Detection of palm fruit maturity level in the grading process through image recognition and fuzzy inference system to improve quality and productivity of crude palm oil (CPO)

A Syaifuddin, L N A Mualifah, L Hidayat and A M Abadi

Department of Mathematics, Faculty of Mathematics and Natural Science, Universitas Negeri Yogyakarta, Yogyakarta, Indonesia

E-mail: akhmadsf44@gmail.com

Abstract. The agricultural sub-sector that has the potential to develop is the plantation sub-sector. The plantation sub-sector contributes around 3.57% in GDP. Plantation sub-sector commodities whose growth is very rapid are oil palm which is the basic ingredient in making palm oil (CPO). The world need for oil and vegetable fats which continues to increase, must be supported by an increase in palm oil production. Determination / grading of oil palm fruit that is still done manually is certainly not effective if done for the large-scale oil palm industry, so there needs to be innovation in the process of selecting / grading palm fruit to obtain the right maturity of palm fruit. This research uses research & development methods. In general, this research detects the maturity level of palm fruit from its outer color. Test the colour of the outer palm fruit based on image taking. Parameters were formed based on the matrix and used the image process that is obtained as input values on the fuzzy inference system. Hardware integration with a digital image processing computing system is done so that the resulting data can be directly processed and recognized in real time.

The recognition system for oil palm fruit maturity can be done by: sample data fulfillment, pre-processing of sample images, extraction of image data using the GLCM method, determining extraction features using regression methods, dividing data into training data and test data, doing clustering, building fuzzy systems use training data, testing the accuracy of the system. Obtained accuracy for training data is 73.07% and testing data is 71.4%.

1. Introduction

As an agrarian country, the agricultural sector in Indonesia has an important role in national economic growth. Palm oil is an agricultural sector commodity, especially the plantation sub-sector whose production growth is very rapid. In 2017 it is estimated that oil palm production in Indonesia amounted to 35.36 million tons, this estimated number increased 6.41% from oil palm production in 2016 [1]. Palm oil is the basic ingredient of making crude palm oil (CPO). CPO is used for various industrial needs such as food, cosmetics, animal feed, etc.

The increasing world demand for vegetable oils and fats must be supported by increased production of palm oil as raw material for vegetable oils. In 2015, around 26.47 million tons of CPO production in Indonesia were exported to various countries in the world, the amount of palm oil exports increased by around 15.62% compared to exports in the previous year [2].

CPO with good quality has a high oil yield and low free fatty acids which are obtained from oil palm fruits with the right level of maturity. In the oil palm industry, the maturity level of fresh fruit
bunches is classified into three fractions, namely the unripe fraction for raw fruit, ripe for ready-to-harvest ripe fruit, and overripe for ripe fruit [3].

The process of processing oil palm fruit into CPO begins with harvesting oil palm fruit, cutting the fruit stem, weighing, and grading. At present most of the grading of oil palm fruit is still done manually, namely sorting the oil palm fruit by someone who is already an expert. This is certainly not effective if done for the large-scale oil palm industry.

The process of palm oil production needs to be improved along with increasing demand for palm oil (CPO) by increasing the quality of CPO and increasing the processing of palm fruit. So there needs to be innovation in the process of selecting / grading oil palm fruit to obtain oil palm fruit with the right level of maturity and with efficient time. Research on oil palm grading was carried out by Azrifirwan [4] by developing an oil palm bunch grading machine based on non-destructive inspection techniques with a control system and sensors to record the image of oil palm bunches. This grading machine has an effective working capacity of 21.4 tons / hour, a maximum palm oil bunch load of 300 kg, an accuracy rate of 95%, and tool deflection from 0 to 0.5 cm. Rohendar [5] offers a non-destructive solution by combining digital image processing with fuzzy to determine the level of maturity of palm. The results obtained have an accuracy rate of 86.6% for crude and overripe oil palm fruit, while for 100% ripe fruit.

