

The Prediction of Chlorophyll Content in African Leaves (*Vernonia amygdalina Del.*) Using *Flatbed Scanner* and Optimised Artificial Neural Network

Retno Damayanti*, Nurul Rachma, Dimas Firmanda Al Riza and Yusuf Hendrawan

Department of Agricultural Engineering and Biosystem, Faculty of Agricultural Technology, Universitas Brawijaya, Jl. Veteran, Malang, East Java, 65145 Indonesia

ABSTRACT

African leaves (*Vernonia amygdalina Del.*) is a nutrient-rich plant that has been widely used as a herbal plant. African leaves contain chlorophyll which identify compounds produced by a plant, such as flavonoids and phenols. Chlorophyll testing can be carried out non-destructively by using the SPAD 502 chlorophyll meter. However, it is quite expensive, so that another non-destructive method is developed, namely digital image analysis. Relationships between chlorophyll content and leaf image colour indices in the RGB, HSV, HSL, and Lab* space are examined. The objectives of this study are 1) to analyse the relationship between texture parameters of red, green, blue, grey, hue, saturation(HSL), lightness (HSL), saturation(HSV), value(HSV), L*, a*, and b* against the chlorophyll content in African leaves using a flatbed scanner (HP DeskJet 2130 Series); and 2) built a model to predict chlorophyll content in African leaves using optimised ANN through a feature selection process by using several filter methods. The best ANN topologies are 10-30-40-1 (10 input nodes, 40 nodes in hidden layer 1, 30 nodes in hidden layer 2, and 1 output node) with a trainlm on the learning function, tansig on the hidden layer, and purelin on the output layer. The selected topology produces MSE training of 0.0007 with

R training 0.9981 and the lowest validation MSE of 0.012 with R validation of 0.967. With these results, it can be concluded that the ANN model can be potentially used as a model for predicting chlorophyll content in African leaves.

Keywords: African leaves, artificial neural network, chlorophyll, flatbed scanner

ARTICLE INFO

Article history:

Received: 01 April 2021

Accepted: 12 July 2021

Published: 08 October 2021

DOI: <https://doi.org/10.47836/pjst.29.4.15>

E-mail addresses:

damayanti@ub.ac.id (Retno Damayanti)

nrlrachma@student.ub.ac.id (Nurul Rachma)

dimasfirmanda@ub.ac.id (Dimas Firmanda Al Riza)

yusufhendrawan@gmail.com (Yusuf Hendrawan)

* Corresponding author

INTRODUCTION

African leaves (*Vernonia amygdalina*) is a medicinal plant and belongs to the Asteraceae family. African leaf is also called bitter leaf since it has a bitter smell and taste (Danladi et al., 2018). African leaf can grow to a height of more than 10 m with a stem diameter of 40 cm, and its bark will be grey to brown. The leaves are green, elliptical in shape, about 4-15 × 1-4 cm in size, 0.2-4 cm long petiole, and serrated at the edges. The cultivation of African leaves is usually carried out through stem cuttings from the parent plants that are more than one year old (Nursuhaili et al., 2019). As a medicinal plant, African leaves are widely used as an antifeedant, antischistosomal, antiplasmodial, antioxidant, anti-inflammatory, antidiabetic and anticancer drug (Oyeyemi et al., 2017). Chlorophyll is a green leafy substance in plants that absorb and convert light energy into chemical energy (Pavlovic et al., 2014). Chlorophyll absorbs light, especially red and blue, from the visible light spectrum (Luimstra et al., 2018). Green colour (around 550 nm) is not absorbed but is reflected to give chlorophyll a special colour (Rajalakshmi & Narasimhan, 2013).

The results of the previous research stated that chlorophyll content and leaf ageing levels could be used to identify the content of compounds produced by plants, including flavonoids, phenols (Abdulkadir et al., 2015), nitrogen (Cartelat et al., 2005), and can be used as a parameter for the quality of green vegetables during storage (Limantara et al., 2015). Common chlorophyll conventional testing is conducted by extracting a chlorophyll before it then measures by UV-VIS spectrophotometry. Besides than that, utilising a chlorophyll meter is another method to measure chlorophyll content (Uddling et al., 2007). Chlorophyll meter measures transmitted light of the leaves at 650 nm and 940 nm then processes it, using a microprocessor, to measure the amount of chlorophyll content (Borhan et al., 2017). Even though several conventional chlorophyll testing has shown a proven result, there are several drawbacks such as being destructive, requires a lengthy process and being costly. Therefore, to overcome these issues, this research proposes another method that is non-destructive, rapid and cheaper computer vision assessment; this includes colour and texture analysis (Hendrawan et al., 2019a). Digital colour analysis for plant leaf colour has become an increasingly popular and effective method for evaluating foliar nutrition and health in response to environmental stresses (Hu et al., 2010). In this study, the computer vision approach uses a flatbed scanner for image acquisition. The machine vision method is very effective because it will not destruct the object. A flatbed scanner is considered fast, assessable on the market, cheap, and not dependent on external light conditions (Dalen, 2006). The flatbed scanner has recently been widely used for image analysis purposes, including rice classification (Dalen, 2006), leaf surface area estimation (Kaur et al., 2014), prediction of chlorophyll content in potato plants (Yadav et al., 2010), and testing the purity of rice seeds (Widiastuti et al., 2018).

Detecting green leaves in a natural environment is more difficult since the leaves are similar to the background in colour. Leaf colour is a good indicator of plant health, and it

can be used to predict the leaf's chlorophyll. The chlorophyll can be predicted by analysing the distribution of colour components [Red (R), Green (G), Blue (B), Hue(H), Saturation (S), and Intensity (I)] of the leaf images. Texture analysis using the grey level co-occurrence matrix (GLCM) method is dominant at the grey level, The GLCM method is developed on colour images, known as the colour co-occurrence matrix (CCM) method. Based on the research of (Hendrawan et al., 2019b), colour texture analysis (CCM) provides additional image characteristics above the grey level representation (GLCM). It has been proven to measure biological objects effectively. Artificial neural network (ANN) modelling with the backpropagation (BP) learning algorithm is utilised in applying this method. The BP algorithm is a simple iterative learning algorithm that works well on complex data. During the training, weights are arranged iteratively to minimise errors (Setti & Anjar, 2018). Image analysis and the ANN method have now been widely applied to identify chlorophyll. Grunenfelder et al. (2006) used colour indices to assess the chlorophyll development and greening of fresh market potatoes. Barman and Choudhury (In Press) analyse the distribution of colour to predict the chlorophyll of citrus leaf. Hassanijalilian et al. (2020) estimate chlorophyll of soybean leaves infield with smartphone digital imaging and machine learning. Mohan and Gupta (2019) predict chlorophyll content on rice and obtained more efficient results than using linear regression models Samli et al. (2014) applied ANN to predict chlorophyll concentrations *a*, Gupta and Pattanayak (2017) used ANN to project chlorophyll content in potato plants non-invasively, Damayanti et al. (2020) predicted chlorophyll content in cassava leaves using ANN, and also Peng and Yi (2019) which has predicted chlorophyll content in pomegranate leaves based on digital image analysis and ANN.

However, there is no research identification a chlorophyll African leave by using machine vision yet. This study aims to: 1) analyse the relationship between the texture of the feature of red, green, blue, grey, hue, saturation_(HSL), lightness_(HSL), saturation_(HSV), value_(HSV), L*, a*, and b* on the chlorophyll content of African leaves using a flatbed scanner (HP DeskJet 2130 Series); and 2) develop a model to predict chlorophyll content in African leaves using ANN which has been optimised through the feature selection process by using multiple filter methods.

MATERIAL AND METHODS

In order to process the data, Lenovo Z40-72 Laptop was utilised with AMD A10-7300 processor specifications (1.9 GHz/Turbo 3.2 GHz) and 4 GB DDR3 @ 1600 MHz memory capacity; flatbed scanner of HP DeskJet 2130 series to acquire digital image imagery; chlorophyll meter SPAD 502 with a measuring area of 2×3 mm and a maximum sample thickness of 1.2 mm for *real-time* measurement of chlorophyll content of leaves as comparative data; Paint Microsoft software to convert image formats to bitmap (BMP)

and to change image resolution; self-built software based on Visual Basic 6.0 (Microsoft) for image feature extraction; Waikato Environment for Knowledge Analysis (WEKA 3.8) software for the feature selection process; and Matlab R2014a software for the ANN design process. African leaves, as samples, were obtained from the Merjosari area, Lowokwaru District, Malang City, East Java, Indonesia. The difference in leaf colour texture is a parameter to determine chlorophyll content. This study's leaf colour texture difference can be categorised into three parts: 1) the shoot, 2) the middle, and 3) the base. 3 to 4 leaves from the top and bottom of the plant were taken, while the middle leaves were taken from the leaves between the top and bottom of the plant. Twenty leaves were taken for each category, so that there are 60 African leaves were needed in this study.

