Machine Learning-based Prototype to Improve the Classification of Olive Status in the Agri-food Industry

Oscar Mendoza-Escudero[1], Sebastián Sarmiento-Curay[1], José Velásquez-Costa[1], Javier Torres-Zavala[1] and Baldomero Mendez-Pallares[2]*

[1]*Faculty of Engineering, Universidad Peruana de Ciencias Aplicadas, Lima, Peru*
[2]*Faculty of Engineering, Universidad Piloto de Colombia, Colombia*
**pcinjosv@upc.edu.pe*

Keywords: Machine Learning, Raspberry Pi 5, Agri-food industry

Abstract

The manual classification of olives in the agri-food industry is a slow and costly process with high variability, affecting product quality and efficiency. This study presented an automated sorting system integrating machine vision, image processing, and real-time control to improve classification accuracy and speed. The prototype included a conveyor belt controlled by a Raspberry Pi 5, an ESP32-CAM with an OV2640 camera for image capture, a servomotor for defective olive ejection, and an LCD for monitoring parameters. The model was trained using 300 olive images expanded to 1800 samples through data augmentation techniques such as rotations, brightness variations, and translations, improving supervised learning robustness and defect detection reliability. The synchronized ejection mechanism automatically removed non-compliant olives through Raspberry Pi GPIO control. Experimental results showed an average processing time of 10 ms per olive. The system achieved 99.17% precision, 98.83% sensitivity, 99.05% accuracy, and an F1-Score of 89.99%, demonstrating strong classification performance. This proposal optimizes post-harvest processes and provides a scalable solution for agricultural classification systems applicable to different crops.

1 Introduction

In the agri-food industry, quality control remains a critical challenge, particularly in the classification of products such as olives, where manual sorting introduces subjectivity that generates economic losses that negatively affect competitiveness. At an international level, several studies have demonstrated that automating this process leads to significant improvements. For instance, the application of artificial intelligence algorithms supported by RGB images enables simultaneous and proactive fruit selection based on oil-quality-related characteristics [1], while other approaches employ hyperspectral imaging to identify olives with high precision attributes [2].

Recent research has focused on lightweight deep learning architectures for olive sorting, showing notable improvements in accuracy and processing speed compared to conventional methods [3]. Additionally, RGB-based classification prototypes have demonstrated the feasibility of automated fruit selection by linking visual analysis with chemical and sensory quality parameters, helping overcome system limitations and improve competitiveness [4].

Early fruit classification systems in agribusiness relied on conventional RGB imaging, achieving limited accuracy, particularly for black olives, which highlighted the constraints of image-only approaches [5]. To overcome these limitations, hybrid methods integrating sensors such as NIR with partial least squares regression (PLSR) achieved accuracies between 99.47% and 99.62% [6]. More recently, machine learning and deep learning techniques have demonstrated superior performance, reaching 99.6% accuracy with YOLOv8 for fruit sorting and up to 99.8% in olive detection, 93.5% in ripeness classification, and 78% in defect detection using Mask R-CNN-based systems [7].

2. Literature Review

Regarding the literature review, A structured literature review was conducted using the Web of Science and Mendeley databases, focusing on high-quality Q1 and Q2 publications from the last three years. Keywords related to computer vision algorithms, optical sensors, software, and hardware were used to identify relevant studies addressing efficiency limitations in post-harvest olive classification system.

2.1 Machine vision algorithms

Currently, computer **vision** algorithms represent one of the most relevant advances in the automation of sorting processes in agribusiness, as they enable the transformation of traditional, manual and subjective activities into digitized, reproducible, and scalable procedures. In this context, Ortenzi et al. demonstrated that rapid artificial vision based methods allow the objective determination of olive ripeness levels,[8].

2.2 Optical sensors

Advances in algorithms have been complemented by innovations in optical sensors, which serve as the main data source for computer vision systems. Conventional RGB cameras have been widely used to correlate fruit surface color with internal quality parameters, showing strong agreement with laboratory analyses [9]. However, the incorporation of advanced technologies such as NIR and hyperspectral sensors has expanded analytical capabilities by enabling the detection of non-visible attributes, including chemical composition and contaminants [10], [11].

2.3 Hardware

For the inspection of small fruits on conveyor belts, image quality depends primarily on controlled lighting and synchronization between image acquisition and conveyor motion, rather than on increasing camera resolution without proper optical design [12]. A systematic review published in Sensors reports that RGB cameras are used in 84% of inspection systems, emphasizing the relevance of lighting solutions such as ring, dome, or backlighting [12].

