TENCON 2023 - 2023 IEEE Region 10 Conference (TENCON)
31 Oct - 3 Nov 2023. Chiang Mai, Thailand

A Machine Learning-based Approach for Accurate Size Classification of Pineapple (ananas comosus)

Robert G. de Luna
Polytechnic University of the Philippines
Sto. Tomas City, Batangas, Philippines
rgdeluna@pup.edu.ph

Verna C. Magnaye
Polytechnic University of the Philippines
Sto. Tomas City, Batangas, Philippines
vcmagnaye@pup.edu.ph

Rose Anne L. Reaño
Polytechnic University of the Philippines
Sto. Tomas City, Batangas, Philippines
ralreano@pup.edu.ph

Karina L. Enriquez
Polytechnic University of the Philippines
Sto. Tomas City, Batangas, Philippines
klenriquez@pup.edu.ph

Rai Racel Armando
Polytechnic University of the Philippines
Sto. Tomas City, Batangas, Philippines
rrvarmando@iskolarngbayan.pup.edu.ph

Mark Louie Bocalbos
Polytechnic University of the Philippines
Sto. Tomas City, Batangas, Philippines
mlcbocalbos@iskolarngbayan.pup.edu.ph

Jamaica Fernandez
Polytechnic University of the Philippines
Sto. Tomas City, Batangas, Philippines
jamcfernandez@iskolarngbayan.pup.edu.ph

Krystel Anne Malacaman
Polytechnic University of the Philippines
Sto. Tomas City, Batangas, Philippines
karmalacaman@iskolarngbayan.pup.edu.ph

Jesirie Natividad
Polytechnic University of the Philippines
Sto. Tomas City, Batangas, Philippines
jnnatividad@iskolarngbayan.pup.edu.ph

Jan Jadrien Ramos
Polytechnic University of the Philippines
Sto. Tomas City, Batangas, Philippines
janjadrienramos@iskolarngbayan.pup.edu.ph

Shaina Marie Salcedo
Polytechnic University of the Philippines
Sto. Tomas City, Batangas, Philippines
smvsalcedo@iskolarngbayan.pup.edu.ph

Abstract— **Pineapple's size is very crucial in determining its market value. Size sorting is commonly done via visual inspection, which is usually subject to inconsistency and errors. Errors due to failed sorting may either lead to wastage or loss, or mispricing. This study presents incorporation of the machine learning techniques like Logistic Regression, K-Nearest Neighbors, Support Vector Machine, Decision Tree, and Random Forest in classifying pineapple sizes as small, medium, and large using the extracted features of images processed via OpenCV libraries as well as Python Programming. A total of 300 pineapples of different sizes were captured and processed to extract features such as the area, width, height, enclosed-circle radius, and perimeter. The models were optimized using GridSearchCV and were evaluated using accuracy and F1 score metrics. Based on the results, SVM was found to be the most suited classification model, having an optimized training and testing accuracy of 95.67 % and 96.67 %, respectively, and an F1 score of 96.67 %.**

Keywords— pineapples, size classification, machine learning, image processing

I. INTRODUCTION

The Philippines is the top global producer of Pineapple (Ananas comosus), locally known as Pinya [1][2]. It has been the source of livelihood for Filipino farmers, particularly in Camarines Norte and Northern Mindanao [3]. However, the industry is also faced with challenges particularly, post-harvest wastage and losses [4], amounting to a 40% loss of the total production [5]. Rough handling and failed sorting cause these losses [6]. Sorting is a crucial process because sorting according to size serves as the basis of pricing each size classification of pineapple.

Currently, farmers manually grade and sort pineapples according to their size, weight, and color, based on their experiences. These laborious, intensive manual estimation through visual inspection are not reliable [7] as factors like fatigue and quality of eyesight, and uncalibrated weighing scales, could affect the results, even more, it is time-consuming to accomplish. These could lead to inaccuracy and inconsistency in classifying pineapples, whereas it calls for a systematic, accurate, and an efficient classification system based on widely used methods.

This study primarily aims to develop a pineapple size classification model, particularly, the Smooth Cayenne pineapple variety, that will aid our local farmers in a systematized and efficient sorting. Proponents believe that with this model, pineapple wastage due to mishandling, and losses due to failed sorting and mispricing will be prevented.

