ESANN'2009 proceedings, European Symposium on Artificial Neural Networks - Advances in Computational Intelligence and Learning. Bruges (Belgium), 22-24 April 2009, d-side publi., ISBN 2-930307-09-9.

Comparison between linear discrimination analysis and support vector machine for detecting pesticide on spinach leaf by hyperspectral imaging with excitation-emission matrix

Mizuki Tsuta[1], Gamal El Masry[1,2], Takehiro Sugiyama[1,3], Kaori Fujita[1] and Junichi Sugiyama[1]

1- National Food Research Institute - National Agriculture and Food Research Organization
2-1-12 Kan-nondai, Tsukuba, Ibaraki 305-8642- JAPAN
2- Faculty of Agriculture - Agricultural Engineering Department
Suez Canal University, Ismailia P.O. Box 41522 - Egypt
3- Graduate School of Agricultural and Life Sciences – The University of Tokyo
1-1-1 Yayoi, Bunkyo, Tokyo 113-8657- JAPAN

Abstract. The performances of support vector machine (SVM) and linear discrimination analysis (LDA) for detecting pesticide on spinach leaf were investigated. Fluorescence images of spinach leaves without any treatment, treated with pure water and methamidophos solution were taken under 561 different wavelength conditions to acquire hyperspectral excitation-emission matrix (EEM) data. Then LDA and SVM were applied to EEMs of pixels randomly sampled from the data for classification of treatment. Misclassification rates were 18.8% and 9.9% for LDA and SVM respectively. Also, the obtained results revealed that when SVM applied pixel-wise of hyperspectral data, methamidophos treated leaves could be distinguished visibly from the others.

1 Introduction

In Japan, violations of the restriction level for pesticide residue of spinach have been reported [1]. Recently, hundreds of people were sickened by frozen dumplings imported from China contaminated by methamidophos, which is a prohibited pesticide in Japan [2]. Consumers in Japan are now worried about food contamination by pesticide and therefore effective pesticide detection methods are in urgent demand.

Excitation-emission matrix (EEM) has been applied to the measurement of chemical compounds such as pollutants in sea water [3] and pesticide [4]. Tsuta et al. [5] have reported visualization of internal structure of soybean seeds by hyperspectral imaging with EEM (HSI-EEM). This method has the advantage that no pretreatments such as staining are necessary and the location of the target compounds can be visualized.

Support vector machine (SVM) has been reported to show high accuracy in classification problems when compared to conventional methods such as linear discrimination analysis (LDA) [6-8]. Accordingly a rapid and simple pesticide

detection technique, with which the location of pesticide in a sample is visualized, can be achieved by combining HSI-EEM and these classification methods.

The objective of this research was to develop a visualization method for methamidophos residue on spinach leaf by HSI-EEM with classification methods. The performances of SVM and LDA for hyperspectral EEM data were also investigated.

2 Materials and Methods

2.1 Pesticide Solution

A standard solution of 500 ppm was prepared for spinach treatment by dissolving 10 mg methamidophos standard (Wako Pure Chemical Industries, Ltd., Japan) in 100 ml pure deionized water allowing the solution to sit for 30 min before usage.

2.2 Sample Preparation

Sixteen leaves of spinach were cut into 20-mm-square samples. 40 μl of pure water was dropped on the surface of four leaves. In the same way, other four leaves were treated with the methamidophos solution. Eight leaves left were kept untreated as control samples. The treated leaves were kept in the room temperature for 20 min for the treating solutions to evaporate or be absorbed inside the cellular structure of the samples.

2.3 Imaging Apparatus

To acquire an EEM at any point of a sample, a measurement system named EEM Imaging System (EEM-IS) was developed, as shown in figure 1. EEM-IS is composed of a spectral illuminator and spectral imaging system.

