Colour Feature Extraction Techniques for Real Time System of Oil Palm Fresh Fruit Bunch Maturity Grading

Meftah Salem M Alfatni 2,*, Abdul Rashid Mohamed Shariff 1, Osama M. Ben Saaed 2, Atia Mahmod Albbbah 2 and Aouache Mustapha 3

1 Faculty of Engineering, Universiti Putra Malaysia, 43400, Serdang, Selangor, Malaysia.
2 Faculty of Information Technology, Alasmarya Islamic University, Zliten, Libya.
3 Centre de développement des technologies avancées, CDTA, Cité 20 août 1956 Baba Hassen, Alger, Algérie.
*Email: meftah_salem@asmarya.edu.ly

Abstract. According to the natural ripeness phenomena of oil palm fruit, the oil palm fresh fruit bunch (FFB) ripeness is divided into three different categories called, under ripe, ripe, and overripe. The current method in the mills for oil palm FFB grading is manually using human graders. However, the manual grading method is subjective and prone to human errors. Thus, accurate classification of oil palm FFB with fast and easy process is necessary in oil palm farms and in the palm oil industry to grade high-quality products, especially when classifying a large amount of fruits. In this paper, an image processing algorithm scenario represented by image acquisition, segmentation, and feature extraction is implemented in real time system for oil palm FFB maturity grading. This study presents an automated inspection for the oil palm FFB using an external grading system processing based on the colour feature extraction technologies; namely, color histogram and statistical color feature (mean and standard deviation), were used to extract the oil palm FFB image’s color features. An artificial neural network (ANN) classifier as supervised machine learning technique is applied to train and test the system performance accuracy based on area under curve (AUC) of the receiver operation characteristic (ROC) as classifier performance evaluation. The result show that the colour histogram technique based on the ANN classifier was observed to be a more robust and accurate classifier, with overall 94% correct ripeness classification accuracy, for oil palm FFB maturity grading compared to the other techniques applied and tested in this study.

1. Introduction

The quality of agricultural products plays a key role in the food industry’s quality. The traditional quality assessment method for agricultural products is boring and expensive as well as giving subjective inconsistent evaluation results due to the effectiveness by human physiological factors [1]. External grading systems offer innovative solutions for industrial automation [2]. Advances in computer technologies used in the various methods and technologies utilized by many researchers around the world led to the development of novel equipment and methods for agricultural inspection, such as an automatic sorting system [3, 4]. Application of automatic sorting systems for assessment has greater than before in current years. Depending on the nature of application, a machine learning system can be used to perform external or internal grading [5-7].

The external grading system device is a combination of hardware and software [2, 5, 8]. The hardware includes different components such as a computer system, sensing device (CCD camera), frame grabber with RGB cable and feeding conveyor system [9, 10]. Whereas the software contains
robust algorithm based on image processing and one of the development software such as (MATLAB, JAVA, visual C#, C++, Visual Basic or python program language) and one of the operating system software such as (Windows, Mac, or Linux) [11-13].

The current study discusses an automated inspection system of oil palm fresh fruit bunch (FFB) based on the colour histogram and statistical colour feature (mean and standard deviation) integrated with ANN classifier technique. The image processing techniques play key roles in the oil palm inspection system by performing pre-processing, segmentation, and feature measurement.

2. Materials and Methods
The oil palm FFB maturity inspection system classifies the fruit ripeness rendering to three different classes of fruit maturity based on the Malaysian Palm Oil Board (MPOB)’s expert standard for oil palm FFB maturity called overripe, ripe, and under ripe [14-16].

Figure 1 shows the fruit image acquisition step, which involves capturing the fruit image using an appropriate charge coupled device (CCD). In this study, CCD Color Camera is utilized to acquire the oil palm FFB image. The camera was positioned upright on top of the housing through which the fruit samples were conveyed via the feeding device. The captured images are sent directly to the processing unit for image processing. An acquisition system including RGB camera cable and the frame grabber of NIPCI-8285, IEEE-1395 based on Vision 2009 are used to capture color oil palm FFB images, which are then sent to the processing unit (computer). The housing provides a controlled environment with constant, stable, and defused tubes of LED light with lighting control. An illumination filter (optical lens) is used in the path of the light beam to make the light beam selectively controllable and allow it to pass through a prearranged visible wavelength band, capturing the features, parameters, and properties of the fruits. These features provide the system with a fixed environment for the image. Figure 1 shows the housing material of the oil palm FFB used to avoid outside distortion factors based on the controlled illumination system where the fruit images were captured.

