COMPUTER SCIENCE

Development of an artificial vision algorithm using image processing techniques to assess the characteristics of export-grade bananas

Ryan Abraham León León[*] and **Martin Antonio Renteria Davila**

Facultad de Ingeniería, Universidad Privada del Norte, 0599-0141, Trujillo, La Libertad, Perú. *Author for correspondence. E-mail: ryan.leon@upn.edu.pe

ABSTRACT. The main objective of this study is to develop a computer vision algorithm capable of identifying export-quality bananas, simplifying the quality control process. Unlike previous research, such as that of Figueroa and Roa (2016), which focused on different fruits using segmentation and clustering techniques, this work integrates a controlled lighting system and advanced neural networks trained with TensorFlow. These innovative approaches ensure an average accuracy of 97.22%, surpassing previous standards in the field and demonstrating their applicability in industrial settings.

Keywords: Python; artificial vision; processing; algorithm; export.

Introduction

Banana exports are an essential component of the Peruvian economy, contributing significantly to the country's income and employment (Gallego et al., 2018). This sector faces several challenges, including the need to ensure the quality and consistency of products destined for international markets (Villaverde et al., 2013). In particular, the identification and selection of defective bananas is a central problem that can negatively impact the competitiveness and reputation of the Peruvian banana industry globally (Benito et al., 2015). The lack of standardized and automated methods to carry out this task entails inherent risks, such as inconsistency in the selection criteria, subjectivity in the evaluations and the possibility of overlooking defects or irregularities that can compromise the quality of the final product (Torres et al., 2015). This problem becomes even more relevant in the context of a globalized and highly competitive market, where consumers increasingly demand high quality, fresh and defect-free products (Negrín et al., 2013). The ability to meet these demands and maintain standards of excellence in banana production and export thus becomes an imperative to ensure the viability and sustainable growth of the Peruvian banana sector in the long term (Challenger et al., 2014).

In this sense, the development of innovative and technologically advanced solutions is presented as a key strategy to effectively address this problem (Semenovich et al., 2020). The application of computer vision and image processing techniques emerges as a promising option to improve the methods for assessing the quality of bananas for export (Figueredo & Ballesteros, 2015). The use of computer vision and image analysis provides solutions in various fields, such as industry, science, and security (Raschka et al., 2020). Using automated algorithms and systems, it is possible to analyze a wide range of characteristics objectively and accurately, from color and texture to shape and size, thus facilitating the identification and classification of defective bananas efficiently and reliably (Viera, 2017).

On the one hand, according to the study entitled 'Maduración controlada y color en bananos', it establishes the need to analyze the diameter and filling of banana fruits during ripening and thus determine the maturity stages suitable for export, using methods such as the 'filling index', which is the weight of the inner fruit of the first or second hand, divided by its length. The results of an evaluation by sampling and mathematical calculation show that the ripening of banana fruits is characterized by changes in diameter and filling, which influences their optimum state for different uses (Obregon & Cachay, 2017).

On the other hand, in the search to address similar problems in the agricultural industry with technological tools, several attempts have been made. For example, in the study entitled 'Desarrollo de un Sistema de Visión

Artificial para Evaluar el Estado de Madurez de las Granadillas', a successful approach to assess fruit maturity through image analysis is presented. In this study, 90 images of granadillas at different stages of ripeness were captured using a 5-megapixel Raspberry Pi camera. These images were subjected to a preprocessing process to remove noise, followed by segmentation, characterization and classification of the pomegranate seeds using the K-means clustering algorithm. The validation results yielded a high accuracy rate of 93% (Figueroa & Roa, 2016).

Also, in the paper 'Clasificación de Imágenes de Frutas Utilizando Inteligencia Artificial', the effectiveness of artificial intelligence in accurately classifying fruits based on features extracted from images is demonstrated. In this study, fruit images were acquired, and a preprocessing process was applied. Then, a data set (dataset) including the images and their corresponding masks was created and divided into two information vectors: a feature vector and a label vector. A supervised learning model was used, with 70% of the data for training and the remaining 30% for model evaluation. During this stage, the algorithm performed predictions and classifications of unlabeled images, obtaining satisfactory results in the categorization of fruits (Heras, 2017).

In the last decade, the field of computer vision applied to fruit classification has seen significant advancements. However, there are still critical areas that require attention to improve the precision, efficiency, and adaptability of these tools in specific contexts. In particular, the banana export industry faces unique challenges that have not been fully addressed in previous research.

Studies by Figueroa and Roa (2016) and Heras (2017) developed computer vision tools to identify the ripeness stage of fruits like granadilla and classify other fruits in general. However, these systems are not specifically designed for bananas, overlooking unique features such as the impact of black spots on export quality. Moreover, most existing systems focus on general ripeness or classification criteria without considering specific quality criteria required in international markets, such as the detection of visual defects in bananas that affect their acceptance for export. Additionally, there is no widely accepted standard in the agricultural sector that combines advanced image processing techniques and neural networks for banana evaluation, resulting in inconsistent and subjective classification methods. Although some systems have shown precision in controlled environments, their implementation in the banana industry is limited by the lack of robustness against variations in lighting, size, and position of the fruits.

