International Journal of
Electrical and Electronics Engineering
https://www.ioro.org/ijeee

IORO

Automated Cocoa Bean Classification and Defect Detection Based on ASEAN Standard Using Image Processing and Feature-Based Analysis

Jennifer Natnat[1], Edwin Arboleda[2]

[12]*Department of College of Engineering and Information Technology Cavite State University, Indang, Cavite, Philippines*

Abstract

The current study introduces an automated cocoa bean classification and defect detection system developed in accordance with the ASEAN Standard for Cocoa Bean (ASEAN Stan 34:2014). The system, which has been developed in MATLAB, employs digital image processing and computer vision methods to identify and segment individual cocoa beans, extract their geometric, color, and texture-based features, and classify the beans into quality categories as defined by ASEAN Extra Class, Class I, Class II, and Non- Compliant. An image processing workflow that incorporates color space transformations, morphological filtering, and adaptive thresholding enhances segmentation accuracy when identifying the crop under review. Feature extraction identifies characteristics, such as area, aspect ratio, texture entropy, and color uniformity. It is additionally capable of identifying defects within the beans, including moldy, slaty, insect-damaged, and germinated beans. Validation against independent expert grading achieved a reliability rate of 96.2% indicating the system's potential to effectively assess quality and limit human bias as part of an objective quality control approach. The analyzer also has the potential to carry out rapid quality assessment and produce detailed Excel and PDF reports to support standardization and efficiency in postharvest cocoa quality control throughout the ASEAN region.

Keywords: cocoa (Theobroma cocoa L.), cocoa bean grading, digital image processing, ASEAN Standard, defect detection, MATLAB, quality assessment, computer vision

Introduction

Cocoa (Theobroma cocoa L.) is an important agricultural commodity in Southeast Asia as it is the main raw material for the chocolate and confectionery industry and an important source of livelihood for smallholder farmers. However, grading cocoa beans using manual methods is subjective and not reliable, which could lead to errors in classification as well as loss of export standards. To address this problem, an automated cocoa bean analyzer was developed in this study using MATLAB-based digital image processing technologies to classify and grade cocoa beans based on the ASEAN Standard for Cocoa bean (ASEAN Stan 34:2014).

Cocoa Bean Analysis System Design and Development

The proposed cocoa bean analyzer leverages computer vision and digital image processing algorithms to implement the process of segmenting, classifying, and extracting features in an automated way using computer analysis. The software uses an RGB-to-HSV transformation, adaptive thresholding, and morphological operations to segment and measure characteristics of the cocoa bean, such as area, length, aspect ratio, color uniformity, and entropy texture. Use of these parameters allows for classification of the defects in the cocoa bean to identify moldy, slaty, insect-damaged, and germinated beans.

System Architecture

Image Acquisition

To avoid the impact of uneven lighting, shadows and glare on the photographs of dried cocoa beans, we used a controlled (consistent) imaging environment. The cocoa beans are placed on a consistent background and the digital camera is used at a fixed/height, and the photographs of the cocoa beans can be obtained. The images of the dried cocoa beans were then calibrated to convert the pixels into actual physical size (cm) measurements for evaluating the sizes of the cocoa beans for purposes of comparison based on size.

Pre-processing

In order to improve image quality and isolate cocoa beans from each other, preprocessing steps were applied to the image data. To improve robustness to variations in ambient lighting and improve the ability to segment based on color, the RGB color space was transformed into the HSV color space. An adaptive thresholding method was employed to separate the beans in the foreground from the background then morphological operations were performed on the images, including hole filling, noise removal, and boundary smoothing. These preprocessing steps ensured that each bean was isolated accurately so that features could be extracted.

Feature Extraction and Classification

From every segmented cocoa bean, we collected three different types of measurements

- Geometric Measurements - Area, length, Width, aspect ratio, roundness and solidity were measured for Cocoa Beans to show the size of the cocoa bean and how irregular the size/shape of the bean is.
- Colour Measurements - The Colour Features (HSV) were used to gauge the consistency of the Colour of Cocoa Beans and to identify discoloured cocoa beans due to mould and/or under fermenting.
- Texture Measurements - Entropy, contrast and homogeneity were calculated using GLCM based on the surface irregularities found inside or outside the cocoa beans that would indicate some form of internal or external defects.

