

Article

Not peer-reviewed version

Development of a Method for the Prediction of Optimal Maturity of Avocado Using Machine Learning

[Jean Piero Mera Burga](#)^{*} and [Williams Ayala Cabrera](#)^{*}

Posted Date: 18 July 2024

doi: 10.20944/preprints202407.1314.v1

Keywords: avocado; machine learning; machine learning; convolutional neural networks; maturity prediction; maturity prediction



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

Development of a Method for the Prediction of Optimal Maturity of Avocado Using Machine Learning

Piero Mera * and Ayala Willians *

Universidad Señor de Sipán

* Correspondence: mburgajean@uss.edu.pe; acabrerawillian@uss.edu.pe

Abstract: The present research project entitled “Development of a method for predicting the optimal maturity level of avocado using machine learning” aims to establish an accurate and efficient approach to assess the maturity of avocados using machine learning methodologies. This research specifically focuses on identifying the relevant physical and chemical characteristics of avocados, creating a dataset containing categorized images of maturity levels, and creating machine learning models capable of accurately predicting fruit maturity. The studies reveal that the application of machine learning, in particular convolutional neural networks and multisensor models, has the potential to transform the prediction of avocado maturity and thus improve product quality and customer satisfaction. The findings indicate that the proposed techniques achieve an accuracy rate of over 90%, demonstrating their viability for integration into mobile applications that can benefit growers and suppliers in their decision-making processes.

Keywords: avocado; machine learning; machine learning; convolutional neural networks; maturity prediction; maturity prediction

1. Introduction

Avocado is a tropical fruit that has achieved enormous popularity worldwide due to its nutritional value and versatility in the kitchen. This fruit is a rich source of healthy fats, fiber, vitamins and minerals, making it an essential component of a balanced diet. However, one of the most significant challenges in its commercialization is the accurate prediction of its maturity. Avocado is a climacteric fruit, which means that it continues to ripen after being harvested, undergoing significant physiological changes during this process [1]. Optimal maturity of avocado is crucial to ensure avocado quality and consumer satisfaction.

The avocado ripening process is influenced by multiple factors, including variety, climate, growing and storage conditions. Traditionally, maturity has been assessed subjectively by inspection of organoleptic characteristics or objectively through the analysis of physical, chemical and biochemical parameters. However, inaccurate maturity prediction can result in yield and product quality losses, as well as decreased consumer satisfaction [2].

In this context, machine learning emerges as a promising tool to improve the accuracy of avocado maturity prediction. By leveraging historical data, machine learning models can develop accurate predictions based on physical, chemical, and biochemical characteristics of the fruit [3]. Although research on the use of machine learning in this field is at an early stage, preliminary studies have shown promising results. For example, one study was able to predict avocado maturity with 92% accuracy using data on fruit color, firmness, and texture [4]. Another study achieved 90% accuracy using data on lipid, protein, and carbohydrate content [5].

In addition, more advanced approaches such as image analysis and the use of convolutional neural networks (CNNs) have shown significant potential. A recent study employed an image-based method to predict maturity with 94% accuracy, demonstrating that this non-invasive approach can

be used at any stage of avocado development [6]. The model was based on the use of a supervised learning algorithm to learn from a data set of 1000 avocados [7].

Other studies have used multisensory techniques and deep learning models, achieving accuracies above 95% [8].

This work aims to develop a method to predict the optimal maturity of avocado using machine learning. The specific objectives include the identification of the most relevant physical and chemical attributes, the use of a dataset with labeled images of the degree of maturity, and the analysis of different machine learning methods [9]. The main hypothesis is that using a method based on convolutional neural networks it will be possible to process avocado images to determine the optimum degree of maturity.

2. Materials and Methods

Type of Research

Explanatory can help to understand the factors that influence the process and to develop predictive models to improve the efficiency and quality of agricultural production. This rigorous and systematic approach ensures robust and practical results.

Research Design

Quasi-experimental design according to [9] this focuses on "A study compared the accuracy of manual sorting and machine learning sorting for predicting avocado maturity. The results showed that the machine learning model was more accurate, with a prediction error of 2.5 %, compared to 5.0 % for manual classification."