This study introduces an innovative approach in the field of computer vision applied to the classification of export bananas, making significant advancements in several key aspects that have not been addressed in the scientific literature. Unlike previous research focused on generic fruits or granadillas, this study is specifically designed to address the unique challenges of bananas intended for export. This includes the precise detection of visible defects such as black spots and the evaluation of color shades, critical parameters in international quality standards. This work combines traditional methods, such as converting RGB images to grayscale, with modern neural network techniques using TensorFlow, enabling more precise and robust detection of relevant features in bananas. Additionally, one of the study's main contributions is the design of a controlled lighting system, which reduces lighting variations and enhances the quality of captured images, an aspect that previous research has not fully addressed.

The algorithm developed in this study achieves an average accuracy of 97.22%, surpassing the results of previous studies in the field. Moreover, it achieves 100% accuracy in the final testing days, demonstrating its effectiveness in both controlled and industrial contexts. The tool is designed with a practical focus, allowing its implementation in automated classification lines at banana export facilities. This aspect adds direct value to the agricultural industry, reducing human errors, improving classification consistency, and increasing operational efficiency. This study not only refines existing approaches in computer vision but also sets a precedent in the design of systems specifically for the banana industry. The combination of advanced image processing techniques, deep learning algorithms, and their applicability in industrial environments represents a significant and innovative contribution to the field.

This study stands out by addressing the unique challenges of the banana industry, combining image preprocessing techniques with advanced neural networks. Compared to the work of Figueroa and Roa (2016), which achieved 92.6% accuracy in fruit ripeness using clustering, and Heras (2017), who classified fruits with supervised learning, our approach improves accuracy by 5% and tailors the criteria specifically to the demands of the international banana market. Additionally, the proposed system optimizes capture conditions through controlled lighting, an aspect not explored in previous studies.

Methodology

To conduct our research, we have selected Python 3.11 as the main programming language. Additionally, we have set up a conducive environment for implementation and study control using several key components. First, we have integrated a powerful LED light source that provides illumination equivalent to a 100-watt light bulb. This ensures adequate lighting for our image capture needs. For image capture, we have employed a quad-camera setup with a 48-megapixel primary sensor, along with additional 8-megapixel sensors for wide-angle shots, 2-megapixel sensors for macro photography, and another 2-megapixel sensor for background blur (bokeh) effects. This camera has been chosen for its compatibility at both hardware and software levels, enabling effective integration with our processing platform. Furthermore, we have designed a structure that incorporates mounting points for the lighting, camera, and the subject of study, which in this case is the bananas. For ease of access and manipulation, we have incorporated a curtain-type entry in this structure. As an intermediary for Python programming and data processing, we have used a laptop equipped with an Intel Core i5 processor and 8 GB of RAM. This configuration provides us with the necessary performance to conduct our investigations efficiently.

Two image collection techniques were employed for this study: the first consisted of extracting images from online sources, while the second involved using the camera in the previously mentioned environment. In total, one thousand images of 'Plátano Manzano' (Musa balbisiana) at their physiological ripeness stage were collected using the RGB (Red, Green, Blue) color format. Figure 1 shows the three stages of banana ripening in chronological order. Stage 1 corresponds to the initial harvest stage, and to advance to the subsequent stages, 8 to 10 days must elapse in each case. The choice of physiological ripeness was based on experimental research, as it is considered that the optimal harvest time occurs in the ninth week after the emergence of the banana flower when the banana is in its initial ripeness state suitable for international marketing (Arrieta et al., 2006).
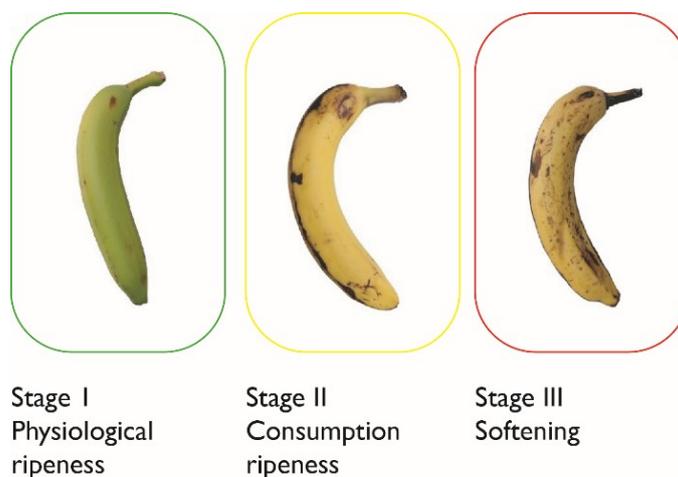

Figure 1. Banana ripening stages.

