Linker: (The Journal of Emerging Research in Agriculture, Fisheries, and Forestry)
Volume 4, Issue 2
ISSN (Print): 3082-3625 ISSN (Online): 2815-2018
DOI: https://doi.org/10.65141/jeraff.v4i2.n4

Eggplant Variety Classification Using Image Processing

Hanna Christine Alegre[1], **Edwin R. Arboleda**[2], **Drew Gwen M. Arboleda**[3], **James Patrick Dones**[4], **John Mark Panganiban**[5]
Department of Computer, Electronics and Electrical Engineering, Cavite State University, Indang, 4109, Cavite, Philippines [1,2,4,5]
Cavite State University Laboratory Science High School, Cavite State University, Indang, 4109, Cavite, Philippines [3]
edwin.r.arboleda@cvsu.edu.ph

RESEARCH ARTICLE INFORMATION	ABSTRACT

Received: May 26, 2023
Reviewed: November 20, 2024
Accepted: December 02, 2024
Published: December 31, 2024

Image processing is a powerful technique used in various fields, including agriculture, to enhance images and extract valuable information. This study focuses on utilizing image processing to determine crop parameters such as area, perimeter, and volume, thereby facilitating the assessment of crop quality. The integration of fuzzy logic concepts and the k-nearest neighbors (KNN) classifier further enhances image processing outcomes. The objective of the study was to identify eggplant varieties based on seed analysis using image processing, fuzzy logic, and KNN algorithms. The methodology involved enhancing image quality and extracting essential features from eggplant seeds, including area, perimeter, equivalent diameter, and roundness, which are used for subsequent fuzzy logic and KNN classification. The results demonstrate a successful application of image enhancement and feature extraction on eggplant seed images. The proposed approach accurately identified eggplant varieties using fuzzy logic and KNN algorithms. In conclusion, the extracted features are crucial for the classification process. Future research should expand the dataset to include a wider range of eggplant varieties, explore alternative machine-learning techniques, and consider variations in lighting conditions and seed sizes. These recommendations aim to improve the accuracy and robustness of the classification model. This research holds promise for broader applications in agriculture, crop classification, plant disease detection, and quality control in food production. The integration of image processing, fuzzy logic, and KNN presents valuable opportunities for advancements in various industries.

Keywords: *classification, eggplant variety, fuzzy logic, fuzzy rules, image processing, MATLAB*

Introduction

Eggplant, scientifically known as Solanum melongena, is a widely cultivated vegetable in Asia, known by various names such as brinjal and aubergine. It belongs to the Solanaceae family, which includes other well-known plants like potatoes, tomatoes, and peppers. In the Philippines, eggplant production contributes to over 30% of the total volume of vegetable production (Andaya et al., 2019). The physical characteristics of eggplants typically feature a dark purple outer skin, although certain varieties may exhibit striped, pale purple, or even white flesh. The interior of the eggplant is generally white or cream-colored and contains seeds. While the long and purple Philippine eggplant variety is the most common, there exists a diverse range of eggplant types (Arboleda et al., 2021). This study focused on analyzing and comparing the Philippine and Indian eggplant varieties, both of which have commercially available seeds.

Philippine eggplants are typically characterized by their long and thin shape, showcasing dark green or purple hues with occasional light purple flushes (Austria et al., 2019). In contrast, Indian eggplants are smaller in size, ranging from thumbnail to hen's egg-sized, and exhibit a dark reddish-purple color.

Historically, eggplant was introduced to Europe by the Arabs and later brought to Africa by the Persians. The Spaniards further spread its cultivation to the New World, including the Philippines. Visual inspection alone proves to be a time-consuming method for identifying different eggplant varieties due to their long growth period. To address this challenge, the current study aimed to classify the two eggplant varieties based on their seeds, utilizing advanced techniques such as image processing, fuzzy logic, and the KNN classifier. Fuzzy logic, being a branch of artificial intelligence, offers an effective means of handling vagueness and unstructured decision-making in agricultural applications. The KNN classifier, on the other hand, operates by comparing similarity measures, storing available cases, and classifying new cases accordingly (Hautea et al., 2016; Lajom et al., 2024).

In recent years, computer vision and image processing have garnered increasing attention from researchers due to their versatile applications across various domains. With this in mind, this paper conducted a comprehensive study focusing on the comparison and identification of two vegetable seeds, employing fuzzy logic and image processing techniques. The images captured were subjected to extensive processing, and three morphological features were utilized as rules for the fuzzy inference system. Encouragingly, the results demonstrated that the fuzzy inference system performed as anticipated, even with only two true values available for each characteristic (Loresco & Daios, 2018). By harnessing the potential of image processing, fuzzy logic, and the

KNN classifier, this paper aimed to successfully classify different types of eggplants based on their distinctive seed characteristics.