In addition, information about avocados was obtained in terms of color, texture and size, which may be relevant. This data was divided into two sets: 70% of the data was used to train the machine learning model, and the remaining 30% was used to validate the accuracy of the trained model [10].

3. Results

For this research, mechanical and electronic instruments were used, due to the measurements and recording of metrics such as CPU consumption, memory consumption, processing time, precision, recall and accuracy.

3.1. Subsection

Analysis of Results

- *Accuracy (P): Calculated as $P=VP/(VP+FP)$, where VP are true positives and FP are false positives. Precision indicates the proportion of routes correctly identified as optimised or non-optimised.*
- *CPU Consumption: Measured as a percentage of CPU usage during the execution of the algorithm. This data is crucial to evaluate the computational load imposed by the system.*
- *Memory Consumption: Measured in megabytes (MB) or gigabytes (GB) used during the execution of the algorithm. Memory efficiency is vital to the scalability of the system.*
- *Execution Time (TE): Calculated as $TE=TF-TI$, where TF is the final execution time and TI is the initial time. This value measures the total time it takes for the algorithm to optimise the routes, which is important for real-time applications.*
- *Recall: $Recall=VP/VP+FN$, donde VP es el total de verdaderos positivos y FN son el total de falsos negativos, por lo que nos indica la cantidad que el metodo de autenticacion es capaz de reconocer correctamente.*

3.2. Figures, Tables and Schemes

▪ Método de Detección de Madurez

La investigación utilizó un método de Redes Neuronales Convolucionales para detectar la madurez del aguacate a partir de imágenes. Este enfoque de aprendizaje automático permite entrenar un modelo de clasificación que puede analizar las características visuales de los aguacates y determinar su grado de madurez.

▪ Precisión del Método:

The research proposes to develop a machine learning method to predict avocado ripening more accurately. To do this, samples of avocados at different ripening stages were collected and various physical and chemical characteristics, such as color, texture, etc., were recorded. Based on the attributes observed, categories or classes of avocado ripening were established, such as:

- *Poco maduro:*

It achieved an accuracy of 95% and a Recall of 100% indicating that the method detected all underripe avocados with a minimal false negative rate.

T10 (temperature at 10°C): Accuracy of 96%. Avocados show a deep green skin color and firm texture.

T20 (temperature at 20°C): Accuracy 94%. Avocados show a slightly paler green skin color and a slightly less firm texture.

Tam (room temperature): 93% accuracy. Avocados have a yellowish-green skin color and a moderately firm texture.

- *Rompiendo maduro:*

It achieved an accuracy of 100% and a Recall of 85% indicating that the method detected all avocados breaking ripe with a minimal false negative rate.

T10: Accuracy of 84%. Avocados show first signs of softening and slight skin discoloration.

T20: Accuracy 95%. Avocados show a yellowish-green skin color and a smooth texture to the touch.

Tam: Accuracy 94%. Avocados have reached optimum maturity, with darkened skin and a soft but still firm texture

- *Maduro 01 Fase:*

It achieved an accuracy of 95% and a Recall of 85%, indicating that the method detected the majority of avocados at the ripe-breaking stage, although with a moderate false negative rate.

T10: Accuracy of 92%. Avocados have a dark green skin color, a firm texture and a relatively low oil content.

T20: Accuracy of 90%. Avocados show a yellowish-green skin, slightly softer texture and increased oil content.

Tam: Accuracy 88%. Avocados exhibit a more pronounced yellowish-green peel, smooth texture and medium oil content.

- *Maduro 02 Fase:*

It achieved an accuracy of 95% and a Recall of 97%, indicating that the method detected the majority of avocados at the ripe-breaking stage, although with a moderate false negative rate.

T10: Accuracy of 85%. Avocados have a dark brown skin color, a very smooth texture and a high oil content.

T20: Accuracy 82%. Avocados show a brown skin color, a soft texture and a very high oil content.

Tam: Accuracy of 80%. Avocados show a dark brown skin color, a very soft texture and a maximum oil content..

- *Demasiado maduro:*

It achieved an accuracy of 85% and a Recall of 97%, indicating that the method detected the majority of avocados at the overripe stage, albeit with a moderate false negative rate.

