Journal of Artificial Intelligence and Engineering Applications

Website: https://ioinformatic.org/

15th October 2023. Vol. 3. No. 1; e-ISSN : 2808-4519

Measuring the Maturity Level of Oil Palm Fruit For CPO Production Based on Color With Using the LVQ Method

Abdullah Hamid [1], Rusmin Saragih [2], Siswan Syahputra [3]

[1,2,3] *Information Systems, STMIK KAPUTAMA*
Jl. Veterans No. 4A-9A, Binjai, North Sumatra, Indonesia
ah49192@gmail.com [1], evitha12014@gmail.com [2], studentsyahputra90@gmail.com [3]

Abstract

Palm oil is a very important commodity besides oil and gas which also has a fairly good export value. The palm oil produced must be supported by the quality standards set by SNI. The level of maturity when harvesting oil palm fruit greatly influences the quality of Crude Palm Oil (CPO) production, which is crude palm oil that has a reddish color obtained from extraction or from the pressing process of oil palm fruit flesh. In fact, in the field of oil palm fruit harvest, there are still many oil palm fruit that are not ripe enough and can even be said to be still raw, entering the CPO production process. Determination of the maturity level of oil palm fruit is generally determined based on the amount of loose fruit and color, so handling the harvest of oil palm fruit is an important activity in improving the quality of CPO. It is necessary to build a system capable of managing and processing palm fruit images to measure the maturity level of the palm fruit to be produced. To obtain the right level of accuracy, this research uses the Learning Vector Quantization (LVQ) method. LVQ is a method for conducting supervised competitive layer learning. From the results of trials conducted, it is proven that the system can measure very ripe oil palm fruit with HSV values (0.052209; 0.896021; 0.791114).

Keywords: *CPO (Crude Palm Oil), Palm Oil, LVQ (Learning Vector Qualization).*

1. Introduction

Palm oil is a very important commodity besides oil and gas which also has a fairly good export value. Therefore, there is a need for supervision to maintain the quality and quantity of these commodities. The palm oil produced must be supported by the quality standards set by SNI. Some of the criteria for palm oil that are needed are having a reddish color, good taste and smell, can be stored for a long time, easy to purify and the level of hydrolysis in the formation of Free Fatty Acids (ALB) produced is low, water and impurities in palm oil. In order to obtain maximum results in both quality and quantity, in the processing of palm oil in factories starting from the processing stage to stockpiling must pay attention to and maintain the quality standards that apply to the company.

The level of maturity when harvesting oil palm fruit also greatly influences the quality of Crude Palm Oil (CPO) production, which is crude palm oil that has a reddish color obtained from extraction or from the pressing process of oil palm fruit flesh. Facts in the field that occur, the yield of oil palm fruit is still found a lot of oil palm fruit that is not ripe enough and can even be said to be still raw into the CPO production process. Meanwhile, the determination of the quality of the CPO produced is highly dependent on the maturity level of the oil palm fruit. Determination of the maturity level of oil palm fruit is generally determined based on the amount of loose fruit and color, so handling the harvest of oil palm fruit is an important activity in improving the quality of CPO.

From the problems above, it is necessary to build a system or use an application that is capable of managing and processing palm fruit images to measure the maturity level of the palm fruit to be produced into CPO. With this system, it is hoped that it will be able to reduce the harvest of oil palm fruit that is not yet suitable for production into CPO, so that CPO production can produce the desired production quality. In measuring the maturity level of oil palm fruit, a feature extraction of oil palm fruit images will be used, namely HSV (Hue, Saturation, Value). To obtain the right level of accuracy, this research uses the Learning Vector Quantization (LVQ) method. LVQ is a method for conducting supervised competitive layer learning. A competitive layer will automatically learn to classify input vectors. The classes obtained as a result of this competitive layer depend only on the distance between the input vectors. If the 2 input vectors are close to the same, then the competitive layer will put the two input vectors into the same class.

Several systems using the LVQ method have been widely studied, including the Application of Learning Vector Quantization in Classifying Tomato Fruit Maturity Levels Based on Fruit Color. The highest accuracy reached 87.25% [1]. While research entitled Implementation of

Learning Vector Quantization (LVQ) for Natural Ripe Mango Identification Model. The results obtained using the Learning Vector Quantization (LVQ) algorithm are very effective in being able to distinguish between tree ripe mangoes and ripe mangoes with calcium carbide properly. The results obtained in the best training process, namely showing an accuracy of 95.83%, are the results of the average. With an average of each mango, that is, for tree-ripe mangoes, the accuracy obtained was 91.67% and for ripe mangoes with calcium carbide, the accuracy was 100% [2].

