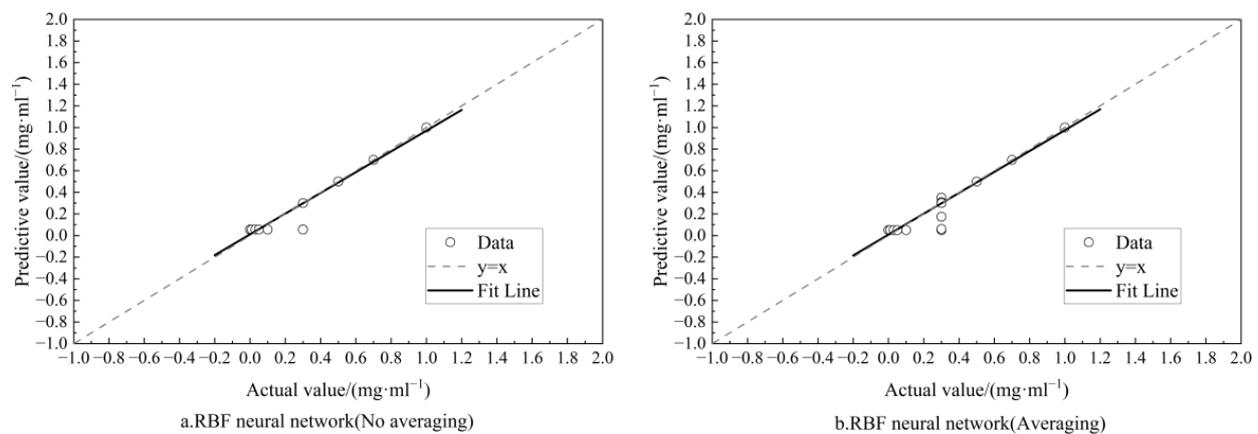


Figure 8. Comparison of prediction effect of RBF neural network model.

After data averaging, the model's root-mean-square error is also substantially decreased. The decrease in the root mean square error of the RBF neural network model is from 0.067 to 0.061, a difference of 0.004. The model of linear regression reduced by 0.026, from 0.197 to 0.171. Between the two models, the root-mean-square error of the other models was reduced, and the quadratic polynomial regression model fell from 0.141 to 0.102, a decrease of 0.039. The model of exponential regression fell from 0.155 to 0.106, a reduction of 0.049.

Theoretically, averaging will cause the data to gravitate toward its mean, reducing the negative influence of data clutter on model fitting and the process and improving the model fitting degree.

After data averaging, the model fit degree of numerous prediction models developed in this study will be enhanced to varied degrees, and the model's root-mean-square error will be greatly lowered. Since this finding is consistent with experimental predictions, it is possible to enhance the model's degree of fit by averaging the data.

Comparing the data of each regression model horizontally reveals that the quadratic polynomial regression model is superior to the other three. The coefficient of determination is 0.908, while the square root of the mean error is 0.102. In contrast, in the actual prediction, due to the nature of the quadratic polynomial, the anticipated value will be negative (solid point in Figure 7) when the expected concentration is near zero, which is incongruous with reality. Negative predictive values are also observed for linear regression (solid points in Figure 5). Not so with the exponential regression prediction model, whose predictive power is second only to that of the quadratic regression model. On the basis of the detection card, the exponential regression model is the optimal solution to the problem of identifying pesticide residues quantitatively.

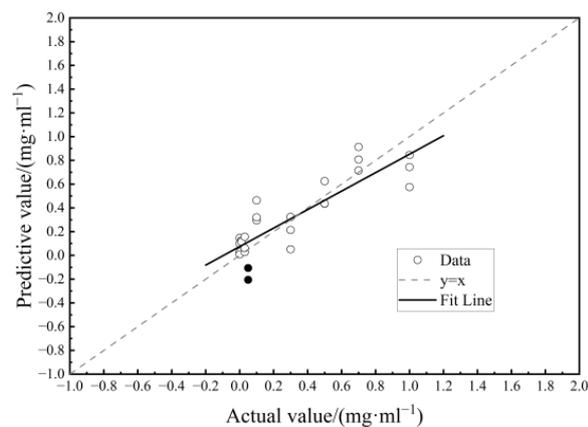
3.6. Verification of Quantitative Identification Model for Pesticide Residues

The outcomes of putting the data from the test set into the four prediction models to obtain the expected value are depicted in Figure 8. Calculations, induction, and sorting were used to determine the root mean square errors, and the resulting information is presented in Table 3.

