E3S Web of Conferences **532**, 02008 (2024) https://doi.org/10.1051/e3sconf/202453202008
CIIA 2024

Fruit Grading based on Deep Learning and Active Vision System

Henry O. Velesaca[1,2,*], *Patricia L.*Suárez[1], *Dario* Carpio[1], and *Angel* Sappa[1,3]

[1]ESPOL Polytechnic University, Escuela Superior Politécnica del Litoral, ESPOL, Guayaquil-Ecuador
[2]Software Engineering Department, University of Granada, 18014, Granada, Spain
[3]Computer Vision Center, 08193-Bellaterra, Barcelona, Spain

Abstract. This paper presents a low-cost computer vision-based solution to obtain the size of fruits without contact. It consists of a low-cost webcam and a cross-shaped laser beam rigidly assembled. The proposed approach acquires and processes the images in real-time. Due to the low computational cost of the proposed algorithm, a robust solution is obtained using a frame redundancy approach, which consists in processing several frames of the same scene and hence computing a robust estimation of the fruit size. The proposed solution is evaluated with different tropical fruits (e.g., banana, avocado, dragon fruit, mamey, papaya, and taxo). Obtained results show on mean average percentage error (MAPE) below 1.50% in the computed sizes.

1 Introduction

The worldwide production of fruits for human consumption is one of the main bases of the agri-food industry. The agricultural industry is the primary sector of the economy of several countries, representing in some cases close to 10% of their Gross Domestic Product (GDP). In 2020, approximately 890 million tons of fresh fruit are produced globally. This data not only meant an increase of nearly five million tons compared to 2019 but also represented the highest production volume in the last 20 years For example, for the case study of this work, the banana is considered, which is one of the most important staple foods in tropical areas and its production for sale in local markets is, together with dairy production and horticulture, one of the few activities that provide family units with regular income throughout the year. To achieve high-quality standards, following the global trade protocol that establishes the type and quality of fruits, some approaches have recently been proposed that attempt to carry out rigorous and automated inspection of fruit size e.g., [1–3].

The automatic fruit grading process has been already tackled for citrus, apples, and other rounded-shaped fruits using machine vision systems. Nowadays, different commercial grading machines are available for such a kind of rounded shape fruit. These kinds of fruits present features (e.g., symmetry, regular shapes, single fruits) that make the automation process feasible. However, there are also other types of fruits whose measurement process requires additional steps to determine their grade.

In the particular case of the automatic evaluation of the tropical fruit grading within the post-harvest process, there is great interest on the part of the scientific community. In general,

*Corresponding author: hvelesac@espol.edu.ec

the estimation of the fruit grading is a feature necessary for the location of a certain production lot in the world market. This problem becomes more challenging in the banana case since the banana is sold in bunches of banana fingers—referred to as the hand of bananas. Banana caliber measurements are obtained using hand tools whose value is subjectively determined depending largely on trained operators. This manual and highly important process within the value chain operation requires a lot of time and cannot guarantee the integrity of the obtained data due to the difference in the evaluation capacity of each operator e.g., [4–6].

Due to the importance of obtaining fruit characteristics such as size, color, and defects, among others, for determining the price of the product in local and international markets following the standards and guidelines of fair trade, several works have been proposed to obtain these values automatically; the goal is to remove the subjectivity of operator when doing the measurements. For example, the authors [7] present a machine vision system for automatic quality grading of fruit where oranges, peaches, and apples are used as a case study and quality attributes such as size, color, stem location, and detection of external blemishes are obtained. On the other hand, the author [8] presents a computer vision system for the determination of external quality characteristics such as size, shape, intensity, defects, and flabbiness for date fruits. In addition, the authors [4] implemented an automatic algorithm based on computer vision to determine three indicators of banana size, namely length, ventral straight length, and arch height, which are used to determine banana quality. Contrary to previous approaches, the work presented by the authors [9–11], use a hyperspectral system to inspect the internal and external quality of fruits and vegetables.

The current work proposes a low-cost computer vision-based approach for fruit caliber estimation. It works by using images from a commercial webcam together with a laser beam. The proposed pipeline consists first of calibrating the camera to compute intrinsic parameters; secondly, the extrinsic parameters that relate the camera reference system with the laser beam reference system are estimated; thirdly, once intrinsic and extrinsic parameters are computed, image processing algorithms are implemented to extract the cross-shaped pattern from the input images; finally, with all previous information, the caliber of the fruit is estimated.

The manuscript is organized as follows. Section 2 presents works related to the vision-based fruit grading problem. Section 3 presents details of the pipeline proposed for estimating fruit size without contact. Experimental results with different fruits are depicted in Section 4. Finally, conclusions and future work are given in Section 5.

2 Related Work

This research work presents an innovative approach for estimating fruit caliber using a non-contact method. Unlike traditional systems that require a predefined distance between the measurement apparatus and the fruit, our proposed system functions effectively across varying distances and in uncontrolled lighting conditions, making it suitable for different types of fruits. The system employs a low-cost setup, consisting of a webcam and a laser beam, to facilitate this task.

