informatics

Article

Estimation of Mango Fruit Production Using Image Analysis and Machine Learning Algorithms

Liliana Arcila-Diaz *, Heber I. Mejia-Cabrera * and Juan Arcila-Diaz

Escuela de Ingeniería de Sistemas, Universidad Señor de Sipán, Chiclayo 14000, Peru; diarcilaju@uss.edu.pe
* Correspondence: adiazlilianat@uss.edu.pe (L.A.-D.); hmejiac@uss.edu.pe (H.I.M.-C.)

Abstract: Mango production is fundamental to the agricultural economy, generating income and employment in various communities. Accurate estimation of its production optimizes the planning and logistics of harvesting; traditionally, manual methods are inefficient and prone to errors. Currently, machine learning, by handling large volumes of data, emerges as an innovative solution to enhance the precision of mango production estimation. This study presents an analysis of mango fruit detection using machine learning algorithms, specifically YOLO version 8 and Faster R-CNN. The present study employs a dataset consisting of 212 original images, annotated with a total of 9604 labels, which has been expanded to include 2449 additional images and 116,654 annotations. This significant increase in dataset size notably enhances the robustness and generalization capacity of the model. The YOLO-trained model achieves an accuracy of 96.72%, a recall of 77.4%, and an F1 Score of 86%, compared to the results of Faster R-CNN, which are 98.57%, 63.80%, and 77.46%, respectively. YOLO demonstrates greater efficiency, being faster in training, consuming less memory, and utilizing fewer CPU resources. Furthermore, this study has developed a web application with a user interface that facilitates the uploading of images from mango trees considered samples. The YOLO-trained model detects the fruits of each tree in the representative sample and uses extrapolation techniques to estimate the total number of fruits across the entire population of mango trees.

Keywords: detection of fruits; production estimation; YOLO; faster R-CNN; extrapolation

Citation: Arcila-Diaz, L.; Mejia-Cabrera, H.I.; Arcila-Diaz, J. Estimation of Mango Fruit Production Using Image Analysis and Machine Learning Algorithms. *Informatics* **2024**, *11*, 87. https://doi.org/10.3390/informatics11040087

Academic Editor: Remo Pareschi

Received: 16 July 2024
Revised: 17 October 2024
Accepted: 23 October 2024
Published: 16 November 2024

1. Introduction

According to the Food and Agriculture Organization (FAO), among tropical fruits, mango is the crop with the highest global production. This tropical fruit is not only valued for its unique flavor and nutritional benefits but also plays a crucial role in generating income and providing economic stability in various regions around the world. As one of the most commercially traded fruit crops, mango represents a significant source of foreign exchange for exporting countries, contributing substantially to the trade balance. Additionally, the mango sector provides employment to numerous communities, from local farmers to workers within the production and distribution chain [1].

Mango production is an agricultural activity that represents a significant source of income for farmers. In 2022, Peru achieved a mango production quantity of 474,000 tons, cultivated over an area of 30,800 hectares. The primary producing regions were as follows: Piura with 19,867 hectares (64.6% of the total), followed by Lambayeque with 4311 hectares (14%), Áncash with 1843 hectares (6%), and finally Cajamarca with 1160 hectares (4%). The Kent variety of mango predominated in exports, accounting for 94% of the shipments [2].

Accurate estimation of fruit crop production is essential for the agricultural industry, as it enables farmers to efficiently plan the harvests, distribution, and marketing of products. This estimation underpins the organization and logistics necessary for collection, storage planning, inventory control, and market supply. Traditionally, this task has been performed using manual methods, which are labor-intensive, costly, and often imprecise due to the natural variability in the growth and yield of plants and trees [3,4].

Automated estimation of yield in fruit and vegetable fields facilitates more efficient management of resources during harvesting, storage, and transportation [5]. In the absence of automated systems for yield estimation in fruit crops, orchard managers must rely on manual fruit counting. This method involves randomly selecting several tree branches, counting the fruits present, and then extrapolating these results to the entire orchard [6].

In recent years, advancements in information technologies and the availability of large volumes of data have encouraged the adoption of machine learning techniques across various fields, including agriculture. In addition, machine learning offers a promising alternative for fruit tree production estimation, as it can handle large datasets, identify complex patterns, and enhance the accuracy of predictions [4].

Scientific research in this field has explored various methodologies and machine learning algorithms to address the challenge of fruit production estimation [7]. These approaches include regression models, decision trees, artificial neural networks [8,9], and deep learning techniques [10], each with its own specific advantages and limitations. Studies have demonstrated that these methods can integrate diverse data sources. These sources include satellite imagery, multispectral images [11], climatic data, soil characteristics, and physiological variables of the trees. Such integration helps generate more accurate and reliable estimates. Similar research has been conducted in various regions around the world [12,13], with the aim of providing advanced information to facilitate the logistical organization of this fruit. This is particularly relevant given the short duration between physiological maturity and ripening on the tree, as well as the relatively brief shelf life of the product post-harvest [14].

Accurate estimation of fruit crop production not only contributes to better agricultural planning but also optimizes the supply chain, reduces waste, and enhances the sustainability of agricultural systems. Therefore, the implementation of machine learning techniques in mango production estimation emerges as an innovative approach, enabling the development of applications for use in agriculture within the communities of technologically developing countries.

2. Related Works

In the work by [13], a model based on convolutional neural networks (CNNs) was developed to estimate citrus production using RGB images captured by an unmanned aerial vehicle (UAV). The model was trained over three annual production seasons (2017, 2018, and 2019) on a 4-hectare plot. For this purpose, 20 trees were randomly selected from a total of 1654 trees in the plot, and photographs were taken from the left and right sides of each tree. During the three seasons, the model demonstrated favorable results, achieving an accuracy of 96% and an error rate of between 4% and 6%, compared to an error rate of between 8% and 11% achieved by a specialist technician. Despite these favorable results, the model encountered difficulties in identifying fruits hidden among the branches.

