COMPUTER APPLICATION ON PATTERN RECOGNITION FOR PALM-DATE GRADING

Ewida, E.H[*], M. N. El Awady[**], M. A. Rashwan[***], M. Z. ElAttar[****]

ABSTRACT

The current procedure in the palm date factories is grading the palm date fruit manually using human graders. This method is inconsistent because each grader has his own technique and may vary from one person to another. Hence, this affects the quality and quantity of the fruit that can be extracted. In this research, a new model for automated grading system of palm date fruit is developed using the RGB color, area, and texture of the fruit and artificial fuzzy logic is used as a classifier. The color model was based on different color intensities. The purpose of this sorting system is to distinguish between three different classes of palm date fruit which are (class A) the best grade among (class B) and (class C). The grading system uses a computer and an ordinary scanner to analyze and interpret images correspondent to human eye and mind. The computer program is developed using MATLAB language: image processing toolbox for the image processing part, and fuzzy logic toolbox for the classification part. The calculation is based on the fruit size, the mean color intensity based on RGB color model, and the texture. The decision making process used fuzzy logic to train the data and make the classification for the palm date fruit. A total of 570 date samples were used to build and train the system. The program developed has been able to distinguish the three different classes of palm date fruit automatically; with 99.8% of overall efficiency. This paper provides a very good technique to standardize the palm date fruit grading system over a large area.

Keywords: Pattern recognition, Image processing, Palm date fruit, Classification, Grading, and Fuzzy Logic.

*** PhD student Agric. Eng. Dep. Faculty of agriculture, Ain shams univ.**
**** Emir. Prof. Agric. Eng. Dep. Faculty of agriculture, Ain shams univ.**
***** Prof. communication Dep. Faculty of engineering, Cairo univ.**
****** Teacher Agric. Eng. Dep. Faculty of agriculture, Ain shams univ.**

I. INTRODUCTION

Automated grading of agricultural products has been getting special interest recently, as the demand for higher quality food products has increased. Market grade of quality food products is determined based on their multiple features: flavor, texture and appearance. While flavor may be measured using chemical compounds to determine sweetness or acidity, texture properties, such as firmness and mouth feel are difficult to measure. In automated fruit grading, appearance (shape, color, size, and texture) is generally utilized to classify the fruit.

The date (*Phoenix Dactylifera, L*) is one of the oldest fruits in the Arab world and is still thought of as highly nutritious food, containing all the basic elements required for a balanced diet. According to the FAOSTAT (2010), Egypt is the largest date producing country in the world as it produces something around 1.3 million tons/year. FAOSTAT (2010) stated also that date fruits take the tenth place of production value in Egypt and the fourteenth place production quantity in Egypt.

The Arabic names for the various stages of development of dates are universally used. The *Kimiri* (الكمري) refers to the stage when the dates are young and green in color. The *Bisr* (البسر) is the stage at which dates begin to change color and reach maximum weight and size. The *Khalal* (الخلال) is the stage when the fruit begins to soften when partially-ripe. The *Rutab* (الرطب) defines the stage when the fruit turns all soft (totally ripe), with high moisture content and acquires darkness. Finally, *Tamr* (التمر) is the stage where fruit moisture content becomes less than 30%.

In this study, the size, color and texture will be the main considerations for grading. (Siwi سيوي) date fruits in *"Tamr* stage" will be the verity used in this work. The aim is to develop a computer application which can classify date fruit automatically through visual means.

II. REVIEW OF LITERATURES

"The decision-making processes of a human being are somewhat related to the recognition of patterns so, the goal of pattern recognition is to clarify these complicated mechanisms of decision-making processes and to automate these functions using computers." Fukunaga (1990)

Duda et al., (2001) "pattern recognition" is a field concerned with machine recognition of meaningful regularities in noisy or complex environments.

Pavlidis (1977) defined "the word pattern is derived from the same root as the word patron and, in its original use, means something which is set up as a perfect example to be imitated. Thus pattern recognition means the identification of the model which a given object was made after.

Gonzalez and Thomason (1978) suggested that "Pattern recognition" can be defined as the categorization of input data into identifiable classes via the extraction of significant features or attributes of the data from a background of irrelevant detail.

Bezdek (1981) stated that "pattern recognition is a search for structure in data"

Schalkoff (1992) stated that "Pattern recognition (PR) is the science that concerns the description or classification (recognition) of measurements."

