

Morphological Based Seedling Comparison Of Natural And Hybrid Lemon

Florante B. Vidad III, Edwin R. Arboleda, Jesusimo L. Dioses Jr. and Rhowel M. Dellosa

Abstract: Experimental lemon seedling classifier was developed for comparing bunches of natural lemon seeds and hybrid lemon seeds by using an image processing technique. The experimental system involved the use of a 16 MP camera phone to capture the images of lemon seeds in bunches, which were subsequently converted into digital data using a laptop. The analysis was performed with MATLAB R2012A, an image processing software, which determined the size parameters of each lemon seeds. In this research, the size of the other 60 natural lemon seeds had obtained by the image processing method (IPM) and compare with the size of 60 hybrid lemon seeds using the Fuzzy Logic Algorithm (FLA) and k-Nearest Neighbor (KNN) algorithm. An automatic algorithm based on artificial intelligence fuzzy logic to determine three size indicators of lemon seeds, namely area, perimeter and equivalent diameter, respectively, was developed in this article. 20 samples of test data were used for data mining of the KNN classifier. The study aimed to help farmers and fruit-bearing tree planters to classify seeds of fruits like lemon easily through an image processing technique. It further aims to test the effectivity of the image processing technique in determining the morphological differences of seeds of some plants like lemon, its subspecies, and hybrids. This paper shows that FLA was the most effective way of classifying the variety of lemon seeds than KNN.

Index Terms: Alemaw, Fuzzy Logic, Image Processing, k-Nearest Neighbor, Lemon Seeds, MatLab, Neuro-fuzzy

1. INTRODUCTION

Triacetic acid is naturally concentrated in citrus fruit, the most popular leading acid citrus fruit is Lemon. The first Lemon tree was planted in America in 1493. It contains a lot of vitamins like Vitamin C and B6 and Niacin [1]. It has a lot of purpose in human life, hence, it is an important medicinal plant that has anticancer activities and antibacterial that can have extracted from its different parts [2]. Its largest producer was from Italy & United States, it is normally grown in the cool and low-temperature area. It can help in sickness and malnutrition, its peel, seeds, and leaves are also used for other medical and experimental purposes [3]. It is high in antioxidant and detoxifier, sickness like Asthma, Colds, Constipation and Sore throat are some that lemon can function with [4]. Some types of lemon Figure 1 show the two samples used on this paper where the large one is the alemow and yellow ones are the natural lemons. In this paper, the morphological process was used for data acquisition. Value by the pixel is based on the experimental assessing of each image with some lemons. The morphological process is done by choosing the shape and size of the area, hence, it should be carefully done because it is sensitive to exact shapes in the input image. By selecting the size and shape of the area, one can make a hypothesis about the morphological process that is sensitive to exact shapes in the input image. After extracting the structures of the images, it will undergo a fuzzy logic algorithm. Better understand by "True or False" or the degrees of truth, either 1 or 0. Also, Fuzzy logic supports the transaction of problems with imprecise, inaccurate and vague information showing the accuracy of fuzzy methods in digital image processing[5]–[9].

In this study, the purpose is to create a system that will aid a person to compare the natural lemon and hybrid lemon through seeds. With the combination of image processing and fuzzy logic, classifying the varieties of lemon through seeds will become automated and more efficient since the structure of the seeds is almost alike. The study aimed to help farmers and fruit-bearing tree planters to classify seeds of fruits like lemon easily through an image processing technique. Its further aims to test the effectivity of the image processing technique in determining the morphological differences of seeds of some plants like lemon, its subspecies, and hybrids.