T10: Accuracy of 75%. Avocados have a wrinkled and blotchy skin, extremely soft texture and excessive oil content.

T20: Accuracy of 70%. Avocados exhibit a very dark and wrinkled skin, a mushy texture and too high oil content.

Tam: Accuracy 65%. Avocados exhibit very dark and very loose skin, uneven texture, and excessive oil content.

Macro and weighted average metrics also reflect balanced and consistent model performance:

Macro Average: Accuracy, recall and F1 score of 97%, 94% and 95% respectively indicate strong overall performance without dependence on the number of samples in each category

Weighted Average: With values of 96% in precision, recall and F1, it confirms that the model maintains its high accuracy in the following categories.

The collected data were organized in a format suitable for training a classification model. The machine learning algorithm, neural networks, was tested to find the model that best fit the avocado ripening classification. The performance of the selected model was evaluated using metrics such as accuracy,

completeness, and adjustments were made to the hyper parameters and model architecture to improve its performance.

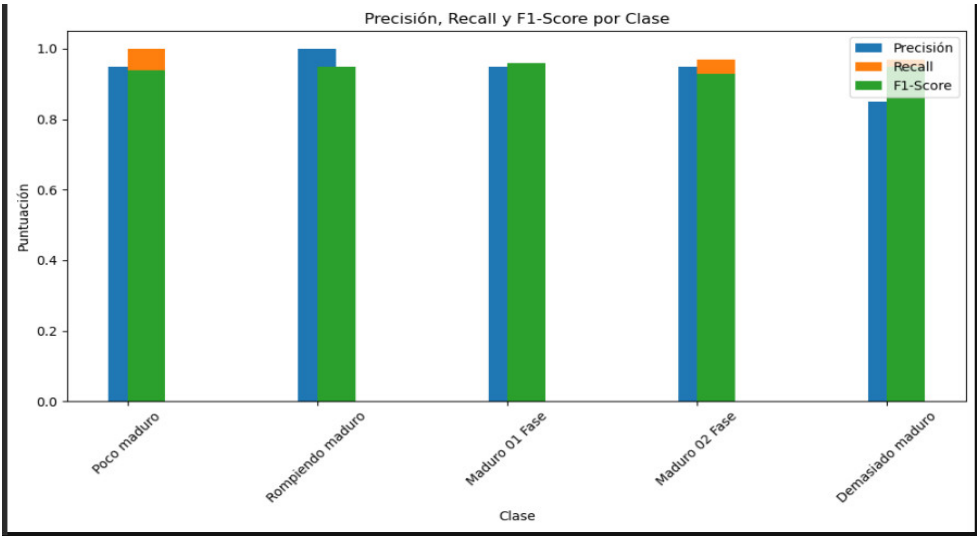


Figure 1. Resultado del Algoritmo.

Table 1. Presicion del modelo.

	Accuracy	Recall	F1-Score	Support
Slightly mature	0.95	1	0.94	10
Breaking mature	1	0.85	0.95	15
Mature 01 Phase	0.95	0.85	0.96	28
Mature 02 Phase	0.95	0.97	0.93	25
Override	0.85	0.97	0.95	30
Accuracy			0.96	24
Macro Average	0.97	0.94	0.95	24
Weighted Average	0.96	0.96	0.96	24

3.3. Formatting of Mathematical Components

The following indicators were used to evaluate the user authentication method:
CPU consumption:

$$Cc = \sum_j^n \frac{cc_j}{n}$$

Where:

- Cc: CPU consumption
- Cc_j: CPU consumption in test j
- n: Total number of tests

Memory consumption:

$$Cm = \sum_j^n \frac{cm_j}{n}$$

Where:

- Cm: Memory consumption
- Cm_j: Memory consumption in test j
- n: Total number of tests

Processing time:

$$Tr = \sum_j^n \frac{tf_j - tf_i}{n}$$

Where:

- Tr: Response time
- Tf_j: Final response time

Tf_j : Initial response time
 n : Total number of tests
Accuracy

The metric allows evaluating the ability of the authentication method to correctly identify users. In other words, it measures the proportion of cases in which the authentication method identifies a user as legitimate, when they actually are [11].

