https://doi.org/10.1590/1678-992X-2024-0156

ISSN 1678-992X

Scientia Agricola

Research article

Deep learning to classify the ripeness of coffee fruit in the mechanized harvesting process

Marco Antonio Zanella[1*], Mikel Barrio-Conde[2], Jaime Gomez-Gil[2], Javier Manuel Aguiar-Perez[2], María Ángeles Pérez-Juárez[2], Fabio Moreira da Silva[3]

[1]Universidade Federal Rural de Pernambuco – Depto. de Engenharia Agrícola, Rua Dom Manoel de Medeiros s/n – 52171-900 – Recife, PE – Brasil.
[2]Universidad de Valladolid – Departamento de Teoría de la Señal y Comunicaciones e Ingeniería Telemática – ETSI Telecomunicación, Paseo de Belén, 15 – 47011 – Valladolid – Spain.
[3]Universidade Federal de Lavras – Depto. de Engenharia Agrícola, Trevo Rotatório Professor Edmir Sá Santos s/n – 37203-202 – Lavras, MG – Brasil.
*Corresponding author <marco.zanella@ufrpe.br>

Edited by: Thiago Libório Romanelli

Received July 12, 2024
Accepted November 12, 2024

ABSTRACT: The coffee industry is a vital sector of global agriculture. Coffee is one of the most widely traded plant products in the world. Coffee fruit ripeness affects the taste and aroma of the final brewed beverage, coffee farms' overall yield and economic viability. Traditional methods of assessing coffee fruit ripeness, which rely on manual inspection by skilled workers, are labor-intensive, time-consuming, and prone to subjective interpretation. In this study, we have used the YOLOv9 (You Only Look Once) algorithm that outperformed previous versions particularly by using a new lightweight network architecture called the gelan-c model. The objective of this study was to identify and classify quickly and accurately the degree of ripeness of the harvested coffee fruits into the following classes: unripe, ripe-red, ripe-yellow, and overripe. The images were captured during harvesting with a commercial harvester in a coffee farm in the southern region of the state of Minas Gerais, Brazil. Data augmentation was performed to increase the dataset in terms of images and bounding boxes. Detection performance was obtained for image sizes between 128 and 640 px. The best performance was achieved with an image size of 640 px, reaching a precision level of 99 %, a recall of 98.5 %, an F1-Score of 98.75 %, a mAP@0.5 of 99.25 %, and a mAP@0.5:0.95 of about 85 % during the validation phase. Our study significantly outperforms previous studies on fruit classification in terms of models used, data augmentation strategies, and overall performance.

Keywords: YOLO, coffee farming, fruit detection, precision agriculture

Introduction

Coffee (*Coffea arabica* L.) is one of the most traded plant products in the world. Brazil is the largest producer and a significant exporter (Santos et al., 2023). Fruit ripeness affects the taste and aroma of the final brewed beverage and the overall yield and economic viability of coffee farms.

Coffee harvesting is a critical process in coffee production, and enhancing its performance can significantly reduce overall production costs (Souza et al., 2023). This process aims to harvest as many ripe coffee berries as possible without removing the unripe ones. However, traditional methods of assessing coffee fruit ripeness rely on manual inspection by skilled workers, are labor-intensive, time-consuming, and prone to subjective interpretation. This information is used to set up the harvester and the harvesting process.

The You Only Look Once (YOLO) algorithm, first introduced by Redmon et al. (2016), has evolved through several iterations, each bringing improvements in performance. YOLOv3 introduced the Darknet-53 backbone network, while YOLOv4 enhanced performance with features such as Weighted-Residual-Connections and Cross-Stage-Partial-connections (Bochkovskiy et al., 2020). YOLOv5 brought faster processing speeds and reduced model sizes, whereas YOLOv7 integrated Extended Efficient Layer Aggregation Network modules, improved learning efficiency (Wang et al., 2022). YOLOv9 was enhanced by a new lightweight network architecture known as the Generalized Efficient Layer Aggregation Network (GELAN) (Wang et al., 2024).

Several studies have demonstrated the effectiveness of YOLO in agricultural settings (Cuong et al., 2022; Fu et al., 2022; Li et al., 2023b; Santana et al., 2023). Moreover, a study by Eron et al. (2024) compared versions of YOLO. Although EfficientNet showed slightly higher accuracy in some cases, it required significantly more training time, making it less practical for real-time applications. The main objective of this work is 1) to obtain images of coffee fruits during the mechanized harvesting process and 2) to use the YOLOv9 algorithm to identify and classify the degree of ripeness of harvested coffee fruits into four classes.

