Available online at www.ijournalse.org

Emerging Science Journal

(ISSN: 2610-9182)

Vol. 6, No. 4, August, 2022

External Features-Based Approach to Date Grading and Analysis with Image Processing

Shabana Habib [1], Ishrat Khan [2], Suliman Aladhadh [1], Muhammad Islam [3*], Sheroz Khan [3]

[1] *Department of Information Technology, College of Computer, Qassim University, Buraydah, 52571, Saudi Arabia.*

[2] *Software Engineering Department College of Computer and Information Sciences, Prince Sultan University, Saudi Arabia.*

[3] *Department of Electrical Engineering, College of Engineering and Information Technology, Onaizah Colleges, Onaizah, 56447, Saudi Arabia.*

Abstract

The analysis and classification of dates is based on their external features: size, appearance, shape, and colour. The process is currently performed manually after harvesting as part of the post-harvesting process. Grading manually is tedious because it usually results in time delays, product quality risks, and it is associated with time and cost delays as well. Although the use of computers and information technology has seen tremendous growth in many small and large sectors, it has been in its infancy in the cultivation of fruit and dates. Using image processing algorithms, we can enhance human vision capabilities through analysis and make images easier to comprehend. A major objective of computer vision-based algorithms for classifying and sorting of dates is to make the procedure fully automated by minimizing the manual component involved in the process. This paper presents an image processing-based algorithm that uses machine learning techniques to extract the characteristics of colour intensity and colour homogeneity, allowing us to grade images in a more timely and automated manner. In order to obtain the results, we extracted the appearance of the date images based on an image processing algorithm. It is used as a validation element for the results that the quality of dates-fruit images can be evaluated through the prior selection process in both separate and in groups. This study has managed to achieve a rate of 95% accuracy in data classification.

Keywords:

Feature Extraction;
Segmentation;
Threshold;
Classification;
Edge detection.

Article History:

Received:	11	December	2021
Revised:	06	April	2022
Accepted:	23	April	2022
Available online:	29	May	2022

1- Introduction

Palm trees are among the oldest plants on earth and have long been considered an important source of sustenance in most desert areas. As one of the ten largest producers of dates in the world, Saudi Arabia produces nearly 400 varieties of dates in large quantities and is one of the top ten countries that produce dates [1]. The Kingdom of Saudi Arabia is considered one of the most important countries when it comes to the production of dates. In the Kingdom of Saudi Arabia, the number of date trees has increased from 13 million in 1990 to over 25 million in 2011 [2]. Statistics show that, since 2005, the annual production rate has been around 1100 thousand tons [3]. Electrical impedance spectroscopy (EIS) is a powerful technique that has been used in a wide variety of applications, from the characterization of food products [4]. Essentially, computer vision is a descriptive algorithm for describing images and is useful for describing things, such as fruit [5], where an automatic characterization system has recently been used by a system to determine fruit quality, size, and ripeness based on their appearance. A combination of sensor techniques and computer statistics

* **CONTACT**: m.islam@oc.edu.sa

DOI: http://dx.doi.org/10.28991/ESJ-2022-06-04-03

has been used to check the quality and control of food products [6]. Image processing base features extraction is used to determine the quality rating [7]. An automated classification method for date fruits, which utilizes both computer vision and pattern recognition, is presented in [8]. Date fruit can be classified by using image analysis and pattern recognition in order to classify it. Differences in view, distance, and lighting exposure all present obstacles to performing a reliable classification, and as a result, classifications are not reliable. In order to make a successful classification, it is important to deal with similarities and differences between classes cautiously. Therefore, various studies have been performed on the recognition and classification of fruits based on the visual features extracted from the images [9].

Similarly, in this paper, we attempt to classify the fruits of dates in terms of shape and quality by using external characteristics. On the other hand, it is common for the majority of fruit quality detection and grading systems to be inefficient from the standpoint of speed, cost, and complexity. Hence, a low-cost and high-speed fruit-size estimation and fruit classification system should be developed by creating an image processing algorithm in MATLAB tools based on the RGB color scheme, determining each fruit maturity level. In order to derive the required statistics and results from image analysis, it is necessary to use image processing software such as MATLAB to derive the required algorithms and to facilitate the identification and classification of fruit quality, which has otherwise been performed through the use of the naked human eye. Even though the manual procedure for identification and distinguishing chromatic differences can be accurate and reliable, it takes a considerable amount of time and effort. The purpose of this paper is to demonstrate how computer imaging can be used to obtain data from color images taken by a color digital camera, consisting of three colors: Red, Green, and Blue. There are approximately three colors in relation to the intensity and contrasts of each, which can be used to derive millions of different colors, guaranteeing the best in furnishing an accurate classification of date fruits.

Dates are considered to be the most popular fruit all over the world, especially in the Middle East peninsula. Currently, there are over 40 different types of dates, and more than 400 varieties of dates, all of which offer a wide range of taste, color, and shape that contribute to grading dates with proportionately varying prices. Dates have a lot of health benefits and are delicious. Since dates are an important source of vitamin C, and so it has become necessary to classify them into different grades in order to increase their nutritional value. Dates are produced in such large quantities that it is difficult for human vision to distinguish between them as shown in Figure 1.

