Automated Classification of Cocoa Bean Fermentation Levels Using Computer Vision

Juan Suarez, Juan Espinosa, Kebin Contreras, and Jorge Bacca
Universidad Industrial de Santander, Bucaramanga, Colombia
jbacquin@uis.edu.co

2025 XXV Symposium of Image, Signal Processing, and Artificial Vision (STSIVA) | 979-8-3315-3808-8/25/$31.00 ©2025 IEEE | DOI: 10.1109/STSIVA66383.2025.11156348

***Abstract*—This study presents an automated system for classifying the fermentation levels of cocoa beans using convolutional neural networks, specifically employing YOLO-based object detection models. RGB images of cocoa beans, which were cut using a guillotine to expose their internal structure, were analyzed and manually labeled by experts according to the NTC 1252:2021 standard. A dataset of 19 high-resolution images, containing approximately 1,850 annotated beans, was used for both training and evaluation. Four versions of YOLOv8 (n, s, m, l) were tested, with YOLOv8m demonstrating the best overall performance, achieving an Intersection over Union (IoU) of 0.6522, accuracy of 0.6817, recall of 0.6558, and an F1-score of 0.6685. Comparative tests with earlier YOLO versions (YOLOv5 to YOLOv7) confirmed YOLOv8m as the most efficient model for this task. In addition, it achieved a competitive inference time of 89.87 ms per image. These results highlight the potential of deep learning and computer vision techniques to automate the classification of cocoa bean fermentation levels, providing a faster, more objective alternative to traditional manual inspection methods.**

***Index Terms*—Cocoa, Cocoa fermentation, Classification, Image processing, Computer vision, Machine learning models, YOLO.**

I. Introduction

Cocoa bean is a strategic product for the Colombian economy, directly impacting more than 25,000 families dedicated to its production [1]. During the last four decades, its cultivation has experienced sustained growth, with approximately 95% of the grains traded in global markets [2]. Within the chocolate industry, the quality of cocoa directly influences the sensory characteristics of the final product, such as flavor, aroma and texture [3]. Consequently, an accurate classification of cocoa beans is a requirement to meet the quality standards demanded by buyers and to ensure competitiveness in the international market [4].

A determining factor in cocoa quality is the level of fermentation of the beans. Fermentation is essential to develop the compounds responsible for the sensory profile of chocolate, allowing the formation of aroma and flavor precursors [5]. In Colombia, the evaluation of the fermentation level of cocoa beans is conducted according to the technical standard NTC 1252:2021, which involves the visual inspection of cut cocoa beans using a guillotine [6]. This standard assesses characteristics such as the color of the cotyledon (ranging from purple to brown, with well-fermented beans typically showing a uniform brown color) and the shape or structure of the bean, including the presence of slaty, moldy, or insect-damaged beans, which indicate poor fermentation or contamination [7]. However, this method has limitations, as it depends on the experience of the evaluator and may generate inconsistencies in the classification [8]. In addition, the lack of a standardized protocol and the limited availability of experts make it difficult to apply the process homogeneously in different producing regions.

To address these shortcomings, this study proposes a computer vision-based methodology for the automatic classification of cocoa bean fermentation levels. A custom dataset was created using high-resolution RGB images of cocoa beans, annotated manually by cocoa quality experts following the NTC 1252:2021 standard. In total, 1,850 individual annotations were obtained from 19 images. These labeled instances were used to train and evaluate object detection models from the YOLO family. By fine-tuning pre-trained YOLO architectures, particularly YOLOv8, the system was able to detect and classify beans into three fermentation categories with promising performance, laying the groundwork for real-time, automated quality control in cocoa processing environments [9].

II. Background

A. Norm NTC1252:2021

Colombian Technical Standard NTC 1252:2021 establishes the quality requirements for dried and fermented cocoa beans marketed in Colombia. This standard is a guide to guarantee the standardization of cocoa quality in the production, marketing and export processes [10]. The main criteria evaluated include the level of grain fermentation, moisture content, the presence of physical defects (such as moldy, sprouted or contaminated grains) and physical characteristics such as grain size and weight [11].