In the study by [15], an efficient method for citrus recognition was proposed through the optimization of the state-of-the-art detector You Only Look Once version 4 (YOLOv4) [16]. A Kinect V2 camera was used to capture RGB images of citrus trees. To automatically select the number and size of the initial bounding boxes from these images, the Canopy and K-Means++ algorithms were employed. An enhanced YOLOv4 network structure was proposed to optimize the detection of smaller citrus fruits in complex environments. Finally, the trained network model underwent sparse training, which involved pruning non-essential channels or layers and fine-tuning the parameters of the pruned model to recover recognition accuracy. In the experiments conducted, the results indicate that the enhanced YOLOv4 detector is effective in detecting the different growth stages of citrus fruits in natural environments, with an average precision increase of 3.15% (from 92.89% to 96.04%). This result surpasses that of the original YOLOv4, YOLOv3, and Faster R-CNN. The model's efficiency is evident as the average detection time of the model is 0.06 s per frame at a resolution of 1920×1080.

In the study by [12], the challenge of detecting and counting mangoes under occlusion conditions—where the mango fruit is partially or fully obscured by leaves, stems, or other tree objects—was addressed. A method was proposed that evaluated color filtering and specific fruit characteristics, such as surface homogeneity. The efficiency of the method was highlighted through the use of information extracted from blobs (large binary objects) created after histogram filtering, including weighting, gradient topology assessment, and hierarchical clustering, without the need to pre-determine the number of clusters. Using 150 mango images divided into 30 images for training and 120 images for testing, the results demonstrated that this method could detect mangoes with precision and error rates of up to 97.53% and 0.72%, respectively.

In Table 1, relevant research pertaining to fruit counting and harvest estimation is presented, elucidating the equipment employed for data acquisition and the techniques utilized for estimation.

Table 1. Techniques and devices for fruit counting and harvest estimation.

Fruit	Technique	Data	Device	Advantages and Disadvantages of Device
Apple [17]	YOLO v4-tiny	RGB images	Kinect V2 Sensor Resolution: 1920 × 1080 pixels	The device has a high resolution and wide field of view, allowing for precise capture of fruit details. However, there are limitations for outdoor use as sunlight can affect the accuracy of the depth sensor.
Citrus [18]	YOLO v5-CS	RGB images	DJI MAVIC Air2, with SLR camera (Panasonic DMC-G7) and a mobile phone Honor 20	They provide high-quality images with excellent resolution, sharpness, and control over parameters such as exposure, focus, and aperture. In this study, they were used in uncontrolled environments, which reduces image quality.
Grape [19]	Artificial neural network (ANN)	Satellite images	Landsat 8 OLI time series	The OLI sensor features nine spectral bands covering different ranges of the spectrum. However, its spatial resolution of 30 m may be insufficient for studies requiring detailed analysis of small objects.
Melon [20]	RetinaNet	RGB images	UAV Phantom 4 Pro equipped with a DJI FC6310 RGB color camera, at 15 m above the field with a resolution of 5472 × 3648 pixels	This device enables detailed analysis and efficient coverage at 15 m above the field, offering good stability and ease of use. However, its performance is dependent on weather conditions, and image processing requires substantial storage resources and computational capacity.
Apple [21]	YOLO v5	RGB images	UAV and a Raspberry Pi camera	The Raspberry Pi camera is low-cost and highly customizable, facilitating its integration with fruit detection algorithms such as YOLOv5. However, the Raspberry Pi camera's image quality is limited compared to more advanced cameras, which may affect the accuracy of fruit detection.
Strawberry [8]	R-CNN	RGB images	Smartphone with a resolution of 4608 × 3456 in JPG format	This device captures high-resolution images, allowing for clear detail of the fruits, and the JPG format produces lightweight files. However, the required proximity for image capture (0.2–0.3 m) could limit efficiency, as multiple shots would be needed to cover a large area of the crop.

Table 1. *Cont.*

Fruit	Technique	Data	Device	Advantages and Disadvantages of Device
Strawberry [11]	Linear regression	Multispectral imaging	RPA Phantom 4 Pro Mapir Survey3W camera Resolution: 12 megapixels	This device captures multispectral images, offering detailed data on strawberry health and growth. With a 12-megapixel resolution and aerial coverage, it allows for precise crop assessment. However, it requires substantial storage capacity and processing power for image capture.
Pepper [22]	DeepSORT	RGB images	D435i RGB-D camera	It captures both color images and depth data, improving accuracy in detecting fruit maturity and counting by providing distance and shape information. However, this integration adds complexity and requires more computational power.
Mango [9]	Faster R-CNN	RGB images	Smartphone resolution: 1920 × 1080 pixels	This device is easy to use for farmers, although the images may exhibit low quality under variable lighting conditions.

3. Materials and Methods

Figure 1 illustrates the process followed for the labeling of images, data augmentation, the training of mango fruit detection models, the development of the fruit counting algorithm, and the application for data extrapolation and estimation of mango fruit production.