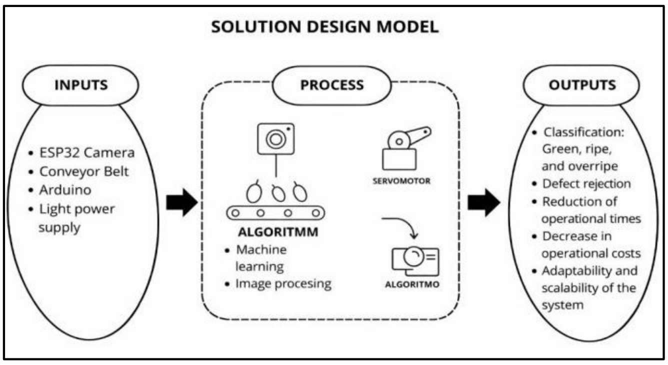


Fig. 1 Solution Desing Model

3. Methodology

3.1 Foundation

The proposed solution employs a three-layer architecture integrating computer vision, machine learning, and embedded control for automated olive sorting. As shown in Fig. 1, the system comprises optical acquisition, intelligent processing, and mechanical ejection control, ensuring a continuous and synchronized identification process.

The intelligent processing layer, implemented using the OpenCV library, extracts morphological and chromatic features such as texture, size, and color to train a model that classifies olives into ripe and overripe categories. This algorithmic approach enables objective and reproducible evaluation while reducing human error in **operations [13].**

3.2 Proposed model

The proposed model is an automated olive classification system based on computer vision using low-cost hardware and open-source software. A Python–OpenCV algorithm processes images captured by an ESP32 camera on a conveyor system to classify olives as ripe or overripe. Using a dataset expanded from 300 to 1,800 images, the system achieved an average accuracy of 96.8% with a processing time of 35–40 ms per fruit.

3.3 Performance metrics

To evaluate the effectiveness of the proposed model, performance was quantified using confusion matrix metrics: True Positives (TP), False Positives (FP), False Negatives (FN), and True Negatives (TN). TP and TN represent correct classifications of positive and negative cases, respectively, while FP corresponds to negative cases incorrectly classified as positive, and FN to positive cases not detected by the system.

According to [24], Precision measures the ability of the model to correctly identify positive instances by minimizing FP, while Sensitivity evaluates its capacity to detect all positive cases by reducing FN.

Those following metrics are defined as follows:
Precision: Precision evaluates the model's ability to correctly recognize positive instances, minimizing the occurrence of false positives.

- $$Precision = \frac{TP}{TP+FP} \quad (1)$$

Recall: Evaluate how effective the model is at detecting all positive occurrences, thereby reducing the number of false negatives

- $$Recall = \frac{TP}{TP+FN} \quad (2)$$

Accuracy: The percentage of correctly classified cases is known as accuracy and represents a fundamental indicator for measuring model performance.

- $$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (3)$$

F1-Score: The F1-Score provides a single value that allows you to evaluate a model's performance by balancing accuracy and sensitivity.

- $$F1 - Score = \frac{2*Precision * Recall}{Precision+Recall} \quad (4)$$

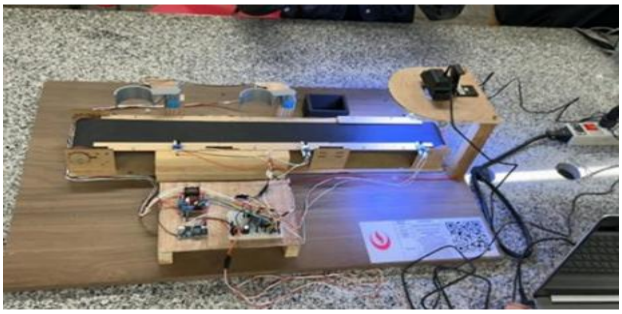
Together, these indicators will facilitate an accurate assessment of the performance of the proposed model, strengthening internal control mechanisms on quality and continuous improvement.