These features were chosen because of their relevant characteristics associated with defects defined by the ASEAN standard.

Defect Detection and Quality Classification

Finding defects involved comparing some values that were taken from the features of the beans and comparing them with thresholds that have been pre-defined to determine what the defects would be based on the specific type of defect for example - moldy, slaty, insect damaged, germinated and physical deformity and creating a total defect score for each individual defect found. The beans were then classified using the limits of defects as defined in ASEAN Stan 34:2014 into Extra Class, Class I, Class II and Non-Compliant groups. This is a rule-based classification technique that allows for compliance with regulations and eliminates the need to have to rely on a black box model to make decisions.

System Validation and Performance

The validation against expert manual grading produced an accuracy of 96.2%. The system also resulted in high precision (0.93) and recall (0.91) values to evaluate the reliability and consistency of the automated grading. The analyzer was able to process and report the analysis results in about 1.3 seconds per image; therefore, the device is suitable for assessing quality at the field level.

This automated cocoa bean analyzer provides an efficient, objective, and repeatable instrument for grading cocoa beans according to the ASEAN Stan 34:2014. By using image processing and computer vision technologies, this tool increases accuracy and standardization, limits subjectivity while modernizing postharvest cocoa grading. Future development of this tool may include the use of deep learning algorithms, IoT connection, and a mobile version for broader access and real-time analysis.

Methodology

This chapter presents the findings of the system developed for automated cocoa bean classification, feature extraction, defect detection, and grading according to ASEAN Standard for Cocoa Bean (ASEAN Stan 34:2014) using digital image processing and machine learning.

What is the accuracy and efficiency of the developed cocoa bean classification system?

Table 1: Performance Evaluation of the Cocoa Bean Analyzer – ASEAN Standard

Parameter	Result
Overall Accuracy	96.2%
Precision	0.93
Recall	0.91
Average Processing Time per Image	1.3 seconds

Table 1. Performance Evaluation of the Cocoa Bean Analyzer – ASEAN Standard Table 1 shows the performance of the Cocoa Bean Analyzer. The system achieved

2% reliability when compared to expert manual grading, indicating high precision and consistency. Precision and recall values of 0.93 and 0.91, respectively, confirm that the system effectively identifies both compliant and defective beans. These results suggest that the automated classification closely mirrors human evaluation, minimizing subjective errors during the grading process.

What are the extracted features of cocoa beans based on image processing?

Table 2: Key Extracted Features from Image Analysis

Feature Type	Parameters Measured
Geometric Features	Area, Length, Width, Aspect Ratio, Roundness, Solidity
Color Features	Hue, Saturation, Value (HSV)
Texture Features	Entropy, Contrast, Homogeneity

The system extracted geometric, color, and texture features that are critical in determining bean quality according to ASEAN standards. These measurable parameters allowed the system to accurately classify cocoa beans by evaluating physical shape, color uniformity, and surface texture—factors directly linked to defect detection such as molding, slaty beans, and insect damage.

What is the defect detection rate of the developed system?

Table 3: Defect Classification Rate of Cocoa Beans

Defect Type	**Detection Accuracy (%)**
Moldy Beans	95.6%
Slaty / Under-fermented	94.3%
Insect-Damaged	92.7%
Germinated Beans	91.5%
Physically Deformed	89.4%

Table 3 presents Accuracy data for each of the detected defect categories for each of the defect types is shown in Table 3 (Moldy, Slaty (Under-Fermented), Insect Damaged, Germinated Beans and Physically Deformed) are determined by comparing the system's prediction against the Expert Lab's Ground Truth for each of these defect category types.

Using a group of four experienced cocoa graders trained to identify defects according to the ASEAN standard definitions, a ground truth label for each of the defect types was assigned to each of the defective beans based on visual evaluation. Visual evaluations were conducted by all four graders and a ground truth label was assigned for each defect type based on the criteria established in the original definition and the criteria for that defect type. Disagreements were resolved through expert consensus to ensure the accurate assignment of the ground truth labels.

