CAROTENE CONTENT ANALYSIS USING SPECTROPHOTOMETER AND ARTIFICIAL NEURAL NETWORK

Arfah Binti Saman
Pn Zuriati Binti Janin
Faculty of Electrical Engineering
Universiti Teknologi Mara
40450 Shah Alam

Abstract

This paper presents a study to investigate the carotene content of palm oil fruit using spectrophotometer and *Artificial Neural Network* (ANN). The study confined only three different ripeness of palm oil fruit on *Tenera* species that are known as ripe, over-ripe and under-ripe. 10 samples of each ripeness of fruit are analyzed and scanned to capture image spectrum through spectrophotometer. Then, in turn that image will be used in the analysis of backpropagation neural network in determining the highest carotene from ripeness of palm oil fruit. The result shows that ripe is the best ripeness of palm oil which contain highest carotene through errors analysis from actual and target outputs out of 30 samples of difference ripeness each. From that analysis, it can be conclude that smaller error represent more quality carotene content of palm oil.

Keywords - Spectrophotometer, neural network and carotene content.

1. INTRODUCTION

The carotenoids, whose name is derived from the fact that they constitute the major pigment in the carrot root, Daucus carota, are undoubtedly among the most widespread and important pigments in living organism. The reason why many of researchers had spent much times to study the existing of the carotene content, as a Tenera over a decade is the unique characteristics of the carotene itself. The current planting material in Malaysia is a cross of the Dura and Pisifera varieties known as Tenera, all belonging to the Elais Guineensis species. Currently, studies on carotene content become more interesting and they are focusing on health and nutritional aspect. Since they are likely to importance and value, their recovery from palm oil and its by-product is important. In fact, crude palm oil is the world's richest natural plant source of carotenes in terms of retinol (provitamin A) equivalent. It contains about 15 to 300 times as many retinols equivalents as carrots, leafy green vegetables and tomatoes, which are considered to have significant quantities of pro-vitamin A and also content vitamin E activity. Of the vegetable oils (soy-bean oil, olive oil, barley oil, and linseed oil) that are widely consumed palm oil contains the highest known concentration of agriculturally derived carotenoids. The total carotenoids in palm oil are usually determined by ultraviolet-visible spectroscopy at 446 nanometers as parts per million (ppm) of β-carotene. Then, analysis shows that α and β-carotenes constitutes approximately 90% of the total carotenoids content.

2. METHODOLOGY

The process of developing this project was divided into two main parts. The first part is to obtain an absorption spectrum for different types of samples of oil palm by using spectrophotometer from 200nm to 600nm wavelength. The second part is the analysis of backpropagation neural network to determine the highest of carotene from ripeness of fruit as a Tenera species. Figure 2.1 illustrates the flow chart on how the project was developed.

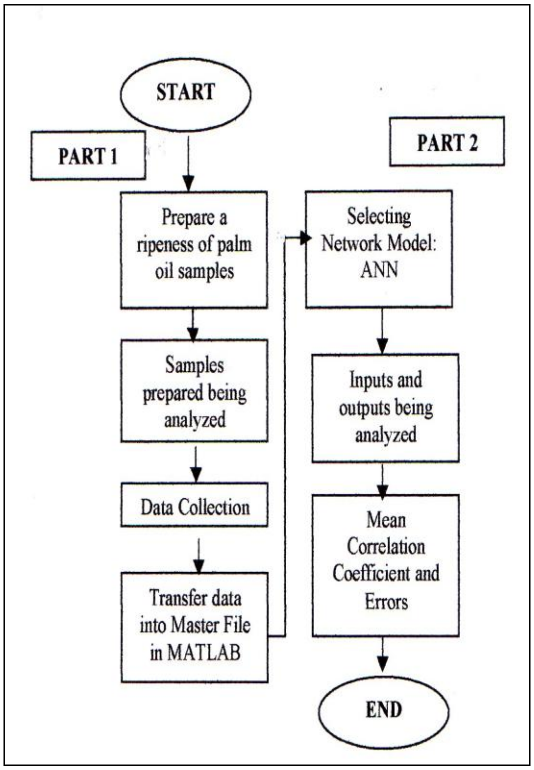

Figure 2.1: Flow Chart for Overall Process

2.1 Samples for Analysis

In part 1, the experiment is progressed at laboratory in MPOB Serdang. The first step is preparation of samples for analysis. Melt the sample at 50°C - 60°C and homogenize thoroughly before taking a test portion. Filter through a fast filter paper if the sample contains impurities or is not clear.