Directional lightning was employed during the digital image capture phase, which involves pointing the camera towards the object and providing illumination from the same direction as the camera, with a low angle of incidence, as shown in Figure 2. This ensures that the camera receives the maximum amount of reflected light. The light was placed at 25 cm from the fruit, while the camera was placed 30 cm, ensuring higher light intensity to enhance fruit visibility.

Using this image acquisition technique, the fruit detection neuron was trained with one thousand images of green bananas at their physiological ripeness stage.

The software was created using IDLE, an integrated environment in Python that allows multiple windows to be open simultaneously for editing. Furthermore, Python 3.11 was chosen as the programming language due to its readability and clarity, as it requires fewer lines of code compared to other languages and has an extensive standard library that covers all tasks. (Hug, 2020). OpenCV and Tensor Flow libraries were used in the project. The selection of OpenCV was due to its ability to integrate computer vision into the project through image manipulation. The reason for using the second library, TensorFlow, lies in its capacity to train the neural network and achieve pattern recognition.
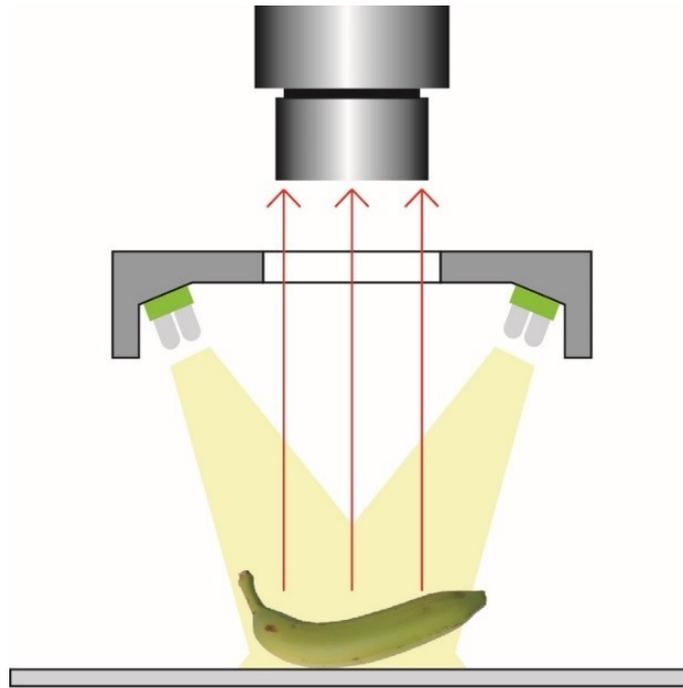
Figure 2. Control system for banana sample collection.

Figure 3 illustrates the process the algorithm undergoes, applying various artificial vision techniques that consider crucial factors such as color, area, edges, and more. At the end, the algorithm manages to detect the black spots on the banana, determining whether the product meets the requirements for export.
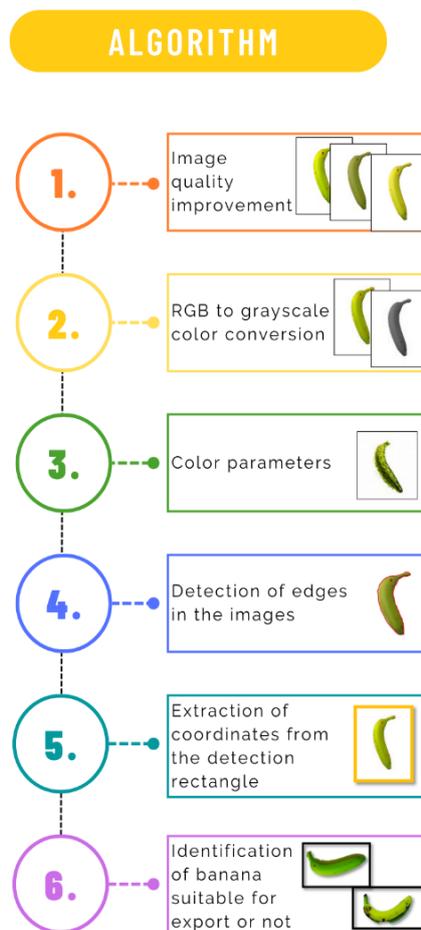

Figure 3. Algorithm Performance.

Image quality improvement

Initially, image quality enhancement was conducted to eliminate or reduce imperfections that appeared after image capture. This included correcting quality-related problems, poor illumination, and minimizing sudden changes in the color of adjacent pixels. In this process, tests were performed, and a convolution filter designed to remove undesired elements based on the properties of the figure. This filter calculates the mean value of pixels in a predefined rectangular region.

$$\mu = \frac{1}{N \times M} \sum_{i=1}^{i=n1 \times n2} V(n1, n2) \tag{1}$$

In equation (1) we present the formula which helps to improve the quality of the images, where n1 and n2 are all the pixels defined by the NxM window. Also, we work with a convolution matrix shown in equation (2) where hc produces an area-averaged filter.

$$h_c = \frac{1}{16} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix} \tag{2}$$

Therefore, in Python programming, the the code lines in Figure 4 are used to condition the image appropriately for further processing. This involves the inclusion of an additional dimension and the normalization of the pixel values in the image (Artola, 2019).

```
image = np.expand_dims(frame, axis=0)
image = imaginative('float32') / 255.0
```

Figure 4. Code for image quality improvement.