2.1. Colour Feature Extraction Techniques
The colour model combines different algorithms and routines based on image processing methods and statistical techniques, such as colour histograms and statistical colour feature techniques (mean and
standard deviation). This combination allows for studying the colour behaviour of the fruit through different types and ripeness of the FFB.

Colour offers valued information for inspecting and assessing the ripeness and freshness of the agriculture crop [10]. The electromagnetic spectral differences deliver an exclusive key for grading systems and image processing, in which the true colour contain red, green, and blue bands. The agricultural crop quality cannot be estimated just based on its shape or pattern because fruits with the same level of quality may have different shapes and patterns. To settle this issue, the machine vision system need to be able to analyse the fruit’s colour prior to achieve its value by using the RGB colour model based on its intensity. In this study, colour histogram and statistical colour feature (mean and standard deviation), are used to extract the colour features of the oil palm FFB images.

2.1.1. Colour histogram. The image’s colour histogram is built by accumulation the quantity of pixels in each colour. An picture colour histogram alludes to the likelihood mass work of the picture power [17]. The factual pixel level (SPL) highlight vector was computed from a trimmed locale concentrated colour histogram with an indicated rectangle that spoken to the picture locale. The colour histogram of a picture alludes to a colour histogram of the pixel concentrated values. The colour histogram may be a chart that shown the density of pixels in each image’s changed escalated esteem. There are 256 diverse conceivable force for an 8-bit dark scale picture, and the colour histogram will graphically show 256 numbers that appear the conveyance of the pixels among the greyscale values. The colour histogram of the grey-level pixel within the picture is given by equation 1.

\[ P(r_i) = \frac{n(r_i)}{n} \]  

(1)

Where, \( P(r_i) \) and \( n(r_i) \) are the likelihood and number of events of the grey value within the region, respectively, and \( n \) is the overall number of pixels within the locale. After the picture concentrated colour histogram was gotten from the picture locale, grey pixel levels inside the locale can be misused by computing the SPL features as follows:

1. Mean (\( m \)) grey values of the pixel within the picture can be computed utilizing equation 2.

\[ m = \frac{1}{n} \sum_{i=0}^{L-1} (r_i p(r_i)) \]  

(2)

where \( n \) is the overall number of grey values within the picture with 0, 1... L-1

2. Variance and central moments in the area is calculated by means of equation 3.

\[ \mu_n = \sum_{i=0}^{L-1} (p(r_i)(r_i - m))^n \]  

(3)

where the second central moment \( \mu_2 \) is the fluctuation of the locale, and the other moments as third and fourth is calculated, respectively, for \( n = 3 \) and \( n = 4 \).

3. Total energy (\( E \)) of the pixel’s grey value is known by equation 4.

\[ E = \sum_{i=0}^{L-1} (p(r_i))^2 \]  

(4)

4. Equation 5. express the scattering of grey values represented by measured information as the Entropy (\( En \)).
5. Skewness ($S$) is a measure of the lack of symmetry and its zero value when the colour histogram is symmetrical. Skewness ($S$) will be negative on the off chance that the cleared out hand tail is longer and positive on the off chance that the correct hand tail is longer. Skewness for data $y_1, y_2, y_3, ..., y_n$ is given by equation 6.

$$S = \frac{1}{n} \sum_{i=1}^{n} (y_i - \bar{y})^3 / (n-1)s^3$$

where, $\bar{y}$ is the mean, $s$ is the standard deviation, and $n$ is the quantity of data points.

Mathematically, the colour histogram is shaped by determining the colours inside the any image by computing the quantity of pixels in each colour. The colour chart technique is very simple, has low level complexity, and has produced good results in practice. The number of bins per colour component has been fixed to 256, and the dimension of each colour histogram is 2563. Figure 2 illustrates the mechanism used for obtaining the image colour histogram for the three colour-image channels of the three ROIs in the image to be used for system training.