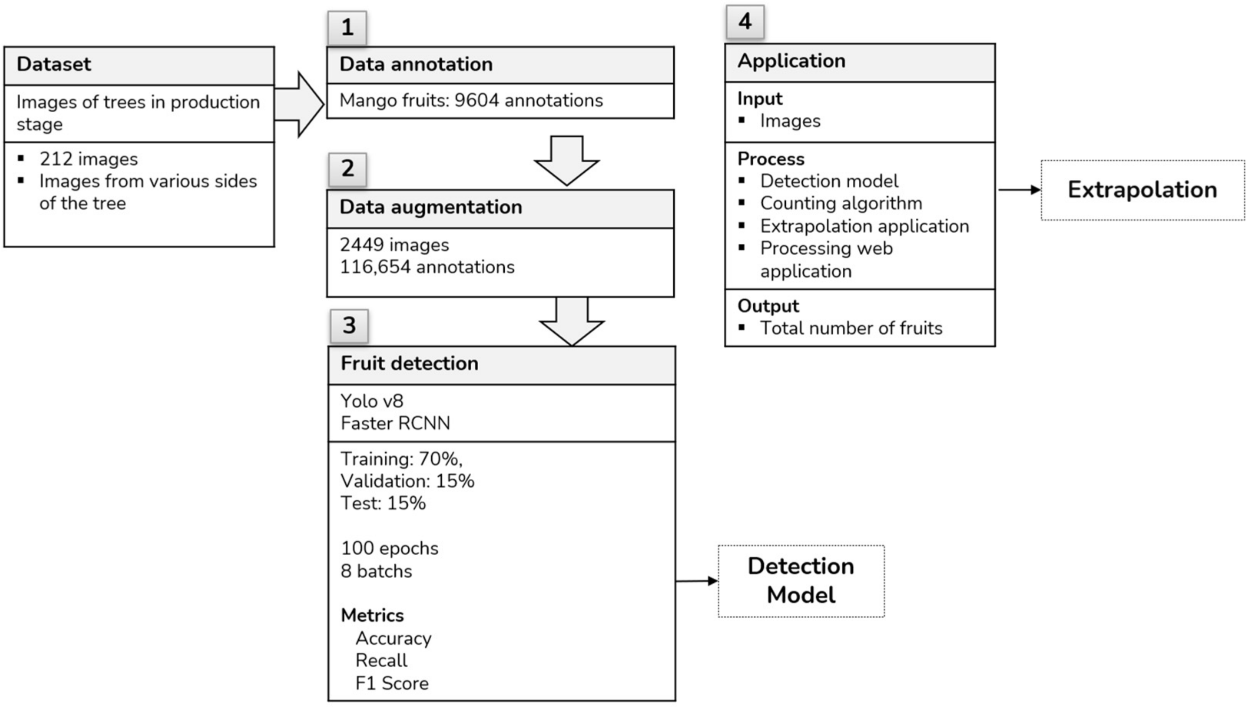

Figure 1. Method for estimating mango fruit production.

3.1. Training Environment

To train the fruit detection model, the Google Colaboratory service has been utilized, and its features are listed in Table 2.

Table 2. The training environment for the models.

Service	Features	Value
Google Collaboratory GPU NVIDIA T4	Processor speed	8 virtual CPUs at 3.7 GHz
	Storage	78.2 GB
	RAM	12.7 GB
	GPU	GDDR6 of 15 GB

3.2. Dataset

Public datasets have been utilized; the first one is presented by [23], which contains 49 images of mango trees of the *Mangifera indica* variety in the production stage. These are JPG images with dimensions of 4000 × 3000 pixels, acquired using a high-resolution RGB camera. These images were captured between 1:00 PM and 3:00 PM on a bright afternoon in April, during the tropical summer. The experiment was conducted in a mango orchard located in the village of Mudimadagu, with a latitude of 13.56 N and a longitude of 78.36 E, within the Rayalpad subdivision of the Srinivaspur taluk, a region known for mango cultivation in southeastern India.

The second dataset was published on GitHub and contains 123 images of mango trees in the production stage. These are JPG images with dimensions of 640 × 640 pixels. Another dataset consists of 40 images with a resolution of 612 × 408 pixels in JPG format. Sample images from this dataset are shown in Figure 2.

Figure 2. Sample images from the fruit tree production stage dataset.

3.2.1. Data Annotation

The images in the dataset have been manually labeled (ground truth) by the researchers using the LabelImg program, adhering to the format required by YOLO and annotating the location (bounding boxes) of the mango fruit or fruits on the tree. Figure 3 shows two sample images from the dataset with the mango fruits labeled.

In the YOLO format, bounding box coordinates are typically defined as the coordinates of the center of the bounding box, along with its width and height relative to the dimensions of the image. To use the Faster R-CNN algorithm, considering image size, annotations have been converted to the Pascal VOC format [24]. The coordinates in YOLO format are (x_{center}, y_{center}, w, h), where x_{center} and y_{center} are the coordinates of the center of the bounding box, and w and h are the width and height, all normalized in the range [0, 1].

Figure 3. Labeled sample images from the fruit tree production stage dataset.

To convert these coordinates to the Pascal VOC format ($x_{min}, y_{min}, x_{max}, y_{max}$), we use the following formulas:

$$x_{min} = x_{center} - \frac{w}{2} \quad (1)$$

$$y_{min} = y_{center} - \frac{h}{2} \quad (2)$$

$$x_{max} = x_{center} + \frac{w}{2} \quad (3)$$

$$y_{max} = y_{center} + \frac{h}{2} \quad (4)$$

3.2.2. Data Augmentation

To expand the training dataset, data augmentation has been employed, which is an effective method to combat overfitting [25]. The augmentation process includes applying various geometric transformations and distortions to the images, such as scaling, rotation, random cropping, vertical flipping, horizontal flipping, and contrast enhancement, to increase the variety of images for model training and improve accuracy.

In this study, data augmentation has been performed using progressive rotations ranging from $-5°$ to $+5°$, simulating different viewing angles of the fruits on the tree (Equation (5)). Table 3 presents the number of resulting images.

$$rotation = I_{rot}(x\prime, y\prime) = I\ (x\ cos(\theta) - y\ sin(\theta), x\ sin(\theta) + y\ cos(\theta)), \quad (5)$$

where θ is the angle of rotation.

Table 3. Dataset details.

Original Images		**Data Augmentation**	
Images	**Annotations**	**Images**	**Annotations**
212	9604	2449	116,654

3.3. Fruit Detection

In this study, two object detection methods were compared to detect mango fruits, and their efficiency was measured. Using the dataset, two models were trained: YOLO version 8 and Faster R-CNN. The dataset was divided into three subsets: 70% for training, 15% for validation, and 15% for testing. Both models were trained for 100 epochs with a batch size of 8.

3.3.1. Yolo Version 8

The first model was trained using YOLO Version 8. YOLO is a family of object detection models that divides the input image into a grid and simultaneously predicts multiple bounding boxes and their corresponding class probabilities [26].

Image Division into a Grid: The input image is divided into a grid of $S \times S$ cells. In this study, images with a width of 4000 pixels and a height of 3000 pixels are used, so for these images, a reduction factor of 32 is applied (a typical reduction factor for YOLO).