4. Results

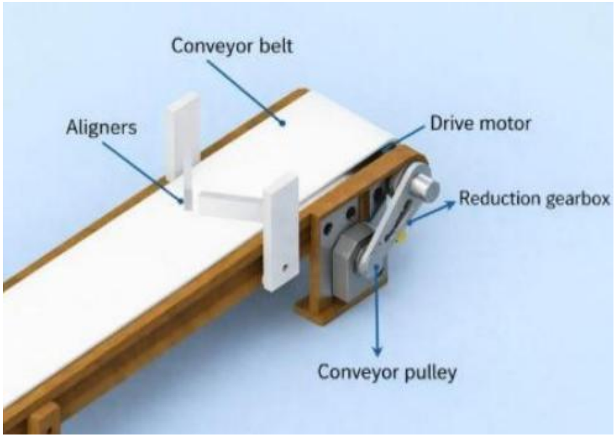
To validate the research, the accuracy, consistency, and verifiability of the results are ensured through the correct application of the proposed improvement model. The validation process describes the operational context, the implementation phases, and the updating of indicator values used to assess model performance.

4.1 Prototype development

A single-phase automated system for olive classification based on preservation state was developed and validated through prototype implementation. A scaled prototype incorporating infrared (IR) sensors was constructed and experimentally tested under controlled conditions to evaluate classification accuracy, system performance, and electronic response efficiency should be embedded within the text at the appropriate point. A maximum of four subfigures is allowed per figure.



a



b

Fig. 2 (a) Pneumatic system prototype (b) Proposed pre-alignment mechanism

4.2 Model indicators

Prior to experimental validation, the proposed model was compared with classification algorithms reported in the literature to identify the approach that provides the best balance between precision and sensitivity.

Table 1 Default Parameters

Model	Precision	Sensibility	Exactitude	F1-Score
K-Nearest Neighbors	86.05%	93.57%	92.89%	89.65%
YOLOv8 Nano	83.85	93.47%	91.88%	88.31%
YOLOv8 Smal	85.06%	94.64%	92.79%	89.49%
YOLOv5 Smal	97.70%	96.10%	-	96.89%

Based on the obtained results, the proposed model achieved the highest performance, followed by YOLOv5-small. However, YOLOv5-small was discarded due to its higher response time (41.2 ms) compared to YOLOv8-small, which operates within a 10–12 ms range. Therefore, a comparative evaluation was based on optimized training epochs. The proposed model produced a confusion matrix with TP = 147, FP = 1, FN = 2, and TN = 150 from 300 experimental samples [21]. These results were used to compute performance metrics for the comparison of the evaluated models.

Table 2 Optimal Outcomes for YOLOnanov8

Epochs	Mean Precision	Mean Recall	Mean Accuracy	Mean F1-Score
25	91.39%	96.50%	95.10%	93.88%
50	94.65%	98.37%	97.61%	96.47%
75	95.86%	97.20%	97.66%	96.53%
100	96.33%	97.20%	97.81%	96.76%

Table 3 Optimal Outcomes for K-Nearest Neighbours

Epochs	Mean Precision	Mean Recall	Mean Accuracy	Mean F1-Score
25	92.45%	97.56%	96.30%	94.94%
50	95.53%	98.95%	97.90%	97.21%
75	97.56%	98.84%	98.46%	98.20%
100	99.17%	98.83%	99.05%	89.99%

As shown in Tables II and III, the proposed model is compared with YOLOv8-nano, with both models trained and optimized over multiple epochs to ensure convergence and result stability. Overall, the developed system outperformed YOLOv8-nano across all evaluated metrics [22]. The proposed model achieved a Precision of 99.17%, exceeding YOLOv8-nano by 2.84%, and a Recall of 98.83%, representing a 1.63% improvement. In terms of Accuracy, the system reached 99.05%, compared to 97.81% for YOLOv8-nano, indicating greater classification consistency. Finally, the F1-score reached 89.99%, surpassing YOLOv8-nano by 3.23%, confirming the overall effectiveness of the proposed approach.

The olive classification system was evaluated under controlled conditions simulating a post-harvest sorting line, processing 300 olives equally distributed between ripe and overripe classes. The system achieved a Precision of 99.17%, Sensitivity of 98.83%, Accuracy of 99.05%, and an F1-score of 98.99%, demonstrating robust performance under chromatic and morphological variability and outperforming conventional RGB-based systems [23].

