Data and Metadata. 2023; 2:175
doi: 10.56294/dm2023175

ORIGINAL

Use of Convolutional Neural Networks (CNN) to recognize the quality of oranges in Peru by 2023

Uso de Redes Neuronales convolucional (CNN) para el reconocimiento de la calidad de las naranjas en el Perú para el año 2023

Franklin Moza-Villalobos[1], Juan Natividad-Villanueva[1], Brian Meneses-Claudio[2]

[1]Universidad Tecnológica del Perú, Facultad de Ingeniería. Lima, Perú.
[2]Universidad Tecnológica del Perú, Facultad de Negocios. Lima, Perú.

Cite as: Moza-Villalobos F, Natividad-Villanueva J, Meneses-Claudio B. Use of Convolutional Neural Networks (CNN) to recognize the quality of oranges in Peru by 2023. Data and Metadata. 2023;2:175.https://doi.org/10.56294/dm2023175

Submitted: 31-08-2023 **Revised:** 27-10-2023 **Accepted:** 28-12-2023 **Published:** 29-12-2023

Editor: Prof. Dr. Javier González Argote

ABSTRACT

Introduction: the agricultural sector in Peru has witnessed a notable increase in the production of oranges, which has promoted the essential use of convolutional neural networks (CNN). The ability to interpret images by visual artificial intelligence has been fundamental for the analysis and processing of these images, especially in the detection and classification of fruits, standing out in the specific case of oranges.
Objective: conduct a systematic literature review (RSL) to evaluate the neural networks used in the classification of oranges in Peru.
Method: an RSL was carried out using the PICO strategy to search the Scopus database. The selection criteria included studies that used convolutional neural networks to classify the quality status of oranges in the Peruvian context.
Results: all the studies reviewed were based on the use of convolutional neural networks (CNN) for fruit classification, using various architectures and techniques. Some studies focused on a single specific fruit, while others addressed the classification of multiple types of fruits, highlighting the importance of the number and variety of images for training the networks.
Conclusions: convolutional neural networks show effectiveness in orange classification, but the quality of the images and the variety of data are essential to improve accuracy.

Keywords: Convolutional Neural Networks; Artificial Intelligence; Fruit; PICO.

RESUMEN

Introducción: el ámbito agrícola en el Perú ha presenciado un notable aumento en la producción de naranjas, lo que ha fomentado el empleo esencial de las redes neuronales convolucionales (CNN). La capacidad de interpretación de imágenes por parte de la inteligencia artificial visual ha resultado fundamental para el análisis y procesamiento de estas imágenes, especialmente en la detección y clasificación de frutas, destacándose en el caso específico de las naranjas
Objetivo: realizar una revisión sistemática de la literatura (RSL) para evaluar las redes neuronales utilizadas en la clasificación de naranjas en Perú.
Método: se llevó a cabo una RSL utilizando la estrategia PICO para la búsqueda en la base de datos Scopus. Los criterios de selección incluyeron estudios que emplearan redes neuronales convolucionales para clasificar el estado de calidad de las naranjas en el contexto peruano.
Resultados: todos los estudios revisados se basaron en el uso de redes neuronales convolucionales (CNN) para la clasificación de frutas, empleando diversas arquitecturas y técnicas. Algunos estudios se centraron en una sola fruta específica, mientras que otros abordaron la clasificación de múltiples tipos de frutas, destacando la importancia de la cantidad y variedad de imágenes para el entrenamiento de las redes.
Conclusiones: las redes neuronales convolucionales muestran eficacia en la clasificación de naranjas, pero la calidad de las imágenes y la variedad de datos son fundamentales para mejorar la precisión.

Palabras clave: Redes Neuronales Convolucionales; Inteligencia Artificial; Frutas; PICO.