The scope of this study was to classify eggplant varieties based on seed analysis using image processing, fuzzy logic, and the KNN algorithm. The research aimed to accurately identify Philippine and Indian eggplant varieties and investigate potential applications in agriculture, crop classification, plant disease detection, and quality control in food production. However, it is important to acknowledge certain limitations in the study. While the focus was on two specific eggplant varieties, future research should include a wider range of varieties to develop a more comprehensive classification model. Moreover, the limitations of the camera's resolution and image quality may have an impact on the accuracy of image processing and feature extraction. The study assumed consistent lighting conditions and seed sizes, but variations in these factors may pose challenges and affect classification accuracy. Future research can explore alternative machine learning techniques and expand the dataset to enhance the performance and generalization capabilities of the classification model. Additionally, using more recent versions or alternative image-processing software tools may offer additional functionalities and improve the analysis. It is crucial for future research to provide detailed information on the ethical considerations and protocols implemented. Furthermore, the application of image processing, fuzzy logic, and KNN classification in other areas such as plant disease detection and quality control can be explored as separate research avenues.

Methods

Three systems were utilized in this research: image processing, fuzzy logic system, and k-nearest neighbors (KNN). Fuzzy logic allows the combination of linguistic knowledge and numerical data in a systematic manner, extending classical rule-based systems based on IF-THEN rules. Fuzzy rule-based systems are built upon the fuzzy concept introduced by Zadeh in 1965, capturing human expert reasoning in production rules and addressing real-life problems in various domains, including control, prediction, inference, data mining, bioinformatics, robotics, and speech recognition. On the other hand, KNN is a lazy learning non-parametric algorithm that requires all the training data during the testing phase, providing valuable information for classification (Rancapan et al., 2019).

Image processing involves performing operations on an image, including analysis and manipulation, to enhance its quality. It is a form of signal processing where the input is an image, and the output can be another image or image characteristics/features. In recent times, image processing has emerged as a rapidly growing technology with diverse applications across various fields of study.

By integrating these methodologies, the study successfully combined image processing, fuzzy logic, and KNN to develop a comprehensive approach for the classification of eggplant seeds. The image processing techniques enhanced the quality of the captured images, enabling the extraction of essential features. The extracted features were then used as inputs for both the fuzzy logic system and the KNN algorithm, facilitating the classification process.

Overall, the chosen methodologies offered a robust framework for the analysis and classification of eggplant seeds, capitalizing on the strengths of image processing, fuzzy logic, and KNN. This integrated approach holds promise for various applications in agriculture, seed classification, and related fields.

Eggplant Seeds

The two varieties of eggplant used were from the supermarket. Since there are only two varieties available, this study had only two classes of eggplant. Samples were arranged twenty each and placed on a white background and then captured using an 8-megapixel cellphone camera placed 10 inches directly above the samples. Captured images were then stored in JPEG files and were used for the image processing.

Image Processing

The JPEG file was loaded into MATLAB, specifically version 2020, for image processing. The code used in the image processing is depicted in Figure 1, with the area being the only variable depending on the eggplant variety. To process an image, the 'vislabels' function must be installed beforehand.


Command Window

```
a=imread('a1.jpg');
b=rgb2gray(a);
c=300-b;
d=im2bw(c);
imshow(d)
e=imfill(d,'holes');
imshow(e)
f=bwlabel(e);

figure,imshow(f),title ('Labelled Image in black and  white')
vislabels(f), title('Each object labelled')
g=regionprops(f, 'Area', 'BoundingBox');
p=regionprops(f, 'Perimeter', 'BoundingBox');
z=regionprops(f, 'EquivDiameter', 'BoundingBox');
area_values=[g.Area]

idx=find((1000 <=area_values)&(area_values<=15000))
h=ismember(f,idx);
figure,imshow(h),title('Area between 1000  and 15000')


%For area
x=g(idx);
%For perimeter
y=p(idx);
%for Equivalent Diameter
w=z(idx);
```

Figure 1. *Code for the Image Processing*

Classification Process

The image processing results yielded four seed features: area, perimeter, equivalent diameter, and roundness. These features were utilized in the fuzzy logic and k-nearest neighbors (KNN) classification processes. KNN was chosen as the classification technique due to its widespread usage and effectiveness in classification tasks.