Color conversion from RGB to Grayscale

To change the colors from RGB to grayscale, a grayscale representation was created as a transformation of the RGB color space. Initially, the image is read in the RGB format, so it is necessary to perform this color conversion (Chernov et al., 2015).

This color conversion was conducted with the purpose of achieving a more precise distinction between light and dark tones, which would allow a more accurate identification of the banana irregularities, particularly in relation to the black spots that some of them may present. The RGB values are transformed to grayscale using the formula described in equation 3, which reflects the relative perception of the average brightness of red, green, and blue light.

$$NTSC: 0.299 \cdot R + 0.587 \cdot G + 0.114 \cdot B \tag{3}$$

Therefore, Figure 5 shows the code line used for the color to grayscale transformation in Python.

```
gray = cv2.cvtColor(frame, cv2.COLOR_BGR2GRAY)
```

Figure 5. RGB to Grayscale color transformation code.

Color parameters

Likewise, a set of conditions were applied to perform saturation and tone, which are shown in equation 4.

$$S = \left\{ \begin{array}{c} if(c == 0)0 \\ else\ \frac{c}{v} \end{array} \right\} \tag{4}$$

$$H = \left\{ \begin{array}{c} if(c == 0)0 \\ else\ if(V == R)1 + \frac{G-B}{6c} \\ else\ if(V == G)\frac{2}{6} + \frac{B-R}{6c} \\ else\ \frac{4}{6} + \frac{R-G}{6c} \end{array} \right\} \tag{5}$$

Additionally, a range of values was included that establishes the parameters of banana color, specifically the color black. Using the code shown in Figure 6, Python will be able to identify ripened bananas through the presence of black color, which can appear in some of these bananas.

athreshold= cv2.threshold(gray, 100, 255, cv2.THRESH_BINARY_INV)

Figure 6. Code for defining the color of the banana.

Edge and Contour Detection

An edge enhancement filter was employed for the purpose of making the edges of the image more distinct by enhancing the points that are highlighted more than others, which increases the presence of high-frequency noise. Equation 5 presents the matrix used in this process.

$$h_i = \begin{Bmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{Bmatrix} \tag{6}$$

Figure 7 shows the functions used to extract the contours.

aFindContours = findContours()

aDrawContours = drawContours()

Figure 7. Code for contour extraction.

Area extraction

Then, the area was extracted, using the vector that defines its contour. The area was determined by counting the number of pixels inside the contour (Dominguez, 2021). We can see that in equation 6 the formula for the calculation of the area is mentioned:

$$Area = \sum_{h=min(rows)}^{max(rows)} max(column(h)) - min(column(h)) \tag{7}$$

Figure 8 shows the programming lines used for the extraction of the area.

```
for Contour in Contours:
area = cv2.con
if area > 1000:
 x, y, w, h = cv2.boundingRect(Contour)
cv2.rectangle(frame, (x, y), (x+w, y+h), (0, 255, 0), 2) tourArea(Contour)
```

Figure 8. Area extraction code.

Extraction of coordinates from the detection rectangle

We proceeded to extract the coordinates of the rectangle responsible for the detection, with the purpose of cutting out and using the appropriate pixels. This network will use the previous classification together with the x (height) and y (width) positions obtained in its output convolutional layers to identify the bananas, defining the coordinates XMIN, YMIN, XMAX and YMAX. To perform this task, a bounding box is implemented, as illustrated in Figure 9.
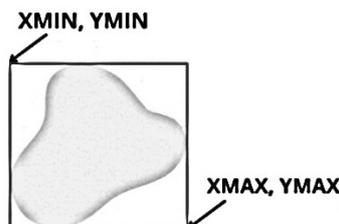
Figure 9. Square that fits the object outline.

In Figure 10, the following lines are used in Python code. The first line is responsible for adjusting the size of the box to a format suitable for prediction. The second line is used to display the frame in a window, and the last lines are used to display the predicted label inside the frame.

```
frame = cv2.resize(frame, (200, 200))
cv2.imshow('Banana Detection', frame)
if black_points > 0:
cv2.putText(frame, 'Suitable Banana', (x, y-10),
cv2.FONT_HERSHEY_SIMPLEX, 0.9, (0, 0, 255), 2)
else:
cv2.putText(frame, 'Not Suitable Banana', (x, y-10),
cv2.FONT_HERSHEY_SIMPLEX, 0.9, (0, 255, 0), 2)
```

Figure 10. Coordinate extraction code.

