IFERP

ISSN (Online) 2395-2717

International Journal of Engineering Research in Electrical and Electronic Engineering (IJEREEE)

Volume 11 Issue 07 July 2024

Detect Level of Pesticide Residue on Butterhead Lettuce using Near-Infrared (NIR) Hyperspectral Imaging

[1] Huy N. Nguyen, [2] Uyen D. Nguyen, [3] Quoc T. Huynh, [4] Khang P. Tran, [5] Chi D. Ngo

[1][2][3][4] School of Electrical Engineering, International University, Ho Chi Minh City, Vietnam
[1][2][3][4] Vietnam National University, Ho Chi Minh City, Vietnam
[5] 5D Agriculture Solutions Joint Stock Company, Ho Chi Minh City, Vietnam
Email: [1] nguyenngochuy3001@gmail.com, [2] nduyen@hcmiu.edu.vn, [3] htquoc@hcmiu.edu.vn,
[4] khang.tran852456@gmail.com, [5] chingoduc@gmail.com

Abstract*— Near-infrared (NIR) hyperspectral imaging method had been utilized for the characterization of pesticides on Butterhead lettuce leaves. For this study, butterhead lettuce leaves were contaminated with three different concentrations of pesticide. Imidacloprid was used as the contaminated pesticide with a concentration of 1:1000, 5:1000, and 10:1000. The hyperspectral data of each contaminated leaf was collected and recorded by the OCI-F series of BaySpec hyperspectral camera. The region of interest (ROIs) of the leaves were selected by CubeCreator from Bayspec. The wavelength range of the camera is from 400 to 955 nm, with a resolution of 9 nm. The Standard Normal Variate (SNV) algorithm was applied to scale the NIR hyperspectral data. Based on the different shapes of the NIR spectrum between 530 to 590 nm, the Root Mean Square Error (RMSE) was used to detect and characterize the pesticide residue. The Confusion matrix was implemented to calculate the accuracy of the algorithm. The proposed algorithm to characterize the pesticide residue on Butterhead lettuce leaves produced an accuracy of approximately 90.25%.

Index Terms*— Bayspec, OCI-F series, Hyperspectral, Lactuca sativa, pesticide, Imidacloprid, Standard Normal Validate (SNV), Root Mean Square Error (RMSE).

I. INTRODUCTION

Butterhead Lettuce (Lactuca sativa var. Capitata) is a very popular vegetable in many temperate climate countries such as the US and Japan. In the past, Butterhead lettuce was not popular in Vietnam because it is difficult to grow and mature Butterhead lettuce in tropical climates. Butterhead lettuce contains more iron than other types of lettuce. Additionally, butterhead provides lots of Vitamin A and Vitamin K, as well as some manganese, folate, and potassium. Recently, Butterhead lettuce was able to flourish in the tropical climate by the technique of Root Zone Cooling (RZC) see ref link. Therefore, Butterhead lettuce has become a popular vegetable to grow by Vietnamese farmers [1].

There are many different types of diseases associated with Butterhead Lettuce. One of the common types of disease is Thrips (Thysanoptera). To limit Thrips, farmers often spray pesticides (Imidacloprid) on the lettuce leaves. The existence of pesticide residues on lettuce leaves is a major problem in exporting to foreign markets or human consumption. For a human, long-term consumption of vegetables containing pesticide residues can induce various diseases and may even lead to cancer and genetic mutations [2]. Knowing the quantity of pesticide residue is one of the key factors to judge butterhead lettuce quality.

In addition to the hyperspectral imaging method, there are many other mature and feasible methods for the detection of pesticide residues on fruits and vegetables, (e. g. gas chromatography and high-performance liquid chromatography), but most of these methods rely on chemical reagents and cause damage to the sample [3]. However, hyperspectral imaging method has dual advantages, without chemical reagents and non-damaging the sample, of detecting pesticide residues. Hyperspectral imaging detection process is a non-destructive and non-polluting method. In addition, the sample does not require pretreatment, therefore hyperspectral imaging detection is a widely selected method in modern agricultural detection and analysis [4].