Cell Prediction: Each cell of the grid predicts a fixed number of bounding boxes B and their corresponding confidence scores. Additionally, each cell predicts a probability value for the class "mango".

Loss Calculation: the YOLO loss function includes several components: the error in bounding box coordinates (x, y, w, h), the error in bounding box confidence scores, and the error in class probabilities.

3.3.2. Faster R-CNN

This method combines region proposal and object detection into a single model, significantly improving speed and accuracy compared to previous methods such as R-CNN and Fast R-CNN [27]. The general process of Faster R-CNN includes three phases:

Feature Extraction: to extract features from each input image, the deep convolutional network ResNet50 [28] is employed.

Region Proposals: the Region Proposal Network (RPN) is a fully convolutional network that generates region proposals, which are potential locations where objects may be present.

Refinement and Classification: The proposals generated by the RPN are used to extract regions of interest (RoIs) from the feature map. These RoIs are normalized to a fixed dimension and passed through a pooling layer (RoI Pooling or RoI Align), which converts them into fixed-size feature maps.

4. Results

During this study, the efficiency of the algorithms in terms of training and the performance of the trained models were assessed.

4.1. Efficiency

To evaluate the efficiency of the two algorithms developed for mango detection in trees, three key parameters were measured: memory consumption, CPU usage, and processing time. The results are presented in Table 4.

Table 4. Efficiency of algorithms used.

Parameter	Yolo Version 8	Faster R-CNN
Storage Usage (GB)	32	32
Processing Time (s)	8046.65	13,551.23
RAM Consumption (MB)	5302.20	7450.30
GPU Memory Consumption (GB)	13.9	16
CPU Usage (%)	58.70	84.75

4.2. Performance

Using the dataset, two models were trained: YOLO version 8 and Faster R-CNN. The dataset was divided into three subsets: 70% for training, 15% for validation, and 15% for testing. Both models were trained for 100 epochs with a batch size of 8.

The detection model trained with YOLO version 8 demonstrates strong performance in identifying mango fruits within images of mango trees during the production stage. With an accuracy of 96.72%, the model achieves a high proportion of correct predictions among all detections made. A recall of 77.4% reflects the model's ability to correctly identify a significant portion of the positive instances within the dataset. Finally, the F1 Score of 86% balances precision and recall, indicating a robust overall performance.

The detection model trained with Faster R-CNN achieves a precision of 98.57%, a recall of 63.80%, and an F1 Score of 77.46%. This model exhibits a lower performance compared to the model trained with YOLO. In Figure 4, the performance of the model trained with YOLO version 8 and Faster R-CNN over 100 epochs is shown.

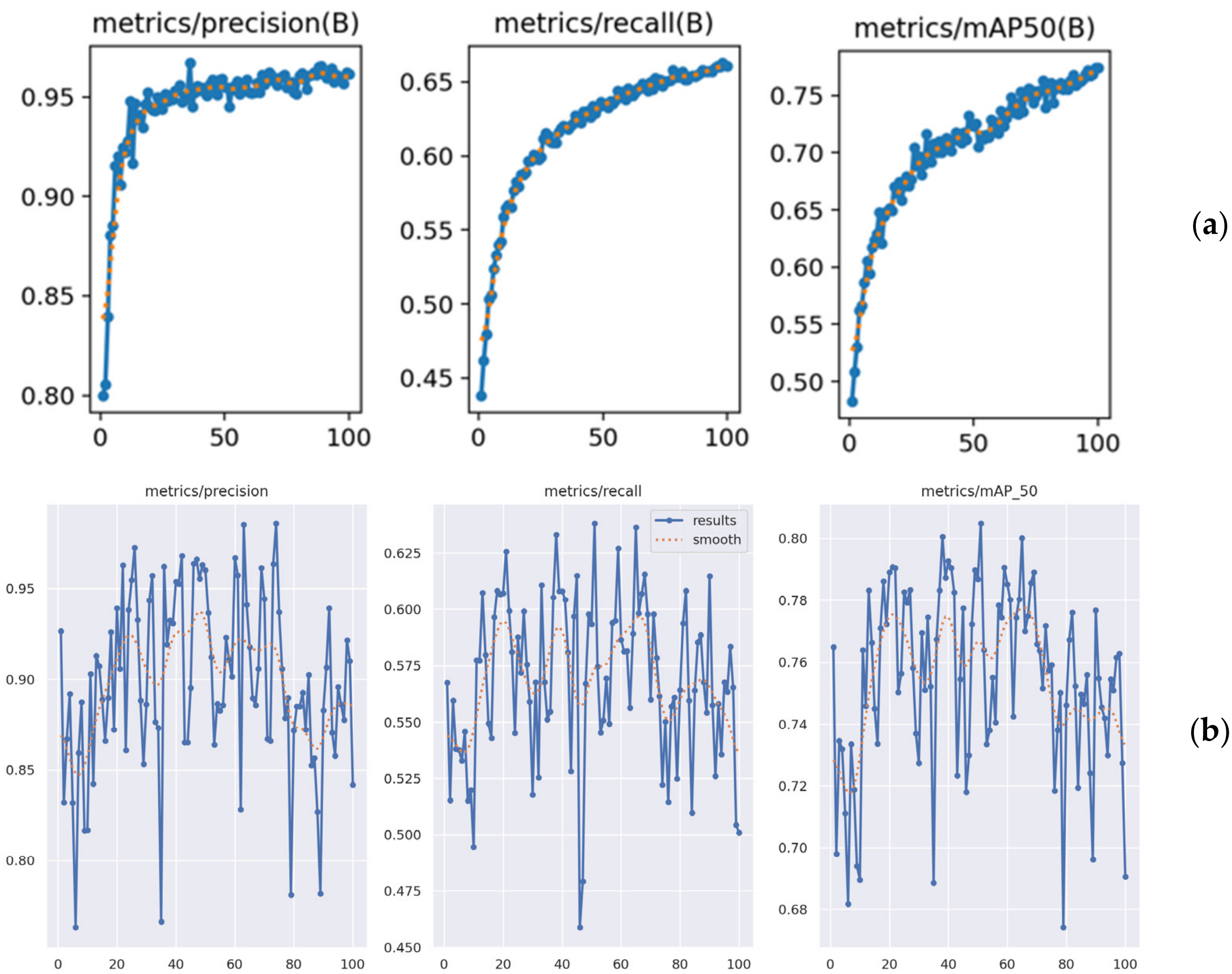

Figure 4. The performance of the model trained over 100 epochs: (**a**) YOLO version 8; (**b**) Faster R-CNN.