To sum up, the algorithm has three different stages, in the first stage, the algorithm receives the images of bananas captured with the camera and collected from the web. Then, it proceeds to the conditioning stage, which involves improving the quality of the images, converting the colors from RGB to HSV, and extracting the edges and areas needed for training the neuron. In the last stage, the algorithm generates a result that can be 'Banana not suitable' if the fruit has black spots or other signs of ripeness that indicate its discard, or 'Banana suitable' if the fruit is in optimal conditions for export.

Attributes of the data used

To ensure a rigorous and effective analysis, the study utilized a carefully selected dataset with attributes that allow for precise evaluation of the relevant characteristics of bananas intended for export. The key attributes of the data used are detailed below:

Data source

Direct Image Capture: The images were captured using a controlled lighting system and a high-resolution camera with a 48-megapixel main sensor, combined with additional sensors for wide-angle and macro detail capture.

External Data Sources: Additionally, images were gathered from online databases, ensuring a wide variety in terms of quality, lighting, and banana positioning.

Relevant attributes

Visual characteristics of the bananas:

Color: Images were captured in RGB format, but a conversion to grayscale was performed to improve the detection of dark tones, such as black spots, which indicate unsuitable ripening.

Black Spots: The images included bananas with varying degrees of black spots, as these are a crucial criterion for determining export suitability.

Green Tone: Images were classified based on the predominant green tone, which indicates the ideal ripeness for export.

Geometric properties

Size: The dimensions of the banana (length and width) were considered by extracting edges and contours.

Area: Pixel analysis was used to calculate the area of the bananas, which is important for identifying appropriately sized fruit.

Capture conditions

Lighting: Lighting was standardized using 100-watt LED lights to minimize shadows and highlight important details on the banana's surface.

Capture Distance and Angle: The camera was positioned 30cm away from the banana and at a fixed angle to ensure consistency in the captures.

Format and resolution

Image Resolution: The captured images had a resolution of 48 megapixels, providing sufficient detail for analysis.

Data Format: The standard image formats (JPEG and PNG) were used, compatible with image processing libraries like OpenCV.

Results

To carry out the evaluation of the results, the process begins with a crucial phase that consists of determining the size of the population. This stage is essential as it provides the basis for calculating the necessary sample size, a fundamental element in the research. The reason behind this initiative is to identify the appropriate quantity of bananas to be studied, which, in turns, allows the exact estimation of a specific parameter with the desired level of confidence.

It is important to emphasize that the annual banana production in each hectare of crop is set at approximately 300 bunches, which results in an average of 25 bunches per month, as previously documented. Considering that, on average, each 'Platano Manzano' bunch harbors about 35 to 40 units (Rivera, 2012), it can be calculated that each monthly crop is composed of about 1000 bananas. This figure is considered as the population of interest for this study.

Sample size calculation

$$n = \frac{N \times Z_\infty^2 \times p \times q}{e^2 \times (N-1) + Z_\infty^2 \times p \times q} \tag{8}$$

n = *Sample size*
N = *Population size*
Z = *Statistical parameter*
e = *Maximum error estimate*
p = *Probability that event occurs*
q = *Probability that the event does not occur*

Using the above formula, we proceed to calculate the sample size, which is explained in equations 8 and 9, considering a population of 1000 bananas.

$$n = \frac{1000 \times 1.96_\infty^2 \times 0.9 \times 0.1}{0.04^2 \times (1000-1) + 1.96_\infty^2 \times 0.9 \times 0.1} = 178 \tag{9}$$

$$\frac{178}{5} = 36 \tag{10}$$

In addition, equation number 7 is provided, which has a fundamental role in the calculation of the required sample size. Through a careful development process, the value 'n' was determined, which corresponds to a total of 178 units of bananas.

It is relevant to mention that during the study, the evaluation of the quantity of bananas during a period of 5 days is contemplated, with a total of 36 bananas selected in each lot daily. This process ensures the collection of solid and representative data that contribute significantly to the results of the research.

Accuracy calculation of the algorithm

To conduct a thorough evaluation of the accuracy of the algorithm used in this study, a series of data collection activities were carried out over a period of five consecutive days (Figure 11). During this process, a total of 36 bananas were imaged on each day, and the algorithm was responsible for identifying those that met the criteria for export suitability. This included the determination of bananas in a state of physiological ripeness, characterized by the display of an optimal green color, as well as the detection of those that were unsuitable due to the presence of black spots or yellow tones.

Table 1. Banana evaluation results for 5 days.