Previous studies have used hyperspectral signatures to detect the level of pesticide residue. REN Zhan-qi, RAO Zhen-hong, and JI Hai-yan [5] used Hyperspectral image technology to identify spinach leaves containing different concentrations of dimethoate pesticide residues. The authors selected 4 regions of interest (ROIs) and calculated the average spectral curves of spinach leaves ROIs with different concentrations. Moreover, they used machine learning algorithms and the Chi-square test algorithms to find the best model for discriminating spinach leaves with different concentrations of dimethoate pesticide residues. The results show that using the Chi-Squared Test combined with four kinds of machine learning classification algorithms to select the best band and the optimal LDA discriminant model. The 8 characteristic wavelengths selected were: 1445.8, 1449, 1452.3, 1455.5, 1458.7, 1462, 1465.2, and 1468.4 nm, which

ISSN (Online) 2395-2717

International Journal of Engineering Research in Electrical and Electronic Engineering (IJEREEE)

Volume 11 Issue 07 July 2024

reduced the running time of the model. The results showed that spinach leaves containing different concentrations of dimethoate pesticide residues could be accurately identified based on hyperspectral imaging.

Weixin Ye 1, Tianying Yan. et al. [6] have employed hyperspectral imaging in conjunction with machine learning techniques to detect and quantify pesticide residue levels in grapes. By leveraging advanced technologies, the research aims to enhance the accuracy and efficiency of pesticide assessment in grape cultivation. In this investigation, visible/near-infrared (Vis-NIR) (376–1044 nm) and near-infrared (NIR) (915–1699 nm) hyperspectral imaging systems (HISs) were employed for the assessment of pesticide residue levels. Four concentrations of pesticides were applied to three distinct grape varieties. Moreover, the authors used several algorithms to establish classification models for different concentrations of pesticide on grapes, including logistic regression (LR), support vector machine (SVM), random forest (RF), convolutional neural network (CNN), and residual neural network (ResNet). The authors had found the NIR spectra performed better than those of Vis-NIR spectra. The accuracy for NIR and Vis-NIR were 93% and 97%, respectively.

In this paper, hyperspectral imaging method is used to recognize the pesticide residue on the butterhead lettuce leaves. Specifically, we will focus on detecting the level of pesticide residue on butterhead lettuce leaves. The Standard Normal Variate (SNV) [7] and Root Mean Square Errors (RMSE) [8] methods are both applied in this study. The SNV method is used to improve spectral linearity and prediction accuracy. To determine the levels of contamination of pesticide (Imidacloprid), the RMSE is employed to compare the Unknown Model Datasets against the Known Model Datasets.

The goal of this paper is to use the hyperspectral images generated by the OCI-F of BaySpec hyperspectral camera to detect the levels of pesticide residue on Butterhead lettuce leaves.

II. METHODOLOGY

1) Equipment

Butterhead lettuce plants were grown in the greenhouse in Ho Chi Minh City. The experimental pesticide was imidacloprid, which was diluted to 0:1000 (water only), 1:1000, 5:1000, and 10:1000 solutions by water. The solutions were sprayed evenly onto the surfaces of 100 butterhead leaves. We divided the samples into 4 groups, including 3 groups with 25 leaves sprayed with each concentration and 1 group with 25 leaves sprayed with pure water. The leaves were left to stand for 12 hours before the hyperspectral image was collected.

Hyperspectral images were collected by the OCI-F hyperspectral camera. The OCI™-F push-broom hyperspectral cameras of Bayspec are high-performance, ultra-miniaturized hyperspectral imaging engines, packed with SuperSpeed USB 3.0 interface. The BaySpec camera (shown in Figure 1) used in this research has a range of 400–955 nm, which covers the entire visible (VIS) spectrum and some of the near-infrared (NIR) region. We only collect data from the spectral range of 450–955 nm due to the excessive noise from the spectral range of 400-449 nm [9].