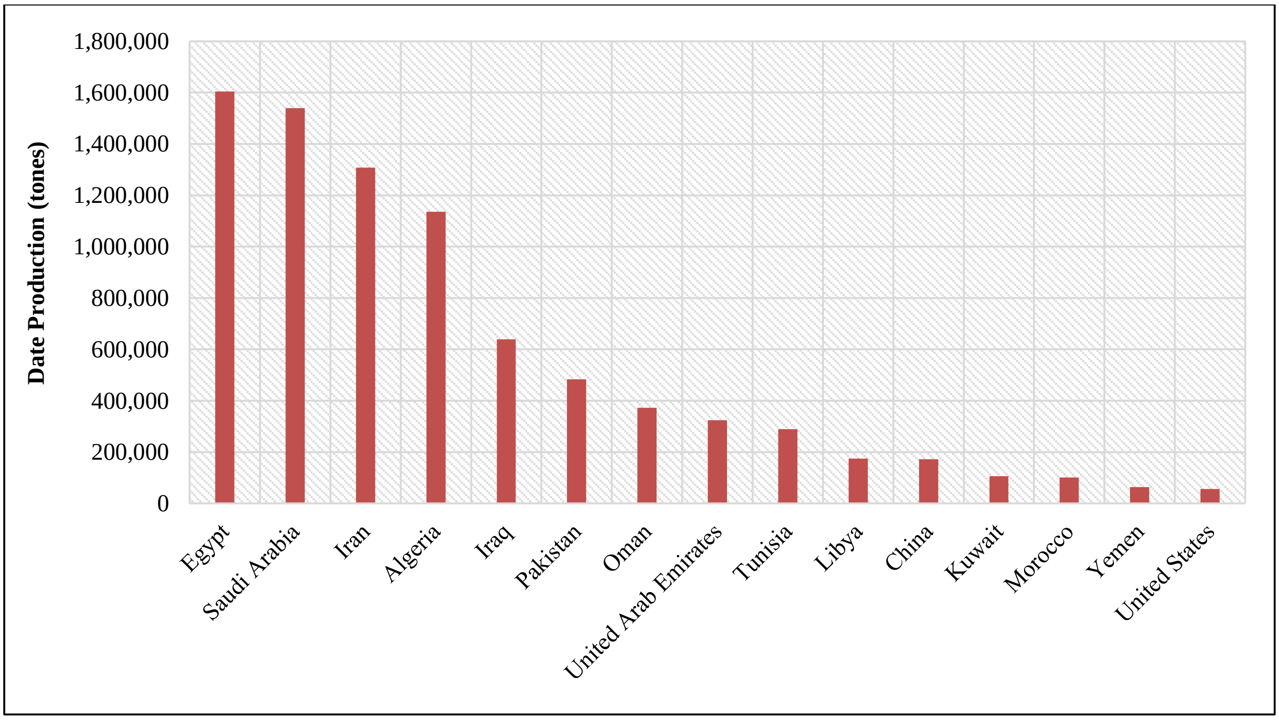


Figure 1. Top Countries in Dates Production (FAO 2019) [10]

A wide variety of date fruit is cultivated across the world, and it is used to make food products, medical products, and cosmetics for different purposes [11]. The distinction between date fruit types is a difficult task, due to the fact that they have different nutritional values, different consumption times, different prices, and varying quality [12]. It has been shown that computer vision systems can be used quickly to distinguish between the quality, size, and variety of dates without the need for an expert [13]. Thus, the process from the stage of production to the stage of consumption can be made shorter [14]. In the agricultural industry, it is possible to increase efficiency and productivity by using this type of agricultural technology [15, 16].

There has been a lot of research from different countries in the world, most of them using software tools and hardware devices to do different types of classification. The purpose of these researchers is to facilitate the classification process of dates used in agriculture. This paper describes the background of the computer vision algorithm and the principle of RGB imaging. It is proposed that a framework and block diagram of an automated system for categorizing dates be included in the paper. According to the date fruit classification system, there are mainly four aspects that make up the external characteristics of date fruits: color, size, absence of defects and physical characteristics such as being firm or strong. Computer vision techniques have been used to measure the various internal and surface characteristics of dates. Based on previous studies, the average classification accuracy (three grades) for Sukkry variety of dates has been 88% and for Maneefi variety of dates it has been 93%, while using a multilayer neural network trained with a back-propagation algorithm. In this study, we developed an automatic classification system for date fruits by using reflective near-infrared imaging to obtain reflective images of the upper surfaces of dates. The class percentage has been calculated from the images and is applied as a method of classifying the fruits. In total, the accuracy of the system from 79% to 95% for different grades [17]. Cracks on dates' surface are detected with the use of a RGB camera with a classification accuracy ranging from 58% to 78% in the three-class model (no crack, low crack, high crack), and from 75% to 88% in the two-class model (with-crack and without crack) using statistical classifiers. The RGB components of all the images are analyzed using an algorithm developed in the MATLAB programming language. It is important to note that in each image, a portion of the dates' sample has been separated from the background using a simple threshold value-based method combined with the use of morphological processing.

2- Material and Methods

This section will describe the process of acquiring date fruit images for features extraction, as well as logistic regression (LR), intensity-based algorithms, and stacking methods used for classification. Moreover, the performance metrics of the classification results will also be provided in this section. It is generally considered that a date classification system generally consists of five basic components: 1) Fruit placer, 2) Image capturing by camera, 3) fruit grading (frame grabber or digitizer), 4) sorting boxes, and 5) covey belt to carry the dates as illustrated in Figure 2.

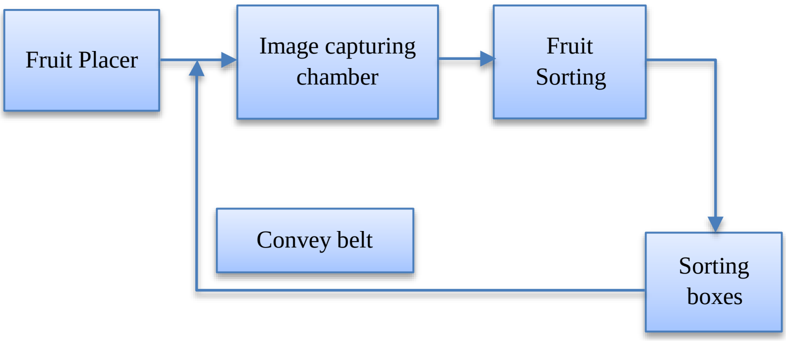


Figure 1. Block diagram of Dates Grading System

3-1- Sample Collection

During this study, we selected Sukkry dates, one of the most commercially produced dates in the world. Samples of dates were collected from different dates' fields located in the Al-qassim region, Saudi Arabia. Those samples were mixed with the samples collected from other fields. A diverse set of samples was selected from which three grades of dates were extracted (high class, middle class and low class). Figure 3 illustrates how a date quality expert verified the quality of the graded samples. There were a total of 120 dates (40 dates for each class) that were analyzed and used for quality estimation.