II. REVIEW OF RELATED WORKS

Studies designed to better facilitate the classification, recognition, and counting of pineapples have been initiated. Pineapple detection and size determination has been ventured using the SVM technique where Speed-Up Robust Features (SURF) [8], and OpenCV Library [9] were used to conduct feature extraction. Pineapple fruit crown identification, detection, and counting also used machine learning techniques and implemented feature selection via ANOVA to decrease dimensionality while improving and optimizing the classification accuracy [10]. Meanwhile, acoustic spectroscopy was also used to classify pineapples using the drum and meat sound quality of pineapples with an accuracy of 0.97 [11].

The color of pineapple scales was also studied to identify the maturity of pineapple using fuzzy logic-based classifiers [12]. The size and weight of pineapples can also be predicted with acceptable accuracy using the Artificial Neural Network (ANN)-calibrated low-cost vision system using MATLAB-based hardware and software utilities [7].

Machine Learning algorithms including ANN, SVM, and K-Nearest Neighbor (KNN) have been used in size classification and grading of other agricultural products such as gooseberry fruit's ripeness [13], citrus and dragon fruit [14] [15], and strawberry shapes [16]. In a study on tomato grading and size classification [17][18][19], geometrical features specifically the area, the perimeter, and enclosed-circle radius were extracted. KNN and SVM were found to be the best models with an accuracy of 97.5% and 95% accuracy respectively. A similar study on tangerine size classification [20] also proved that KNN and SVM were the best models. Furthermore, SVM was able to yield 100% accuracy in fruit classification, where a system was developed using MATLAB for image processing of bananas and apples [21].

Fuzzy logic, ANN, SVM, and an Adaptive Network-based Interference System (ANFIS) were used to obtain an analysis of different fruit diseases, and ANFIS was found to be the most accurate [22]. Random Forest was proven to be the most accurate model in identifying fruit diseases versus KNN and Decision Tree [23], and also in banana size classification versus ANN and SVM [24]. Meanwhile, a study on cherry tomatoes' mass and volume prediction showed that RBF-SVM performed better compared to other models showing accuracy results of 0.9706 and 0.9694, for 2D and other features, respectively [25].

Additionally, image processing together with machine learning [26][27] was also used in classifying some fruits where the proponents utilized image acquisition [28], LabVIEW, KNN, DT, Naive Bayes, Random Forest, and MLP as algorithms. The system success rate with this method was 95% for Esmek quince and 86% for Esme quince [29].

Deep Learning techniques have also been utilized in processing agricultural products such as fruit recognition [30], potato size classification [31], apple fruit size estimation [32][33], mango quality and size detection [34][35], orange quality grading [36], olive fruit grading [37], onion sorting and grading [38], okra grading [39], and calamansi size classification.

Certain initiatives to facilitate the classification and recognition of pineapples were conducted such as type classification, crown recognition, and maturity grading. However, there is no study that's exclusive and definite on the classification of pineapples based on size. A better machine learning technique, a camera with a higher resolution, and a larger dataset, should be used to further improve the results. Existing standards were set to classify pineapple sizes based on weight, but the pineapple local market commonly sells pineapples based on the visual output that varies on the perspective of the observer, thus, indicating the need for an objective and systematized pineapple size classification based on computer vision using the machine learning algorithms.

Machine Learning techniques have been utilized in most agricultural classifications, but they are lacking in terms of cross-validation, optimization methods, and performance matrices. Cross-validation is important to prevent bias in the data, while optimization should be utilized to improve the model's accuracy using hyperparameter tuning. Performance matrices could have also been used to evaluate the error and accuracy of the built system. Hence, the proponents decided to use machine learning techniques, apply optimization using the GridSearchCV, and evaluate the models' performance using certain performance matrices.

In this study, the proponents utilized the OpenCV Library and Python programming for image processing and feature extraction where the samples were classified into small, medium, and large classes. The extracted features include the area, width, height, enclosed-circle radius, and perimeter. A self-generated dataset with 300 pineapple samples was created using 100 samples for each size classification. As determined in related studies as the most accurate models for classification, the KNN, SVM, LR, Decision Tree, and Random Forest were machine learning algorithms used by the proponents in developing the Pineapple Size Classification Model.

III. Methodology

A. Dataset Description

This study used 300 images of Smooth Cayenne pineapples categorized as small, medium, and large. Each size classification is comprised of 100 samples distributed as 80 and 20 for the training and the testing set, respectively. The dataset was gathered on a pineapple farm in San Agustin, in the City of Santo Tomas, Batangas using an improvised capturing box and a 50-megapixel Android camera. Natural lighting and a capturing distance of 75 cm from the pineapple were the main setups for image acquisition. Fig. 1 shows the pineapple samples consisting of three different sizes.