Figure 1: Hyperspectral camera

2) Preparation samples

In this study, Butterhead lettuce plants were cultivated in a hydroponics greenhouse. Hydroponics, a soilless cultivation method, was employed for optimal control over nutrient delivery to the plants. Butterhead lettuce samples were selected at a growth stage of approximately 30 days, ensuring a representative snapshot of the plant's development. Hydroponics was chosen as the cultivation method to eliminate soil interference in pesticide residue analysis. The lettuce plants were grown in nutrient-rich water solutions within a controlled greenhouse environment. Pesticide treatment was administered through a spraying method to simulate real-world conditions. Imidacloprid, a common pesticide, was applied to the lettuce leaves. The dosage and application were carefully controlled to ensure uniform distribution across the sample set.

After the leaves were left to stand for 12 hours, each concentration group of leaf samples were scanned by the OCI™-F push-broom hyperspectral camera to generate the hyperspectral images. For each concentration group, there are 25 leaf samples. For each leaf sample, the hyperspectral camera generated 64 spectral images of different wavelengths. The wavelength of the spectral images ranges from 400 to 955 nm, with an average resolution of 9 nm. CubeCreator software was used to select the Region of Interest (ROIs) of each leaf sample. There are 8 ROIs per leaf sample (show in Figure 2). The 8 different ROIs were deliberately chosen to exclude the main stem. The size of each ROI is 25x25 pixels. The average spectral value of each ROIs is calculated. Therefore, for each ROI, there are 64 points raw spectral curve is collected. For each concentration group with 25 leaf samples and with 8 ROIs per leaf sample, there were 200 raw spectral curves gathered. As the results, there were 800 raw spectral curves collected for all 4 concentration groups.

IFERP
Explore Your Research Journey..

ISSN (Online) 2395-2717

International Journal of Engineering Research in Electrical and Electronic Engineering (IJEREEE)

Volume 11 Issue 07 July 2024

Figure 2: ROIs selection for leaf samples.

3) Data analysis

For the preparation of the Known Model Datasets, we used 100 raw spectral curves of each concentration group. Based on the SNV method, we calculated SNV values of 100 raw spectral curves and generated 100 SNV spectral curves. There is a total of 400 SNV spectral curves for the preparation of the Known Model Datasets.

- *Standard Normal Variate (SNV) method*: After SNV processing, each SNV spectral curve will have a mean of 0 and a standard deviation of 1. Eq. 1 is the equation that is used to perform SNV processing.

$$w_i = \sqrt{\frac{\sum_{j=1}^{n}(X_{i,j}-\bar{x}_i)^2}{(n-1)}} + \delta^{-1} \tag{1}$$

Where n is the sequence number of each wavelength, $X_{i,j}$ is the value of the jth wavelength for the ith raw spectral curve, and δ is a user-definable offset.

- *Calculate Average Data*: Each Known Model Dataset of each concentration group has 100 SNV spectral curves. We used Eq.2 to calculate the average spectral curve for each concentration group. The result is an average spectral curve generated from 100 SNV spectral curves. As the results, there are 4 average spectral curves generated.

$$\bar{x}_i = \frac{\sum_{j=1}^{n} X_{i,j}}{n} \tag{2}$$

Where j is number of SNV spectral curves, i is concentration group, $X_{i,j}$ is point raw spectral curve, n is the total number of SNV spectral curves in each concentration group, $\bar{x}_i$ is the average spectral curve of each concentration group.

Figure 3 shows the average spectral curves of butterhead lettuce leaves for each concentration group from the wavelength range of 450 to 955 nm. Clearly, the average spectral curves shift in the wavelength range of 530 to 590 nm has the largest distance between each concentration group. The largest difference is approximately 0.048 in NIR reflection value (see Figure 4).