---

**2.1.2. A Statistical colour feature.** Colour histograms have been widely and successfully used in grading system applications such as oil palm FFB ripeness grading [17]. However, spatial information was not included in the colour histogram results, and thus, statistical colour features were extracted by using the colour feature algorithm such that both colour and spatial information can be integrated. Statistical colour features (mean and standard deviation) of pixel colour values were used to classify colour images in this work, and were found to be a basic however capable strategy for extracting colour features. The colour feature statistical mean ($\overline{p}$) is described by equation 7.

$$\overline{p} = \frac{1}{q} \sum_{i=1}^{q} p_i$$

---

![Figure 2](image-url)
The standard deviation (S) is described by equation 8.

\[ S = \frac{1}{q-1} \sum_{i=1}^{q} (P_i - \bar{P})^2 \]  

(8)

where q signifies the total of pixels in each image. The measurements were used as the feature values of object pixel colours that were used as training data, as illustrated in Figure 3.

Figure 3. Methodology of image RGB statistical colour features extraction used for system training: (a) ROI1, (b) ROI2, and (c) ROI3.

2.2. System classification (decision-making)

Image sorting that eventually leads to the recognizable proof of the FFB maturity level is the ultimate image-processing step of the oil palm FFB classifying system. In image classification, the impacts of utilizing color histogram and statistical color feature extraction procedures with supervised machine-learning strategies are assessed. The artificial neural network (ANN) classifier is utilized with the three different region of interests (ROIs) including ROI1, ROI2, and ROI3. Training and testing are implemented as classification part.

2.2.1. Artificial neural network (ANN). The ANN supervised machine learning classifier forward back propagation model was utilized in this study. [8] presented the structure and design of ANN classifier. In the current study the ANN classifier built on the different input image sizes described above, the ANN structures were constructed as [40 20 1], [35 20 1], and [30 15 1] models for classes 1, 2, and 3, respectively. Nodes 40, 35, and 30 at the first input layer and nodes 20, 20, and 15 at the second hidden layer connected with node 1 as the output for the three classes that will be indicated to the FFB ripeness of the grading system testing.

2.2.2. An Assessment of Classifier performance. The area under (ROC) curve are utilized for classifier performance assessment based on the control of positive predictive value (PPV) and negative predictive value (NPV) as described by equation 9. [18].

\[ PPV = \frac{TP}{TP + FN}, \quad NPV = \frac{FP}{TN + FP} \]  

(9)
Where where TP is the number of true positives (correctly classified maturity); TN is the number of true negatives (wrongly classified maturity); FP is the number of false positives (ripeness classified as not ripeness); and FN is the number of false negatives (not ripeness classified as ripeness).

In the experiments performed in this study, 180 images in the training set and 270 images in the testing set were divided into three ripeness groups, namely, under ripe, ripe, and overripe. In order to avoid the natural image distortion at random FFB image parts as well as investigate the system performance with different size of entire FFB image, all the FFB images were segmented and analysed based on three different image sizes called region of interest (ROIs), ROI1 (300 × 300 pixels), ROI2 (50 × 50 pixels), and ROI3 (100 × 100 pixels). The transformed images based on the color histogram and statistical color feature extracted from the FFB sample images were investigated to obtain the accuracy of the technique with the ANN machines-learning classifier.

Furthermore, the MATLAB programs were used for all grading system image processing steps and experimental results.

3. Results and Discussion

The maturity of oil palm FFB called over ripe, ripe and under ripe was properly classified based on the robust algorithm procedure of the oil palm maturity grading system. The threshold points between true positive rate and false positive rate is plotted and displayed thorough a series of the ROC curve points. The area under the ROC curve is considered a successful degree of the inalienable legitimacy of a grading system test.

3.1. FFB ripeness grading system

To classify the FFB ripeness classes, several feature extraction techniques were used. These techniques included statistical colour feature and colour histogram for the colour model. In addition, the FFB ripeness decision-making process was built based on ANN supervised machine-learning classifier.

The effectiveness of the statistical colour feature (mean and standard deviation) and colour histogram feature extraction techniques has applied and verified with the ANN classifier. The execution of the three distinctive feature extraction strategies based on the three distinctive image ROIs were evaluated and compared with each other for high performance accuracy.

The experiments were conducted. In experiment 1, the statistical colour features were extracted, and the classification was implemented by using ANN as shown in Figure 9. The same process was implemented in experiment 2 with colour histogram features and ANN classifiers, as shown in Figure 4.