Results and Discussion

In this study system, the inputs had fuzzy sets according to the range that they fell in. But first, pictures of seeds were taken and underwent image processing. Although there are many varieties of beans and peas, they have certain ranges of areas they fall in. Aside from area, perimeter, equivalent diameter, and roundness were also acquired through image processing.

Image Enhancement

The image was first read in MATLAB using the code 'imread.' Then the original image was complemented to make the subject in color white, while the background in color black. After the image was complemented, the holes were filled in each subject and labeled. The labeled object shows that they are the seeds in the picture. Lastly, after the image was labeled, it was then converted into black and white with less noise so the subject can be easily seen and the parameters of features needed can easily be obtained.

Figure 2. *Original Image*

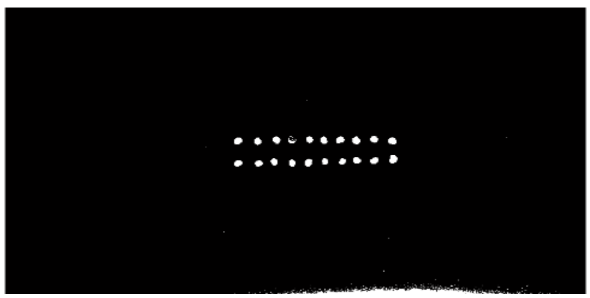

Figure 3. *Original Image in Black and White*

Figure 4. *Filled Images Labeled*

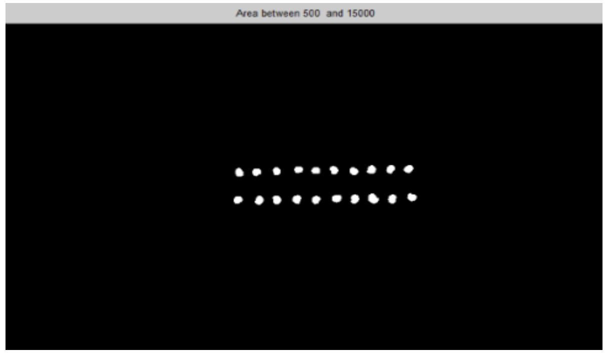

Figure 5. *Filled Images in Black and White*

Parameters Extraction

In Figure 4, wherein each seed was already labeled, the extraction of the parameters such as the area, the perimeter, and the equivalent distance was easily obtainable. While the roundness of the seeds was calculated using the formula R = ((4*Area*pie)/ (Perimeter^2)) *100. The parameters obtained in each of the data sets are summarized in Table 1.

Table 1. Summary of the Range of Each Seed Feature of the Two Eggplant Varieties

Features	Philippine Variety	Indian Variety
Area	1218 - 2193	1820-3005
Perimeter	130.8112- 178.7523	159.1932- 225.9655
Equivalent Diameter	30.6122-52.8414	48.1838- 60.8554
Roundness	86.2474-91.3115	73.9556- 90.8654

Table 2 shows the mean of the values of the parameters of the 60 seeds of the two varieties of the eggplant. It can be concluded that the Indian variety of eggplant has a larger area, perimeter, and equivalent diameter than the Philippine variety.

Table 2. Mean Values of the Parameters of Each Seed Feature of the Two Eggplant Varieties

Features	Philippine Variety	Indian Variety
Area	1550.15	2240.5838
Perimeter	147.8558	181.1173
Equivalent Diameter	43.9381	53.2556
Roundness	88.8458	85.6599

Classification of the Variety of the Eggplant Using Fuzzy Logic

After the ranges of different features of each variety of eggplant were obtained, the fuzzy logic system was then applied. Figures 6 to 9 show the fuzzy logic system of each feature.