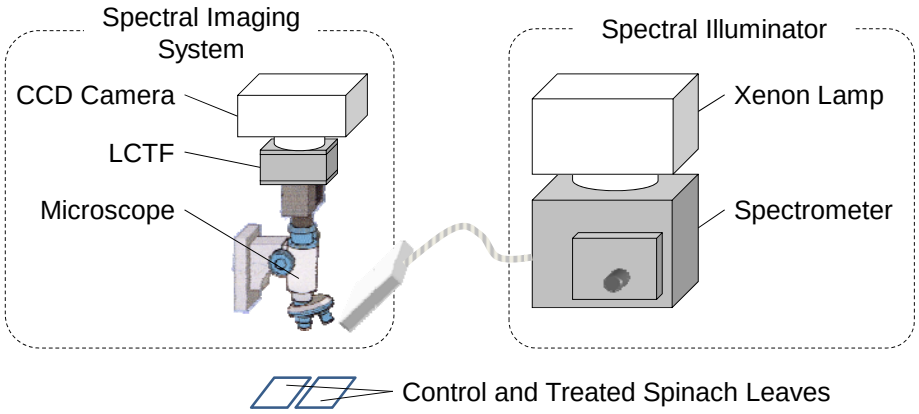

Fig. 1: Schematic Image of EEM-IS.

The spectral illuminator consists of a xenon lamp and grating spectrometer, which were both custom-made by Soma Optics, Inc., Tokyo, Japan. It illuminates the

surface of the sample at a specific wavelength from 200 to 1000 nm with a wavelength error of ±1 nm and a band width of 10 nm. The spectral imaging system is composed of a fluorescence microscope (BXFM, Olympus, Japan), a liquid crystal tunable filter (LCTF; VS-VIS2-10-MC-35, Cambridge Research and Instrumentation, Inc., USA), and a monochrome CCD camera (ORCA-ER-1394, Hamamatsu Photonics K.K., Japan). A 2× objective lens was attached to the microscope and the field of view was 9.47 × 7.22 mm. The wavelength of the light transmitted through the LCTF can be selected from 400 to 720 nm by changing the voltage applied to the LCTF, with a wavelength error of 1.25 nm and a bandwidth of 20 nm. The CCD camera has a 1344 × 1024 pixel resolution (1.3 megapixels) and its gradation is 12 bits (4096 steps).

When the spectral images of the sample surface are taken using the EEM-IS, with the wavelength of illumination by the spectral illuminator fixed and the transmission wavelengths of the LCTF continuously changed, an emission spectrum of each pixel can be acquired. Then, by acquiring emission spectra at different illumination wavelengths, an EEM of each pixel can be measured.

2.4 **Acquisition of Hyperspectral EEM**

Two leaves, one control and the other, water-treated, were set in EEM-IS so that both of them could be observed in the same view. Their fluorescence images were taken at various excitation and emission wavelengths. First, the excitation wavelength of the spectral illuminator was fixed and the fluorescence images of the sample were taken while the emission wavelength of the LCTF changed successively, starting at the wavelength 50 nm larger than the fixed excitation wavelength and ending at 720 nm at 10 nm intervals. Next, the excitation wavelength was set to a value 10 nm larger than the previous one, and the imaging of the sample with emission wavelength scanning was carried out. These processes were repeated in the excitation wavelength range from 350 to 670 nm and the emission wavelength range from 400 to 720 nm at 10 nm intervals, and 561 images were acquired, as shown in figure 2a. The exposure period for the imaging was set to 1.0s and the binning mode of 8 × 8 was applied (that is, 64 pixels of the CCD camera were combined to function as one pixel and gain 64 times higher sensitivity). The size of the images was 168 × 128 pixels. A similar process was applied to four pairs of control and water-treated leaves and four other pairs of control and methamidophos-treated leaves, so that eight hyperspectral EEM data were acquired.

2.5 **Data Sampling**

A region of interest (ROI) surrounding each leaf was manually set in the hyperspectral EEM data as shown in figure 2b. 10% of pixels within ROIs were selected randomly and their EEM was sampled for analysis. The total number of selected pixels was 13,391. These pixels were also assigned specific index according to the treatment (control=0, water-treated=1 and methamidophos treated=2) so that a data set of 13,391 by 562 (561 wavelength conditions plus index) matrix was composed as shown in figure 2c. Then the data set were divided into halves to form calibration and validation sets as shown in figure 2d.