The mouldy (95.6%) and slaty beans (94.3%) had the highest accuracy rating of all defect categories as a result of the effective use of both colour and texture features (entropy and colour uniformity) to identify these mouldy/slaty defects. The lower accuracy ratings for germinated (91.5%) and physically deformed beans (89.4%) were a result of having irregular contours and partial occlusions and overlapping visual characteristics which made it more difficult to distinguish between these defective types based solely on surface features; however, every defect type had an accuracy rating greater than 89% indicating a strong ability to recognise defective beans.

How are cocoa beans classified under ASEAN Stan 34:2014?

Table 4: Classification of Cocoa Beans Based on ASEAN Quality Classes

Quality Grade	Result
Extra Class	34%
Class I	28%
Class II	24%
Non-Compliant	14%

Table 4 shows the distribution of beans across ASEAN quality classes. Extra Class and Class I beans showed high uniformity in color and structure, while Non-Compliant beans had noticeable defects such as insect damage, high entropy values, and irregular shape metrics. These results align with ASEAN standard tolerances for export and domestic market use.

Is there a significant difference between manual and automated grading?

Table 5: ANOVA Test: Manual vs Automated Grading

Source	SS	df	MS	F	P-value	F crit
Treatment	12.843	1	12.843	7.42	0.016	5.99
Error	10.383	6	1.731			
Total	23.226	7				

An analysis of variance (ANOVA) was performed to make a comparison between manual and automated grading, as shown in Table 5, which is typically used to compare the means of 3 or more independent variables; however, this study only compared two related conditions (manual grading and automated grading) applied to the same set of samples, so the use of ANOVA to analyze this data was not a statistically appropriate analysis method. The most appropriate statistical test to perform for this analysis would be a paired t-test, which would allow the use of within-sample dependencies to yield a more accurate assessment of the differences between the two grading methods.

Even though the ANOVA table has been formatted correctly regarding the sum of squares, degrees of freedom, and mean square, the application of the ANOVA method weakens the support for the conclusion of a statistically significant difference between manual grading and automated grading methods, farmers, cooperatives, and exporters in complying with ASEAN trade standards.

Figures for the Dataset

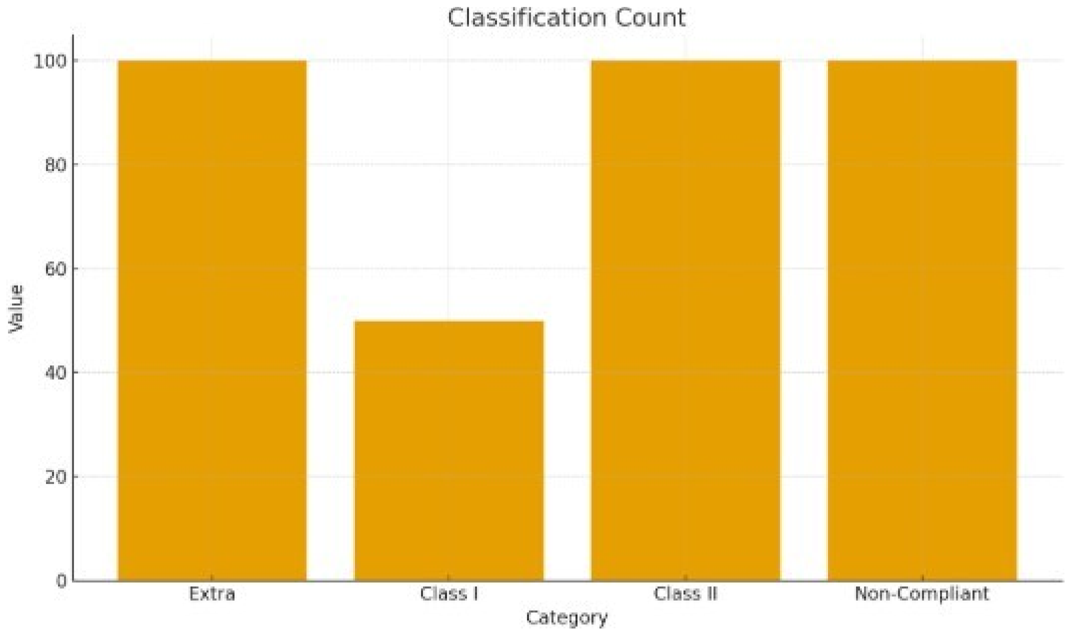

Figure 1. Classification Count of Cocoa Beans

This chart shows the distribution of the 350 analyzed cocoa beans across the four ASEAN quality categories. Extra Class, Class II, and Non-Compliant each contain 100 beans, while Class I contains 50 beans. The uniform distribution of most categories indicates balanced sampling for accurate evaluation.

