A study on the oil palm fresh fruit bunch (FFB) ripeness detection by using Hue, Saturation and Intensity (HSI) approach

Muhammad Kashfi Shabdin[2], Abdul Rashid Mohamed Shariff[1, *], Mohd Nazrul Azlan Johari[1], Nor Kamilah Saat[2], Zulkifly Abbas[2]

[1]Department of Biological and Agricultural Engineering, Faculty of Engineering, Universiti Putra Malaysia, Serdang, Selangor, Malaysia

[2]Department of Physics, Faculty of Science, Universiti Putra Malaysia, Serdang, Selangor, Malaysia

*Email: rashidpls@upm.edu.my

Abstract. To increase the quality of palm oil means to accurately grade the oil palm fresh fruit bunches (FFB) for processing. In this paper, HSI color model was used to determine the relationship between FFB's color with the underipe and ripe category so that the grading system could be accurately done. From the analysis manipulation, a formula was generated and applied to the data obtained. The by linear regression in the data shows an average success rate at 45% accuracy for oil palm ripeness detection. Artificial Neural Network (ANN) however return a better accuracy result for both underipe and ripe categories which are 60% and 80% respectively. This yield an overall accuracy of 70%. This can be increased more by improving the grading system.

1. Introduction

Quality is the most important factor in the oil palm production because high quality products were believed to be the key to success in today's competitive market. In oil palm productions, quality is always determining by its texture, shape and color. These features of oil palm Fresh Fruit Bunches (FFB) is always been observed by the human's vision which leads to inconsistencies and inaccurate. Besides, this job is tiresome and time consuming. These faults lead research on methods to make this system computerized using machine vision based technologies by using the imaging technique.

Oil palm fruit as shown in figure 1 is one of the major agricultural product exports by Malaysia. Palm oil has become the ingredient in the making of margarine, candles, soaps, domestic frying oil and snack foods [1]. Oil palm FFB is very mutual in Malaysia. Grading of oil palm fruit conventionally observed by human's vision but today, many types of research has been carried out to find the correlation between oil content in the oil palm fruit against the color of the oil palm fruit [2].

Figure 1. Three categories of oil palm: left is underipe, the center is ripe, right is overipe.

In this research, oil palm is utilized as sample for testing the improved fresh fruit bunch (FFB) grading system because this machine is fabricated especially to assist the conventional oil palm grading system. This machine uses the imaging technique in getting important information regarding the feature of the FFBs. With a closed environment and balanced illumination, a good image could be captured thus analyzed. Imaging technique can utilize features and color of the fruit. Therefore, in this study, colors (HSI) were chosen as the main parameter because the fabrication of the machine already utilized all the features and not the colors of the FFBs. The scope of work will only cover the two main category of oil palm which is underipe and ripe. This is because these two categories are hard to differentiate between each other and are the main categories that are in need to be segregated. Overipe category, however, is easier to differentiate when compared with ripe. Overipe is not in concern because it will be processed along with ripe fruit but underipe has to be segregated because processing the fruit will be costly and yield low quantity palm oil.

Color provides valuable information in estimating the ripeness of the oil palm FFB. Color is one of the most significant criteria related to fruit ripeness identification [2]. Light reflected from an object determines the color of the object thus these variations provide a base for image processing and analysis. Red, Green and Blue (RGB) are the main components in the color code. The other color is the additive of these three main colors. Although a human's perception of color is a subjective process, the physical nature of color still can be distinct by various experiments and results which explained by the frequent response of the brain to the stimuli when in contact with lights, naturally produced when incoming light reacts with several types of Cone cells in the eye.

According to [3], when a beam of light comprising of many wavelengths approach the eye, its response is not given by the sum of the responses that would be produced by each one of the component wavelengths acting alone. Although there are many different distributions of wavelengths that can be observed, our characterization of the appearances of these distributions can be expressed using a relatively small number of parameters. That is, there are many different combinations of wavelengths which appear to produce the same visual color.