These trained models can process a test image to detect mango fruits in images of mango trees in the production stage. Figure 5 presents several test images with predicted labels and confidence scores.

In this study, the images of mango fruits were labeled by the researchers, ensuring that each image contained an exact number of labels corresponding to a specific fruit from the mango tree. To validate the models, 15% of the images were used, and the results obtained were derived from the training process.

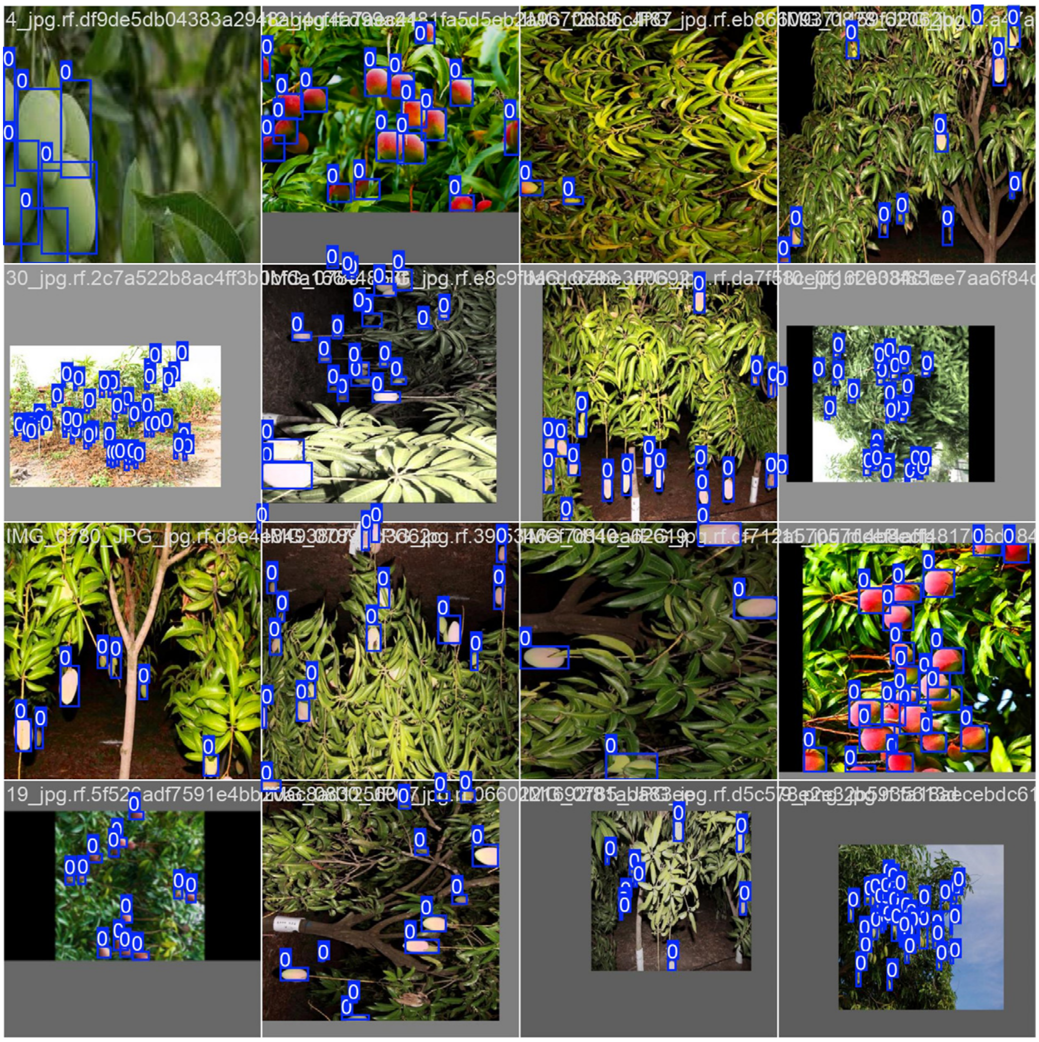

Figure 5. The detection of mango fruits with the trained models.

4.3. Estimation of Production

For counting mangoes, the initial input will include the number of trees to be photographed and the total number of trees. Subsequently, the number of fruits per sampled tree will be determined, and through extrapolation, the total mango production for all trees will be estimated.

Initially, the average number of fruits per tree (x) will be calculated as follows:

$$\overline{X} = \frac{1}{n}\sum\nolimits_{i=1}^{n} X_i \tag{6}$$

where $X_{i,}$ is the number of fruits on tree i.

Then, we will calculate the standard deviation (S) using the following equation:

$$S = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2} \tag{7}$$

The Standard Error of the Mean (SE) will be determined as follows:

$$SE = \frac{s}{\sqrt{n}} \tag{8}$$

Next, the population (N trees) will be extrapolated:

$$L_{inf} = (\overline{X} - t \times SE) \times N \quad (9)$$

$$L_{sup} = (\overline{X} + t \times SE) \times N \quad (10)$$

The confidence interval will be given by the following:

$$\overline{X} \pm t \times SE \quad (11)$$

where t is the critical value from the t-distribution for a specific confidence level (e.g., 95%) and $n - 1$ degrees of freedom.

4.4. Application for Production Estimation

Figure 6 illustrates the architecture of the developed web application, which is designed based on a client–server model. The user interface has been implemented using the JavaScript programming language, supplemented with HTML and CSS for the creation of the visual environment. This interface facilitates the execution of necessary requests to the server, allowing for data entry and the submission of images of the trees for subsequent processing. The data are sent from the user interface to the server via requests using the Hypertext Transfer Protocol (HTTP) in JavaScript Object Notation (JSON) format.

