

Training of a Convolutional Neural Network for the Classification of Coffee Fruit State

Fabricio Varela Marín Alvaro^{1,*}, *Timur Gadzhievich Aygumov*², and *Movsar Musaevich Matygov*³

¹FI, Universidad Tecnológica Centroamericana (UNITEC), San Pedro Sula, Honduras

²Dagestan State Technical University, Makhachkala, Russia

³Kadyrov Chechen State University, Grozny, Russia

Abstract. Convolutional neural networks are recognized for their high artificial intelligence capacity, mostly for their identification of objects either in images or videos. Honduras has a strong dependence on coffee cultivation, as it plays a crucial role in both its culture and economy. However, Honduras has endured difficulties linked to aspects such as quality, contamination, processing problems, illnesses, or lack of training. The root of problems often arises during the coffee harvest, and since Honduras is a less developed country, the use of advanced technology is not common. The recurring research aims to train a cnn model by using images and object detection to classify the quality of the coffee fruit. The training of a convolutional neural network was developed by using Robo Flow for the classification of the state of the coffee fruit. Firstly, the coffee farm was visited to see firsthand the fruit in its various stages. Due to the seasonal cuts established on the farms, a second visit was made to collect the coffee fruit. After that, the fruit was taken into a controlled environment to start the photo session. The session concluded with a total of 1702 photographs. The fruits within the images were individually annotated by classifying them into two categories: good and bad. A model result gave a precision, recall and mpa of 97.7%, 94.1% and 98.3%, respectively. The configurations were the use of 600 images, using a ratio of 80:10:10 for the division of training, validation, and testing tasks. Only one adjustment was made in the preprocessing, which consisted of changing the size of the images from 3000x4000 to 640x640. This is considered the best result after carrying out more training to test the alteration of the results of other models by making changes such as the number of images and the types of augmentations. Therefore, the results conclude the fulfillment of the objective.

1 Introduction

Convolutional Neural Networks (cnns) are an artificial intelligence algorithm that has achieved various purposes due to their ability to interact with a visual database such as images, graphs, or videos. Cnns are primarily renowned for their potential in classifying

*Corresponding author: matygov.movsar@gmail.com

objects within the visual field, facial recognition, transfer learning. This technology is not commonly used in industries in underdeveloped countries like Honduras [1]. Farms in Honduras experience seasons for harvesting their coffee fruits [2, 3]. Coffee holds significant importance for Honduras as it is a part of the culture and the economy specially in small communities [4, 5]. Therefore, a cnn model will be trained to identify and classify coffee cherries according to their respective classes. To achieve this objective, a proprietary database of images will be created in a controlled environment. These images will then be input into a cnn for accurate annotations, aiming to obtain favorable results.

1.1 State of the Art

1.1.1 The coffee

The classification of coffee quality may vary since it depends on the standardizations proposed by country, such as Indonesia. However, these customs have a high margin of error since they cause fatigue and irritation for humans [1]. CENICAFE (National Coffee Research Center in Colombia) states that the time it takes for a coffee fruit to mature is approximately 7 to 8 months, as shown in Fig. 1. Their coffee quality classification depends on indicators such as size, color and chemical state. [2]. Research that involves making use of the coffee fruit is complex since it must be carried out in the correct season of the year.

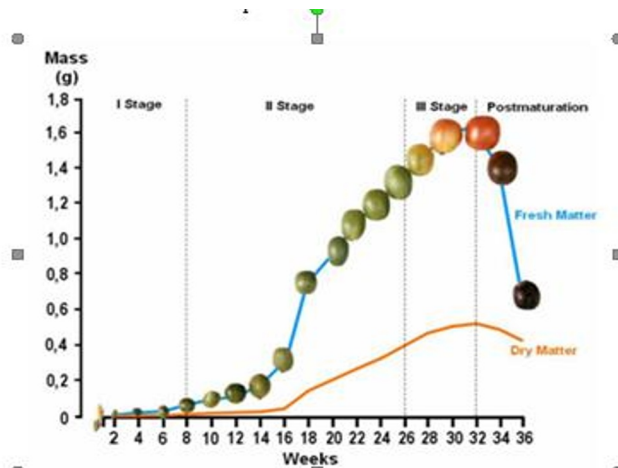


Fig. 1. Graph illustrating the stages of coffee fruit development [3].

1.1.2 Database

According to the study that developed a database for analysis of defects in apples. The selected images were already infected apples with a resolution of 4000x3000 because they used a Fujifil FinePix F60fd camera. They used a white background and sunlight in most of the photos. However, the resolution of the images were adjusted due to the manual change in which they converted the images to 256x256 pixels [4]. Research on a computer vision system to detect sorghum aphid states that the ratio of 80:10:10 in segment division in a database is commonly used to detect defects in a deep learning model [5]. The mentioned configurations are commonly established in various studies that aim to achieve a high accuracy percentage in their models.