An important aspect of the standard to be considered is the classification of the level of fermentation, which determines whether a cocoa bean is well fermented, partially fermented, or poorly fermented [12]. According to the visual criteria outlined in NTC 1252:2021, well-fermented beans typically exhibit a uniform brown color in the cotyledon and a loosened internal structure, indicating good biochemical changes. Partially fermented beans may show a mix of brown and purple coloration, with a more compact texture. In contrast, poorly fermented beans are usually dark purple or slaty in color, with a hard, compact structure, reflecting insufficient fermentation and lower flavor development.

B. YoloV8

1) Architecture: YOLO is a deep neural network architecture developed for object detection tasks in images. Unlike other traditional approaches such as Regions with Fast Convolutional

This work was supported by EcosNord-Minciencias through Project 105765 and Contract 050-2024.

Neuronal Networks (Fast R-CNN), which separate the localization and classification process into multiple stages, YOLO performs both tasks simultaneously in a single pass over the image, enabling it to achieve significantly higher processing speeds. This feature makes YOLO an ideal tool for real-time applications where speed and efficiency are paramount [13].

The YOLO architecture is based on dividing the input image into a grid, and each cell of this grid is responsible for predicting one or more bounding boxes and the probabilities of the classes associated with the detected objects [14].

Mathematically, each cell i predicts B bounding boxes, and for each box, a confidence score is computed as:

$$\text{Confidence}_i = \Pr(\text{Object}_i) \cdot \text{IoU}^{\text{truth}}_{\text{pred},i}, \tag{1}$$

Here, $\Pr(\text{Object}_i)$ is the probability of an object being present in cell i, and $\text{IoU}^{\text{truth}}_{\text{pred},i}$ is the intersection-over-union between the predicted and ground truth bounding boxes.

Additionally, YOLO predicts class conditional probabilities $\Pr(\text{Class}_c|\text{Object}_i)$ for each class c. The final class-specific confidence score is then calculated as:

$$\text{Score}_{i,c} = \Pr(\text{Class}_c|\text{Object}_i) \cdot \text{Confidence}_i, \tag{2}$$

where $\Pr(\text{Class}_c|\text{Object}_i)$ is the conditional probability that the object detected in cell i belongs to class c. This formulation allows YOLO to detect and classify multiple objects in a single forward pass, making it highly efficient for real-time applications.

2) YOLO Loss Function Components: The localization loss quantifies how accurately the predicted bounding boxes match the ground truth boxes. It focuses on minimizing the error in the predicted center coordinates (x, y) and dimensions (w, h) of each bounding box.

The localization loss is defined as:

$$\begin{aligned}\mathcal{L}_{\text{loc}} &= \lambda_{\text{coord}} \sum_{t=0}^{S^2} \sum_{j=0}^{B} \mathbb{1}^{\text{obj}}_{tj} \left[(x_t - \hat{x}_t)^2 + (y_t - \hat{y}_t)^2\right] \\ &+ \lambda_{\text{coord}} \sum_{t=0}^{S^2} \sum_{j=0}^{B} \mathbb{1}^{\text{obj}}_{tj} \left[(\sqrt{w_t} - \sqrt{\hat{w}_t})^2 + (\sqrt{h_t} - \sqrt{\hat{h}_t})^2\right],\end{aligned} \tag{3}$$

here, λ_{coord} is a weighting factor that emphasizes localization accuracy, and $\mathbb{1}^{\text{obj}}_{tj}$ is an indicator function that equals 1 if the object is present in the corresponding grid cell and bounding box predictor.

The confidence loss measures the model's certainty about the presence or absence of an object in a given cell and bounding box. It penalizes incorrect predictions for both detected and undetected objects.

$$\begin{aligned}\mathcal{L}_{\text{conf}} &= \sum_{t=0}^{S^2} \sum_{j=0}^{B} \mathbb{1}^{\text{obj}}_{tj} (C_t - \hat{C}_t)^2 \\ &+ \lambda_{\text{noobj}} \sum_{t=0}^{S^2} \sum_{j=0}^{B} \mathbb{1}^{\text{noobj}}_{tj} (C_t - \hat{C}_t)^2.\end{aligned} \tag{4}$$

In this expression, C_t represents the predicted confidence score, while $\hat{C}_t$ denotes the ground truth. The term λ_{noobj} reduces the impact of cells where no object is present.