Second step is about procedure on how to conduct this experiment. Weight, to nearest 0.0001g, 0.1g of the samples into the 25ml volumetric flask. Dissolve the test portion with a few milliliters of solvent and dilute to the mark. Transfer the solution to a 1cm cuvette and scanned samples. Notice that the spectrophotometer will scan through the sample starting from 600nm until 200nm.

Spectrophotometer UV/VIS, lambda 12 is connected to computer. A disked is used to save image spectrum from the computer as link on the spectrophotometer and this image is employed as the reference pattern spectrum for ANN analysis as a second part.

2.2 Experiment Process (ANN Analysis)

For each of ripeness oil palm only one image used, as a reference and it will be selected more carotene than other 10 samples. Therefore ten samples consist of 401 points each samples of absorption spectrum will be employed for ANN inputs and it's divided into training, and testing set. Three samples of ripeness are selected for the training set. In the backpropagation algorithm, a network with two layers was applied. Neuron in the input layer used tan-sigmoid transfer function *'tansiq'* and output layer applied linear transfer function *'purelin'* [3]. The inputs to the network model are the value of the absorption spectrum extracted from 200nm to 600nm wavelength and the outputs are the pattern spectrum. Different training functions and value of learning rate, momentum constant and number of nodes are varied to find the best possible network model [4]. The ripeness oil palm analysis is continued on the regression analysis, error between actual and target output, and the average of the correlation coefficient (R-value) between the actual outputs and the targets outputs to determine the best ripeness of the network models.

3. RESULTS

3.1 Spectrum from spectrophotometer.

The images of the pattern spectrum capture from spectrophotometer are variable among each types of palm oil and the values of absorption spectrum are different base on fruit of ripeness as shown in Figure 3.1. The X-axis and Y-axis are represents the range of wavelength and the absorption spectrums respectively. The result has come out an absorption spectrum at wavelength at range 200nm to 600nm and point at 446nm is applied. From this point, the carotene content can be determined as shown in Table 3.1. It can be concluded that the ripe fruit is contained more carotene and absorbed highest energy of light that passes through than other types of palm oil. Hence, the maximum value of the absorption spectrums at 446nm was able to predict a good source of carotene. The total carotenoids in palm oil are usually determined by ultraviolet-visible spectroscopy at 446 nanometers as ppm of β-carotene. Palm oil contains the highest known concentration of agriculturally derived carotenoids [5]. The carotene is state more clear if the maximum value of the absorption is bigger.

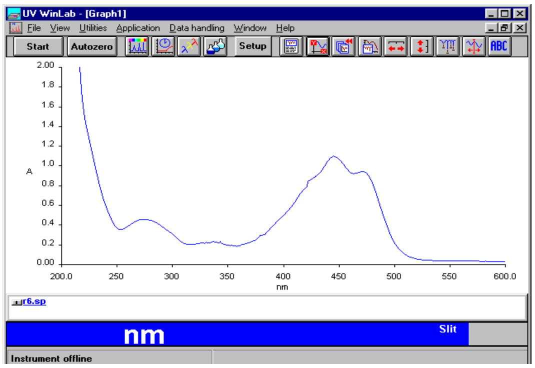

Figure 3.1. Image of the absorption spectrum versus wavelength of ripeness oil palm from spectrophotometer.