Within the value chain of the agricultural industry, the fruit grading process is a fundamental task for determining the marketing price in both national and international markets, which is why the industry has used automated fruit grading systems to streamline the process significantly. For example, the author [12] presents a grading robot system, that automatically provides apples from containers and inspects all sides of the fruits, facilitating the work of the operators in addition to reducing contact with the fruits. In operation, the grading robot is capable of handling 12 fruits simultaneously, quickly and efficiently capturing images of each fruit's bottom as it moves them to the conveyor line. This process is complemented by an additional capture of four side images per fruit, achieved by rotating the fruits 270°

on the suction pads. This thorough imaging process is critical for ensuring comprehensive fruit inspection from all angles. The efficiency of the system is notable, with the manipulators completing a full cycle—from fruit pick-up to image capture to placement on the conveyor—in approximately 4.3 seconds. This speed enables the robot to grade an impressive three fruits per second. Additionally, the authors [13] propose a prototype of an automated fruit grading system to detect the defects on the surface of apples and mangoes, improving the response time and the quality of production batches by not depending on the subjectivity of the operator. Also trying to minimize the contact time between the operator and fruit, the authors [14] propose a computer vision-based fruit grading system for the quality evaluation of tomatoes, the system is designed to avoid contact with very delicate vegetables. Generally, these traditional solutions are supported by the regular and constant shape of the fruits under study (e.g., oranges, apples, grapefruit, among others), in addition to processing a single fruit at a time without considering the scenario of fruits in clusters.

However, these traditional systems often rely on the uniform shape of fruits like oranges and apples and typically handle one fruit at a time, not accounting for clustered fruits. This limitation presents a research gap for developing systems capable of processing diverse fruit shapes and configurations. For example, Shiddiq et al. [15] utilized a computer vision method with a cross-line laser beam to measure the volumes of clustered oil palm fruits, indicating progress toward handling more complex fruit arrangements.

Among the traditional approaches, Jana et al. [16] propose a method for measuring fruit volume and mass, such as those proposed, who used polynomial-based edge detection and integration for volume calculation, show high precision in volume and mass estimation for fruits like potatoes, lemons, and tomatoes. However, there remains a gap in adapting these methods for a broader range of fruit types and in more variable conditions. The proposed approach consists of the usage of traditional segmentation techniques such as thresholding; as a result of this segmentation process, the edge points are obtained, which are approximated using a polynomial expression. The volume is obtained by integrating the polynomial while the mass is computed using the relationship of density, mass, and volume. The technique is validated by comparing the value measured using standard methods of volume and mass estimation of fruits and vegetables. The volume estimation precision for potatoes, lemons, and tomatoes are 92.54%, 88.82%, and 89.02%, respectively, while the mass estimation precision using the proposed approach for potatoes, lemons, and tomatoes are 92.98%, 89.31%, and 88.56% respectively.

Similarly to the approach proposed in the current paper, the authors [15] use a computer vision method with a cross-line laser beam to determine different key points in the image that will serve to measure the volumes of oil palm using a segmentation process and an ellipsoid equation. In another work, the authors [17] implement different algorithms for classifying tomatoes according to the size into small, medium, and large from a given image. The authors use the combination of image processing, thresholding, machine learning, and deep learning techniques, where for the machine learning method, KNN registered the highest accuracy —about 97.50%. Similar to the previous approach, the authors [18] use machine learning techniques for non-destructive internal quality inspection of pear. The authors use a combination of machine learning and X-ray computed tomography. Trained SVMs achieve good classification accuracy ranging between 90.2% and 95.1% depending on the cultivar and the number of features that are used.

In recent years several techniques have been proposed to determine the grade of different fruits. Among the most common techniques are computer vision techniques where, with the help of images and video, different quality parameters such as size, freshness, and defects, among others, can be determined. For example, the authors [19] implemented an automatic electronic vision-based system for sorting and grading fruits, like mango, based on their ma-

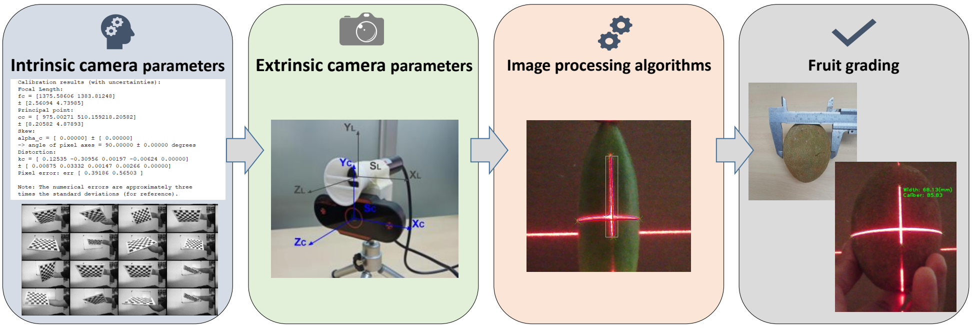

Figure 1. The pipeline of the proposed approach, from the camera calibration till the fruits grading estimation process.

turity level and size. The authors also use fuzzy logic techniques in the proposed work. On the other hand, the authors [20] use non-destructive techniques for the detection of pineapple surface quality, and a visible and near-infrared (VIS/NIR) spectrum-based platform is proposed for optimized detection of pineapple translucency.