Materials and Methods

This section describes the materials and means to achieve the two objectives. First, we describe the image acquisition device used to obtain images of the coffee fruits during the mechanized harvesting and labeling processes. Second, we explain how the YOLO algorithm is used to identify and classify quickly and accurately the degree of ripeness of the harvested coffee fruits into the following classes: unripe, ripe-red, ripe-yellow, and overripe.

Data acquisition and labeling

In certain cases, harvesters have used cameras to improve the harvesting process. For example, a camera was mounted on the harvester's exit spout to generate maturity maps and capture images of all the harvested fruit (Bazame et al., 2021). In a subsequent study, Bazame et al. (2022) estimated the coffee yield using the same system and compared it with data obtained by a device on the coffee harvester. In the present study, the images of the coffee fruits were captured using a device designed to capture higher quality images during the harvesting process and placed in a position different from those of previous studies. The device was designed according to the working conditions of the harvester, avoiding light variation over time and vibration during the harvesting process. This system consisted of (Figure 1):

- Support for coupling the camera to the harvesting machine.
- Anti-vibration mount for the camera.
- GoPro Hero seven video camera at a distance of 0.2 m from the conveyor belt. Images were taken in full HD definition (1920 × 1080), 60 fps, 1080P, ISO 1600 and 1/800 shutter speed.
- LED lighting system with 4 × 3.5 W bulbs.
- Rubber curtains.

The device is designed to be attached to the harvester, positioned on the cross belt, and between the vertical conveyors (Figure 2A). The cross conveyor leads the harvested coffee fruit to the bulk tank or the discharge spout, depending on the needs of the harvesting process (Figure 2B). Thus, the images collected correspond to the coffee fruits stripped from one side of the coffee plants. Following this, in the next pass of the harvester, the images of the fruits corresponding to the other side of the coffee plants in the adjacent planting line are collected.

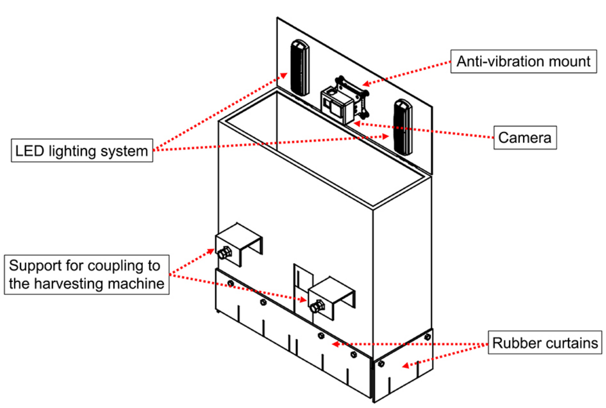

Figure 1 – Equipment used to capture images of harvested coffee fruits.

The use of the device aligns with the requirements of a commercial harvester model equipped with a storage tank for harvested coffee. This harvester model is popular for its ability to enhance autonomy during the harvesting process. Installing the device at the end of the unloading conveyor belt or the entrance to the tank would capture a more comprehensive mixture of fruit ripeness from both sides of the plants, as was demonstrated in the study by Bazame et al. (2021) and can also be used to yield monitoring as shown by Bazame et al. (2022). Considering that the harvester makes the outward pass in a coffee planting line and the return pass in the adjacent planting line, by monitoring one side of each line at a time, it is possible to assess variability in fruit ripeness in a plot with the same sun exposure characteristics, which strongly influence fruit ripeness. This focused approach could provide practical insights for managing ripeness. Therefore, monitoring only one side may lead to an initial approach which assesses variability of the sun-exposed and shaded sides on planting lines.

For this study, eight videos were recorded in full HD (1920 × 1080, 60 fps) with an approximate duration of 10-20 min each. All videos were recorded under the actual working conditions of the coffee harvester, which can include vibration, occlusion, and overlapping of coffee fruits. The frames of all videos were extracted to create a dataset. The images were selected randomly, resulting in 200 Red-Blue-Green (RGB) images of harvested coffee fruits. All images were captured with a GoPro Hero seven video camera at 0.2 m from the conveyor belt, attached to a device designed for use on the harvester, without zoom or flash, and were saved with an image resolution of 96 DPI. All images were