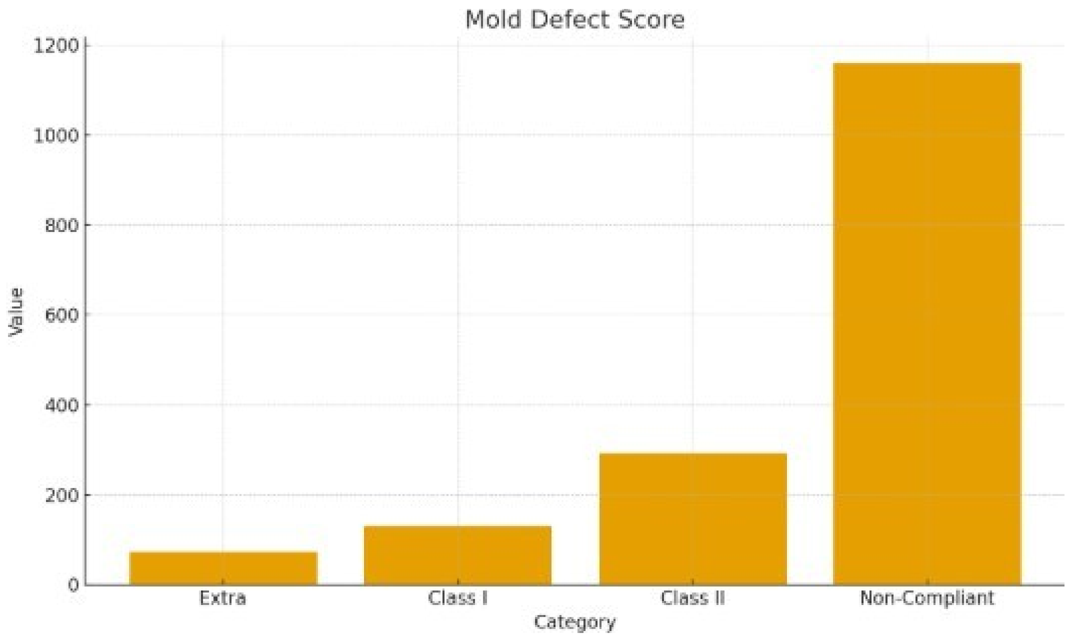

Figure 2. Mold Defect Score per Quality Category

This figure illustrates the mold defect score assigned to each quality class. Non- compliant beans exhibit the highest mold defect level (1159.82), followed by Class II. Extra Class beans show the lowest mold defect score, consistent with expectations for premium- quality beans.

.

