Banana Ripeness Detection and Servings Recommendation System using Artificial Intelligence Techniques

Izzat Hafiz Hadfi and Zeratul Izzah Mohd Yusoh
Computational Intelligence and Technologies Lab, Centre for Advanced Computing Technology, Universiti Teknikal Malaysia Melaka
zeratul@utem.edu.my

***Abstract*—The banana plantation industry is the second most widely cultivated fruit in Malaysia. The current method to detect the ripeness of banana is using chemicals in order to obtain the characteristic of the fruit. This method will harm the fruit and also affects its quality. There are also methods which are non-destructive to the products which use manpower to identify the banana. This method is time-consuming and prone to human error. Thus, the objective of this research is to determine the correct method in assisting users in selecting bananas. This project is proposing a combination of two Artificial Intelligence (AI) techniques, namely Image Processing technique and Fuzzy Logic rules in a knowledge-based system for the solution. Several samples of un-ripe, ripe and over-ripe banana are taken to identify their red, green and blue (RGB) value. The values are then analyzed to create the membership functions for the fuzzy logic. Then a set of fuzzy knowledge-based rules are implemented for the system to determine the ripeness level of the banana. From the result of the ripeness, the system will give recommendations on the fruit including suggested meal preparation and the best before the date to consume the banana. The proposed system contributes to both farmers and customers. As for the farmer, they can pick their best product to be sold to the market. While for the customers, they can choose efficiently their desired banana ripeness by using this system.**

***Index Terms*—Banana Ripeness Detection; Fuzzy Rules; Image Processing; Knowledge-based System.**

I. INTRODUCTION

The banana plantation industry is a large industry in Malaysia where it is the second most widely cultivated fruit in Malaysia [1]. Quality grading is crucial for the farmer to choose the best banana to be sold to the market. The current method to detect the ripeness of a banana is using chemicals in order to gain the parameters needed to tell whether the banana is ripe or not. This method will both harm the fruit and also endanger the health of the consumers themselves [2].

Furthermore, most of the banana plantation industry uses man power for the quality grading phase. With man power, it is not consistent and time-consuming. This is because the person that can tell the ripeness of the banana needs to be an expert. The knowledge of the expert can also be passed, but this can cause inconsistencies.

Limited capability of human can also result in a limited color grading issue. This is because the human's eye is lack of capability to compare the indistinct color range of the banana. As we know, the banana fruit does not only come with the color yellow or green. Some of the fruit have a little amount of green color, some have a little yellow color, and there are also both of the colors in one fruit.

Lastly, it is a big disadvantage for the customer who has a lack of knowledge of the banana to buy one. This is because the customer is not able to pick the right fruit based on their desires. There are some of the fruit condition which needs to wait long until it can be eaten, and some fruit can be eaten raw by cooking it into any delicacies.

By developing the system which is Banana Ripeness Detection and Servings Recommendation System using Artificial Intelligence techniques, it can help to over for the customer who has lack of knowledge these problems faced by the banana plantation industry. By using a computerized method, the quality grading is not only consistent but also a time-saving method. This system uses image processing techniques in order to obtain the banana characteristics. With the implementation of the Fuzzy Logic method, it can help the system to determine whether the banana is ripe or not.

With this system, farmers can ensure the best quality of their product to be sold on the market. It also saves the man power needed for the operation of the company. Besides farmers, the customer also gains some advantages. Customers can also use the system to pick the desired banana to be bought.

The remaining paper is organized as follows. Section II discusses related work. Some backgrounds of AI techniques are presented in Section III. The proposed system is described in Section IV. Section V discusses the results and evaluation. The concluding remark is presented in Section VI.

II. RELATED WORK

This section discusses existing systems related to this research. It focuses on existing solutions for recognizing banana ripeness using the artificial neural network, determination of size and ripeness of banana and other fruits using image processing, and also other related existing literature.

Mustafa et. al [3] proposed a system to help the farmer to choose a banana to be sold on the market and also to help the customers to choose the right banana for them to buy. Key factors to determine the best quality of a banana is to identify its size and its ripeness. Thus, they apply image processing technique in order to full fill these requirements. The size of the banana is determined based on the pixel count of the image, while the ripeness is determined based on a formula in which the formula uses the total number of pixels of the image.