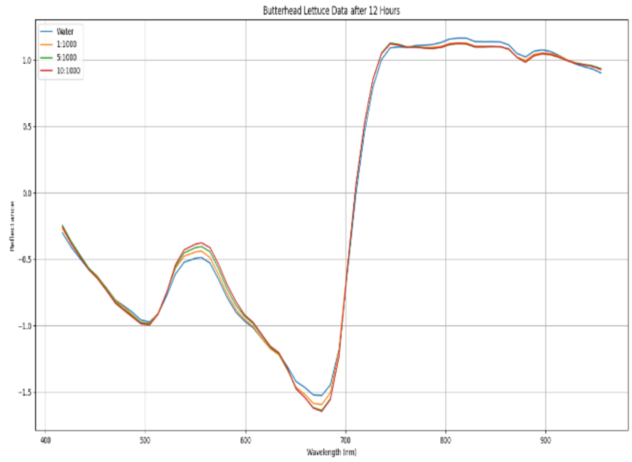

Figure 3: Average spectral curves of each concentration group from the wavelength range of 450 to 955 nm.

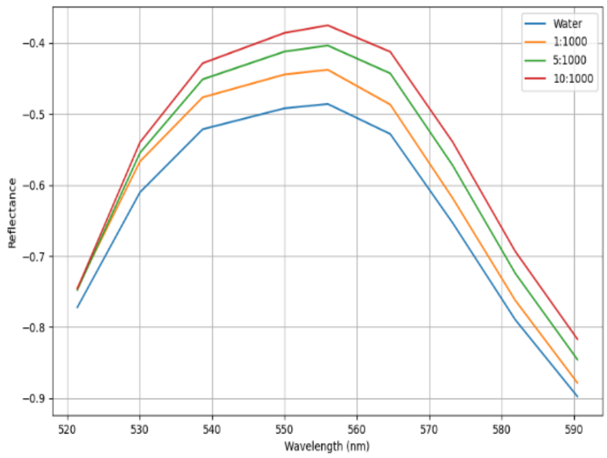

Figure 4: Average spectral curves of each concentration group from the wavelength range of 530 nm to 590 nm.

4) Pesticide Detection Algorithm

In this section, the Unknown Model Datasets were used to test the algorithm. The Unknown Model Datasets include 100 unknown raw spectral curves remaining of each group (a total of 400 raw spectral curves). The SNV method is applied to generate 400 unknown SNV spectral curves from 400 raw spectral curves. Each unknown SNV spectral curve is used to perform the RMSE test with the Known Model Datasets.

The flowchart of the pesticide detection algorithm is summarized in Figure 5.

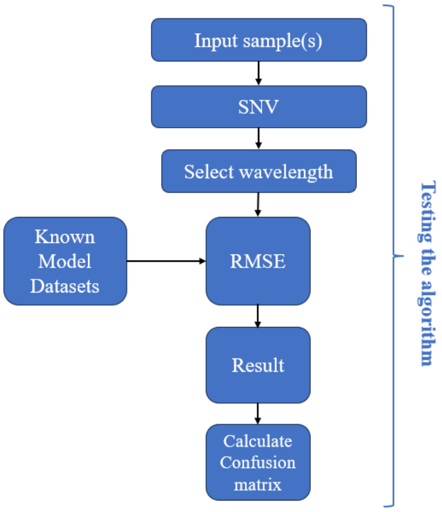

Figure 5: The flowchart of the pesticide detection algorithm.

- *Select wavelength:* As described in section 3, the optimum wavelength range of the Known Model Datasets are identified with the largest difference in the NIR reflection value. The wavelength range for the calculation of the RMSE is between 530 to 590 nm.
- *Root Mean Square Error (RMSE):* The RMSE is applied to calculate the distance of the unknown SNV spectral curves and the Known Model Datasets in the wavelength range of 530 to 590 nm. The RMSE, which measures the difference between the unknown spectral curves versus the Known Model Datasets. Eq. 3 is the equation that is used to perform the RMSE detection process.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(X_{obs,\ i} - X_{model,\ i})^2}{n}} \qquad (3)$$

Where n is the sequence number of each wavelength in the range from 530 to 590 nm, X_{obs} is observed values, and X_{model} is modelled values at time/place.