2. Research methodology

There are several problem solving methods that will be discussed in this study which will be described as below.

2.1. Image processing

Digital Image Processing (Digital Image Processing) is a field of science that studies how an image is formed, processed, and analyzed so as to produce information that can be understood by humans [3] .

Analog images are generated by analog image acquisition tools, for example the human eye and analog cameras. Images captured by the human eye and photos or films captured by analog cameras are examples of analog images. The image has quality with a very good level of detail (resolution) but has weaknesses, including that it cannot be stored, processed, and duplicated on a computer.

2.2. Learning Vector Quantization (LVQ)

The Learning Vector Quantization (LVQ) method is a training method for learning in a supervised competitive layer with a single layer network architecture. The classes obtained as a result of this competitive layer depend only on the distance between the input vectors. If the two input vectors are close to the same, then the competitive layer will put the two input vectors into the same class [4].

A classification method where each output unit represents a class. LVQ is used for grouping where the number of groups has been determined by the architecture (the target/class has been determined). LVQ is a neural network which is a supervised competitive learning algorithm version of the Kohonen Self-Organizing Map (SOM) algorithm. The aim of this algorithm is to approximate the distribution of vector classes to minimize classification errors.

2.3. Definition of Oil Palm

Oil palm is a monocot plant that does not have a taproot. Radicles (potential roots) in seedlings continue to grow elongated downwards for six months continuously and the roots reach 15 meters in length. Oil palm plants have leaves (fronds) that resemble bird or chicken feathers. At the base of the leaf sheath are formed two rows of very sharp and hard thorns on both sides. Children leaves (foliage leaflet) are arranged in two rows to the tip of the leaf. In the middle of each leaf child a stick is formed as a leaf bone. Oil palm plants that are three years old have started to mature and start producing male or female flowers. The male flowers are oval in shape, while the female flowers are slightly round. Each type of oil palm has a different size and seed weight. African dura seeds are 2-3 cm long and weigh an average of 4 grams, so that in 1 kg there are 250 seeds. Dura deli seeds weigh 13 grams per seed, and African tenera seeds weigh an average of 2 grams per seed [5].

3. Application of the Learning Vector Quantization (LVQ) Method

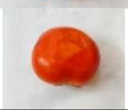
The application of the method is needed in solving a problem in a research and assessment process. To identify a data, of course, accurate data analysis must be carried out in the process of identifying a data. Many methods used in the data identification process have been carried out by many researchers. This research was conducted to measure the maturity level of oil palm fruit for CPO production based on color using the Learning Vector Quantization algorithm. The steps in applying the Learning Vector Quantization method are as follows.

3.1. Determine Input Data, Training Data and Target Data

In calculations using the Learning Vector Quantization algorithm, image feature extraction data from oil palm fruit cannot be used, the image feature extraction used is Hue, Saturation, value (HSV). The HSV value is obtained by extracting image features using Matlab software. The oil palm image data that will be analyzed is as in the following table:

Table 1: Input data

No	Image	Target
1		Ripe (2)
2		Ripe (2)

No	Image	Target
3		Ripe (2)
4		raw (1)
5		raw (1)
6		raw (1)
7		raw (1)
8		Very Ripe (3)
9		Very Ripe (3)
10		Very Ripe (3)

From the image data above, the results of the extraction of image feature values can be seen in the table below.

Table 2: Image Value Feature Extraction Data

No.	Hue	saturation	Value	Target
1	0.259079	0.734066	0.601649	2
2	0.127415	0.839965	0.707763	2
3	0.256974	0.734302	0.596348	2
4	0.471839	0.520787	0.445418	1
5	0.298390	0.596090	0.516828	1
6	0.111216	0.611003	0.618157	1
7	0.116804	0.551994	0.611786	1
8	0.036969	0.835557	0.779227	3
9	0.051759	0.894254	0.791505	3
10	0.036073	0.831630	0.777427	3

The three inputs from the data above will be initialized as data weights, namely data for serial numbers 4, 1 and 8 (W):

Table 3: Weight Data

No.	Hue	saturation	Value	Target
1	0.471839	0.520787	0.445418	1
2	0.259079	0.734066	0.601649	2
3	0.036969	0.835557	0.779227	3