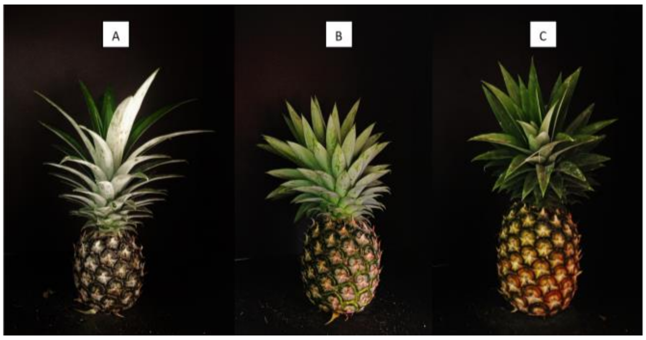


Fig. 1. Pineapple Samples (A) small, (B) medium, (C) large.

B. System Overview

Displayed in Fig. 2 is the overview of the system consisting of the process of image acquisition, image processing, feature extraction, model creation then evaluation.

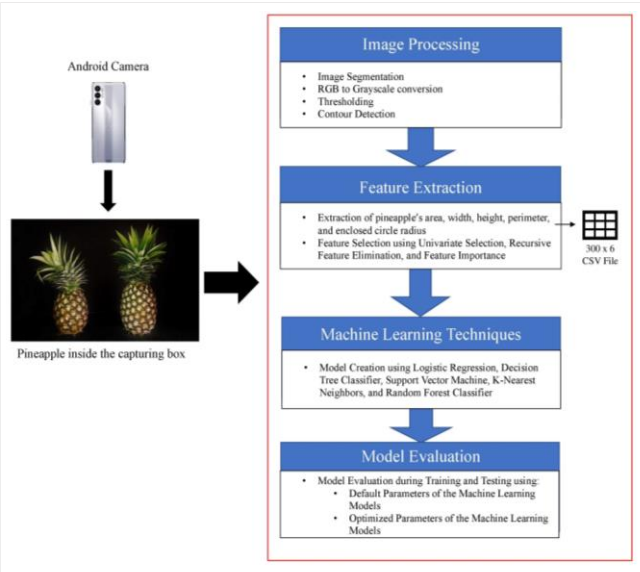


Fig. 2. System Overview

In image acquisition, pineapples were captured in an improvised capturing box with a black background 75cm far from the camera, with the pineapples placed 23 cm apart from each other. The samples were then subjected to image processing using OpenCV libraries and Python programming to enhance the image and extract features. The image consisting of two pineapples was first segmented to produce two images of a single pineapple and then cropped to include only the fruit part. The images were then converted from RGB images to grayscale and applied with thresholding to define the contour of the subject. The contour was used as the basis of geometrical features such as the area, the perimeter, the enclosed-circle radius, width, and height. A comma-separated value (CSV) file was made from the extracted dataset consisting of 300 rows and 6 columns. The rows corresponds to the number of samples, and the columns represent the features, including the area, perimeter, enclosed-circle radius, width, and height, as well as the class. Feature selection was applied to the data to determine the three most significant features that can predict the class and be utilized in the creation of different machine learning (ML) models.

Each model was created using default parameters and was optimized using GridSearchCV to improve their performance. The models were assessed using F1 score metrics and classification accuracy to determine which model is the best fit for the size classification task. The model's F1 score for each class was tabulated to evaluate which of the models classify classes correctly and to determine which of the sizes the models have difficulty classifying.

C. *Development of the Size Classification Models*

The proponents adopt five supervised ML algorithms commonly used for classification tasks, namely LR, DTC, SVM, KNN, and RFC.

Logistic Regression is a simple and widely used linear classification algorithm that models the probability of the target variable using a logistic function. In LR, each category has its threshold value, and the category with the highest probability score is assumed as the output class for the given sample. Another algorithm for classification tasks is the DTC, which is a flowchart-like structure that divides the data into subsections, according to the most informative features, creating a tree-like formation where each leaf node correlates to a class label. The process of making predictions using decision trees is easier to interpret and understand compared to other algorithms. Moreover, the RFC is an ensemble learning algorithm with improved accuracy. It combines multiple decision trees. The results of the trees are synthesized to make a final prediction. Another popular classification algorithm is the SVM which uses a kernel function to transform the input data into higher dimensional space and then locates an ideal hyperplane that accurately separates data into different classes. Furthermore, KNN is known to be a non-parametric model that categorizes data according to the majority class of its k-nearest neighbors in the feature space. It is effective for size classification tasks because it can handle both categorical and numerical data, and can work well with both linear and nonlinear data.