1.1.3 Formulas for the metrics to be considered.

These formulas demonstrate common results after conducting evaluations on the CNN model. Precision shows the percentage of correct observations that were identified as positive. Recall signals the estimation of observations that were classified as positive. The F1 score demonstrates the harmonic mean of the precision and recall results [6].

$$Precision = \frac{True\ positives}{True\ positives + False\ positives} \tag{1}$$

$$Recall = \frac{True\ positives}{True\ positives + False\ negatives} \tag{2}$$

$$Mpa = 2 * \frac{Recall * Precision}{Recall + Precision} \tag{3}$$

1.1.4 Previous Studies.

The bar chart from Fig. 2 shows that articles have been utilized using modern technology, which enables the identification of key points that could be currently employed and those that have been omitted.

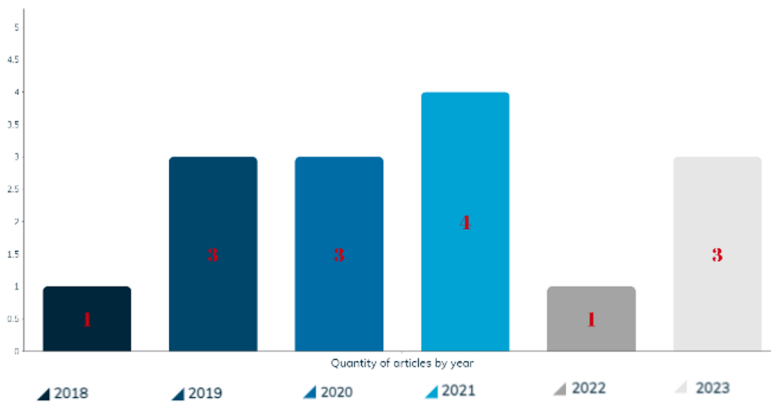


Fig. 2. Year of publication of the articles.

The selected articles from Fig. 3 demonstrate a positive pattern in their results based on the techniques they implemented to achieve the desired percentage. Indicating that a percentage equal to or greater than 90% is considered a fulfillment of their research.

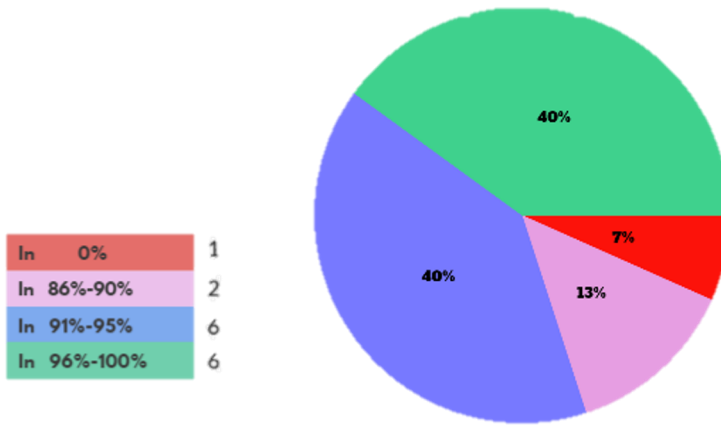


Fig. 3. Pie chart depicting the accuracy percentages of the articles read.

Finally, in Fig. 4, it is observed that, according to the selected articles, the majority focused on the agricultural aspect. However, this does not imply that the techniques they employed cannot be applied to the food-related aspect. Both can complement each other to enhance classification techniques.



Fig. 3. Food and agricultural sector sectors

2 Figures and tables

The research relies on a strict plan that must be managed with precision to fulfill the main objective. Below, in Fig. 5, a flowchart is presented, indicating the activities that will be carried out.

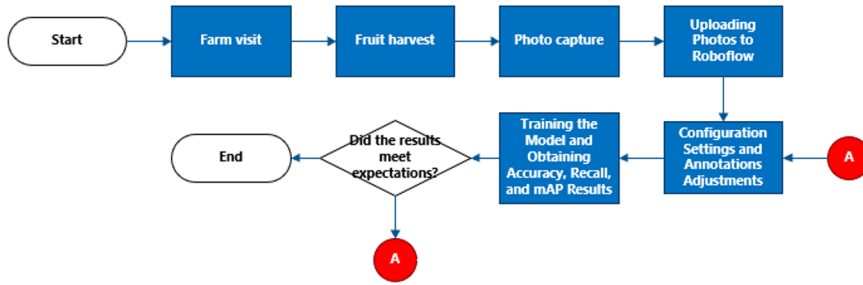


Fig. 5. Flowchart

Transportation and Harvesting. A visit to a coffee farm with a coffee fruit plantation is necessary. At the location, the coffee beans will be obtained to isolate them and place them on a cleared table for image capture. However, branch cutting is not a common practice within the coffee industry. Therefore, it will be necessary to collect the beans, place them in a bag or container, and transport them to a well-lit room. The farm to be visited is located near Lake Yojoa.

