Grading of Maturity of Oil Palm Fruit Based on Visible and NIR Band

Goh Jia Quan (1), Abdul Rashid Bin Mohamed Shariff (1), Nazmi Mat Nawi (1)

[1] Faculty of Engineering, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia

Email: gs53069@student.upm.edu.my; rashidpls@upm.edu.my; nazmimat@upm.edu.my

ABSTRACT: This paper reports the grading of oil palm fresh fruit bunch by using optical spectrometer called Oil Palm Ripeness Detector (OPRiD). OPRiD consists of four sensors and eight light sources which is UV (365nm), Blue (460nm), Green (523nm), Amber (590nm), Red (623nm), Deep Red (660nm), Far Red (735nm) and Near Infra-Red (850nm). At least 30 samples of oil palm fresh fruit bunches (FFB) from three categories, that is unripe, ripe and over ripe were subjected to being scanned by the OPRiD. The data were analyzed and compared using machine learning algorithms. The classification results of the OPRID experiment shows an overall accuracy of 83.8% and the Far-Red band was identified as the most significant band. Secondary parameters were obtained using vegetation indices. Non-linear vegetation index (NLI) produced the highest accuracy among the vegetation indices applied, which is 86.8%. Simple Logistic algorithm was determined to be the best algorithms to classify oil palm fruit maturity levels. This research will be valuable in shaping to future of non-destructive oil palm FFB maturity grading. It will be of interest to researchers and professionals in remote sensing, image analysis, mechatronics, hardware engineers, public and private sector researchers and decision makers.

KEY WORDS: oil palm, maturity, ripeness, spectrometer, OPRiD

1. INTRODUCTION

Oil palm tree (*Elaeis guineensis jacq.*) originates from West Africa was brought to Malaya by the British in the early 1870's. During the early period, it was planted as an ornamental plant. Later, this crop turned into commercial planting and became foundation of palm oil industry in Malaysia. The cultivation of oil palm increased at a fast pace in early 1960s under the government's agricultural diversification programme, which was introduced to reduce the country's economic dependence on rubber and tin (MPOC, 2020). Since then, oil palm industry has grown into one of the most important economy income sources of the country. In 2016, the agricultural sector contributed 8.1% of the total Malaysian Gross Domestic Product (GDP), equivalent to RM89.5 million, and the oil palm sector alone constituted 37.9% of the agricultural GDP (Department of Statistics, 2019). Malaysia ranked second in the export of oil palm worth USD 9.7 billion in 2017. The production of palm oil in Malaysia also ranked second in the same year, totaling 21 million metric tons.

Oil palm fruit can be categorized into three categories which is unripe, ripe and over ripe. Malaysia Palm Oil Palm Board (MPOB) has developed standards to assist in ripeness determination based on total number of empty sockets and mesocarp color. The standard is summarized in Table 1.

Table 1: Grading standard of oil palm fresh fruit bunch

Total number of empty fruitlet sockets	Mesocarp color		
	Yellow	Yellowish/orange	Orange
0	Unripe	Unripe	Ripe
0-10	Unripe	Under-ripe	Ripe
>10	Unripe	Ripe	Ripe

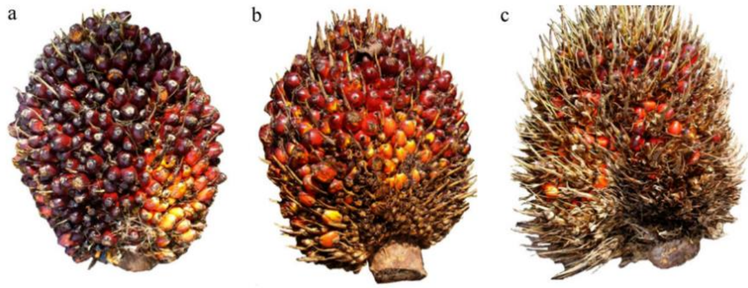

Figure 1. Colour and condition of (a) unripe, (b) ripe and (c) over ripe oil palm fresh fruit bunch (Hafiz, 2011)

However, the fruit maturity classification by using human eyes is inefficient and inaccurate. Manual grading of oil palm FFB into ripeness categories is a very difficult and tedious task, even for an expert grader. Besides, human perceptions to color are often inconsistent, influence by their physical and psychological state (Makky, 2016).