- *Confusion matrix*: After the RMSE calculation process, all RMSE values will be entered into the Confusion matrix to calculate the accuracy of the system. The equation to calculate the accuracy is described in Eq. 4.

$$Accuracy = \frac{TP+TN}{N} \qquad (4)$$

Where TP (True Positive) refers to sample belonging to the positive class being classified correcctly. TN (True Negative) refers to a sample belonging to the negative class being classified correctly. N is the number of all unknown SNV spectral curves. Accuracy is the index to evaluate the performance of the pesticide detection algorithm.

III. RESULTS

We performed two different tests: Positive and Negative Classification and Concentration Classification.

1) Positive and Negative Classification: Using binary confusion matrix.

In this test, we identified the unknown SNV spectral curves as Positive (leaf samples with water-only) and Negative (leaf samples with pesticide residue mixed 1:1000, 5:1000, and 10:1000). To select the threshold with highest accuracy, we chose threshold at 0.0271 for the accuracy of 90.25% (see Figure 6). In Figure 6, the x-axis represents thresholds and the y-axis represents accuracy. Figure 7 shows the confusion matrix at threshold 0.0271.

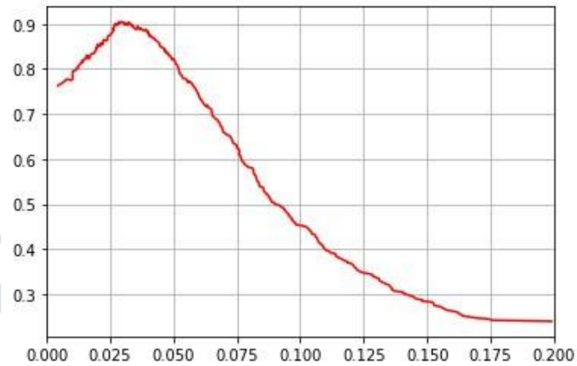

Figure 6: Accuracy of algorithm.

At Threshold 0.02718040

	Predicted Positive	Predicted Negative
Actual Positive	78	21
Actual Negative	18	283

Figure 7: Confusion matrix of threshold 0.0271.

2) Concentration Classification: Using multi-classes confusion matrix.

In this paper, the NIR spectral in range of 530 nm to 590 nm was used to detect the level of pesticide residues. There are 400 unknown leaf samples used to identidfication level of pesticide residue including 100 water-only leaf samples, 300 leaf samples had pesticide residue mixed with water at 1:1000, 5:1000, and 10:1000. Multi-classes confusion matrix was applied to calculate the accuracy of this system. Figure 8 shows the confusion matrix of level classification pesticide residues on butterhead lettuce leaves samples. Based on multi-classes confusion matrix, the accuracy of this system is 75.75%. Table 1 shows the remaining parameters, which include Precision, Recall, and F1-score parameters of this system.

ISSN (Online) 2395-2717
International Journal of Engineering Research in Electrical and Electronic Engineering (IJEREEE)
Volume 11 Issue 07 July 2024

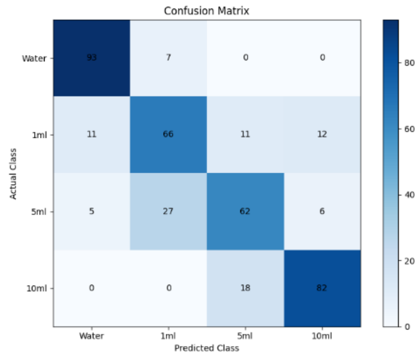

Figure 8: Confusion matrix of concentration classification.

Table 1: Precision, Recall, and F1-score parameters by using multi-class confusion matrix.