Fig 1. Hybrid (Alemow) and Natural Lemon

2 PROCEDURE FOR PAPER SUBMISSION

2.1 Image Capturing

The first step is image acquisition which refers to capturing images in a standard lighted room. In this study, an android phone was used to capture the image of the fruit which done through a high definition camera almost up to 16 megapixels to have a standardized image for better output. Generally, some factors that affect image capturing such as the quantity and quality of lighting, angle, and orientation, distances, physical characteristics of the object and even the temperature [10]. The captured images were processed using software that will extract color features then converted into a grayscale

- Florante B. Vidad III is from Department of Computer and Electronics Engineering, Cavite State University, Indang, Cavite, Philippines
- Edwin R. Arboleda is from Department of Computer and Electronics Engineering, Cavite State University, Indang, Cavite, Philippines
- Jesusimo L. Dioses Jr. is from Isabela State University
- Rhowel M. Dellosa is from Asia Technological School of Science and Arts

image. Figure 2a shows seeds of the lemon bought at Kadiwa, Dasmariñas, Cavite, while figure 2b shows the seeds that came from SM Pala-Pala, Dasmariñas, Cavite. On the other hand, figure 3 presents the hybrid lemon seeds that harvested from the backyard at Vineyard, Robinson, Dasmariñas, Cavite

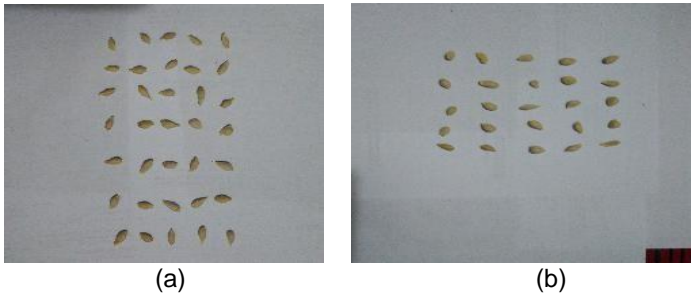


Fig 2. Captured Natural Lemon Seeds

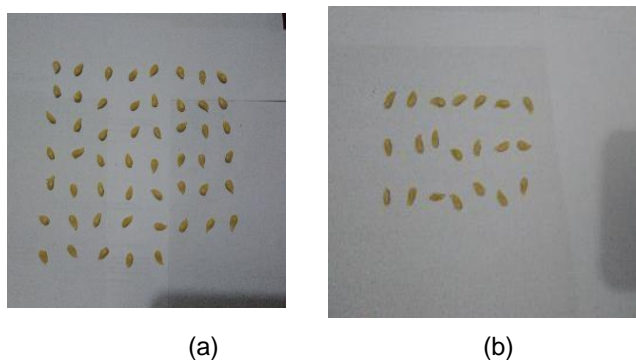


Fig 3. Captured Hybrid Lemon Seed

2.2 Preprocessing the Image

Preprocessing or initial processing of an image is a significant stage of an object recognition. It was done to make the image more suitable than the original image. The first step in the initial processing was to crop the image to retain the central part of the images. The image will also undergo filtering to correct area color properties. This image will then be translated into binary images having either 0 or 1 pixel's intensity. The forefront carries black pixels with 1 intensity while the backdrop carries white pixels with 0 intensity. The second step is to remove imperfections in the image and provide information on the form and structure of the image morphological operations such as erosion and dilation used [11].

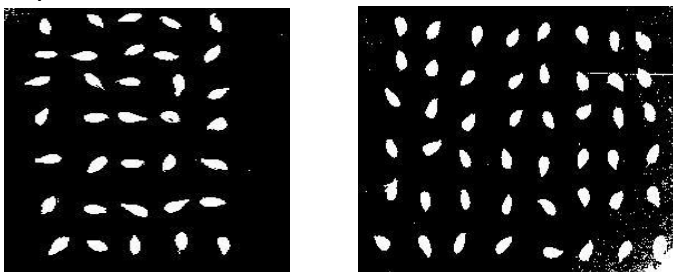


Fig 4. Processed Image

2.3 Image Segmentation

In a world of Digital Image Processing and its analysis, the most commonly used and remained a fundamental step in many image analysis applications in Image Segmentation. In this process, the system separates the real-world object or the foreground from the background. The target region depends on the needed output of the specific applications and most likely adapts to the operator's experience and the subjective reasoning. It is very useful in locating objects and boundaries in an image. and processing. Image segmentation aids in image analysis from image processing to obtain further image understanding [12]. Figure 5 shows the segmentation of an image with the used of vislabel function.