Therefore, a fast and accurate detecting method with help of instruments is needed. Throughout the years, many researchers developed different approaches to assess fruit maturity level. Few of them have been used for on-tree fruit quality inspection process, while as others are best suited for lab level applications. However, Chauhan et al. in 2017 stated that non-destructive methods (NDM) are more effective than conventional method as NDM are mainly based on physical properties which correlate well with certain quality factors of crops. Besides, NDM do not rupture the fruit tissue, can be used to assess internal variable of fruits.

The pioneer in the field of computer gaming and artificial intelligence, Arthur Samuel introduced the term “Machine Learning” in 1959 and defined it as a “field of study that gives computers the ability to learn without being explicitly programmed” (Ivan, 2017). Machine learning are classified into three main categories, which is supervised learning (task-driven), unsupervised learning (data-driven) and reinforcement learning (learn from errors). Commonly used machine learning algorithms are such as linear regression, logistic regression, decision tree, support vector machine, Naïve Bayes, Random Forest and others (Sunil, 2017).

Electromagnetic spectrum is the distribution of electromagnetic radiation. The region on this spectrum with the highest energy (so the shortest wavelengths) are gamma rays and the region with the lowest energy (so the longest wavelengths) are radio waves. The visible region of the spectrum has wavelengths from about 400-700 nm. Near-infrared light generally refers to light within the wavenumber range of 12,500 to 4,000 cm^{-1} (wavelengths from 800 to 2,500 nm) (NASA, 2013).

In this report, we focus on studying grading of oil palm fruit ripeness by using NDM devices, which is Oil Palm Ripeness Detector (OPRiD). Machine learning algorithms were investigated in classifying classes of oil palm FFBs. In order to test whether the classification accuracy can be improved, secondary parameters were computed by applying vegetation indices (VI).

2. METHODOLOGY

2.1 Device

OPRID is a device designed based on these technology (measures reflected EMR in AU form) for determining the degree of maturity and color differences of sensed surfaces (oil palm bunches). OPRID device has high radiometric resolution, thus the device can distinguish minor changes on the surfaces color (Alhadi, 2017). The OPRID consists of four sensors working in different bandwidth and lighting source (LED) in UV (365nm), Blue (460nm), Green (523nm), Amber (590nm), Red (623nm), Deep Red (660nm), Far Red (735nm) and Near Infra-Red (850nm). Main control interface allows the user to select band combination and read the results in arbitrary unit (AU). 25.9V 5000m Ah li-ion battery serves as battery which supply power for LED and main control interface.

The sensing mechanism of OPRiD is described as followed. The surface of object is illuminated by selected LED and the reflected energy is collected by the sensors. The collected energy is filtered, and the relevant energy will be converted to AU. The output will be displayed on the screen of the device. The output data can be saved in a flash drive to be analyzed accordingly.

2.2 Study area and data collection

The data in this experiment were collected at a Palm Oil Mill of a large oil palm plantation company. At least thirty FFB samples from three categories of ripeness which is unripe, ripe and over ripe were scanned using OPRiD.

The FFB samples are classified according to different ripeness (unripe, ripe and overripe) by human grader. The collected data is recorded in light intensity form. They were transferred and arranged accordingly in an Excel sheet. Every FFB's reflectance data was recorded horizontally across the Excel sheet and their ripeness was recorded at the end of the row. From the vertically column, thirty-two model were built by combining the reflectance data of the color band and sensor. For example, UVS1, BlueS1, GreenS1…, FRedS4, NIRedS4. S1 represents sensor 1, S2 represents sensor 2, S3 represents sensor 3 and S4 represents sensor 4. These models were recorded as primary parameters. Secondary parameters were built from the foundation of primary parameters by applying different types of vegetation indices. Vegetation Indices (VIs) are combinations of surface reflectance at two or more wavelengths designed to highlight a particular property of vegetation. They are derived using the reflectance properties of vegetation. Each of the VIs is designed to accentuate a vegetation property (Yusuf, 2011).