While the remaining 7 inputs will be used as data to be trained:

Table 4: Training Data

No.	Hue	saturation	Value	Target
1	0.127415	0.839965	0.707763	2
2	0.256974	0.734302	0.596348	2
3	0.298390	0.596090	0.516828	1
4	0.111216	0.611003	0.618157	1
5	0.116804	0.551994	0.611786	1
6	0.051759	0.894254	0.791505	3
7	0.036073	0.831630	0.777427	3

3.4. Determine Learning Rate and Maximum Epoch

In this calculation analysis, the initial value will be chosen Learning Rate (α) = 0.05, with a reduction of $0.1 * \alpha$, and maximum epoch (MaxEpoch) = 1 by calculating the Learning Vector Quantization method as follows:

1st data: (0.127415; 0.839965; 0.707763)

1st weight distance

$=\sqrt{(0.127415-0.471839)^2+(0.839965-0.520787)^2+(0.707763-0.445418)^2}$

= 0.537892

Distance of the 2nd weight

$=\sqrt{(0.127415-0.259079)^2+(0.839965-0.734066)^2+(0.707763-0.601649)^2}$

= 0.199525

3rd weight spacing

$=\sqrt{(0.127415-0.036969)^2+(0.839965-0.835557)^2+(0.707763-0.779227)^2}$

= 0.115356

The smallest distance is the 3rd w

The 1st data target is 2

Since the 1st data target ≠ 1st w, then the new 3rd w is:

$W_j(new) = W_j(old) + \alpha (X_i - W_j(old))$

= 0.036969 + 0.05 * (0.041491 - 0.036969) = 0.041491

= 0.835557 + 0.05 * (0.835777 - 0.835557) = 0.835777

= 0.779227 + 0.05 * (0.775654 - 0.779227) = 0.775654

W3(baru) = (0,041491; 0,835777; 0,775654)

Data 2: (0.256974; 0.734302; 0.596348)

1st weight distance

$=\sqrt{(0.256974-0.471839)^2+(0.734302-0.520787)^2+(0.596348-0.445418)^2}$

= 0.338431

Distance of the 2nd weight

$=\sqrt{(0.256974-0.259079)^2+(0.734302-0.734066)^2+(0.596348-0.601649)^2}$

= 0.296592

3rd weight spacing

$=\sqrt{(0.256974-0.041491)^2+(0.734302-0.835777)^2+(0.596348-0.775654)^2}$

= 0.298129

The smallest distance is at the 2nd w

The 2nd data target is 2

Because the 2nd data target = 2nd w, then the new 2nd w is:

$W_j(new) = W_j(old) + \alpha (X_i - W_j(old))$

= 0.259079 + 0.05 * (0.256974 - 0.259079) = 0.258974

= 0.734066 + 0.05 * (0.734302 - 0.734066) = 0.734078

= 0.601649 + 0.05 * (0.596348 - 0.601649) = 0.601384

W2(new) = (0.258974; 0.734078; 0.601384) (new) = (0.041491; 0.835777; 0.775654)

Data 3 : (0.298390; 0.596090; 0.516828)

1st weight distance

$=\sqrt{(0.298390-0.471839)^2+(0.596090-0.520787)^2+(0.516828-0.445418)^2}$

= 0.202124

Distance of the 2nd weight

$=\sqrt{(0.298390-0.258974)^2+(0.596090-0.734078)^2+(0.516828-0.601384)^2}$

= 0.166566

3rd weight spacing

$=\sqrt{(0.298390-0.041491)^2+(0.596090-0.835777)^2+(0.516828-0.775654)^2}$

= 0.436392

The smallest distance is at the 2nd w

The 3rd data target is 1

Because the 3rd data target = 2nd w, then the new 2nd w is:

$W_j(new) = W_j(old) + \alpha (X_i - W_j(old))$

= 0.258974 + 0.05 * (0.298390 - 0.258974) = 0.260945

= 0.734078 + 0.05 * (0.596090 - 0.734078) = 0.727179

= 0.601384 + 0.05 * (0.516828 - 0.601384) = 0.597156

W2(new) = (0.260945; 0.727179; 0.597156)

Fourth Data : (0.111216; 0.611003; 0.618157)
1st weight distance
=√(〖(0.111216-0.471839)〗 ^2+ 〖(0.611003-0.520787)〗 ^2+ 〖(0.618157-0.445418)〗 ^2)
= 0.409910