The classification loss evaluates the discrepancy between the predicted class probabilities and the actual class labels for the detected objects.

$$\mathcal{L}_{\text{cls}} = \sum_{t=0}^{S^2} \mathbb{1}^{\text{obj}}_{t} \sum_{c \in \text{classes}} (p_t(c) - \hat{p}_t(c))^2, \tag{5}$$

here, $p_t(c)$ and $\hat{p}_t(c)$ represent the predicted and true class probabilities, respectively, for class c in cell t.

C. Total Loss

The total loss function $\mathcal{L}_{\text{total}}$ is the sum of all three components:

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{loc}} + \mathcal{L}_{\text{conf}} + \mathcal{L}_{\text{cls}}. \tag{6}$$

This composite loss function enables YOLO models to simultaneously optimize for object localization, confidence estimation, and classification during training.

III. Methodology

The proposal methodology consisting of several steps that allowed us to build, train, and evaluate the model in an organized manner. Each step ensured the final system's performance. Specifically, we carried out step a), data collection for dataset creation; step b), labeling and classification of the data; step c), testing and selection of the model architecture; step d), training and hyperparameter optimization; and step e), evaluation and analysis of the obtained metrics (Figure 1).

A. Dataset

To train the model, a database was created with RGB images of dried cocoa beans, both open and closed, as shown in Figure 2. The acquisition process was conducted under uniform lighting conditions to reduce shadows and color distortions that could affect model performance. The images were captured using an acquisition system, the data were organized and labeled according to their level of fermentation: green beans are well fermented, red beans are poorly fermented and blue beans are partially fermented, following the parameters established in the technical standard NTC 1252:2021.

B. Model training and classification

The detection and classification of cocoa beans were carried out using object detection models from the YOLO family. The objective was to detect individual beans and classify them into three categories based on their fermentation level: well fermented (green), poorly fermented (red), and partially fermented (blue), following the NTC 1252:2021 standard.

A pre-trained YOLO model was used as the base architecture, and fine-tuning was performed using a custom dataset of cocoa beans. This approach allowed the model to leverage learned features from large scale datasets while adapting specifically to the characteristics and requirements of cocoa bean classification.

The beans were manually labeled by a group of cocoa quality experts (see Figure 3) using a cocoa guillotine to expose the

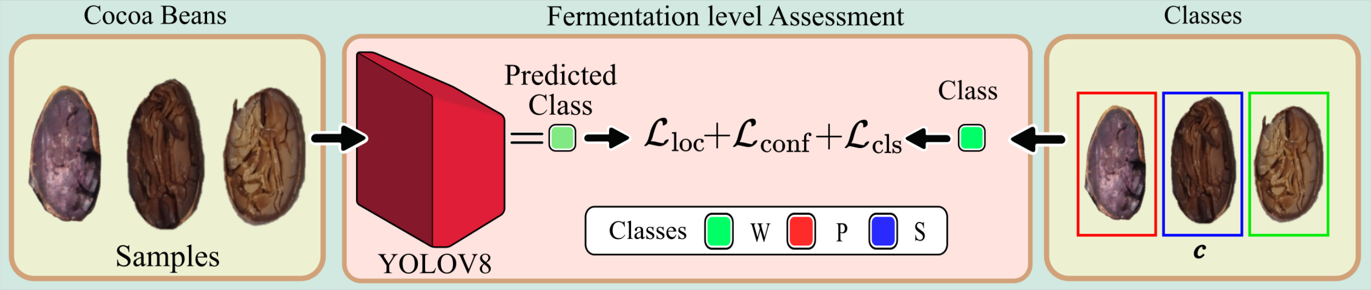

Cocoa Beans
Fermentation level Assessment
Classes
Samples
YOLOV8
Predicted Class
=
$\mathcal{L}_{\text{loc}}+\mathcal{L}_{\text{conf}}+\mathcal{L}_{\text{cls}}$
Class
Classes 🟩 W 🟥 P 🟦 S
c

Figure 1: Illustration of the proposed workflow for fermentation level assessment using YOLOv8. Cocoa bean samples are input into the model, which performs object detection and classification to assign a fermentation level. The loss function used for model optimization combines localization loss, confidence loss, and classification loss. The model predicts one of three classes—well fermented (W, green), poorly fermented (P, red), or partially fermented (S, blue) as indicated by the bounding box colors surrounding the classified beans.

internal structure of each bean. Each annotation was cross-verified by at least two evaluators to ensure accuracy and consistency.