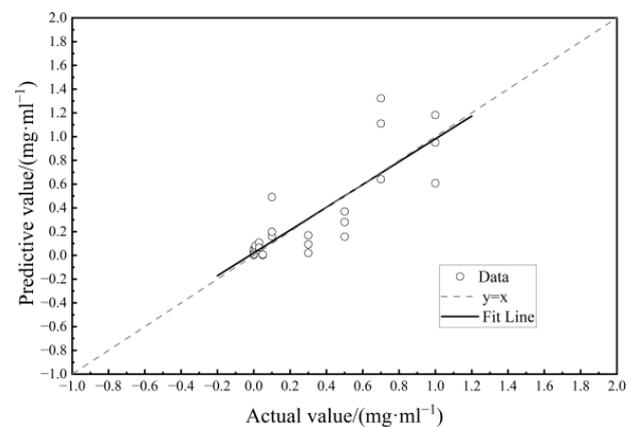
Table 3. Model validation.

Model Name	Linear Regression Model	Quadratic Polynomial Regression Model	Exponential Regression Model	RBF Neural Network Model
Root mean square error	0.173	0.158	0.216	0.452

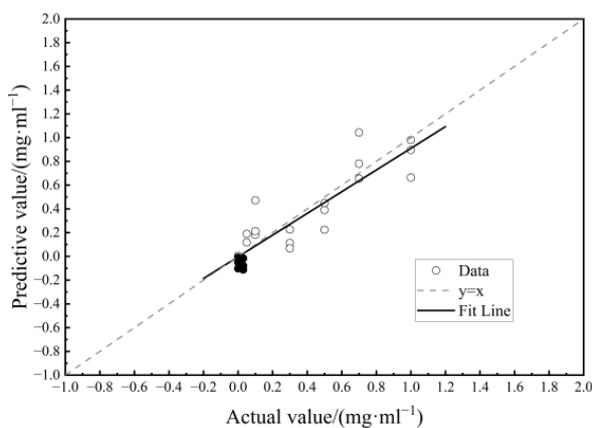
In terms of model verification, the root mean square errors of three traditional models are minimal and within the allowed range. However, RBF performed surprisingly poorly in the test results. Figure 9 demonstrates that the fitting line between the projected value and the actual value in the exponential regression model is closer to the straight line $y = x$, the prediction effect is significantly improved, and the predicted value is not negative. The overall forecast of the quadratic polynomial regression model is somewhat less accurate than that of the exponential regression model, and negative values (solid dots in Figure 9c) are observed when the anticipated concentration is near 0. The linear regression model also gave a negative result, with solid dots in Figure 9a. Although RBF neural network model has no negative value, the testing effect is the worst and cannot be applied.



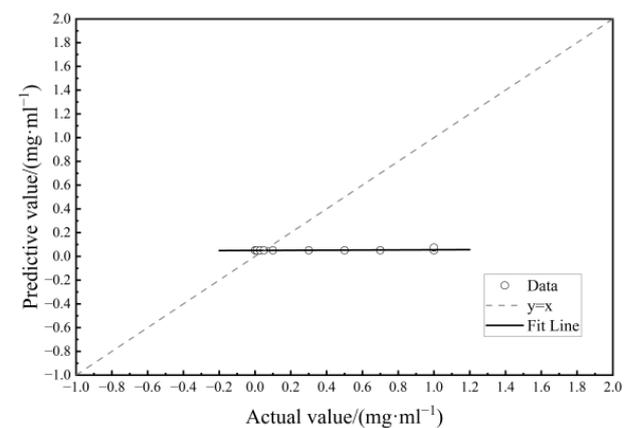
a. Verification of linear regression model



b. Verification of exponential regression model



c. Verification of quadratic polynomial regression model



d. Verification of RBF neural network model

Figure 9. Model validation.