Even though a number of different methods can be used to characterize a combination of wavelengths, it turns out that all of them use either three or four parameters since this small number of parameters is related to the way the eye recognizes color. The simplest triplet of parameters is called hue, saturation and intensity.

If a beam of light is break up into each of its component wavelengths and if the intensity of each component is plot as a function of wavelength, then, the hue is the peak of this plot – the wavelength (or relatively insignificant band of wavelengths) which has the extreme intensity. The hue is generally the single word that can be used to describe a composite color. Hue values range from about 440 nm for violet, 450 nm for blue, up to about 700 nm for red light. The names associated with different hues follow the spectral decomposition of a rainbow: red, orange, yellow, green, blue, and violet. These descriptive colors are associated with ranges of wavelengths rather than with unique values, and some people can see colors outside of this conventional range of wavelengths (ultraviolet with a wavelength shorter than violet or infra-red with a wavelength longer than red).

The saturation of a beam of light is related to the width of the plot of intensity vs. wavelength described above. A completely saturated beam would have only one wavelength and would be called monochromatic, which a completely unsaturated beam would contain all wavelengths in equal fraction and would appear white. An absolutely saturated beam therefore, has a very narrow intensity distribution function (possibly consisting of only one non-zero value in the limit), which a completely unsaturated beam has a very wide distribution function, possibly consisting of a constant value over most or all of the visible spectrum.

The intensity is related to the strength of the light beam. Intensity is very difficult to specify because the apparent brightness and the actual brightness can differ significantly. The intensity is related to the total power in the light beam as measured by some objective instrument (such as a photographic light meter), but the perceived brightness of a light (or lightness of a surface) is strongly influenced by lots of other factors and cannot always be specified objectively.

These parameters are often not independent of each other. For example, the intensity and hue of a standard light bulb are related through the black-body relationships – decreasing the output intensity of a black body also shifts the hue towards longer wavelengths.

The hue, saturation and brightness of a light beam are often specified using a three-dimensional color tree, as shown in Figure 2. The vertical axis of the tree specifies the intensity of the beam, from nothing at the bottom which is black, through gray to some maximum value at the top equivalent of the brightest possible white. At each level of the tree (which corresponds to a given lightness or brightness), we draw a circle whose circumference shows the various pure, fully saturated, monochromatic colors of the rainbow in wavelength order from red to violet. The points on a radius line from the center of the tree to some point on the circumference represent different unsaturated colors formed by mixing some amount of white from the center of the tree with some amount of the color at the endpoint of the line.
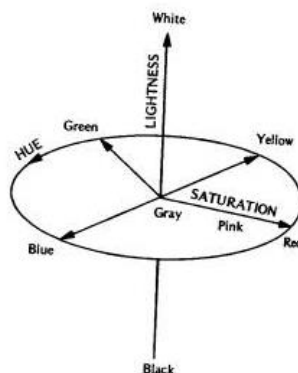
Figure 2. The HSI representation of color [3].

Whereas in oil palm application, shape and pattern cannot be a typical guide for one to estimate the ripeness of FFBs. This is because the fruit may vary in the shape and pattern but share equal quality. In order to overcome this, the grading device should be able to analyze the color of the fruit therefore measures the density of the color by using RGB model. It was believed that color changes resulting from the chemical reactions in the fruit texture can be related to the fruit maturity [4]. This application would decrease the usage of the human grader and minimize the time consume for processing purposes.