Model training was performed in a environment using RGB images at a resolution of 640×640 pixels. each containing approximately 100 cocoa beans, 50 whole beans cut in half to expose their internal structure, this resulted in around 1900 annotated instances across the dataset, each individual bean was annotated and treated as a separate training instance. This allowed the object detection model to learn from approximately 1900 labeled examples, significantly increasing the effective dataset size despite the small number of images. The dataset was split into training and validation sets (90% and 10%, respectively).

Due to computational constraints and the time required to train the model for 300 epochs per fold, k-fold cross validation was not applied. Instead, a fixed 90/10 split was used, ensuring a representative sample of all classes in both training and validation sets.

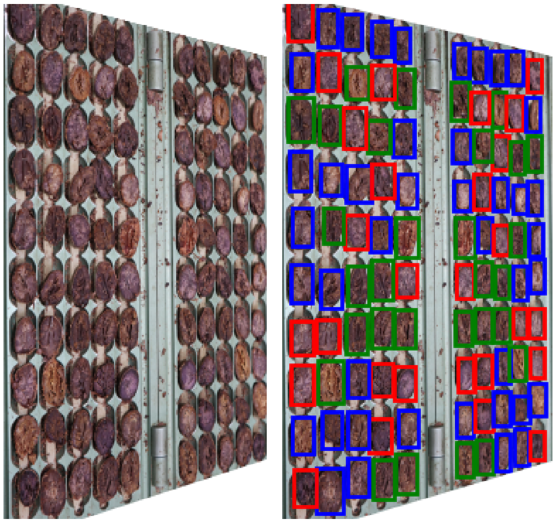
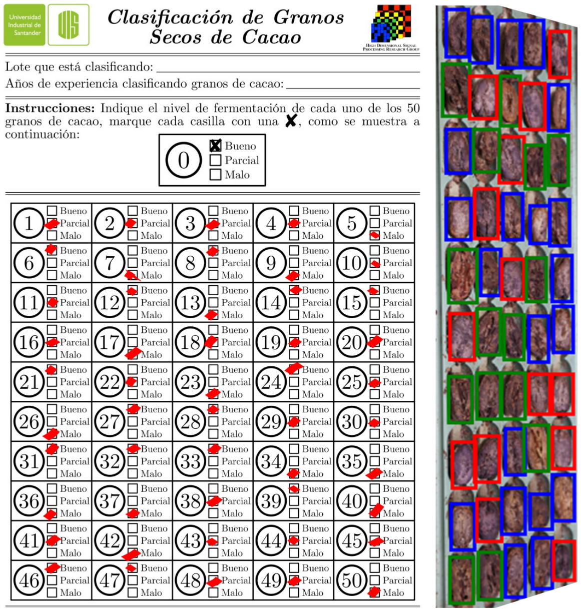


Figure 2: The image on the left shows the original unprocessed view of the cocoa beans. The right image displays the classified beans: green for well-fermented, blue for partially fermented, and red for poorly fermented.

The training process lasted for 300 epochs, with a batch size of 19 and an adaptive learning rate to enhance convergence. Data augmentation techniques such as rotations and brightness changes were applied to improve generalization. Additionally, hyper parameters including learning rate, batch size, and model depth were adjusted to optimize performance on the classification task. These hyperparameters were defined empirically through iterative experimentation. Multiple training sessions were conducted testing different values, and the final configuration was selected based on the combination that consistently yielded the best results across evaluation metrics.

The experiments were carried out using Google Colab with a Python 3 runtime, utilizing a GPU T4 accelerator. The environment provided access to 12.7 GB of RAM, with a system usage of approximately 10.2 GB of RAM and 37 GB of disk at the time of execution.



