Basic and Applied Computing for Digital Innovation (BACODING)
Volume 01, No. 02, August, 2026, pp. 80-88
Received: 08 August 2026; Revised: 21 August 2026; Accepted: 25 August 2026
https://bacoding.id/index.php/journal/
E-ISSN: 3163-5881
DOI:

Research Article

The Impact of Parameter Tuning on The Multi-Class Classification of Copra Quality

Jamie Eduardo C. Rosal[1], Juvy Amor M. Galindo[2]
[1,2] Bachelor of Science in Computer Engineering, Cor Jesu College, Philippines

* **Correspondence:** Email: Jamiewertalmighty@gmail.com

Abstract: This study aims to classify copra quality into four categories: Overcooked, Partially Dried, Perfectly Dried, and Sunog using the Support Vector Machine (SVM) algorithm. The performance of the SVM model with default parameters was compared to the model after parameter tuning to evaluate its effectiveness in handling multiclass classification. The results show that parameter tuning improved the model's performance on several classes. The Partially Dried class achieved the best performance with a precision of 82.0% and a recall of 91.0%, indicating that the model was able to recognize most samples in this class accurately. The Sunog class also showed improvement after tuning, with a precision of 64.5% and a recall of 74.0%. However, the model still struggled with minority classes. For the Perfectly Dried class, only 2 out of 45 samples were correctly classified. In the Overcooked class, all data were misclassified, resulting in 0% for both precision and recall. Overall, this study demonstrates that classification success is not only determined by SVM parameter optimization but also influenced by class imbalance and the similarity of features between copra quality classes. Therefore, in addition to parameter tuning, strategies to address data imbalance are needed to achieve more accurate and balanced classification across all classes.

Keywords: SVM, Copra Quality, Parameter Tuning, Class Imbalance

1. Introduction

Coconut is a vital agricultural commodity supporting the livelihoods of millions of farmers in the Philippines, with approximately 25% of the country's agricultural land dedicated to coconut cultivation. The harvested coconuts are primarily processed into copra, which is subject to specific quality standards in the Philippines. Typically, these quality grades are categorized into four levels: Perfectly Dried, Partially Dried, Overcooked, and Sunog.

Copra, the dried kernel of the coconut, remains the most common export commodity traded by farmers prior to further processing into coconut oil and other secondary derivatives. The commercial value of copra is strongly determined by its quality, which is evaluated based on color, surface texture, dryness level, and the presence of mold or physical damage. However, copra quality grading at both the farmer and collector levels is still performed manually, relying heavily on visual inspection and individual worker experience. This conventional approach is prone to inconsistent grading across evaluators, time-consuming when handling large volumes, and highly susceptible to worker fatigue and subjective judgment. Such inconsistencies ultimately risk financial losses for smallholders, as the selling price is directly tied to the assigned quality grade.

Advances in digital image processing offer promising opportunities to address these limitations. For instance, [1]-[4] implemented YOLOv8 to classify the dryness levels and quality of copra, achieving classification accuracies exceeding 86%. Similarly, [5] applied the InceptionV2 transfer learning model to classify copra dryness, achieving an accuracy of 90%, while [6] obtained an accuracy of 86% using deep learning techniques. Furthermore, [7] utilized a Convolutional Neural Network (CNN) to classify copra maturity levels. Although these studies successfully leveraged deep learning models, such approaches carry inherent limitations, including high computational resource requirements and the need for large training datasets [8] [9]. To overcome these constraints, this study adopts a traditional machine learning approach.

Classifying image data using machine learning models necessitates a feature extraction phase. For copra analysis, texture features are particularly crucial. A widely adopted technique for extracting textual information from images is the Gray Level Co-occurrence Matrix (GLCM), a second-order statistical method that calculates spatial pixel relationships across specified directions and distances. While raw GLCM matrices effectively capture image textures, they cannot be directly utilized for comparative texture analysis; thus, derivative statistical features must be extracted as inputs for classification algorithms. Features such as contrast, correlation, energy, and homogeneity derived from GLCM serve as key discriminative variables between quality classes.

The effectiveness of GLCM for texture-based agricultural and crop image classification has been well documented in prior studies, including applications on coffee beans [10][12], soybeans [13][14], and tomato quality evaluation [15]-[16], all yielding superior results. Therefore, this study performs GLCM feature extraction across directional angles of 0^0, 45^0, 90^0, and 135^0. Subsequently, a Support Vector Machine (SVM) model is employed to classify copra into four distinct quality grades: Perfectly Dried, Partially Dried, Overcooked, and Sunog.