Latest development and application of the basic concepts and technologies associated with image capturing and processing were reviewed. There was an approach done to classify potato chips by using pattern recognition from color digital images which include a) image acquisition, b) preprocessing, c) segmentation, d) feature extraction and e) classification [5]. There were also some of the sensor systems which have been explored for machine vision based FFB grading were using optical RGB cameras [6][7][8][9][10]. In [6], the harvesting machine was designed by using the RGB color intensity where they can segregate between different FFB categories. The charged couple device (CCD) camera was used in this system. Hue, Saturation and Intensity (HSI) were used in [7] to predict the best time for harvesting the FFB and by using this color model with multivariate discriminant analysis, they found that the average misclassification for the vision system at 8%. In [8], visible and near infrared wavelength is used to differentiate the old wheat plant which is infected by the yellow stripe, leaf and stem rust. They claimed to found that the sequential application of Anthocyanin Reflectance index can separate the healthy and infected plant under the laboratory conditions. In [10], the experiment was done by using the RGB color model to differentiate between different categories. The data was then trained by using supervised learning Hebb technique and graded using fuzzy logic. The results showed that the automated grading by using the RGB produced an average of 49% success rate while the neuro-fuzzy logic used achieved an accuracy level of 73.3%. Therefore, the motivation here comes to analyzed different categories of FFB by using HSI color model in order to correctly classify whether the FFB is good to be processed or not.

2. Methodology
In this study, the main concern is the two category which is underipe and ripe. This is because they are the main class that determines the quality and quantity of the oil. Therefore, the data collection and the data analysis will only revive on these two categories. Today, the milling industries have difficulties in defining underipe and ripe category but has less problem in detecting between ripe and overipe category. This is because from the physical appearance both ripe and overipe can be differentiated by the naked eye but the underipe and ripe category is difficult to distinguish.

2.1 Data collection

Malaysian Palm Oil Board (MPOB) has established standards to correctly classify the oil palm FFB which is : 'underipe' with 1 – 9% loose fruits from the bunch, and 'ripe' with 10 – 50% loose fruits from the bunch [2]. In this study, images were captured by using 60 samples of fresh fruit bunches (FFB) that already been categorized according to the mill graders and inspectors before the data collection being carried out.

There are few types of oil palm in Malaysia, but the most planted are Elaeis guineensis (subsps: Nigrescens) which acquired from Pertubuhan Peladang Salak Tinggi, Sepang, Malaysia. The FFB was placed in the FFB grading machine and images were taken from the side of the FFB. The area covered was about 100cm x 100cm. Digital images were captured from overall 30 bunches under each category (underipe and ripe) using a GigE camera. Readings taken was then saved for further analysis.

2.2 Color feature extraction

This paper investigated on oil palm FFB grading using hue saturation and intensity (HSI). The image was taken in the grading system with the closed environment so that no ambient light would disturb the quality of the image. Figure 3 shows the image before and after background extraction. ENVI classic was used to extract the background from the image and then the image was uploaded in the MATLAB software to extract the RGB values. The HSI values were then extracted from the RGB values. The HSI values are then analyzed by using WEKA 3.6 software.
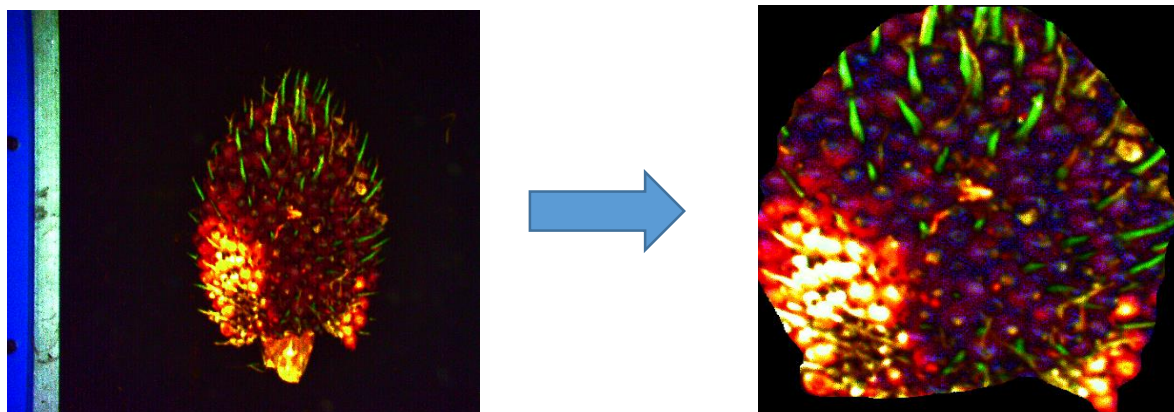

Figure 3. FFB image after background extraction using ENVI 5.1.