Recently, the usage of deep learning-based techniques has been considered to improve the performance of implemented systems. For example, the authors [21] propose the use of a model based on deep learning (e.g, SqueezeNet) together with a non-destructive system for mango sorting and grading, in addition to the use of RGB and thermal images. On the other hand, the authors [22] use a deep neural network such as ResNet50 to determine quality properties such as freshness, tenderness, color, and shape, among others, obtaining an accuracy of 99%. Similar to the previous approach, the authors [23] use the ResNet50 to classify bananas into good and defective classes. The authors also generate a dataset for the implementation of the deep neural network. The results of the evaluation of the designed system have shown a great capacity for generalization when they are tested on images of unseen bananas and obtained an accuracy of 99%. Focusing on the fruit and vegetable size estimation, Zheng et al. [24] propose a system based on a stereo camera that relies on key points detection. On the other hand, Fukui et al. [25] implement a robot that autonomously searches for an appropriate measurement position that uses 3D point cloud for estimating the size of tomatoes. Another example, similar to the previous approach, is Gongal et al. [26] which uses a 3D machine vision system to estimate the size of apples.

Finally, while significant progress has been made in automating fruit grading, existing systems often lack the flexibility to handle unstructured environments and diverse fruit types effectively. Our approach focuses on these gaps by developing a versatile, low-cost grading system capable of operating under varied conditions without direct contact with the fruit. This approach not only preserves the fruits' physical integrity but also enhances fruit grading systems' scalability to accommodate different agricultural settings.

3 Proposed Approach

This section details the different stages of the proposed approach. Firstly, the camera calibration process, both intrinsic and extrinsic parameters, is detailed; then the segmentation algorithm is used for extracting feature points, and finally, the fruit grading estimation is presented. Figure 1 shows the pipeline of the proposed approach. In our approach, the chal-

Table 1. Intrinsic camera parameters—focal length in pixels: F_x (focal on the X axis), F_y (focal on the Y axis); principal points in pixels: C_x (center of the image on the X axis), C_y (center of the image on the y axis)

Image		**Focal length**		**Principal points**	
Width	**Height**	F_x	F_y	C_X	C_y
1920	1080	1375.59	1383.81	975.00	510.16

lenges posed by camera distortions in fruit grading systems are addressed by integrating a comprehensive approach that includes calibration, multi-frame analysis, and empirical adjustments, the system ensures high accuracy and robustness, essential for commercial applications in agriculture where quality and precision directly impact profitability and operational efficiency.

3.1 Intrinsic camera parameters

Intrinsic camera parameters—i.e., focal length, central point, and distortion parameters—are computed by using the pinhole camera model (Eq. 1). The well-known pinhole model describes the mathematical relationship between the coordinates of a point in three-dimensional space and its projection onto the image plane of an ideal pinhole camera [27]. The calibration has been performed by taking several images of a checkerboard calibration pattern and using the camera calibration toolbox from Matlab[1]. The values obtained in this step are very important for subsequent processes since the precision obtained in the final results will depend on these initial values. Table 1 shows the results of the calibration obtained for the low-cost camera. Distortion parameters are not considered.

$$u = C_x + F_x * \frac{X}{Z}, \qquad v = C_y + F_y * \frac{Y}{Z}, \tag{1}$$

where (u, v) are the coordinates on the 2D image plane, where u is the horizontal and v is the vertical coordinate, C_x, C_y are the principal point coordinates of the camera, representing the offset from the top left corner of the image sensor to the optical center, in pixels. Also, (F_x, F_y) are the focal lengths of the camera expressed in pixels, for the x-axis and y-axis respectively, and X, Y, Z are the coordinates of a point in 3D space relative to the camera's coordinate system, where Z is the depth or distance from the camera along its viewing direction.

3.2 Extrinsic camera parameters

To make the extraction of feature points required for fruit grading, a laser beam is used. It has an output power of 5mw with a wavelength of 650nm and an operating voltage of 3 to 5V. The most important feature of the pointer is the laser beam pattern it projects, which is in the shape of a cross. This feature allows easy detection of the boundary of the fruit and extracts key points needed to measure the fruit's caliber. These boundary points are crucial for computing the fruit's size and are detected via the laser's cross pattern projected onto the fruit. The laser delineates clear, measurable boundaries by creating parallel segments on the fruit's surface, which the camera system captures. These captured images are then processed to extract feature points along the laser-generated lines, ensuring precise and consistent measurements across different fruits. As mentioned above the acquisition system consists of a low-cost full

[1] http://robots.stanford.edu/cs223b04/JeanYvesCalib/

HD **webcam** with a USB connection. The laser beam and webcam are rigidly attached trying to minimize the distance between their reference systems. Hence, in addition to the intrinsic camera parameters, the relationship between the reference systems of the camera and laser is estimated by taking several images of a planar surface at different depths and then obtaining the relationships between depth, size of the projected segment in 3D space (mm) and size of the projected segment in 2D space (pixels). Figure 2 *(left)* shows components used in the image acquisition system.

3.3 Image processing algorithms

Once the image acquisition system is calibrated, both intrinsic and extrinsic parameters, the image processing algorithms that allow first extract the cross-shaped image and then detection feature points are implemented. Firstly, the RGB image is segmented to get the laser projection in the fruit surface. Just the red channel is considered to apply the segmentation threshold. Subsequently, the key points, **P1** and **P2** (i.e., the intersection of the laser with the contour of the fruit to be measured), are extracted to be used in future calculations.