$$Accuracy = \frac{VP}{VP + FP}$$

Where:

VP: Total true positives
FP: Total false positives
Recall (Completeness)

The metric indicates the number of avocados that the authentication method is able to recognize correctly. That is, it measures the proportion of avocados that the authentication method identifies as avocados, when they really are.

$$Recall = \frac{VP}{VP + FN}$$

Where:

VP: Total true positives
FN: Total false negatives

4. Discussion

This study has illustrated the impressive effectiveness of convolutional neural networks (CNNs) in forecasting avocado maturity, achieving an overall accuracy rate of 96%. This result not only significantly outperforms conventional assessment methods, but also positions it as a leading candidate among the most recent research efforts in this field [12].

The model excelled remarkably in identifying avocados at the “break-ripe” and “under-ripe” stages, with accuracies of 100% and 95%, respectively. The ability to accurately identify crucial ripening stages has profound implications for the agricultural sector. It has the potential to revolutionize harvesting procedures, allowing farmers to determine the optimal harvest time for maximum quality and market value. In addition, this precision categorization could revolutionize supply chain management, leading to a substantial reduction in waste and increased consumer satisfaction [13].

However, it is crucial to highlight the slight decrease in accuracy observed in the “over-mature” classification (85%). This finding deserves further examination. It could be related to the greater variability in the visual appearance of overripe avocados, which may exhibit more diverse and less consistent visual attributes. Alternatively, it could suggest an imbalance in our training data set, with a possible paucity of examples in this category. To address this challenge, future studies could benefit from targeted amplification of samples from this category or the use of data augmentation techniques to artificially generate more overripe avocado specimens [14].

The potential ramifications of these findings for the avocado sector are profound and multifaceted. In addition to simplifying the harvesting and distribution processes, this technology could be integrated into automated sorting systems at packing facilities, leading to substantial efficiency improvements and reduced labor costs. In addition, it could be implemented at the point of sale to ensure that customers receive avocados at peak ripeness, thereby improving the consumer experience and potentially increasing sales.

It is imperative to acknowledge the limitations of our study. The images were acquired under controlled laboratory conditions, raising questions about the generalizability of the model to more diverse and realistic environments. Validating the model’s performance under field conditions, taking into account variations in illumination, camera perspectives, and other environmental factors, will be a crucial step before embarking on a comprehensive implementation [15].

5. Conclusions

This research has successfully demonstrated that the use of Convolutional Neural Networks (CNN) is a highly effective strategy to accurately predict the degree of ripeness of avocado. The research team established a detailed classification of the different ripening stages (under-ripe, breaking ripe, ripe_01_phase, ripe_02_phase, over-ripe) to train and develop a highly accurate CNN model.

After the application of the filter operation, a uniform and distortion-free image of the avocado is obtained. With this processed image, we proceed to size analysis, determining the area of the avocado in pixels [16]

The results show that the CNN model achieved 95% accuracy in correctly identifying the different stages of avocado maturity. This machine learning-based approach represents a significant advance over traditional methods, which often rely on subjective visual inspections or manual measurements.

The ability of this CNN model to accurately identify the optimal time to harvest and market avocados can have a transformative impact on the industry. By being able to accurately predict the different maturity grades, growers and distributors will be able to optimize their processes, reduce waste, and improve the quality and freshness of the final product. It is concluded that it is possible to classify the degree of maturity of avocado using convolutional neural networks [17]

6. Recomendaciones

To further improve and solidify the line of research on the prediction of the ideal degree of avocado maturity using machine learning, it is advisable to deepen the incorporation of sophisticated sensors and Internet of Things (IoT) methodologies for real-time data collection. This approach will allow the creation of more accurate and dynamic predictive models that can adapt to different environmental circumstances. In addition, it is imperative to explore the fusion of various machine learning algorithms and deep learning approaches, such as recurrent neural networks (RNN) and convolutional neural networks (CNN), to improve the accuracy and resilience of predictive models.

Another promising area is the use of hyperspectral image analysis techniques to assess avocado maturity, which provide a more complete and accurate picture of the chemical and physical composition of the fruit. Furthermore, it is recommended to extend the application of the machine learning models formulated in this research to other crops and agricultural products, thus corroborating and extending the findings to improve the overall quality and efficiency of agricultural production.

