PALM OIL BRUISE DETECTION OF TEXTURE AND SHAPE FEATURES EXPERIMENTAL COMPARISON ON SUPPORT VECTOR MACHINE AND NAÏVE BAYES

Nurbaity Sabri[1*], Anis Amilah Shari[1], Mohd Rahmat Mohd Noordin[1], Nur Nabilah Abu Mangshor[1], Noor Suriana Abu Bakar[1]

[1] *Faculty of Computer and Mathematical Sciences*
Universiti Teknologi Mara UiTM Melaka, Jasin Campus, 77300 Merlimau, Melaka

[*]*Corresponding author: nurbaity_sabri@uitm.edu.my*

Abstract

Palm oil is one of the largest and significant contributions to the Malaysia economy. It is important to improve the quality of this product as defects on palm oil fruit may affect the production of palm oil. Bruise is one of the defects on palm oil fruit where it is unavoidable during the field material activities. It will increase the number of Free Fatty Acid (FFA) and reduce number of palm oil quality. We proposed using a combination of four (4) textures Grey Level Co-occurrence Matrix (GLCM) and six (6) shape features to detect bruise and non-bruise. A comparison between two classifiers named Support Vector Machine (SVM) and Naïve Bayes has been done using the same features. The experiment shows Naïve Bayes classifier achieve 97.5% accuracy compared to SVM with the combination of two types of features. Further study will be done to classify the palm oil bruise into more stages.

Keyword: Palm Oil bruise, Grey Level Co-occurrence Matrix, shape feature, Support Vector Machine, Naïve Bayes.

Introduction

Palm oil is one of the contributors to Malaysia economy where it is one of profitable agriculture product. Many researches have been done to increase and maintain the quality of palm oil (Makky, 2016) which can be determined based on the surface color of palm oil. MPOB (Malaysia Palm Oil Board) has listed three ripeness stage of palm oil grading which are unripe, ripe and overripe. Each stage has their specific colors that differentiate between one another. However, these three categories (ripe, unripe, and overripe) could be easily damage and bruised (Lovely, Mahmud, & Mohamed Halim, 2015). There are two types of oil produced by palm oil fruit which is palm oil and palm oil kernel. The palm oil obtains from mesocorp covered the palm oil kernel. While the palm oil kernel produced by nut or seed of palm oil fruit (Kabutey, et al., 2018). The bruise usually happens on the mesocorp layer where this layer produces a Free Fatty Acid (FFA).

A correlation between FFA and bruise has been done by (Afroza et al., 2015). This research stated that the FFA increase when there are major bruises on the palm oil fruit and it will affect the quality of palm oil (Afroza et al., 2015). FFA derived from the hydrolytic rancidity of triglycerides. It is an acid content of edible fats where it affects the palm oil quality. The higher FFA in palm oil fruit, the lower quality of palm oil produced (Hadi et al., 2009). According to Krisdiarto & Sutiarso (2016), a study of palm oil fruit bruise during the field material handling (harvesting and transportation) and correlation between the FFA content has been done. It also stated that it is important to learn about bruise and cause of bruise on FFB fruit to increase quality of palm oil (Krisdiarto & Sutiarso, 2016).

Current manual identification of bruise has been done by the experts. However, it is still lack of accuracy and time consuming (Dubey et al., 2013). The bruise damage seems to be almost unavoidable during harvesting and it eventually decreases the fruit quality. Thus, there is a need to improve the quality of agricultural products. To achieve this objective, various image analysis techniques is needed due to the demand of the growing population (Ashok & Vinod, 2014). Currently, there are many researches on fruit bruise detection using computer vision such as apple, kiwi, citrus, pear, tomato and many more (Ashok & Vinod, 2014; Qiang, & Mingjie, 2012; Dang et al., 2012). NIR (Near Infrared imaging) technique proposed to detect a bruise on pear fruit (Lee et al., 2014). This research using hyper spectral imaging (waveband) and implement the F-value algorithm for classification. It manages to achieve 92% accuracy on bruise detection. However, this research needs a high specification device which is InGaAS Focal Plane Array (FPA) camera as an image acquisition tools.

Identification of a defect on apples surface using machine vision has been done (Vijayarekha, 2008). Multispectral imaging technique implemented using score and image space as features extraction. A clustering technique has been used to cluster the defect area of apple. However, the bruise image detects incorrectly due to the glossy surface of apples. Another research on apple has been carried out using k-means clustering (Khade et al., 2016). This technique manages to improve the defect segmentation quality precision of apple and its computational time. This system recommends recognition capability for future work. Besides, this research also proposed a combination of texture features with other features to improve accuracy of the proposed system. Disease detection on tomato using the same method (k-means clustering algorithm) has also been proposed (James & Punitha, 2016). This research successfully identifies disease accurately with 90% accuracy achieved. However, more technique is recommended to obtain more accurate result.