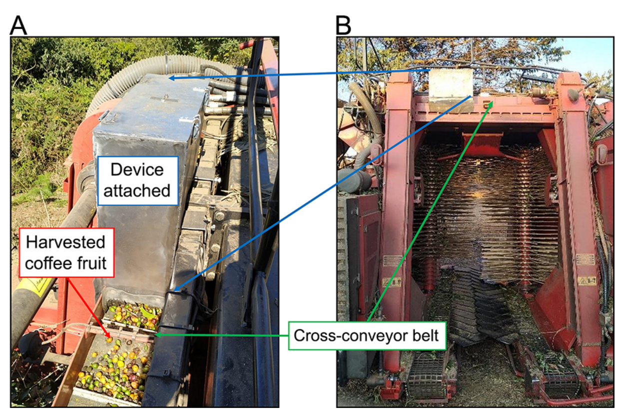

Figure 2 – Device attached to the coffee harvester to collect images. A) Side view of the device attached to the harvester on the cross-conveyor belt for transporting the harvested coffee fruit. B) Rear view of the coffee harvester with the device mounted on top of the machine and on the cross-conveyor belt.

taken during the harvest in a commercial coffee farm located in the municipality of Lavras, in the southern region of the Minas Gerais state – Brazil (21°10′17″ S, 44°58′47″ W, altitude 934 m). The images used in this study were taken between 11 and 20 July 2023. A harvester, model Coffee Express 200, was used to harvest coffee fruits. The study areas were cultivated with Arabica coffee (*C. arabica*) of different cultivars.

For labeling, three stages of ripeness are commonly used to classify the fruits: unripe, ripe. and overripe (Bazame et al., 2021; Khojastehnazhand et al., 2019). Coffee fruits in the unripe and overripe stages were predominantly close to green and brown colors, respectively. However, at the ripe stage, the coffee fruits may be yellow or red, depending on the cultivar of the coffee plant. The Graphical User Interface Label-Studio (https://github.com/HumanSignal/label-studio) was used for bounding box labeling.

Using the YOLO model to detect coffee fruit ripeness

This study employed the YOLOv9 algorithm, a single-stage target recognition deep learning algorithm. This YOLO version offers significant improvements over earlier versions, particularly through the use of the GELAN. This architecture optimizes gradient path planning, allowing for faster and more precise object detection, making it ideal for agricultural applications such as classifying coffee fruit ripeness during mechanized harvesting. Notably, this is the first study applying YOLOv9 and gelan-c specifically in the context of coffee bean ripeness detection. While YOLO models have been extensively studied in other agricultural applications, their use with gelan-c in this domain opens new avenues for real-time, high-precision detection in challenging environments.

To classify the coffee fruits into four ripeness stages (unripe, ripe-red, ripe-yellow, and overripe), the YOLOv9 model was trained on a dataset augmented with bounding boxes for each fruit. This dataset consisted of 440 images and 13,130 bounding boxes. The model's performance was evaluated on images of various sizes (128 to 640 px) and the best results were achieved with a 640-px image size, yielding a 99 % precision level, 98.5 % recall, an F1-Score of 98.75 %, and a mAP@0.5 of 99.25 %.

These results underscore the ability of YOLOv9 with gelan-c to handle real-time agricultural tasks with high precision, particularly in detecting small or overlapping fruits in complex environments.

Performance evaluation

Several experiments were performed to analyze the YOLOv9 algorithm using the configuration shown in Table 1. These experiments varied the image size from 128 × 128 to 640 × 640 px.

Table 1 – Specific parameters of the YOLOv9 (You Only Look Once) algorithm.

Factor	Value
Model	gelan-c
Input Image size	[128, 254, 320, 416, 512, 640]
Images	440
Bounding boxes	13,130
Epochs	300
Batch size	8
Confidence threshold	0.50
Optimizer	SGD
Learning rate	0.01

SGD = stochastic gradient descent.

The original dataset consisted of 63 images with 1,880 bounding boxes. However, data augmentation improves the results, increasing the dataset to 440 images and 13,130 bounding boxes. The different augmentations used in this study are illustrated in Figure 3A and B horizontal flip and rotation between –45° and +45°; vertical flip and brightness adjustment between –30 and +30 % (Figure 3C and D); blur up to 2.5 px and noise up to 3 % of px (Figure 3E and F); original image (Figure 3G). The dataset was randomly divided into three parts: training, validation, and test sets, with the percentages of images being 70, 15, and 15 %, respectively, which is a standard distribution used in similar studies.

The gelan-c model was selected for its efficient and lightweight integration of multiple computational blocks to improve precision and speed.

Results

The performance achieved by the YOLOv9 architecture using the gelan-c model is in Figure 4. It shows five graphs, each corresponding to an experiment during the validation phase. The horizontal axis represents the number of epochs, while the vertical axis represents the measured values, ranging from 0 to 1. These graphs show an excellent detection performance for image sizes between 256 and 640 px, while the performance drops significantly for 128 px, with values decreasing to almost half for most measurements.