4. Conclusions

We compared the effectiveness of four models in quantitatively identifying pesticide residue rapid detection cards. The determination coefficient, root mean square error, and actual prediction of the four models were evaluated. It was determined that the exponential regression model was better suited for the quantitative prediction of pesticide residue concentrations indicated by the detection card. This provides data support and ideas for other scholars to study the rapid identification of pesticide residue rapid detection card technology. The impact of the data averaging strategy on the model fit degree of the linear regression model, quadratic polynomial regression model, exponential regression model, and RBF neural network model is explored. The findings reveal that following data averaging, the determination coefficients of the four models improve to varied degrees due to greater data quality. We analyze that the improvement in data quality is due to the correction of scattered data structures by data averaging. The data averaging idea is a commonly used data preprocessing method used to map data points to values relative to the entire data range in a certain way. As it is based on the overall relative mapping, it does not change the adaptability of the established model to data that have not been data averaged and can play an optimization role in modeling. Through the work of this work, it is proven that our proposed mapping formula can effectively optimize the data structure and modeling process, improve the model fit of the established model, and is of great significance in the field of data preprocessing. It provides data preprocessing methods and ideas for researchers. The root-mean-square error also dropped to varied degrees. This approach may also be used in the modeling process in other domains and serves as a reference for other researchers to handle data. This study's quantitative identification process of pesticide residue based on detection card can give an algorithm basis and data support for the development of image recognition equipment of pesticide residue detection card based on enzyme inhibition technique. Ordinary families, when combined with the use of a detection card, may readily assess pesticide residues in fruits and vegetables at home, with cheap cost, simple operation, and easy marketing.

Author Contributions: Conceptualization, Q.Z. and X.C.; methodology, Y.Z. and X.C.; software, Y.Z.; validation, Y.Z., J.D. and Y.G.; formal analysis, M.Z.; investigation, Y.D.; resources, Q.Z. and X.C.; data curation, Y.Z. and X.C.; writing—original draft preparation, Y.Z.; writing—review and editing, Y.Z. and X.C.; visualization, Y.Z. and X.C.; supervision, H.T.; project administration, Q.Z. and X.C. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Joint Funds of the Zhejiang Provincial Natural Science Foundation of China (LGN21C030001), and the Quzhou science and technology plan project (2018K38).

Data Availability Statement: No new data were created.

Acknowledgments: This work was supported from the Basic Public Research Project of Zhejiang Province (LGN21C030001), and the Quzhou science and technology plan project (2018K38).

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

References

1. Ping, H.; Wang, B.H.; Li, C.; Li, Y.; Ha, X.J.; Jia, W.S.; Li, B.R.; Ma, Z.H. Potential health risk of pesticide residues in greenhouse vegetables under modern urban agriculture: A case study in Beijing, China. *J. Food Compos. Anal.* **2022**, *105*, 104222. [[CrossRef](#)]
2. Omeje, J.S.; Asegbeloyin, J.N.; Ihedioha, J.N.; Ekere, N.R.; Ochonogor, A.E.; Abugu, H.O.; Alum, O.L. Monitoring of pesticide residues in fresh fruits and vegetables available in Nigerian markets and assessment of their associated health risks. *Environ. Monit. Assess.* **2022**, *194*, 516. [[CrossRef](#)]
3. Chauhan, P.; Sundari, S.K. Harmful Effects of Organophosphate Pesticides: A Study using Yeast Model System. *Indo Glob. J. Pharm. Sci.* **2017**, *7*, 40. [[CrossRef](#)]
4. Seo, S.-H.; Choi, S.-D.; Stuart, B.; Chang, Y.-S. Health risk assessment of exposure to organochlorine pesticides in the general population in Seoul, Korea over 12 years: A cross-sectional epidemiological study. *J. Hazard. Mater.* **2022**, *424*, 127381. [[CrossRef](#)]
5. Mojtaba, S.; Sadegh, A.M.; Jhalukpreya, S. Health Risk of Potato Farmers Exposed to Overuse of Chemical Pesticides in Iran. *Saf. Health Work.* **2021**, *13*, 23–31.