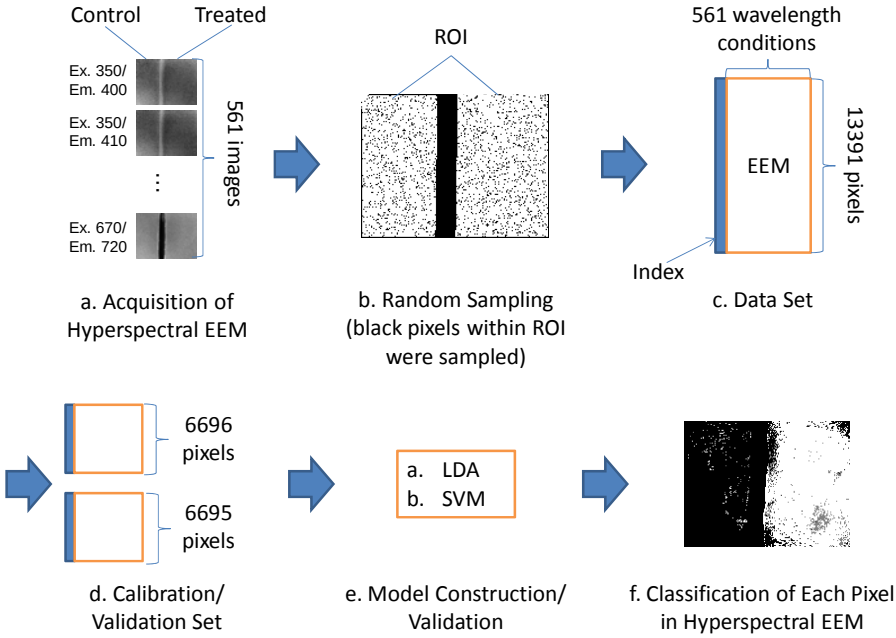

Fig. 2: Experimental Procedure.

2.6 Model Construction and Validation

As shown in figure 2e, LDA was applied to the calibration set to develop a discrimination model. Then the model applied to the validation set for the prediction of the index of each pixel and the misclassification ratio was calculated. In the same way, SVM was applied to the calibration set using radial basis kernel with the gamma value of 1/ 561 and the cost value of 1. The acquired model was applied to the validation set. Then the misclassification ratio of SVM was calculated. Numerical analysis software (Matlab R2007b, The MathWorks, Inc., USA) and statistical software (R 2.0.8, R Development Core Team, Austria) with SVM package (e1071, E. Dimitriadou et al, Austria) were utilized for LDA and SVM, respectively.

2.7 Model Application to Hyperspectral EEM Data

The LDA model was applied to EEM of each pixel in the hyperspectral EEM data for the classification of the treatment. Then each pixel was assigned a specific color according to the predicted index; that is, 0=black, 1=gray and 2=white. As a result, a classification image, in which location of the control and treated leaves could be visualized, was developed as shown in figure 2f. In the same way, the classification image based on SVM was acquired. On the assumption that every part in a single leaf had same truth label (i.e. control, water or methamidophos), misclassification ratio on the classification images was also calculated by dividing the number of misclassified pixels by that of the pixels of all leaves in the images.

3 Results and Discussion

As shown in table 1, misclassification ratio of SVM for the classification of validation set was 0.099, while that of LDA was 0.188. As for the classification images, there were only a few misclassified pixels in the image based on SVM and the methamidophos treated leaves could be clearly distinguished from others as shown in figure 3. The pixels of the gap between two leaves were classified as control and seemed to have little influence on the results. It seemed that the image based on SVM reflected the actual treatment more accurately than LDA based one. Misclassification ratio of SVM on the images also showed a lower value than that of LDA as shown in table 2. These results indicated that SVM was superior to LDA in detecting pesticide using hyperspectral EEM data.

		Predicted					
		LDA			SVM		
		Control	Water	Methamidophos	Control	Water	Methamidophos
Actual	Control	2692	191	157	2927	163	131
	Water	125	1416	335	43	1576	117
	Methamidophos	122	328	1329	70	137	1531
			LDA error:	0.188		SVM error:	0.099

Table 1:Confusion matrix of Actual and Predicted Treatment of Pixels in Validation Set with LDA (left) and SVM (right).