To feed a data into these classifiers, a feature needs to be choose based on the image characteristic. A combination of texture and shape features has been done using Multi linear perceptron to recognize plant leaf which can produces a convincing result where it has been as effective method on leaves recognition (Chaki et al., 2015). A combination of texture and shape features has also been implemented to identify breast cancer. This research achieves 92% of accuracy using Backpropagate neural network classifier. It also shows that using combination of texture and shape features increase the speed of neural network learning capability (Helwan & Abiyev, 2016).

Therefore, this research will investigate the implementation of palm oil bruise detection using image processing technique. Combination of texture and shape features is proposed and experimental comparison between Support Vector Machine (SVM) and Naive Bayes to classify the features will be done.

Materials and Methods

Image processing has been used in many areas of research such as in food industry, medical processing, textiles, engineering and many more (Jamil, & Awang, 2014). There are four component of image processing used in this research which is image acquisition, segmentation, features extraction and classification. These elements are important to produce a desired result (Shih, 2010). This research focuses on two classes which are bruise and no bruise.

Image Acquisition

This research using 120 images of four stages palm oil bruise which bruise (53 images) and no bruise (67 images). Image collected from Sime Darby oil palm mills at Jasin, Melaka with the help of palm oil fruit expert using mobile phone.

Image Segmentation

The images of a palm oil surface are clustered according to their color. K-means clustering is implements due to its fast and straightforward algorithm. The process starts with the selection of initial centroids. These points are assigned according to the closest cluster. Next, the new centroid is computed for each cluster. This process will continue until all the colors are clustered. The palm oil bruise is segmented using k-means clustering where the surfaces of palm oil fruit (non-bruise region) are removed.

Shape Features

Shape features used to extract the boundary and region descriptor of the bruise region. An external method has been implemented to extract the boundary of palm oil bruise. Meanwhile, region descriptor extracts the pixel of internal bruise region (Mohana, & Prabhakar, 2014) where the shape features using six shape metrics which are area, perimeter, major axis length, minor axis length, eccentricity and equidiameter to extract the features date has been used. The same technique has been implemented to extract the features of palm oil bruise.

Area (A)

The area of segmented pixel of palm oil bruise is calculated using Eq. 1 where its shows the sum of *b (i,j)* represent the total number of pixel in *i* row and *j* column. Meanwhile, variable *M, N* represents row and column of the image.

$$Area = \sum_{i,j=0}^{M,N} b(i,j) \tag{1}$$

Perimeter (P)

The pixel boundary of palm oil bruise is calculated where the distance between each adjoining pair of pixels is accumulated. In Eq. 2, the *Ed (i,j)* is the total of boundary pixels of the region.

$$Perimeter = \sum_{i,j=0}^{M,N} E_d(i,j) \tag{2}$$

Major-axis Length

Major-axis length is the length of the major axis of the region of interest. It holds the maximum diameter of the palm oil bruise region. Eq.3 shows the max value of diameter where *Pi, Pj* represent the points on the boundary of *D* distance.

$$max = \max_{i,j}[D(p_i p_j)] \tag{3}$$

Minor-axis Length

The *min* variable represents the length of the minor axis of the palm oil bruise region. It holds the minimum diameter value of the region as shown in Eq. 4.

$$min = \min_{i,j}[D(p_i p_j)] \tag{4}$$

Eccentricity (E)

The eccentricity variable measures the aspect ratio of the distance between lengths of major axis to the length of minor axis.

Equidiameter
This feature represents the diameter of the region of interest. It shows the same circle diameter in the area region. Eq.5 below shows the calculation of equidiameter of the palm oil bruise represented by *ED*.

$$ED = \sqrt{(4 * A/\pi)}, \tag{5}$$

Texture Features
Gray-level co-occurrence matrix (GLCM) is a texture features where it to describe the physical composition of the bruise surface. From the image provided, the image of palm oil bruise has a fiber texture. Meanwhile, a clean and shining surface is for non-bruise image. The segmented image of palm oil bruise will be converted to grayscale. The contrast, correlation, energy and homogeneity are extracted from segmented image. Eq. 6 until Eq. 9 shows the equation of GLCM (Gour and Patil, 2016):

Contrast
The intensity contrast between pixel and its neighbor of the segmented bruise are shown in Eq. 6 shows. It shows a value of contrast where *Pi,j* is the co-occurrence matrix and *N* is the size of the matrix.

$$Contrast = \sum_{i,j=0}^{N-1} P_{i,j}(i-j)^2 \tag{6}$$

Correlation
It is the measurement of the correlated pixel with the neighbor of the entire segmented image. The formula of the *correlation* can be computed as Eq. 7 to find the correlation between textures on palm oil bruise.