Figure 2. Dates grading quality experts

3-2- Image Acquisition

Images of different dates samples were collected in the following categories: dates that have been fully ripened, dates that have not yet fully ripened, and incipient dates just beginning to ripen, that have not yet been completely ripe. There is also the inclusion of dates that contain multiple levels of cracks, and based on the study and measurement of the extent of computer vision when sorting dates and categorizing them as ready to be sold or not. In order to capture the dates pieces for this set, a Nikon D3200 camera was used, with a resolution of 5184 × 3456 pixels. The camera was mounted 20 cm above the samples being collected.

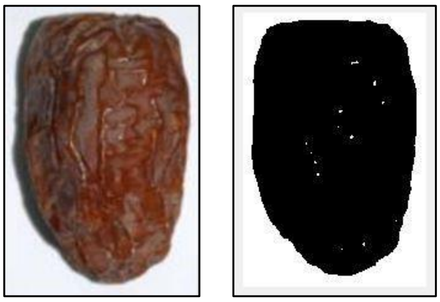
3-3- Image Analysis

The obtained images were analyzed using MATLAB software (version 9.2.0 MathWorks U.A). The images were segmented using the threshold method in order separate them from the background. Considering the skin and variegated size of the variety of dates are not the same as the normal color and size, based on the characteristics of the skin and size extracted. A description of the features that were extracted is given in Table 1.

Table 1. Feature extraction for image segmentation

Features	Description
Red Intensity	Intensity of red component in an image
Green Intensity	Intensity of green component in an image
Blue Intensity	Intensity of red component in an image
Gray Intensity	Grayscale intensity from 0-255
Brown Area	Area extracted from image using threshold
Green Area	Area extracted from image using threshold
Crack Area	Area extracted from image using binary and threshold

After the images have been captured or taken, they are transferred to a computer to be processed by the MATLAB software. The captured images are converted into digital images to improve the image quality by eliminating unwanted distortions referred to as "problems" as shown in Figure 4. The input data is then converted into digital format by the image processing software, which performs the necessary operations on the digital data. The pre-processing of an image is described as the processing of converting it to black and white images as shown in Figure 4, which are represented by numbers ranging from 0 to 255, where 0 is represents black and 255 represents white.



Original image Pre-processing image

Figure 3. Segmented images using pre-processing

3- Segmentation

It is a process which consists of cutting an image, adding features, analyzing them, as well as dividing an image into regions that have a strong correlation with objects or areas of interest using the principle of matrix analysis. It is referred to either as threshold-based segmentation or region-based segmentation.

4-1- Threshold-Based Segmentation

As shown in Figure 5, it is used to characterize images while claiming to be based on the constant reflection or light absorption of the surface of the object. A threshold was used in this method to separate the different skin images from the set of images obtained from the different fields. We chose five different images of skin shapes and sizes in order to identify the different periods of ripening. In the next step, the intensities of the different ripened images were then measured using MATLAB's data cursor tool. We used the calculated valued to separate images that were at different stages of ripening. Finally, the ripened area of different skin shapes was calculated by the pixel-shifting of the foreground portion in the threshold image. The procedure for running this method can be used to determine color maturity and feature-based defects related to skin defects and damages via the sorting such as shown in Figure 5.

Original Image Extracting Image

Figure 4. Extracting image using Threshold

4-2- Texture Feature

As a result of this method, the intensity range of each pixel is separated by calculation that can be applied to the pixel set of the extracted image. The purpose of texture feature extraction is used to enhance or accentuate specific aspects of an image so that those aspects can be taken into account and recognized. In order to convert the ripened shape of each class into effective value, five different ripened shapes (high class, middle class and low class) were selected. A texture value that contains a significant amount of information that is useful for image recognition can be calculated. The image may be completely duplicated, or a set of small variations in different categories is selected depending on the size of the image, may be selected. Despite the idea that texture can be roughly defined as a combination of some of the characteristics of an innate image, such as including smoothness, roughness, fineness, granulation, randomness, lineation, the state of being in a hammock, etc., a strictly scientific definition has not yet been determined.

4-3- Classification

In the classification step, we can classify dates into different classes based on the dates quality, using threshold and edge detection algorithms. Once the quality of the dates has been established, they are ready to be sold or returned to the factory. As can be seen in Figure 6, dates are classified into different categories based on their quality.

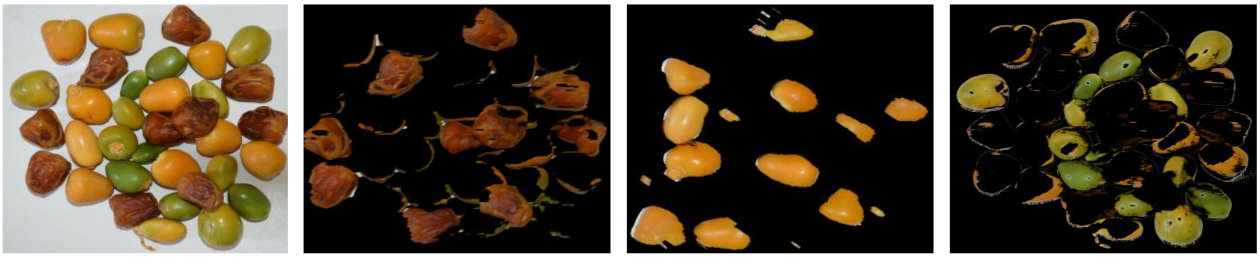
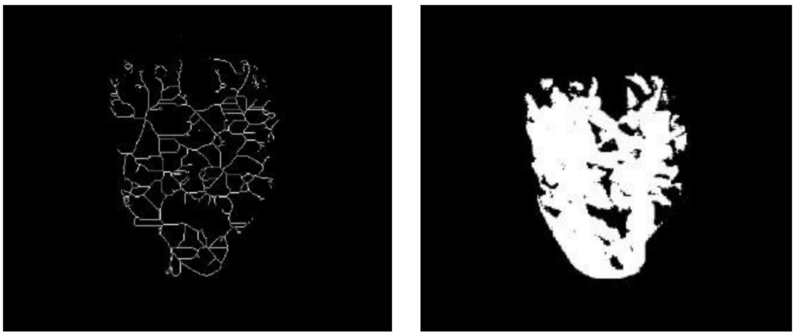