Figure 2 *(a)* shows components used for the image acquisition, in addition the camera reference system (X_C, Y_C, Z_C, S_C) and the laser beam reference system (X_L, Y_L, Z_L, S_L) are presented. Figure 2 *(b)* shows an illustration of the points that would be obtained after the segmentation when the case study is considered. The extraction of key points without the usage of the last pointer is also evaluated, however, it is not that feasible to place the acquisition system in the correct position (i.e., centered concerning the fruit shape and orthogonal to the plane tangent at the fruit surface in the laser projection point). The **P1** and **P2** points are the extreme points at the image plane, however, these points do not correspond with the left and right points in the 3D space required to measure the fruit caliber. This difference is considered in the next stage, once the fruit grading is estimated. It should be noticed that the usage of the laser beam is additionally required in the case of the hand of bananas—the technician using the proposed approach needs some visual feedback about the banana's finger being measured.

The depth value (i.e., Z in Eq. 1) is important to be used within the pinhole model and subsequently obtain the caliber of the fruit. To estimate the depth value (i.e., distance from the fruit to the camera) shown in Fig. 4, the technique MiDaS [28] is used. MiDaS is a deep learning-based approach that outputs the relative depth from the camera to the objects in the scene taking as input a single RGB image (see Fig. 3), to convert from relative to absolute depth a regression model must be found, this achieves positioning several objects in the scene at know depths and finding the linear transformation that best maps the relative depth to the absolute depth [28]. The depth estimation is performed using distances ranging from 10 to 50 cm.

3.4 Fruit grading

Finally, with camera intrinsic and extrinsic parameters together with feature points detected in the image plane, the fruit grading estimation is performed. It just consists of first estimating the distance from the camera to the fruit under study using the relationship obtained in Section 3.3, after obtaining the depth value, the width of the fruit, in millimeters, is estimated by using the intrinsic camera calibration information.

To obtain a robust measurement, the fruit grading is not obtained from the measurement of a single frame but is estimated from a set of 90 consecutive estimates (the system works at a frame rate of 30 frames per second). To carry out this task, the median and average of these measurements are used as metrics, and as expected, the best results have been obtained

 https://doi.org/10.1051/e3sconf/202453202008

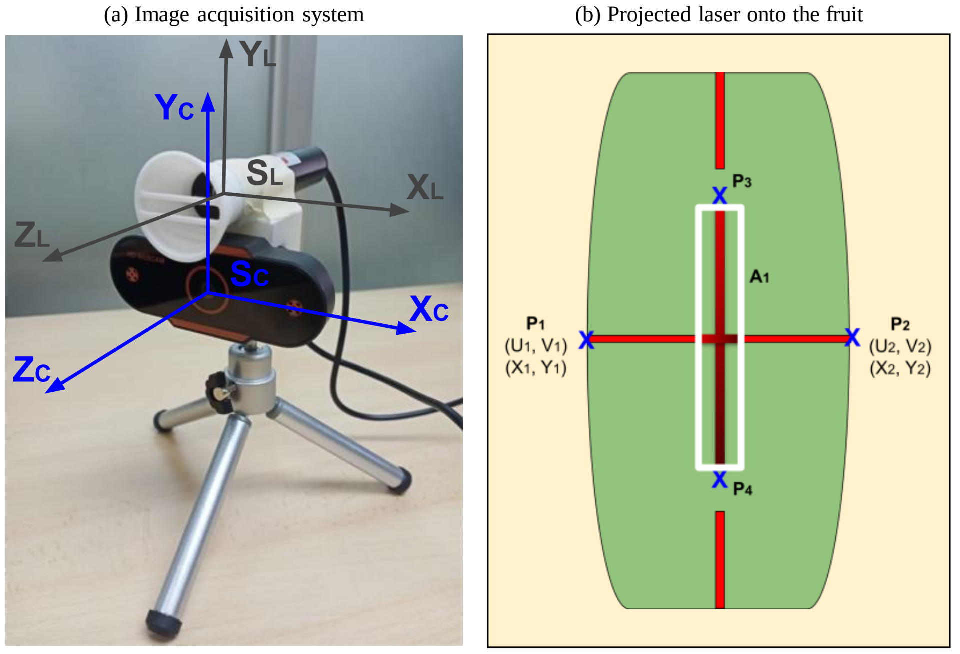

Figure 2. Proposed fruit grading vision system. (*a*) Components used for the image acquisition. (*b*) Illustration of the projected laser (i.e., cross-shaped segments) onto the fruit surface.

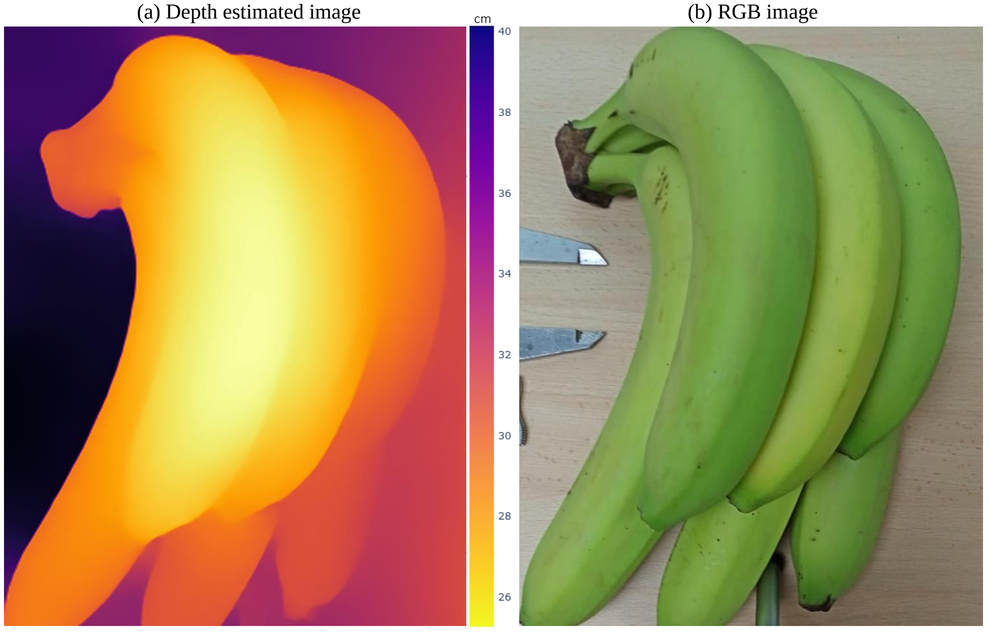

Figure 3. Depth estimated with MiDas [28] and RGB image.