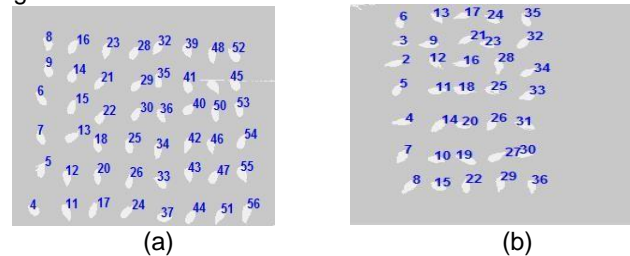


Fig 5. Segmentation of Seeds

2.4 Feature Extraction

Feature extraction is a process of dimensionally reducing the initial raw data set into a set of image features[13][14]. In regards to the isolation, identification, and labeling of regions of common property, image features play an important role. The parameters used were area, equivalent diameter, and perimeter. Table 1 shows the acquired data after being processed.

3 FUZZY LOGIC ALGORITHM

3.1 Fuzzy Logic Controller

This study aims to generate a Fuzzy Logic Algorithm (FLA) for a comparison of natural and hybrid lemon seeds using three parameters. As soon as the model is finished, it will undergo several tests and experimentation to test the accuracy and effectiveness of the system. Using the MATLAB software, the algorithm is processed through the Fuzzy Logic system [15]. There are two important methods under the Fuzzy Inference System (FIS) namely, Takagi-Sugeno Fuzzy Model (TS Method) and Mamdani Fuzzy Inference System. Among the two, the latter suits the study. In this paper, the Mamdani fuzzy logic model is developed to compare the seeds, specified as a pilot of this study and used in the software of the application. For this, the proposed fuzzy model system is presented by modifying the Mamdani fuzzy logic model with another method.

3.2 Proposed Fuzzy Medical Diagnosis System

The system shows another kind of approach to develop the whole framework in classifying the lemon seed. The framework is consisting of the fuzzification, fuzzy inference system, and defuzzification value as shown in Figure 6 [16].

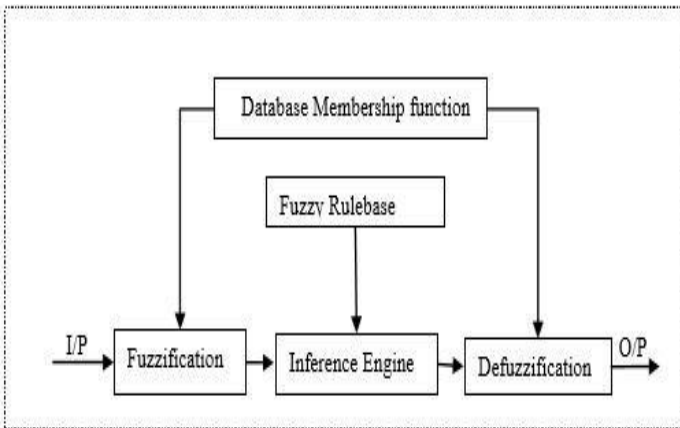


Fig 6. A framework of Fuzzy Expert System

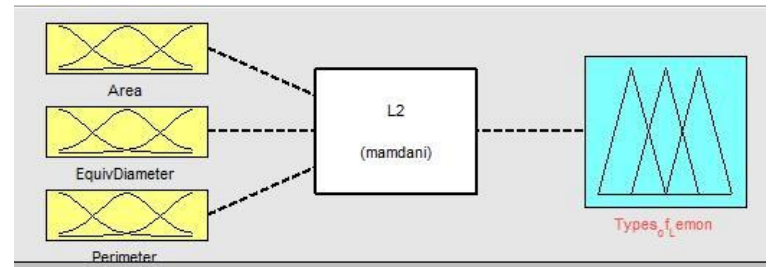


Fig 7. The structure of the FES

In the structure shown above, there are two to three membership functions as input for each parameter and four functions as output. The range of the output membership function is set by the intuition method done by the user.