![Statistical Feature Extraction with ANN Classifier](image)

**Figure 4.** Oil palm FFB ripeness classification based on color feature extraction with ANN.
Figure 4 show that the oil palm FFB maturity sorting is implemented by the ANN classifier with high accuracy results with different image’s sizes ROI1, ROI2, and ROI3 to recognize the statistical features of the dataset using the colour model. As illustrated in Figure 9, the results indicated that statistical colour feature extraction based on the ANN classifier is proficient of execution a greater sorting with ROI1 based on ROC graph, at an AUC test accuracy up to 93% classification rate.

Figure 5 show that the oil palm FFB maturity sorting is implemented by the ANN classifier with high accuracy results with different image’s sizes ROI1, ROI2, and ROI3 to recognize the colour histogram features of the dataset using the colour model. As illustrated in Figure 10, the results indicated that colour histogram feature extraction based on the ANN classifier is proficient of execution a better classification with ROI3 based on ROC graph, at an AUC test accuracy up to 94% classification rate. In general, the FFB maturity classification implemented by the oil palm grading system to distinguish FFB ripeness class features by utilizing statistical colour features and colour histogram demonstrations appeared reasonable execution comes about (all cases AUC > 90%) for FFB maturity classification in colour model datasets

3.2. FFB ripeness classification
The maturity classification task as training and testing for overripe, ripe, and under ripe as close classes, which tested and validated based on the colour model of the real-time oil palm FFB ripeness grading system. The performance of the real-time oil palm FFB maturity classification as training and testing based on the colour model for the three different FFB image ROIs has evaluated. The results has clearly demonstrated in Table 1 and Table 2, correspondingly.

**Table 1.** Results of training computing FFB ripeness classification based on statistical and histogram using ANN, KNN, and SVM.

| Classifier | Image size | Colour feature extraction techniques |
|------------|------------|--------------------------------------|
|            |            | Statistical colour feature          |
|            |            | Colour histogram                     |
|            |            | Training accuracy (%)                |
|            |            | Training accuracy (%)                |
| ANN        | ROI1       | 100                                   |
|            | ROI2       | 100                                   |
|            | ROI3       | 100                                   |

**Figure 5.** Oil palm FFB ripeness classification based on colour histogram with ANN.
Table 2. Results of test computing FFB ripeness classification based on statistical and histogram using ANN, KNN, and SVM

| Classifier | Image size | Statistical colour feature | Colour histogram |
|------------|------------|-----------------------------|------------------|
|            |            | Time (S) | Testing accuracy (%) | Time (S) | Testing accuracy (%) |
| ANN        | ROI1       | 93       | 3                   | 92       | 2.7                  |
|            | ROI2       | 92       | 2.5                 | 92       | 1.4                  |
|            | ROI3       | 92       | 2.65                | 94**     | 1.6**                |

Table 1 and Table 2 indicated the optimal methods and techniques, which are fastest and accurate for oil palm FFB maturity classification. The statistical colour and colour histogram feature extraction technique combined with the ANN supervised machine learning technique as well as applied on the 100x100 pixels FFB image size with ROI3. The results achieved 100% training performance, 93% accuracy, and 1.6 second image processing speed as testing performance for Nigrescens and Oleifera. The colour histogram feature extraction technique combined with the ANN supervised machine learning technique was applied to the 50x50 pixels FFB image size with ROI2. This technique achieved a training performance accuracy of 100%, 93% testing performance accuracy, and 1.4 second image processing speed. For Virescens, the statistical colour feature accurately gave higher training and testing performance accuracy of 100% and 93%, respectively, based on the ANN supervised machine learning for the different oil palm types. However, the results were limited by the slow processing time compared with colour histogram performance as well as the oil palm system objectives.

4. Conclusion
This study presents an automated inspection for the oil palm FFB using an external grading system processing based on the colour feature extraction technologies; namely, color histogram and statistical color feature (mean and standard deviation) to extract the color features of oil palm FFB image. An image processing algorithm scenario represented by image acquisition, segmentation, and feature extraction is implemented in real time system for oil palm FFB maturity grading. Different image region of interests (ROIs) based on different image sizes; ROI1 (300x300) pixels, ROI2 (50x50) pixels, ROI3 (100x100) pixels. An artificial neural network (ANN) classifier as supervised machine learning technique is applied to train and test the system performance accuracy based on area under curve (AUC) of the receiver operation characteristic (ROC) as classifier performance evaluation. The result show that the colour histogram technique based on the ANN classifier with ROI3 was observed to be a more robust and accurate classifier, with overall 94% correct ripeness classification accuracy, for oil palm FFB maturity grading compared to the system based on the ANN with ROI2 and ROI3, which obtained results of 92% each of them.