6. Zhu, S.M.; Zhou, C.N.; He, J.S.; Zhang, X.Y.; Guo, X.S. Rapid colorimetric detection of pesticide residues based on enzyme inhibition method. *Nongye Gongcheng Xuebao* **2014**, *30*, 242–248.
7. He, S.G.; Zhang, S.F.; Zhao, X.; Zhu, X.Y.; Chen, L.S.; Cui, J.N. Highly selective NIR fluorescent probe for acetylcholinesterase and its application in pesticide residues detection. *Chin. Chem. Lett.* **2022**, *33*, 4233–4237. [[CrossRef](#)]
8. Wan, Y.; Wang, H.T.; Zhang, L.; Chen, Y.X.; Li, S.; Zhou, J.; Zhang, Q.; Xia, L.X. Highly stable acetylcholinesterase electrochemical biosensor based on polymerized ionic liquids microgel for pesticides detection. *Microchim. Acta* **2022**, *189*, 300. [[CrossRef](#)] [[PubMed](#)]
9. Silva, W.D.; Ghica, M.E.; Christopher, M.A.B. Choline oxidase inhibition biosensor based on poly(brilliant cresyl blue)—Deep eutectic solvent/carbon nanotube modified electrode for dichlorvos organophosphorus pesticide. *Sens. Actuators B* **2019**, *298*, 126862. [[CrossRef](#)]
10. Jiang, X.S.; Wang, W.Q.; Xu, L.Y.; Lu, L.Q.; Zhou, H.P.; Chen, H.H. Review on rapid detection of pesticide residues in agricultural and food products. *Nongye Gongcheng Xuebao* **2016**, *32*, 267–274.
11. Wang, D.W.; Liu, C.; Zhou, Z.Q.; Wang, P. Recent advances in rapid detection of pesticide residue. *Nongyaoxue Xuebao* **2019**, *21*, 852–864.
12. Chen, F.H.; Zhao, Y.P.; Zhang, S.X.; Wei, S.H.; Ming, A.J.; Mao, C.H. Hydrophobic Wafer-Scale High-Reproducibility SERS Sensor Based on Silicon Nanorods Arrays Decorated with Au Nanoparticles for Pesticide Residue Detection. *Biosensors* **2022**, *12*, 273. [[CrossRef](#)] [[PubMed](#)]
13. Liao, H.J.; Zhang, X.C.; Hu, L.Y.; Zhong, H.C.; Sun, G.Y.; Liu, R.F. Simultaneously Detection of 10 Kinds of Organophosphorus Pesticides Residues in Tea by Gas Chromatography. *Chin. J. Food Hyg.* **2015**, *27*, 41–44.
14. Wu, D.F.; Wang, L.; He, M.M.; Wu, J.J.; Dai, P.Z.; Zhu, L.F. HPLC Determination of Residual Amounts of 19 Carbamate Pesticides in Vegetables. *Lihua Jianyan Huaxue Fence* **2020**, *56*, 142–147.
15. Guo, L.Q.; Liu, P.; Zhang, X.K.; Zhang, L.; Li, Y.J. Rapid Determination of Formaldehyde in Paper Cups by High Performance Liquid Chromatography. *Zhongguo Zaozhi Xuebao* **2019**, *34*, 38–41+52.
16. Ansari, S.; Manvar, T.; Kumar, D.; Gupta, R.; Sivaperumal, P. A Rapid, Efficient, Reliable Method for Simultaneous Determination of Pesticide Residues in Green Chilli by Gas Chromatography-Mass Spectrometry. *Asian J. Chem.* **2022**, *34*, 2379–2385. [[CrossRef](#)]
17. Liu, X.; Liu, Z.T.; Bian, L.L.; Ping, Y.F.; Li, S.H.; Zhang, J.R.; Wang, J.M.; Ann, V.S.; Wang, X. Determination of pesticide residues in chilli and Sichuan pepper by high performance liquid chromatography quadrupole time-of-flight mass spectrometry. *Food Chem.* **2022**, *387*, 132915. [[CrossRef](#)] [[PubMed](#)]
18. Chen, Z.P.; Ma, J.B.; Du, J.H. Self-service detection platform of pesticide residue based on enzyme inhibitory method. *Nongye Gongcheng Xuebao* **2019**, *35*, 289–294.
