International Journal of Artificial Intelligence and Neural Networks – IJAINN

Volume 9 : Issue 1- [ISSN : 2250-3749] - Publication Date: 27 Dec, 2019

Red defect detection for rice quality assurance by using Machine Learning

Phuvin Kongsawat, Nawapat Jamroenrak, Tuanjai Archevapanich, Boonchana Purahong and Attasit Lasakul
Department of Information Engineering, Faculty of Engineering, King Mongkut's Institute of Technology Ladkrabang, Bangkok 10520, Thailand
Faculty of Engineering and Architecture, Rajamangala University of Technology Suvarnabhumi, Thailand

This article presents an inspection system to detect red kernel defect, normally contaminating in white rice product. This contamination causes a reduction in the price rice of 6% approximately. To detect Red defect successfully, a method proposed in this paper was build up on Machine Vision techniques. The method contains three processing steps as follows. Firstly, noise elimination and localization were executed through image processing techniques. After that, RGB image would be transformed to HSV in order to obtain discriminative features. Finally, the pre-processed data was then passed into model training by using both linear and non-linear Support Vector Machines. Apart from that, Logistic regression was then employed to challenge margin maximization ability of the SVMs. The experimental result shows that linear-SVM still yields the highest performance at 86.3% of classification accuracy.

Keywords: Thailand standard for rice, Machine Vision, HSV, Support Vector Machine.

I. Introduction

Thailand has the second most export volume of rice into the global Rice market. The success factor is that there is the Rice standard declared by the ministry of commerce Thailand call [1] "Thailand standard for rice". On the other hand, the existing method being utilized to evaluate rice quality is human judgment which cause time-consuming and mistakes. To solve these problems, machine vision are crucial and required.

In the science techniques there are many methods of machine vision to inspect and assess rice quality which have developed in a few years ago for example [2] the machine vision was used to assess rice quality by using feature extraction which select Major axis, Minor axis, Area and eccentricity then put features to Neural Network PNN Probabilistic Neural Network model in order to train and predict rice quality, [3] presents an assessment of grain quality emphasizing on broken rice, head rice, small broken rice, and large-sized rice discarded in rice production by Least-Square Support Vector Machine (LS-SVM) and Radius Basis Function (RBF) with the accuracy of 98.20%, [4] used machine vision to measure sizing of rice. The article provides accurate result than existing methods both Scion and Feret diameter, [5] apply machine vision and image processing for Chalky kernel detection. The article extracts geometrical features then put the features to train and predict by using Back-propagation Neural Network which come up with 90% of accuracy, [6] utilize Support Vector Machine to classify rice grade (Premium, Grade A, Grade B and Grade C), [7] the machine vision applied to classify cereal grains (barley, rye, oats and wheat) the article extract color features which achieved a classification accuracy of 98.5% for barley, 99.97% for CWRS, 99.93% for oat, and 100% for rye. [8] focused on classify foreign matters, type admixture and brown grain content by using Back-propagation Neural Network. Tests on the system for the training and test sets show accuracy in between 94% to 68% for the four grades. The institute classifies rice into four quality categories according to several parameters.

This article presents the Machine Vision method to detect Red defect Fig 1 which contaminate in white rice Fig 2. The method contains three processing steps as follows. Firstly, noise elimination and localization were executed through image processing techniques. After that, RGB image would be transformed to HSV in order to obtain discriminative features. Then, the pre-processed data was then passed into model training phase. In this phase aim to find appropriate both the amount of data training and classified model. Hence, comparing of accurate result in case of Linear Support Vector Machine and Radius Basis Function Support Vector Machine in order to find separable pattern is either linear or non-linear.