Performance tends to stabilize around 250 epochs from 416 px. The best performance was achieved with an image size of 640 px, reaching a precision level of 99 %, a recall of 98.5 %, an F1-Score of 98.75 %, an mAP@0.5 of 99.25 %, and a mAP@0.5:0.95 of around 85 % during the validation phase.

A random frame from a conveyor belt video showcasing different stages of ripening of coffee fruits is shown in Figure 5A and B. Figure 5A shows the original frame, while Figure 5B illustrates the detection results, including bounding boxes with class names and confidence scores, which are generally high. A graph of the total number of coffee fruits detected per image is shown in Figure 6.

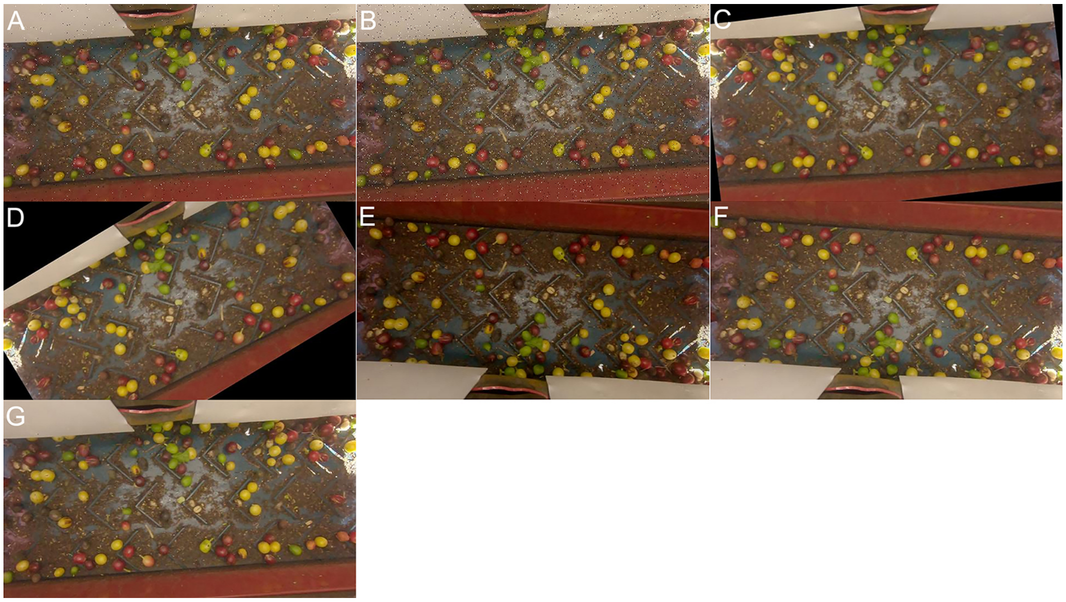

Figure 3 – Images enhanced with data augmentation: A and B) horizontal flip and rotation between –45° and +45°; C and D) vertical flip and brightness adjustment between –30 % and +30 %; E and F) blur up to 2.5 px and noise up to 3 % of px; G) original image.