using the median value. Finally, a correction factor is applied to the value obtained in the previous step. The correction factor depends on the fruit size and curvature and it is used to correct the position of projected feature points used to estimate the fruit grading. It is important to emphasize that the correction factor is a fixed value or differs depending on the fruit; it could depend on the complexity of the classification system and the variability of the fruits being evaluated. In some cases, a single correction factor might be sufficient for all fruits within a certain range of sizes and shapes. However, for more accurate classification or when dealing with highly variable fruits, it may be necessary to have different correction factors for different categories of fruits or even individual fruits. This would involve a more sophisticated approach to determining the appropriate correction factor for each specific case. Figure 4 shows illustrations where it can be appreciated the difference between key point position (i.e., points from ray tracing process) and points used for computing fruit size (i.e., points from parallel segments). The correction factor is empirically obtained.

4 Experimental Results

This section presents experimental results obtained using the proposed pipeline. First, the results of a case study that evaluates the performance of the proposed strategy using bananas are presented. This case study consists of the evaluation of a set of several banana fingers of different sizes, the obtained results are considered a benchmark for this work. Then other types of tropical fruits (e.g., avocado, dragon fruit, mamey, papaya, and taxo) are included in the evaluation.

4.1 Case Study

The results of the evaluation of the proposed system for bananas as a case study are presented in this section. Figure 5 illustrates a banana finger grading from a cluster of bananas using the low-cost system proposed in this work. The first step in the grading is the selection of a banana finger *(left)* from the bunch. The second step is caliber estimation using the proposed platform, the fruit width value and caliber are displayed to the system operator. Finally, the value obtained is validated using a caliper *(right)* to determine the real caliber of the banana. The percentage of error between the physical measurement (measures obtained by five experts) and the measurement estimated from the proposed work in a large set of bananas is depicted in Table 2 *(1st row)*. On average, in the case of bananas, the error is 1.31% with a sample set of 30 bananas. In this illustration, it can be appreciated the challenge of banana finger grading using a caliper. On the one hand, it is a time-consuming process that requires touching the fruit, which can damage the surface of it. These two drawbacks are avoided using the proposed approach.

4.2 Results in other fruits

The case study presented in the previous section is the main motivation of the proposed approach, however, it can be used with other fruits. In order to evaluate its usefulness, different measurements are made on a variety of tropical fruits. Fruits such as avocados, dragon fruit, mamey, papaya, and taxo are considered by their irregular shape and lack of commercial products to compute their grading. Results with an average error of less than 1.5% in fruits such as mamey and taxo are obtained. On the other hand, results with an average error between 7% to 8% in fruits such as avocado, dragon fruit, and papaya are obtained. It should be mentioned that the results are obtained without any specific tuning of the vision system for