		Predicted					
		LDA			SVM		
		Control	Water	Methamidophos	Control	Water	Methamidophos
Actual	Control	15035	874	536	16093	184	168
	Water	1042	8533	1615	1214	9474	502
	Methamidophos	714	1475	8681	862	628	9380
			LDA error:	0.162		SVM error:	0.092

Table 2:Confusion matrix of Actual and Predicted Treatment of Pixels in Hyperspectral EEM Data with LDA (left) and SVM (right).

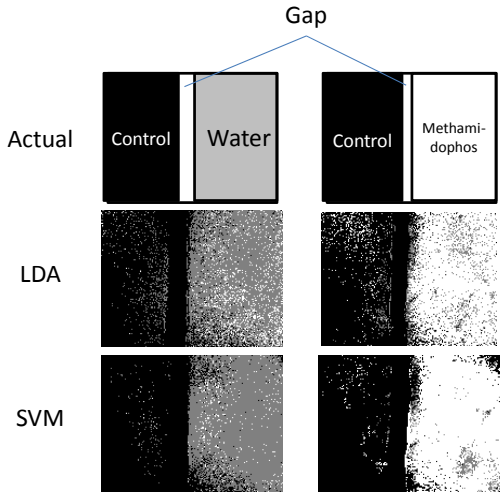

Fig. 3:Classification of Each Pixel in Hyperspectral EEM Data.

4 Conclusion

Application of SVM to hyperspectral EEM data was effective in detection of methamidophos on spinach leaves. In the pixel-wise validation, misclassification ratio for SVM was 0.099 while that for LDA was 0.188. As for the classification of each pixel in images, SVM also showed better results than LDA as the misclassification ratios for SVM and LDA were 0.092 and 0.162, respectively. It was also found that methamidophos treated leaves could be distinguished visibly from others by their color after SVM was applied to hyperspectral EEM data. Therefore, it can be concluded that SVM was superior to LDA and effective in detecting methamidophos using hyperspectral EEM data. Further research is needed for the detection of less concentrated pesticides other than methamidophos and improvement of the measurement speed for online applications.

5 References

[1] H. Satoh, E. Amakawa, H. Kondo, Y. Aoyagi, M. Miyakoda, Y. Yamada, T. Ogiwara and K. Yasuda, Survey of pesticides residues in frozen vegetables imported from China, *Ann. Rep. Tokyo Metr. Inst. P.H.*, 54: 201-203, Tokyo Metropolitan Institute of Public Health, 2003.

[2] M. Fackler, Insecticide-tainted dumplings from China sicken 175 in Japan, in *The New York Times*, The New York Times Company, 2008.

[3] K. S. Booksh, A. R. Muroski and M. L. Myrick, Single-measurement excitation/emission matrix spectrofluorometer for determination of hydrocarbons in ocean water. 2. Calibration and quantitation of naphthalene and styrene, *Anal. Chem.*, 68:3539-3544, American Chemical Society, 1996.

[4] R. D. JiJi, G. A. Cooper and K. S. Booksh, Excitation-emission matrix fluorescence based determination of carbamate pesticides and polycyclic aromatic hydrocarbons, *Analytica Chimica Acta*, 397:61-72, Elsevier, 1999.

[5] M. Tsuta, K. Miyashita, T. Suzuki, S. Nakauchi, Y. Sagara and J. Sugiyama, Three-dimensional visualization of internal structural changes in soybean seeds during germination by excitation-emission matrix imaging, *Transactions of the ASABE*, 50:2127-2136, American Society of Agricultural and Biological Engineers, 2007.

[6] F. Luan, R. S. Zhang, C. Y. Zhao, X. J. Yao, M. C. Liu, Z. D. Hu and B. T. Fan, Classification of the carcinogenicity of n-nitroso compounds based on support vector machines and linear discriminant analysis, *Chem. Res. Toxicol.*, 18:198-203, American Chemical Society, 2005.

[7] R. Judson, F. Elloumi, R. W. Setzer, Z. Li and I. Shah, A comparison of machine learning algorithms for chemical toxicity classification using a simulated multi-scale data model, *BMC Bioinformatics*, 9:241, BioMed Central Ltd., 2008.

[8] M. A. Oskoei and H. S. Hu, Support vector machine-based classification scheme for myoelectric control applied to upper limb, *IEEE Trans. Bio-med. Eng.*, 55:1956-1965, IEEE, 2008.