$$Correlation = \sum_{i,j=0}^{N-1} P_{i,j}\left[\frac{(i-\mu_i)(j-\mu_j)}{\sqrt{(\sigma_i^2)(\sigma_j^2)}}\right] \tag{7}$$

Energy
This angular second moment method is an algorithm to measure the textural uniformity of the palm oil bruise. This algorithm works by extracting the highest value either constant or periodic form of gray-level distribution found in bruise image. Eq. 8 shows the equation for energy.

$$Energy = \sum_{i,j=0}^{N-1} P_{i,j}^2 \tag{8}$$

Homogeneity
The numbers of features produce by the *homogeneity* variables is depends on the *P* matrix where *P* matrix is the a few dominant gray tone transitions as in Eq. 9.

$$Homogeneity = \sum_{i,j=0}^{N-1} \frac{P_{i,j}}{1+(i-j)^2} \tag{9}$$

Table 1 shows the range of extracted values of texture and shape features performed based on

Eq.1 until Eq.9. These tables contain maximum and minimum range of all the features of palm oil bruise. These features are used as an input to a classifier which is Naïve Bayes and SVM.

Confusion Matrix

The performances of these classifiers are measured using confusion matrix. In this research, accuracy is evaluated based on the classifier performance. Confusion matrix consists of components which are True Positive (*TP*), True Negative (*TN*), False Positive (*FP*) and False Negative (*FN*). The accuracy of the classifier is calculated from using the result gain from the confusion matrix table. The Eq. 10 shows the calculation to find the accuracy for each classifier where total of True Positive *(TP)* divided by total of palm oil images which is combination of *TP*, *TN*, *FP* and *FN*. *TP* represent the image correctly detect as bruise and *TN* represent image correctly detect as no bruise. The *FP* incorrectly detect as bruise and *FN* incorrectly detect as no bruise.

$$Classifier\ Accuracy = \frac{\sum TP}{\sum Palm\ Oil's\ images} \times 100 \tag{10}$$

Table 1 Range of Shape and Texture Features

Shape		
Parameter	**Range Value**	
	Bruise (Min – Max)	**No Bruise (Min – Max)**
Area	39379 -138	99-1
Perimeter	2181.72 - 56.12	65.77 - 0
MajorAxisLength	292.1196 -21.02	22.21 – 1.15
MinorAxisLength	230.48 - 6.90	12.72 – 1.15
Eccentricity	0.98 - 0.32	0.99 - 0
EquivDiameter	223.92 - 13.26	11.23 – 1.13
Texture		
Parameter	**Range Value**	
	Bruise (Min – Max)	**No Bruise (Min – Max)**
Contrast	0.06 - 0.01	0.04 - 0.01
Correlation	0.98 - 0.88	1 - 0.91
Energy	0.70 - 0.45	0.58 - 0.47
Homogeneity	1.00 - 0.97	1- 0.98

Results and Discussion

Data collected from the acquisition phase are divided into two categories of training and testing according to percentage which is 80% training and 20% testing. The data divided into 80% training and 20% testing shows classification using extracted features shape and GLCM from Eq. 1 until Eq. 9 on Naïve Bayes obtained a higher performance with 42.5% detected on bruise images and 55% detected for no bruise images with the total overall performance is 97.5%. Meanwhile, SVM shows a poor result with 5.85% detected for bruise images and

55.83% for no bruise images with overall accuracy is 61.66%. It shows that Naïve Bayes has produced a high result compared to SVM. It demonstrates that Naïve Bayes is an effective classifier for the detection of bruises in oil palm fruit.

Table 2 Performance Evaluation between SVM and Naïve Bayes Classifier

Naïve Bayes	**Bruise detection**	**Total images**	**TP Training (80%)**	**TP Testing (20%)**	**Accuracy %**
	Bruise	120	42	9	42.5
	No Bruise	120	52	14	55
	Total accuracy				**97.5**
SVM	**Bruise**	120	0	7	5.83
	No Bruise	120	52	15	55.83
	Total Accuracy				**61.66**

Conclusion

Palm oil bruise detection has been executed in this research and accuracy has been recorded. Detection of palm oil bruise has achieved 97.5% accuracy on detection of the palm oil bruise. Meanwhile, only 2.5% of the samples incorrectly detected. This is due to the image size used as a sample is too small. Besides, the image of palm oil taken may be too glossy and silhouette due to uncontrolled environment during image acquisition process. A combination of six shape features and GLCM has been shown to be appropriate for palm oil bruise detection. The used of Naïve Bayes as classification algorithm to classify between bruise and no bruise images with accuracy more than 90% is well implement with selected features. More samples of palm oil bruise are recommended for future work to improve the accuracy. Besides, more deep analysis to classify the palm oil bruise into more bruise stages will be done.