Figure 5. Segmentation and classification of dates

4-4- Edge Detection

The accuracy of meaningful edge detection is critical the successful classification of many image processing tasks. There is a technique that is used to detect intensity interruptions in a digital image. In this context, edge detection is an essential tool to help in segmenting an image as it will detect the edge of the image. In order to detect cracks in the dates, the intensity values obtained are used as threshold values. In the final step, the crack area is calculated by counting the pixels that cover the foreground area in the threshold image. The operational procedure in this method is used to detect the skin of the image using the image processing algorithm as shown in Figure 7.



Edge detection image Skin detection image

Figure 6. Detection the skin effected images

4-5- Flowchart

The flowchart of the proposed date classification system is shown in Figure 8. The image captured with a coloured background without a specific size is used as input of the system. The image is entered into the system and then pre-processing algorithms are applied in order to distinguish between the image and the foreground. It is in this step that colour image is converted to a black using histogram technique. According to the histogram technique the image is converted into three separate colour levels and then checked for colour intensity {Red (R), Green (G), and Blue (B)}. Finally, a feature extraction algorithm is used applied to classify the images into different categories in order to determine the quality of the dates.

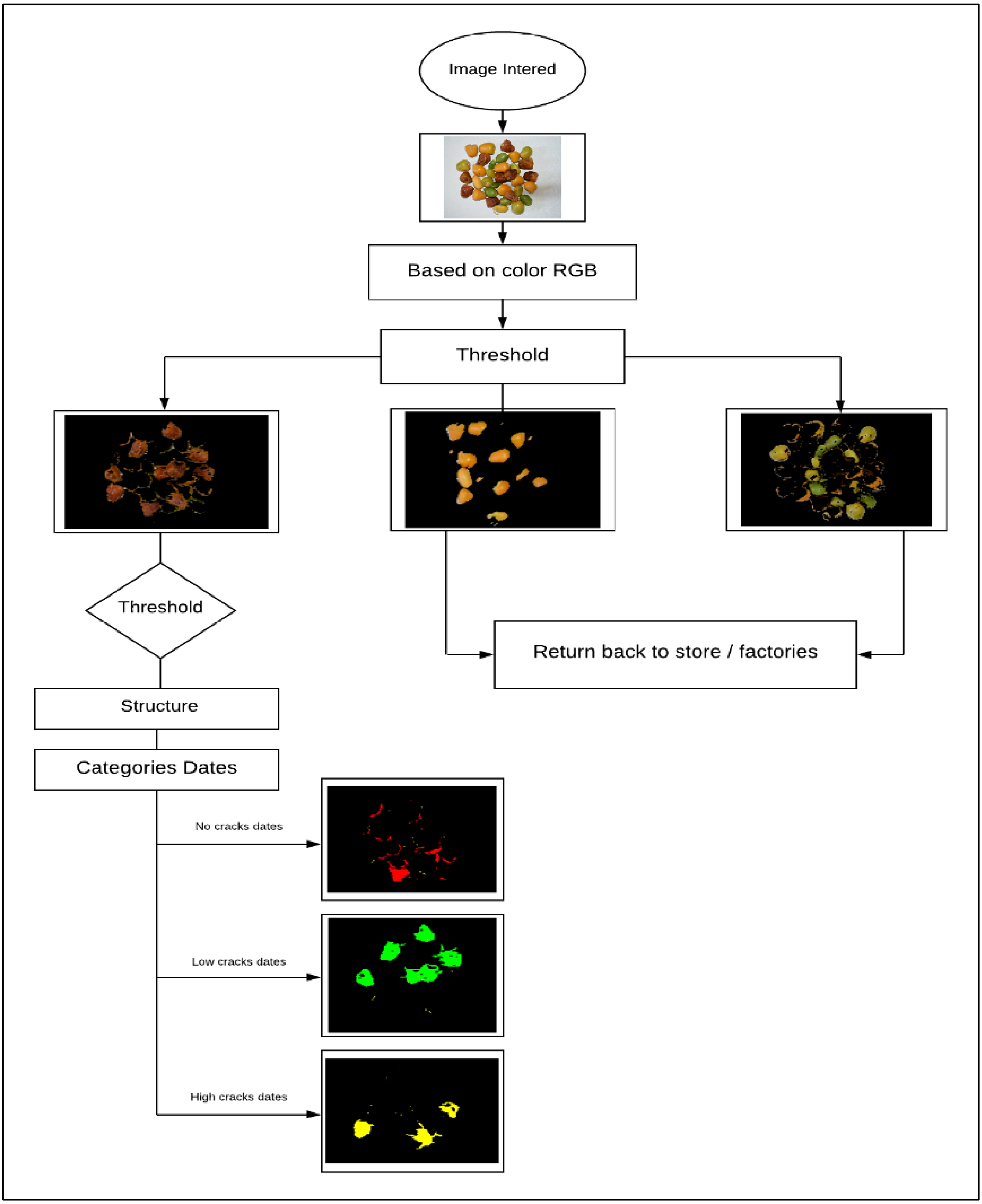


Figure 7. Flow chart describing the classification process

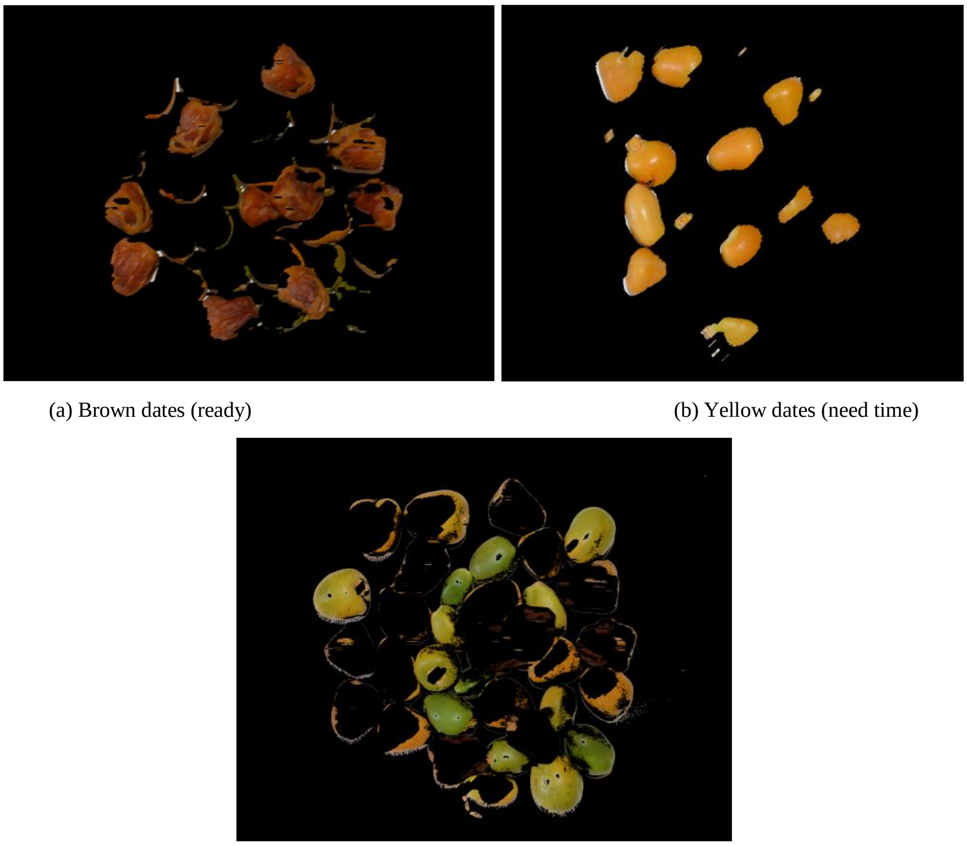
4- Results and Discussion