Photo Collection. The images will be taken using a Samsung Galaxy A21s smartphone, equipped with a 48-megapixel camera capable of capturing high-resolution images. The captured images will feature multiple objects. A minimum of 1000 images will be collected, all with a resolution of 3000 x 4000 pixels, which will be modified in a subsequent step. The photo session will be conducted by holding the phone horizontally. The focal distance of the phone with the fruit on the table will be 25 to 30 cm. The number of objects in the image will vary. It is emphasized that bright light will not be used; only natural light will be employed in capturing the photographs.

Annotations on Images and Configuration Adjustments. The obtained images will be uploaded to the RoboFlow platform. A project file will be created with the object detection category and the YOLOv5 PyTorch format. During the annotation stage, two categories will be established: good and bad. Care will be taken to avoid overfitting in the CNN model, which is an important consideration. Additionally, image processing will be applied, and the database will be augmented within RoboFlow. The image size will be adjusted from 3000x4000 to 640x640 pixels.

Model Training. The model will train, validate, and test the images with an 80:10:10 ratio concerning the obtained images. A significant portion of the images will be allocated to the training section to enable the model to increase its learning.

Fruit Classification. Table 1 and Table 2 illustrate the variety of fruits that will serve as the basis for the images. From the beginning, starting with green coffee fruit, to the end, which is the dried coffee fruit.

Table 1. Table of coffee fruit categories falling under the classification of good condition.












Image	Description	Image	Description
	Red, Fully Ripe		Combination of Yellow and Red, Near the Completion of the Process
	Yellow, Maturation Process		Green, Pre-stage of the Maturation Process

Table 2. Table of coffee fruit categories falling under the classification of poor condition.

Image	Description	Image	Description
	Rotten Fruit		Rotten Fruit and Near Drying State
	Combination of Yellow and Red with Rotten Area		Completely Dry, Later Stage of Fully Ripe Red Fruit
	Fruit with Small Openings and Rotten Areas		Rotten Fruit with Notable Opening of Coffee Beans Inside
	Fruit with Notable Opening		

3 Results and Analysis

Given that coffee fruit is a perishable product, the photo shoots must be conducted immediately as there is a risk of expiration, and further photo capture may become impractical. Therefore, the photo sessions were carried out for two consecutive days, randomly selecting the quantity and type of cherries to be placed on the table. A total of 1702 photos were successfully obtained. Annotations were made based on the classification of which cherries fell into the appropriate categories, namely good and bad from Table 1 and Table 2. Each image was manually annotated using the bounding box tool provided by RoboFlow. Numerous annotations were made to achieve an accurate model result (Fig. 7). Training was conducted with 600 images without augmentation and pre-adjustment to 640x640, using correct annotations to obtain an efficient, precise model capable of accurately detecting and classifying coffee categories. The task division for the model was in a ratio of 80:10:10, with 480 images for training, 60 images for validation, and 60 images for testing (Fig. 8). Resulting in an accuracy of 97.7%, recall of 94.1%, and mAP of 98.3%. After reviewing several images and the model's classification, it can be affirmed that it can detect the properties of the cherries and correctly predict which ones fall into the two established classes.



Fig. 7. Annotations within the image falling into the classification of good and bad

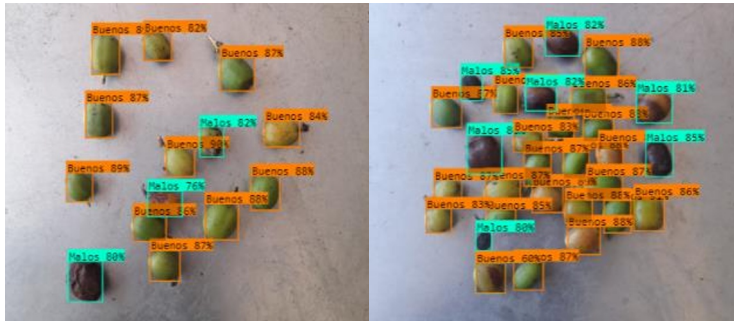


Fig. 8. Test Images of the Model

Subsequently, further training was conducted with different options that can be obtained by applying certain changes to the model. Table 3 shows the results when adding more images, reducing images, and using different types of augmentations.

Table 3. Table comparing the results of each model after training.