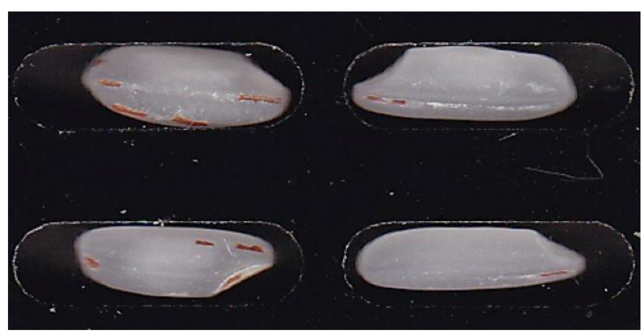

Figure 1. Red Defect Kernels

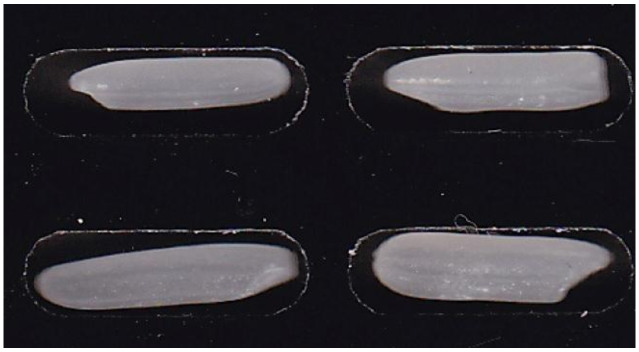

Figure 2. White Kernels

II. Methodology

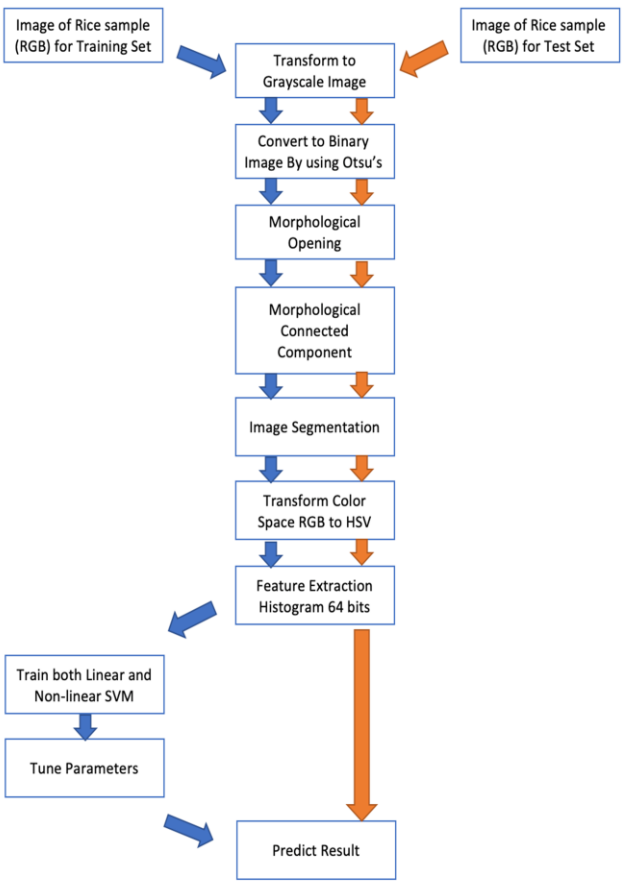

Figure 3. Dataflow diagram for the proposed approach

A. Equipment Preparation

The first step in machine vision acquire properly and certainly input image in order to gain accurate results. Hence, flatbed scanner is required. This article used Canon LIDE 120 flatbed scanner by selecting resolution 600 dpi. The scanner is independence from external light and provide best quality image which appropriate for classification task. The uniform background which is black in color. The rice kernels are spread on a scanner glass randomly. Although the rice kernels were putted into the scanner randomly, and then make sure that they are not overlap and joint each other. The images were captured and stored in JPEG format automatically.

B. Preprocessing

This procedure focus on prepare proper data to train the machine learning. The procedure as follows.