Authors in [4] also apply similar techniques in assessing banana's maturity in determining its ripeness. For their research, banana fruits were collected and selected by its maturity stages from three categories which are under-mature, mature and over-mature. The sample images of the bananas were taken, and the Red Green Blue (RGB) color intensity was calculated using statistical moments obtained from the histogram. The graph is then calculated by its mean, variance, smoothness texture, skewness, and kurtosis. Next, the size of the banana is measured. The fruit was measured by its area and perimeter by measuring the number of pixels in the fruit region and measuring the number of pixels in the boundary region of the fruit respectively. The pixel's measurement is then converted into centimeter in order to ease the farmer. The way to convert this measurement is by putting a cross-section of a banana fruit next to a banana fruit on a graph sheet. Analysis of variance (ANOVA) with Duncan's multiple range test (DMRT) was used to compare the significance of datasets of color mean intensity value, area, perimeter, major axis length and minor axis length between each banana group. The datasets were processed using two classifiers algorithm which is a box and whisker plot technique. The experimental results are claimed to be better than the existing system.

Saad et. al [5] used the artificial neural network to recognize the ripeness of a banana. They used supervised learning rules for the learning rules. Supervised learning is where the data is trained from a set of examples. The examples will be the input, and then the output will be compared with a target which is the correct output. The learning rules are then used to adjust the weights and biases of the network in order to move output as close as to the target. The project starts by gathering 3 sets of the unripe, ripe, and overripe of the banana image. The images are then being resized and apply the process of extraction of RGB component. Each of the components contains pixel values ranging from 0 to 255, and then it is divided into 3 groups of intensity. The values of the pixels are then exported into Microsoft Excels and graph is plotted based on the pixel values. After the network has been trained, the optimization process is applied in order to choose the Smallest Mean Error value. In the process, it uses to try and error in order to find the most suitable number of hidden nodes. It is found that 7 hidden layers will be used to get the most optimized results. The result of the system is as the expected results. The system can successfully recognize between unripe banana, ripe banana, and overripe banana.

There are also other types of fruit that used a similar technique in assessing its ripeness. Authors in [6] developed a system to determine a ripe watermelon by extracting its color feature and process the data using supervised learning in Neural Network. While authors in [7] using image acquisition and image processing techniques in detecting the ripeness of tomatoes.

To summarize, image processing techniques are proven to be a good approach in determining the ripeness level of the fruit. There are few problems in this existing system. Firstly, is that the system can only predict the category of the fruit which is unripe, ripe and overripe. The system cannot tell the percentage of the ripeness. This is because it only uses the artificial neural network to predict the ripeness. In addition, bananas have phases for its ripeness. The banana can be still unripe but will be ripe in a few days more. Without using fuzzy logic, these values cannot be obtained. The system is also lack of recommendation for the users. These gaps will be addressed by the proposed system presented in this paper.

III. Related Background And AI Techniques

This section will explain in general the domain background as well as AI techniques that are used in this research, which are, image processing, fuzzy logic rules, and knowledge-based system.

A. Problem Domain: Banana

Bananas are one of the most widely consumed fruits in the world [8]. In Malaysia, the banana is the second cultivated plant after durian [1]. As we can see in Malaysia, there are many products that are made of banana. The fruit itself can be made into many delicacies. Bananas have their ripening stages; mostly people only think that a green banana means that it is till unripe and yellow means that it is ripe and ready to be eaten. There are a few aspects that need to be considered before confirming a banana is ripe or not. The factors are such as size, color, and the brown spots formation on the fruit. Sometimes people consider the perfect yellow color of the banana is ripe banana and other people consider the brown spots on a banana are rotten spots. This is wrong, the rotten spots that formed on a banana indicate that sugar content has risen in the ripening process.