2. Method and materials

2.1 Research steps

This study utilizes a publicly available dataset. The method employed is SVM, while the extracted feature is the Gray Level Co-occurrence Matrix (GLCM). The research steps are illustrated in Figure 1.

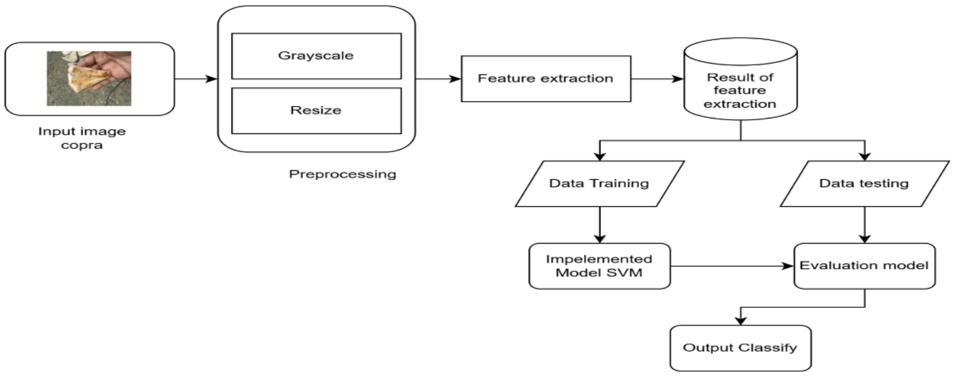

Figure 1. Copra quality classification steps

The diagram above illustrates the workflow of the copra quality classification model using GLCM and SVM. The following is an explanation of each stage:

1. Copra image: Copra images are captured as input data; the image format used is JPG.
2. Image preprocessing: images are converted to grayscale, resized to uniform dimensions, and filtered to reduce noise, thereby making the texture more clearly discernible.
3. GLCM (Gray Level Co-occurrence Matrix) feature extraction: Texture features such as contrast, correlation, energy, and homogeneity are calculated from the image's co-occurrence matrix. These features represent the surface texture of the copra. The Gray Level Co-occurrence Matrix (GLCM) is utilized in image processing to extract texture features. GLCM calculates probability values based on the neighbor relationship between two pixels at specific distances and angular orientations. It is a matrix that describes the frequency of occurrence of pixel pairs with specific intensities at defined distances and directions within an image [17]. Values for GLCM parameters [18]—specifically correlation, Angular Second Moment (ASM), dissimilarity, homogeneity, contrast, and energy—are derived from angles of 0°, 45°, 90°, and 135°. Each value is obtained using Equations (1–6).

$$\text{Correlation} = \frac{\sum_i \sum_j (ij)\, P\,(i,j) \;-\; \mu_i\, \mu_j}{\sigma_i\, \sigma_j} \tag{1}$$

$$\text{ASM} = \sum_i \sum_j P\,(i,j\,)^2 \tag{2}$$

$$\text{Dissimalirity} = \sum_i \sum_j |i - j|\,.\,P\,(i,j) \tag{3}$$

$$\text{Homogenity} = \frac{\sum_i \sum_j P\,(i,j)}{1 + |i - j|} \tag{4}$$

$$\text{Contrast} = \sum_i \sum_j |i - j|^2\, P\,(i,j) \tag{5}$$

$$\text{Energy} = \sqrt{ASM} \tag{6}$$

4. SVM Classification: GLCM features serve as input vectors for a trained Support Vector Machine model to separate quality classes based on an optimal hyperplane.
5. Output: The final result consists of copra quality labels: Perfectly Dried, Partially Dried, Overcooked, and Sunog.

2.2 SVM Algorithm

The Support Vector Machine (SVM) is a classification algorithm that falls under the category of supervised learning. SVM operates by finding an optimal separating line (hyperplane) to separate two groups of data. Developed by Vapnik in the 1990s, this algorithm has demonstrated excellent performance across various classification tasks, including credit scoring. The SVM model functions by identifying a line (hyperplane) that separates two data groups. The hyperplane represents the optimal dividing line between two classes. Finding the hyperplane involves determining the hyperplane margin and identifying maximum points. The margin is defined as the distance between the nearest data points of two different classes—points known as support vectors. The hyperplane serves as the decision boundary separating two classes within the feature space. In an n-dimensional space, the hyperplane is an (n-1)-dimensional subspace. The linear SVM classification hyperplane is denoted as:

$$w \cdot x + b = 0 \tag{7}$$

Where w is the weight vector perpendicular to the hyperplane, x is the input feature vector, and b is the bias that determines the hyperplane's position relative to the coordinate origin.