The real-time measurement, using SPAD 502 chlorofilmeter, is considered for comparative study. The leaves are clipped to SPAD 502 chlorofilmeter, and the sensor will provide a chlorophyll content index (CCI) to detect chlorophyll. Simultaneously, chlorophyll measurement is done after image data acquisition by flatbed scanner HP DeskJet 2130 series. This scanner uses a CIS sensor (CMOS Image Sensor) which bring more benefits such as requires low power and voltage, gives low price and provides detail image because of its smaller pixel size (Shakeri et al., 2012). The flatbed scanner is considered assessable on the market, easy to use, stable when acquiring images and does not depend on external light conditions (Widiastuti et al., 2018) and low cost. The acquired image is saved in bitmap format. The data augmentation process was carried out on 60 images. Data augmentation is one of the processes in image processing to increase the number of samples and improve modelling performance (Shorten & Taghi, 2019). Various augmentation techniques include cutting (*cropping*), rotation, illumination, scaling and colouring (Okafor et al., 2018). In this study, the cutting and rotating images technique of African leaves were used. A total of 60 images that have been acquired were then divided into two parts by using Paint software with an image resolution of 300×300 pixels comprising 120 images of African leaves (Figure 1). Furthermore, the rotation process (at an angle of 0°, 90°, 180° and 270°) from 120 pictures was conducted (Figure 2), providing 480 total images.

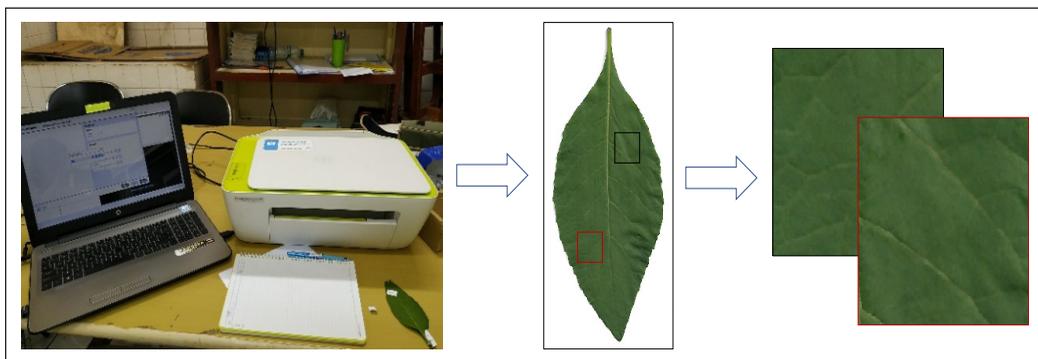


Figure 1. Image acquisition and augmentation process of cutting images of African leaves

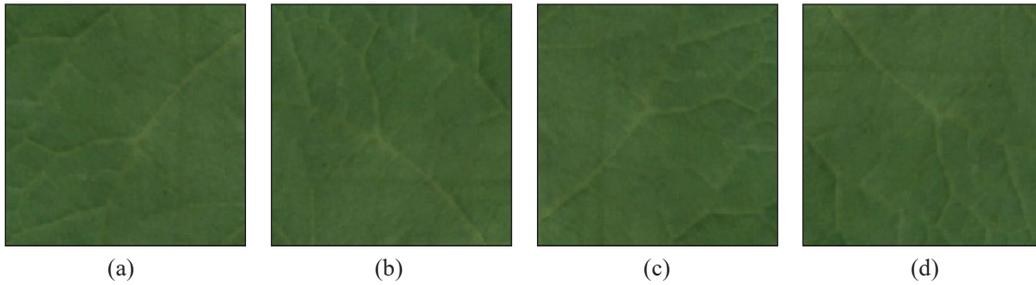


Figure 2. The results of rotating augmentation of African leaf image images on rotation: (a) 0°, (b) 90°, (c) 180°, (d) 270°

The value of the textural features of the acquired image saved in bitmap format will then be extracted. Colour texture analysis is an important method in computer vision that can be used for object recognition, surface defect detection, or pattern recognition (Armi & Shervan, 2019). The results of texture feature extraction are entropy, energy, contrast, homogeneity, sum mean, variance, correlation, maximum probability, inverse difference moment and cluster tendency on each colour-space which includes red, green, blue, grey, hue, saturation_(HSL), lightness_(HSL), saturation_(HSV), value_(HSV), L*, a* and b* (Hendrawan et al., 2018). The ten texture features used in this study are based on Harlick (Haralick et al., 1973) texture Equations 1-10 (Hendrawan & Haruhiko, 2009):

$$Entropy = - \sum_i^M \sum_j^N P [i, j] \log P [i, j] \quad (1)$$

$$Energy = \sum_i^M \sum_j^N P^2 [i, j] \quad (2)$$

$$Contrast = \sum_i^M \sum_j^N (i - j)^2 P [i, j] \quad (3)$$

$$Homogeneity = \sum_i^M \sum_j^N \frac{P [i, j]}{1 + |i - j|} \quad (4)$$

$$Sum\ Mean = \frac{1}{2} \sum_i^M \sum_j^N (iP [i, j] + jP [i, j]) \quad (5)$$

$$Variance = \frac{1}{2} \sum_i^M \sum_j^N ((i - \mu)^2 P [i, j] + (j - \mu)^2 P [i, j]) \quad (6)$$

$$Correlation = \sum_i^M \sum_j^N \frac{(i - \mu)(j - \mu)P [i, j]}{\sigma^2} \quad (7)$$

$$Inverse\ Difference\ Moment = \sum_i^M \sum_j^N \frac{P [i, j]}{|j - k|^k} \quad i \neq j \quad (8)$$

$$Cluster\ Tendency = \sum_i^M \sum_j^N (i + j - 2\mu)^k P [i, j] \quad (9)$$

$$Maximum\ Probability = Max_{i, j}^{M, N} P [i, j] \quad (10)$$

Where: $P_{(i, j)}$ is the element of (i, j) normalised *co-occurrence matrix*, μ and σ are *mean* and the *standard deviation* of pixel elements are as in Equations 11-13:

$$P[i, j] = \frac{N(i, j)}{M} \quad (11)$$

$$\mu = \sum_i^M i \sum_j^N P[i, j] \quad (12)$$

$$\sigma = \sum_i^M (i - \mu)^2 \sum_j^N P[i, j] \quad (13)$$

Where: $N_{(i,j)}$ is the amount calculated in the image with the pixel intensity i followed by the pixel intensity j at a displacement of one pixel to the left, and M is the total number of pixels.

The definition of texture features (Hendrawan et al., 2019b) are as follows: entropy measures the randomness of the grey level distribution; *energy* measures the number of repeated pairs or measures the level of texture uniformity; contrast measures the local contrast in the image; homogeneity measures the local homogeneity of a pixel pair; sum mean calculates the average grey level of a colour image; variance shows the distribution of the grey level distribution; correlation shows the correlation between two pixels in a pixel pair; different inverse moment shows the refinement of the image; cluster tendency measures the grouping of images that has similar grey levels, and maximum probability is the result of the most dominant pixel pair in the image.

From the feature extraction process, 120 texture features were obtained consisting of 10 texture features for red, green, blue, grey, hue, saturation_(HSL), saturation_(HSV), value, lightness, L^* , a^* and b^* then sorted from 1 to 120 orderly. Next, feature selection (a preprocessing image data) using WEKA 3.8 to select ANN input. These methods will delete outliers and chose highly important data features to fit the requirement of the learning function (Garner, 1995). In this research, feature selection uses the filter method consisting of Chi-Squared Attribute Evaluator, Correlation Attribute Evaluator, ReliefF Attribute Evaluator and Gain Attribute Evaluator. Feature selection is considered an effective and efficient procedure for optimising data mining and machine learning. The purpose of feature selection is to reduce irrelevant, redundant, and noise features to get better learning performance, produce higher accuracy and produce a simpler model (Wang et al., 2016). The method does not rely on learning algorithms and data characteristics, and this is not affected by other features but solely based on rank (Li et al., 2017). The output from feature selection was then used as ANN input to produce the model.

Network topology was the design, and the lowest mean square error (MSE) validation data was set. Four hundred eighty digital images were divided into 75% (360 images) and 25% (120 images) for training data and validation data, respectively, to prevent overfitting (Xu & Royston, 2018). Input data, taken from feature selection, were then modelled using ANN (Matlab R2014a) to predict the chlorophyll content of African leaves. However, after feature selection performing another preprocessing data, in this case, was normalisation, is important to standardise the data scale. The input and output data were switched on a scale of -1 to +1 to avoid a significant weight change during the ANN training process.