Distance of the 2nd weight
=√(〖(0.111216-0.260945)〗 ^2+ 〖(0.611003-0.727179)〗 ^2+ 〖(0.618157-0.597156)〗 ^2)
= 0.190674

3rd weight spacing
=√(〖(0.111216-0.041491)〗 ^2+ 〖(0.611003-0.835777)〗 ^2+ 〖(0.618157-0.775654)〗 ^2)
= 0.283179
The smallest distance is at the 2nd w
The 4th data target is 1
Because the 4th data target = 1st w, then the new 1st w is:
Wj(new) = Wj(old) + α (Xi - Wj(old))
= 0.260945 + 0.05 * (0.111216 - 0.260945) = 0.253458
= 0.727179 + 0.05 * (0.611003 - 0.727179) = 0.721370
= 0.597156 + 0.05 * (0.618157 - 0.597156) = 0.598206

W2(new) = (0.253458; 0.721370; 0.598206)
5th Data : (0.116804; 0.551994; 0.611786)
1st weight distance
=√((〖0.116804-0.471839)〗 ^2+ 〖(0.551994-0.520787)〗 ^2+ 〖(0.611786-0.445418)〗 ^2)
= 0.393322

Distance of the 2nd weight
=√(〖(0.116804-0.253458)〗 ^2+ 〖(0.551994-0.721370)〗 ^2+ 〖(0.611786-0.598206)〗 ^2)
= 0.181140

3rd weight spacing
=√(〖(0.116804-0.041491)〗 ^2+ 〖(0.551994-0.835777)〗 ^2+ 〖(0.611786-0.775654)〗 ^2)
= 0.283179

The smallest distance is at the 2nd w
The 5th data target is 1
Because the 5th data target = 2nd w, then the new 2nd w is:
Wj(new) = Wj(old) + α (Xi - Wj(old))

= 0.253458 + 0.05 * (0.116804 - 0.253458) = 0.246626

= 0.721370 + 0.05 * (0.551994 - 0.721370) = 0.712901

= 0.598206 + 0.05 * (0.611786 - 0.598206) = 0.598885

W2(new) = (0.246626; 0.712901; 0.598885)

6th Data : (0.051759; 0.894254; 0.791505)
1st weight distance
=√(〖(0.051759-0.471839)〗 ^2+ 〖(0.894254-0.520787)〗 ^2+ 〖(0.791505-0.445418)〗 ^2)
= 0.660092

Distance of the 2nd weight
=√(〖(0.051759-0.246626)〗 ^2+ 〖(0.894254-0.712901)〗 ^2+ 〖(0.791505-0.598885)〗 ^2)
= 0.328579

3rd weight spacing
=√(〖(0.051759-0.041491)〗 ^2+ 〖(0.894254-0.835777)〗 ^2+ 〖(0.791505-0.775654)〗 ^2)
= 0.061452

The Smallest Distance is at the 3rd w
The 6th data target is 3
Because the 6th data target = 3rd w, then the new 3rd w is:
Wj(new) = Wj(old) + α (Xi - Wj(old))
= 0.041491 + 0.05 * (0.051759 - 0.051759) = 0.042005
= 0.835777+ 0.05 * (0.894254 - 0.835777) = 0.838701
= 0.775654 + 0.05 * (0.791505 - 0.775654) = 0.776446
W3(new) = (0.042005; 0.838701; 0.776446)

7th Data : (0.036073; 0.831630; 0.777427)
1st weight distance

=√(〖(0.036073-0.471839)〗^2+ 〖(0.831630-0.520787)〗^2+ 〖(0.777427-0.445418)〗^2)
= 0.629877

Distance of the 2nd weight
=√(〖(0.036073-0.246626)〗^2+ 〖(0.831630-0.712901)〗^2+ 〖(0.777427-0.598885)〗^2)
= 0.300509

3rd weight spacing
=√(〖(0.036073-0.042005)〗^2+ 〖(0.831630-0.838701)〗^2+ 〖(0.777427-0.776446)〗^2)
= 0.009282