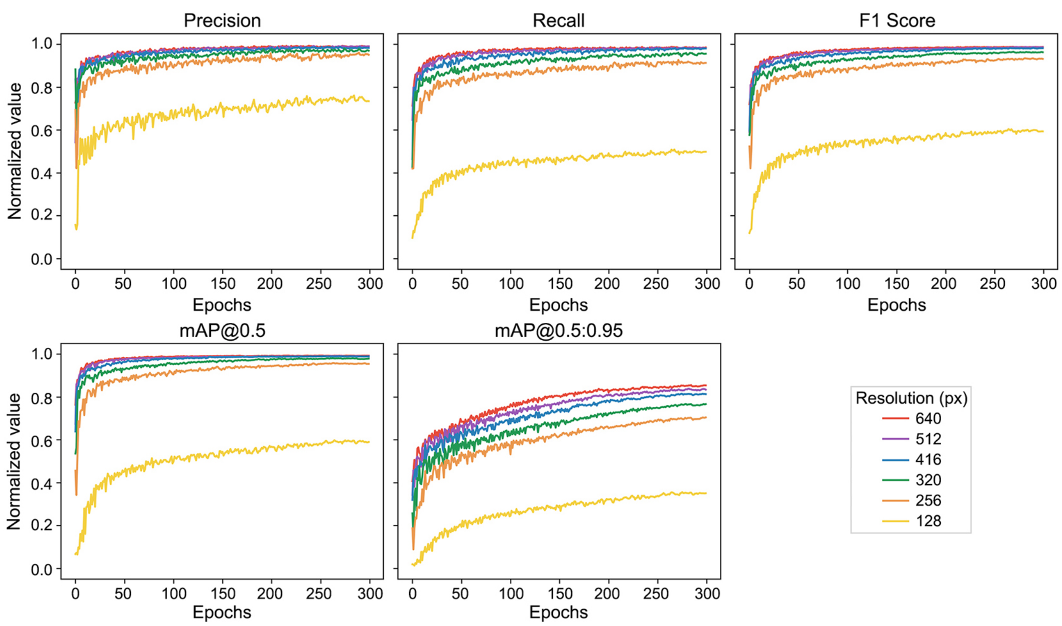

Figure 4 – Performance of the YOLOv9 (You Only Look Once) model with gelan-c during the validation phase.

To evaluate the model using the weights obtained during training, a video was recorded of a moving production line showcasing coffee fruits at various stages of ripeness: overripe, ripe red, ripe yellow, and unripe. Four graphs, each representing detections per frame for different stages of coffee fruit ripening, are contained in Figure 7. Finally, a graph of the percentage of classes detected per frame is shown in Figure 8.

Discussion

Our study provides a detailed analysis of the performance of the YOLOv9 algorithm using the gelan-c model to detect coffee fruit ripening stages. When comparing our results with two previous studies on fruit classification, we observe significant differences in the models employed, data augmentation strategies, and overall performance.

First, we compare our study with Bazame et al. (2021), who used a YOLOv3-tiny model to classify coffee fruits into three ripening stages: unripe, ripe, and overripe. Despite using higher resolution images (up to 896 px) and employing data augmentation to compensate for a small initial dataset, they achieved a maximum precision level of 83 %, a recall of 82 %, an F1-Score of 82 %, and an mAP@0.5 of 83 %. Although data augmentation was applied, their results were limited to using an older model and a smaller, less diverse dataset compared to ours. Our study used the more advanced YOLOv9 algorithm and a significantly larger augmented dataset with an additional ripening class (ripe-yellow). This provided a more comprehensive training set, allowing our model to achieve superior performance metrics by effectively generalizing across a wider variety of data conditions.

Second, an improved YOLOv7-based multi-task deep convolutional neural network, the MTD-YOLOv7 model, was employed to classify cherry tomatoes into four maturity stages (green, turning, ripe, and fully ripe) (Wenbai et al., 2024). Despite employing data augmentation and a dataset of similar size (though slightly smaller), their model achieved a precision of 85.6 %, a recall of 84.7 %, an F1-Score of 85.1 %, and a mAP@0.5 of 84.2 %. In contrast, our study, which used the YOLOv9 algorithm with the gelan-c model and a more extensive and diverse dataset, yielded significantly better results in detecting coffee fruit ripening stages. Our model achieved higher precision, recall, and overall performance metrics, highlighting the effectiveness of our approach in agricultural fruit classification tasks.

However, to our knowledge, no prior studies have applied YOLOv9 with the gelan-c architecture in the context of coffee fruit ripeness detection, making our work a novel contribution to this area of agricultural research. The ability of YOLOv9 to manage complex challenges such as occlusion and overlapping fruits during mechanized harvesting further highlights its suitability for this task. However, this may vary according to crop yield and harvester dynamic. Alternatively, we authors have implemented modules to deal with occlusions. A specific module to detect citrus fruit in occlusion scenarios was developed, demonstrating good performance with a precision level of 90.6 %, a mAP@50 of 83.2 %, and a mAP@50:95 of 60.3 % (Lin et al., 2024). In contrast, our study with YOLOv9 outperformed these models, achieving a precision level of 99 %, a recall of 98.5 %, and an F1-Score of 98.75 %. This underscores the effectiveness of YOLOv9, especially when paired with the gelan-c architecture, which optimizes gradient paths for real-time performance, making it more suitable for mechanized coffee harvesting environments where

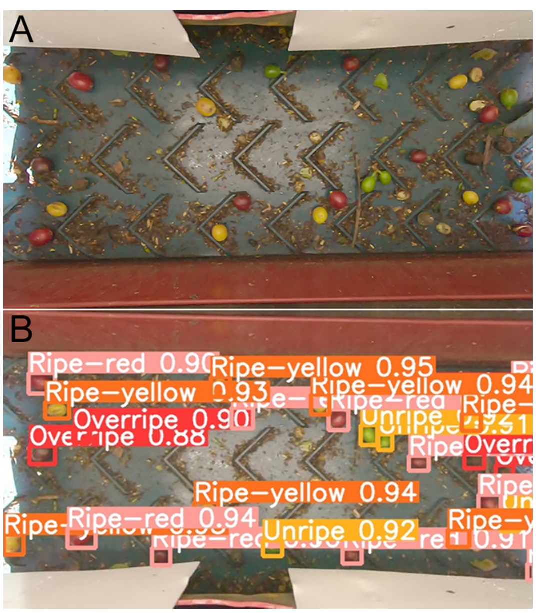

Figure 5 – A random frame with the different ripening stages of coffee fruits: A) Original frame and B) Detection in the original frame.