- Transform RGB color space transform to grayscale image Fig 1.
- Transform grayscale image to binary image by Otsu's method [9]
- Morphological opening [10] method was utilized for noise eliminating.
- Morphological labeling component [10] was used to find each position of rice kernel in image.
- Cropping RGB image of each rice kernel from their position was delivered from previous morphological method rice.
- Transform RGB image to HSV image
- Transform HSV to histogram 64 bits

C. *Feature Extraction*

There are various features in image such as color, geometrical and etc. However, we selected features from 3 histogram intensity 64 bits including Hue: H, Saturation: S and Value: V. And then normalization by divide each bit by their summation of each histogram

D. *Support Vector Machine Preliminary*

A Support Vector Machine: SVM is a discriminative classifier. It was defined by decision boundary as a separating hyperplane. In other words, given labeled training data (Supervised Learning) and Non-parametric model. To optimize the decision boundary with maximum margin of separation between 2 class which made from the support vectors Fig 4. The decision boundary is perpendicular with vector w. It implies that $W^TX^{+}+b= +1$ and $W^TX^{-}+b= -1$. X^{+} denote a positive point with functional margin of 1. X^{-} denote a negative point respectively. b denotes the distance of projection between point and W vector.

The functional margin of resulting classifier M follow equation (1) which aim to maximize margin.