Universidad Industrial de Santander

Clasificación de Granos Secos de Cacao

High Dimensional Signal Processing Research Group

Lote que está clasificando: ____________________

Años de experiencia clasificando granos de cacao: ____________________

Instrucciones: Indique el nivel de fermentación de cada uno de los 50 granos de cacao, marque cada casilla con una ✘, como se muestra a continuación:

0 ☒ Bueno
☐ Parcial
☐ Malo

1	☐ Bueno ☒ Parcial ☐ Malo	2	☐ Bueno ☒ Parcial ☐ Malo	3	☐ Bueno ☒ Parcial ☐ Malo	4	☐ Bueno ☒ Parcial ☐ Malo	5	☐ Bueno ☐ Parcial ☒ Malo
6	☒ Bueno ☐ Parcial ☐ Malo	7	☐ Bueno ☐ Parcial ☒ Malo	8	☒ Bueno ☐ Parcial ☐ Malo	9	☐ Bueno ☐ Parcial ☒ Malo	10	☐ Bueno ☒ Parcial ☐ Malo
11	☐ Bueno ☒ Parcial ☐ Malo	12	☒ Bueno ☐ Parcial ☐ Malo	13	☐ Bueno ☐ Parcial ☒ Malo	14	☒ Bueno ☐ Parcial ☐ Malo	15	☒ Bueno ☐ Parcial ☐ Malo
16	☐ Bueno ☒ Parcial ☐ Malo	17	☐ Bueno ☐ Parcial ☒ Malo	18	☐ Bueno ☒ Parcial ☐ Malo	19	☐ Bueno ☒ Parcial ☐ Malo	20	☐ Bueno ☒ Parcial ☐ Malo
21	☒ Bueno ☐ Parcial ☐ Malo	22	☐ Bueno ☒ Parcial ☐ Malo	23	☐ Bueno ☐ Parcial ☒ Malo	24	☒ Bueno ☐ Parcial ☐ Malo	25	☐ Bueno ☒ Parcial ☐ Malo
26	☐ Bueno ☐ Parcial ☒ Malo	27	☒ Bueno ☐ Parcial ☐ Malo	28	☒ Bueno ☐ Parcial ☐ Malo	29	☐ Bueno ☒ Parcial ☐ Malo	30	☐ Bueno ☒ Parcial ☐ Malo
31	☒ Bueno ☐ Parcial ☐ Malo	32	☒ Bueno ☐ Parcial ☐ Malo	33	☒ Bueno ☐ Parcial ☐ Malo	34	☐ Bueno ☐ Parcial ☒ Malo	35	☐ Bueno ☐ Parcial ☒ Malo
36	☐ Bueno ☐ Parcial ☒ Malo	37	☐ Bueno ☐ Parcial ☒ Malo	38	☐ Bueno ☒ Parcial ☐ Malo	39	☒ Bueno ☐ Parcial ☐ Malo	40	☐ Bueno ☐ Parcial ☒ Malo
41	☐ Bueno ☒ Parcial ☐ Malo	42	☐ Bueno ☐ Parcial ☒ Malo	43	☐ Bueno ☒ Parcial ☐ Malo	44	☐ Bueno ☒ Parcial ☐ Malo	45	☐ Bueno ☒ Parcial ☐ Malo
46	☒ Bueno ☐ Parcial ☐ Malo	47	☒ Bueno ☐ Parcial ☐ Malo	48	☐ Bueno ☒ Parcial ☐ Malo	49	☐ Bueno ☒ Parcial ☐ Malo	50	☐ Bueno ☐ Parcial ☒ Malo

Figure 3: Cocoa bean fermentation classification: The left image shows annotated beans with color-coded quality (green = good, blue = partial, red = bad). The right image displays the table used to classify cocoa bean lots.)

C. Metrics

To evaluate the performance of the model, standard metrics in object detection were used, such as: Precisión, Recall, F1-Score and IoU. Analysis of these metrics identified opportunities for improvement in the model architecture and in the quality of the data set [15].

Precision: Measures the proportion of correctly predicted positive instances, reflecting the model's accuracy in identifying relevant cases.