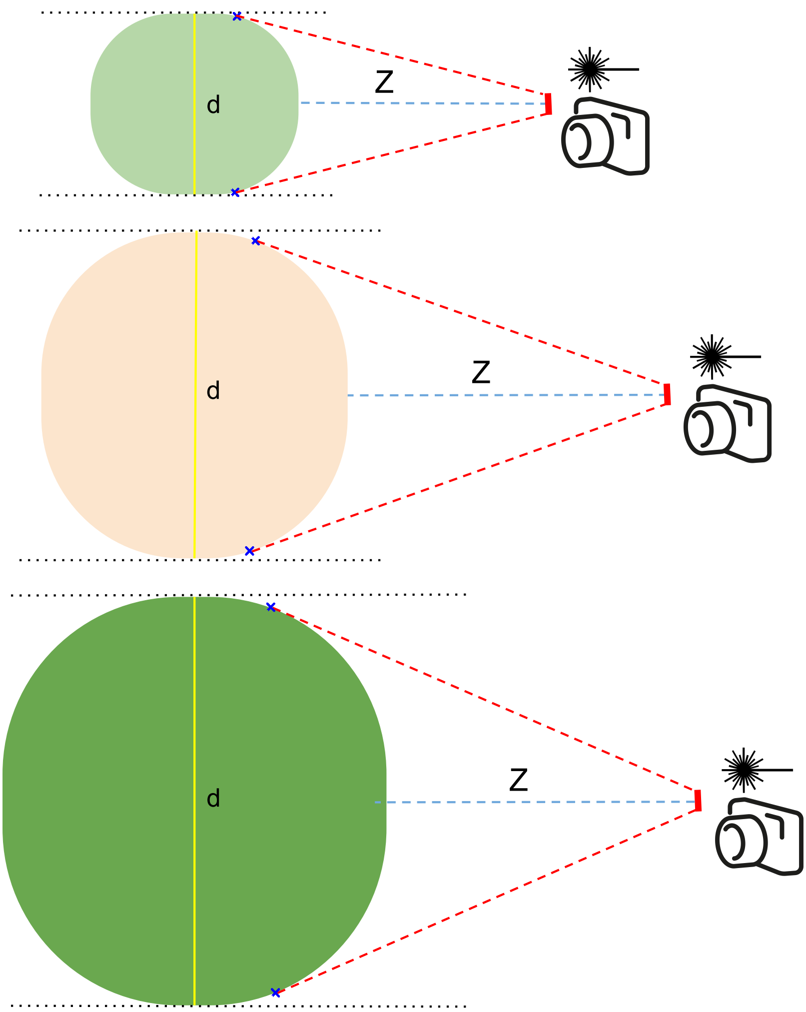

Figure 4. Illustration showing representation schemes of curvatures of different sizes of fruits, with parallel lines indicating manual measurements obtained by a caliper, and points of tangency from the laser projection obtained by the segmentation algorithms.

each particular fruit; in other words, results would be considerably improved if factors such as color, texture, and surface reflection of the fruit are considered. For example, fruits such as taxo and mamey, in which the surface of the peel is opaque and does not reflect light, present similar results to the case study, although the color is different. While brighter surfaces and non-contrasting colors with the red laser beam such as avocado, dragon fruit, and papaya have errors greater than 5% due to the nature of the peeling texture. Based on the results obtained with the fruits under study, it is possible to identify that better results are obtained with fruits whose texture, color, and reflectance are similar to the reference fruit such as bananas, for

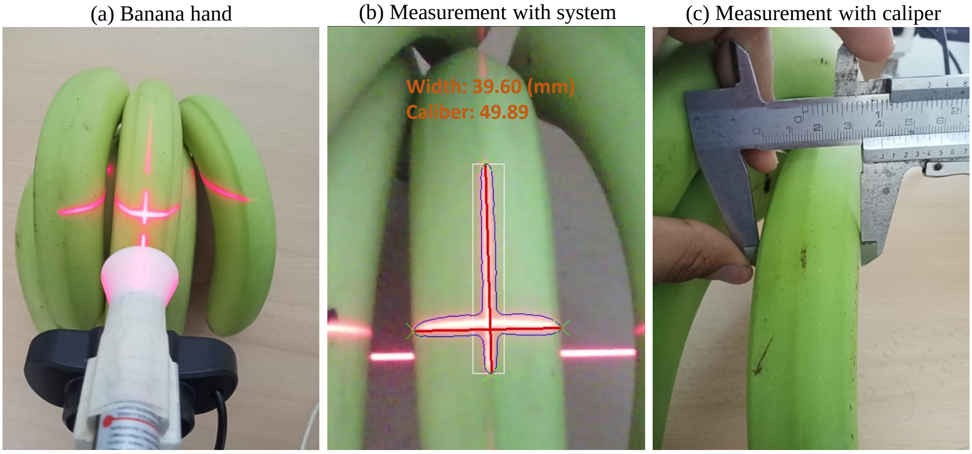

Figure 5. Illustration of a banana finger measurement from a hand of bananas, where it can be appreciated the challenge of fruit grading

example in the case of avocado, dragon fruit and papaya, average errors greater than 7% are obtained because of the surface of the fruit had yellow/orange tones and also had a partially shiny surface, which caused the laser beam to not be seen by the camera and segmentation algorithms not detect the edge correctly. While in the mamey and taxo fruits, the peel did not present this shine and is more similar to the texture of the banana, although with a different color.

Table 2 *(2nd to 6th rows)* shows the experimental results using another type of fruit. Figure 6 shows illustrations of the laser projection and feature points obtained in the evaluation fruits. The values obtained with the proposed approach use a sample of 30 fruits for each study—ground truth values are obtained by five experts. Table 3 presents comparisons between the results of the proposed approach and results from state of the art techniques i.e., [24–26] in different fruits and vegetables using the Mean Average Percentage Errors (MAPE). MAPE calculates the average of the absolute percentage errors between ground truth values and estimated values. It is expressed as a percentage and provides a measure of the average magnitude of errors in forecasting. Although none of these works is indeed to be used on the same fruits of our study, the comparison is made based on the efficiency of each approach on MAPE metric, where it can be observed that with a low-cost monocular system, it is possible to obtain very good results.

5 Conclusions

This paper proposes a low-cost system for non-contact fruit grading using image processing techniques, a cross-shaped laser beam, and a low-cost webcam. Table 2 show the results obtained in the present work. Bananas are considered as a case study and as a reference fruit for the validation of the proposed platform. For this case study, an average error of less than 1.50% is obtained. In addition, other types of fruits are evaluated, obtaining a similar average error of less than 1.50% in mamey and taxo, and an average error between 7% to 8% in fruits

Table 2. Mean Average Percentage Errors (MAPE) obtained for each of the evaluation fruits in the experimental studies.