As can be seen from Table 2, the MATLAB tool has been used to extract the features for the quality of the three grades. The red, blue and green intensities of the high-grade were significantly different from those of other two categories of grades. The intensity of saturation of the dates of the middle class as well as the dates of the low class was similar. There was no difference in the intensity of grayscale between the three classes of dates.

Table 2. Shows comparison of the values of crack

An example of a column heading	High Class	Middle Class	Low Class
Red intensity	5×10^{10}	1.2×10^{10}	1.3×10^{10}
Blue Intensity	4.95×10^{10}	1.3×10^{10}	1.25×10^{10}
Green Intensity	5.25×10^{10}	1.02×10^{10}	1.20×10^{10}
Grayscale Intensity	4.95×10^{10}	5×10^{10}	5.05×10^{10}

In Figure 9, show the representative images of date samples in each grade after applying the extracted color algorithm. The edge effect is also processed on all images to ensure that the dates are of highest quality. The problem was that while trying to remove the edge effect in the image with the filter, many processed photos were thrown away. It would appear that the dates peel has the same intensity corresponding to the same color and, therefore, can be identified as a gradient region with threshold images in the case of depletion to detect edges. It was observed that in some of the samples of dates, the nature of the peel was glossy, giving the dates a similar reflective area and a brightness to what was detected in the grade samples. It can be seen that the percentage of graded samples is shown in Table 3. The accuracy of the high grade was 93%, the accuracy of the middle class was 94.5%, and the accuracy of the low class was 96.3%.



(a) Brown dates (ready)

(b) Yellow dates (need time)

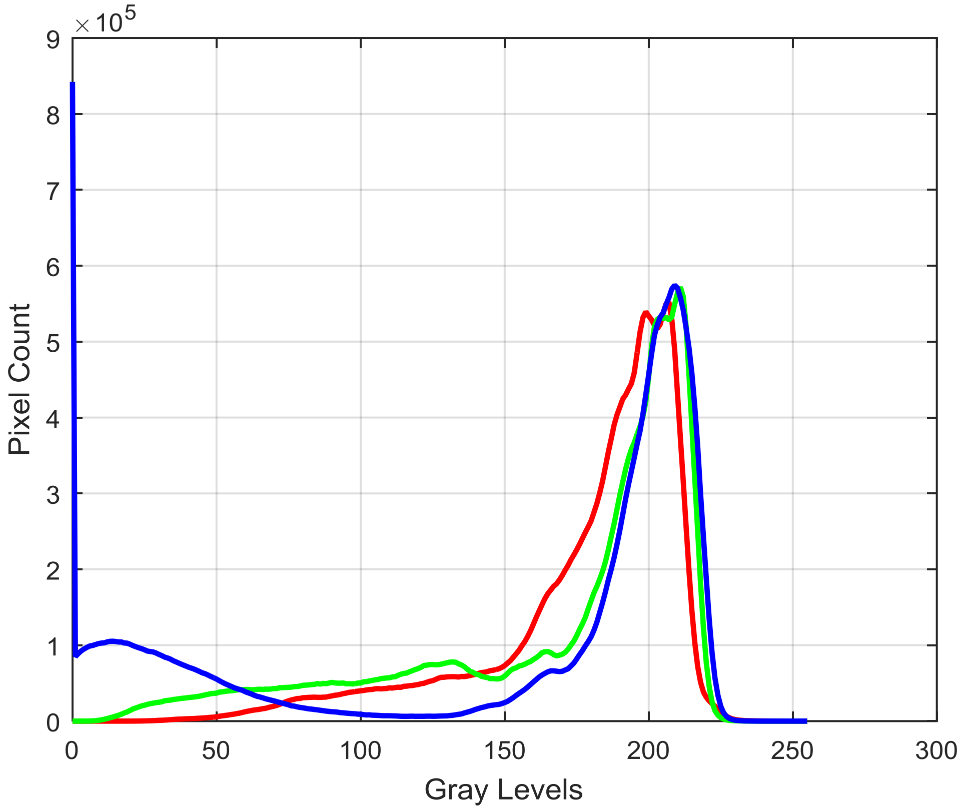
(c) Green dates (back to store)