$$M = \frac{2}{\|W\|} \tag{1}$$

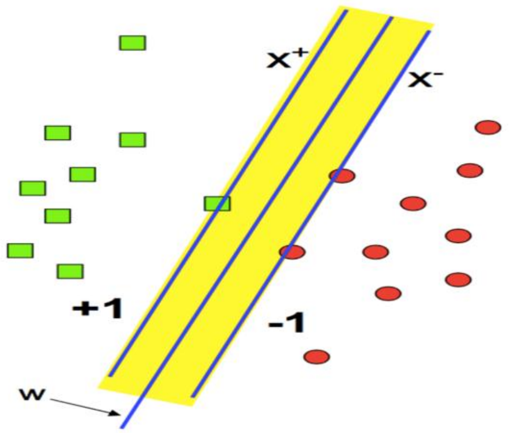

Figure 6. Decision Boundary

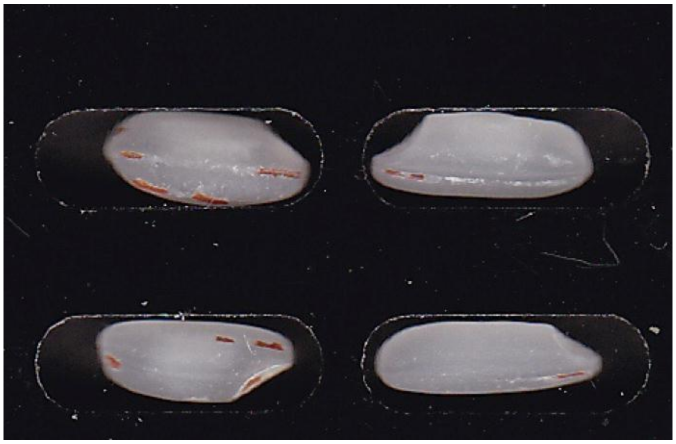

Figure 5. Binary Image

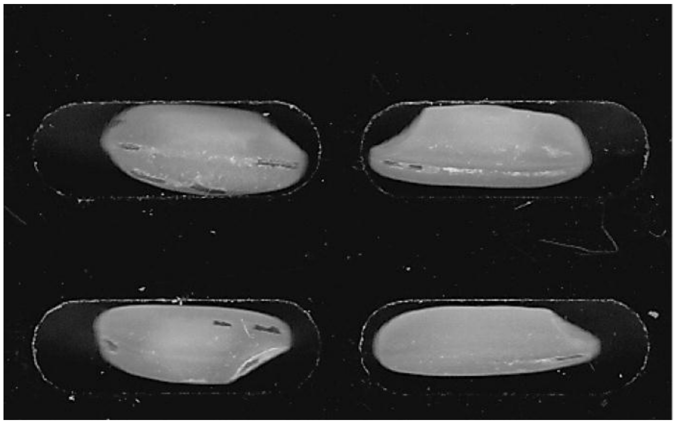

Figure 6. Binary Image

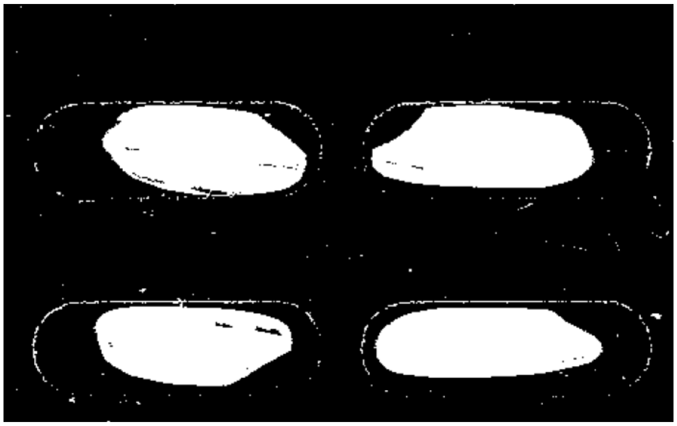

Figure 7. Binary Image

Then, we aim to avoid overfitting. So, we add the slack variable to reduce variance but increase bias as equation (2) which is L2 regularization.

$$\text{Minimize } \frac{\|W\|^2}{2} + C\sum_{i=1}^{n}\xi_i^2 \qquad (2)$$

$$\text{s.t. } y_i(W^T x_i + b) \geq 1-\xi_i,\ i = 1,2,...n$$

Equation (2) is modified into the new equation (3) by its unrestraint dual form

$$\text{Maximize } \sum_{i=1}^{n}\alpha_i - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=}^{n}\alpha_i\alpha_j y_i y_j K(x_i,x_j) \qquad (3)$$

$$s.t.\ 0 \leq \alpha_i \leq C,\ i = 1,2,...n,\ j = 1,2,...n\ \sum_{i=1}^{n}\alpha_i y_i = 0$$

This article used non-linear of kernel SVM as Radial Basis Function: RBF in equation (4)

$$K(x_i,x_j) = e^{\frac{\|x_i - x_j\|^2}{G}} \qquad (4)$$

III. Experiments

A. *Rice Sample*

Rice samples were utilized in this article which were collected from Rice inspector Board of Trade of Thailand. This organization is government agency. They provide the solutions for rice exporters such as rice quality inspection and export permission.

In order to gain proper Rice data to train machine. We selected expert of rice inspectors who have experience at least 25 years for data collection. The rice sample consist both white rice 1250 kernels and red defect rice 1250 kernels

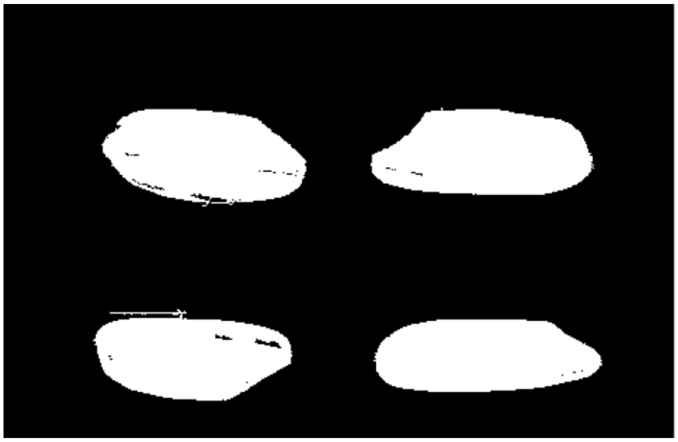

Figure 8. Noise Eliminating Image

B. *Preprocessing*

RGB image Fig.5 was transformed to grayscale image Fig. 6 after that transform grayscale image to binary image Fig. 7 by Otsu's method. Next, the morphological opening was utilized to eliminate speckle noise Fig. 8. Then, morphological labeling was applied to find each rice position in order to localize rice kernels. Each rice image is converted from RGB image to HSV in order to obtain discriminative feature. And then, Gaussian filter was used to this image. The image is smoothed for noise reduction and improving edge detection performance.