B. Image Processing Techniques

Image processing is a process of processing images using mathematical operations using any form of signal processing for which the input is a digital image or a video [9]. Image processing usually focuses on two tasks which are an improvement of an informative image for human interpretation and processing of an image data for storage, transmission, and representation. The input of an image into the image processing platform needs to be a digital image. A digital image is a representation of a two-dimensional image as a finite set of digital values, called picture elements or pixels. Each pixel of an image has its own values. The values are often grey levels which range from 0 to 255 which indicate from black to white. All the values will be stored in matrices composed of M rows and N columns. One image processing platform is Matlab. In Matlab, there are many tools that are already in the software, such as image enhancement, image segmentation, line detection and etc. These tools use mathematical operations to modify the pixel values of the image to give the desired information from the image for the user. With all the tools provided in the Matlab toolbox, the RGB values of the bananas can be obtained easily. Furthermore, this technique is a non-destructive to the fruits

C. Fuzzy Logic

The meaning of fuzzy is not sharp, unclear, imprecise or approximate. Therefore, fuzzy logic is a form of many-valued logic in which the truth values of variables may be any real number between 0 and 1 [10]. Fuzzy logic needs a membership function in order to operate. The membership function is a graph that is a plotted line that defines how each point in the input space is mapped to a membership value or degree of membership which are between 0 and 1. There are many methods of assigning the membership functions such as intuition, inference and rank order. Besides these methods, some other artificial intelligence techniques can be applied such as neural network and genetic algorithms. The first process of fuzzy logic is fuzzification. This is where the membership functions are obtained more than one. The last process is called defuzzification. This process involves converting the fuzzy sets into crisp sets. There are many methods used for defuzzification such as max-membership principle, centroid method, weighted average and etc. Each of these methods has its own advantage and disadvantage based on each situation.

Since detecting the banana ripeness involved imprecise information, image, and knowledge, this technique will be used after the image processing phase in determining the ripeness level.

D. Knowledge-based System

A knowledge-based system is composed essentially of two sub-systems that known as the knowledge base and the inference engine. The knowledge base represents the fact and the rules, while the inference engine is an automated reasoning system. Inference engine will evaluate the current state of the knowledge base, applies relevant rules and deduce the recommendation or results to the users. Many techniques were incorporated into various types of inference engines as the knowledge-based system evolved including Fuzzy Logic Rules.

IV. THE PROPOSED SYSTEM

This system consists of three different phases which are image processing, Fuzzy Inference System, and Knowledge-Based System.

The first phase is the image processing phase. Three bunches of banana from the same species which is the Cavendish banana is first collected. Each of the three bunches has different ripeness category which is under-ripe, ripe, and over-ripe as determined by the expert. From these samples, only 14 bananas from each bunch were used for the data training. While the extra fruit from each bunch is used for the testing. The images of the bananas are taken and go through image processing in order to obtain the RGB value of the banana. Figure 1 shows the samples of banana from a different category. The image will then be converted into three different binary forms which consist of three different channels. The channels are red, green and blue which will be the RGB value of the image. Before counting the total number of pixels, the threshold value of each channel must be set. The threshold values need to be set according to the image histogram. We need to find the position of between two peaks of the image histogram. Then the number of pixels in each image is counted. In this process, the pixels are counted by each row with the selected threshold. Figure 2 shows the sample histogram of Red, Green and Blue channels, and Table 1 depicts the threshold values of each channel from each ripe category.

Figure 1: Sample images of different level of ripeness

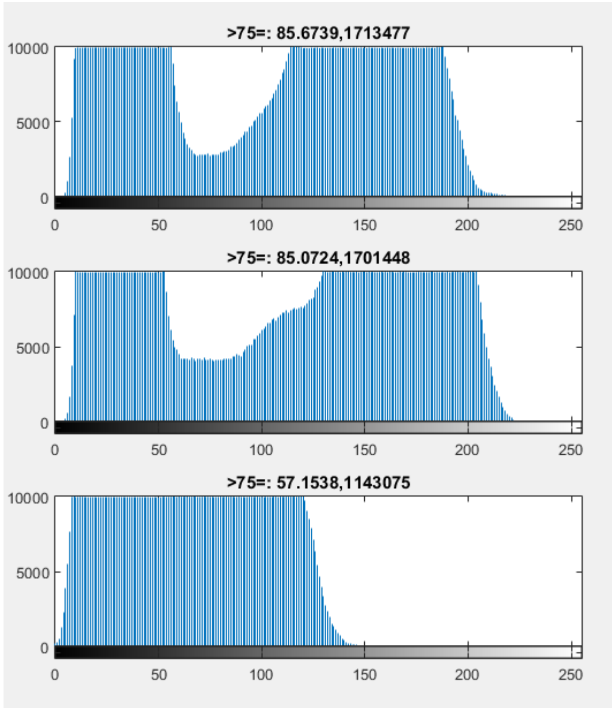

Figure 2: Image histogram of Red, Green and Blue Channel

Table 1
RGB Values of the Bananas

CATEGORY	COLORS					
	Red		Green		Blue	
	Min	Max	Min	Max	Min	Max
Under-ripe	63.0	96.2	62.2	96.8	44.2	82.1
Ripe	61.9	92.4	60.2	92.0	51.6	84.2
Over-ripe	57.9	88.8	55.1	86.0	33.0	80.8
Minimum	57.8	88.8	55.1	85.9	32.9	80.8
Maximum	63.0	96.2	62.2	96.8	51.5	84.2