The Smallest Distance is at the 3rd w
The 7th data target is 3
Since the 7th data target = 3rd w, then the new 3rd w is:
Wj(new) = Wj(old) + α (Xi - Wj(old))
= 0.042005 + 0.05 * (0.036073 - 0.042005) = 0.041708
= 0.838701 + 0.05 * (0.831630 - 0.838701) = 0.838347
= 0.776446 + 0.05 * (0.777427 - 0.776446) = 0.776495
W3(new) = (0.041708; 0.838347; 0.776495)

For the 1st iteration, the final weights are obtained:
W1 = (0.471839; 0.520787; 0.445418)
W2 = (0.246626; 0.712901; 0.598885)
W3 = (0.041708; 0.838347; 0.776495)

Next is to test the new data with image data as follows:

Table 5: Test Image

Image	Image Feature Value Extraction		Information
	Hue saturation Value	: 0.052209 : 0.896021 : 0.791114	Very Ripe (3)

From the data above, first find the input distance to the two weights. The number of the weight with the shortest distance will be the class.
1st weight distance
$=\sqrt{(0{,}052209-0{,}471839)^2+(0{,}896021-0{,}520787)^2+(0{,}791114-0{,}445418)^2}$
= 0.660602

Distance of the 2nd weight
$=\sqrt{(0{,}052209-0{,}246626)^2+(0{,}896021-0{,}712901)^2+(0{,}791114-0{,}598885)^2}$
= 0.329063

3rd weight spacing
$=\sqrt{(0{,}052209-0{,}041708)^2+(0{,}896021-0{,}838347)^2+(0{,}791114-0{,}776495)^2}$
= 0.660602

The smallest distance is in the 3rd weight, so that the input image of the oil palm fruit test in Table III.5 is included in class 3 (very ripe). Thus the test image data of the oil palm fruit are measured as very ripe oil palm fruit with HSV values (0.052209; 0.896021; 0.791114).

4. Conclusion

With a system for measuring the maturity level of oil palm fruit for CPO production based on color using the LVQ method, several conclusions can be drawn, including the following:

1. Helping PT. Perkebunan Nusantara II Sawit Hulu is in the process of harvesting oil palm fruit that has the right level of maturity.
2. Provide color analysis information for laypeople or new employees to make it easier for them to understand the maturity level of oil palm fruit.
3. Minimizing palm oil errors for the production of palm oil CPO, which of course will reduce losses for the company because all palm oil is harvested according to the standard level of maturity.
4. From the results of trials conducted it is proven that the system can measure very ripe oil palm fruit with HSV values (0.052209; 0.896021; 0.791114).

Reference

[1] Sanjaya, S. (2019). Application of Learning Vector Quantization in Classifying Levels of Ripeness of Tomatoes Based on Fruit Color. CoreIT Journal: Journal of Computer Science and Information Technology Research Results, 5(2), 49.
[2] Luthfi Khalid, Jayanta, YW (2020). Implementation of Learning Vector Quantization (Lvq) for Natural Ripe Mango Identification Model. 606–621.
[3] Jepriana, SH & IW (2020). The Concept of Algorithms and Their Applications in the C++ Programming Language. CV. Andi Offset, Yogyakarta.
[4] Usman, W., Damanik, IS, & Hardinata, JT (2020). Artificial Neural Network with Learning Vector Quantization (LVQ) Method in Determining Classification of Types of Tickets Based on Vehicles. Proceedings of the National Seminar on Information Science Research (SENARIS), 1(September), 780.
[5] Dlh.probolinggokab.go.id. (2023).
[6] BW, TA, Hermanto, IGR, & D, RN (2009). Introduction to Balinese Letters Using Modified Direction Feature (MDF) and Learning Vector Quantization (LVQ) Methods. National Conference on Systems and Informatics 2009, 7–12.
[7] Cakra Adipura Wicaksana, DETL (2021). Recognition of Digital Image Signature Vector Patterns Using Regional Division and Learning Vector Quantization (LVQ) Methods. Journal of Control and Network Systems, 5(2), 1–13.
[8] Jepriana, SH & IW (2020). The Concept of Algorithms and Their Applications in the C++ Programming Language. CV. Andi Offset, Yogyakarta.
[9] Nurhayati. (2022). Ensemble Learning Techniques to Increase Prediction Model Accuracy Performance (Selection of Scholarship Recipient Students). Pascal Books, Tangerang.
[10] Rachmat Destriana, Syepry Maulana Husain, Nurdiana Handayani, ATPS (2021). UML Diagram in Creating Firebase Android Applications "Case Study of Garbage Bank Applications. Deepublish Publisher, Sleman.