“Image processing can be classified to three types of computerized processes in this continuum: low-, mid-, high-level processes. Low level processes involve primitive operations such as image preprocessing to reduce noise, contrast enhancement, and image sharpness. A low-level process is characterized by the fact that both its inputs and outputs are images. Mid-level processes on images involve tasks such as segmentation (partitioning an image into regions or objects), description of those objects to reduce them to a form suitable for computer processing and classification (recognition) of individual objects. A mid-level is characterized by the fact that its inputs generally are images, but its outputs are attributes extracted from those images. High-level processing involves “making sense” of an ensemble of recognized objects, as in image analysis.” Gonzalez and woods (2008).

“Pattern recognition: the act of taking in raw data and taking an action based on the category of the pattern.” Duda et al., (2001).

“Date sorting techniques using color vision have been used in many studies. These studies work on different classification factors and verities. Aljanobi (1993) developed some techniques to classify dates. His system classifies date fruit based on color and shape. "

III. MATERIALS AND METHODS

3.1: Materials:

3.1.1: Hardware:

- PC – Desktop computer.
    - Manufacturer: Gateway – E series.
    - Processor: Intel Pentium IV, 2.8 GHz.
    - RAM: 2GB.
- Scanner: Manufacturer: BenQ.

3.1.2: Computer Applications:

- Photoshop CS5.
- AutoCAD 2013.
- MATLAB 2011a.
    - Image processing toolbox.
    - Fuzzy logic toolbox.
- O.S.: Windows 7.
- SPSS. (statistical-analysis software)

3.2: Methods:

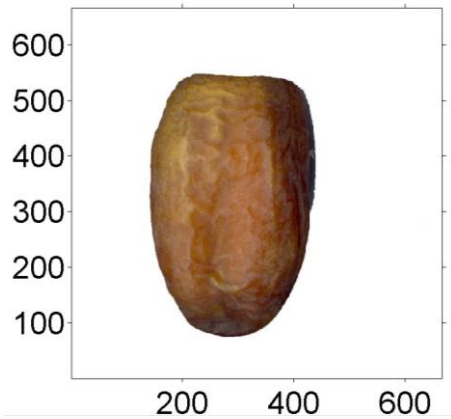
3.2.1: Feature Extraction

3.2.1.1: Fruit Size:

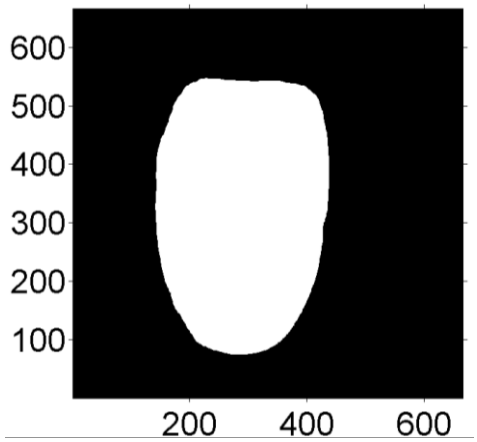
The fruit size is a quality attribute used by farmers. The bigger size fruit is considered of better quality. The size is estimated by calculating the area covered by the fruit image. To compute the area, first, the fruit image is binarized to separate the fruit image from its background using the "Canny edge detecting algorithm" (Canny, 1986). The number of pixels that cover the fruit image were counted and considered as an estimate of size. Fruits have been categorized as (big, medium and small) for fuzzy logic input function.

3.2.1.2: Color model:

The distribution color intensity has been used in the image as an estimate of quality. (AlOhali (2011), AlHomedey (2011) observed that the image of the low quality date is darker than the bright date. The color intensity distribution is obtained from the RGB colored image that is obtained from the relationship: I=(R+G+B)/3 (Gonzalez and Woods, 2008) where ***I***, ***R***, ***G***, ***B*** are: The color ***I***ntensity, ***R***ed, ***G***reen, and ***B***lue values of each pixel. After that, the average color intensity of the fruit was calculated. Fruits have been categorized according to color as (*good*, *fair* and *poor*) for fuzzy logic input function..



A) Original image (Numbers = Pixels).

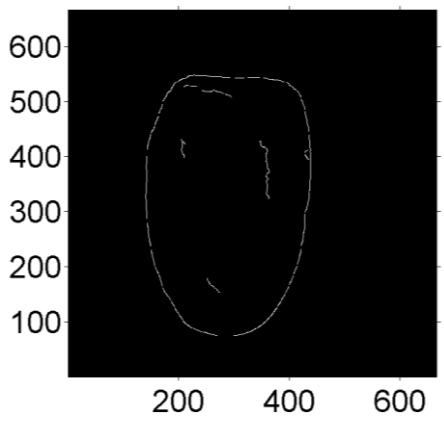


B) Binarized image (Numbers = Pixels).