#	Fruit	% Error by expert					MAPE(%)
		#1	**#2**	**#3**	**#4**	**#5**	
1	Banana	1.10	1.30	1.16	1.58	1.39	1.31
2	Avocado	9.62	6.39	6.62	7.36	5.45	7.09
3	Dragon fruit	7.07	6.01	6.15	5.97	6.85	6.41
4	Mamey	1.26	1.48	1.15	1.21	1.56	1.33
5	Papaya	7.92	7.71	6.36	5.81	8.02	7.16
6	Taxo	0.42	1.40	1.38	1.25	0.97	1.08

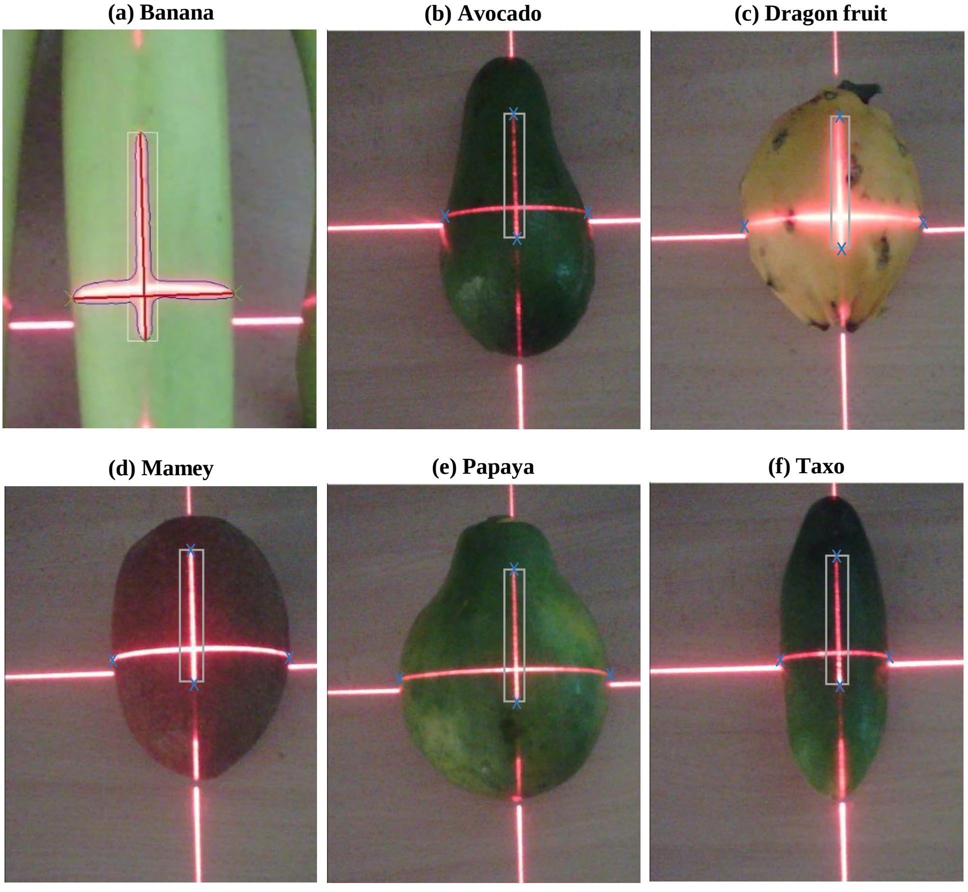

Figure 6. Snapshots of the proposed approach when used with different fruits.

Table 3. Comparisons from results of the state-of-the-art approaches on the size estimation of different fruits.

#	Fruit	Method	MAPE(%)
1	Taxo **(our)**	Monocular + Laser beam	1.08
2	Mamey **(our)**	Monocular + Laser beam	1.33
3	Banana **(our)**	Monocular + Laser beam	1.31
4	Tomato (Zheng et al.[24])	Stereo vision + Key points	1.99
5	Eggplant (Zheng et al.[24])	Stereo vision + Key points	2.15
6	Pepper (Zheng et al.[24])	Stereo vision + Key points	2.84
7	Cucumber (Zheng et al.[24])	Stereo vision + Key points	3.56
8	Apple (Fukui et al. [25])	3D Point cloud	5.00
9	Dragon fruit **(our)**	Monocular + Laser beam	6.41
10	Avocado **(our)**	Monocular + Laser beam	7.09
11	Papaya **(our)**	Monocular + Laser beam	7.16
12	Apple (Gongal et al. [26])	3D Vision system	15.20

such as avocado, dragon fruit, and papaya. Fruits with textures, colors, and reflectance similar to bananas achieve better segmentation results, while those with yellow/orange tones and shiny surfaces, like avocado, dragon fruit, and papaya, incur errors exceeding 7%. In contrast, mamey and taxo fruits, resembling bananas in texture despite different colors, exhibit fewer segmentation errors, highlighting texture similarity's significance. Other determining factors in obtaining a correct size value are the characteristics of the fruit under study and the contrast generated between the surface of the fruit peel and the laser light beam. Because there are fruits that have a natural or artificially added shine, in addition, there are fruits with tones and colors that in the visible spectrum are close to the red color band, a characteristic used in segmentation algorithms. Comparisons with results from state-of-the-art approaches show that the proposed approach reaches better results on the MAPE for the tested fruits. As a future work, it is proposed to improve the segmentation algorithms to delimit with greater precision the cross formed by the laser taking into account the different textures and colors of fruits, and obtain the feature points that allow estimating with greater accuracy the caliber of the greatest variety of fruits.

Acknowledgment

This work has been partially supported by: the ESPOL project "Estimación del diámetro usando imágenes 3D" (CIDIS-05-2021).