Sensitivity analysis was carried out by using variations of the learning rate, momentum, the number of nodes on the hidden layer and the number of hidden layers. This study used 10 types of learning functions to obtain a suitable learning function to predict the chlorophyll content of African leaves.

RESULTS AND DISCUSSION

Based on the measurement of chlorophyll content, the difference of each leaf colour texture reflected its chlorophyll content (Figure 3). Chlorophyll content contributed to the greenish level of the leaf, meaning that the denser the green colour, the more chlorophyll on the leaf. On the contrary, pale green or yellowish-green leaves expressed low chlorophyll content. Therefore, this parameter was important to measure chlorophyll content using computer vision (Barman et al., 2018). Based on Figure 4, it can be seen that the increase chlorophyll content follows the increase of difference in leaf colour. The average chlorophyll content in (1) leaves' shoot was 41.2 CCI; (2) the middle part of the leaves was 46.3 CCI; and (3) leaves' base was 51.2 CCI. Therefore, the insignificant green colour difference on African leaves was equated to the low level of chlorophyll interval. According to Pavlovic et al. (2014), changes in chlorophyll content can be influenced by several factors, including physiological, morphological (age and position of leaves) and abiotic (temperature, relative humidity, and light quality) factors.

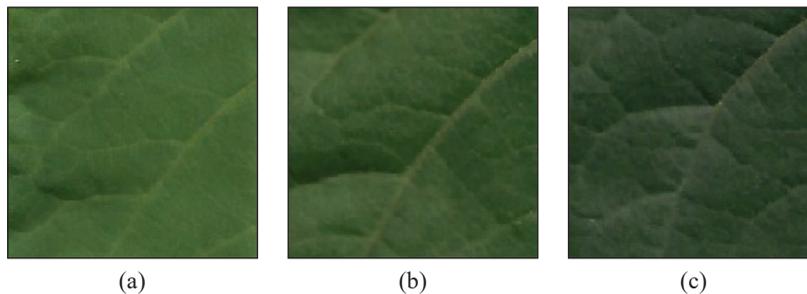


Figure 3. The results of image acquisition of African leaf images: (a) the shoot, (b) the middle, (c) the base

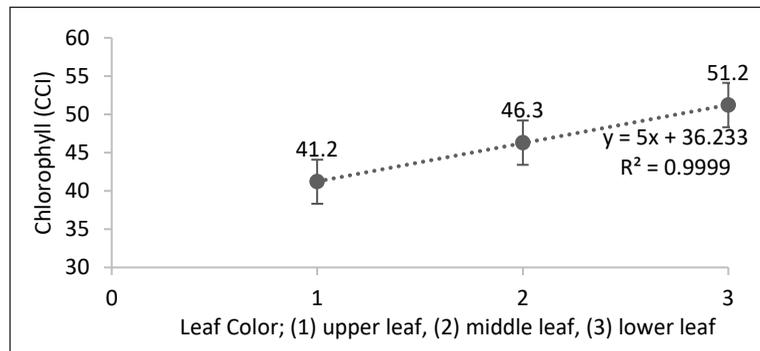


Figure 4. Graph of the relationship between chlorophyll content and leaf colour texture category

The augmented image data was then proceeded to feature extraction, and the results were 120 textural features. However, only selected textural features with the filter method proposed a fast, efficient and non-dependent learning algorithm to predict chlorophyll content in African leaves (Kumar & Rama, 2014; Mendoza et al., 2018). This study utilised the filter method based on four evaluation attributes: chi-squared attribute evaluator, correlation attribute evaluator, ReliefF attribute evaluator, and gains attribute evaluator. The top 10 best feature for each evaluation attribute was chosen. Table 1 shows the feature selection for each evaluation attribute used. Table 1 illustrates $b^*_{(Lab)}$ Energy with the evaluation attribute ReliefF has a strong correlation to chlorophyll content in African leaves with weight accounting for 0.3431. After obtaining the weight and ranking for each feature, the features selected in Table 1 were modelled using ANN using the trial and error method to define the best combination of features-subset which later for chlorophyll prediction; the lowest validation MSE was also considered as a robust model. Trial and error on ANN modelling on preliminary research generated the best ANN topology, consisting of 30 nodes on the first hidden layer, 40 nodes on the second hidden layer, with the learning function of trainlm, tansig activation function in each hidden layer, and purelin on the output layer; learning rate 0.1; and momentum 0.5 (Hendrawan et al., 2019c). Table 2 shows that the entire 120 texture features configuration created higher validation MSE compare with feature selection. Compare with three other evaluation attribute (Table 2), ReliefF produces the lowest validation MSE using 10 ANN input was 0.0025. The highest validation MSE is in the evaluation attribute Gain Ratio with 2 ANN input was 0.7469. Therefore, feature selection was important to optimise the performance of the ANN model. Based on these results, 10 textural features in the evaluation attribute ReliefF were used as input of ANN modelling to predict chlorophyll content in African leaves. Moreover, ReliefF gives more benefits such as dealing with nominal or continuous features, handling lost data and tolerating noise (Mendoza et al., 2018).

From the trial and error results on the feature selection process (Table 2), we obtain the best combination of 10 textural features that have a high correlation against the chlorophyll content of African leaves. The ten textural features include b^* energy, b^* correlation, b^* entropy, $\text{saturation}_{(HSV)}$ sum mean, $\text{saturation}_{(HSL)}$ sum mean, blue sum mean, sum mean hue, maximum hue probability, correlation hue, and red correlation. Figure 5a shows the relationship of chlorophyll content to b^* energy. The results stated that the value of b^* energy decrease with the increase in chlorophyll content. Another thing, the low b^* energy value indicated that the uniformity of texture on African leaves was also low. Figure 5a also tells that the value of b^* energy has a negative correlation with a determination coefficient of 0.0297. As for the relationship of chlorophyll content to the b^* correlation (Figure 5b), b^* correlation increase with the increase of chlorophyll content, the increase of b^* correlation showed that the correlation among pixels on African leaves was high.

Table 1
Result of feature selection using the filter method

No.	Attribute Evaluator	Search Method	Textural features	Weight	Rank
1.	<i>Chi-Squared Attribute Evaluator</i>	<i>Ranker</i>	b* Energy	1568.4034	1
			b* Entropy	1040.8304	2
			b* Correlation	1004.3086	3
			Saturation _(HSV) Energy	558.246	4
			Hue Energy	496.8733	5
			Hue Entropy	476.624	6
			a* Entropy	388.9211	7
			a* Energy	373.5219	8
			Saturation _(HSV) Sum Mean	373.3694	9
			a* Maximum Probability	354.2468	10
2.	<i>Correlation Attribute Evaluator</i>	<i>Ranker</i>	b* Energy	0.1554	1
			b* Entropy	0.1187	2
			Hue Entropy	0.0881	3
			b* Inverse	0.0862	4
			Saturation _(HSV) Sum Mean	0.0857	5
			Hue Energy	0.0827	6
			b* Correlation	0.0822	7
			a* Maximum Probability	0.0777	8
			Red Entropy	0.0762	9
			Red Maximum Probability	0.076	10
3.	<i>ReliefF Attribute Evaluator</i>	<i>Ranker</i>	b* Energy	0.3431	1
			b* Correlation	0.2717	2
			b* Entropy	0.2643	3
			Saturation _(HSV) Sum Mean	0.2543	4
			Saturation _(HSL) Sum Mean	0.2205	5
			Blue Sum Mean	0.211	6
			Hue Sum Mean	0.2055	7
			Maximum hue probability	0.205	8
			Hue Correlation	0.2026	9
			Red Correlation	0.2007	10
4.	<i>Gain Ratio Attribute Evaluator</i>	<i>Ranker</i>	b* Energy	0.4125	1
			b* Entropy	0.319	2
			b* Correlation	0.2209	3
			Hue Entropy	0.1669	4
			Saturation _(HSV) Sum Mean	0.1604	5
			Blue Maximum Probability	0.1571	6
			b* Inverse	0.1559	7
			Hue Energy	0.1534	8
			Red Maximum Probability	0.1483	9
			Saturation _(HSV) Energy	0.1453	10