References
[1] Wu, D. and D.-W. Sun, Colour measurements by computer vision for food quality control–A review. Trends in Food Science & Technology, 2013. 29(1): p. 5-20.
[2] Alfatni, M.S., et al., Recent Methods and Techniques of External Grading Systems for Agricultural Crops Quality Inspection - Review. International Journal of Food Engineering, 2011. 7(3): p. 1-40.
[3] Saeed, O.M.B., et al., Classification of oil palm fresh fruit bunches based on their maturity using portable four-band sensor system. Computers and Electronics in Agriculture, 2012. 82: p. 55–60.
[4] Misron, N., et al., Fruit Battery with Charging Concept for Oil Palm Maturity Sensor. Sensors, 2020. 20(226): p. 2-14.
[5] Alfatni, M.S.M., et al., Real-time oil palm FFB ripeness grading system based on ANN, KNN and SVM classifiers. *IOP Conf. Series: Earth and Environmental Science*, 2018. 169: p. 1-5.

[6] Ariffin, N.I.K., et al., Fruits Recognition based on Texture Features and K-Nearest Neighbor. *International Journal of Engineering & Technology*, 2018. 7(4.31): p. 452-458.

[7] Riese, F.M., S. Keller, and S. Hinz, Supervised and Semi-Supervised Self-Organizing Maps for Regression and Classification Focusing on Hyperspectral Data. *Remote Sens*, 2020. 12(7): p. 1-23.

[8] Zhao, W., et al., Hyperspectral images classification with convolutional neural network and textural feature using limited training samples. *Remote Sensing Letters* 2019 10(5): p. 449-458.

[9] Zemmour, E., P. Kurtser, and Y. Edan, Automatic Parameter Tuning for Adaptive Thresholding in Fruit Detection. *Sensors*, 2019. 19(2130): p. 1-21.

[10] Boukar, O., et al., Food Non-Destructive Quality Evaluation Using Color Image Analysis System. *International Journal of Engineering Research and Advanced Technology*, 2019. 5(8): p. 60-65.

[11] Guo, R., et al., Pixel-Wise Classification Method for High Resolution Remote Sensing Imagery Using Deep Neural Networks. *ISPRS Int. J. Geo-Inf*, 2018. 7(110): p. 1-23.

[12] De-la-Torre, M., et al., Multivariate Analysis and Machine Learning for Ripeness Classification of Cape Gooseberry Fruits. *Processes*, 2019. 7(928): p. 1-16.

[13] Alfatni, M.S.M., et al., Oil palm fresh fruit bunch ripeness classification based on rule-based expert system of ROI image processing technique results. *Earth and Environmental Science*, 2014. 20: p. 1-8.

[14] MPOB, ed. *Oil Palm Fruit Grading Manual*. Second Edition ed. 2003, Malaysian Palm Oil Board, Ministry of Primary Industries: Bandar Baru Bangi. 40.

[15] Alfatni, M.S.M., et al., Oil palm fresh fruit bunch ripeness classification based on rule-based expert system of ROI image processing technique results. *IOP Conference Series: Earth and Environmental Science*, 2014. 20(1): p. 012018.

[16] Alfatni, M.S.M., et al., Oil Palm Fruit Ripeness Grading System Based on BGLAM Feature Extraction and SVM, KNN and ANN Supervised Machine learning. *Journal of Advanced Science Letters*, 2013. 19(9): p. 2612-2614.

[17] Gomashe, A.S. and R.R. Keole, A Novel Approach of Color Histogram Based Image Search/Retrieval. *International Journal of Computer Science and Mobile Computing*, 2015. 4(6): p. 57-65.

[18] Tilaki, K.H., Receiver Operating Characteristic (ROC) Curve Analysis for Medical Diagnostic Test Evaluation. *Caspian J Intern Med*, 2013 4(2): p. 627-635.