C. *Feature Extraction*

we selected features from 3 histogram intensity 64 bits including Hue: H, Saturation: S and Value: V. And then normalization by divide each bit by their summation of each histogram

D. *Pattern of Data Distribution*

we created the experiment to understand pattern of data distribution by using both Linear and Radius Basis Function SVM. The accurate result will reply the pattern of distribution. If the distribution was linear, Linear SVM will provide precise result than RBF SVM. On the other hand, Radius Basis Function SVM will provide accurate result than Liner SVM in case of non-linear distribution

IV. Experimental Results

A. *Comparison of rice selling price*

According to Thai Rice Exporters Association, the price depends on rice quality which was appraise by three factors (broken rice, size of rice, defects). This article focuses on defect factor. By selecting crucial selling factor is red defect. This factor knocks off 6% of price when the rice product was sold. Therefore, this factor is crucial to considerate in order to gain appropriated profit.

B. *Classifier performance*

Firstly, we divide the rice data into three groups consisting of training set with 1190 kernels (595 Red kernels and 595 of White kernels), test set containing 511 kernels (263 Red kernels and 248 of White kernels), blinded test set with 600 kernels (300 of Red kernels and 300 of White kernels). After that, we applied Linear SVM to train red-kernel classifiers throughout the training set. Next, we tuned the value of C parameter by varying in a list which includes 1, 10, 100, 1000, 10000 and 100000 to obtain the best accurate result. The tuning process shows that the C at 1000 is suitable for this task. As a result, the model performs on the test set with 96.67 of accuracy. Then, we utilized the training set to train Radial-Basis-Function SVM. We varied C and G parameters to gain the best result that the C values are varied in {1, 10, 100, 1000, 10000, 100000, 1000000} and G value varied in {0.1, 0.01, 0.0001, 0.00001, 0.000001, 0.0000001}. According to the experimental result, we found that the values of C and G appropriating with this task are 100000 and 0.0001 respectively. The model with these optimal parameter values illustrates 95.6% accuracy when it was evaluated with test set.

Then, we utilized the blind test set to evaluate the performance of model. Linear SVM perform 86.37 % of accurate result which accurate than Radius Basis Function SVM perform 84.38%.

TABLE I. PRICE OF WHITE RICE

	Prices of white rice		
	Type	***At least Percent of contamination***	***Price USD***
1	White rice 100%	0.5 %	446
2	White rice 25%	7 %	420

TABLE II. CONFUSION MATRIX OF LINEAR SVM (TEST SET)

Red Defect Classification		**Actual**	
		Red kernel	White kernel
Prediction	Red Kernel	256 (TP)	9 (FP)
	White kernel	8 (FN)	238 (TN)

TABLE III. CONFUSION MATRIX OF RADIUS BASIS FUNCTION SVM (TEST SET)

Red Defect Classification		**Actual**	
		Red kernel	White kernel
Prediction	Red Kernel	254 (TP)	11 (FP)
	White kernel	9 (FN)	237 (TN)

TABLE IV. CONFUSION MATRIX OF LINEAR SVM (UNSEEN SET)

Red Defect Classification		**Actual**	
		Red kernel	White kernel
Prediction	Red Kernel	289 (TP)	12 (FP)

Red Defect Classification		Actual	
		Red kernel	White kernel
	White kernel	70 (FN)	231 (TN)

TABLE V. CONFUSION MATRIX OF RADIUS BASIS FUNCTION SVM (UNSEEN SET)

Red Defect Classification		Actual	
		Red kernel	White kernel
Prediction	Red Kernel	254 (TP)	11 (FP)
	White kernel	9 (FN)	237 (TN)

V. Performance Improvement

From the experimental result, Linear SVM performed better than Radius Basis Function. So, it can be concluded that the pattern of data distribution is linear. Therefore, we improved the result by using a parametric model for linear-separable data as Logistic Regression. Then, we utilized the training set to train machine learning model. After that, the testing set was passed to the model for performance assessment. which resulted in 95.7% of accuracy. Moreover, to testify model capability in facing other unseen samples.