References

1. M. Khojastehnazhand, M. Omid, A. Tabatabaeefar et al., Development of a lemon sorting system based on color and size, African Journal of Plant Science **4**, 122 (2010).
2. X. Liming, Z. Yanchao, Automated strawberry grading system based on image processing, Computers and Electronics in Agriculture **71**, 32 (2010).
3. H. Maeda, T. Akagi, R. Tao, Quantitative characterization of fruit shape and its differentiation pattern in diverse persimmon (diospyros kaki) cultivars, Scientia Horticulturae **228**, 41 (2018).
4. M.H. Hu, Q.L. Dong, P.K. Malakar, B.L. Liu, G.K. Jaganathan, Determining banana size based on computer vision, Int. Journal of Food Properties **18**, 508 (2015).
5. J. Ehiem, K. Simonyan, Physical properties of wild mango fruit and nut, International Agrophysics **26** (2012).

6. M.R. Satpute, S.M. Jagdale, Automatic fruit quality inspection system, in *Int. Conf. on Inventive Computation Technologies* (IEEE, 2016), Vol. 1, pp. 1–4
7. J. Blasco, N. Aleixos, E. Moltó, Machine vision system for automatic quality grading of fruit, Biosystems Engineering **85**, 415 (2003).
8. Y. Al Ohali, Computer vision based date fruit grading system: Design and implementation, Journal of King Saud University-Computer and Information Sciences **23**, 29 (2011).
9. G. ElMasry, N. Wang, A. ElSayed, M. Ngadi, Hyperspectral imaging for nondestructive determination of some quality attributes for strawberry, Journal of Food Engineering **81**, 98 (2007).
10. G. ElMasry, N. Wang, C. Vigneault, J. Qiao, A. ElSayed, Early detection of apple bruises on different background colors using hyperspectral imaging, LWT-Food Science and Technology **41**, 337 (2008).
11. S. Cubero, N. Aleixos, E. Moltó, J. Gómez-Sanchis, J. Blasco, Advances in machine vision applications for automatic inspection and quality evaluation of fruits and vegetables, Food and Bioprocess Technology **4**, 487 (2011).
12. N. Kondo, Robotization in fruit grading system, Sensing and Instrumentation for Food Quality and Safety **3**, 81 (2009).
13. M.A. Ali, K.W. Thai, Automated fruit grading system, in *Int. Symposium in Robotics and Manufacturing Automation* (IEEE, 2017), pp. 1–6
14. M.P. Arakeri et al., Computer vision based fruit grading system for quality evaluation of tomato in agriculture industry, Procedia Computer Science **79**, 426 (2016).
15. M. Shiddiq, Herman, D.S. Arief, O.P. Yesyaf, C. Lase, I.R. Husein, S. Ayunda, Volume measurement of oil palm fresh fruit bunches using computer vision, in *AIP Conference Proceedings* (AIP Publishing LLC, 2022), Vol. 2663, p. 030007
16. S. Jana, R. Parekh, B. Sarkar, A de novo approach for automatic volume and mass estimation of fruits and vegetables, Optik **200**, 1 (2020).
17. R.G. de Luna, E.P. Dadios, A.A. Bandala, R.R.P. Vicerra, Size classification of tomato fruit using thresholding, machine learning, and deep learning techniques, Journal of Agricultural Science **41**, 586 (2019).
18. T. Van De Looverbosch, M.H.R. Bhuiyan, P. Verboven, M. Dierick, D. Van Loo, J. De Beenbouwer, J. Sijbers, B. Nicolaï, Nondestructive internal quality inspection of pear fruit by x-ray ct using machine learning, Food Control **113**, 1 (2020).
19. C. Nandi, B. Tudu, C. Koley, in *Sensing technology: current status and future trends II* (Springer, 2014), pp. 27–46
20. S. Xu, J. Ren, H. Lu, X. Wang, X. Sun, X. Liang, Nondestructive detection and grading of flesh translucency in pineapples with visible and near-infrared spectroscopy, Postharvest Biology and Technology **192**, 1 (2022).
21. V. Bhole, A. Kumar, Mango quality grading using deep learning technique: perspectives from agriculture and food industry, in *Proceedings of the 21st Annual Conf. on Information Technology education* (2020), pp. 180–186
22. M.M. Raikar, S. Meena, C. Kuchanur, S. Girraddi, P. Benagi, Classification and grading of okra-ladies finger using deep learning, Procedia Computer Science **171**, 2380 (2020).
23. A. Helwan, M.K. Sallam Ma'aitah, R.H. Abiyev, S. Uzelaltinbulat, B. Sonyel, Deep learning based on residual networks for automatic sorting of bananas, Journal of Food Quality **2021**, 1 (2021).
24. B. Zheng, G. Sun, Z. Meng, R. Nan, Vegetable size measurement based on stereo camera and keypoints detection, Sensors **22**, 1617 (2022).

25. R. Fukui, K. Kawae, S. Warisawa, Development of a tomato volume estimating robot that autonomously searches an appropriate measurement position–basic feasibility study using a tomato bed mock-up–, Journal of Robotics and Mechatronics **30**, 173 (2018).
26. A. Gongal, M. Karkee, S. Amatya, Apple fruit size estimation using a 3d machine vision system, Information Processing in Agriculture **5**, 498 (2018).
27. K.M. Dawson-Howe, D. Vernon, Simple pinhole camera calibration, Int. Journal of Imaging Systems and Technology **5**, 1 (1994).
28. R. Ranftl, K. Lasinger, D. Hafner, K. Schindler, V. Koltun, Towards robust monocular depth estimation: Mixing datasets for zero-shot cross-dataset transfer (2019)