2.3 Statistical test

Selected data was classified and analyzed using a machine learning software, Waikato Environment for Knowledge Analysis (Weka) developed at the University of Waikato, New Zealand. Weka is tried and tested open source machine learning software that can be accessed through a graphical user interface, standard terminal applications, or a Java API. The classifiers algorithms are divided into Bayesian classifiers, trees, rules, functions, lazy classifiers and a final miscellaneous category. Some commonly used algorithms are Logistics, Simple Logistics, J48, Logistic Model Tree, Random Forest and Random Tree. These classification algorithms were chosen to perform on the selected attributes. The algorithms that produced the highest accuracy reading for each model were also recorded.

3. RESULTS AND DISCUSSION

Among 32 models that were built, only eleven models were remained after the models that did not capture any reflectance were removed. These eleven models are GreenS1, RedS2, DRedS2, FRedS3, BlueS4, GreenS4, AmberS4, RedS4, DRedS4, FRedS4, NIRedS4. Their classification accuracy was recorded in table below.

Table 2: Classification accuracy test results

No.	Models	Correct %	Algorithms
1.	GreenS1	58.8	Simple Logistic
2.	RedS2	54.4	Simple Logistic, Logistic Model Tree (LMT)
3.	DRedS2	55.9	Simple Logistic
4.	FRedS3	80.9	Simple Logistic
5.	BlueS4	48.5	Bagging
6.	GreenS4	52.9	LMT
7.	AmberS4	50	Simple Logistics
8.	RedS4	52.9	Simple Logistic, LMT
9.	DRedS4	58.8	Logistic
10	FRedS4	83.8	Logistic, Simple Logistic, LMT
11.	NIRedS4	82.4	Logistic

The result in the table shows that FRedS4 has the highest classification accuracy of 83.8% and followed by IRedS4 with 82.4% accuracy. BlueS4 was identified as the least significant band as it only has 48.5% accuracy. In the system of OPRiD, sensor 4 is a sensor that collect raw reflectance without any filter. Thus, the result suggested that reflectance without any disturbance from filter is the most suitable data for classification. On the other hand, a combination model that comprises of all the above models in the table achieved an accuracy of 85.3%, by applying Simple Logistic as the classifier. Among 11 model's classification, Simple Logistic able to produce highest accuracy for seven times. This made it become the best algorithm to classify the oil palm fruit ripeness.

Comparing this result to Alhadi's research using the same device in 2017, some models obtained higher accuracy, and some obtained lower accuracy. GreenS1, RedS2, DRedS2, GreenS4, AmberS4, RedS4 and DRedS4 obtained lower accuracy. Meanwhile, FRedS3, BlueS4, FRedS4 and IRedS4 have higher accuracy. The reason for the differences in accuracies between the two experiments could be due to different maturity colour characteristics of the oil palm fresh fruit bunches varieties at the two different locations of the experiments. The details were recorded in Table 3.

Several vegetation indices were applied onto the dataset to produce a secondary parameter for classification. These vegetation indices were mainly used maximize sensitivity to the vegetation characteristics while minimizing confounding factors such as soil background reflectance, directional, or atmospheric effects. The vegetation indices formula and the classification results were shown in the Table 4.

Table 3: Comparison between two research

Models	Alhadi's research's accuracy	This research's accuracy
GreenS1	61.90	58.80
RedS2	59.22	54.40
DRedS2	73.80	55.90
FRedS3	50.00	80.90
BlueS4	52.38	48.50
GreenS4	57.14	52.90
AmberS4	64.28	50.00
RedS4	73.80	52.90
DRedS4	69.05	58.80
FRedS4	57.14	83.80
NIRedS4	38.10	82.40

Table 4: Classification accuracy test results of vegetation indices in OPRiD experiment.