Figure 8. Classification images with their ripened period

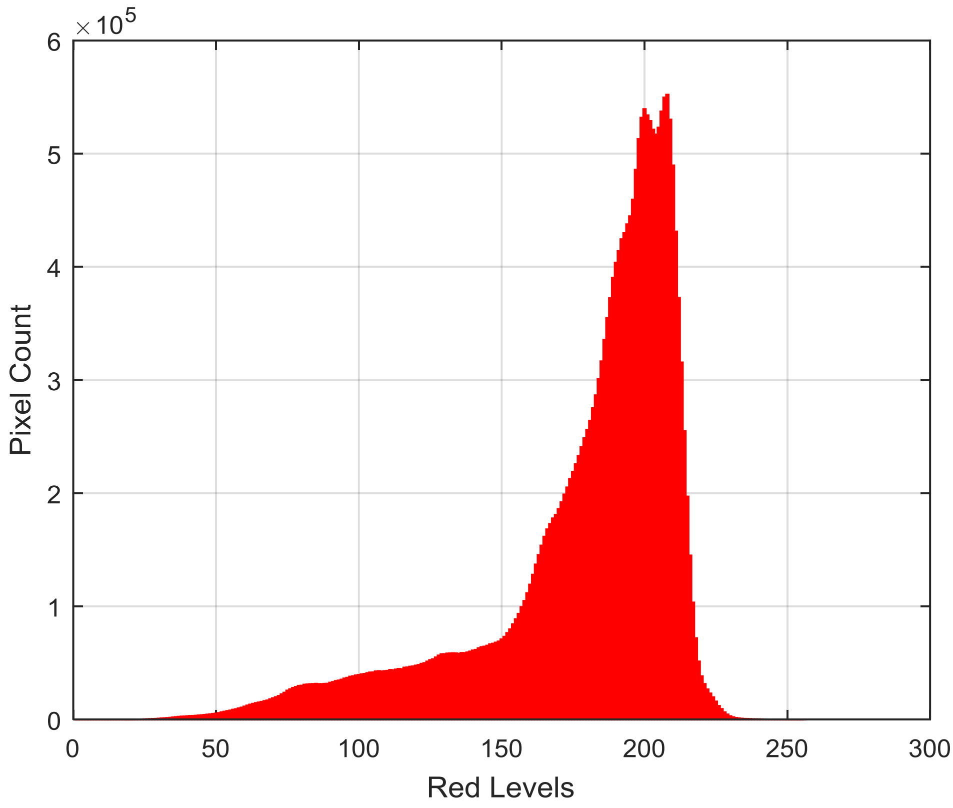
Table 3. Percentage of graded samples

Grade Class	Intensity Value
High class	8901±126000 = 7.06%
Medium class	4001±89000 = 4.5%
Low class	1500±40000 = 3.75%

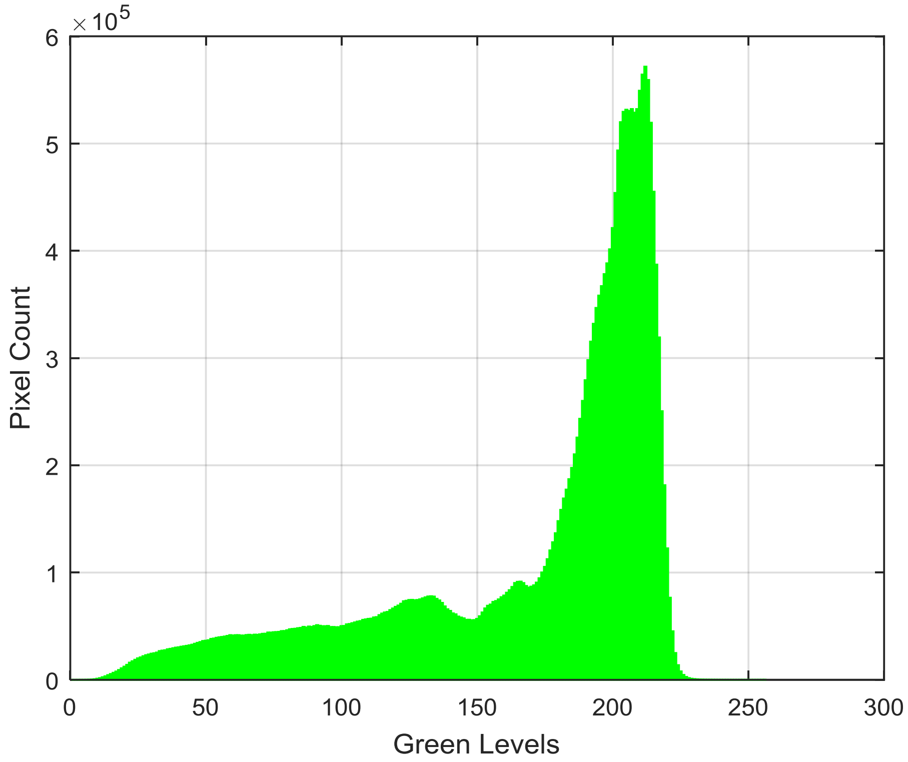
A histogram of all the RGB bands combined together in Figure 10 shows the number of pixels at the virtual and gray levels on the horizontal axis, and the RGB histograms for each color are shown on the vertical axis. The horizontal axis represents the pixel density and the vertical axis represents the average pixel size. Figure 10 shows histograms of the performance of each category.