3. Results & Discussion

Results are divided into two parts, which are the analysis by using WEKA software and artificial neural network analysis. The WEKA analysis shows different results compared to the ANN analysis. This is perhaps the methods of running the analysis is different from each other. WEKA software uses cross validation to validate all data but ANN analysis uses training and testing data to validate all sample.

3.1 Analysis by WEKA software.

The analysis of the data in the WEKA 3.6 software uses cross validation to determine the correlation coefficient of the data set. Figure 4 shows the correlation coefficient for the category and the HSI. From the correlation, we can indicate that the accuracy for the underipe and the ripe category by using the HSI

model is an average of 45%. The mean absolute error is 38% and the root means squared error is 44% showing that the data is averagely correlated but can be improved.
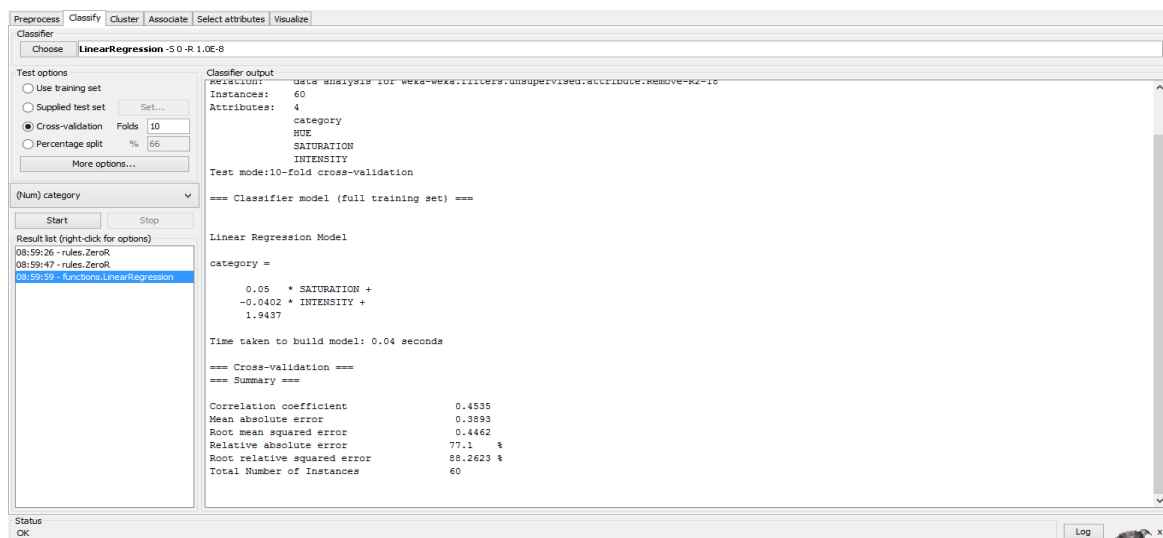

Figure 4. The correlation coefficient for the category with HSI.

Table 1 shows the difference in accuracy shown by different type of combination for the HSI data. Oil palm category was taken as the independent variable whereas HSI as dependent variables. For combination number one, the category is taken into cross validation with Hue and produce 17% accuracy, which is the lowest accuracy compared to other combinations. The highest average accuracy obtained was for combination number five, which is category taken with saturation and intensity. The accuracy was 45%. Combination number five also shows the lowest mean absolute error and root mean squared error which is 38% and 44% respectively. Combination number six shows the highest mean absolute error for 52% and for root mean squared error, combination number one and six shows the highest at 53%.