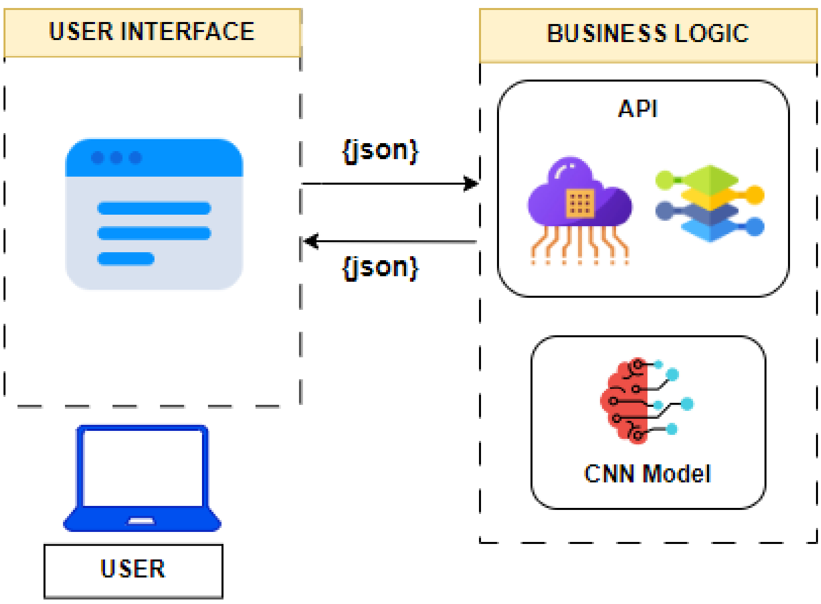

Figure 6. Software architecture of application.

The business layer has been developed in Python, utilizing the Flask framework. This layer is responsible for receiving requests through endpoints and performing the necessary processing. Within this business logic resides the trained model, which is designed to detect mango fruits on the trees. This model takes an input image and, as a result, provides the number of mango fruits identified within the image.

Initially, the application has been deployed in a local environment for testing purposes. Upon accessing the web application, users can specify the total number of mango trees present in their cultivation area and must also input the number of trees that will constitute the representative sample (Figure 7a). For each tree, the user is required to capture photographs that include all visible fruits on the tree; these images must be specifically selected for each tree, and the application will automatically display the number of fruits identified for that tree (Figure 7b). Finally, the application presents the results of the extrapolation performed, indicating the average range of mango fruits present in the crop (Figure 7c).

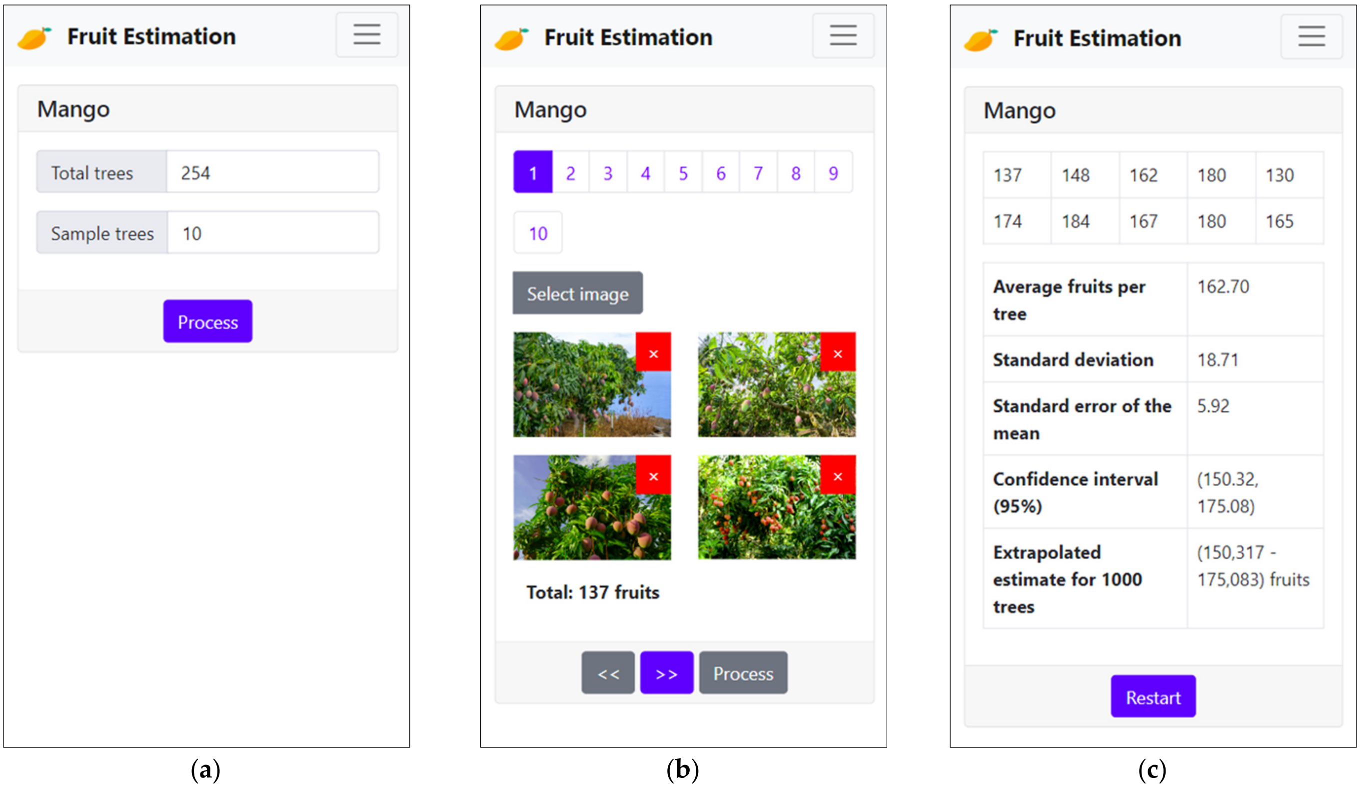

Figure 7. Application interface: (**a**) initial data entry; (**b**) image upload; (**c**) estimation results.