Table 3.1: The carotene content for different types of ripeness

Type of fruit	Weight (g)	Wavelength absorption at 446nm	Carotene (ppm)
Ripe	0.1044	1.095200	1004.46
Over Ripe	0.1019	0.566280	532.10
Under Ripe	0.1023	0.431280	403.67

3.2 Training Function.

The best training function of carotene oil palm analysis is the *'trainlm'-Levenberg-Marquardt* since it reached the minimum value of MSE during the training process as shown in Table 3.2.

The summary of the optimum value for learning rate, momentum constant and number of effective nodes and training function of each types of carotene is presented in Table 3.3. In the backpropagation network, learning rate is varied accordingly from 0.01 until 1. If the learning rate is made too large, the algorithm becomes unstable. If the learning rate is too small, the algorithm takes a long times to converge [4].

Table 3.2: Effect of training functions in network model

Training Function	Ripe-MSE	Over Ripe-MSE	Unripe-MSE
trainlm	1.86E-08	2.00E-08	4.76E-10
trainbfg	3.68E-05	2.52E-06	6.99E-06
trainrp	0.000366815	8.92E-05	0.000126372
trainscg	4.17E-05	1.44E-05	3.48E-05
traincgb	2.71E-05	1.77E-05	4.88E-05
traincgf	8.35E-06	2.64E-05	2.35E-05
traincgp	5.93E-05	2.34E-05	5.88E-05

Table 3.3: Optimum value of the training parameters

Fruit palm oil / Training Parameters	Ripe	Over ripe	Under ripe
Training function	**Trainlm**	**Trainlm**	**Trainlm**
Learning rate	0.5	0.5	0.7
Momentum constant	0.1	0.4	0.8
Number of nodes	10	5	8

3.3 The Training Output.

Figure 3.2 through Figure 3.4 shows the result between the target output and the actual output of the ripe, over ripe and under ripe carotene respectively when backpropagation neural network is applied in the training process. For each types of ripeness, a value for the target outputs is fixed from one sample as the highest amount of carotene and this values is extracted at the time image is captured to make sure there is only one reference pattern spectrum that neural network has to follow. At the training output, the result shows that the target output (blue line) equals with the actual output (green line) since these two lines are propagated at the same points. The, it can be said that the backpropagation neural network with optimum values of training parameters was able to train the training process with higher level of performance of the network in order to match the actual outputs with the desired outputs (target output). Hence this network model was able to improve the actual outputs of the carotene instead of applying the model with no optimum values.

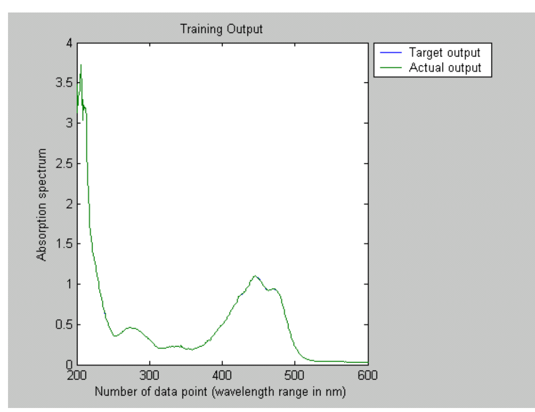

Figure 3.2: The actual and the target outputs of ripe oil palm.

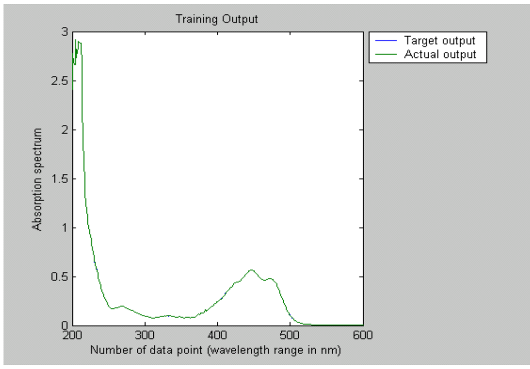

Figure 3.3: The target and actual outputs of over ripe oil palm.