3.3 Identifying Parameters

It is essential to take into consideration of all parameters that can be measured using MatLab that plays a

Table 1. Parameters and its Corresponding Functions

1. If (Area is Low) and (EquivDiameter is Low) and (Perimeter is Low) then (Natural_or_Hybrid is Natural) (1)
2. If (Area is Low) and (EquivDiameter is High) and (Perimeter is Low) then (Natural_or_Hybrid is Natural) (1)
3. If (Area is Low) and (EquivDiameter is High) and (Perimeter is High) then (Natural_or_Hybrid is Natural) (1)
4. If (Area is High) and (EquivDiameter is Low) and (Perimeter is Low) then (Natural_or_Hybrid is Hybrid) (1)
5. If (Area is High) and (EquivDiameter is High) and (Perimeter is Low) then (Natural_or_Hybrid is Hybrid) (1)
6. If (Area is High) and (EquivDiameter is High) and (Perimeter is High) then (Natural_or_Hybrid is Hybrid) (1)

a significant role in determining the difference between natural and hybrid lemon seeds. The input and output parameters and values are shown in Table 1 [17].

3.4 Choice of Membership Functions

Membership functions are graphical forms that help us visualize the fuzzy sets [18]. It contains all the information in the fuzzy set regardless of the set being discrete or continuous. This technique solves practical problems through experience [19]. It also classifies the fuzzy sets. These are normal fuzzy set, subnormal fuzzy set, convex fuzzy set and non-convex fuzzy set. The first mentioned set is used in this paper. There are different methods of assigning membership values to elements [20]. In this study, it only requires intuition or by using some local procedures or algorithms. Intuition, according to Tech-wonders, doesn't need an introduction. It is based on humans being capable of developing membership functions with their intelligence and understanding [21].

3.5. Fuzzification

Fuzzification is the conversion of a clear input value to a fuzzy value, performed by using the information in the knowledge base [22]. This study uses the linguistic variables high (H) and low (L) for the fuzzification process of the factors given. The system is modeled with lower and higher values of each parameter but these values are not an averaged value. The FIS editor used for classification was shown in Figure 7.

3.6. Fuzzy Rule Base A total of 6 rules were formed shown in figure 8.

Fig 8. Fuzzy rules

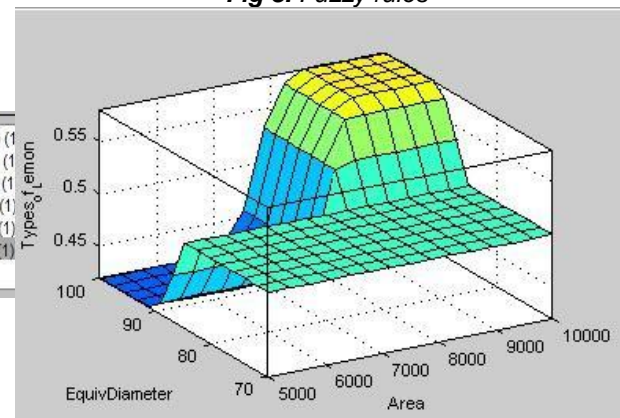


Fig 9. Surface Output of Fuzzy rules

3.7. Structure of the Fuzzy Factors The used factors' memberships are shown in Fig. 10 – Fig. 12. This value of the area, equivalent diameter and perimeter were measured using the Matlab Software. The range for the area was from 5000 to 10000, the equivalent diameter was from, 70 to 100 and the perimeter was from 400 to 800. All of the range was intuited base on the gathered data.