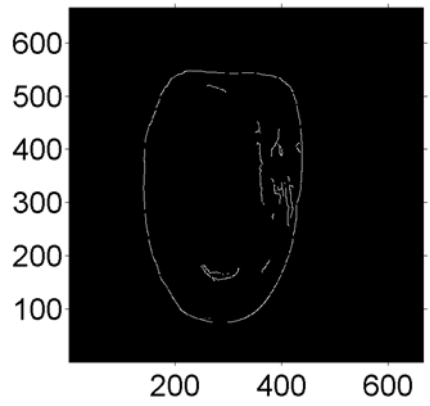
Fig 2: Size calculating technique.

3.2.1.3: Texture:

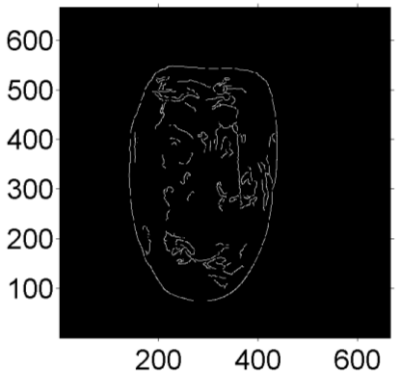
(AlOhali, 2011, AlHomedey, 2011) observed that the better quality date has bigger number of wrinkles. The number of edges inside the image was considered as the number of wrinkles. To determine the texture, the image is binarized and edges are extracted using the "Canny edge detecting algorithm" fig. 3: notice that the background is turned black by the MATLAB program. The number of pixels that cover the fruit wrinkles were counted for the RGB pixels separately as shown in figures 3 A, B, C and the average is considered as an estimate of texture. Fruits has been categorized as (good, fair and poor) for fuzzy logic input function.



A) Wrinkles of red pixels (Numbers = Pixels).



B) Wrinkles of green pixels (Numbers = Pixels).



C) Wrinkles of blue pixels (Numbers = Pixels).

Fig 3: texture calculating technique.

3.2.2 Classification:

Fuzzy Logic algorithm is chosen for classifying the date fruit into (A, B, and C), categories. This research used an already used fuzzy logic algorithm of (AlHomedey, 2011). This technique is selected because it represents a good approach for interpreting the decision making process of human to the computer. Currently, the date fruit grading is done based on the experience of the human grader. This method needs to be replaced with a new approach that is able to standardize the grading process. The process consists of 3 main steps: defining the input and output in "Membership Function Editor", set the fuzzy rules in "Rule Editor" and obtaining the output for each rule in "Rule and Surface Viewer".

The system has three input functions (size, color, and texture) and one output function (grading) fig. 4. These membership functions were built on the Gaussian distribution curve, since it gives the best results