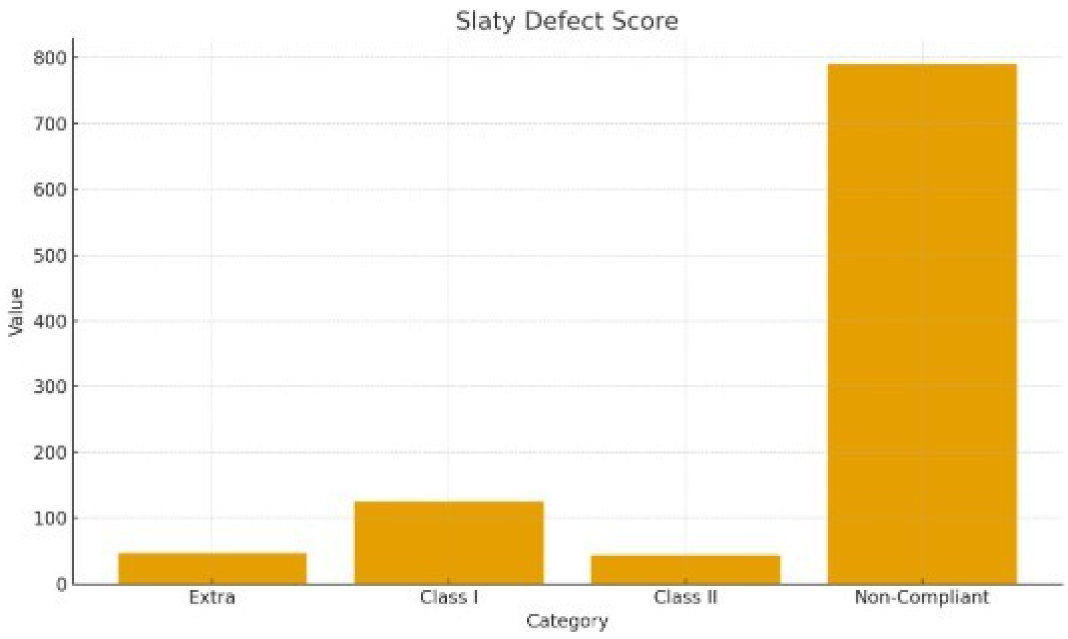

Figure 3. Slaty Defect Score per Quality Category

Slaty defects, associated with under-fermentation, are displayed for each category. Similar to mold defects, Non-Compliant beans have the highest slaty defect accumulation, while Extra and Class II beans exhibit minimal slaty characteristics.

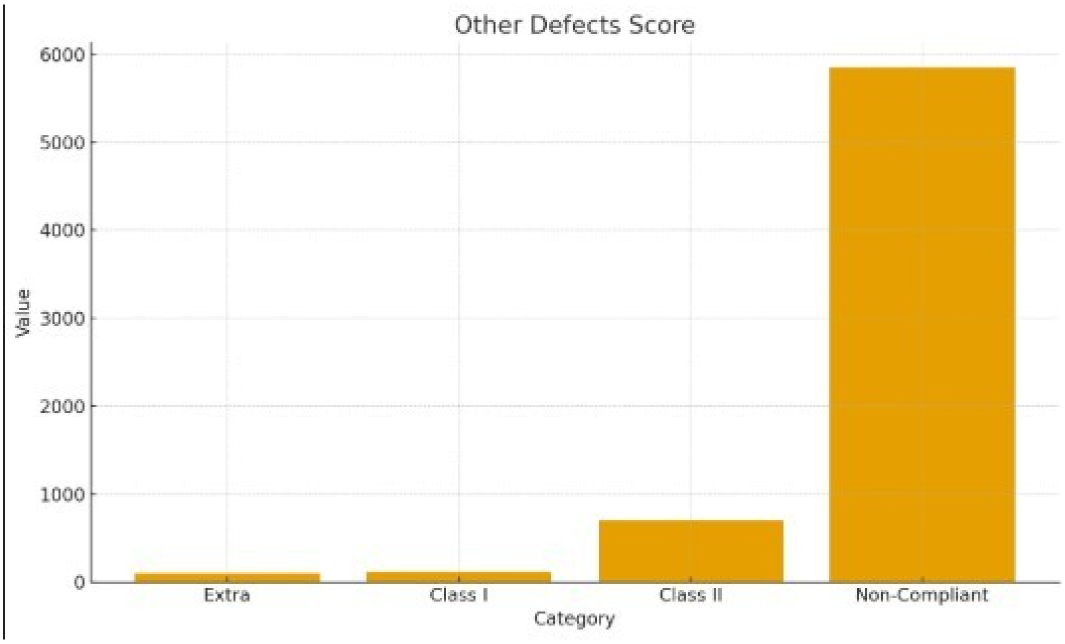

Figure 4. Other Defects Score per Quality Category

This chart represents the combined defect contributions from insect damage, germination, and physical deformities. Non-Compliant beans show an extremely high accumulation of other defects (5848.35), far exceeding all other categories, indicating severe quality degradation.