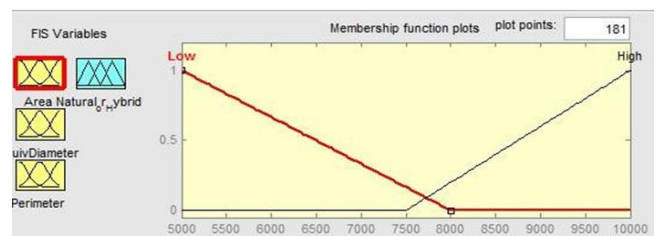


Fig. 10. The membership function of Area

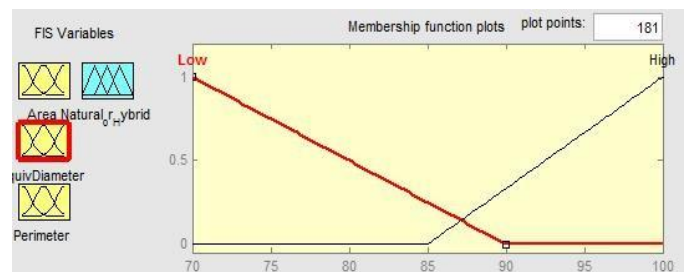


Fig. 11. The membership function of Area

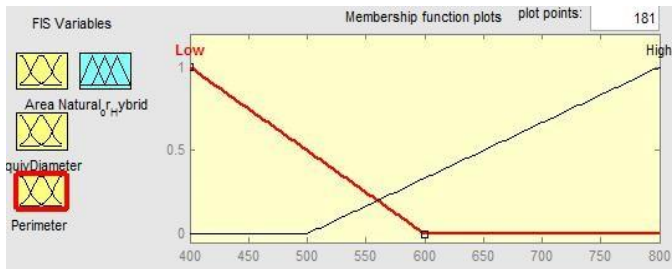


Fig. 12. The membership function of Area

4 K NEAREST NEIGHBOR

In classifying two points, the nearest neighbor decision rule takes place to classify the nearest set from the previously classified points [23]. Oftentimes, it considers more than one neighbor into account. This technique is called the K Nearest Neighbor (KNN) [24]. Though it is known as the simplest form of the techniques, KNN rule has been continuously studied and improved upon by researchers. What makes this algorithm interesting to researchers is the computational simplicity it gives and the good results it brings [25]. This rule is commonly used in text categorization [26], pattern recognition [27],

Area	EquivDuameter	Perimeter	Area	EquivDuameter	Perimeter
9187	108.1538	579.6123	5788	85.8459	455.2864
8630	104.8239	593.0265	5302	82.1627	412.6173
8508	104.0804	492.1148	6211	88.9274	408.7595
3425	66.0367	1956	5569	84.2061	466.3574
8391	103.3622	481.1442	8236	102.4031	561.127
8712	105.3208	659.595	8082	101.4412	697.8377
8222	102.3161	652.1808	7492	97.6684	435.8305
6515	91.0777	557.2691	5865	86.415	424.6589
8108	101.6043	595.5118	6513	91.0638	546.8843
7521	97.8572	588.3402	7003	94.4272	482.8427
7028	94.5956	526.2986	6982	94.2855	403.5879
6434	90.5098	497.4285	6457	90.6714	423.5462
8868	106.2595	652.524	7361	96.8107	587.0265
9097	107.6228	566.7838	5944	86.995	403.3452
8257	102.5336	524.0559	5285	82.0309	388.7595
9444	109.6562	730.3646	7197	95.7262	497.0854
7270	96.2105	465.2864	6145	88.4537	541.7128
8222	102.3161	652.1808	5476	83.5001	441.3452
6515	91.0777	557.2691	6699	92.3549	544.8843
8108	101.6043	595.5118	5984	87.2873	491.2864

object recognition [28], and ranking models [29] applications.

Fig 13. Gathered Training Data

Using MatLab Software, the KNN algorithm was conducted. 60 samples for natural lemon seeds and 60 samples of hybrid lemon seeds for test data were used in this paper. While 20 samples each of it was gathered for training data. Table 2a shows the training data for natural lemon seeds while table 2b shows the training data for hybrid lemon seeds.