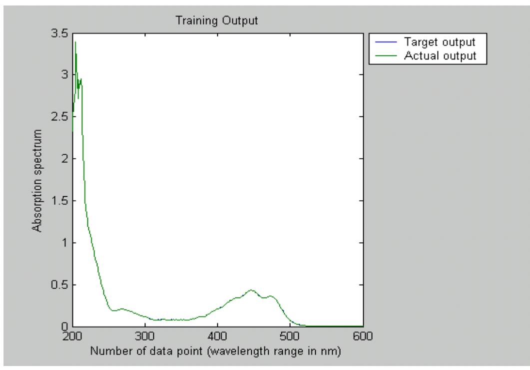

Figure 3.4: The actual and the target outputs of under ripe oil palm.

3.4 Error between actual and target output.

Figure 3.5 through Figure 3.7 shows the error between the target output and the actual output for each ripeness of oil palm. Table 3.4 show that the error at point 446nm as a reference to determined the carotene content. From this error, it's can conclude that the accurate of network model. The smaller error, -0.0002 for ripe, –0.0005 over ripe oil palm while the under ripe oil palm is large error, 1.1217. From the calculation determined the carotene content, type of ripe oil palm is more amount carotene than the over ripe oil palm. It is because ripe fruit is the better fruit containing the carotene, while the over ripe fruit is less a carotene because it is not fresh compared the ripe fruit and has more worst and useless part.

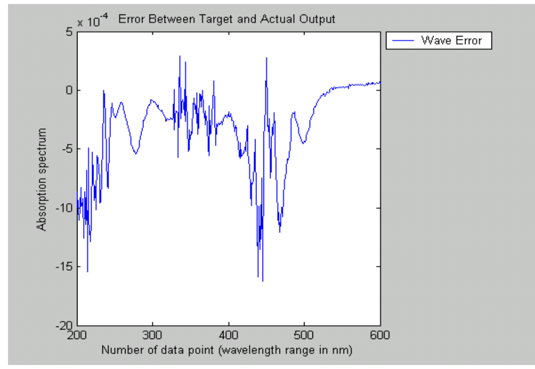

Figure 3.5: Error between the actual and the target outputs -ripe oil palm.

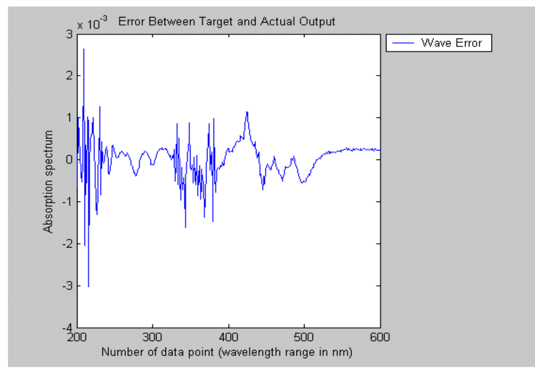

Figure 3.6: Error between the actual and the target outputs -over ripe oil palm.

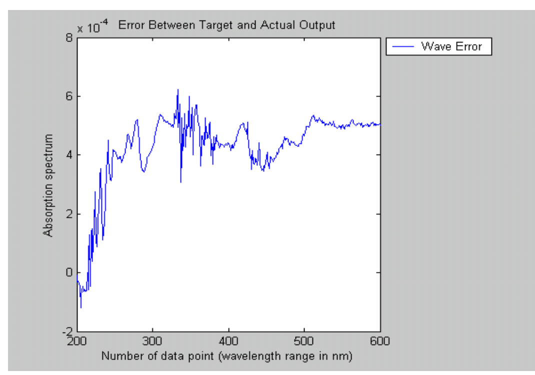

Figure 3.7: Error between the actual and the target outputs -under ripe oil palm.

Table 3.4: Error between target and actual output at point 446 nm.