3. Result And Discussion

3.1. Dataset

The data used in this study consists of the COPRAA dataset, comprising 2,248 images. We categorized this data into four copra condition classes. We observed a significant disparity in the number of samples across classes: the "Partially Dried" class was the most dominant with 1,500 images (66.7%), followed by "Sunog" with 481 images (21.4%) and "Perfectly Dried" with 225 images (10.0%), while the "Overcooked" class had the fewest samples, with only 42 images (1.9%). This data distribution imbalance is a crucial factor to consider when evaluating the model's performance on specific classes.

Table 1. Classes and Number of Images

No	Class	Number of images	Percentage
1	Partially Dried	1.500	66,7%
2	Sunog	481	21,4%
3	Perfectly Dried	225	10,0%
4	Overcooked	42	1,9%
Total		2.248	100%

3.2. Result of classify

Regarding copra quality classification, the study involved testing using SVM with default settings and hyperparameter tuning. The results for the default SVM configuration are shown in Figure 2.

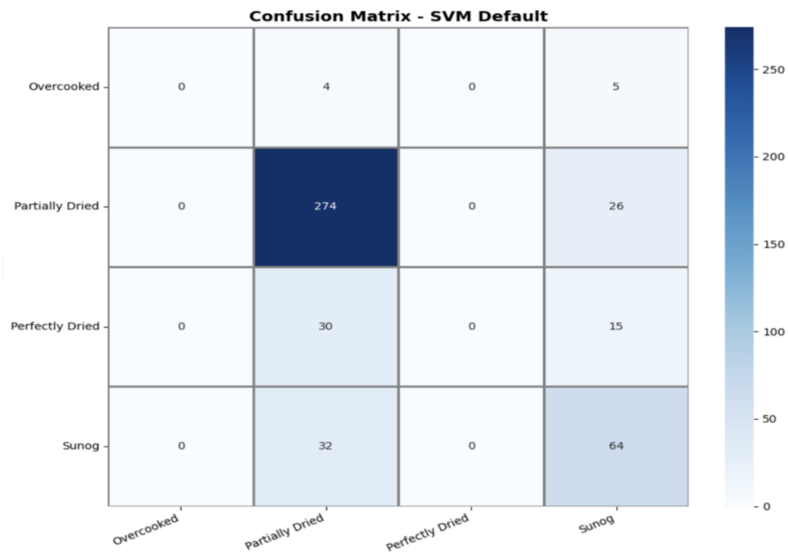

Figure 2. SVM classification results without parameter tuning

Based on testing results using the Support Vector Machine (SVM) algorithm with default parameters, varying classification success rates were observed across the different copra quality classes. The results indicate that the model recognizes the "Partially Dried" class very effectively but struggles to distinguish between the "Overcooked" and "Perfectly Dried" classes.

For the "Partially Dried" class, the model correctly classified 274 out of 300 data points, while 26 were misclassified. The high success rate in this class suggests that the data characteristics are sufficiently consistent for the SVM model to learn them well. It is highly likely that the features used exhibit clear patterns, allowing the separating boundary (hyperplane) established by the SVM to effectively distinguish this class from others.

In contrast to the "Partially Dried" class, results for the "Sunog" class showed reasonably good performance, though significant misclassification errors persisted. Out of 96 data points, the model correctly classified 64, while 32 were misclassified. This indicates that the "Sunog" class shares characteristics with other classes, causing some data points to lie near the decision boundary. Consequently, the model struggled to determine the correct class for certain samples. Meanwhile, the lowest performance was observed in the "Overcooked" and "Perfectly Dried" classes; the model failed to correctly classify a single data point in either category. All 9 "Overcooked" samples and all 45 "Perfectly Dried" samples were misclassified into other classes. This situation suggests that the default SVM parameters were unable to establish suitable separating boundaries for these two classes. Furthermore, the very small number of samples in the "Overcooked" class compared to others caused the model to lean toward the patterns of the majority classes. This class imbalance can diminish the model's ability to recognize minority classes.