Table 2
Trial and error ANN performance using textural feature selection

No.	Attribute Evaluator	Search Method	Input	MSE training	MSE validation
1.	-	-	All features (120 inputs)	0.0094	0.5193
2.	Chi-Squared Attribute Evaluator	Ranker	Feature rank 1-2	0.0097	0.7465
			Feature rank 1-3	0.0097	0.0818
			Feature rank 1-4	0.0093	0.0174
			Feature rank 1-5	0.0089	0.0136
			Feature rank 1-6	0.0063	0.0121
			Feature rank 1-7	0.0098	0.0149
			Feature rank 1-8	0.0058	0.0117
			Feature rank 1-9	0.0060	0.0076
			Feature rank 1-10	0.0033	0.0046
			3.	Correlation Attribute Evaluator	Ranker
Feature rank 1-3	0.0098	0.0063			
Feature rank 1-4	0.0099	0.1236			
Feature rank 1-5	0.0090	0.0254			
Feature rank 1-6	0.0066	0.0237			
Feature rank 1-7	0.0055	0.0249			
Feature rank 1-8	0.0090	0.0165			
Feature rank 1-9	0.0094	0.0166			
Feature rank 1-10	0.0092	0.0166			
4.	ReliefF Attribute Evaluator	Ranker			
			Feature rank 1-3	0.0098	0.0624
			Feature rank 1-4	0.0090	0.0172
			Feature rank 1-5	0.0081	0.0159
			Feature rank 1-6	0.0099	0.0125
			Feature rank 1-7	0.0016	0.0026
			Feature rank 1-8	0.0083	0.0092
			Feature rank 1-9	0.0099	0.0108
			Feature rank 1-10	0.0015	0.0025
			5.	Gain Ratio Attribute Evaluator	Ranker
Feature rank 1-3	0.0097	0.0818			
Feature rank 1-4	0.0099	0.0217			
Feature rank 1-5	0.0091	0.0160			
Feature rank 1-6	0.0064	0.0099			
Feature rank 1-7	0.0088	0.0225			
Feature rank 1-8	0.0077	0.0211			
Feature rank 1-9	0.0092	0.0265			
Feature rank 1-10	0.0069	0.0220			

Moreover, b^* correlation has a positive correlation with a coefficient of determination of 0.1593. When it comes to the relationship between b^* entropy and chlorophyll (Figure 5c), b^* entropy increases with the increase of chlorophyll content. The higher b^* entropy value indicated that the roughness of the texture of the African leaves was high. From Figure 5c, it can also be seen that b^* entropy has a positive correlation with a determination coefficient of 0.0332.

In Figure 5d, it can be seen that the relationship of chlorophyll content to sum mean saturation_(HSV). The results showed that the sum mean saturation_(HSV) decrease with the increased chlorophyll content, the decreasing sum mean saturation_(HSV) shows that the level of grey saturation_(HSV) in African leaves was low. From Figure 5d, it can also be seen that sum mean saturation_(HSV) has a negative correlation with a determination coefficient of 0.3449.

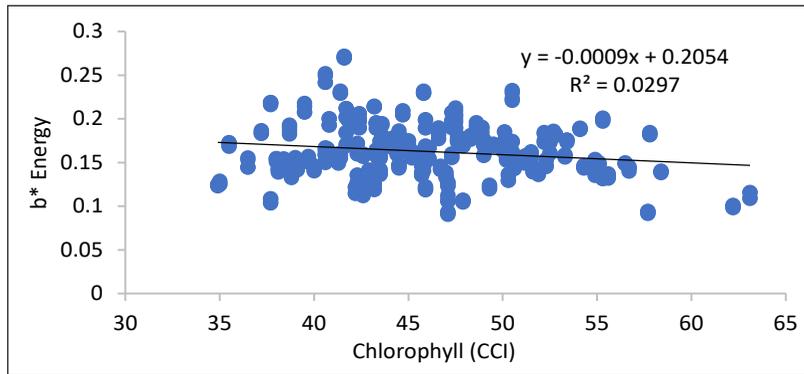
In Figure 5e, it can be seen the relationship of chlorophyll content to sum mean saturation_(HSL). The results showed that the sum mean saturation_(HSL) decreased with the increase in chlorophyll content. The decreasing sum mean saturation_(HSL) indicates that the saturation_(HSL) of the grey level of African leaves is low. From Figure 5e, it can also be seen that sum mean saturation_(HSL) has a negative correlation with a determination coefficient of 0.3376.

In Figure 5f, it can be seen the relationship of chlorophyll content to blue sum mean. The results showed that the value of the sum mean blue decreased with the increase in chlorophyll content. Decreasing the blue_(RGB) sum mean indicates that the level of blue in African leaves is low. From Figure 5f, it can also be seen that the blue sum mean has a negative correlation with a determination coefficient of 0.208.

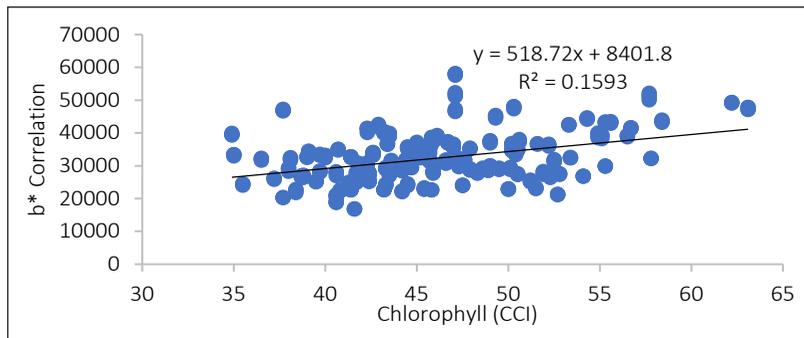
Figure 5g, it can be seen the relationship of chlorophyll content to hue sum mean. The results showed that the hue sum mean value decrease with the increase in chlorophyll content. The decrease of the hue sum mean showed that the grey level of hue on African leaves is low. From Figure 5g, it can also be seen that the hue sum mean has a negative correlation with a determination coefficient of 0.0557.

In Figure 5h, it can be seen the relationship of chlorophyll content to the maximum hue probability. The results show that the maximum hue probability value decreases with the increase in chlorophyll content. The decrease of maximum hue probability value indicated that the dominant hue pixel pair is getting lower or irregular on African leaves. In Figure 5h, it can also be seen that the maximum hue probability has a negative correlation with a coefficient of determination of 0.1022. In Figure 5i, it can be seen the relationship of chlorophyll content to hue correlation. The results showed that the value of hue correlation increased with the increase in chlorophyll content. The increase of hue correlation showed that the correlation between hue on African leaves was high. From Figure 5i, it can also be seen that the hue correlation has a positive correlation with a determination coefficient of 0.1648.

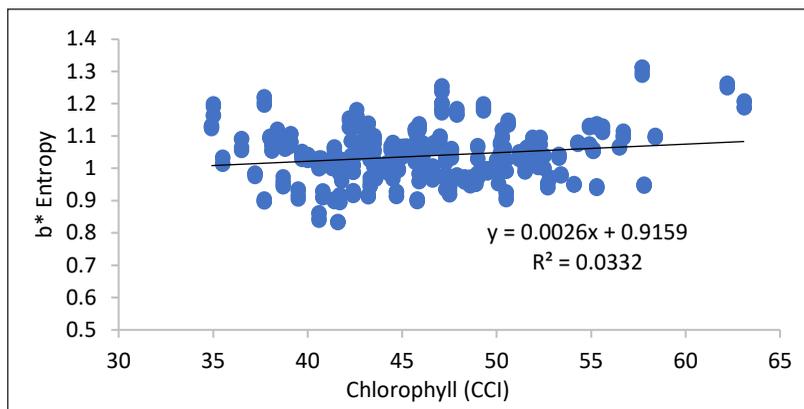
In Figure 5j, it can be seen the relationship of chlorophyll content to red correlation. The results showed that the red correlation value increase with the increase in chlorophyll content. The increasing red correlation value showed that the correlation between red on African leaves was high. Hendrawan and Murase (2009) have stated that during the photosynthesis process, plants absorb red wavelengths to reflect more green wavelengths.