	Day 1	Day 2	Day 3	Day 4	Day 5	Mean
Correct identification	33	35	35	36	36	
Incorrect identification	3	1	1	0	0	
Total	36	36	36	36	36	
Algorithm accuracy	91,67%	97,22%	97,22%	100%	100%	97,22%

Table 1 presents the results obtained during the first day of evaluation, in which the program was able to correctly identify 33 bananas, either at a ripening stage suitable or unsuitable for export. However, it was observed that on this day, 3 bananas were not properly detected by the program. On the second and third days, tests were carried out with the same number of samples, and in both cases 35 bananas were correctly

identified, with only one identification error on each of these days. On the fourth and fifth day of evaluation, the system demonstrated a high level of accuracy by correctly identifying all 36 bananas submitted.

In addition, Table 1 provides a visual representation of the performance of the algorithm, expressed in terms of percentage accuracy. The results show a value of 91.67% for the first day, an impressive 97.22% accuracy for both the second and third days, and 100% accuracy for the last two days of evaluation. In this context, the average accuracy of the algorithm was calculated for all days evaluated, yielding a solid 97.22%. These results strongly confirm the system's capability for application in industrial environments and its effectiveness in identifying bananas suitable for export.
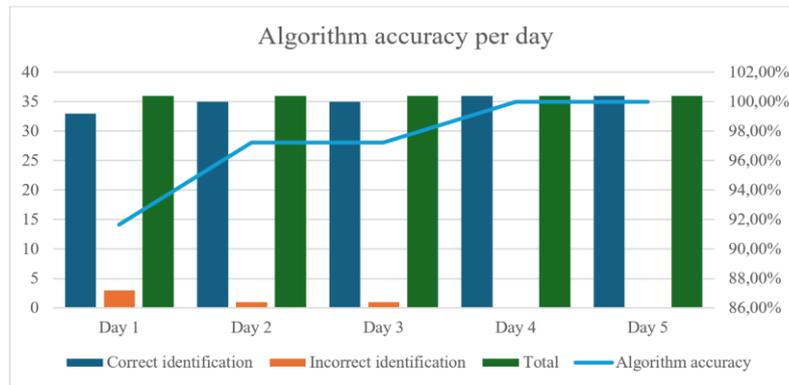

Figure 11. Algorithm accuracy analysis.

Figure 12 summarizes graphically the accuracy analysis of the algorithm used in the identification of bananas suitable for export. It is clearly seen how, as the number of days of evaluation increased, the accuracy of the algorithm results experienced a gradual and significant increase. Finally, an accuracy rate of 100% was achieved for the last two days evaluated, supporting the robustness of the algorithm in this context, with an overall average of 97.22% for all days evaluated, confirming its effectiveness in industrial applications.
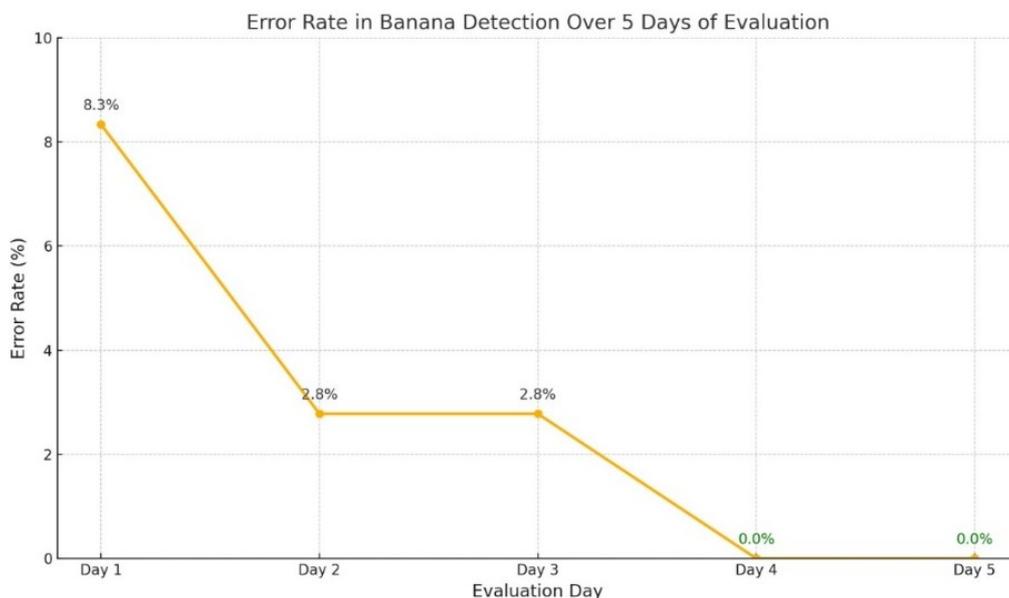

Figure 12. Banana detection error rate.

The graph shows how the banana detection error rate improved over the five days of evaluation. On Day 1, the error was 8.3%, with 3 bananas misclassified out of 36. For Days 2 and 3, the error dropped to 2.8%, where only 1 banana was misclassified. Finally, on Days 4 and 5, the algorithm achieved perfect 100% accuracy, with no errors. This progressive improvement demonstrates that the system was fine-tuned and stabilized over time, achieving reliable and consistent performance in detecting bananas suitable for export. The graph helps to clearly visualize this progress, highlighting both the progress achieved and the areas that may require adjustment in the early stages.