To develop the models, the proponents utilized the Python programming language. Different training-testing ratios of the dataset were performed and the models' performance for each set was tabulated. The training set was used in model construction, while the testing set was employed to evaluate and validate the built models. Moreover, the GridSearchCV is used to optimize the models in finding the best set of hyperparameters that will produce the best results.

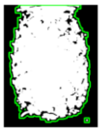
To assess the model's performance, the classification accuracy as well as the F1 score metrics were used. Accuracy measures the ratio of accurately classified samples to the overall number of samples in a dataset while the F1 score is a metric that combines both precision and recall and provides a more balanced analysis of the model's performance. The accuracy will determine the best model to implement in the size classification task, while the F1 score will determine which of the models can classify the samples into their correct sizes.

IV. Results and Discussions

A. *Dataset Generation for Machine Learning*

Pineapple images were gathered using an improvised capturing box and a 50 megapixels resolution Android camera. The captured images were processed using the OpenCV libraries of Python to transform the image and extract geometrical features. The image processes applied to all 300 samples were shown in Table I.

TABLE I. Sample Results of the Image Processing and Feature Extraction

Process	Result
Image Acquisition	 Original Pineapple Image
Image Segmentation	 Segmented Image  Cropped Image
RGB to Grayscale Conversion	 RGB Image  Grayscale Image
Thresholding	 Grayscale Image  Binary Image
Contour Detection	 Binary Image  Contour Image
Feature Extraction	 Extracted Features Area: 416780.50 Width: 594 Height: 895 Perimeter: 5717.86 Enclosed circle radius: 467.14

The generated values are compiled and saved into CSV file format to create a dataset of 300 rows by 6 columns, where each row represents a sample, and the columns represent the features and class label. The descriptive analysis of the dataset is presented in Table II.

TABLE II. DESCRIPTIVE ANALYSIS OF THE DATASET

	area	perimeter	enclosed_circle_radius	width	height
count	300	300	300	300	300
mean	305715.55	4401.10	382.21	588.09	695.09
Std	96716.56	1162.74	69.84	86.34	135.80
min	128718.50	1813.09	237.23	410	411
max	491310	7790.49	569.41	764	1009

The dataset has five features, namely area, perimeter, enclosed-circle radius, width, and height. The dataset has no missing values, with each column having a 300-value count. The features have a numerical float data type, while the target column, class, has three categorical values specifically small, medium, and large. These categorical values were converted into numerical values using the Label Encoder. The features were also normalized using the MinMax Scaler to ensure that all features are on a similar scale.

Fig. 3. Results of the feature selection methods

Features are then subjected to a feature selection process. This method helps to eliminate redundant and less important features and reduce training time by using smaller datasets, which can improve the performance of the models. Three different methods for feature selection were employed: Univariate Selection, Recursive Feature Elimination, and Feature Importance. The outcome of each method is shown in Fig. 3. After analyzing the results of the feature selection methods, the top three features—area, width, and height—were identified and selected to be used for the model creation.

B. Model Performance for Pineapple Size Classification

The researchers developed five different machine learning models for the size classification task, including LR, DTC, SVM, KNN, and RFC. The dataset was split into 80 % training data and 20 % testing data. The 10-fold stratified cross-validation and hold-out validation techniques were used to assess the models' performance during training and testing phases, respectively. Moreover, by tuning the hyperparameters into their optimal value to achieve the best possible performance from each model, the models were optimized using the GridSearchCV to improve their performance.

To determine the best machine learning model for the size classification task, the accuracies and F1 scores of each model were tabulated and compared to identify which model is best suited for the size classification task.

Table III summarizes the results of the training and testing accuracy of each model using default and optimized parameters for the standard training-testing ratio of 80-20. Moreover, the F1 score performance of each model for 80-20 data splitting was also presented in Table IV.