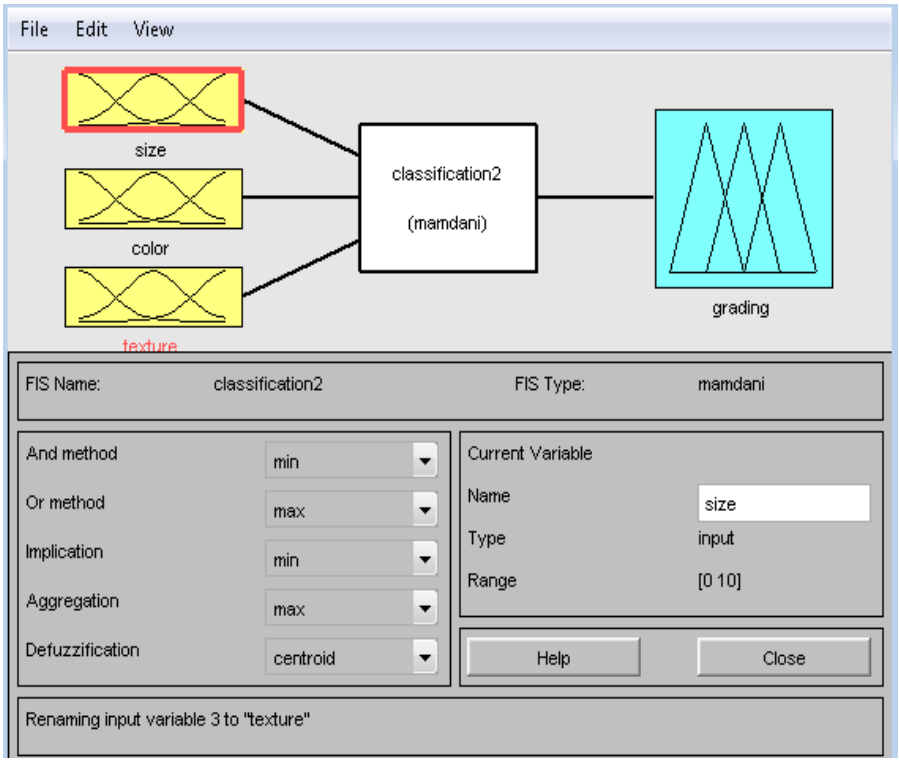


Fig. 4: The fuzzy logic inference system.(original scale)

compared to other shapes. A total of 27 rule statements were created in order to classify the date fruit categories. Example of the rules are illustrated in table 1.

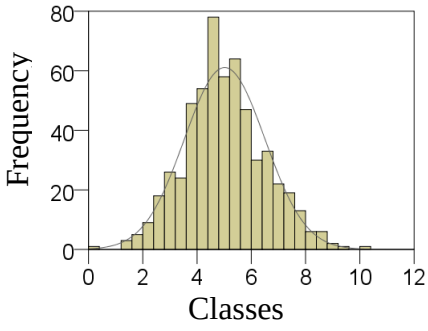
Table 1: Sample of fuzzy set rules for classifications.

If (size is medium) And (color is good) And (texture is good) Then (grading is grade_A).	If (size is big) And (color is fair) And (texture is poor) Then (grading is grade_B).
If (size is small) And (color is fair) And (texture is poor) Then (grading is grade_C).	If (size is medium) And (color is fair) And (texture is good) Then (grading is grade_B).

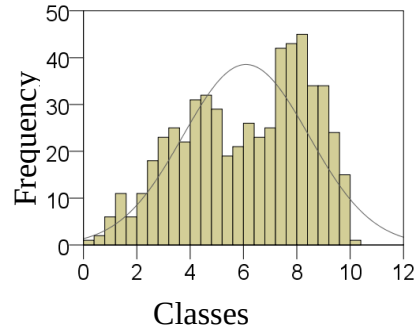
3.3: Date samples:

A total of 570 date samples were used to build and train the system. These samples cover all grading difference as shown in figures 5 A, B, C. These samples consisted of 2 groups. The first group was collected from the local market from several sources comprizing of 305 date sample, and 265 date sample for the second group. The second group was dried by the researcher using the direct method (in air, under sun). So, it was darker than the first group. The difference between these two groups is obvious as double class, in figure 5B. (Siwi سيوي) date fruits in *"Tamr* stage" is popular and will be the variety used in this work. The values of each input function were converted into ratios. Ratios from 0- 10 were obtained by using the following function “f(x)”.

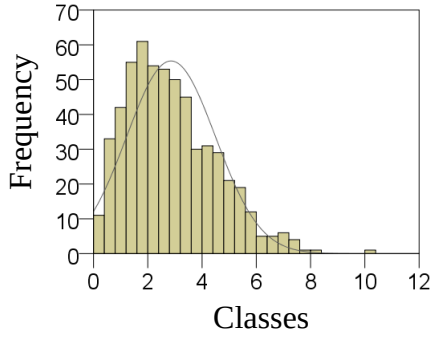
$$f(x) = \frac{current_{value} - minimum_{value}}{maximum_{value} - minimum_{value}} \times 10$$



A) the frequency chart of size



B) the frequency chart of color



C) the frequency chart of texture

Fig 5: The frequency charts of input functions

IV. RESULTS AND DISCUSSION

The output of the system shows that the pattern recognition system has succeeded in extracting features of all 570 samples but only one exception. The classification error was less than 0.2%. The confidence of the computer application is 99.8%. This application was designed to use any of classification methods with sample classifier of (AlHomedey, 2011) fuzzy logic.

Figure 6 shows the output of the classifier. The output values ranged from 1.5-8.5. For this sample, grade *A* was 105 date fruits with output ratio of more than 4.9. Grade *C* was 107 date fruit with output ratio of less than 2.7. The in-between samples were taken as grade *B*.