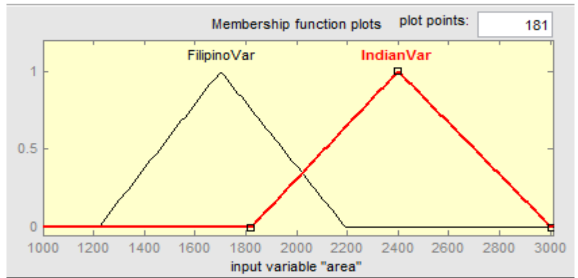

Figure 6. *Area as Input*

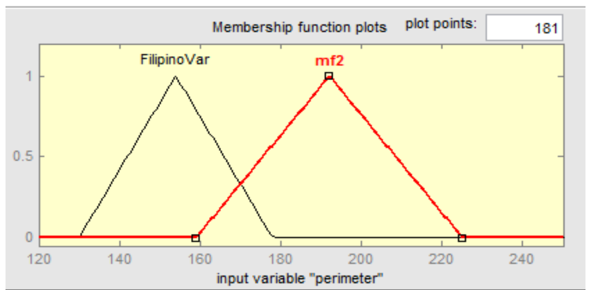

Figure 7. *Perimeter as Input*

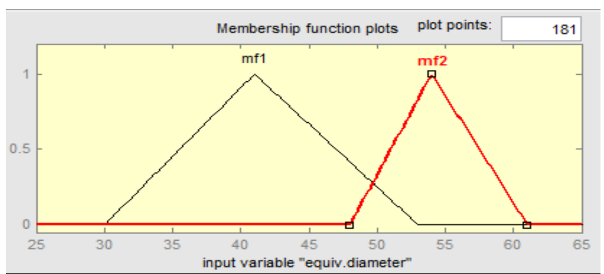

Figure 8. *Equivalent Diameter as Input*

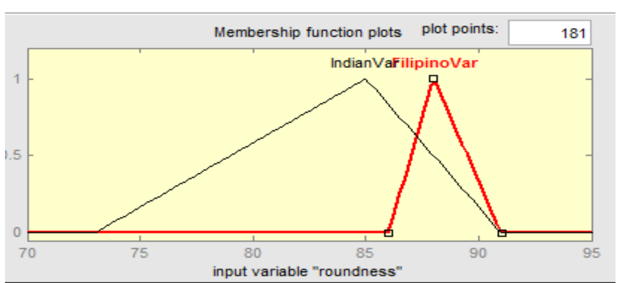

Figure 9. *Roundness as Input*

K-nearest neighbor (KNN) was utilized as a classification technique in the study to classify the seeds of eggplants. This technique requires training data and test data. In this particular research, 60 training data points were utilized for the Philippine variety eggplant, and an additional 60 training data points were used for the Indian variety eggplant. Moreover, 20 test data points were employed for each variety to evaluate the effectiveness of the classification. The following figure illustrates the results of the classification of the test data.

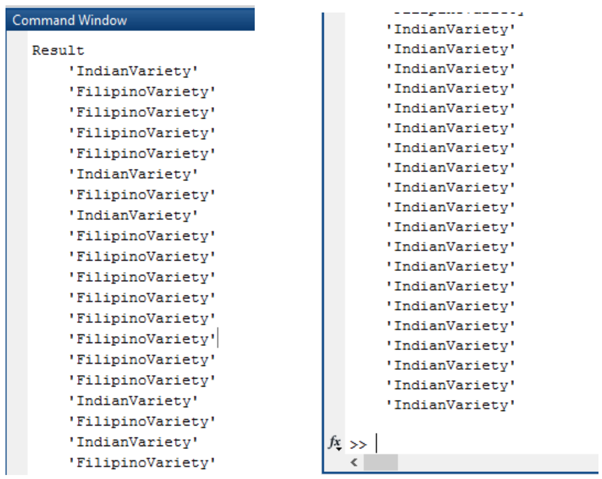

Figure 10. *Result of the Classification of the Test Data Using KNN*

Table 3. Summary of the Percentage of Right Classification Using KNN

K	Percentage of Right Classification	
	Philippine Variety	**Indian Variety**
1	75%	100%

Table 3 shows that in the Indian variety classification, there is no error in classifying while in the Philippine variety, there is a 25% error in classifying the eggplant seeds.

Conclusion and Future Works

This study highlighted the successful application of image processing, fuzzy logic, and k-nearest neighbors (KNN) in the identification of eggplant varieties based on seed characteristics. The results showed significant differences between the Indian and Philippine eggplant varieties, with notable variations in seed features such as area, perimeter, and equivalent diameter. The fuzzy logic system and KNN classifier achieved high accuracy rates in classifying the eggplant seeds, with the Indian variety exhibiting 100% accuracy and the Philippine variety encountering a slightly higher error rate of 25%.

Despite the promising findings, it is crucial to acknowledge the limitations of this study. To further improve the classification accuracy, future research should encompass a wider range of eggplant varieties and consider utilizing updated software tools or alternative image processing software for enhanced analysis and feature extraction.

Moreover, the implications of this research extend beyond eggplant classification, offering potential applications in other crops and plant species. This innovation holds promise for advancing agricultural research and crop classification. The integration of image processing, fuzzy logic, and KNN classifiers also holds significant potential in areas such as plant disease detection, quality control in food production, and pattern recognition across diverse industries.