TABLE III. ACCURACY PERFORMANCE OF THE SIZE CLASSIFICATION MODELS FOR 80-20 DATA SPLITTING

Model	**Cross-Validation**		**Hold-out Validation**	
	Default Parameter Accuracy (%)	***Optimized Parameter Accuracy (%)***	***Default Parameter Accuracy (%)***	***Optimized Parameter Accuracy (%)***
LR	94.67	95.33	90.00	96.67
DTC	92.67	95.33	90.00	93.33
SVM	95.33	95.67	96.67	96.67
KNN	95.67	96.00	93.33	93.33
RFC	94.33	95.33	96.67	95.00

Based on Table III, the KNN has the highest training accuracy using both the default and optimized parameters, followed by SVM, LR, RFC, and DTC, respectively. Moreover, a significant increase in the model's accuracy during the training phase after optimization can be observed. On the other hand, the accuracy of the models during testing also improved after performing optimization, except for the RFC, in which the testing accuracy using the default parameter is better than the result using the optimized parameters. Furthermore, the testing performance of SVM using default and optimized parameters is consistent and considered the highest compared with other models. Based on the model's accuracy performance, SVM is the best fit for the size classification task as its accuracy during training and testing phase remained consistent.

Similarly, as presented in Table IV, the KNN also has the highest F1 score during training phase using default and optimized parameters, followed by SVM. A significant increase in models' performance after optimization can be observed. Same with the accuracy performance, the SVM yields the highest F1 score in the testing phase for both the default and optimized model. The consistency of SVM's F1 score performance in training and testing concludes that the model classifies most of the samples into their correct sizes which indicates that it is a good model to implement in the size classification of pineapples.

TABLE IV. F1 SCORE PERFORMANCE OF THE SIZE CLASSIFICATION MODELS FOR THE 80-20 DATA SPLITTING

Model	Cross-Validation		Hold-out Validation	
	Default Parameter F1 Score (%)	***Optimized Parameter F1 Score (%)***	***Default Parameter F1 Score (%)***	***Optimized Parameter F1 Score (%)***
LR	94.60	95.31	89.90	96.67
DTC	92.57	95.32	89.76	93.32
SVM	95.28	95.62	96.67	96.67
KNN	95.62	95.95	93.32	93.32
RFC	94.35	95.32	96.67	95.00

Presented in Table V is the summary of the F1 score performance per size using the optimized machine learning models. In addition, the confusion matrix of each optimized model for the 80-20 training-testing ratio is shown in Fig.4.

TABLE V. F1 SCORE PERFORMANCE PER SIZE USING THE OPTIMIZED MACHINE LEARNING MODELS (IN %)

Size	LR, %	DTC, %	SVM, %	KNN, %	RFC, %
Large	100	100	100	100	100
Medium	95	89.47	95	89.47	92.68
Small	95	90.48	95	90.48	92.31
No. of Items with F1 score above 90%	3	2	3	2	3
No. of Items with F1 score below 90%	0	1	0	1	0

All models predicted large pineapples correctly, with a perfect 100 % F1 score for each model. Furthermore, the F1 score for small sizes is favorable, with the models having above 90 % performance. However, both DTC and KNN have an F1 score below 90 % for medium size. This result indicates that medium-sized pineapples are the most challenging to classify. These may be due to the geometrical features of the medium-sized pineapple being very close to the small-sized pineapple.

Based on the confusion matrix shown in Fig. 4, all models correctly predicted the large pineapples. However, when predicting the small and medium-sized pineapples, the models exhibited a degree of confusion. A possible reason is that the medium-sized pineapple shares geometric properties that closely resemble those of the small-sized pineapple.

V. CONCLUSIONS

The developed pineapple size classification using the SVM, RFC, and LR models combined with computer vision techniques achieved a consistently high optimized training and testing accuracy and F1 score. The image samples were acquired using an improvised capturing box and a 50-megapixel Android camera and were processed using OpenCV libraries of Python to extract features and generate a dataset. Using selected features, such as the area, width, and height, machine learning models were built in which the