Table 1. The difference in accuracy shown by different HSI combination.

number	combination	accuracy	mean absolute error	root mean squared error
1	category, H	17%	51%	53%
2	category, S	25%	45%	48%
3	category, I	34%	51%	52%
4	category, H,S	18%	46%	50%
5	category, S,I	45%	38%	44%
6	category, H,I	20%	52%	53%

Figure 5 shows the difference in accuracy by different combination as explained in Table 1. The accuracy for the Hue is lower than when independent category variable is combined with saturation and intensity. This situation happens perhaps because of the lightness that is a shift to a longer distribution wavelengths. This cause it to be a non – monochromatic and unsaturated beam as mentioned by [3].
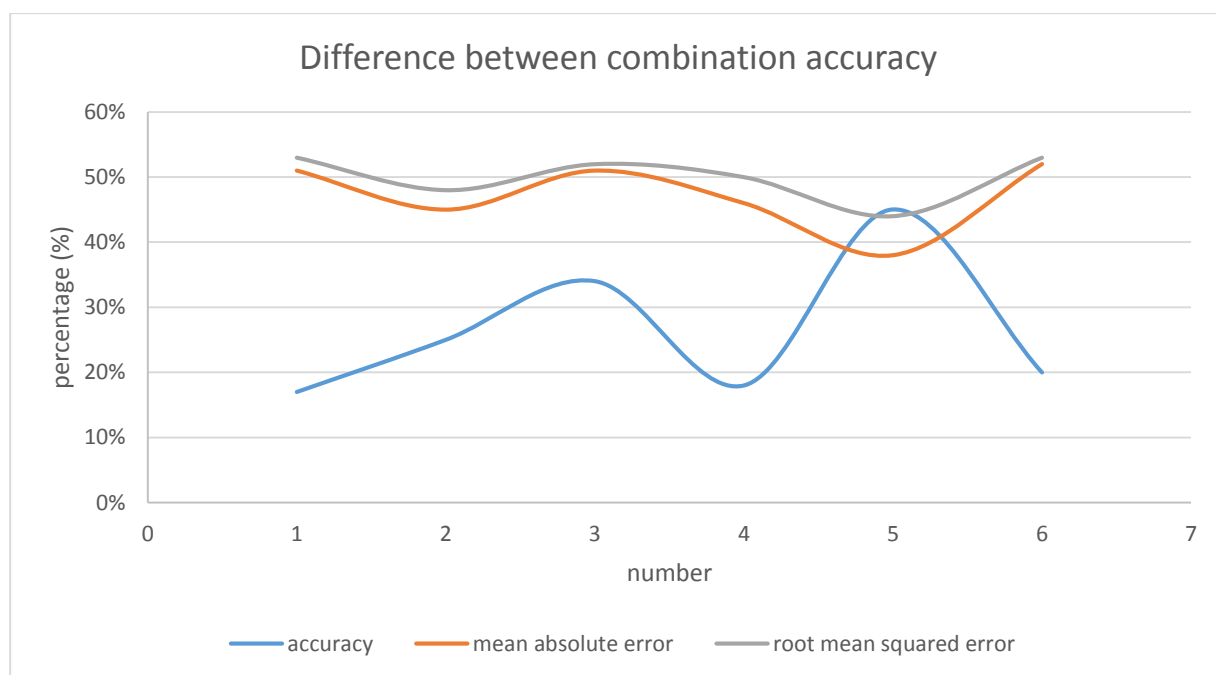

Figure 5. The accuracy, mean absolute error and root mean squared error of the HSI combination.