Beyond data imbalance, there may be similarities in feature characteristics among the "Overcooked," "Perfectly Dried," and other classes. If the feature distributions overlap, an SVM using default parameters will struggle to identify a hyperplane capable of optimally separating each class. The use of default parameters implies that key parameters—such as the kernel, C, and gamma—have not been tailored to the data's characteristics, preventing the model from achieving optimal performance.
Overall, the classification results indicate that the default SVM successfully identified the class with dominant patterns (Partially Dried) and yielded reasonably good results for the Sunog class. However, the model remained ineffective at identifying the Overcooked and Perfectly Dried classes. Consequently, model optimization is required; this entails tuning SVM parameters, balancing the number of samples across classes, and selecting or extracting more representative features. These measures are expected to enhance the model's ability to accurately distinguish between all copra quality classes and achieve more balanced classification performance.

3.3. Result of parameter tuning implementation

During the model development phase, we did not simply rely on the standard settings for the Support Vector Machine (SVM) algorithm. We performed hyperparameter tuning to identify the optimal parameter combination, thereby enhancing the system's accuracy in classifying copra quality. The grid search algorithm was employed to determine the best parameters. The results identified the optimal parameters as C=100, gamma='scale', and kernel='rbf'. These results are presented in the form of a confusion matrix in Figure 3.

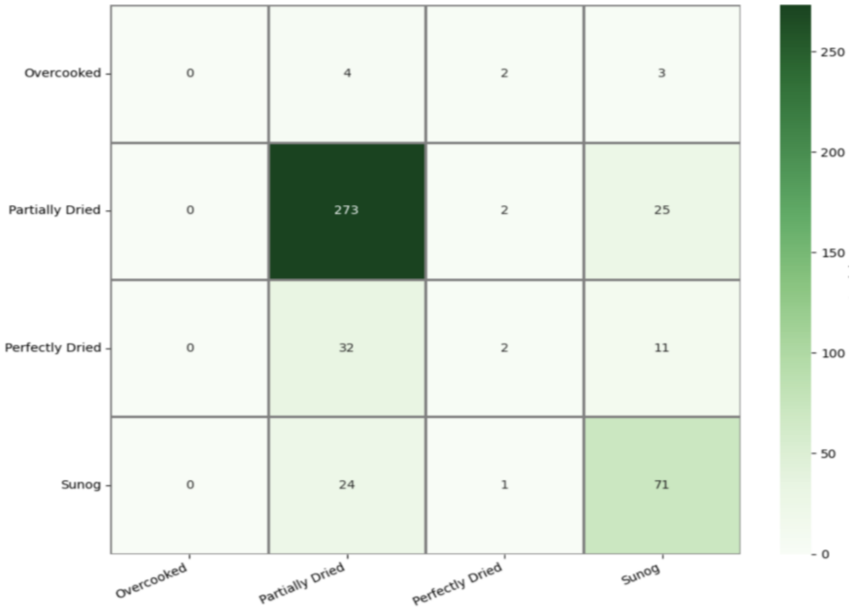

Figure 3. results of confusion matrix parameter tuning

Testing results using the Support Vector Machine (SVM) algorithm—following parameter tuning—showed performance improvements in several classes compared to using default parameters. Parameter tuning aims to identify the optimal parameter combination, enabling the model to establish a more effective separating boundary (hyperplane) between classes. However, the results indicate that performance improvements were not uniform across all copra quality classes.

The "Partially Dried" class demonstrated the best performance. Out of 300 data points, the model correctly classified 273, with only 27 misclassifications. A precision of 82.0% indicates that the majority of predictions for the "Partially Dried" class were accurate, while a recall of 91.0% shows that the model successfully identified nearly all data points belonging to this class. These results suggest that the visual characteristics and features associated with the "Partially Dried" class exhibit consistent patterns, making them easily learnable for the SVM.

The "Sunog" class also showed reasonably good performance. Out of 96 data points, 71 were correctly classified, while 25 were misclassified. A precision of 64.5% and a recall of 74.0% indicate that the model successfully recognized most "Sunog" samples, although some data points were incorrectly predicted as "Partially Dried." This suggests a similarity in characteristics between the two classes, causing some data points to lie near the decision boundary established by the SVM.