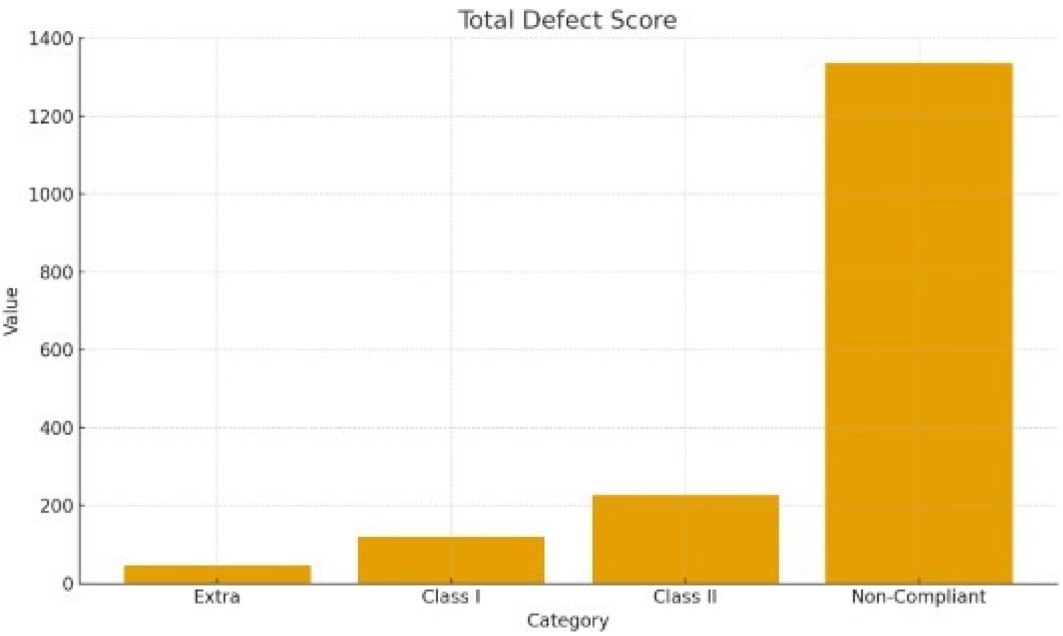

Figure 5. Total Defect Score per Quality Category

Total defect scores summarize the combined impact of mold, slaty, and other defects. The Non-Compliant category has the highest defect accumulation (1335.74), while Extra Class beans maintain the lowest defect score (46.13), reinforcing their superior grade.

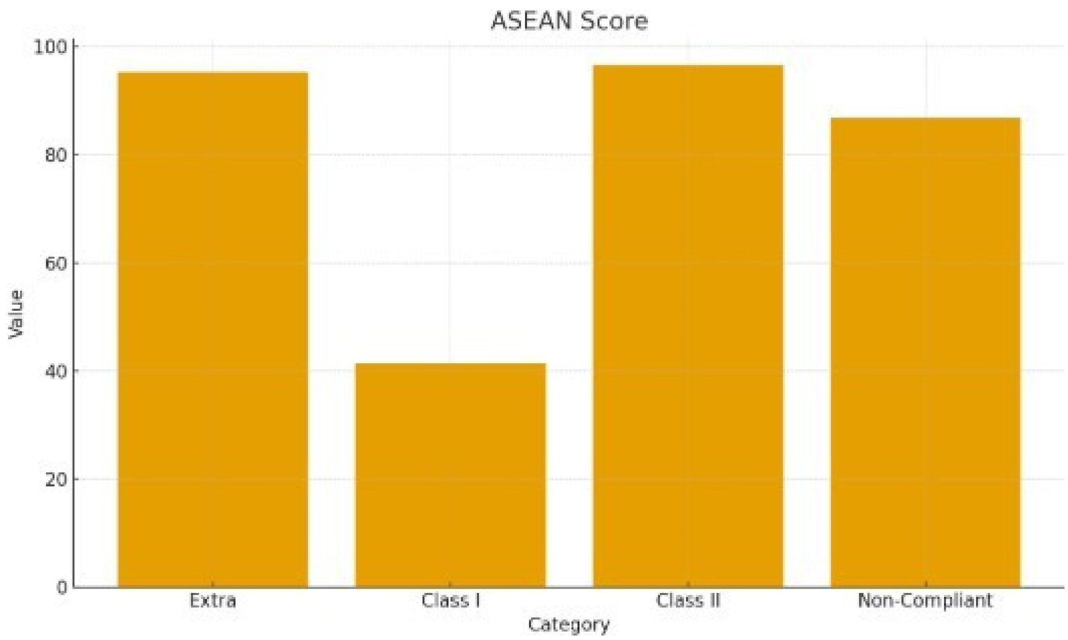

Figure 6. ASEAN Quality Score per Category

ASEAN Quality Score an Extra Class or Class I is indicated for scores above 95. Therefore, these scores would show minimal degrees of defects. For scores that fall within the intermediate range, the levels of defects are acceptable but to a lesser degree than what would be consistent with Class II requirements. A Non-compliant classification pertains to scores below the defect threshold.