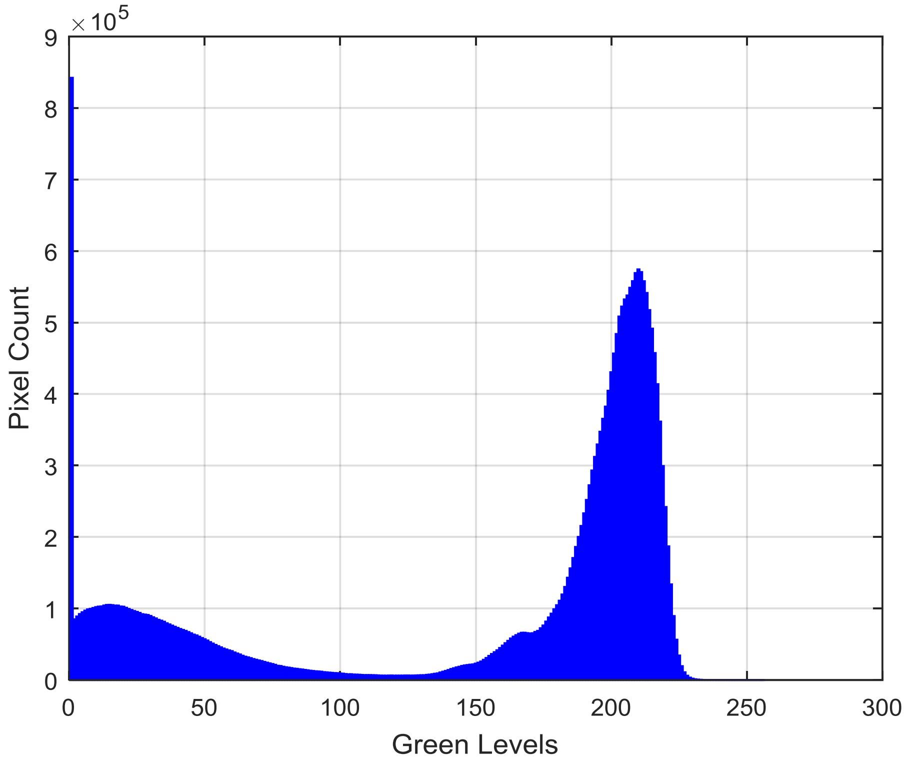
(a) All Bands



(b) Red Band



(c) Green Band

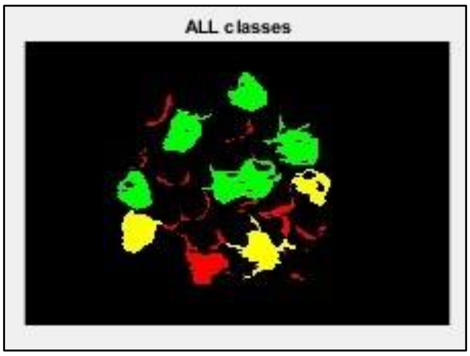


(d) Blue Band

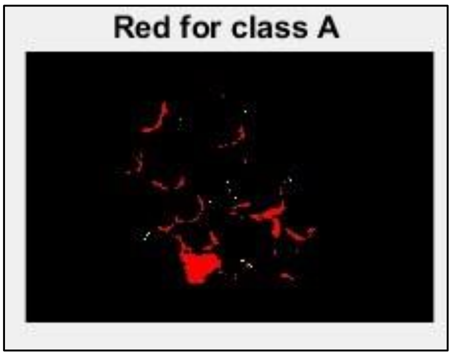
Figure 9. Histogram based performance of each class

The color segmentation matrix shows that, according to the color segmentation matrix, Sukkary and Barhee have the highest classification success, while mixed dates have the lowest classification success. It is also used in image processing to evaluate the performance of a classification problem that is carried out with the help of a histogram. As can be seen from the histogram (Figure 10) presented for each color fruit in the study, the x-axis displays the color ratio and the y-axis displays the pixel intensity ratio. Based on the histograms, it is evident that the successful performance is high in Sukkary, and the Barhee date fruits are the best when it comes to successful performance.

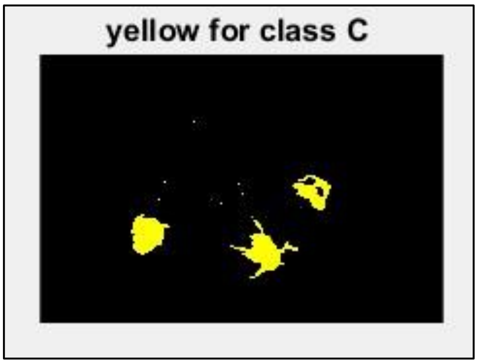
After the program has extracted only the brown color from the image, it selects based on the combination of the texture and all three classes. The results shown in Figure 11 show how successful this approach is in achieving classification.



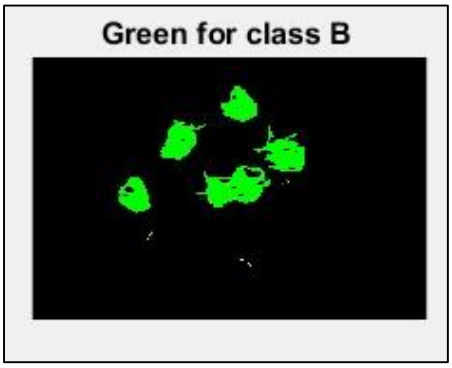
(a) All three types



(b) High Class A



(c) Middle Class



(d) Lower Class

Figure 10. Different classified results

5- Conclusion

The purpose of this study was to determine if it is possible to classify date fruits automatically without requiring time-consuming and complex physical measurements. Based on the review of previous studies, it is seen that basic machine learning methods have been used in the classification process of a date fruit in order to extract more features. In addition to this, a more effective stacking method created by combining these two methods was found to yield better results. In terms of classification studies, it has been found that high-performance results can be obtained not only with common image processing methods but also with new stacking methods that can be created by combining two or more of these methods. The proposed approach has been shown to achieve high classification rates with excellent classification accuracy. During the study, it was found that the best accuracy of 93% was achieved for classifying the high class, 94.5% for classifying the middle class, and 96.3 % for classifying the low class. In contrast, the performance efficiency can be as high as 100% if infrared or ultraviolet cameras are used to detect the skin of the dates. The results of this study can be used to create a date fruit classification program that can be presented to the users with the help of software on a mobile phone. Through the smartphone application, consumers will be able to be informed about the type of date fruit they are buying by categorizing anyone that is sold over the counter. As more features are extracted in classification studies, it is reported that success rates can rise in the classification of not only date fruits, but also other vegetables, fruits, legumes, or anything else for that matter.