Recall: Indicates the model's ability to detect all actual positive instances, highlighting its sensitivity to the target class.

F1-Score: Represents the harmonic mean of precision and recall, useful for assessing overall performance, especially with imbalanced data.

IoU (Intersection over Union): Quantifies the overlap between the predicted and actual object locations, evaluating spatial accuracy in object detection.

$$\text{Precision} = \frac{TP}{TP + FP}, \tag{7}$$

$$\text{Recall} = \frac{TP}{TP + FN}, \tag{8}$$

$$F_1 = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}, \tag{9}$$

$$\text{IoU} = \frac{A_{\text{pred}} \cap A_{\text{true}}}{A_{\text{pred}} \cup A_{\text{true}}}, \tag{10}$$

where: TP (True Positives) are correctly detected objects; FP (False Positives) are incorrect detections; FN (False Negatives) are real objects that were not detected; A_{pred} is the area of the predicted bounding box; A_{true} is the area of the ground truth box; $\cap$ represents the intersection between both areas, and $\cup$ their union.

IV. Implementation and Results

To evaluate the feasibility of the proposed system, several YOLO models were trained and tested using images of open cocoa beans captured under controlled conditions and labeled according to NTC 1252:2021. The goal was to automatically detect the beans and classify their level of fermentation as well-fermented, poorly fermented, or partially fermented, evaluating both the accuracy and efficiency of the system.

As illustrated in Figure 4, the "partial" class achieved the highest area under the Precision–Recall curve (0.755), indicating that the model most reliably distinguishes partially fermented beans. The "poor" class followed in performance (0.631), while the "well" class had the lowest value (0.584)), suggesting that well-fermented beans are the most challenging to consistently recognize. This highlights the need to increase and diversify the training samples for well-fermented beans.

The precision-recall curve shows the performance of the model across different thresholds. The YOLOv8m model achieved the best trade-off between precision and recall, maintaining higher values throughout the curve compared to other tested versions.

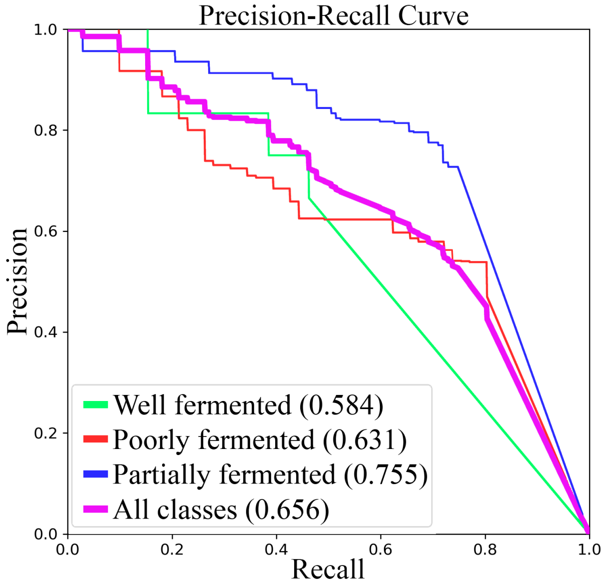


Figure 4: Precision-Recall curve by class: "partially" (0.755), 'Poorly' (0.631), "Well" (0.584) and average of all classes (blue line), showing the optimal trade-off of YOLOv8.

An initial comparative evaluation was carried out between medium-sized models from different YOLO versions, ranging from version 5 to the latest, YOLOv11. As shown in Table I, the yolov8m model stood out for its superior overall performance. This result suggests that version 8 continues to hold an advantage in specific tasks such as the visual classification of cocoa beans.

Subsequently, four variants of YOLOv8 (n, s, m, l) were tested to determine which offered the best balance between performance and computational efficiency. The results, summarized in Table II, once again highlight the strong performance of yolov8m, which achieved a precision of 0.6817, an IoU@0.5 of 0.6522, and an F1-score of 0.6685. These values demonstrate its ability to accurately identify fermented and unfermented cocoa beans, achieving a solid balance between recall (0.6558), precision, and overall accuracy. Additionally, yolov8m maintained this high performance with an inference time of 89.87 milliseconds per image, indicating an efficient processing time suitable for real-time applications.