5 RESULTS AND ANALYSIS

5.1 Evaluation by Using Mamdani-Type FIS

The proposed FIS for the comparison of natural existing lemon seeds and hybrid lemon seeds consists of three inputs (Area, Equivalent Diameter, and Perimeter) as shown in Figure 2. In this final step, the input is an aggregated fuzzy set which produces a single numerical coefficient that returns the midpoint of the cumulated area that can be found below the curvature. Defuzzification is a process of turning the output of

a fuzzy inference into a clearer and cleaner result. In this experiment, the Mamdani type fuzzy inference systems process the input using triangle membership function (TMF) and rules as previously described in figure 12. In this stage, we can determine the truth degrees (a) of the rules with the aid of the minimum and then by taking the maximum between working rules. For example, for Area =8640, EquivDiameter= 100, and Perimeter= 656.3, rule 6 will be fired and we will obtain:

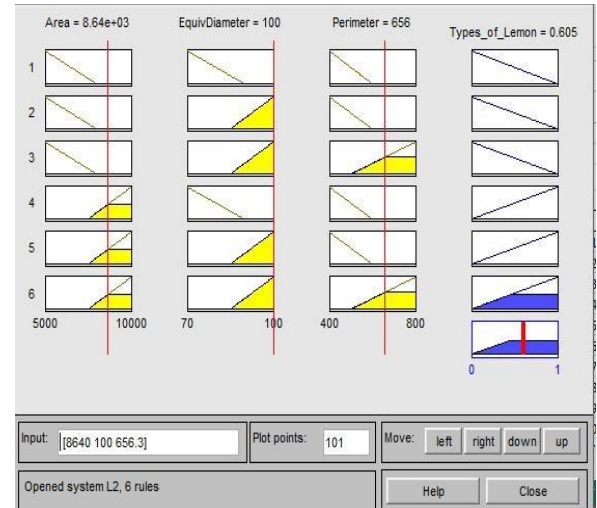


Fig. 14. The membership function of Area

From Fig. 14, the value of Types of Lemon = 0.605 which was the centroid. This means that the type of that seed was hybrid based on the formulated rule 6. Based on the gathered output, 13 out of 60 were classified as a hybrid seed on natural samples while 22 out of 60 were determined as natural seed based on the hybrid sample. The overall accuracy of the FLA was 70.333%.

5.2 Evaluation by Using K Nearest Neighbor (KNN) Classifier

Classification is the process of grouping samples that possesses same characteristics[30][31]. K-nearest neighbor algorithm or KNN is the most basic algorithms in machine learning. This method classifies the object by determining the majority vote of its neighbors, then it will then assign the class according to the most common amongst its k - nearest neighbors where a positive integer k is ideally small. If the results showed that k = 1, then it will be assigned to the single nearest neighbor class. After tabulating all the gathered data, the samples undergo a KNN classifier. This will classify the seeds if it was a natural lemon seed or a hybrid lemon seed. A range from 1 to 10 of k was used for its neighboring. Figure 14 displays the result of the KNN classifier at K= 10. Based on the given result, 13 out of 60 samples of natural lemon had been classified as hybrid seeds. For this result, a total of 21.67% of error was done. While on the other hand, 25 out of 60 samples of hybrid lemon seeds had been classified as a natural lemon seed. This will give a 41.67% error of classifying the types of seed. The accuracy for overall data of KNN classifier was 68.333%.

```

>> class=knnclassify(A,B,G,10);
>> disp('Result');
Result
>> disp(class)
'Natural' 'Natural' 'Natural'
'Natural' 'Natural' 'Natural'
'Hybrid' 'Natural' 'Natural'
'Natural' 'Natural' 'Natural'
'Hybrid' 'Natural' 'Natural'
'Natural' 'Natural' 'Natural'
'Hybrid' 'Natural' 'Natural'
'Natural' 'Natural' 'Natural'
'Natural' 'Natural' 'Natural'
'Natural' 'Natural' 'Natural'
'Hybrid' 'Natural' 'Natural'
'Natural' 'Natural' 'Natural'
'Hybrid' 'Natural' 'Natural'
'Natural' 'Natural' 'Natural'
'Hybrid' 'Natural' 'Natural'
'Natural' 'Natural' 'Natural'
'Hybrid' 'Natural' 'Natural'
'Natural' 'Natural' 'Natural'
'Hybrid' 'Natural' 'Natural'
'Natural' 'Natural' 'Natural'
'Hybrid' 'Natural' 'Natural'
'Natural' 'Natural' 'Natural'
'Hybrid' 'Natural' 'Natural'