Group	Precision	Recall	F1-score
Water	0.8532	0.93	0.89
1ml	0.66	0.66	0.66
5ml	0.6813	0.62	0.6492
10ml	0.82	0.82	0.82

IV. CONCLUSION

Many studies have been based on the hyperspectral imaging technique to recognize the level of pesticide residues on vegetables. In this research, the program pesticide residue detection was developed by using Python code. The percentage level of butterhead lettuce leaf samples was accurately detected using the Standard Normal Validate method and Root Mean Square Error algorithm. Diluted mixtures of pesticide and water with three ratios of 1:1000, 5:100, and 10:1000 was used to spray on butterhead lettuce leaf samples. A total of 400 raw spectral curves were used to create Known Model datasets. After the Known Model datasets were created, the wavelength in the range of 530 nm to 590 nm were used for testing the detection algorithm. A total of 400 unknown raw spectral curves were used to determine positive or negative pesticide residues, the system's accuracy at the threshold of 0.0271 was 90.25%. In addition, a multi-classes confusion matrix was used to classify the concentration of pesticide residues. The accuracy of this system when used 400 unknown raw spectral curves is 75.75%.

Hence, the quantitative detection of mixed pesticide residues on lettuce leaves through hyperspectral techniques is viable, offering a methodological framework for future non-destructive detection of mixed pesticide residues on various other vegetables.

ACKNOWLEDGMENT

This research is funded by Mr. Ngo Duc Chi, from the 5D Agriculture Solutions Joint Stock Company.

REFERENCES

[1]. Fazlil Ilahi, W.F., Ahmad, D. & Husain, M.C. Effects of root zone cooling on butterhead lettuce grown in tropical conditions in a coir-perlite mixture. Hortic. Environ. Biotechnol. 58, 1–4 (2017).
[2]. Gosselin, R. E., Hodge, H. C., Smith, R. P., & Gleason, M. N. (1976). Clinical toxicology of commercial products (pp. 229-233).Baltimore: Williams & Wilkins.
[3]. Li, X.T., Wang, J.H., Zhu, D.Z., Pan, L.G., and Ma, Z.H.(2011). Research progress of fast detection methods of fruits and vegetables pesticide residues. Transactions of the Chinese Society of Agricultural Engineering, 27(s2), 363-370.
[4]. Feng, Y.Z., and Sun, D.W. (2012). Application of hyperspectral imaging in food safety inspection and control: a review. Critical Reviews in Food Science & Nutrition, 52(11), 1039-1058.
[5]. REN Zhan-qi, RAO Zhen-hong, JI Hai-yan, Identification of Different Concentrations Pesticide Residues of Dimethoate on Spinach Leaves by Hyperspectral Image Technology, IFAC-PapersOnLine, Volume 51, Issue 17, 2018, Pages 758-763.
[6]. Ye, W.; Yan, T.; Zhang, C.; Duan, L.; Chen, W.; Song, H.; Zhang, Y.; Xu, W.; Gao, P. Detection of Pesticide Residue Level in Grape Using Hyperspectral Imaging with Machine Learning. Foods 2022, 11, 1609.
[7]. Samadi, S Wajizah and A A Munawar 2018 IOP Conf. Ser.: Earth Environ. Sci. 122 012112. Fast and simultaneous prediction of animal feed nutritive values using near infrared reflectance spectroscopy.
[8]. Ronak Jain, Richa U Sharma, Mapping of Mineral Zones using the Spectral Feature Fitting Method in Jahazpur belt, Rajasthan, India, *Indian Institute of Remote Sensing, 4-*Kalidas Road, *Indian Institute of Remote Sensing, ISRO*,Dehradun – 248001, India,.
[9]. Ewing, J., Oommen, T., Jayakumar, P., & Alger, R. (2020). Utilizing hyperspectral remote sensing for soil gradation. Remote Sensing, 12(20), 1-13.