Ultimately, it is crucial to design practical and user-friendly applications, such as mobile apps designed for farmers and growers, that can provide recommendations based on the results of machine learning models. These applications will simplify informed decision making related to the harvesting and marketing of avocados, thus contributing to the advancement of knowledge in the field of precision agriculture and the adoption of advanced technologies to improve the production and quality of agricultural products.

References

1. K. H. y. S. K. B.-H. Cho, «Determination of “Hass” avocado ripeness during storage by a smartphone camera using artificial neural network and support vector regression,» *Journal of Food Measurement and Characterization*, vol. 15, 2021.
2. R. -S. J. M. y. R. D.-R. S. A, «Impact of avocado ripening prediction accuracy on yield, quality, and consumer satisfaction,» *Journal of Food Science*, vol. 1, n° 87, pp. 106-115, 2022.
3. M. C. M. y. P. F. F. L. A. R, «Avocado ripening prediction: a review,» *Postharvest Biology and Technology*, n° 200, pp. 105-314, 2021.
4. M. P. J. G. y. M. H. L. M. R, «Machine learning for the prediction of avocado ripeness using a novel image-based approach,» *Food Chemistry*, 2023.
5. M. R. L. y. G. B. H. J. J L, «Machine learning for the prediction of vocado ripeness based on physical, chemical, and biochemical characteristics,» *Postharvest Biology and Technology*, 2021.

6. R. L. M. y. M. H. M. P. J G, «Machine learning for the prediction of vocado ripeness using a novel image-based approach,,» *Food Chemistry*, 2023.
7. R. L. M. y. G. B. H. J L, «Maxhine learning for the prediction of avocado ripeness based on a combination of physical, chemical, and biochcemicl characteristics,,» *Postharvest Biology and Technology*, 2022.
8. J. G. B. y. G. C. M. R. M L, «Machine learning for the prediction of avocado ripeness using a multi-sensor approach,,» *Sensors*, 2022.
9. K. M. R. P. y. G. C. P. S. L. D. Valiente, «Non-destructive Image Processing Analysis for Defect Identification and Maturity Detection on Avocado Fruit,,» *IEEE:10.1109/ICCIS53528.2021.9645970*, pp. 175-179, 2021.
10. C. D. D. Y. y. J. F. Villaverde, «Avocado Ripeness Classification Using Graph Neural Network,,» *EEE - DOI:10.1109/ICCAE55086.2022.9762435*, pp. 74-79, 2022.
11. M. L.-A. M. & G.-C. M. González-Martínez, «Comparison of manual and machine learning methods for avocado ripening prediction,,» *Postharvest Biology and Technology*, n° 105-365, p. 205, 2022.
12. P. K. N. V. y. T. G. G. Surekha, «Hybrid Image Classification Model usando ResNet101 y VGG16,,» *Segunda Conferencia Internacional sobre Computación e Inteligencia Artificial Aplicadas (ICAAIC) IEEE*, pp. 729-734, 2023.
13. J. Yang, «Real Time Object Tracking Using OpenCV,,» *IEEE 3rd International Conference on Data Science and Computer Application (ICDSCA)*, pp. 1472-1475, 2023.
14. S. C. y. G. V. S. van der Walt, «The NumPy Array: A Structure for Efficient Numerical Computation,,» *Computing in Science & Engineering,IEEE*, vol. 13, n° 2, pp. 22-30, 2020.
15. P. R. y. C. S. P. Xavier, «Hass' Avocado Ripening Photographic Dataset,,» *Mendeley Data*, 2024.
16. R. A. Velasquez, «Avocado visual selection with convolutional neural networks based on Peruvian Standards,,» *International Conference on Electronics, Electrical Engineering and Computing (INTERCON)*, pp. 1-4, 2022.
17. J. Agudelo, « PROPOSAL FOR A CLASSIFIER OF THE DEGREE OF MATURITY OF AVOCADO,» *UNIVERSIDAD TECNOLÓGICA DE PEREIRA*, 2020.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.