Number of Photos	%Precision	%Recall	%mAP	Type of augmentation x3
50	94.5%	96.3%	96.6%	Without
600	97.7%	94.1%	98.3%	Without
1000	95.4%	95.6%	98.4%	Without
100	93.2%	97.9%	96.9%	Rotation 25%
100	90.1%	91.8%	95.8%	Saturation 50%
100	85.2%	93.7%	96.2%	Sound 5%
100	94.7%	96.5%	98.6%	Difumination
100	90.4%	93.9%	96.6%	2.5%
				Everything

Additional training was conducted to assess whether the quantity of images could be a factor in the metric results. It can be affirmed that capturing images in a controlled environment significantly aids the model in achieving adequate results. Notably, with training on just 50 images, the model achieved over 90% in both precision and recall. Additionally, training with data augmentation were performed to identify which parameter resembles the results obtained with the 600-image model, as it is considered to have achieved optimal outcomes.

4 Conclusion

With the coffee harvest on the explored farm, it was possible to obtain all kinds of fruits that exhibit physical traits found in other farms. Likewise, each type of fruit was successfully classified, indicating whether it was in an adequate or deteriorated state. The development of a database with images proved crucial for model training. The proper division of tasks, such as training, validation, and testing, was fundamental in achieving a model with favorable metrics. Due to the acquisition of the necessary quantity of images, which amounted to 1702, favorable results were obtained. The favorable outcome was achieved through the capture of images in a controlled environment, utilizing natural light, a focal distance of 25 to 30 cm, and manual annotations based on the fruit category.

References

1. M. E. Perdomo, M. Cardona, D.M. Castro, W.M. Mejia, Literature Review on Artificial Intelligence Implementation in the Honduran Agricultural Sector. 2022 IEEE International Conference on Machine Learning and Applied Network Technologies (ICMLANT), Soyapango, El Salvador, 1-5 (2022)
2. M.E. Perdomo, I. Magomedov, O. Anokhina, Literary Review: Coffee Technologies. BIO Web Conf., **57**, 05003 (2023)
3. J.L. Ordoñez-Avila, M.G. Martínez-Rangel, Proposed Open Systems Model and Intelligent Agriculture in Coffee Farms within the Harvesting Process, en Proceedings of the 2nd LACCEI International Multiconference on Entrepreneurship, Innovation and Regional Development (LEIRD 2022): “Exponential Technologies and Global Challenges: Moving toward a new culture of entrepreneurship and innovation for sustainable development”, Latin American and Caribbean Consortium of Engineering Institutions (2022)
4. J.L. Ordoñez-Avila, M.E. Perdomo, D. Barahona, H. Perdomo, G. Sosa, M.G. Martínez-Rangel, Proposed Model for Coffee Cooperative in Santa Bárbara, Honduras. en Proceedings of the 2nd LACCEI International Multiconference on Entrepreneurship, Innovation and Regional Development (LEIRD 2022): “Exponential Technologies and Global Challenges: Moving toward a new culture of entrepreneurship and innovation for sustainable development”, Latin American and Caribbean Consortium of Engineering Institutions (2022)
5. M.E. Perdomo, Relationship between coffee quality methods and the coffee price in Honduras. en Proceedings of the 2nd LACCEI International Multiconference on Entrepreneurship, Innovation and Regional Development (LEIRD 2022): “Exponential Technologies and Global Challenges: Moving toward a new culture of entrepreneurship and innovation for sustainable development”, Latin American and Caribbean Consortium of Engineering Institutions (2022)
6. T.A. Heryanto, I.G.B.B. Nugraha, Classification of Coffee Beans Defect Using Mask Region-based Convolutional Neural Network. 2022 International Conference on Information Technology Systems and Innovation (ICITSI), 333-339 (2022)
7. M.S. Fuentes, N.A.L. Zelaya, J.L.O. Avila, Coffee Fruit Recognition Using Artificial Vision and neural NETWORKS. 2020 5th International Conference on Control and Robotics Engineering (ICCRE), 224-228 (2020)
8. J.R. Sanz-Urbe, P.J. Ramos-Giraldo, C.E. Oliveros-Tascon, Algorithm to Identify Maturation Stages of Coffee Fruits. Advances in Electrical and Electronics Engineering - IAENG Special Edition of the World Congress on Engineering and Computer Science 2008, 167-174 (2008)

9. A.G. Alharbi, M. Arif, Detection And Classification Of Apple Diseases using Convolutional Neural Networks. 2020 2nd International Conference on Computer and Information Sciences (ICCIS), 1-6 (2020)
10. I. Grijalva, B.J. Spiesman, B. McCornack, Computer vision model for sorghum aphid detection using deep learning. Journal of Agriculture and Food Research, **13**, 100652 (2023)
11. A. Chaitanya, J. Shetty, P. Chiplunkar, Food Image Classification and Data Extraction Using Convolutional Neural Network and Web Crawlers. Procedia Computer Science, **218**, 143-152 (2023)