Discussion

The precision of the algorithm was calculated over five days using a dataset of 36 bananas each day, evaluating its ability to correctly classify the bananas as suitable or unsuitable for export. Precision was defined as the proportion of correctly classified bananas relative to the total evaluated. The results showed an increase in precision, reaching 91.67% on the first day, improving to 97.22% on the second and third days, and achieving 100% on the last two days. The average precision was 97.22%. This increase in precision is attributed to improvements in lighting conditions and system stability after initial adjustments. The initial errors were caused by variations in lighting and the difficulty of detecting minor defects, but these issues were corrected using advanced image filters. Days 4 and 5, with 100% precision, demonstrated the robustness of the algorithm under controlled conditions. Compared to previous research, the average precision of 97.22% surpasses the results reported by Figueroa and Roa (2016) (92.6%), highlighting the impact of modern techniques such as neural networks and advanced image preprocessing. High precision is crucial for the industry, as it reduces errors in the export process and improves competitiveness in international markets. The algorithm has proven to be highly effective in classifying bananas suitable for export, demonstrating its practical applicability and contribution to the advancement of artificial vision in the agricultural industry.

The tool developed in this study has been validated by comparing it with previous research in the field of artificial vision applied to fruit evaluation. In particular, the system for identification of ripeness stage in fruits described by Figueroa and Roa (2016) was taken as a reference, as well as the artificial intelligence-based fruit image classifier proposed by Heras (2017). Both systems present relevant methodological similarities that allow for a robust comparison. The selection of previous references included the work of Figueroa and Roa (2016), who employed segmentation and clustering techniques, using OpenCV and the K-means algorithm to classify passion fruit images according to their ripeness stage. We also considered the study by Heras (2017), who used supervised learning to classify fruits by feature extraction, proposing a robust methodology for comparative analysis. For direct comparison, the same images used in this study were applied to the system developed by Figueroa and Roa (2016) to evaluate its performance in classifying bananas according to their export grade.

Likewise, the process described by Heras (2017) was replicated, adapting it to the characteristics of bananas, with the objective of obtaining a comparison of results between both systems. For validation, key metrics such as accuracy, sensitivity and specificity were calculated. The results obtained, presented in Table 1 show a significant improvement in the performance of the algorithm developed in this study. The average accuracy reached 97.22%, compared to 92.6% of Figueroa and Roa's system. In addition, a 5% reduction in classification errors was achieved in our system compared to 10% in Heras' system. The improved performance is attributed to several key factors, including the implementation of advanced image processing techniques, such as the conversion from RGB to grayscale, the use of color-specific parameters, and the training of a neural model using TensorFlow. In addition, the use of a controlled illumination system and the accurate extraction of features such as areas and edges contributed significantly to the accuracy of the algorithm. The comparative analysis shows that the developed tool is not only able to match the performance of previous tools but surpasses them in accuracy and efficiency. This progress is due to the implementation of image preprocessing techniques that improve the quality of the input data, the integration of neural networks for classification, which allows greater generalization capability, and the specific design for banana classification, adapting parameters and criteria relevant to this type of agricultural product. Validation based on previous tools supports the robustness and effectiveness of the system presented, positioning it as an advanced solution for the classification of bananas for export.

Conclusion

In this study, a computer vision algorithm was developed with the main objective of identifying bananas suitable for export, focusing on key parameters such as detecting black spots, achieving the optimal green tone, and assessing fruit size. The results validate the algorithm's effectiveness, with an average accuracy of 97.22%, reaching 100% during the last days of testing. This demonstrates the system's ability to enhance consistency and efficiency in industrial sorting processes.

The implementation of the algorithm, supported by advanced image processing techniques and neural networks trained in TensorFlow, enabled the objective and reproducible evaluation of critical banana

characteristics. Additionally, the design of a controlled lighting system contributed to improving image quality, minimizing the impact of environmental variations.

The developed system represents a practical and scalable solution for the industry, reducing human errors and optimizing the selection process. This results in significant economic benefits for exporting companies by ensuring that only bananas meeting international standards are sent to the external market.

Despite the promising results, an error rate of 2.78% was detected in the initial tests, suggesting the need to optimize the model to better adapt to variations in environmental conditions and samples. For future research, it is recommended to expand the dataset and explore the use of pre-trained neural networks such as ResNet or MobileNet, as well as advanced image preprocessing techniques.

The proposed algorithm achieves its objective of automating and optimizing the selection of export-quality bananas, establishing a new standard in the application of computer vision for the agroindustry. Its large-scale implementation will contribute to improving the competitiveness of the Peruvian banana sector in the global market.