Conversely, the "Perfectly Dried" class showed poor performance. Out of 45 data points, only two were correctly classified, while the remaining 43 were misclassified. A precision of 28.6% and a recall of 4.4% indicate that the model was largely unable to recognize this class. According to the confusion matrix, the majority of "Perfectly Dried" data points were predicted as "Partially Dried" (32 points), with the remainder predicted as "Sunog" (11 points). This situation indicates that the "Perfectly Dried" class shares significant characteristics with the other two classes, making it difficult for the model's decision boundaries to effectively distinguish between them.

The "Overcooked" class remains the most difficult to identify. None of the nine test samples were correctly classified, resulting in 0% precision and recall. Instead, all "Overcooked" samples were misclassified as other classes: four as "Partially Dried," two as "Perfectly Dried," and three as "Sunog." These results demonstrate that the model has not effectively learned the characteristics of the "Overcooked" class. A primary reason for this is the scarcity of "Overcooked" data compared to other classes, causing the model to focus more on learning patterns from the majority classes. Additionally, similarities in color or texture characteristics across classes hinder the model's ability to establish clear classification boundaries.

Comparing the results before and after parameter tuning reveals an improvement for the "Sunog" class—with correct classifications rising from 64 to 71—indicating enhanced recognition capability for this class. However, performance for the "Partially Dried" class remained relatively stable, showing a slight decline from 274 to 273 correct classifications, while the "Perfectly Dried" class saw a marginal increase from zero to two correct classifications. Meanwhile, the "Overcooked" class showed no improvement, as all samples remained misclassified.

Overall, these results indicate that parameter tuning successfully improved model performance for certain classes—particularly "Sunog" and, to a lesser extent, "Perfectly Dried"—but failed to resolve the core issues affecting minority classes, specifically "Overcooked." This demonstrates that model performance is influenced not only by the choice of SVM parameters but also by data distribution, sample sizes per class, and the degree of feature similarity between classes. Consequently, future research should consider implementing data balancing techniques (such as oversampling or class weighting), employing more discriminative feature extraction methods, or combining SVM with other optimization techniques to achieve more balanced and accurate recognition across all copra quality classes.

4. Conclusion

Research results indicate that applying the Support Vector Machine (SVM) algorithm to classify copra quality allows for the differentiation of four quality classes: Overcooked, Partially Dried, Perfectly Dried, and Sunog. Following parameter tuning, the model demonstrated improved performance in certain classes compared to using default parameters, although this improvement was not uniform across all classes.

The Partially Dried class exhibited the best performance, achieving a precision of 82.0% and a recall of 91.0%, demonstrating the model's ability to effectively recognize the majority of samples in this category. The Sunog class also showed performance gains, with precision and recall values of 64.5% and 74.0%, respectively, indicating an enhanced ability to identify this class following the tuning process.

However, the model continued to struggle with classifying the Perfectly Dried and, particularly, the Overcooked classes. For the Perfectly Dried class, only 2 out of 45 data points were correctly identified, while all data points in the Overcooked class were misclassified, resulting in precision and recall values of 0%. These results suggest that parameter tuning alone is insufficient to improve the model's ability to recognize minority classes.

Overall, this study demonstrates that classification success is influenced not only by the SVM parameter optimization process but also by class imbalance and the similarity of feature characteristics across copra quality grades. Consequently, while parameter tuning successfully improved performance for some classes, further refinements are required for the model to classify all categories with greater accuracy and balance.

5. Reference

[1]. Joshua Joemer D. Bagunu, et al., (2025) "Classification of the Stage of Dryness of Copra via Computer Vision Using Image Processing Techniques" Proceding International Conference on Computer Systems and Technologies (CompSysTech). 136 – 140. Ruse; Bulgaria Doi: 10.1109/CompSysTech65493.2025.11136960

[2]. P. Kanaga Priya, et al., (2025) "Real-Time Coconut Copra Classification Using YOLOv8 and ESP32 CAM". Proceding International Conference on Emerging Information Technology and Engineering Solutions (EITES). 885 – 889. Pune, India. Doi: 10.1109/EITES66543.2025.00011.

[3]. Usman, U. (2026). Improving the accuracy of classification and detection of coconut fruit quality based on the integration of the EfficientNet-YOLO hybrid model with spatial-coordinate attention. International Journal of Intelligent Engineering and Systems, 19(1), 164–179. https://doi.org/10.22266/ijies2026.0131.11

[4]. Espiritu J. B. B, et al. (2025) " Development of a Computer Vision Application for Copra Meat Cooking Detection and Classification Using YOLOv8 Architectur" *Proceedings of the 2025 14th International Conference on Software and Computer Applications (ICSCA),* ACM. https://doi.org/10.1145/3731806.3731838.