5. Discussion

The estimation of fruit production and the counting of fruits are based on the analysis of images as input data. Some studies employ high-resolution images, which enable the capture of precise details of the fruits; however, this approach necessitates greater storage and data processing capacity [17,20]. Other studies utilize images obtained through a UAV [11,18,21]; however, applying these results requires farmers to invest in equipment, often incurring excessive costs. Our research, similar to that of [8,9], utilizes images captured by smartphones, thereby facilitating a more economical and feasible adoption by farmers.

In the work by [13], a convolutional neural network (CNN) was employed to estimate citrus production using images captured by drones. Although the authors achieved an accuracy of 96% and an error rate of 4–6%, they encountered difficulties in identifying occluded fruits. Our research focuses on estimating mango production and utilizes the YOLO model, achieving an accuracy of 96.72%. However, the model presented in [13] faced challenges in identifying occluded fruits, a limitation that may also be pertinent to our study, where the occlusion of fruits by leaves or branches could affect detection accuracy.

In the study by [12], mango detection under occlusion conditions was addressed using color filters and fruit-specific characteristics, achieving a precision of 97.53% and an error rate of 0.72%. Although our method does not specifically focus on occlusion conditions, it is more generalizable and adaptable to different environments, making it robust for various applications.

While the study by [9] employs the ResNet-50 model, focusing on a more traditional convolutional neural network approach for yield estimation, our research leverages advanced models that facilitate real-time detection, which is crucial in dynamic agricultural environments. Although both studies aim to estimate agricultural production, our research not only emphasizes precision, achieving an accuracy of 96.72%, but also prioritizes practical application by providing farmers with a service to estimate their mango production.

In comparison to the work presented in [12], our research implements a more robust approach utilizing data augmentation techniques. The methodological framework employed in our study is based on advanced object detection models, specifically YOLOv8 and Faster R-CNN, which facilitate real-time analysis and enhance the ability to identify objects under varying lighting conditions and background complexities. Our research not only advances the accuracy of mango detection but also provides a framework for the practical implementation of these technologies in agriculture.

6. Conclusions

The dataset used in this study comprised a total of 212 original images, each with 9604 annotations. After applying data augmentation techniques, 2449 additional images were generated, resulting in 116,654 annotations. This represents a substantial increase in the amount of data available for training, which is crucial for enhancing the robustness and generalization of detection models. The significant increase in the number of images and annotations strengthens the model's ability to learn from a greater diversity of examples, thereby optimizing its performance in detection and classification tasks.

The YOLO algorithm outperforms Faster R-CNN in terms of processing time, memory consumption, and CPU usage during training. These advantages make YOLO the preferred choice for rapid development environments with resource constraints.

The results obtained from the evaluation of the YOLO and Faster R-CNN models demonstrate that YOLO outperforms Faster R-CNN, achieving an accuracy of 96.72% and an F1 Score of 86%.

Extrapolation was used to calculate the total number of fruits in mango trees based on the quantity observed in a representative sample of trees. This technique allows for an accurate estimation of the total number of fruits in a larger population from the data collected in the sample, providing a more comprehensive view of mango production in the studied trees.

Finally, a web application was developed that allows farmers to upload images of their trees and receive an estimation of mango production. While this application represents progress in the integration of technologies for agriculture, it is concluded that such solutions must evolve into accessible services for farmers, enabling them to benefit from these estimations in a consistent and efficient manner.

This project represents an initial effort in our region to estimate mango production, a crop of significant local importance. In future work, a dataset of mango images from the area will be developed, ensuring that the images are taken from a consistent distance between the capture device and the tree. Regarding image processing, the latest versions of YOLO will be utilized, and other object detection architectures will be tested. The application developed in this study will be available to all farmers through cloud computing services, ensuring broad accessibility and scalability.

Author Contributions: Conceptualization, L.A.-D.; methodology, L.A.-D.; software, L.A.-D.; validation, L.A.-D.; formal analysis, H.I.M.-C.; investigation, L.A.-D. and J.A.-D.; data curation, L.A.-D.; writing—original draft preparation, L.A.-D., J.A.-D. and H.I.M.-C.; writing—review and editing, H.I.M.-C. and J.A.-D.; supervision, H.I.M.-C. All authors have read and agreed to the published version of the manuscript.