3.2 Analysis by artificial neural network (ANN).
The artificial neural network allows the user to input the training data and testing data. Training data are used to train the system to recognize the sample features so that it can be tested accurately. Figure 6 shows the data obtained when the system is trained. the R value for the training data is 0.8 and this shows a near correlation between the sample and the output. Validation and test result however, shows better R value which is 0.99 which means that the sample has a high correlation with the output. From the regression, the R value for all correlation is 0.85. This suggests that this neural network can be used for the training and testing purposes.
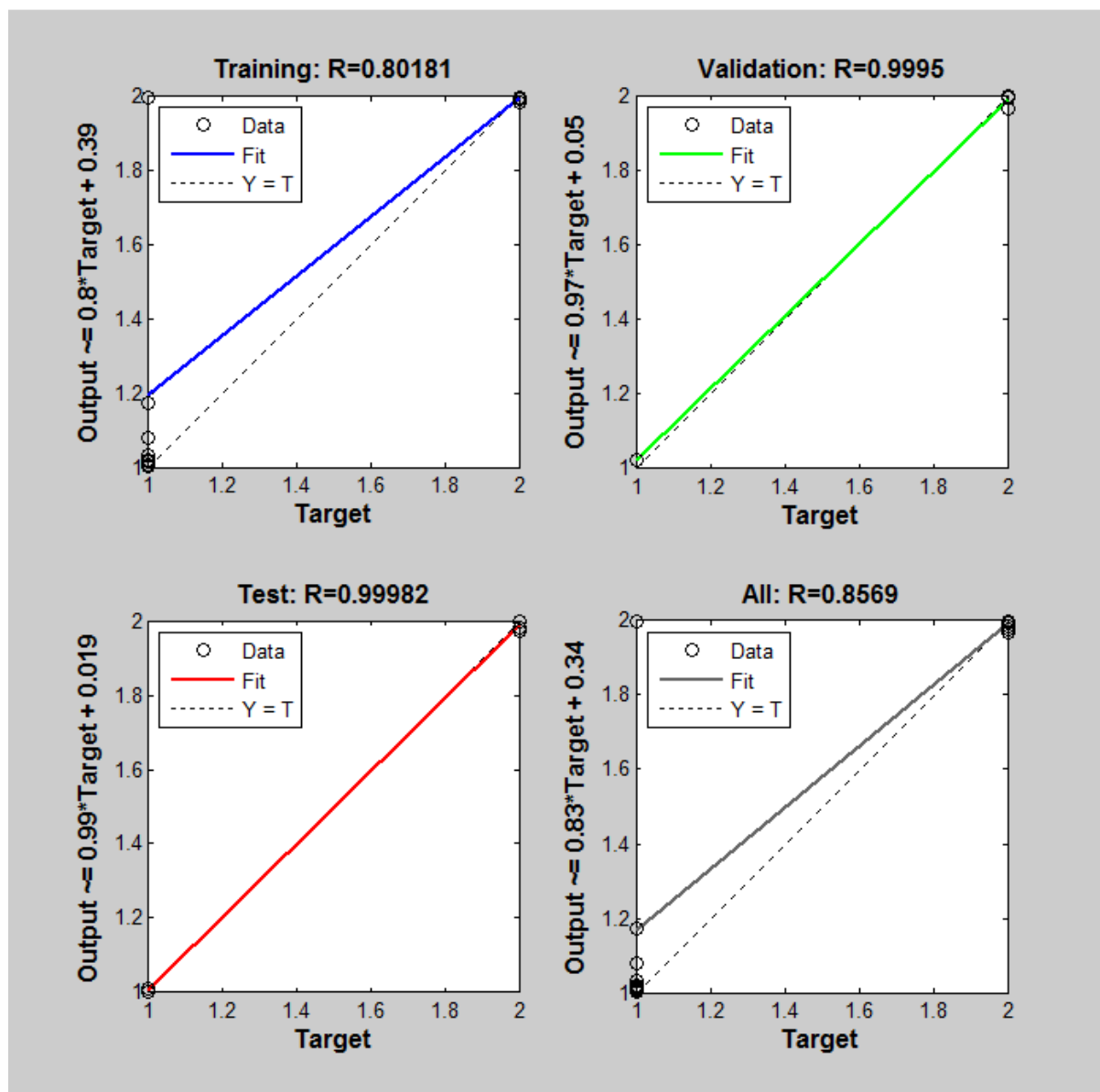
Figure 6. The training, testing, validation and overall regression of the ANN analysis.