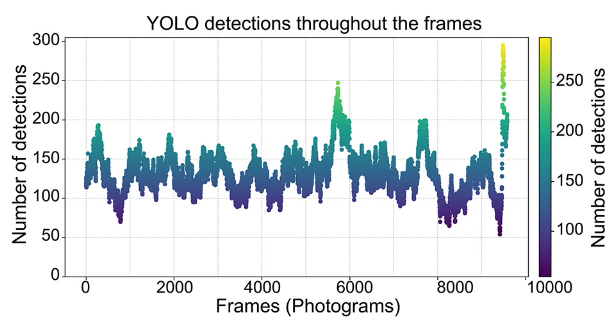

Figure 6 – Total detections per frame. YOLO = You Only Look Once.

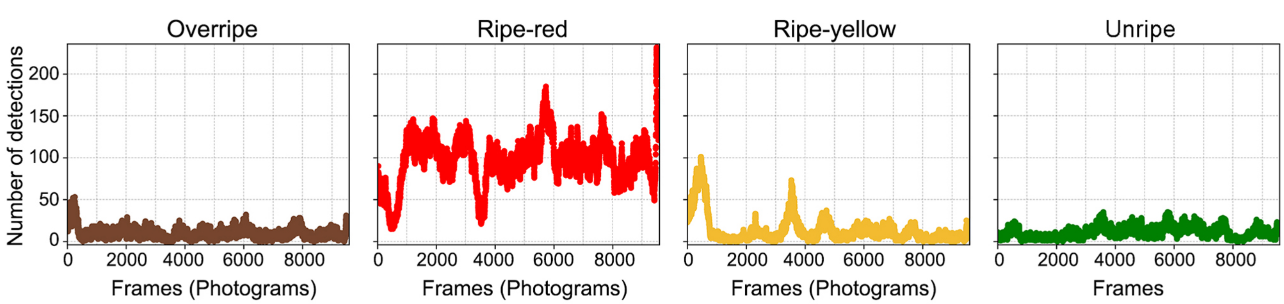

Figure 7 – Each detected class per frame.

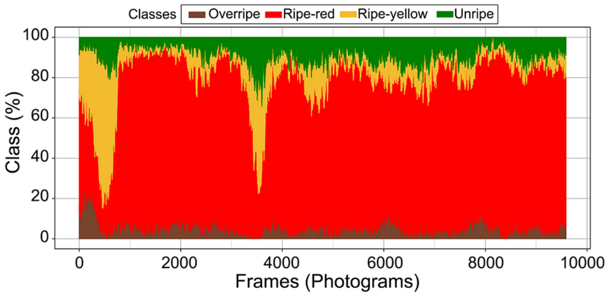

Figure 8 – Percentage of classes detected per frame.

accuracy and speed are essential. Our mAP@0.5 of 99.25 % and mAP@0.5:0.95 of 85 % further highlight the superior capability of YOLOv9 in handling intricate details like occlusion and overlapping berries, challenges frequently encountered in agricultural settings.

Additionally, Li et al. (2023a) used YOLOv7-CS, based on YOLOv7, to estimate bayberry yield. Compared with Single-Shot Detector, Region-based Convolutional Neural Network (Faster-RCNN), Deconvolutional Single-Shot Detector, and YOLOv7X target detection algorithms, YOLOv7-CS increased the mAP 0.5 by 35.52, 56.74, 12.36, and 7.05 %, respectively. These results demonstrated YOLO's greater capacity than other models. As regards the different versions of YOLO, Bazame et al. (2023) used four versions: YOLOv4, YOLOv4-tiny, YOLOv3, and YOLOv3-tiny to classify coffee fruits on tree branches into tree classes. The average precision level varies from 76 to 84 % according to the tree classes. YOLOv5, YOLOv7, and YOLOv5m6 were compared to classify four ripeness stages of coffee fruits, with the best results using YOLOv7 achieving a precision level of 85.2 %, a recall of 87.1 %, and a mAP@0.5 of 90.4 % (Eron et al., 2024). Despite these respectable results, our study surpasses Eron et al. (2024) findings in all key metrics, with a 14 % higher precision level and an 8.85 % higher mAP@0.5, demonstrating that the YOLOv9 with the gelan-c model offers significant advancements in terms of both accuracy and generalizability across different conditions.