[5]. Juvy Amor M. Galindo, et al. (2023) " Detection and Classification of Copra Meat Dryness Using Faster Region-Based Convolutional Neural Network with Inception v2 Architecture" Proceding 2023 IEEE International Conference on Artificial Intelligence in Engineering and Technology (IICAIET). 225 – 230. Kota Kinabalu, Malaysia. Doi: 10.1109/IICAIET59451.2023.10291856

[6]. Rufo I. Marasigan Jr, Alvin Sarraga Alon, Mon Arjay F. Malbog, and Joshua S. Gulmatico. (2020)., "Copra Meat Classification using Convolutional Neural Network". International Journal of Emerging Trends in Engineering Research. Vol 8 (2). Pp, 435 – 439. Doi: 10.30534/ijeter/2020/30822020

[7]. Luther Alexander Latumakulita, Frangky J. Paat, Glenn Budiman, Dedie Tooy, Mayko Edison Koibur, Noorul Islam, (2024) "A Convolutional Neural Network-based Intelligence System for the Identification of Copra Maturity Levels" vol 8 (3). Pp, 1381 – 1389. Doi: 10.62527/joiv.8.3.2574

[8]. K. Sri Poojitha. K., et al., "Automated Copra Grading System Using AI" *International Journal of Research in Engineering, Science and Management*. Vol 3 (4), pp, 230 – 234.

[9]. Gabi, C. A. G., et al. (2025). Quality assessment of copra using deep learning and image processing techniques. *Tra Vinh University Journal of Science, 15*(1). https://doi.org/10.35382/tvujs.15.1.2025.100

[10]. A. Septiarini, H. Hamdani, A. E. Burhandeny, D. Nurcahyono, and S. E. Priyatna, "Image analysis for classifying coffee bean quality using a multi-feature and machine learning approach," *IAES International Journal of Artificial Intelligence*, vol. 13, no. 4, pp. 4241–4248, 2024, doi: 10.11591/ijai.v13.i4.pp4241-4248.

[11]. R. E. G. Rivas, P. L. L. Bertarini, and H. Fernandes, "Automated Coffee Roast Level Classification Using Machine Learning and Deep Learning Models," *J. Food Sci.*, vol. 90, no. 9, pp. 1–11, 2025, doi: 10.1111/1750-3841.70532.

[12]. C. Kuruppuarachchi, M. Sher, M. Roman, M. Maimaitijiang, A. Zahid, and K. S. P. Amaratunga, "Grading of Specialty-Grade Coffea arabica Beans Using Digital Imaging and Machine Learning," *Food Anal. Methods*, vol. 19, no. 1, 2026, doi: 10.1007/s12161-025-02961-1.

[13]. G. da Silva André, P. C. Coradi, L. P. R. Teodoro, and P. E. Teodoro, "Predicting the quality of soybean seeds stored in different environments and packaging using machine learning," *Sci. Rep.*, vol. 12, no. 1, pp. 1–13, 2022, doi: 10.1038/s41598-022-12863-5.

[14]. S. Jitanan and P. Chimlek, "Quality grading of soybean seeds using image analysis," *International Journal of Electrical and Computer Engineering*, vol. 9, no. 5, pp. 3495–3503, 2019, doi: 10.11591/ijece.v9i5.pp3495-3503.

[15]. S. A. Shehab and M. Abdelfattah, "Classifying the safety and quality of tomatoes based on genetic algorithms integrated with Gray Wolf optimizer," *Journal of Food Measurement and Characterization*, 2025, doi: 10.1007/s11694-025-03695-8.

[16]. K. Goyal, P. Kumar, and K. Verma, "Tomato ripeness and shelf-life prediction system using machine learning," *Journal of Food Measurement and Characterization*, vol. 18, no. 4, pp. 2715–2730, 2024, doi: 10.1007/s11694-023-02349-x.

[17]. R. Chityala and S. Pudipeddi, Image Processing and Acquisition using Python. CRC Press, 2021. [Online]. Available: https://lccn.loc.gov/2020015556

[18]. W. Burger and M. J. Burge, *Digital Image Processing. Texts in Computer Science*. 2016.