After few times of training the data to fit in the equation, the testing of the data is carried out so that the accuracy and efficiency of the system can be tested. The results are shown in Figure 7. Hue, Saturation and Intensity was plotted on the same axis while the result is on the secondary axis. The yellow dot shows the system categorizing the FFB into their correct classification while the red dot shows incorrect classification. For the ANN system to work, earlier it was set into the training system that for any result with number one, that means underipe, whereas if the result comes out number two, that means ripe. 20 samples were taken into account for testing where sample number one – ten was underipe and 11 – 20 was ripe. The classification shows that the accuracy for underipe is 60% because of four incorrect classifications and for ripe, the accuracy is 80% because of two incorrect classifications. This makes the overall accuracy of 70% for the ANN system.
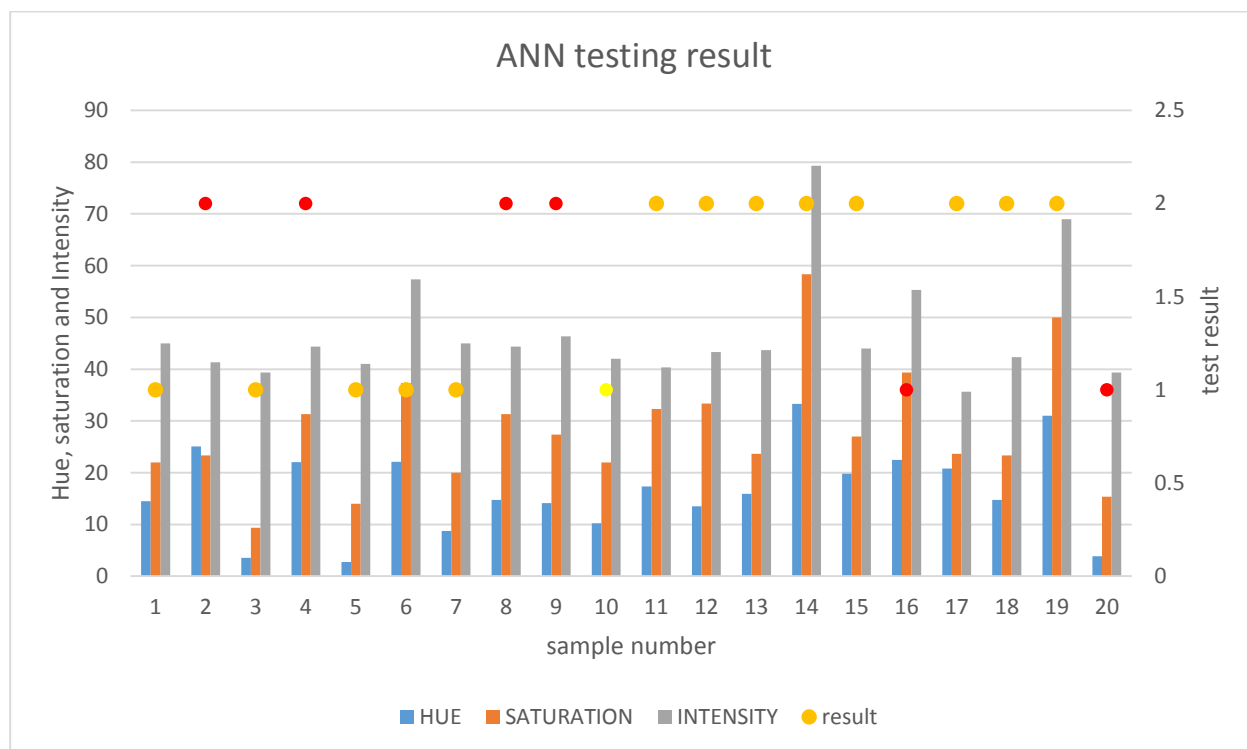
Figure 7. The testing result of the ANN analysis.