(a)

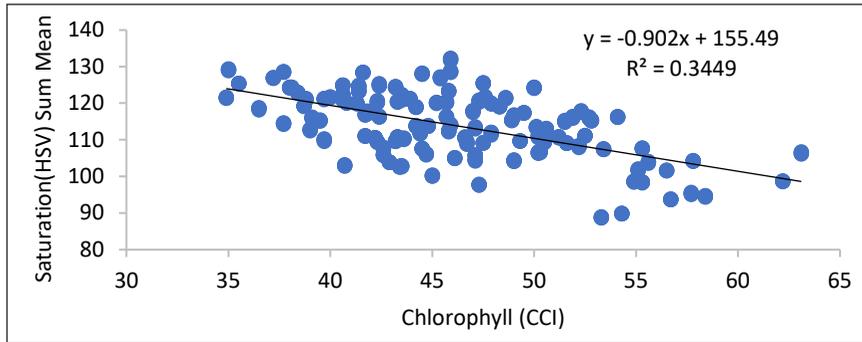


(b)

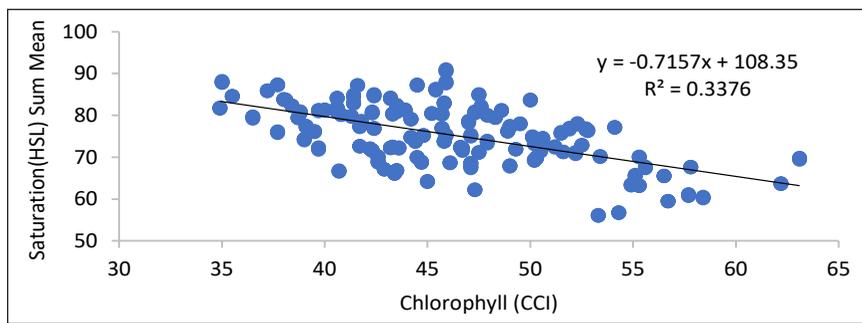


(c)

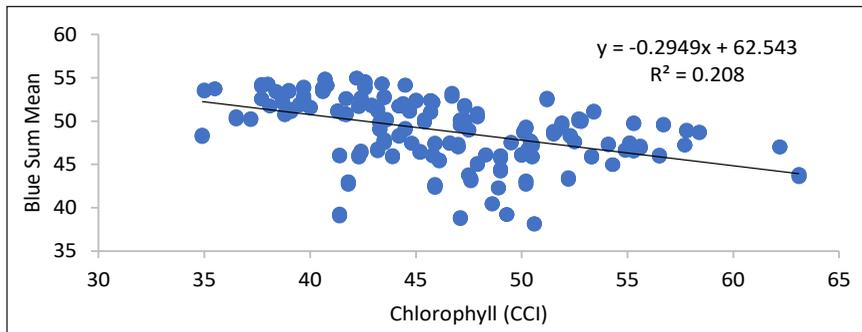
Figure 5. The relationship between chlorophyll content on the 10 best textural features: (a) b* energy, (b) b* correlation, and (c) b* entropy,



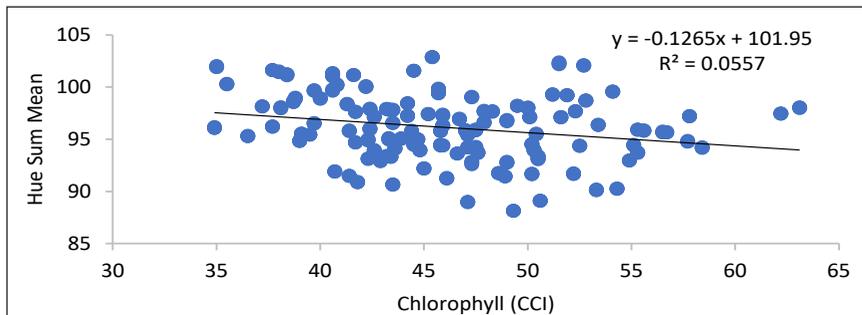
(d)



(e)

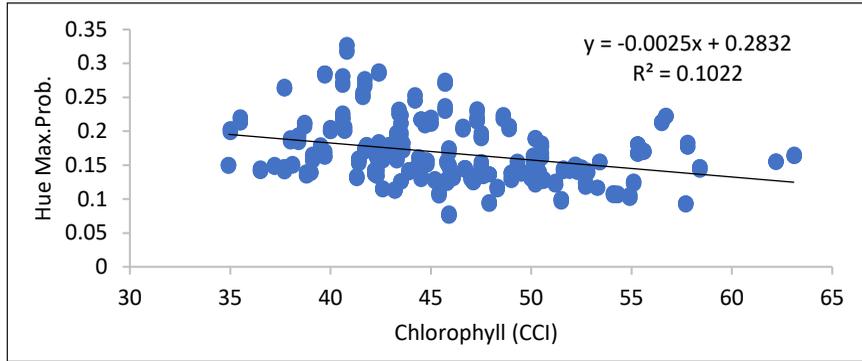


(f)

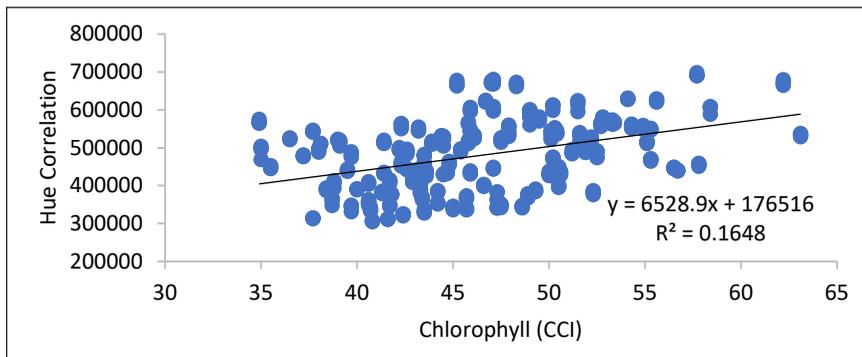


(g)

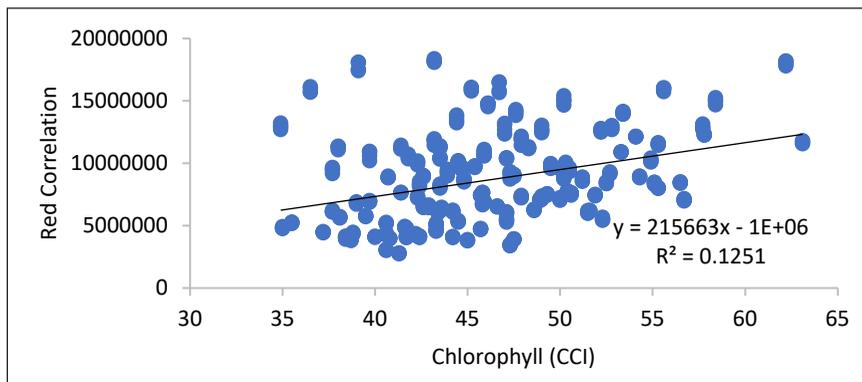
Figure 5. (continue) The relationship between chlorophyll content on the 10 best textural features: (d) saturation_(HSV) sum mean and (e) saturation_(HSL) sum mean, (f) blue sum mean, and (g) hue sum mean



(h)



(i)



(j)

Figure 5. (continue) The relationship between chlorophyll content on the 10 best textural features: (h) maximum hue probability, (i) hue correlation, and (j) red correlation.

Thus, the more red is absorbed, the higher the green colour is formed. From Figure 5j, it can also be seen that the red correlation has a positive correlation with a determination coefficient of 0.1251. Based on the linearity of the relationship between chlorophyll content on textural features, the resulting coefficient of determination is low.

In Table 3, it can be seen that the trial and error results on the learning function. Table 3 shows that trainlm (Lavenberg-Marquardt) is more accurate than other learning functions with the smallest MSE validation, which was 0.0032, and the largest validation R was 0.9925. Trainlm updates weights and biases based on Jacobian Matrix calculations. Trainlm is a simpler and more efficient learning function because the solution to matrix equations is a built-in function. The next sensitivity analysis process is the selection of activation function.

Table 3
Trial and error on the learning function selection

No.	Learning Function	R Training	R Validation	MSE Training	MSE Validation
1	Traincgb (Conjugate Gradient BP with Powell – Beale Restart)	0.9719	0.9716	0.0096	0.0099
2	Traincgf (Conjugate BP with Fletcher Reeves Update)	0.9711	0.9704	0.0099	0.0103
3	Traincgp (Conjugate Gradient BP with Polak Ribiere Update)	0.9709	0.9706	0.0100	0.0104
4	Traingd (Gradient Descent BP)	0.9794	0.9781	0.0100	0.0114
5	Traingda (Gradient Descent with Adaptive Learning Rate BP)	0.9708	0.9699	0.0100	0.0106
6	Traingdm (Gradient Descent with momentum Adaptive Learning)	0.9709	0.9702	0.0100	0.0103
7	Traingdx (Gradient Descent with Momementum Adaptive Learning)	0.9707	0.9698	0.0100	0.0105
8	Trainlm (Lavenberg Marquadt BP)	0.9934	0.9925	0.0034	0.0032
9	Trainoss (One Step Secant BP)	0.9708	0.9700	0.0100	0.0104
10	Trainrp (Resilient BP)	0.9712	0.9672	0.0098	0.0113
11	Trainscg (Scaled Conjugate Gradient BP)	0.9708	0.9708	0.0100	0.0101

The activation function is a function used to determine the output based on the input. In the ANN-BP algorithm, three common activation functions are the purelin function (linear), logsig (binary sigmoid) function, and tansig (bipolar sigmoid) function. Thus, the activation function can significantly improve the performance of ANN modelling. In trial and error, the activation function is randomly paired to obtain the lowest validation MSE.