Smaller models (YOLOv8n, YOLOv8s) achieved high recall

Table I: Results of the YOLO-m models. Several YOLO versions were evaluated, from YOLOv5 to YOLOv11. The table shows performance metrics including IoU, Accuracy, Recall, and F1-Score. The best results obtained among the tested models are highlighted in bold.

Model	IoU@0.5	Accuracy	Recall	F1-Score	Time [ms]
yolov5m	0.6722	0.6037	0.6156	0.6096	57.59
yolov8m	**0.6522**	**0.6817**	**0.6558**	**0.6685**	**89.87**
yolov5m	0.6722	0.6037	0.6156	0.6096	57.59
yolov9m	0.5750	0.6111	0.5864	0.5985	116.06
yolov10m	0.6422	0.5390	0.7012	0.6095	38.11
yolov11m	0.6867	0.6460	0.6069	0.6258	53.22

Table II: Results of the YOLOv8 models. Several YOLOv8 variants were evaluated, including yolov8n, yolov8s, yolov8m, and yolov8l. The table presents the performance metrics: Intersection over Union (IoU), Accuracy, Recall, and F1-Score. The best results obtained among the tested configurations are highlighted in bold.

Model	IoU@0.5	Accuracy	Recall	F1-Score	Time [ms]
yolov8n	0.5436	0.4508	0.7411	0.5606	39.41
yolov8s	0.5946	0.4844	0.7266	0.5813	68.71
yolov8m	**0.6522**	**0.6817**	**0.6558**	**0.6685**	**89.87**
yolov8l	0.7013	0.4951	0.7821	0.6064	77.87

values but at the cost of precision, leading to more false detections. In contrast, YOLOv8m achieved a solid balance with a precision of 0.6817, recall of 0.6558, and an F1-score of 0.6685, making it a suitable option for applications where both classification accuracy and detection coverage are critical. Moreover, it maintained this performance with an inference time of 89.87 milliseconds per image, offering a good trade-off between accuracy and computational efficiency.

The experiments performed demonstrate that the use of RGB images combined with deep learning models such as YOLOv8 allows for highly accurate classification of fermentation levels in cocoa beans, as illustrated in Figure 5. The proposed solution not only overcomes the limitations of the manual approach, but also lays the foundation for the implementation of automatic quality control systems in industrial or rural environments, with the possibility of real-time integration.

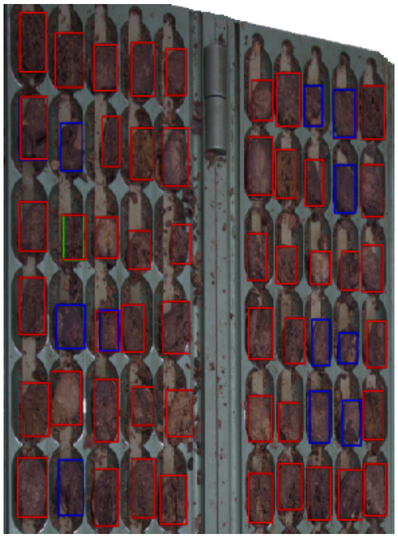


Figure 5: Detection and classification of beans by the model: red boxes = poorly fermented, blue = partially fermented, green = well fermented.

V. Conclusions

YOLOv8m achieved the highest performance in the automatic classification of cocoa bean fermentation levels, with an IoU of 0.6522, accuracy of 0.6817, recall of 0.6558, and an F1-score of 0.6685. It outperformed previous YOLO versions (v5 to v7) and other YOLOv8 variants, demonstrating a solid balance between detection coverage and classification accuracy. Additionally, it maintained this performance with an inference time of 89.87 milliseconds per image, highlighting its computational efficiency. The ”partially fermented” class obtained the highest area under the precision–recall curve (0.755), followed by the ”poorly fermented” class (0.631), while the ”well-fermented” class was the most difficult to identify (0.584), indicating the need to improve dataset representation for this category. The results support the use of computer vision models with RGB images to replace traditional visual inspection, offering a more objective and replicable system. Future work should include expanding the dataset with more cocoa varieties and environmental conditions, integrating multispectral or hyperspectral imaging to capture internal features [16], and evaluating the model under field conditions to support its deployment in automated quality control systems.