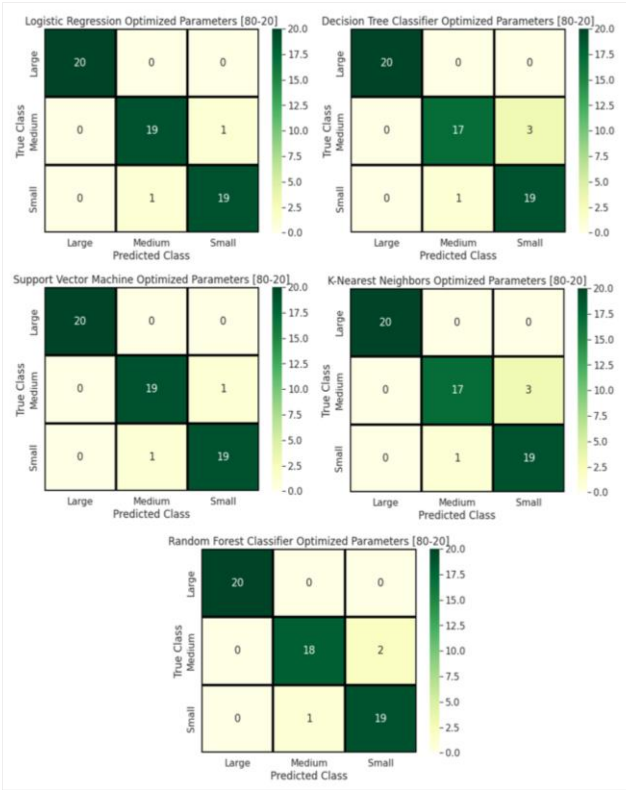


Fig. 4. Confusion matrix of each optimized models

SVM model performed consistently, with an optimized training and testing accuracy of 95.67 % and 96.67 %, and an F1 score of 96.67 %. Similarly, the performances of LR and RFC models are also notably high with both having a training accuracy of 95.33 % and testing accuracy of 96.67 % for LR while 95 % for RFC. These models achieved an F1 score of 96.67 % for LR and 95 % for RFC while DTC and KNN struggled to classify medium-sized pineapples, having an F1 score lower than 90 % for the said size.

VI. ACKNOWLEDGMENT

The authors would like to acknowledge the support and resources provided by Polytechnic University of the Philippines (PUP). The facilities, funding, and access to relevant literature and research materials have greatly contributed to the success of this project.

REFERENCES

[1] Statista. (2023, January 6). Global pineapple production by leading countries 2021. https://www.statista.com/statistics/298517/global-pineapple-production-by-leading-countries/

[2] Food Philippine Administrator. (2022, February 8). An overview on Philippine pineapples. Food Philippines. https://foodphilippines.com/story/an-overview-on-philippine-pineapples/

[3] Philippine Statistic Authority. (2022). Major Fruit Crops Quarterly Bulletin, July-September 2022. Retrieved from https://psa.gov.ph/fruits-crops-bulletin/pineapple

[4] Pineapple Export Act of 2022. S.B 1256, 19th Cong. (2022). https://legacy.senate.gov.ph/lis/bill_res.aspx?congress=19&q=SBN-1256

[5] Queen Pineapple Industry Profile. (n.d.). PCAARRD's Industry Strategic Science and Technology Plans. Retrieved January 19, 2023, from https://ispweb.pcaarrd.dost.gov.ph/queen-pineapple/

[6] Wahab, N. F., & Khairuddin, F. (2020). Risk Identifications of Postharvest Handling in Pineapple Crop that Affecting the Fruits

Quality. International Journal of Academic Research in Business and Social Sciences, 10(9). https://doi.org/10.6007/ijarbss/v10-i9/7858

[7] Al-kaf, H. A. G., Chia, K. S., & mubin, F. R. B. A. (2020, May 1). The Size And Weight Prediction For Intact Pineapples Using A Low Cost Vision System. IEEE Xplore. https://doi.org/10.1109/ZINC50678.2020.9161770

[8] Woods, N. C., Abuh, E. O. & Robert, A. B. C. (2019). Development of a Pineapple Fruit Recognition and Counting System using Digital Farm Image, Afr. J. Comp. & ICT, 12(2), 131-139.: https://www.semanticscholar.org/paper/Development-of-a-Pineapple-Fruit-Recognition-and-Woods-Abuh/8ad1e8400fd0de3db98f61e1c596cccd4aa35bb9

[9] Harris, J. Object detection and size determination of pineapple fruit input to a juicing factory in Eastern Cape, South Africa: A deep learning approach. (2022, September 21). https://doi.org/10.21203/rs.3.rs-1942268/v2

[10] Wan Nurazwin Syazwani, R., Muhammad Asraf, H., Megat Syahirul Amin, M. A., & Nur Dalila, K. A. (2022). Automated image identification, detection and fruit counting of top-view pineapple crown using machine learning. Alexandria Engineering Journal, 61(2), 1265–1276. https://doi.org/10.1016/j.aej.2021.06.053