In Table 4, the selection of activation function can be known by learning function using trainlm. Common activation functions were tansig and logsig for the hidden layer and purelin for the output layer. Moreover, this statement is in line with the result of the research, where tansig and purelin was chosen for hidden layers and output layer, respectively. It provided the lowest validation MSE (0.0032) and the highest validation R (0.9925). A study from Jaber et al. (2019) stated that the configuration of tansig and purelin resulted in the decrease of MSE training and validation equates with the increase

in epoch with the best performance on epoch 2487, which resulted in R training of 0.9995 and R validation of 0.99994.

The determination of the best ANN topology is determined by a combination of input variables, output variables, learning functions, activation functions, learning rate, momentum, number of the hidden layer and nodes in the hidden layer, which is appropriate so that it produces low MSE values and high determination coefficient (Jaber et al., 2019). However, a standard method has not been found to produce the best combination for each study by now. Therefore, to get the best topology, it is necessary to do *trial and error* on a certain subject (Kato et al., 2015). The results of *trial and error* to predict chlorophyll content in African leaves can be seen in Table 5. Table 5 shows the best topology results comprising 10-40-30-1 (10 input variables, 40 nodes on hidden layer 1, 30 nodes in

Table 4
Trial and error to determine activation function

Learning function	Activation function			R Training	R Validation	MSE Training	MSE Validation
	Hidden Layer 1	Hidden Layer 2	Output Layer				
Trainlm	Tansig	Tansig	Purelin	0.9934	0.9925	0.0034	0.0032
	Tansig	Tansig	Tansig	0.9780	0.9764	0.0078	0.0085
	Tansig	Tansig	Logsig	0.8050	0.8086	0.1647	0.1658
	Logsig	Logsig	Purelin	0.9735	0.9727	0.0093	0.0094
	Logsig	Logsig	Tansig	0.9778	0.9753	0.0078	0.0085
	Logsig	Logsig	Logsig	0.8050	0.8067	0.1647	0.1658

Table 5
Trial and error to determine learning rate and momentum

Learning Rate	Momentum	ANN Structure	R Training	R Validation	MSE Training	MSE Validation
0.1	0.5	10>>30>>1	0.9757	0.9739	0.0085	0.0089
		10>>40>>1	0.9799	0.9797	0.0078	0.0073
		10>>30>>40>>1	0.9945	0.9933	0.0020	0.0025
		10>>40>>30>>1	0.9873	0.9856	0.0055	0.0066
		10>>40>>40>>1	0.9771	0.9744	0.0085	0.0094
	0.9	10>>30>>1	0.9789	0.9765	0.0075	0.0086
		10>>40>>1	0.9796	0.9783	0.0080	0.0080
		10>>30>>40>>1	0.9908	0.9895	0.0034	0.0038
		10>>40>>30>>1	0.9981	0.9967	0.0007	0.0012
		10>>40>>40>>1	0.9808	0.9795	0.0068	0.0075

the hidden layer 2, and 1 variable output, which resulted in the lowest validation MSE compared to other topology designs. Therefore, the best ANN topology model is used to predict chlorophyll content in African leaves. The lowest validation MSE is 0.0012, with an R of 0.9967, meaning that the correlation reaches 99.67% between the input and output variables. The results also showed that two hidden layers topology provided higher accuracy than one hidden layer.

Figure 6 presents a graph of the relationship between epoch and MSE in the training process to predict chlorophyll content in African leaves. The graph shows the learning process during training from the best topology was 10-40-30-1. The maximum number of iterations used was 10000 with an MSE goal was 0.01, where learning will stop when reaching either the 10000th iteration or 0.01 of the MSE goal. The convergent learning process in the 8th iteration with the resulting MSE training value of 0.00068112. It shows that learning reaches the MSE goal in the 8th iteration within 4 seconds. According to Damayanti et al. (2020), a good learning process decreased the MSE but iterations, so the learning graph will show a decrease in linear lines. Determination of epoch and goals based on previous research (Hendrawan et al., 2019c) stated that validated MSE with the goal of 0.01 was considerably accurate for predicting objective functions. Nevertheless, extremely low MSE can cause overfitting.

In Figure 7a, the blue line shows the regression plot of the training data simulation result. In contrast, the blue line in Figure 7b shows the regression plot of the validation data simulation result. Figure 7b shows the regression plot of the simulation results of the validation data. The two graphs show the distribution of data close to the linear fit line, which shows the accuracy of prediction as they are closer to the actual value; the closer the R-value, the more robust the correlation. For example, in Figure 7a, the R was 0.99806, which means a correlation of 99.806%.

In Figure 7b, the correlation coefficient is 0.99671, which means a correlation of 99.671%. According to Schober et al. (2018), there were several R criteria: 0.00 to 0.10 can be ignored, 0.10 to 0.39 was weak, 0.40 to 0.69 was moderate, 0.70-0.89 was strong, and 0.90-1.00 was robust. Therefore, the research results showed a robust correlation between textural features (input) and chlorophyll content (output). Figure 8 shows the best topology results with the lowest validation MSE to predict chlorophyll content in African leaves.

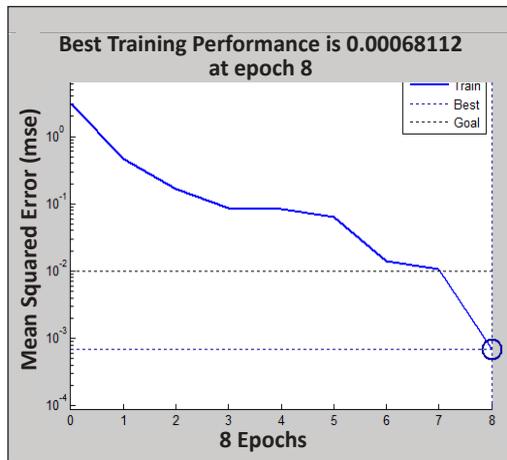


Figure 6. Learning process of ANN model to predict chlorophyll content in African leaves

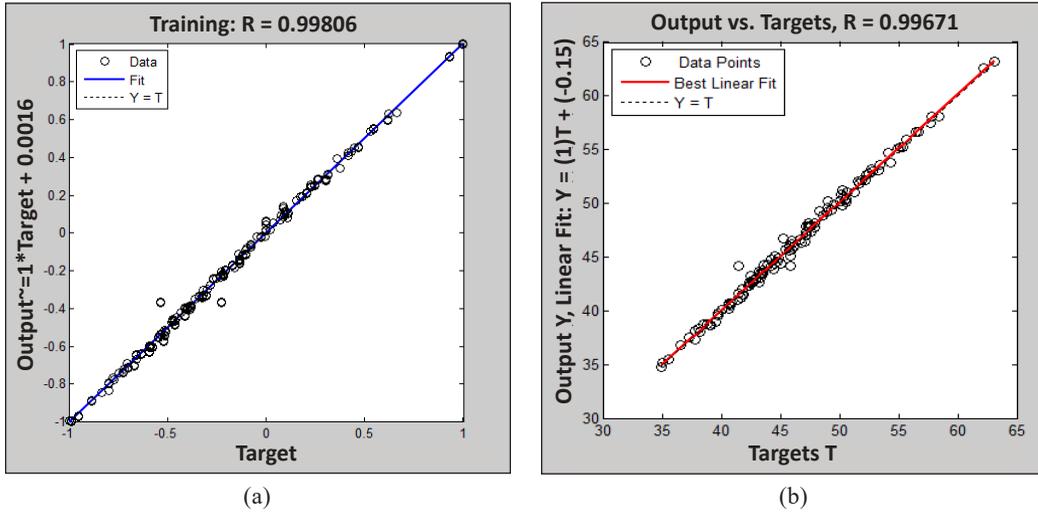


Figure 7. Simulation results of regression plots: (a) data training; (b) data validation