Thus, while EfficientNet delivers strong accuracy in general-purpose tasks, its performance is constrained in agricultural applications, particularly on account of its limited efficiency on Advanced RISC Machines architectures and slow training times. On the other hand, both YOLOv7 and YOLOv8, though fast and efficient, struggle to capture the intricate details required for precise coffee ripeness classification. YOLOv9, with its optimized architecture, is better suited for real-time applications, making it a more robust choice for mechanized coffee harvesting.

In summary, our study demonstrates that the choice of model and data augmentation strategies can significantly impact the performance of fruit detection algorithms. Our approach using YOLOv9 and the gelan-c model, together with an extended and diverse dataset, represents a significant advance in the accurate classification of coffee ripening stages.

The coffee industry is a key sector of global agriculture. The ability to accurately determine fruit ripeness is important because it affects the taste and aroma of the final brewed beverage, as well as the overall yield and economic viability of coffee farms. Mechanized coffee harvesting aims to harvest as many ripe coffee fruits as possible without removing the unripe ones. In this study, the YOLOv9 algorithm with the gelan-c model was successfully used to identify and classify quickly and precisely the ripeness of harvested coffee fruits into the following classes: unripe, ripe-red, ripe-yellow, and overripe. The highest detection performances were obtained for image sizes between 256 and 640 px. Specifically, the best performance was achieved with an image size of 640 px, reaching a precision level of 99 %, a recall of 98.5 %, an F1-Score of 98.75 %, a mAP@0.5 of 99.25 %, and a mAP@0.5:0.95 of around 85 % during the validation phase. It is also important to emphasize that our study significantly outperforms some previous studies on fruit classification in terms of models used, data augmentation strategies, and overall performance, as discussed in the previous section.

Future work could enhance the present study in several ways. First, the integration of multispectral or hyperspectral imaging techniques could be explored to improve the model's ability to detect subtle variations in fruit ripening stages, such as distinguishing between early and late unripe stages. These techniques could enhance classification accuracy by providing additional spectral data that standard RGB images cannot capture. Secondly, expanding the dataset to include a broader range of environmental conditions, as well as addressing occlusion and overlapping of coffee berries would increase the model's robustness. By incorporating data from diverse lighting conditions, angles, and background complexities, the model could better generalize across varied agricultural environments, especially in regions with different climate conditions. Finally, other machine learning algorithms, or newer versions of the YOLO algorithm as they become available, could be employed to compare the results of the present study with what could be achieved using these alternative techniques. Additionally, the ripeness data generated could be utilized to improve coffee plantation management practices and to establish correlations with beverage quality parameters.

Acknowledgments

The authors would like to thank the Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) for the scholarship for the first author, to the managers of the commercial coffee farm where all the images used for this study were taken and to the EU Horizon 2020 Research and Innovation Program which partly supported this work under the Marie Sklodowska-Curie grant agreement No 101008297 (the European Union Commission is not responsible for any use that may be made of the information it contains), and the Institute of Neuroscience of Castilla y León, University of Salamanca, Spain.

Authors' Contributions

Conceptualization: Barrio-Conde M, Zanella MA, Aguiar-Perez JM, Gomez-Gil J, Silva FM. **Formal analysis:** Barrio-Conde M, Zanella MA. **Investigation:** Barrio-Conde M, Zanella MA. **Methodology:** Barrio-Conde M, Zanella MA, Pérez-Juárez MA. **Project administration:** Aguiar-Perez JM, Gomez-Gil J, Silva FM. **Resources:** Barrio-Conde M, Zanella MA, Aguiar-Perez JM, Gomez-Gil J, Silva FM. **Software:** Barrio-Conde M, Zanella MA. **Supervision:** Aguiar-Perez JM, Gomez-Gil J. **Validation:** Aguiar-Perez JM, Gomez-Gil J, Pérez-Juárez MA. **Visualization:** Barrio-Conde M, Zanella MA. **Writing-original draft:** Barrio-Conde M, Zanella MA, Pérez-Juárez MA. **Writing-review & editing:** Pérez-Juárez MA, Aguiar-Perez JM, Gomez-Gil J.

Conflict of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability statement

The authors declare that all relevant data are included in the article.

Declaration of use of AI Technologies

The authors declare that no artificial intelligence tools have been used in the writing of this paper.