19. Ranveer, S.A.; Harshitha, C.G.; Dasriy, V.; Nimisha, T.; Naresh, K.; Raghu, H.V. Assessment of developed paper strip based sensor with pesticide residues in different dairy environmental samples. *Curr. Res. Food Sci.* **2023**, *6*, 100416. [[CrossRef](#)]
20. Jia, B.H.; Zhang, P.Q.; Luo, Y.Q.; Chen, L. Rapid determination of pesticide residues based on machine vision. *Comput. Eng. Appl.* **2017**, *53*, 231–234.
21. Dutta, M.K.; Sengar, N.; Minhas, N.; Sarkar, B.; Goon, A.; Banerjee, K. Image processing based classification of grapes after pesticide exposure. *LWT-Food Sci. Technol.* **2016**, *72*, 368–376. [[CrossRef](#)]
22. Yan, L.; Lu, A.X.; Luan, Y.X.; Ren, D.; Wang, J.H. Rapid detection of pesticide residues based on image processing. *Shipin Anquan Zhiliang Jiance Xuebao* **2014**, *5*, 748–753.
23. Fanfani, M.; Iuliani, M.; Bellavia, F.; Colombo, C.; Piva, A. A vision-based fully automated approach to robust image cropping detection. *Signal Process. Image Commun.* **2020**, *80*, 115629. [[CrossRef](#)]
24. Almazaydeh, L.; Salah, M.; Misto, M.; Nairok, E.; Alsayed, S.; Elleithy, K. Speed-limit Signs Detection and Recognition Based on HSV Color Model. *Asian J. Sci. Res.* **2016**, *9*, 219–222. [[CrossRef](#)]
25. Hu, X.Y.; Li, F.Z. Study on Classification Corn Seedling and Weed Based on RGB Model and HSV Model. *Agric. For. Fish.* **2017**, *6*, 49. [[CrossRef](#)]
26. Gamage, K.M.; Anil, B.; Eun, K.N.; Tae, K.H. Measurement of Overlapping Leaf Area of Ice Plants Using Digital Image Processing Technique. *Agriculture* **2022**, *12*, 1321.
27. Anastasia, U.; Aleksandra, K.; Dmitrii, S.; Ksenia, E.; Evgeny, M.; Nikita, R. Computer vision-based platform for apple leaves segmentation in field conditions to support digital phenotyping. *Comput. Electron. Agric.* **2022**, *201*, 107269.
28. Patil, M.A.; Manohar, M. Enhanced radial basis function neural network for tomato plant disease leaf image segmentation. *Ecol. Inform.* **2022**, *70*, 101752. [[CrossRef](#)]
29. Lu, Z.A.; Qi, L.J.; Zhang, H.; Wan, J.J.; Zhou, J.R. Image Segmentation of UAV Fruit Tree Canopy in a Natural Illumination Environment. *Agriculture* **2022**, *12*, 1039. [[CrossRef](#)]
30. Song, Q.Q.; Yang, G.P. Lane Line Detection Based on HSV Color Threshold Segmentation. *Comput. Digit. Eng.* **2021**, *49*, 1895–1898.
31. Zhao, H.; Gan, Z.W.; Xiao, M. The method of judging outliers in multivariate statistical data. *J. Cent. China Norm. Univ. (Nat. Sci.)* **2003**, *49*, 133–137.
32. Gao, L.Z.; Lu, C.Y.; Guo, G.D.; Zhang, X.; Lin, S. Quantum K-nearest neighbors classification algorithm based on Mahalanobis distance. *Front. Phys.* **2022**, *10*, 1–6. [[CrossRef](#)]
33. Dashdondov, K.; Kim, M.H. Mahalanobis Distance Based Multivariate Outlier Detection to Improve Performance of Hypertension Prediction. *Neural Process. Lett.* **2021**, *28*, 1–13.