2. The Preliminaries

2.1. Digital Image
Digital image is a matrix where row and column indices represent a point on the image and matrix elements express intensity at that point. Generally in applications of image processing, digital images can be divided into three, namely color images, black and white images, and binary images [6].

2.2. Feature Extraction
Gray Level Co-occurrence Matrix (GLCM) is one of the statistical methods that can be used for texture analysis. The co-occurrence matrix is formed from an image by looking at pairs of pixels that have certain intensity. The values computed by GLCM can describe the features of the image, so it can be a basis of texture classification [7]. The use of this method is based on the hypothesis that in a texture there will be a recurrence of configuration or gray level pair. In the co-occurrence matrix, there are eleven texture characteristics that can be obtained from an image that is used as a differentiator between images with a certain class, with another class [8].

2.3. Backward Method
The backward method is a step backward, all X variables are regressed with the Y variable, eliminating the X variable based on the values of the smallest or the least variable X in the model is also determined by the table variable. This method is a good method because in this method the response variable behavior is explained carefully by selecting explanatory variables from the many explanatory variables available in the data [9].

2.4. Fuzzy C-Means Clustering
Fuzzy clustering is one technique to determine the optimum cluster in a vector space based on the formal form of Euclidian for the distance between vectors [10]. There are several categories of data clustering, one of which is Fuzzy C-means Clustering (FCM). FCM is a data clustering technique in which the existence of each data point in a cluster is determined by the degree of membership. The output of the FCM is a row of cluster centers and some degree of membership for each data point. This information can be used to build a fuzzy inference system.

2.5. Fuzzy Logic
Fuzzy logic is a mathematical discipline that we use every day and helps us to reach the structure in which we interpret our own behaviors [11]. It needs to be done if the antecedent is more than one preposition [12]. In fuzzy set theory, the role of the degree of membership as a determinant of the
existence of elements in a set is very important. Membership value or degree of membership is the main characteristic of reasoning with fuzzy logic [10]. The membership function is a curve that shows the mapping of input points into membership values that have intervals between 0 and 1. Functions that can be used include linear, triangular, trapezoidal and Gauss [10]. In this study the membership function used is gauss, which is a bell-shaped curve with a degree of membership 1 located at the center of the domain, and curve width as in the following picture.

\[ G(x; \sigma; \gamma) = e^{\frac{(x-\gamma)^2}{2\sigma^2}} \]  

Figure 1. Gauss Curve

2.6. Fuzzy System
Fuzzy systems are linguistic decryptions (in the form of fuzzy If-Then rules) about processes that can be combined into a model of those processes [13]. Fuzzy system consists of several components, namely fuzzification, fuzzy rules, fuzzy inference, and defuzzification.

2.7. Graphical User Interface (GUI)
One of the MATLAB applications is the development of mathematical and computational algorithms, data visualization, the development of graphical systems, and the creation of a Graphical User Interface (GUI). A GUI is a graphical display in one or more windows that contain a collection of controls called components. The advantage of the GUI is that the user does not have to create scripts or type commands to complete the task, other than that the user does not need to understand the details of how the task is completed [14].

Figure 2. Graphical User Interface Basic Display

3. Research Method
In general, this research detects the level of maturity of oil palm fruit from the outside color. Testing the outside color of oil palm fruit bunches based on the image capture and the parameters are formed based on the matrix of the image processing. This research was conduct from October 2018 to April 2019, the research procedures are as follows:
1. Fulfillment of sample data starts from the taking of the palm image and sorting the image data.
2. Image processing includes cropping, resizing, gray scaling, and extraction of image data with GLCM method.
3. Determine the features extraction using backward regression method. The features extraction are used as the input variables in the fuzzy system.
4. Divide data into training data and testing data, and then clustering the training data using FCM method.
5. Construct the fuzzy system using the training data and test the system accuracy.
6. Model construction with GUI.

4. Results and Discussion

4.1. Image Pre-Processing
Preprocessing is done by changing the RGB image to form a grayscale image and image re-sizing. The images of pre-processing are shown in Figure 3.