[11] Huang, T.-W., Bhat, S. A., Huang, N.-F., Chang, C.-Y., Chan, P.-C., & Elepano, A. R. (2022). Artificial Intelligence-Based Real-Time Pineapple Quality Classification Using Acoustic Spectroscopy. Agriculture, 12(2), 129. https://doi.org/10.3390/agriculture12020129

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

[13] Castro, W., Oblitas, J., De-La-Torre, M., Cotrina, C., Bazan, K., & Avila-George, H. (2019). Classification of Cape Gooseberry Fruit According to its Level of Ripeness Using Machine Learning Techniques and Different Color Spaces. IEEE Access, 7, 27389–27400. https://doi.org/10.1109/access.2019.2898223

[14] Sugumar, D., Harshavarthan, V., Kavisri, S., Aezhisai Vallavi, M. S., & Vanathi, P. T. (2019). Citrus Classification and Grading Using Machine Learning Algorithms. International Journal of Innovative Technology and Exploring Engineering (IJITEE), 8(10). http://doi.org/10.35940/ijitee.J9349.0881019

[15] Patil, P. U., Lande, S. B., Nagalkar, V. J., Nikam, S. B., & Wakchaure, G. C. (2021). Grading and sorting technique of dragon fruits using machine learning algorithms. Journal of Agriculture and Food Research, 4, 100118. https://doi.org/10.1016/j.jafr.2021.100118

[16] Ishikawa, T., Hayashi, A., Nagamatsu, S., Kyutoku, Y., Dan, I., Wada, T., Oku, K., Saeki, Y., Uto, T., Tanabata, T., Isobe, S., & Kochi, N. (2018). Classification of Strawberry Fruit Shape By Machine Learning. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLII-2, 463–470. https://doi.org/10.5194/isprs-archives-xlii-2-463-2018

[17] Tamakuwala, S., Lavji, J., & Patel, R. (2018). Quality Identification of Tomato Using Image Processing Techniques. International Journal of Electrical, Electronics and Data Communication, 6, 2321–2950. https://www.semanticscholar.org/paper/Quality-identification-of-tomato-using-image-Tamakuwala-Lavji/b40804969a0bf56af09a04df59d6bcb83f6964c1

[18] Sari, M. I., Fajar, R., Gunawan, T., & Handayani, R. (2022). The Use of Image Processing and Sensor in Tomato Sorting Machine by Color, Size, and Weight. JOIV : International Journal on Informatics Visualization, 6(1-2), 244. https://doi.org/10.30630/joiv.6.1-2.944

[19] de Luna, R. G., Dadios, E. P., Bandala, A. A., & Vicerra, R. P. (2019). Size Classification of Tomato Fruit Using Thresholding, Machine Learning, and Deep Learning Techniques. AGRIVITA Journal of Agricultural Science, 41(3), 586–596. http://doi.org/10.17503/agrivita.v41i3.2435

[20] Panprasittikit, J. & Amnatchotiphan, K. (2021). Thai Tangerine Size Classification via Computer Vision. 2021 IEEE 8th International Conference on Industrial Engineering and Applications (ICIEA). https://doi.org/10.1109/ICIEA52957.2021.9436701

[21] Chithra, P. L. & Henila, M. (2019). Fruits Classification Using Image Processing Techniques. International Journal of Computer Sciences and Engineering, 7(5), 131-135. https://doi.org/10.26438/ijcse/v7si5.14

[22] Sekar, R., & Scholar, U. (2018). Fruit Classification System Using Computer Vision: A Review. International Journal of Trend in Research and Development, 5(1), 2394–9333. https://doi.org/10.31219/osf.io/kt75d

[23] Benlachmi, Y., Airej, A. E., & Hasnaoui, M. L. (2022). Fruits Disease Classification using Machine Learning Techniques. Indonesian Journal of Electrical Engineering and Informatics (IJEEI), 10(4), 917–929. https://doi.org/10.52549/ijeei.v10i4.3907

[24] Piedad, E. J., Larada. J. I., Pojas, G. J., & Ferrer, L. V. V. (2018). Postharvest classification of banana (Musa acuminata) using tier-based machine learning. Postharvest Biology and Technology, 145, 93–100. https://doi.org/10.1016/j.postharvbio.2018.06.004