TABLE VI. CONFUSION MATRIX OF LOGISTIC CLASSIFICATION

Red Defect Classification		Actual	
		Red kernel	White kernel
Prediction	Red Kernel	224 (TP)	76 (FP)
	White kernel	8 (FN)	292 (TN)

The blinded test set were used to evaluate the realistic performance, which is 86% of accuracy. Hence, we can summarize that first data set tends not to cover the whole of data distribution for this task. To conclude, the classification performance cannot be improved unless the more samples are collected.

VI. Conclusions

In this study, we aim to detect red defect which contaminate in white rice product. This contamination knocks off 6 % of price. The experiment was executed by using both linear and non-linear SVM. The experimental result show that the linear SVM suitable with this task than non-linear. The Linear SVM provide 86.37 % of accuracy. From the result, we can conclude that the pattern of data distribution is linear. After that, we employed the non-parametric model as Logistic Regression. It provides 86% of accuracy. It is able to summarize that first data set tends not to cover the whole of data distribution. To unleash the better performance the more samples are collected

Acknowledgment

Authors appreciate to thank Rice Inspector Department Board of Trade of Thailand. They supported knowledge and rice samples in order to improve the technology and move forward rice industry.

References

[1] Thai Standard of Rice , Ministry of Commerce Thailand, 2016.

[2] V.S.Kolkure, B.N.Shaikh, "Identification And Quality Testing Of Rice
Grains Using Image Processing and Neural network", International Journal of Recent Trends in Engineering & Research (IJRTER), 2017.

[3] Dollawat Ngampak and Punpiti Piamsa-nga, "Image Analysis of Broken Rice Grains of Khao Dawk Mali Rice", Knowledge and Smart Technology (KST), 2015.

[4] P. Kongsawat, S. Chivapreecha and T. Sato, " Quality Assessment of Thai Rice Kernels Using Low Cost Digital Image Processing System", International Workshop on Advanced Image Technology (IWAIT), 2018.

[5] Bhupinder Verma, "Image Processing Techniques for Grading & Classification of Rice," Computer and Communication Technology (ICCCT), 2010

[6] Harpreet Kaur, Baljit Singh, "Classification and Grading Rice Using Multi-Class SVM", International Journal of Scientific and Research Publications, Volume 3, Issue 4, April 2013.

[7] H.K. Mebatsion, J. Paliwal, D.S. Jayas, "Automatic classification of non-touching cereal grains in digital images using limited morphological and color features" Elsevier ,Computers and Electronics in Agriculture 90,2013.

[8] L.A.I.Pabamalie, H.L.Premaratne, "A Grain Quality Classification System", Institute of Electrical and Electronics Engineers, 2010.

[9] Vala, H.J., " A review on Otsu image segmentation algorithm", Int. J. Adv. Res. Comput. Eng. Technol. (IJARCET) **2**(2), 387–389 , 2013.

[10] Gonzalez, Eddins and Wood, Digital Image Processing Using Matlab (2nd Edition), GATESMARK. 2009.

About Author (s):

Mr. Phuvin Kongsawat. As a master student of Information Engineering, Faculty of Engineering, King Mongkut's Institute of Technology Ladkrabang. CEO of Tech Startup (Easy Rice Digital Technology) in Thailand. Interesting about Machine Learning, Signal Processing, Management and Supply Chain

Mr. Nawapat Jamroenrak. As a r student of Information Engineering, Fa of Engineering, King Mongkut's Institu Technology Ladkrabang. COO of Startup (Easy Rice Digital Techno Interesting about Machine Lea Robotics, Mechanical engineering Embedded System.

Asst.Prof.Dr. Tuanjai Archevapanich. As a Lecturer of Faculty of Engineering and Architecture, Rajamangala University of Technology Suvarnabhumi. Interesting about Antenna and Engineering Education.

Asst.Prof. Boonchana Purahong. As a Lecturer of Information Engineering, Faculty of Engineering, King Mongkut's Institute of Technology Ladkrabang. Interesting about Machine Learning, Robotics, Image Processing and Embedded System.

Assco.Prof.Dr. Attasit Lasakul. As a Lecturer of Information Engineering, Faculty of Engineering, King Mongkut's Institute of Technology Ladkrabang. Interesting about Machine Learning, Signal and Image Processing and Microprocessor