![Pre-processing process](image)

Figure 3. Pre-processing process (a) RGB image of overripe palm fruit, (b) Grayscale image of overripe palm fruit, (c) RGB image of ripe palm fruit, (d) Grayscale image of ripe palm fruit, (e) RGB image of unripe palm fruit and (f) Grayscale image of unripe palm fruit.

4.2. Features Extraction
This process aims to obtain quantitative data from an image or image. The feature extraction method used is the Gray Level Co-occurrence matrix (GLCM). The extraction process using the GLCM
method produces 22 parameters or variables including uniformity, entropy, dissimilarity, contrast/inertia, inverse difference, correlation, homogeneity, autocorrelation, cluster shade, cluster prominence, maximum probability, sum of squares, sum average, sum variance, sum entropy, difference variance, difference entropy, information measures of correlation (1), information measures of correlation (2), maximal correlation coefficient, inverse difference normalized (INN), and information difference normalized moment (IDN).

4.3 Input Selection
The next step is to select the input using backward regression. The independent variable used is obtained through image extraction while the dependent variable is the classification of maturity level. This backward regression process is carried out with the help of SPSS software. The results obtained are eight specific or significant variables for GLCM extraction including uniformity (X1), inverse difference (X2), correlation (X3), homogeneity (X4), sum of squares (X5), difference entropy (X6), information measures of correlation (2) (X7), and maximal correlation coefficient (X8). Then the eight variables will be input into each fuzzy model built from training data, while the other variables are not used.

4.4 Clustering
The clustering process is done by using the FCM method to obtain cluster centers that are used to build rules or fuzzy rules. Determination of the optimum number of clusters is done by trial and error so as to obtain a high accuracy value. The optimum number of clusters is 8 clusters. Furthermore, the data was clustered as many as 8 clusters to obtain a cluster center that was used to build rules in the fuzzy inference system.

4.5 Fuzzy System
This study uses eight inputs which are defined as eight fuzzy sets according to the optimum number of clusters with gaussian membership. While output is divided into 3 fuzzy sets, namely 1 for raw palm, 2 for mature palm, and 3 for overcooked palm. The next step is forming a fuzzy rule based on the cluster center. Each cluster is a rule, so there are eight rules. Rule is formed using an if-then rule with input variables as antecedent and output (classification) as consequent. In this study, the first order Sugeno system is used, so that the consequences of the rule are linear equations. The linear equation is then solved by the singular value decomposition method. The final step in the formation of a fuzzy system is defuzzification. In this study, the defuzzification method used is the weight average method.

4.6 Accuracy Analysis of Results
The model that has been formed from the training data is then tested by calculating the accuracy of all data (training data and testing data) to determine the accuracy of a model. Test the accuracy of the results of the 78 training data defuzzification.

\[
\text{accuracy} = \frac{\text{amount of data correct}}{\text{amount of all data}} \times 100\% = \frac{57}{78} \times 100\% = 73.07\%.
\]

Defuzzification results in training data have a high accuracy due to the process of forming a model based on training data. Therefore it is necessary to test the model that has been formed for data testing. Test the accuracy results of testing data defuzzification.

\[
\text{accuracy} = \frac{\text{amount of data correct}}{\text{amount of all data}} \times 100\% = \frac{15}{21} \times 100\% = 71.4\%.
\]

4.7 Graphical User Interface (GUI) Display
This GUI display can be directly used to identify the maturity level of a palm oil fruit image input. Selecting image of the Import button is the beginning. Figure 5 uses the input of the Palm Oil fruit image from the training data with mature classification. The initial image or image is inserted into the GUI through the pre-processing stage, namely grayscale and resize, the results of which can be seen
below the original image, then the image is extracted using the GLCM method. The GLCM method produces 22 variables but because it has been selected only 8 variables are used with the values we can see next to the image of the Palm fruit which will later be used as a fuzzy system. Then the processing is done using a fuzzy system that has been made and the results obtained are in the form of a mature palm image. The results of the GUI design are in accordance with the initial design using a fuzzy system that has been tested for accuracy.