```

Fig. 15. k-Nearest Neighbor at k=10

6 CONCLUSION

An attempt was made in this study to apply both image processing, fuzzy logic and neuro-fuzzy for comparing the varieties of lemon, natural and hybrid (Alemow), that was found in the Philippines. The study was divided into two parts: a morphological technique for data gathering and a fuzzy logic algorithm and a knearest neighbor algorithm based on the information gathered by the first part. All of the processes was implemented in the Matlab software. From the two classifiers, the fuzzy logic algorithm was a more efficient and more effective classifier for varieties of lemon seeds than the KNN algorithm due to its proposed rules. The 41.67% error of classifying the Alemaw (lemon/pomelo hybrid) was due to its hybrid structure.

REFERENCES

- [1] M. J. Dhanavade, C. B. Jalkute, J. S. Ghosh, and K. D. Sonawane, "Study Antimicrobial Activity of Lemon (Citrus lemon L.) Peel Extract," *Br. J. Pharmacol. Toxicol.*, no. August, p. 1, 2011.
- [2] M. H. M. Khan and N. M. Khan, "Analysing factors affecting hand biometrics during image capture," *Procedia Comput. Sci.*, vol. 32, pp. 521–528, 2014.
- [3] M. Khojastehnazhand, V. Mohammadi, and S. Minaei, "Maturity detection and volume estimation of apricot using image processing technique," *Sci. Hortic. (Amsterdam)*, vol. 251, no. January, pp. 247–251, 2019.
- [4] X. Luo, X. Li, Z. Wang, and J. Liang, "Discriminant autoencoder for feature extraction in fault diagnosis," *Chemom. Intell. Lab. Syst.*, vol. 192, no. November 2018, p. 103814, 2019.
- [5] M. J. L. Jacoba and E. R. Arboleda, "Identification of rain cloud amount, rain intensity and possible chance of flooding using image processing and fuzzy logic," *ARPN J. Eng. Appl. Sci.*, vol. 14, no. 14, pp. 2563–2568, 2019.