References

[1] J. Puello-Mendez, P. Meza-Castellar, L. Cortés, L. Bossa, E. Sanjuan, H. Lambis-Miranda, and L. Villamizar, “Comparative study of solar drying of cocoa beans: Two methods used in colombian rural areas,” *Chemical Engineering Transactions*, vol. 57, 2017.

[2] Swiss Platform for Sustainable Cocoa. (2024) Cocoa facts and figures. Accessed: 2025-03-31. [Online]. Available: https://www.kakaoplattform.ch/about-cocoa/cocoa-facts-and-figures

[3] A. C. Aprotosoaie, S. V. Luca, and A. Miron, “Flavor chemistry of cocoa and cocoa products—an overview,” *Comprehensive Reviews in Food Science and Food Safety*, vol. 15, no. 1, pp. 73–91, 2016.

[4] J. Cadby and T. Araki, “Towards ethical chocolate: multicriterial identifiers, pricing structures, and the role of the specialty cacao industry in sustainable development,” *SN Business & Economics*, vol. 1, no. 3, p. 44, 2021.

[5] M. Santander Muñoz, J. Rodríguez Cortina, F. E. Vaillant, and S. Escobar Parra, “An overview of the physical and biochemical transformation of cocoa seeds to beans and to chocolate: Flavor formation,” *Critical reviews in food science and nutrition*, vol. 60, no. 10, pp. 1593–1613, 2020.

[6] A. F. R. González, G. A. G. García, P. A. Polanía-Hincapié, L. J. López, and J. C. Suárez, “Fermentation and its effect on the physicochemical and sensory attributes of cocoa beans in the colombian amazon,” *Plos one*, vol. 19, no. 10, p. e0306680, 2024.

[7] D. A. Sukha, “The grading and quality of dried cocoa beans,” in *Drying and Roasting of Cocoa and Coffee*. CRC Press, 2019, pp. 89–139.

[8] K. Sánchez, J. Bacca, L. Arévalo-Sánchez, H. Arguello, and S. Castillo, “Classification of cocoa beans based on their level of fermentation using spectral information,” *TecnoLógicas*, vol. 24, no. 50, pp. 172–188, 2021.

[9] S. N. Khonina, N. L. Kazanskiy, I. V. Oseledets, A. V. Nikonorov, and M. A. Butt, “Synergy between artificial intelligence and hyperspectral imagining—a review,” *Technologies*, vol. 12, no. 9, p. 163, 2024.

[10] E. Guevara, A. E. Rojas, and H. Florez, “Technology platform for the information management of theobroma cacao crops based on the colombian technical standard 5811.” *Engineering Letters*, vol. 30, no. 1, 2022.

[11] E. Subroto, M. Djali, R. Indiarto, E. Lembong, and N. Baiti, “Microbiological activity affects post-harvest quality of cocoa (theobroma cacao l.) beans,” *Horticulturae*, vol. 9, no. 7, p. 805, 2023.

[12] S. S. Thompson, K. B. Miller, A. S. Lopez, and N. Camu, “Cocoa and coffee,” *Food microbiology: fundamentals and frontiers*, pp. 881–899, 2012.

[13] Viso Suite. (2023) Yolo explained: Everything you need to know. Accessed: Apr. 9, 2025. [Online]. Available: https://viso.ai/computer-vision/yolo-explained/

[14] T. Diwan, G. Anirudh, and J. V. Tembhurne, “Object detection using yolo: Challenges, architectural successors, datasets and applications,” *multimedia Tools and Applications*, vol. 82, no. 6, pp. 9243–9275, 2023.

[15] K. Contreras, B. Monroy, and J. Bacca, “High dynamic range modulo imaging for robust object detection in autonomous driving,” *arXiv preprint arXiv:2504.11472*, 2025.

[16] J. Bacca, E. Martinez, and H. Arguello, “Computational spectral imaging: a contemporary overview,” *Journal of the Optical Society of America A*, vol. 40, no. 4, pp. C115–C125, 2023.