6- Declarations

6-1- Author Contributions

Conceptualization, S.H. and M.I.; methodology, S.H.; software, S.K.; validation, M.I., S.A. and S.K.; formal analysis, S.H. and I.K.; investigation, S.A.; resources, S.K. and I.K.; data curation, S.K.; writing—original draft preparation, S.K.; writing—review and editing, S.K. and I.K.; visualization, S.H. and I.K.; supervision, S.A.; project administration, S.A.; funding acquisition, S.H., S.A. All authors have read and agreed to the published version of the manuscript.

6-2- Data Availability Statement

The data presented in this study are available on request from the corresponding author.

6-3- Funding & Acknowledgements

The authors received no financial support for the research, authorship, and/or publication of this article. The authors extend their appreciation to the Research and Innovation, Qassim University and Unaizah Colleges, Saudi Arabia and to the Research and Innovation Center, Prince Sultan University Riyadh Saudi Arabia for their support. The authors would like to acknowledge the support of Prince Sultan University for paying the Article Processing Charges (APC) of this publication.

6-4- Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this manuscript. In addition, the ethical issues, including plagiarism, informed consent, misconduct, data fabrication and/or falsification, double publication and/or submission, and redundancies have been completely observed by the authors.

7- References

[1] Altaheri, H., Alsulaiman, M., & Muhammad, G. (2019). Date Fruit Classification for Robotic Harvesting in a Natural Environment Using Deep Learning. IEEE Access, 7, 117115–117133. doi:10.1109/ACCESS.2019.2936536.

[2] Al-Rahbi, S., Manickavasagan, A., Al-Yahyai, R., Khriji, L., & Alahakoon, P. (2013). Detecting surface cracks on dates using color imaging technique. Food Science and Technology Research, 19(5), 795–804. doi:10.3136/fstr.19.795.

[3] Grossi, M., & Riccò, B. (2017). Electrical impedance spectroscopy (EIS) for biological analysis and food characterization: A review. Journal of Sensors and Sensor Systems, 6(2), 303–325. doi:10.5194/jsss-6-303-2017.

[4] Nasiri, A., Taheri-Garavand, A., & Zhang, Y.-D. (2019). Image-based deep learning automated sorting of date fruit. Postharvest Biology and Technology, 153, 133–141. doi:10.1016/j.postharvbio.2019.04.003.

[5] AKodagali, J., & Balaji, S. (2012). Computer Vision and Image Analysis based Techniques for Automatic Characterization of Fruits-A Review. International Journal of Computer Applications, 50(6). doi:10.5120/7773-0856.

[6] Ruiz-Altisent, M., Ruiz-Garcia, L., Moreda, G. P., Lu, R., Hernandez-Sanchez, N., Correa, E. C., Diezma, B., Nicolaï, B., & García-Ramos, J. (2010). Sensors for product characterization and quality of specialty crops-A review. Computers and Electronics in Agriculture, 74(2), 176–194. doi:10.1016/j.compag.2010.07.002.

[7] Marji, K.A. (2018). Fusion Approach for Dates Fruit Classification. International Journal of Computer Applications, 181(2), 17–20.

[8] Haidar, A., Dong, H., & Mavridis, N. (2012, October). Image-based date fruit classification. In 2012 IV International Congress on Ultra-Modern Telecommunications and Control Systems, 357-363, IEEE. doi:10.1109/ICUMT.2012.6459693.

[9] Alzu'bi, R., Anushya, A., Hamed, E., Al Sha'ar, E. A., & Vincy, B. A. (2018). Dates fruits classification using SVM. AIP Conference Proceedings, 1952(1), 020078, AIP Publishing LLC. doi:10.1063/1.5032040.

[10] F.A.O. (2019). Food and Agriculture Organization of the United Nations. Dates Production. Available online: https://www.fao.org/home/en (accessed on May 2022).

[11] Abdulrahman S. Alturki, Muhammed Islam, Mohammed F. Alsharekh, Mohammed S. Almanee, Anwar H. Ibrahim. (2020). Date Fruits Grading and Sorting Classification Algorithm Using Colors and Shape Features. International Journal of Engineering Research and Technology, 13(8), 1917-1920.

[12] Voulodimos, A., Doulamis, N., Doulamis, A., & Protopapadakis, E. (2018). Deep Learning for Computer Vision: A Brief Review. Computational Intelligence and Neuroscience, 1–13. doi:10.1155/2018/7068349.

[13] Behera, S. K., Rath, A. K., Mahapatra, A., & Sethy, P. K. (2020). Identification, classification & grading of fruits using machine learning & computer intelligence: a review. Journal of Ambient Intelligence and Humanized Computing, 1–11. doi:10.1007/s12652-020-01865-8.

[14] Liakos, K. G., Busato, P., Moshou, D., Pearson, S., & Bochtis, D. (2018). Machine learning in agriculture: A review. Sensors, 18(8), 2674. doi:10.3390/s18082674.

[15] Korohou, T., Okinda, C., Li, H., Cao, Y., Nyalala, I., Huo, L., Potcho, M., Li, X., & Ding, Q. (2020). Wheat Grain Yield Estimation Based on Image Morphological Properties and Wheat Biomass. Journal of Sensors, 2020, 11. doi:10.1155/2020/1571936.

[16] Ziafati Bagherzadeh, S. H., & Toosizadeh, S. (2022). Eye Tracking Algorithm Based on Multi Model Kalman Filter. HighTech and Innovation Journal, 3(1), 15–27. doi:10.28991/hij-2022-03-01-02.

[17] Tharwat, A. (2020). Classification assessment methods. Applied Computing and Informatics, 17(1), 168–192. doi:10.1016/j.aci.2018.08.003.