Acknowledgement

Authors would like to thank Universiti Teknologi MARA Cawangan Melaka and Centre of Vision and Algorithm Analytics Research Group for the supports throughout this research.

Conflict of interests

All authors in this paper declare there is no conflict of interests.

References

Afroza, K., Mahmud B. F., & Halim M. R. (2015). Determination of the Harvesting factors of Independent Smallholder Oil Palm Estate. *Journal of Advanced & Applied Sciences (JAAS), 3*(4), 136–144.

Ashok, V., & Vinod, D. S. (2014). Using K-Means Cluster and Fuzzy C Means for Defect. *International Journal of Computer Engineering & Technology*, *5*(9), 11–19.

Chaki, J., Parekh, R., & Bhattacharya, S. (2015). Plant Leaf Recognition Using Texture and Shape Features with Neural Classifiers. *Pattern Recognition Letters*, *58,* 61-68.

Dang, H., Kim, I., Cho, B., & Kim, M. S. (2012). Detection of Bruise Damage of Pear Using Hyperspectral Imagery. 12th *International Conference on Control, Automation and Systems*, 1258–1260.

Dubey, S. R., Dixit, P., Singh, N., & Gupta, J. P. (2013). Infected Fruit Part Detection using K-Means Clustering Segmentation Technique. *International Journal of Interactive Multimedia and Artificial Intelligence*, *2*(2), 65.

Hadi S. D., Ahmad, F. B. Akande, (2009). Determination of the bruise indexes of oil palm fruits. *Journal of Food Engineering, 95*(2), pp. 322–326.

Helwan, A., & Abiyev, R. (2016). Shape and Texture Features for the Identification of Breast Cancer. *In Proceedings of the World Congress on Engineering and Computer Science*, *2,* 19-21.

James, G. M., & Punitha, S. . (2016). Tomato Disease Segmentation using K-Means Clustering. *International Journal of Computer Applications*, *144*(5), 25–29.

Jamil, N., & Awang, J. (2014). Practical Digital Image Processing with MATLAB. Kuala Lumpur: UiTM Press, UiTM 2014.

Kabutey, A., Herak, D., Mizera, C., & Hrabe, P. (2018). Mathematical Description of Loading Curves And Deformation Energy Of Bulk Oil Palm Kernels. *Agronomy Research, 16,* 1686-1697.

Khade, S., Pandhare, P., Navale, S., Patil, K., & Gaikwad, V. (2016). Fruit Quality Evaluation Using K-Means Clustering. *Proceedings of 49th IRF International Conference,* (February), 11–15.

Krisdiarto W., A., & Sutiarso, L. (2016). Study on Oil Palm Fresh Fruit Bunch Bruise in Harvesting and Transportation to Quality. *Makara Journal of Technology*, *20*(2), 67–72.

Lee, W. H., Kim, M. S., Lee, H., Delwiche, S. R., Bae, H., Kim, D. Y., & Cho, B. K. (2014). Hyperspectral Near-Infrared Imaging For The Detection Of Physical Damages Of Pear. *Journal of Food Engineering, 130*, 1–7.

Lovely, A. K., Mahmud, F., & Mohamed Halim, R. (2015). Determination of the Harvesting factors of Independent Smallholder Oil Palm Estate. *Journal of Advanced & Applied Sciences (JAAS)*, *3*(4), 136-144.

Makky, M. (2016). Trend In Non-Destructive Quality Inspections For Oil Palm Fresh Fruits Bunch. *International Food Research Journal, 23*, 81-90.

Mohana , S., & Prabhakar , C. (2014). A novel technique for grading of dates using shape and texture features. Machine Learning and Applications. *An International Journal (MLAIJ).*

Qiang, L., & Mingjie, T. (2012). Detection of Hidden Bruise on Kiwi fruit Using Hyperspectral Imaging and Parallelepiped Classification. *Procedia Environmental Sciences, 12(Icese 2011),* 1172–1179. https://doi.org/10.1016/j.proenv.2012.01.404

Gour, S. and Patil, P. B. (2016). A Novel Machine Learning Approach To Recognize Household Objects, *in International conference on Signal Processing, Communication, Power and Embedded System (SCOPES*), 69–73.

Shih, F. Y. (2010). Image Processing and Pattern Recognition. *A John Wiley & Sons.* New Jersey: Wiley & Sons, Inc.

Vijayarekha, K. (2008). Multivariate Image Analysis For Defect Identification Of Apple Fruit Images. *Proceedings - 34th Annual Conference of the IEEE Industrial Electronics Society,* IECON 2008, 1499–1503.