References

Arrieta, A. J., Baquero, U. M., & Barrera, J. L. (2006). Caracterización fisicoquímica del proceso de maduración del plátano 'papocho' (musa abb simmonds). *Repositorio Institucional UNAL*. https://repositorio.unal.edu.co/handle/unal/33144

Artola, A. (2019). *Clasificación de imágenes usando redes neuronales convolucionales en Python* [Dissertação de Mestrado, Universidad de Sevilla]. Repositorio de la Universidad de Sevilla. https://hdl.handle.net/11441/89506

Benito, P., Arellanes, N., & Pérez, M. E. (2015). Color and state of maturity of fruit of the husk tomato. *Agronomía Mesoamericana, 27*(1), 115–130. https://doi.org/10.15517/am.v27i1.21891

Challenger, I., Díaz, Y., & Becerra, R. (2014). El lenguaje de programación Python/The programming language Python. *Revista Ciencias Holguin, 10*(3), 1–15. https://www.redalyc.org/pdf/1815/181531232001.pdf

Chernov, V., Alander, J., & Bochko, V. (2015). Integer-based accurate conversion between RGB and HSV color spaces. *Computers & Electrical Engineering, 46*, 39–48. https://doi.org/10.1016/j.compeleceng.2015.08.005

Dominguez, T. (2021). *Visión artificial: Aplicaciones prácticas con Open CV-Python*. Anaya Multimedia.

Figueredo, G., & Ballesteros, J. (2015). Identificación del estado de madurez de las frutas con redes neuronales artificiales, una revisión. *Revista Ciencia y Agricultura, 12*(2), 65–75. https://doi.org/10.19053/01228420.4811

Figueroa, D., & Roa, E. (2016). Sistema de visión artificial para la identificación del estado de madurez de frutas (granadilla). *Redes de Ingeniería, 7*(1), 77–86. http://dx.doi.org/10.14483/udistrital.jour.redes.2016.1.a08

Gallego, V., Herrera, A., & Toro, M. (2018). *Competitividad del sector hortofrutícola en relación con la exportación del plátano en el departamento de Risaralda en los años 2018-2021* [Trabalho de conclusão de curso, Universidad Católica de Pereira]. Repositorio Institucional. http://hdl.handle.net/10785/12483

Heras, D. (2017). Clasificador de imágenes de frutas basado en inteligencia artificial. *Killkana Técnica, 1*(2), 21–28. https://doi.org/10.26871/killkana_tecnica.v1i2.79

Hug, H. (2020). Surprise: A Python library for recommender systems. *Journal of Open Source Software, 5*(52), 2174. https://doi.org/10.21105/joss.02174

Negrín, L., Barros, R., Da Silva, A., Figueiredo, A., & Cárdenas, N. (2013). Comportamiento mecánico de dos variedades de mango (Mangifera Indica) bajo compresión axial. *Revista Ciencias Técnicas Agropecuarias, 22*(1), 1–8. https://www.redalyc.org/articulo.oa?id=93225718002

Obregon, F., & Cachay, L. (2017). *Maduración controlada y color en bananos* [Tese de Doutorado, Universidad Nacional de San Martín]. Repositorio de la Universidad Nacional de San Martín. http://hdl.handle.net/11458/2499

Raschka, S., Patterson, J., & Nolet, C. (2020). Machine learning in Python: main developments and technology trends in data science, machine learning, and artificial intelligence. *Informatics, 11*(4), 193. https://doi.org/10.3390/info11040193

Rivera, M. (2012). *Manual práctico para el cultivo sustentable del plátano*. Universidad de Puerto Rico. https://www.uprm.edu/cms/index.php?a=file&fid=15184

Semenovich, M., Zotikovna, E., Vasilievna, T., Anatolyevich, M., & Alekseyevna, R. (2020). Desarrollo de un curso de programación para estudiantes de una institución de educación superior de formación docente utilizando el lenguaje de programación Python. *Propósitos y Representaciones, 8*(3), 1–12. http://dx.doi.org/10.20511/pyr2020.v8n3.484

Torres, R., Montes, E., Perez, O., & Andrade, R. (2015). Influencia del Color y Estados de Madurez sobre la Textura de Frutas Tropicales (Mango, Papaya y Plátano). *Información Tecnológica, 26*(3), 85–94. http://dx.doi.org/10.4067/S0718-07642015000300008

Viera, G. (2017). *Procesamiento de imágenes usando OpenCV aplicado en Raspberry Pi para la clasificación del cacao* [Tese de Doutorado, Universidad de Piura]. Repositorio da Universidad de Piura. https://hdl.handle.net/11042/2916

Villaverde, J., Oliveira, L., Vilela, C., Domínguez, R., Freitas, N., Cordeiro, N., Freire, C., & Silvestre, A. (2013). Compuestos de alto valor a partir de la piel inmadura de varias especies de Musa cultivadas en la isla de Madeira (Portugal). *Revista Cultivos y Productos Industriales, 46*, 49–57. https://doi.org/10.1016/j.indcrop.2012.06.037