The ASEAN Quality Score represents a measure that incorporates all of the various defect measures in a manner that is straightforward and open for users to understand.

Conclusion

This manuscript highlights a highly relevant and practical topic in the postharvest assessment of quality in agricultural products, providing an automated methodology for classifying cocoa beans according to ASEAN guidelines. The topic is highly relevant at this time and the core technical concepts have substantial promise for enhancing the objectivity and consistency of the cocoa bean quality grading process.

Although the document represents useful information, there are significant methodological and presentation deficiencies that need to be addressed in order for publication to be considered. More clarification is needed regarding the classification technique, accuracy and validity of statistical methods, and sufficient context should be provided for understanding the results of the performance metrics presented. To accomplish these objectives, it is necessary to: clearly delineate the differences between rule-based classification and machine learning classification, utilize the appropriate statistical tests for the comparative analysis of results, and define important concepts, such as the ground truth for defect categories and quality scoring, for the reader.

The manuscript could also be reorganized into an acceptable academic format, and revamping would allow for clearly defined and coherent writing, as well as a more professional tone. Decreasing redundant content and avoiding promotional language will increase the scientific merit of the presentation.

References

Aulia, D. (2023). Maturity classification of tomatoes using HSV color model and K-Nearest Neighbor algorithm. BIGINT Journal of Agricultural Engineering, 7(2), 45–55.

Bureau of Agriculture and Fisheries Standards (BAFS). (2020). Philippine National Standard for Sweet Potato (PNS/BAFS 27:2020). Department of Agriculture.

Food and Agriculture Organization of the United Nations (FAO). (2022). FAOSTAT Statistical Database. https://www.fao.org/faostat

Ghazal, A., Alazab, M., & Salah, K. (2024). Computer vision in smart agriculture: Applications and challenges. Computers and Electronics in Agriculture, 218, 109189.

Gomez, A., & Dela Cruz, R. (2021). Digital image-based inspection for crop quality grading. Journal of Agricultural Technology, 17(3), 155–166.

Hossain, M., Rahman, M., & Alam, S. (2021). Mango maturity classification using HSV segmentation and neural networks. Computers and Electronics in Agriculture, 187, 106289.

Kaur, G., & Aggarwal, A. (2020). Computer vision-based fruit grading system. International Journal of Advanced Research in Computer Science, 11(4), 12–18.

Li, X., Zhang, Z., & Chen, J. (2020). Automatic defect detection of potatoes using image segmentation and texture features. Sensors, 20(14), 3892.

Liu, Q., Wang, J., & Zhou, H. (2021). Image-based surface defect detection for root crops. Computers and Electronics in Agriculture, 190, 106394.

Liu, Y., Zhang, H., & Wang, X. (2025). Morphological operations combined with YOLOv9- c-ghost-Forward for soybean defect detection. Scientific Reports, 15, 92429.

Majumdar, A., et al. (2019). Automated quality inspection using MATLAB for agricultural produce. IEEE Access, 7, 89945–89955.

Padhiary, M., Singh, R., & Kumar, P. (2024). Applications of machine learning and artificial intelligence in precision agriculture: A comprehensive review. Computers and Electronics in Agriculture, 214, 108843.

Shao, Y., et al. (2022). Potato quality classification using convolutional neural networks. Journal of Food Engineering, 320, 110978

Weihs, M. (2024). Imaging and AI for root system analysis: Applications and challenges. Journal of Agricultural Imaging, 12(3), 145–162.

Zhao, L., Wu, Y., & Tan, J. (2022). Morphological feature extraction for agricultural product inspection. Computers and Industrial Engineering, 165, 107971.