The next phase is to implement a fuzzy inference system (FIS) to the system. Firstly, the FIS needs to have its membership function. This membership function is developed from the sample data that were provided. The sample data were obtained by analyzing the RGB value of each three different ripeness category of the banana. The data were then sorted. The membership function is developed based on the data that were collected from the banana sample. In this system, the inputs for the membership function are Red, Green, and Blue. Each membership function consists of three different linguistic values which are low, medium and high. Next for the output is Ripeness, it also consists of three linguistic values which are under-ripe, ripe and over-ripe. Then the process of defuzzification will be carried out by using the centroid method. Table 2 shows the membership function for each channel in FIS.

Table 2
Membership Function for the FIS

CATEGORY	RED	GREEN	BLUE
Low	0-60	0-60	0-45
Medium	50-95	50-95	35-82
High	90-100	90-100	75-100

The KBS phase of this system consists of three different sections of KBS which are ripeness, servings and shelf life. The ripeness section is to predict the ripeness of the banana from the percentage value that was obtained from the FIS. For the servings, the section is to give a recommendation to the user on how to serve the banana based on the ripeness of the banana and from the preferences of the user. The last one is the shelf life; this section will give a recommendation to the user on the shelf life of the banana. The knowledge for the KBS was obtained from various sources including human experts, articles and Internet resources. This knowledge is converted to rules by using Mockler chart and decision tables. The system will give a recommendation of the servings and shelf life based on users' input and matched it with the rules. Figure 3 shows an example of Mockler chart for banana shelf life recommendation.

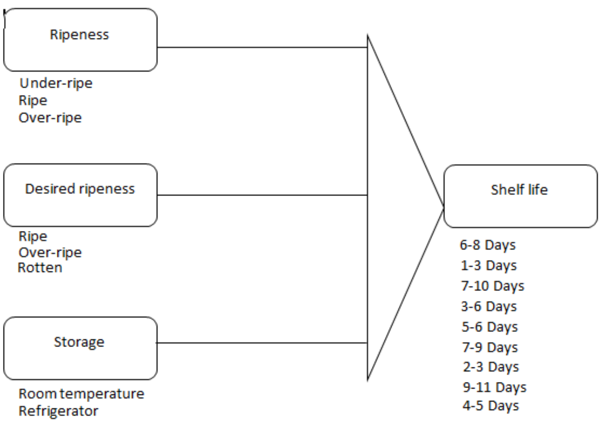

Figure 3: Mockler chart for banana shelf life

All these three phases are executed in sequence in order to get the output. Each phase has a different input. For the image processing phase, the input is only the image selected by the user. The image can be in any type of files such as Joint Photographic Experts Group (JPEG) and Portable Network Graphics (PNG). While for the FIS phase, it accepts three different inputs. These three inputs are the total number of pixels that are processed during the image processing phase. The three inputs are the total number of red pixels, the total number of green pixels, and the total number of blue pixels. For the Knowledge-Based System, it acquires user input such as ripeness percentage, servings for the banana, texture of the banana, the age of the banana, desired ripeness of the banana and banana storage. The system will then display the level of ripeness of the banana, servings recommendation of the banana and shelf life of the banana.

V. Results and Discussion

The evaluation plan consists of two parts: 1) evaluating the results of the ripeness level and shelf life, and 2) evaluating feedback from users. For the first part, two experts in the field are recruited. These experts have experienced banana sellers. The experts will determine the level of ripeness of a bunch of bananas, and the result is compared with the system's results. The experts will also give feedback on the system's recommendation of the shelf-life of a banana.

The second part is conducted with five users, in which all testers are the university's staff and undergraduate students. The second part focuses on a recommendation from the system, particularly on the servings and banana's shelf life. In addition, the usability of the system is also being evaluated. An analysis form was prepared to analyse all the reviews submitted by the testers.