Type of Ripeness	Absorption at 446 nm	
	Basic Point	Error in ANN
Ripe	1.095200	-0.0002
Over ripe	0.566280	-0.0005
Under ripe	0.431280	0.1217

3.5 Mean Square Error.

Figure 3.8 and Figure 3.9 shows the MSE versus training epoch of ripe and over ripe carotene-both the training data MSE and validation data MSE curves are shown. The MSE of ripe and over ripe carotene reached 1.9135e-008 and 1.82864e-008 respectively before a goal stop occurred. The regression analysis between the network response and the corresponding target for each ripeness. There was a perfect correlation between the target and actual outputs since the correlation coefficient-R-value is equal to 1. Therefore the performance of the trained network can be determined by the MSE, error between the actual and target output and the regression analysis

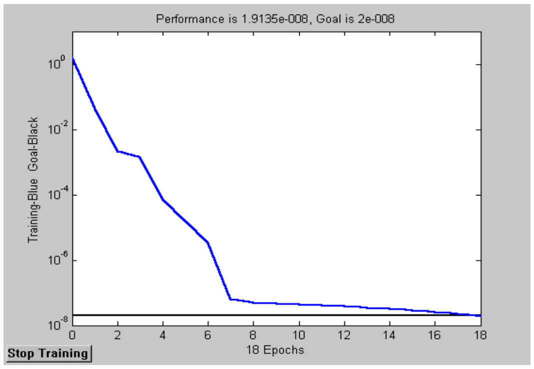

Figure 3.8: MSE versus training epochs-ripe oil palm.

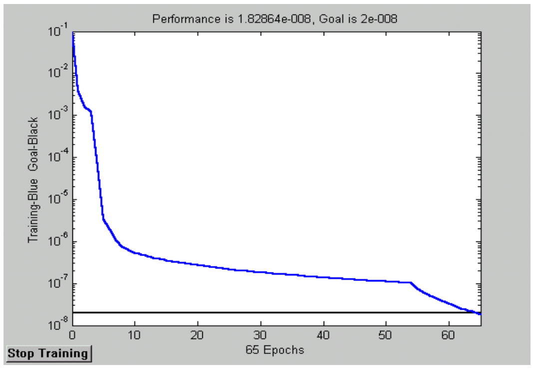

Figure 3.9: MSE versus training epochs-over ripe oil palm.

3.7 The Regression Analysis.

Figure 3.10 and Figure 3.11 shows the regression analysis between the network response and the corresponding targets of ripe-over ripe and ripe-under ripe of oil palm. Once the optimum parameters are identified, the training process is repeated for 10 times to find the correct ripeness of oil palm and Mean Correlation Coefficient-R-value as performed it shown in Table 3.5. The correlation coefficient (R-value) between the outputs and the targets is observed from the trained network.

Table 3.4 and Table 3.6 show error between target and actual output at point 446nm. Hence, the error for ripeness analysis is equal to smaller value, it can be said that network model with optimum parameters value was perfect of ripeness.

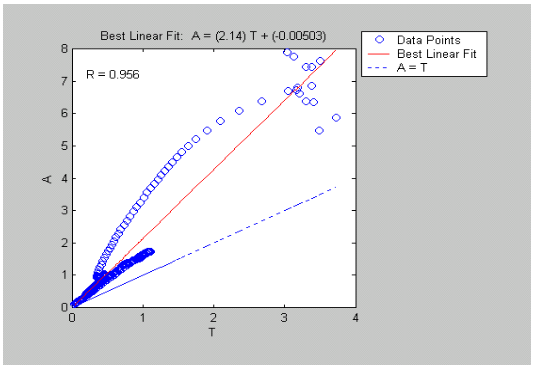

Figure 3.10 Regression analyses between the targets and actual outputs of ripe versus over ripe oil palm.