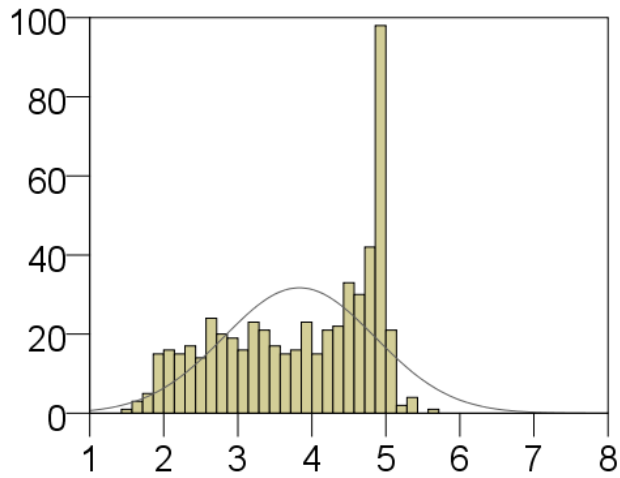
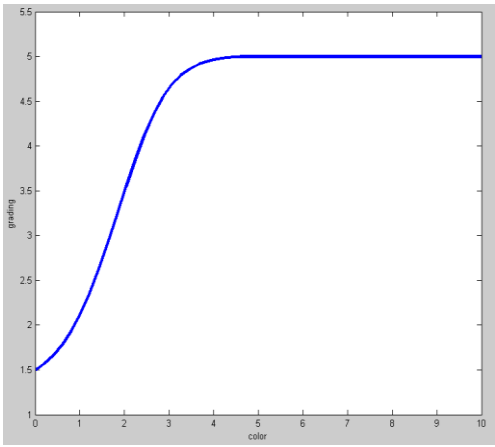


Fig 6: The frequency chart of output function.

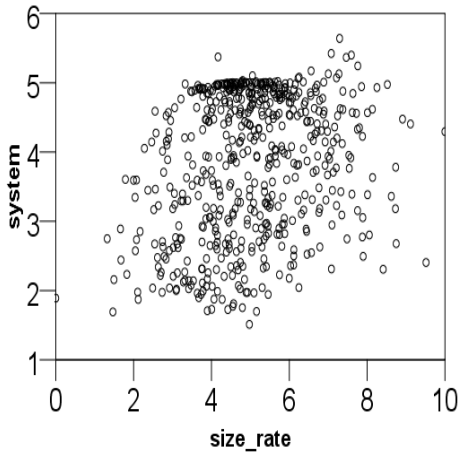
Fuzzy logic is a very complex algorithm. It cannot be understood by crisp logic or mathematical formulation. "Chart 7A" is the theoretical relationship between any input function and the output function of the classifier system. The output values normally range from 1.5 to 8.5. However this range is reduced in the relationship between one input function and the output function.

Charts 7 B, C, D represents the relationship between size, color, and texture input function and the output function respectively. From these

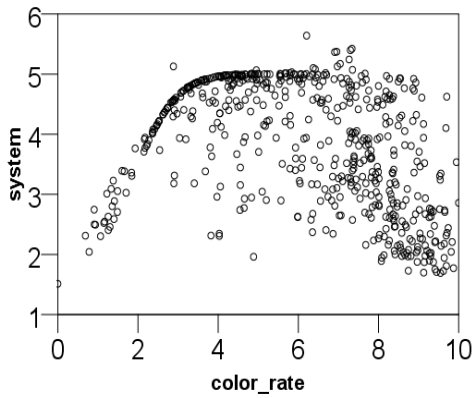
charts it could be noted that the texture is the most important input function in this classifier method after that the color input function and at last the size.



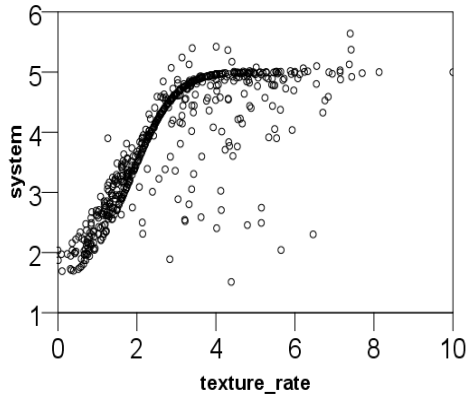
A) Theoretical relationship between any input function and the output function.



B) The relationship between size input function and the output function.



C) The relationship between color input function and the output function.



D) The relationship between texture input function and the output function.