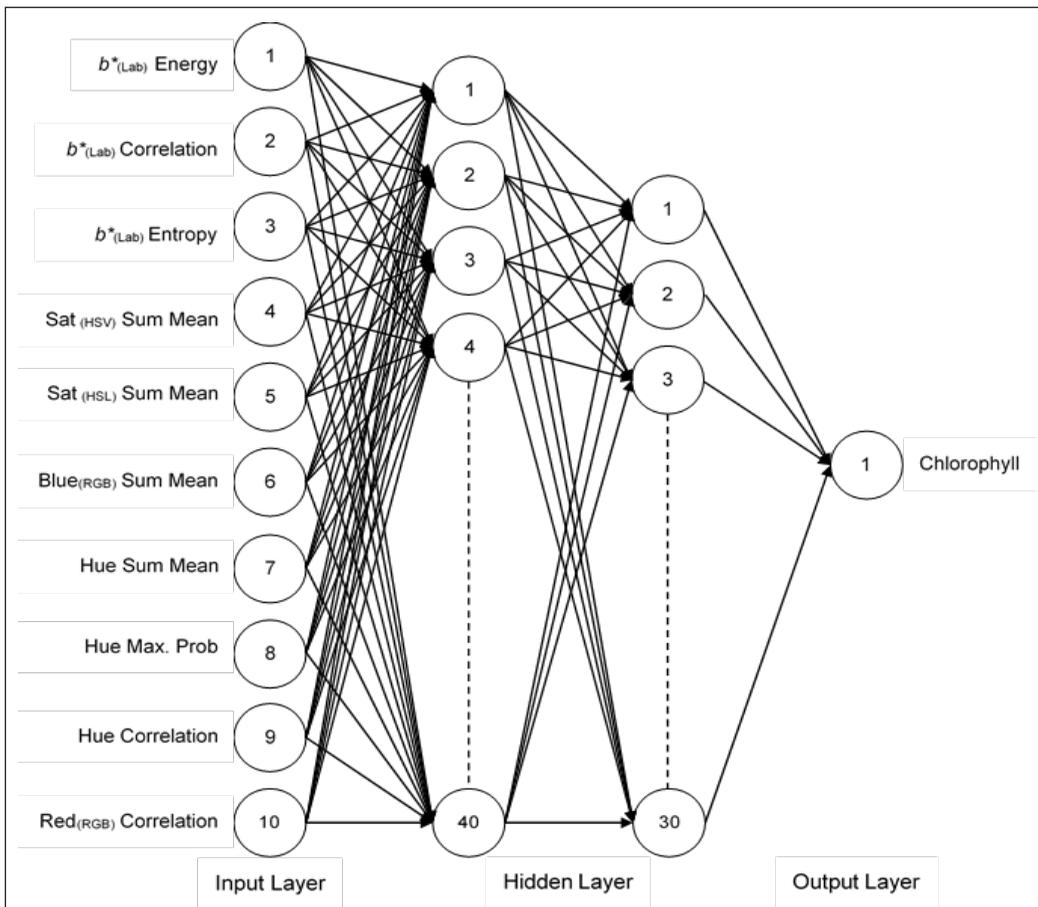


Figure 8. ANN-BP structure to predict chlorophyll content in African leaves

CONCLUSION

The best combination of 10 textural features input in ANN modelling to predict chlorophyll content in African leaves, among others, b* energy, b* correlation, b* entropy, saturation_(HSV) sum mean, saturation_(HSL) sum mean, blue sum mean, hue sum mean, maximum hue probability, hue correlation, and red correlation. Those ten textural features were obtained from selecting 120 textural features using feature selection with filter method ReliefF. Based on the research results, the best ANN topology was 10-40-30-1 (10 input layers, 40 nodes in hidden layer 1, 30 nodes in hidden layer 2, and 1 output layer). The selected learning function was trainlm, while the best activation function uses tansig in the hidden layer and purelin in the output layer. The selected ANN topology produced low MSE training of 0.0007 with R training of 0.9981, and the lowest validation MSE was 0.0012 with R validation of 0.9967. From these results, it can be concluded that there was a robust correlation of 99.67% between textural features (input) and chlorophyll content (output). Therefore, the topology can be potentially used as a model for predicting chlorophyll content in African leaves.

ACKNOWLEDGEMENT

The authors wish to acknowledge support from the Department of Agricultural Engineering, Faculty of Agricultural Technology, Universitas Brawijaya, for the assistance given during the research. This study was funded by the Associate Professor Research Grant Program (Hibah Doktor Lektor Kepala 2020), Universitas Brawijaya, the Ministry of Research, Technology, and Higher Education of the Republic of Indonesia.

REFERENCE

- Abdulkadir, A. R., Sarwar, M. J., & Dhiya, D. Z. (2015). Effect of chlorophyll content and maturity on total phenolic, total flavonoid contents and antioxidant activity of *Moringa oleifera* Leaf (Miracle Tree). *Journal of Chemical and Pharmaceutical Research*, 7(5), 1147-1152.
- Armi, L., & Shervan, F. E. (2019). Texture image analysis and texture classification methods - A review. *International Online Journal of Image Processing and Pattern Recognition*, 2(1), 1-29.
- Barman, U., & Choudury, R. D., (In Press). Smartphone image based digital chlorophyll meter to estimate the value of citrus leaves chlorophyll using linear regression, LMBP-ANN and SCGBP-ANN. *Journal of King Saud University – Computer and Information Sciences*. <https://doi.org/10.1016/j.jksuci.2020.01.005>
- Barman, U., Ridip, D. C., Arunav, S., Susmita, D., Bijon, K. D., Barna, P. M., & Golap, G. B. (2018). Estimation of chlorophyll using image processing. *International Journal of Recent Scientific Research*, 9(3), 24850-24853. <https://doi.org/10.24327/IJRSR>
- Borhan, M. S., Panigrahi, S., Satter, M. A., & Gu, H. (2017). Evaluation of computer imaging technique for predicting the SPAD readings in potato leaves. *Information Processing in Agriculture*, 4(4), 275-282. <https://doi.org/10.1016/j.inpa.2017.07.005>

- Cartelat, A., Cerovic, Z. G., Goulas, Y., Meyer, S., Lelarge, C., Prioul, J. L., Barbottin, A., Jeuffroy, M. H., Gate, P., Agati, G., & Moya, I. (2005). Optically assessed contents of leaf polyphenolics and chlorophyll as indicators of nitrogen deficiency in wheat (*Triticum aestivum* L.). *Field Crops Research*, *91*, 35-49. <https://doi.org/10.1016/j.fcr.2004.05.002>
- Dalen, G. V. (2006). Characterisation of rice using flatbed scanning and image analysis. *Food Policy, Control, and Research*, *6*, 149-186.
- Damayanti, R., Sandra, & Dahlena, E. (2020). The artificial neural network to predict chlorophyll content of cassava (*Manihot esculenta*) leaf. In *IOP Conference Series: Earth and Environmental Science* (Vol. 475, No. 1, p. 012012). IOP Publishing. <https://doi.org/10.1088/1755-1315/475/1/012012>
- Danladi, S., Muhammad, A. H., Idris, A. M., & Umar, I. I. (2018). *Vernonia amygdalina* Del: A mini review. *Research Journal of Pharmacy and Technology*, *11*(9), 4187-4190. <https://doi.org/10.5958/0974-360X.2018.00768.0>
- Garner, S. R. (1995, April 18-21). WEKA: The waikato environment for knowledge analysis. In *Proceedings of the New Zealand computer science research students conference* (Vol. 1995, pp. 57-64). University of Waikato, Hamilton.
- Grunenfelder, L., Hiller, L. K., & Knowles, R. (2006). Color indices for the assessment of chlorophyll development and greening of fresh market potatoes. *Postharvest Biology and Technology*, *40*(1), 73-81. <https://doi.org/10.1016/j.postharvbio.2005.12.018>
- Gupta, S. D., & Pattanayak, A. K. (2017). Intelligent image analysis (IIA) using artificial neural network (ANN) for non-invasive estimation of chlorophyll content in micropropagated plants of potato. *In Vitro Cellular & Developmental Biology-Plant*, *53*, 520-526. <https://doi.org/10.1007/s11627-017-9825-6>
- Haralick, R. M., Shanmugam, K., & Dinstein, I. H. (1973). Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics*, *3*(6), 610-621. <https://doi.org/10.1109/TSMC.1973.4309314>
- Hassanijalilian, O., Igathinathane, C., Doetkott, C., Bajwa, S., Nowatzki, J., & Esmaceli, S. A. H. (2020). Chlorophyll estimation in soybean leaves infield with smartphone digital imaging and machine learning. *Computer and Electronics in Agriculture*, *174*, 1-12. <https://doi.org/10.1016/j.compag.2020.105433>
- Hendrawan, Y., Amini, A., Maharani, D. M., & Sandra. (2019a). Intelligent non-invasive sensing method in identifying coconut (*Coco nucifera* var. Ebunea) ripeness using computer vision and artificial neural network. *Pertanika Journal of Science & Technology*, *27*(3), 1317-1339.
- Hendrawan, Y., Fauzi, M. R., Khoirunnisa, N. S., Andreane, M., Hartianti, P. O., Halim, T. D., & Umam, C. (2019b). Development of colour co-occurrence matrix (CCM) texture analysis for biosensing. *IOP Conference Series: Earth and Environmental Science*, *230*, 1-8. <https://doi.org/10.1088/1755-1315/230/1/012022>
- Hendrawan, Y., Widyaningtyas S., & Sucipto, S. (2019c). Computer vision for purity, phenol, and pH detection of Luwak Coffee green bean. *TELKOMNIKA*, *17*(6), 3073-3085. <http://dx.doi.org/10.12928/telkomnika.v17i6.12689>
- Hendrawan, Y., & Haruhiko, M. (2009). Precision irrigation for sunagoke moss production using intelligent image analysis. *Environmental Control in Biology*, *47*, 21-36. <https://doi.org/10.2525/ecb.47.21>