No.	Vegetation indices	Formula	Correct %
1	Normalized Difference Vegetation Index, NDVI	(NIR-Red)/(NIR+Red)	73.5
2	Difference Vegetation Index, DVI	NIR-Red	85.3
3	Simple Ratio	(NIR/Red)	67.6
4	Green Vegetation Index	(NIR-Green)/(NIR+Green)	70.6
5	Infrared percentage vegetation index, IPVI	$\frac{\mathrm{NIR}}{\mathrm{NIR}+\mathrm{Red}}$	73.5
6	Non-linear vegetation index, NLI	$\frac{\mathrm{NIR}^2-\mathrm{Red}}{\mathrm{NIR}^2+\mathrm{Red}}$	86.8
7	Renormalized difference vegetation index, RDVI	$\frac{\mathrm{NIR}^2-\mathrm{Red}}{\sqrt{\mathrm{NIR}+\mathrm{Red}}}$	83.8
8	Modified simple ratio, MSR	$\frac{\frac{\mathrm{NIR}}{\mathrm{Red}}-1}{\sqrt{\frac{\mathrm{NIR}}{\mathrm{Red}}}+1}$	72.5

By using vegetation indices as secondary parameter, the classification accuracy was able to be elevated to 86.8%, which is slightly higher than the highest accuracy that was produced by using FredS4 model. This result was achieved by Non-linear vegetation index (NLI). Besides, difference vegetation index (DVI) and renormalized difference vegetation index (RDVI) also produced high accuracy of 85.3% and RDVI obtained 83.8%. Simple Ratio obtained the lowest accuracy of 67.6%. VIs does not improve the accuracy of the classification of oil palm ripeness using a portable four-band sensor (Osama, 2011), however this research showed that OPRID was able to overcome the limitations of the earlier work and allows VIs to improve the accuracy of the classification of oil palm ripeness.

4. CONCLUSION

This research studied the application of visible and NIR bands of OPRiD and its ability to determine the oil palm ripeness level. The FRedS4 model was identified as the best model for classification with 83.8% accuracy. The analysis of vegetation indices also indicated that Non-Linear Vegetation Index (NLI) is the best VI to classify oil palm ripeness, with an 86.8% accuracy. Besides, Simple Logistic was identified as the best algorithm. Thus, the device can be concluded as a useful sensor for oil palm fresh fruit bunch classification. The result can be valuable for future researcher and market to study further into non-destructive method of oil palm fruit maturity grading.

5. REFERENCE

Aiad, A, “Oil Palm Fresh Fruit Bunches Maturity Classification and Oil Analysis Correlation,” M.S. thesis, Fac. Eng., UPM, June 2017

Chauhan, O. P., Lakshmi, S., Pandey, A. K., Ravi, N., Gopalan, N., & Sharma, R. K., 2017. Non-destructive Quality Monitoring of Fresh Fruits and Vegetables. *Defence Life Science Journal*, *2*(2), pp. 103-110.

Department of Statistics Malaysia, 2019. Selected Agricultural Indicators, Malaysia, 2019 [Press release]. November 29, 2019.

Hafiz, M., Rashid. A., Amiruddin, M., 2012. Determination of oil palm fresh fruit bunch ripeness—Based on flavonoids and anthocyanin content, *Industrial Crops and Products*, Volume 36, pp.466-475.

Ivan, R., 2017. What is Machine Learning?. Retrieved October 5, 2020 from https://www.cognizantsoftvision.com/

Makky, M., 2016. A Portable Low-cost Non-destructive Ripeness Inspection for Oil Palm FFB. *Agriculture and Agricultural Science Procedia,* Volume 9, pp. 230-240.

MPOC, na. The Oil Palm Tree, Retrieved October 9, 2020, from http://mpoc.org.my/the-oil-palm-tree/

National Aeronautics and Space Administration, 2013. The Electromagnetic Spectrum. Retrieved October 5, 2020 from https://imagine.gsfc.nasa.gov/science/toolbox/emspectrum1.html

Ray, S. 2017. Commonly used Machine Learning Algorithms (with Python and R Codes). Retrieved October 5, 2020 from https://www.analyticsvidhya.com/blog/2017/09/common-machine-learning-algorithms/

Saeed, Osama & Sankaran, Sindhuja & shariff, rashid & Shafri, Helmi & Ehsani, Reza & Alfatni, Meftah & Hafiz, Mohd, 2012. Classification of oil palm fresh fruit bunches based on their maturity using portable four-band sensor system. *Computers and Electronics in Agriculture*, Volume 82, pp. 55–60.

Yusuf, B., 2011. Application of hyperspectral imaging sensor to differentiate between the moisture and reflectance of healthy and infected tobacco leaves. *African Journal of Agricultural Reseearch*, Volume 6, pp. 6267-6280.