Performance tests reported an average processing time of 10 ms per image (σ = 1.4 ms), confirming the feasibility of real-time inference using the ESP32-CAM in low-cost environments [24]. Maximum power consumption during active inference was 1.8 W, lower than Raspberry Pi 4 and Jetson Nano–based systems, in line with previous findings [25]. Compared to YOLOv8 and MobileNet-V3 models, which report accuracy between 95% and 98% but require higher computational resources [26], the proposed system prioritizes efficiency and accessibility [27].

5. Conclusion

The OpenCV-based model developed using supervised learning achieved an average response time of 10 ms, ensuring photometric stability and reliable real-time processing. The integration of RGB segmentation, synchronized control, and stabilized lighting enabled consistent and reproducible ripeness detection. Although deep learning models such as YOLOv8 can achieve higher accuracy, the results demonstrate that classical machine learning approaches implemented with OpenCV provide competitive performance with significantly lower computational requirements. This enables the deployment of efficient, sustainable, and replicable agricultural classification systems, particularly in resource-constrained environments.

6. Acknowledgements

To the Dirección de Investigación of the Universidad Peruana de Ciencias Aplicadas for the support provided in the completion of this research work through the incentive UPC-EXPOST-2026-1

7. References

[1]Figorilli,S.,Violino,S.,Moscovini,L.,Ortenzi,L.,Salvucci,G.,Vasta,S.,Tocci,F.,Costa,C., & Toscano,P., "Olive Fruit Selection through AI Algorithms and RGB Imaging," Foods, vol. 11, no. 6, p. 906, 2022. https://doi.org/10.3390/foods11213391

[2] Domínguez-Cid S.Larios, D., Barbancho,J., Molina,F., Guerra, J., & León, C., "Identification of Olives Using In-Field Hyperspectral Imaging with Lightweight Models," Sensors, vol. 24, no. 3, p. 982, 2024. https://doi.org/10.3390/s24051370

[3] Saedi, S., Rezai, M., & Khosravi, H., "Dual-path lightweight CNN for sorting and grading table olives," Computers Industrial Engineering, https://doi.org/10.1016/j.postharvbio.2024.113054

[4] Violino, S., Moscovini, L., Costa, C., Del Re, P., Giansante, L., Troscano, P., Tocci, F., Vasta, S., Manganiello,R.., Ortenzi, L., & Pallottino, F., "Superior EVOO Quality Production: An RGB Sorting Prototype for Olive Classification," Foods, vol. 11, no. 23, p. 3798, 2022. https://doi.org/10.3390/foods11182917

[5] Lázaro, M., Madueño, A., Pascual, A., Molina, J., Ruiz, A., Madueño, J., Justicia, M., & Baena, M., "Deep learning in olive pitting machines by computer vision," Computers and Electronics in Agriculture, vol. 175, p. 105606, 2020. https://doi.org/10.1016/j.compag.2020.105304

[6] Wold, J.,Sanden,K., Skaret,J.,Carlehog,M. ,Tjaland, M., & Harsen,A., "Non contact Non-contact interactance NIR spectroscopy for estimating TSS and sensory sweetness in conveyor-belt transported cherry tomatoes (Lycopersicon esculentum 'Piccolo'),2025. https://doi.org/10.1016/j.saa.2025.125962

[7] Moya,V. ,Guerra, M.,Pazmiño,K.,Abedrabbo,F.,Chicaiza, F. & Pozo,D. , "Tomato classification with YOLOv8:Enhancing automated sorting and quality assessment", Smart Agricultural Technology. https://doi.org/10.1016/j.atech.2025.101221

[8] Ortenzi,L., Figorilli,S., Costa,C., Pallottino, F.,Violino, S., Pagano,M., Imperi,G., Manganiello,R., Lanza,B., & Antonucci,F., "A machine vision rapid method to determine the ripeness degree of olive lots," Sensors, vol. 21, no. 9, p. 2940, 2021. https://doi.org/10.1007/s11042-023-14668-8.

[9] Ezenarro,A., Álvarez, M.,Barrios,L., & Rodríguez,J., "Chemometrics-Assisted Colour Histogram-based Analytical Systems (CACHAS) for maturity index estimation in olives," Journal of Food Engineering, vol. 349, p. 111567, 2023. https://doi.org/10.1007/s11042-023-14668-8.

[10] Montanaro,G., Dichio, R., Xiloyannis,C., & Todorovic, M.,"Predicting phenolic content and oil yield in olives using RGB imaging and backpropagation neural networks," Computers and Electronics in Agriculture, vol. 209, p. 107815, 2023. https://doi.org/10.1007/s11042-023-14668-8.