Fig 7: The relationship between every input function and the output function.

V. CONCLUSION

The date fruit grading system can be improved by adding more training with images from various additional sources. The pattern recognition computer application have a strong confidence value 99.8%. The study has proven that, by using fuzzy logic algorithm, the accuracy of date fruit grading is quite high and to improve it the relationships between input and output function have to be reordered to make the most important input function is color the size and the last is texture.

VI. REFERENCES

AlHomedey. A (2011). Fuzzy Image Analysis and Classification of Agricultural Produce: A Case Study of Dates (*Phoenix Dactylifera*). Master thesis, Southern Illinois University

ALJanobi, A. (1993). Machine Vision Inspection of Date Fruits. Ph.D. Thesis, Oklahoma State University.

AlOhali, Y., (2011). Computer vision based date fruit grading system: Design and implementation. Journal of King Saud University – Computer and Information Sciences 23, pp29–36.

Bezdek, J. C. (1981). Pattern Recognition with Fuzzy Objective Function Algorithms, Plenum, NY.

Canny, J. F. (1986). A computational Approach to Edge Detection. IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 8, pp. 679-698.

Duda, R., P. Hart, D. Stork, (2001). Pattern Classification 2nd edition, Wiley Interscience.

FAOSTAT (2010) – Statistics Database. Retrieved 1/9/2012, from http://faostat.fao.org.

Fukunaga, K. (1991). Statistical Pattern Recognition, second edition (first edition, 1972), Academic Press, San Diego, CA.

Gonzalez, R. C. and M. G. Thomason. (1978). Syntactic Pattern Recognition: An Introduction, Addison Wesley, Reading.

Gonzalez, R. C. and R.E. Woods. (2008). Digital image Processing 3rd edition. Pearson Prentice Hall.

Pavlidis, T. (1980). Structural Pattern Recognition, Springer-Verlag, NY.

Schalkoff, R. (1992). Pattern Recognition: Statistical, Structural and Neural Approaches, John Wiley and Sons, NY.

الملخص العربي

تطبيق حاسوبي للتعرف على الأنماط في تدريج التمور

م. اسماعيل عويضة*، د. نبيل العوضي، د. محسن رشوان***، د. محمود العطار*****

يعتبر التدريج اليدوي هو الإجراء الحالي في مصانع التمور،و لكن هذه الطريقة غير متناسقة لأن كل إنسان لديه أسلوبه الخاص ويمكن أن تختلف من شخص إلى آخر. وبالتالي ، فإن هذا يؤثر على نوعية وكمية الفاكهة التي يمكن استخراجها. في هذا البحث تم تطوير نموذج جديد لنظام التدريج الآلي للتمورباستخدام اللون،والمساحة، وتعريجات السطح للثمرة.وتم استخدم المنطق الضبابي الاصطناعي باعتباره المصنف.والغرض من هذا النظام هو التدريج للتمييز بين ثلاث فئات مختلفة من نخيل التمر الفاكهة التي هي فئة (A) الدرجة الافضل يليها الدرجة (B) و أخيرا الدرجة (C). نظام التدريج يستخدم جهاز كمبيوتر و ماسح ضوئي عادي لتحليل وتفسير الصور كمكافئ للعين البشرية والعقل. تم تطوير برنامج الكمبيوتر باستخدام لغة : MATLAB. ويستند حساب فئة الثمرة على مساحة الثمرة، و متوسط كثافة اللون على أساس نموذج الألوانRGB ، و التعريجات على سطح الثمرة. تستخدم عملية صنع القرار المنطق الضبابي لتدريب البيانات وإجراء تصنيف التمور . تم استخدام ما مجموعه ٥٧٠ ثمرة لبناء و تدريب النظام. وقد تم تطوير هذا البرنامج قادرا على التمييز بين ثلاث فئات مختلفة من التمر تلقائيا. اثبت النظام كفاءة بنسبة ٩٩,٨ ٪. تقدم هذه الورقة نظاما جديدا لتدريج ثمار التمر على نطاق واسع.

*** طالب دراسات عليا – دكتوراة – قسم هندسة زراعية – كلية الزراعة – جامعة عين شمس.**
**** أستاذ الهندسة الزراعية غير المتفرغ – كلية الزراعة – جامعة عين شمس.**
***** أستاذ الاتصالات– كلية الهندسة – جامعة القاهرة.**
****** مدرس الهندسة الزراعية – كلية الزراعة – جامعة عين شمس.**