From the test results of the first part of the evaluation, almost all the test sample achieved its expected output. The only module that is sometimes faulty is the fuzzy inference system module. The FIS sometimes failed to predict the correct ripeness percentage of the banana. This is due to inconsistent of the image acquisition process. The image acquisition needs to be done in a controlled environment room with no excessive light exposure. However, this is not the case for this research.

The second part of the evaluation is collecting users' feedback on (a) interface of the system, (b) whether the questions/options asked by the systems are clear and in a sensible sequence, (c) the system's performance as well as (d) the users' appreciation of the system.

Most testers agree that the interface of the system is easy to use. However, one user commented on the lack of explanation on some terms using in the system, i.e., banana's life. The feedback for (b) is generally good. While feedbacks on (c) and (d) indicates that users are satisfied with the results of the ripeness level and shelf life. On the other hand, users want more servings options included in the knowledge base, such that the recommendation for servings are varies and more interesting.

VI. Conclusion

As a conclusion, the banana plantation industry needs to focus on their quality management in order to survive the competition among the other agriculture industry sectors. The current quality grading method uses man power which is time-consuming and inaccurate. There are also methods using chemicals to show the ripeness of the banana. This method is accurate, but it loses the quality of the banana and can endanger the health of the consumer. Thus the objective of this research is to determine the correct method in assisting users in selecting bananas. This project is proposing a combination of two Artificial Intelligence techniques, namely Image Processing technique and Fuzzy Logic rules in a knowledge-based system. The type of banana which is used in this project is the Cavendish banana. The target users of the system are the farmer and also the customers. Based on the evaluation, the system able to determine the ripeness of the banana correctly as well as give good recommendations for the users. For future work, the rules in FIS can be improved by optimizing it in order to get better results.

Acknowledgment

This work is funded by the Universiti Teknikal Malaysia Melaka through the PJP Research Grant Scheme, PJP/2016/SILTERRA/FTMK/C-ACT/S01509.

REFERENCES

[1] Banana Farming Information Guide. (n.d) in Asia Farming. Retrieved May 23, 2017, from http://www.asiafarming.com/banana-farming-information-guide/

[2] Ibrahim, N. U. A., Abd Aziz, S., & Nawi, N. M. (2017). Novel Impedance Measurement Technique for Soluble Solid Content Determination of Banana. *Pertanika Journal Science & Technology*, 25 (2), pp. 519 – 526

[3] Mustafa, N. B. A., Fuad, N. A., Ahmed, S. K., Abidin, A. A. Z., Ali, Z., Yit, W. B., & Sharrif, Z. A. M. (2008, August). Image processing of an agriculture produce: Determination of size and ripeness of a banana. In International Symposium on Information Technology, 2008. ITSim 2008. (Vol. 1, pp. 1-7). IEEE.

[4] Prabha, D. S., & Kumar, J. S. (2015). Assessment of banana fruit maturity by image processing technique. *Journal of food science and technology*, 52(3), pp. 1316-1327.

[5] Saad, H., Ismail, A. P., Othman, N., Jusoh, M. H., fadzlina Naim, N., & Ahmad, N. A. (2009, November). Recognizing the ripeness of bananas using artificial neural network based on histogram approach. In 2009 IEEE International Conference on Signal and Image Processing Applications (ICSIPA). IEEE.

[6] Shah Rizam, M. S. B., Farah Yasmin, A. R., Ahmad Ihsan, M. Y., & Shazana, K. (2009). Non-destructive watermelon ripeness determination using image processing and artificial neural network (ANN). *International Journal of Electrical and Computer Engineering*, 4(6).

[7] Monavar, H. M., Alimardani, R., & Omid, M. (2011). Detection of red ripe tomatoes on stem using Image Processing Techniques. *Journal of American Science*, 7(7).

[8] Schuh, S. (2014). Fitness & Clean Eating Motivation. Retrieved May 23, 2017, from https://www.pinterest.com/pin/374361787748592789/

[9] Tinku Acharya and Ajoy K. Ray (2006). Image Processing - Principles and Applications. Wiley InterScience.

[10] Negnevitsky, Michael. Artificial intelligence: a guide to intelligent systems. Pearson Education, 2005.