4. Conclusion

This study evaluated the applicability of a new GigE camera with HSI color model in determining the oil palm underipe and ripe maturity in a closed environment. The incident beam from the illumination in the closed environment is in low intensity thus not enough causes the beam to be unsaturated and non – monochromatic. This affects the image of the FFBs consequently produced the highest accuracy obtained for the analysis is at only an average of 45%. This average result was obtained by applying the linear regression from the WEKA software. Artificial Neural Network (ANN) however brings about higher accuracy in oil palm FFB ripeness detection. The underipe and ripe categories were taken into account and yield an acceptable accuracy result of 60% and 80% respectively. Those result then yield an overall accuracy of 70%. ANN shows more reliable results than WEKA and it could be utilised with the HSI measurement for oil palm. Furthermore, this could be a great analysis if the sample condition and the illumination in the system can be increased to a higher level but this would introduced noise into the system. Further research and analysis will combine the automated grading system that can be tested with a larger data set and thus more accurate analysis.

5. Acknowledgement

6. References

[1]   N. Fadilah, J. Mohamad-Saleh, Z. A. Halim, H. Ibrahim, and S. S. S. Ali, "Intelligent color vision system for ripeness classification of oil palm fresh fruit bunch," *Sensors (Switzerland)*, vol. 12, pp. 14179–14195, 2012.

[2]   M. S. M. Alfatni, a R. M. Shariff, H. Z. M. Shafri, O. M. B. Saaed, and O. M. Eshanta, "Oil palm fruit bunch grading system using red, green and blue digital number," *Journal of Applied Sciences*, vol. 8. pp. 1444–1452, 2008.

[3]   Judah Levine, "Light and Color. Topic 45. Hue, Saturation and Intensity." [Online]. Available: http://www.colorado.edu/physics/phys1230/phys1230_fa01/topic45.html. [Accessed: 22-Mar-2016].

[4]   F.-R. A. Thomas S.Y. Choong, Saad Abbas, Abdul Rashid Shariff, Rohaya Halim, Mohd Halim S. Ismail, Robiah Yunus, Salmiaton Ali, "Digital Image Processing of Palm Oil Fruits," *Int. J. Food Eng.*, vol. 2, no. 2, pp. 1556–3758, 2006.

[5]   J. M. A. F. P Edreschi, D. Mery, F. Mendoza, "Classification of Potato Chips using Pattern Recognition," *J. Food Sci.*, vol. 69, no. 6, pp. E264–E270, 2004.

[6]   Wan Ishak Wan Ismail, Mohd Zohadie Bardaie, Abdul Malik Abdul Halim, "Optical properties for mechanical harvesting of oil palm FFB," *J. Oil Palm Res.*, vol. 12, no. 2, pp. 38–45, 2000.

[7]   Albdullah, M.Z., Guan, L.C., Mohamed, A.M.D., Noor, M.A.M., "Color vision system for ripeness inspection of oil palm Elaeis Guineensis.," *J. Food Process. Preserv.*, vol. 26, pp. 213–235, 2002.

[8]   Devadas, R., Lamb, D.W., Simpfendorfer, S., Backhouse, D., "Evaluating ten spectral vegetation indices for identifying rust infection in individual wheat leaves," *Precis. Agric.*, vol. 10, pp. 459–470, 2009.

[9]   A. Alfatni, M.S.M., Shariff, A.R.M., Shafri, H.Z.M., Saaed, O.M.B., Eshanta, O.M. and M., "Automated oil palm fruit bunch grading system using density of color (RGB).," *Proc. 7th Saudi Eng. Conf. King Saud Univ. Riyadh*, pp. 299–313, 2007.

[10]   S. Jamil, N., Mohamed, A., Adbullah, "Automated grading of palm oil fresh fruit bunches (FFB) using neuro-fuzzy technique.," *Proc. 2009 IEEE Int. Conf. Soft Comput. Pattern Recognition, Malacca, Malaysia,* pp. 245–249, 2009.