In addition, expanding the dataset, exploring alternative machine learning techniques, incorporating additional image features, and investigating the sensitivity of the proposed approach to different lighting conditions or seed sizes are recommended. Moreover, the application of image processing, fuzzy logic, and KNN classification in
plant species classification holds significant promise, and future works can extend this approach to different crops and plants, advancing agricultural research, quality control in food production, and the detection of plant diseases.

In conclusion, this study provides compelling evidence of the effectiveness of image processing, fuzzy logic, and KNN classification in accurately identifying different eggplant seed varieties. To maximize the impact of this research, future work should address the identified limitations by exploring additional eggplant varieties, updating software tools, and extending the proposed approach to encompass different crops and domains. These collective efforts will undoubtedly contribute to the progress of crop classification and the advancement of agricultural technology.

References

[1] Andaya, A. E., Arboleda, E. R., Andilab, A. A., & Dellosa, R. M. (2019). Meat marbling scoring using image processing with fuzzy logic-based classifier. *International Journal of Scientific and Technology Research, 8*(8), 1442–1445.

[2] Arboleda, E. R., Jesus, C. L. T. D., & Tia, L. M. S. (2021). Pineapple maturity classifier using image processing and fuzzy logic. *IAES International Journal of Artificial Intelligence, 10*(4), 830–838. https://doi.org/10.11591/ijai.v10.i4.pp830-838

[3] Austria, V. A. C., Arboleda, E. R., & Galas, E. M. (2019). Image processing of clean and dirty dishes to design and construct a fuzzy logic dishwasher. *International Journal of Scientific and Technology Research, 8*(12), 2470–2473.

[4] Bae, R. E. T., Arboleda, E. R., Andilab, A., & Dellosa, R. M. (2019). Implementation of template matching, fuzzy logic, and k-nearest neighbor classifier on Philippine banknote recognition system. *International Journal of Scientific and Technology Research, 8*(8), 1451–1453.

[5] Baluran, C. I. A., Arboleda, E. R., Dizon, M. G., & Dellosa, R. M. (2019). Crab gender classification using image processing, fuzzy logic, and k-nearest neighbor (KNN) classifier. *International Journal of Scientific and Technology Research, 8*(10), 1458–1462.

[6] Dioses, J. L., Jr. (2024, June 29). Mung bean variety classification using KNN and image processing technique. In *Proceedings of the 2024 IEEE International Conference on Automatic Control and Intelligent Systems (I2CACIS)* (pp. 71). IEEE. https://doi.org/10.1109/I2CACIS61270.2024.10649868

[7] Garcia, J. A. A., Arboleda, E. R., & Galas, E. M. (2020). Identification of visually similar vegetable seeds using image processing and fuzzy logic. *International Journal of Scientific and Technology Research, 9*(2), 4925–4928.

[8] Hautea, D., Narciso, J., Sagadraca, R., Buerano, C., Hautea, R., Abad, J., & Obsioma, A. (2016). Field performance of Bt eggplants (Solanum melongena L.) in the Philippines: Cry1Ac expression and control of the eggplant fruit and shoot borer (Leucinodes orbonalis Guenée). *PLoS ONE, 11*(6), e0157498. https://doi.org/10.1371/journal.pone.0157498

[9] Lajom, M. P., Remigio, J. P., Arboleda, E., & Sacala, R. J. R. (2024). Design and development of eggplant fruit and shoot borer (Leucinodes orbonalis) detector using near-infrared spectroscopy. *Journal of Engineering and Sustainable Development, 28*(4), 439–454. https://doi.org/10.31272/jeasd.28.4.3

[10] Loresco, J. M., & Dadios, E. P. (2018, July). *Fuzzy logic-based identification method for eggplant class.* In *Proceedings of the 5th International Conference on Communication and Computer Engineering (ICOCOE 2018)*, Swiss-Garden Hotel Melaka, Melaka, Malaysia. https://www.researchgate.net/publication/334318199_Fuzzy_Logic-Based_Identification_Method_for_Eggplant_Class

[11] Rancapan, J. G., Arboleda, E. R., Dioses, J. L., Jr., & Dellosa, R. M. (2019). *Egg fertility detection using image processing and fuzzy logic. International Journal of Scientific and Technology Research, 8*(10), 3228–3230. https://www.researchgate.net/publication/336991910_Egg_Fertility_Detection_Using_Image_Processing_And_Fuzzy_Logic

Conflict of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.