34. Xiao, X.L.; Chen, C.M. Establishment and Verification of Multivariate Linear Regression Model for Prediction of Ethanol Concentration. *Infrared Technol.* **2018**, *32*, 115–120.
35. Qiu, R.C.; Miao, Y.L.; Zhang, M.; Li, H.; Sun, H. Modeling and verification of maize biomass based on linear regression analysis. *Trans. Chin. Soc. Agric. Eng.* **2018**, *34*, 131–137.
36. Alharbi, R.; Alageel, N.; Alsayil, M.; Alharbi, R.; Alhakamy, A. Prediction of Oil Production through Linear Regression Model and Big Data Tools. *Int. J. Adv. Comput. Sci. Appl. (IJACSA)* **2022**, *13*. [[CrossRef](#)]
37. Abid, N.; Saleem, U.; Ahmad, S.Z.; Azhar, A.; Asad, A.; Shahid, I.M.; Khalid, H.; Muhammad, S.; Munawar, S.; Usman, B.M. Estimation and Forecasting of Rice Yield Using Phenology-Based Algorithm and Linear Regression Model on Sentinel-II Satel-lite Data. *Agriculture* **2021**, *11*, 1026.
38. Liao, D.Q.; Wu, Y.; Chai, C.C. Construction of Nonlinear Regression Model Based on Relationship between Air Pollutants and Meteorological Elements. *Meteorol. Sci. Technol.* **2020**, *48*, 871–876.
39. Xu, Y.C.; Wang, X.Y.; Yin, X.; Yue, R.C.; Hu, Z.X.; Sun, J.B. Potato Processing Quality Characteristics Prediction Based on Multivariate Nonlinear Regression Analysis. *Trans. Chin. Soc. Agric. Mach.* **2018**, *49*, 366–373.
40. Daniel, C.S.; Ordoñez Salanueva, C.A.; Salvador, S.M.; Campos, J.E.; Alma, O.S.; Flores Ortiz, C.M. Quantifying Cardinal Temperatures of Chia (*Salvia hispanica* L.) Using Non-Linear Regression Models. *Plants* **2022**, *11*, 1142.
41. Cai, X.W.; Bao, Y.Q.; Hu, M.F.; Liu, J.B.; Zhu, J.M. Simulation and Prediction of Fungal Community Evolution Based on RBF Neural Network. *Comput. Math. Methods Med.* **2021**, *2021*, 7918192. [[CrossRef](#)] [[PubMed](#)]
42. Efi, Y.; Amit, F.; Victor, L.; Shabtai, C.; Meir, T.; Josef, T. Estimating Evapotranspiration of Greenhouse Banana Plantations Using Artificial Neural Network and Multiple Linear Regression Models. *Water* **2022**, *14*, 1130.
43. Li, Q.Q.; Wang, C.G.; Yue, T.X.; Li, B.; Yang, J. Method for spatial variety of soil organic matter based on radial basis function neural network. *Trans. CSAE* **2010**, *26*, 87–93.
44. Bhusan, C.B.; Mohan, P.B.; Asit, R.; Sudipta, J.; Ambika, S.; Sujata, M.; Jeetendranath, P.; Kumar, N.P.; Chandra, P.P.; Sanghamitra, N. Implementation of multilayer perceptron (MLP) and radial basis function (RBF) neural networks for predicting Shatavarin IV content in *Asparagus racemosus* accessions. *Ind. Crops Prod.* **2023**, *191*, 115968.
45. Xie, Y.F.; Yu, J.J.; Xie, S.W.; Huang, T.W.; Gui, W. H On-line prediction of ferrous ion concentration in goethite process based on self-adjusting structure RBF neural network. *Neural Netw.* **2019**, *116*, 1–10. [[CrossRef](#)] [[PubMed](#)]
46. Liu, A.Q.; Zhao, D.F. Dispersion Model of Volatile Organic Compounds Based on RBF Neural Network. *J. Phys. Conf. Ser.* **2019**, *1237*, 052024.

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.