References

Bazame HC, Molin JP, Althoff D, Martello M. 2021. Detection, classification, and mapping of coffee fruits during harvest with computer vision. Computers and Electronics in Agriculture 183: 106066. https://doi.org/10.1016/j.compag.2021.106066

Bazame HC, Molin JP, Althoff D, Martello M, Corrêdo LP. 2022. Mapping coffee yield with computer vision. Precision Agriculture 23: 2373-2387. https://doi.org/10.1007/s11119-022-09924-0

Bazame HC, Molin JP, Althoff D, Martello M. 2023. Detection of coffee fruits on tree branches using computer vision. Scientia Agricola 80: e20220064. https://doi.org/10.1590/1678-992X-2022-0064

Bochkovskiy A, Wang CY, Liao HYM. 2020. YOLOv4: optimal speed and accuracy of object detection. Available at: http://arxiv.org/abs/2004.10934 [Accessed May 10, 2024]

Cuong NHH, Trinh TH, Meesad P, Nguyen TT. 2022. Improved YOLO object detection algorithm to detect ripe pineapple phase. Journal of Intelligent & Fuzzy Systems 43: 1365-1381. https://doi.org/10.3233/JIFS-213251

Eron F, Noman M, Oliveira RR, Chalfun-Junior A. 2024. Computer vision-aided intelligent monitoring of coffee: towards sustainable coffee production. Scientia Horticulturae 327: 112847. https://doi.org/10.1016/j.scienta.2024.112847

Fu L, Yang Z, Wu F, Zou X, Lin J, Cao Y, et al. 2022. YOLO-Banana: a lightweight neural network for rapid detection of banana bunches and stalks in the natural environment. Agronomy 12: 391. https://doi.org/10.3390/agronomy12020391

Khojastehnazhand M, Mohammadi V, Minaei S. 2019. Maturity detection and volume estimation of apricot using image processing technique. Scientia Horticulturae 251: 247-251. https://doi.org/10.1016/j.scienta.2019.03.033

Li S, Tao T, Zhang Y, Li M, Qu H. 2023a. YOLO v7-CS: A YOLO v7-based model for lightweight bayberry target detection count. Agronomy 13: 2952. https://doi.org/10.3390/agronomy13122952

Li H, Gu Z, He D, Wang X, Huang J, Mo Y, et al. 2023b. A lightweight improved YOLOv5s model and its deployment for detecting pitaya fruits in daytime and nighttime light-supplement environments. Computers and Electronics in Agriculture 220: 108914. https://doi.org/10.1016/j.compag.2024.108914

Lin Y, Huang Z, Liang Y, Liu Y, Jiang W. 2024. AG-YOLO: a rapid citrus fruit detection algorithm with global context fusion. Agriculture 14: 114. https://doi.org/10.3390/agriculture14010114

Redmon J, Divvala S, Girshick R, Farhadi A. 2016. You only look once: unified, real-time object detection. Computer Vision and Pattern Recognition. Available at: http://arxiv.org/abs/1506.02640 [Accessed May 10, 2024]

Santana LS, Ferraz GAS, Santos GHR, Bento NL, Faria RO. 2023. Identification and counting of coffee trees based on convolutional neural network applied to RGB images obtained by RPA. Sustainability 15: 820. https://doi.org/10.3390/su15010820

Santos VP, Ribeiro PCC, Rodrigues LB. 2023. Sustainability assessment of coffee production in Brazil. Environmental Science and Pollution Research 30: 11099-11118. https://doi.org/10.1007/s11356-022-22922-z

Souza GS, Bouzan AMB, Infantini MB, Silva SA, Almeida RF. 2023. Mechanized harvesting of conilon coffee plants using a self-propelled machine. Revista Brasileira de Engenharia Agrícola e Ambiental 27: 140-148. https://doi.org/10.1590/1807-1929/agriambi.v27n2p140-148

Wang CY, Bochkovskiy A, Liao HYM. 2022. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. Computer Vision and Pattern Recognition: arXiv:2207.02696v1. https://doi.org/10.48550/arXiv.2207.02696

Wang CY, Yeh IH, Liao HYM. 2024. YOLOv9: learning what you want to learn using programmable gradient information. Computer Vision and Pattern Recognition: arXiv:2402.13616v2. https://doi.org/10.48550/arXiv.2402.13616

Wenbai C, Mengchen L, ChunJiang Z, Xingxu L, Yiqun W. 2024. MTD-YOLO: multi-task deep convolutional neural network for cherry tomato fruit bunch maturity detection. Computers and Electronics in Agriculture 216: 108533. https://doi.org/10.1016/j.compag.2023.108533