[25] Naranjo-Torres, J., Mora, M., Hernández-García, R., Barrientos, R. J., Fredes, C., & Valenzuela, A. (2020). A Review of Convolutional Neural Network Applied to Fruit Image Processing. Applied Sciences, 10(10), 3443. https://doi.org/10.3390/app10103443

[26] Hameed, K., Chai, D., & Rassau, A. (2018). A comprehensive review of fruit and vegetable classification techniques. Image and Vision Computing, 80, 24-44. https://doi.org/10.1016/j.imavis.2018.09.016

[27] Femling, F., Olsson, A., & Alonso-Fernandez, F. (2018). Fruit and Vegetable Identification Using Machine Learning for Retail Applications. 2018 14th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS). https://doi.org/10.1109/sitis.2018.00013

[28] Nyalala, I., Okinda, C., Nyalala, L., Makange, N., Chao, Q., Chao, L., Yousaf, K., & Chen, K. (2019). Tomato volume and mass estimation using computer vision and machine learning algorithms: Cherry tomato model. Journal of Food Engineering, 263, 288–298. https://doi.org/10.1016/j.jfoodeng.2019.07.012

[29] Mureşan, H., & Oltean, M. (2018). Fruit recognition from images using deep learning. Acta Universitatis Sapientiae, Informatica, 10(1), 26–42. https://doi.org/10.2478/ausi-2018-0002

[30] Gerdan Koç, D., & Vatandaş, M. (2021). Classification of Some Fruits using Image Processing and Machine Learning. Turkish Journal of Agriculture - Food Science and Technology, 9(12), 2189–2196. https://doi.org/10.24925/turjaf.v9i12.2189-2196.4445

[31] Quilloy, E., Rommel, P., Sanchez, P., Luther, J., Manuel, & Renovalles, E. (n.d.). Development of Machine Vision System for Size Classification of Potatoes (Solanum tuberosum L.). Retrieved January 20, 2023, from https://www.researchgate.net/publication/338684851_Development_of_Machine_Vision_System_for_Size_Classification_of_Potatoes_Solanum_tuberosum_L

[32] Mirbod, O., Choi, D., Heinemann, P. H., Marini, R. P., & He, L. (2023). On-tree apple fruit size estimation using stereo vision with deep learning-based occlusion handling. Biosystems Engineering, 226, 27–42. https://doi.org/10.1016/j.biosystemseng.2022.12.008

[33] Gongal, A., Karkee, M., & Amatya, S. (2018). Apple fruit size estimation using a 3D machine vision system. Information Processing in Agriculture, 5(4), 498–503. https://doi.org/10.1016/j.inpa.2018.06.002

[34] Nguyen, T. T., Nguyen, D. T., & Huynh, T. C. (2020) Sorting and Classification of Mangoes based on Artificial Intelligence. International Journal of Machine Learning and Computing, 10(2), 374–380. https://doi.org/10.18178/ijmlc.2020.10.2.945

[35] Santi, K., Behera, Kumar Sethy, P., & Rath, A. (n.d.). Image Processing Based Detection & Size Estimation of Fruit on Mango Tree Canopies. Retrieved January 20, 2023, from https://www.researchgate.net/publication/327535431_Image_Processing_Based_Detection_Size_Estimation_of_Fruit_on_Mango_Tree_Canopies

[36] Ifmalinda, Putri, R. E., & Rasinta, I. (2022). Estimation of Size, Volume and Weight of Oranges Using Digital Images Processing. IOP Conference Series: Earth and Environmental Science, 1059(1), 012016. https://doi.org/10.1088/1755-1315/1059/1/012016

[37] Ponce, J. M., Aquino, A., Millan, B., & Andujar, J. M. (2019). Automatic Counting and Individual Size and Mass Estimation of Olive-Fruits Through Computer Vision Techniques. IEEE Access, 7, 59451–59465. https://doi.org/10.1109/access.2019.2915169

[38] Paymode, A., Mohite, J., Shinde, U., & Malode, V. (2021). Artificial Intelligence for Agriculture: A technique of Vegetables Crop Onion Sorting and Grading Using Deep Learning. 6, 2456-0774. https://doi.org/10.51319/2456-0774.2021.4.0004

[39] Raikar, M. M., S M, M., Kuchanur, C., Girraddi, S., & Benagi, P. (2020). Classification and Grading of Okra-ladies finger using Deep Learning. Procedia Computer Science, 171, 2380–2389. https://doi.org/10.1016/j.procs.2020.04.258