- [6] K. E. N. T. Rabe, E. R. Arboleda, and R. M. Dellosa, "Fuzzy Logic Based Vehicular Congestion Estimation Monitoring System Using Image Processing And KNN Classifier," *Int. J. Sci. Technol. Res.*, vol. 8, no. 08, pp. 1377–1380, 2019.
- [7] A. E. Andaya, E. R. Arboleda, A. A. Andilab, and R. M. Dellosa, "Meat Marbling Scoring Using Image Processing with Fuzzy Logic Based Classifier," *Int. J. Sci. Technol. Res.*, vol. 8, no. 08, pp. 1442–1445, 2019.
- [8] C. L. Macalalad, E. R. Arboleda, A. A. Andilab, and R. M. Dellosa, "Morphological Based Grain Comparison of Three Rice Grain Variety," *Int. Journal Sci. Technol. Res.*, vol. 8, no. 08, pp. 1446–1450, 2019.
- [9] R. E. T. Bae, E. R. Arboleda, A. Andilab, and R. M. Dellosa, "Implementation Of Template Matching , Fuzzy Logic And K Nearest Neighbor Classifier On Philippine Banknote Recognition System," *Int. J. Sci. Technol. Res.*, vol. 8, no. 08, pp. 1451–1453, 2019.
- [10] M. Mohanapriya, L. Ramaswamy, and R. Rajendran, "Health and Medicinal Properties of Lemon (Citrus Limonum)," *Int. J. Ayurvedic Herb. Med.*, vol. 1, no. 3, pp. 1095–1100, 2013.
- [11] Z. Peng, S. Qu, and Q. Li, "Interactive image segmentation using geodesic appearance overlap graph cut," *Signal Process. Image Commun.*, vol. 78, no. 1, pp. 159–170, 2019.
- [12] K. L. Penniston, S. Y. Nakada, R. P. Holmes, and D. G. Assimos, "Quantitative assessment of citric acid in lemon juice, lime juice, and commercially-available fruit juice products," *J. Endourol.*, vol. 22, no. 3, pp. 567–570, 2008.
- [13] E. R. Arboleda, A. C. Fajardo, and R. P. Medina, "Classification of Coffee Bean Species Using Image Processing , Artificial Neural Network and K Nearest Neighbors," *2018 IEEE Int. Conf. Innov. Res. Dev.*, no. May, pp. 1–5, 2018.
- [14] E. R. Arboleda, "Comparing Performances of Data Mining Algorithms for Classification of Green Coffee Beans," *Int. J. Eng. Adv. Technol.*, vol. 8, no. 5, pp. 1563–1567, 2019.
- [15] I. H. Altas and A. M. Sharaf, "A Generalized Direct Approach for Designing Fuzzy Logic Controllers in Matlab/Simulink GUI Environment," *Int. J. Inf. Technol. Intell. Comput. Int. J. IT&IC*, vol. 1, no. 4, 2007.
- [16] S. SushilSikchi, S. Sikchi, and A. M. S., "Fuzzy Expert Systems (FES) for Medical Diagnosis," *Int. J. Comput. Appl.*, vol. 63, no. 11, pp. 7–16, 2013.
- [17] A. M. Andrew, *An Introduction to Digital Image Processing with Matlab Notes for SCM2511 Image Processing 1 Semester 1*. 2004.
- [18] eMath, "eMathTeacher: Mamdani's Fuzzy Inference Method." .
- [19] Tutorialspoint, "Fuzzy Logic - Membership Function." .
- [20] J. Verkuilen, "Assigning membership in a fuzzy set analysis," *Sociol. Methods Res.*, vol. 33, no. 4, pp. 462–496, 2005.
- [21] Tech-wonders, "Intuition and Inference Methods of Assigning Membership Values to Elements." .
- [22] E. Kayacan and M. A. Khanesar, "Fundamentals of Type-1 Fuzzy Logic Theory," in *Fuzzy Neural Networks for Real Time Control Applications*, 2016, pp. 25–35.
- [23] J. M. Keller, M. R. Gray, and J. Givens, "A Fuzzy K-

- Nearest Neighbor Algorithm," IEEE Trans. Syst. Man. Cybern., vol. 15, no. 4, pp. 580–585, 1985.
- [24] T. . Cover and P. . Hart, "Nearest Neighbor Pattern Classification," IEEE Trans. Inf. Theory, vol. 13, no. 1, pp. 118–119, 1967.
- [25] B. V. Dasarathy, "Nosing Around the Neighborhood: A New System Structure and Classification Rule for Recognition in Partially Exposed Environments," IEEE Trans. Pattern Anal. Mach. Intell., vol. PAMI-2, no. 1, pp. 67–71, 1980.
- [26] G. Toker and Ö. Kirmemiş, "Text Categorization Using k-Nearest Neighbor Classification," Middle East Tech. Univ., 2013.
- [27] B. R. Kowalski and C. F. Bender, "The K-Nearest Neighbor Classification Rule (Pattern Recognition) Applied to Nuclear Magnetic Resonance Spectral Interpretation," Anal. Chem., vol. 44, no. 8, pp. 1405–1411, 1972.
- [28] F. Bajramovic, F. Mattern, N. Butko, and J. Denzler, "A comparison of nearest neighbor search algorithms for generic object recognition," 2006.
- [29] X. Geng, A. Arnold, T. Y. Liu, H. Li, T. Qin, and H. Y. Shum, "Query dependent ranking using K-nearest neighbor," in ACM SIGIR 2008 - 31st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Proceedings, 2008, pp. 115–122.
- [30] E. R. Arboleda, "Discrimination of civet coffee using near infrared spectroscopy and artificial neural network," Int. J. Adv. Comput. Res., vol. 8, no. 39, pp. 324–334, 2018.
- [31] E. R. Arboleda, A. C. Fajardo, and R. P. Medina, "An Image Processing Technique for Coffee Black Beans Identification," in 2018 IEEE International Conference on Innovative Research and Development (ICIRD), 2018, no. May, pp. 1–5.