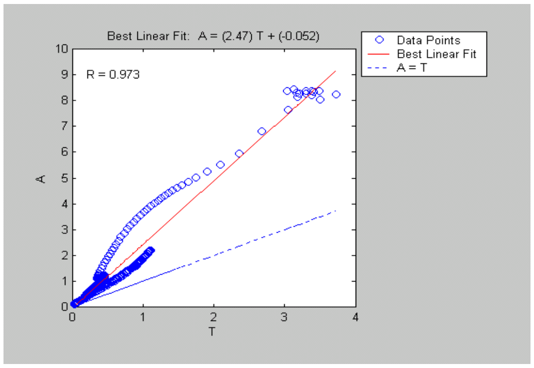

Figure 3.11: Regression analyses between the targets and actual outputs of ripe versus under ripe oil palm.

Table 3.5: Mean Correlation Coefficient (R-value) between target and actual output of ripeness

Trial	Correlation Coefficient (R-value)	
	Ripe versus Over ripe	Ripe versus Unripe
1	0.897	0.924
2	0.925	0.898
3	0.811	0.878
4	0.940	0.944
5	0.942	0.973
6	0.955	0.939
7	0.953	0.764
8	0.902	0.965
9	0.956	0.961
10	0.875	0.949
% R value	91.56	91.95

Hence, the MCC for ripeness analysis is equal to:

(91.56+91.95)/2 = 91.76%

The result of the regression line shows that the correlation (R value) between targets and actual outputs of each of the carotene oil palm was good since the values are between 0.764 to 0.973. Therefore it can be said that the network model with optimum parameters value was able to analyze each ripeness of oil palm according to the pattern spectrum.

Table 3.5: Error between target and actual output at point 446nm

Trial	Error between target output and Actual output	
	Ripe versus Over ripe	Ripe versus Unripe
1	1.2821	0.9915
2	1.1833	0.6087
3	0.7101	1.1176
4	1.1534	0.9633
5	1.2037	1.1146
6	1.1765	1.0582
7	1.2571	0.6695
8	1.5946	0.9804
9	1.1726	1.2124
10	1.2422	0.9079
Mean Error value	1.1976	0.9624

4. DISCUSSION AND CONCLUSION

The analysis of carotene content from ripeness oil palm in this project was presents successfully. Spectrophotometer is the ideal devices to produce the pattern spectrum of ripeness by absorbing the amount energy of light in a certain wavelength. Three different of pattern spectrum was determined for ripe, over ripe and under ripe oil palm. The maximum value of absorption spectrum at point 446nm for each types of ripeness was able to predict the condition of the carotene content and this value shows that the highest carotene content in ripe oil palm. Hence the ripe oil palm is state more clear than other ripeness.

From the analysis, the neural network models with the optimum values of training parameters and 'trainlm' training function were able to improve the network performances and hence the accurate desired outputs of trained network can be determined. Finally the network model was able to perform well in determined ripeness analysis with smaller error value, -0.0002 and high carotene content in ripe of oil palm fruit.

5. FUTURE DEVELOPMENT

The study of the carotene content analysis can be further analyzed by focusing for one bunch fruit of Tenera species. It can be investigate for Dura and Pisifera species too. Other software such as Fuzzy Logic also can be used for analysis.

6. ACKNOWLEDGEMENTS

I would like to take this opportunity to thank to my project supervisor, Puan Zuriati Janin for her support, guidance, ideas and suggestions during the progress of this project. Also a special thank to En Razali bin Abdul Hadi for his co-operation.

7. REFERENCES

[1] Laurene Fausett, "*Fundamentals of Neural Networks*" Architectures, Algorithms and Applications,

[2] Malaysia, "*The Oil Palm in M*alaya", 1966.

[3] The MathWorks-Neural Network Toolbox-Supported Network Architectures and Training and Learning Functions.http://www.mathworks.com/backpropagation/MSE (10 Jan. 2004)

[4] The MathWorks-Neural Network Toolbox-Supported Network Architectures and Training and Learning Functions.http://www.mathworks.com/backpropagation/transferfunction (10 Jan. 2004)

[5] http://www.burotrop.org/bulletin/14/oilpal.pdf

[6] Siew Wai Lin, Tang thin Sue and Tan Yew Ai, "*PORIM TEST METHOD*", 1995.

[7] http://www.dacs.dtic.mill/techs/neural/html