![Detection of Palm Fruit Maturity Level using Image Recognition and Fuzzy Inference System](image)

**Figure 4.** The GUI for Maturity Identification of Coconut Palm using Fuzzy Inference System

Many factors can affect the results of this fuzzy system. Starting from the most basic is image data, image pre-processing stage carried out, the selected image extraction method, the variable selection method used as input, the number of variables used as input, the shape or architecture arrangement of the fuzzy system used to process data. These things will certainly affect the results, accuracy or quality of the fuzzy system.

5. Conclusion
The recognition system of coconut palm fruit maturity can be done by: fulfilling sample data, pre-processing the sample image, extracting image data using the GLCM method, determining extraction features using the regression method, dividing the data into training data and testing data, clustering process, construct the fuzzy system using training data, testing the accuracy of the system. The accuracy for training data is 73.07% and testing data is 71.4%.

References
[1] Direktorat Jendral Perkebunan 2016 *Statistik Perkebunan Indonesia 2015 – 2017 Kelapa Sawit* (Jakarta: Kementerian Pertanian Indonesia)
[2] BPS 2017 *Statistik Kelapa Sawit 2016* (Jakarta: Badan Pusat Statistika Indonesia)
[3] Shiddiq, Minarni and Umami R 2016 Analysis of relation between fluorescence intensity and
ripeness levels of loosed palm oil fruits *Jurnal Penelitian Sains* **18** p 84-88

[4] Azrifiwan 2008 *Pengembangan Mesin Sortasi Tandan Kelapa Sawit Otomatis Berbasis Teknik Pemeriksaan Non Destruktif Untuk Meningkatkan Produksi dan Kualitas CPO* Laporan Penelitian Hibah Bersaing Perguruan Tinggi 2017 Universitas Andalas

[5] Rohendar and Ransi 2013 Pendugaan tingkat kematangan buah kelapa sawit (elaeis guineensis jacq) berdasarkan pengolahan citra menggunakan metode logika fuzzy *Thesis Faculty of Agriculture University Syiah Kuala Darussalam Banda Aceh*

[6] Nataliana, Decy, et al Pengenalan plat nomor kendaraan dalam sebuah citra menggunakan jaringan saraf tiruan *Jurnal Itenas Library*

[7] Abadi A M, Wustqa D U and Pamungkas L R 2017 *Detection of Lung Cancer Using Radiograph Images Enhancement and Radial Basis Function Classifier* International Congress on Image and Signal Processing, BioMedical Engineering and Informatics

[8] Klir, George J, Ute C and Bo Y 1997 *Fuzzy Set Theory Foundations and Applications* (London: Prentice Hall, Inc.)

[9] Desrina R, Mardiningsih and Bu’ulolo F 2013 Menetukan model persamaan regresi linear berganda dengan metode backward (kasus penyalahgunaan narkoba di tanah karo) *Sainsia Matematika* 285 – 297

[10] Sri K and Purnomo H 2013 *Aplikasi Logika Fuzzy untuk Pendukung Keputusan 2nd ed* (Yogyakarta: Graha Pustaka)

[11] Permatasari D, Azizah I N, Hadiat H L and Abadi A M 2017 *Classification of Toddler Nutritional Status Using Fuzzy Inference System (FIS)* International Conference on Research, Implementation, and Education of Mathematics and Science p 1-9

[12] Zen H N H, Trimartanti L W, Abidin Z and Abadi A M 2017 *Determining Hydrocarbon Prospective Zone Using The Combination of Qualitative Analysis and Fuzzy Logic Method* System Engineering Society of China and Springer-Verlag Berlin Heidelberg p 1-12

[13] Wang and Li- Xin 1997 *A Course in Fuzzy Systems and Control* (London: Prentice- Hall International)

[14] MathWorks 2013 *MATLAB Creating Graphical User Interfaces R2013b* (America: MathWorks, Inc.)