- Hendrawan, Y., Sakti, I. M., Wibisono, Y., Rachmawati, M., & Sandra. (2018). Image analysis using color co-occurrence matrix textural features for predicting nitrogen content in spinach. *TELKOMNIKA*, 16(6), 2711-2723. <http://dx.doi.org/10.12928/telkomnika.v16i6.10326>
- Hu, H., Liu, H. Q., Zhu, J. H., Yao, X. G., Zhang, X. B., & Zheng, K. F. (2010). Assesment of chlorophyll content based on image color analysis, comparison with SPAD-502. In *2010 2nd International Conference on Information Engineering and Computer Science* (pp. 1-3). IEEE Publishing. <https://doi.org/10.1109/ICIECS.2010.5678413>
- Jaber, A. A., Ahmed, A. M. S., & Hussein, F. M. A. (2019). Prediction of hourly cooling energy consumption of educational buildings using artificial neural network. *International Journal on Advanced Science Engineering Information Technology*, 9(1), 159-166. <https://doi.org/10.18517/ijaseit.9.1.7351>
- Kato, J., Hiroya, H., Shinsuke, T., Kenjiro, T., & Takashi, K. (2015). Analytical sensitivity in topology optimization for elastoplastic composites. *Structural and Multidisciplinary Optimization*, 52(3), 507-526. <https://doi.org/10.1007/s00158-015-1246-8>
- Kaur, G., Salim, D., Amandeep, S. B., & Derminder, S. (2014). Scanner image analysis to estimate leaf area. *International Journal of Computer Application*, 107(3), 5-10. <https://doi.org/10.5120/18729-9963>
- Kumar, C. S., & Rama, R. J. S. (2014). Application of ranking based attribute selection filters to perform automated evaluation of descriptive answers through sequential minimal optimization models. *ICTACT Journal on Soft Computing*, 5(1), 860-868. <https://doi.org/10.21917/IJSC.2014.0122>
- Li, J., Kewei, C., Suhang, W., Fred, M., Robert, P. T., Jiliang, T., & Huan, L. (2017). Feature selection: A data perspective. *ACM Computing Surveys*, 50(6), 94:1-94:45. <https://doi.org/10.1145/3136625>
- Limantara, L., Martin, D., Renny, I., Indriatmoko., & Tatas, H. P. B. (2015). Analysis on the chlorophyll content of commercial green leafy vegetables. *Procedia Chemistry*, 14, 225-231. <https://doi.org/10.1016/j.proche.2015.03.032>
- Luimstra, V. M., Schuurmans, J. M., Antonie, M. V., Klass, J. H., Jef, H., & Hans, C. P. M. (2018). Blue light reduce photosynthetic efficiency of cyanobacteria through an imbalance between photosystems I and II. *Photosynthesis Research*, 138(2), 177-189. <https://doi.org/10.1007/s11120-018-0561-5>
- Mendoza, R. J. P., Daniel, R., & Luis, D. M. (2018). Distributed reliefF-based feature selection in spark. *Knowledge and Information System*, 57, 1-20. <https://doi.org/10.1007/s10115-017-1145-y>
- Mohan, P. J., & Gupta, S. D. (2019). Intelligent image analysis for retrieval of leaf chlorophyll content of rice from digital images of smartphone under natural light. *Photosynthetica*, 57(2), 388-398. <https://doi.org/10.32615/ps.2019.046>
- Nursuhaili, A. B., Nur, A. S. P., Martini, M. Y., Azizah, M., & Mahmud, T. M. M. (2019). A review: Medicina valur, agronomic practices and post-harvest handlings of *Vernonia amygdalina*. *Food Research*, 3(5), 380-390. [https://doi.org/10.26656/fr.2017.3\(5\).306](https://doi.org/10.26656/fr.2017.3(5).306)
- Okafor, E., Lambert, S., & Marco, A. W. (2018). An analysis of rotation matrix and colour constancy data augmentation in classifying image of animals. *Journal of Information and Telecommunication*, 2(4), 465-491. <https://doi.org/10.1080/24751839.2018.1479932>

- Oyeyemi, I. T., Akinbiyi, A. A., Aderiike, A., Abimbola, O. A., & Oyetunde, T. O. (2017). *Vernonia amygdalina*: A folkloric herb with anthelmintic properties. *Beni-Suef University Journal of Basic and Applied Sciences*, 7(1), 43-49. <https://doi.org/10.1016/j.bjbas.2017.07.007>
- Pavlovic, D., Bogdan, N., Sanja, D., Hadi, W., Ana, A., & Dragana, M. (2014). Chlorophyll as a measure of plant health: Agroecological aspects. *Pestic Phytomed*, 29(1), 21-34. <https://doi.org/10.2298/PIF1401021P>
- Peng, Y., & Yi, W. (2019). Prediction of the chlorophyll content in pomegranate leaves based on digital image processing technology and stacked sparse autoencoder. *International Journal of Food Properties*, 22(1), 1720-1732. <https://doi.org/10.1080/10942912.2019.1675692>
- Rajalakshmi, K., & Narasimhan, B. (2013). Extraction and estimation of chlorophyll from medicinal plants. *International Journal of Science and Research*, 4(11), 209-212. <https://doi.org/10.21275/v4i11.nov151021>
- Samli, R., Nuket, S., Selcuk, S., & Vildan, Z. K. (2014). Applying artificial neural networks for the estimation of chlorophyll-a concentrations along the Istanbul Coast. *Polish Journal of Environmental Studies*, 23(4), 1281-1287.
- Schober, P., Christa, B., & Lothar, A. S. (2018). Correlation coefficients: Appropriate use and interpretation. *Anesthesia & Analgesia*, 126(5), 1763-1768. <https://doi.org/10.1213/ANE.0000000000002864>
- Setti, S., & Anjar, W. (2018). Analysis of backpropagation algorithm in predicting number of internet users in the world. *Jurnal Online Informatika*, 3(2), 110-115. <https://doi.org/10.15575/join.v3i2.205>
- Shakeri, M., Mohammad, M. A., Nasima, S., Mamun, M., & Syedul, M. A. (2012). Advanced CMOS based image sensors. *Australian Journal of Basic and Applied Sciences*, 6(7), 62-72.
- Shorten, C., & Taghi, M. K. (2019). A survey on image data augmentation for deep learning. *Journal of Big Data*, 6(60), 1-48. <https://doi.org/10.1186/s40537-019-0197-0>
- Uddling, J., Alfredsson, J. G., Piikki, K., & Pleijel, H. (2007). Evaluating the relationship between leaf chlorophyll concentration and SPAD-502 chlorophyll meter readings. *Photosynthesis Research*, 91(1), 37-46. <https://doi.org/10.1007/s11120-006-9077-5>
- Wang S., Tang J., & Liu H. (2016). Feature selection. In C. Sammut & G. Webb (Eds.), *Encyclopedia of Machine Learning and Data Mining* (pp. 1-9). Springer.
- Widiastuti, M. L., Aris, H., Endah, R. P., & Satriyas, I. (2018). Digital image analysis using flatbed scanning system for purity testing of rice seed and confirmation by grow out test. *Indonesian Journal of Agricultural Science*, 19(2), 49-56. <http://dx.doi.org/10.21082/ijas.v19n2.2018.p49-56>
- Xu, Y., & Royston, G. (2018). On splitting training and validation set: a comparative study of cross-validation, bootstrap and systematic sampling for estimating the generalization performance of supervised learning. *Journal of Analysis and Testing*, 2, 249-262. <https://doi.org/10.1007/s41664-018-0068-2>
- Yadav, S., Yasuomi, I., & Snehasish, D. G. (2010). Estimation of the chlorophyll content of micropropagated potato plants using RGB based image analysis. *Plant Cell Tissue and Organ Culture*, 100, 183-188. <https://doi.org/10.1007/s11240-009-9635-6>