[11] Santelices, I., Cano,S.,Moreira,F., & Peña, A., "Artificial Vision Systems for Fruit Inspection and Classification: Systematic Literature Review",Sensors,2025. https://doi.org/10.3390/s25051524

[12] Feng, Y., Wang, G., Wang, Y., Hu, H., & Wang, S.,"Research on an intelligent seed-sorting method and sorter based on machine vision and lightweight YOLOv5n". Agronomy, 14(9), 1953,2024. https://doi.org/10.3390/agronomy14091953

[13] Rosca,C.,Rădulescu,G., & Stacu,A.," Artificial Intelligence of Things infrastructure for quality control in cast manufacturing environments: Shedding light on industry changes." Applied Sciences, 15(4), 2068,2025. https://doi.org/10.3390/app15042068

[14] León, F., Hernández, L., Contreras, S., & Flores, E.," A low-cost remotely configurable electronic trap for monitoring pest insects." Applied Sciences, 14(22), 10307,2024. https://doi.org/10.3390/app142210307

[15] Wang, X., Gao, H., Ma, X., & Wang, L.," A lightweight hybrid detection system based on the OpenMV vision module for an embedded transportation vehicle." Sensors, 25(18), 5724,2025. https://doi.org/10.3390/s25185724

[16] L. Hou, Z. Yang, C. Jin, and Q. Liu, "Tomato sorting system based on machine vision," Electronics, vol. 13, no.

11, p. 2114, 2024.
https://doi.org/10.3390/electronics13112114
[17] J. Xu, H. Wang, X. Chen, and Y. Cui, "Design and preliminary evaluation of automated sorting equipment for sweetpotatoes based on machine vision," AgriEngineering, vol. 6, no. 3, pp. 3058–3069, 2024.
https://doi.org/10.3390/agriengineering6030175
[18] H. X. Huynh, V. T. Nguyen, D. T. Nguyen, and D. P. Pham, "Design of an IoT ultrasonic–vision-based system for fruit sorting," Computers & Electrical Engineering, vol. 116, p. 109190, 2024. https://doi.org/10.1016/j.iot.2023.101017
[19] C. Shweta and M. Ritika, "Real time purpose identification and trace using deep-learning and computer vision," in Proc. IEEE Sustain. Conf., 2024.
https://doi.org/10.1109/SUSTAINED63638.2024.11074164

[20] A. Odeh and N. Odeh, "OpenCV and its applications in artificial intelligent systems," in Proc. IEEE ICCNS Conf., 2024. https://doi.org/10.1109/ICCNS62192.2024.10776047
[21] H. Hassan, A. Donkol, M. Mabrook, and A. Mabrouk, "Highly efficient machine learning approach for automatic disease and color classification of olive fruits," IEEE Access, 2024. https://doi.org/10.1109/ACCESS.2024.3362294
[22] S. Geng, "Analysis of the different statistical metrics in machine learning," Highlights in Science, Engineering and Technology, vol. 88, pp. 350–356, 2024.
https://doi.org/10.54097/jhq3tv19
[23] A. Reis, M. T. Ferreira, P. Lopes, and E. Silva, "Deep learning-based inspection system for olive ripeness classification using RGB imaging," Computers and Electronics in Agriculture, vol. 212, p. 108167, 2023.
https://doi.org/10.1016/j.compag.2023.108167
[24] N. T. Vo, L. Tran, and P. Nguyen, "Design of a low-power fruit sorting device using ESP32 and machine vision," Sensors, vol. 24, no. 2, p. 792, 2024.
https://doi.org/10.3390/s24020792
[25] Z. Zhang, Y. Sun, and J. Li, "Performance optimization of embedded AI for agricultural machine vision systems," IEEE Access, vol. 12, pp. 154321–154334, 2024.
https://doi.org/10.1109/ACCESS.2024.3412249
[26] H. Chen, F. Liu, and K. Zhou, "YOLOv8-based detection for fruit grading under complex illumination," Agronomy, vol. 14, no. 5, p. 1178, 2024.
https://doi.org/10.3390/agronomy14051178
[27] C. Wang, Y. Gao, and L. Hu, "Enhancing fruit classification via data-augmentation and ensemble CNNs," Computers and Industrial Engineering, vol. 188, p. 109865, 2024. https://doi.org/10.1016/j.cie.2024.109865