Funding: The APC was funded by the Universidad Señor de Sipán (Perú).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors on request.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Agraria.pe. Agroindustria Generó 1.3 Millones de Puestos de Trabajo Entre Enero y Octubre de 2022. Available online: https://agraria.pe/noticias/agroindustria-genero-1-3-millones-de-puestos-de-trabajo-entr-30353 (accessed on 10 January 2023).
2. Agraria.pe. Producción Nacional de Mango Alcanzó las 474.000 Toneladas en 2022. Available online: https://agraria.pe/noticias/produccion-nacional-de-mango-alcanzo-las-474-000-toneladas-e-30987 (accessed on 10 January 2023).
3. Manonmani, S.; Senthilkumar, S.; Govind, U.S.A.; Manivannan, S. Application of Artificial Intelligence in Fruit Production: A Review. *Agric. Sci. Dig. Res. J.* **2022**, *44*, 1–5. [CrossRef]
4. Koirala, A.; Walsh, K.B.; Wang, Z.; McCarthy, C. Deep learning—Method overview and review of use for fruit detection and yield estimation. *Comput. Electron. Agric.* **2019**, *162*, 219–234. [CrossRef]
5. Javaid, M.; Haleem, A.; Khan, I.H.; Suman, R. Understanding the potential applications of Artificial Intelligence in Agriculture Sector. *Adv. Agrochem.* **2023**, *2*, 15–30. [CrossRef]
6. Häni, N.; Roy, P.; Isler, V. A comparative study of fruit detection and counting methods for yield mapping in apple orchards. *J. Field Robot.* **2020**, *37*, 263–282. [CrossRef]
7. AJacques, A.B.; Adamchuk, V.I.; Park, J.; Cloutier, G.; Clark, J.J.; Miller, C. Towards a Machine Vision-Based Yield Monitor for the Counting and Quality Mapping of Shallots. *Front. Robot. AI* **2021**, *8*, 627067. [CrossRef]
8. Li, J.; Zhu, Z.; Liu, H.; Su, Y.; Deng, L. Strawberry R-CNN: Recognition and counting model of strawberry based on improved faster R-CNN. *Ecol. Inform.* **2023**, *77*, 102210. [CrossRef]
9. Vasanth, N.M.; PPandian, R. Fast region based convolutional neural network ResNet-50 model for on tree Mango fruit yield estimation. *Indones. J. Electr. Eng. Comput. Sci.* **2024**, *33*, 1084. [CrossRef]
10. Romero, D.J.G.; Apolo-Apolo, O.E.; Ruíz, M.P. Estimating Optimal Harvest Time and Yield in Tomatoes Using Deep Learning Techniques: A Preliminary Study. In Proceedings of the 2023 IEEE International Workshop on Metrology for Agriculture and Forestry (MetroAgriFor), IEEE, Pisa, Italy, 6–8 November 2023; pp. 81–85. [CrossRef]
11. de Oliveira, L.S.; Castoldi, R.; Martins, G.D.; Medeiros, M.H. Estimation of Strawberry Crop Productivity by Machine Learning Algorithms Using Data from Multispectral Images. *Agronomy* **2023**, *13*, 1229. [CrossRef]
12. Hutagalung, A.; Nugroho, H.; Suheryadi, A.; Yunanto, P.E. Detection and Counting of Mango Fruits in Occluded Condition Using Image Analysis. In Proceedings of the 2017 5th International Conference on Instrumentation, Communications, Information Technology, and Biomedical Engineering (ICICI-BME), Bandung, Indonesia, 6–7 November 2017; pp. 190–195. [CrossRef]
13. Apolo, O.E.A.; Martínez-Guanter, J.; Egea, G.; Pérez-Ruiz, M. *Estimación de Producción en Cítricos Usando Técnicas de Aprendizaje Automático*; Universidad de Zaragoza Repository: Zaragoza, Spain, 2019. [CrossRef]
14. Anderson, N.T.; Walsh, K.B.; Koirala, A.; Wang, Z.; Amaral, M.H.; Dickinson, G.R.; Sinha, P.; Robson, A.J. Estimation of Fruit Load in Australian Mango Orchards Using Machine Vision. *Agronomy* **2021**, *11*, 1711. [CrossRef]
15. Chen, W.; Lu, S.; Liu, B.; Li, G.; Qian, T. Detecting Citrus in Orchard Environment by Using Improved YOLOv4. *Sci. Program.* **2020**, *2020*, 8859237. [CrossRef]
16. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You Only Look Once: Unified, Real-Time Object Detection. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 779–788. [CrossRef]
17. Gao, F.; Fang, W.; Sun, X.; Wu, Z.; Zhao, G.; Li, G.; Li, R.; Fu, L.; Zhang, Q. A novel apple fruit detection and counting methodology based on deep learning and trunk tracking in modern orchard. *Comput. Electron. Agric.* **2022**, *197*, 107000. [CrossRef]
18. Lyu, S.; Li, R.; Zhao, Y.; Li, Z.; Fan, R.; Liu, S. Green Citrus Detection and Counting in Orchards Based on YOLOv5-CS and AI Edge System. *Sensors* **2022**, *22*, 576. [CrossRef]
19. Arab, S.T.; Noguchi, R.; Matsushita, S.; Ahamed, T. Prediction of grape yields from time-series vegetation indices using satellite remote sensing and a machine-learning approach. *Remote Sens. Appl. Soc. Environ.* **2021**, *22*, 100485. [CrossRef]
20. Kalantar, A.; Edan, Y.; Gur, A.; Klapp, I. A deep learning system for single and overall weight estimation of melons using unmanned aerial vehicle images. *Comput. Electron. Agric.* **2020**, *178*, 105748. [CrossRef]
21. Li, Z.Y.; Yang, S.; Shi, D.; Liu, X. Yield Estimation Method of Apple Tree Based on Improved Lightweight YOLOv5. *Smart Agric.* **2021**, *3*, 100. [CrossRef]
22. Escamilla, L.D.V.; Gómez-Espinosa, A.; Cabello, J.A.E.; Cantoral-Ceballos, J.A. Maturity Recognition and Fruit Counting for Sweet Peppers in Greenhouses Using Deep Learning Neural Networks. *Agriculture* **2024**, *14*, 331. [CrossRef]
23. Kestur, R.; Meduri, A.; Narasipura, O. MangoNet: A deep semantic segmentation architecture for a method to detect and count mangoes in an open orchard. *Eng. Appl. Artif. Intell.* **2019**, *77*, 59–69. [CrossRef]
24. CVAT. Pascal VOC. Available online: https://docs.cvat.ai/docs/manual/advanced/formats/format-voc/ (accessed on 20 June 2024).
25. van Dyk, D.A.; Meng, X.-L. The Art of Data Augmentation. *J. Comput. Graph. Stat.* **2001**, *10*, 1–50. [CrossRef]
26. Yu, C.; Shin, Y. An efficient YOLO for ship detection in SAR images via channel shuffled reparameterized convolution blocks and dynamic head. *ICT Express* **2024**, *10*, 673–679. [CrossRef]

27. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *arXiv* **2016**, arXiv:1506.01497. [CrossRef]
28. Rani, R.; Sahoo, J.; Bellamkonda, S.; Kumar, S.; Pippal, S.K. Role of Artificial Intelligence in Agriculture: An Analysis and Advancements with Focus on Plant Diseases. *IEEE Access* **2023**, *11*, 137999–138019. [CrossRef]