INTRODUCCIÓN

Currently, the agricultural sector in Peru has experienced significant growth in the production of oranges, which is why the use of Convolutional Neural Networks (CNN) has become an important tool in it, since thanks to the existence of Artificial Vision has made it possible to interpret images for processing and analysis using said technology. According to Villalobos and Bolt "The use of artificial vision has given notable results, especially in the field of agriculture. Fruit detection is an area that has seen great improvement".[1] Therefore, it can generate great progress in the country's agricultural sector. However, the problem being addressed is that the information on identification, classification and quality control of oranges through neural networks is scarce, as it is a critical factor for the profitability of orange crops in Peru, especially if in large volumes. For the authors Leelavathy et al.[2] "The analysis of orange is generally done by visual examination and observing the size. For large volumes, it cannot be evaluated with human graders, so image preparation is done with quantitative, robust and predictable data". That is, the identification of oranges en masse, both in terms of size and maturity, and among others, may be deficient. In this sense, it is argued that a new RSL is necessary to evaluate and synthesize the current state of knowledge on the classification of oranges using neural networks, because there is not much research and articles on it in the country, and to identify the challenges and opportunities for future research and learning about new data sciences that allow for more efficiency and precision in orange classification systems. Well, data science is today a fundamental tool for the exploitation of data and the generation of knowledge by Alamilla Jiménez et al.[3], and one of these sciences are neural networks that can optimize the various processes that society and human beings have.

METHODS

Firstly, a systematic review of the literature was carried out using the PICO strategy, the question considered was: What types of neural networks for classification can be used to determine the quality status of oranges in Peru? Subquestions were formulated for each component of PICO which were: What are the criteria to determine the quality of oranges? What types of neural networks have been applied? How effective was the automatic classification of oranges in comparison? of manual classification? If the orange is in good condition or not? and relevant keywords were selected such as: fruit producers, fruit growers, Harvesting, Fruit Feature Evaluation, Nutrient deficiency, Fruit Classification, fruit recognition, classification of fruit, Fruit, Fruit Detection, ANN, CNN, artificial neural networks, convolutional neural network, Visual fruit detection, Fruit classification, fruit automated detection, quality level in fruit, fruit quality, CNN, Types of ANN. The search was carried out in the Scopus scientific database. The search equation used in this database was designed to ensure comprehensive coverage of related studies, the equation of which was: "fruit producers" OR "fruit growers" OR Harvesting OR "Fruit Feature Evaluation" OR "Nutrient deficiency" OR " Fruit Classification" OR "fruit recognition", "classification of fruit" OR Fruit OR "Fruit Detection" AND ANN OR CNN OR "artificial neural networks" OR "convolutional neural network" OR "Visual fruit detection" AND "Fruit classification" OR " fruit automated detection" AND "quality level in fruit" OR "fruit quality" OR CNN OR "Types of ANN".

Secondly, clear criteria were established for the selection of articles. The inclusion criteria mainly required that the studies address the use of convolutional neural networks for the classification of the quality status of oranges in the Peruvian context and the most up-to-date articles were taken into consideration. The exclusion criteria were applied to discard studies that did not meet the objectives of the RSL or that did not meet the required quality standards. Some of these criteria considered were: reports other than articles, techniques for classifying fruits other than convolutional neural networks and studies not related to fruit classification using ANN.

Finally, for the study selection process, a total of 93 results were obtained from the Scopus database. Following the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) methodology, zero duplicate articles were eliminated, since only one database was used, which was Scopus. Subsequently, a review of titles, abstracts and keywords was carried out, resulting in 16 articles that met the theme of the RSL. Of these, the full texts were recovered and the established inclusion and exclusion criteria were applied, some of these criteria were: outdated information, thematic irrelevance and lack of credibility, which led to the final selection of 15 articles for analysis. Below is the PRISMA flowchart, which visually represents the process described above:

https://doi.org/10.56294/dm2023175

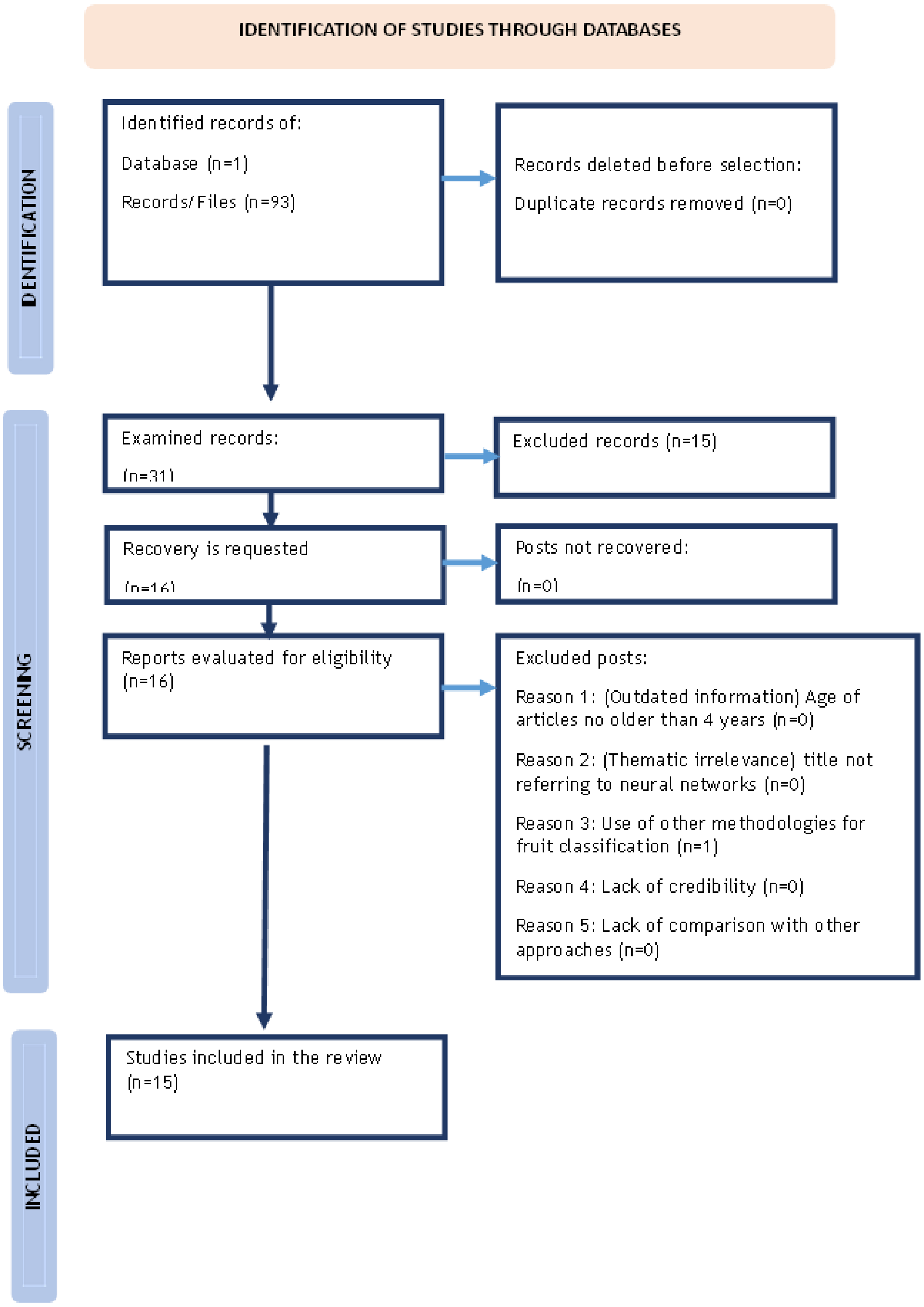

Figure 1. PRISMA

Table 1. Excluded records

Excluded records	
Human	16
Automation tools	62

RESULTS

Once the search for articles was carried out, we investigated what types of neural networks were used in studies related to fruit classification. In addition, the types of fruits targeted by each study were examined, the effectiveness of neural networks was evaluated in comparison with traditional classification techniques or other systems, and the results obtained in each case were analyzed.

According to the analysis carried out on the articles, it was found that all of them were based on the use of convolutional neural networks (CNN) as the main approach for fruit classification, where different characteristics such as architectures, algorithms and techniques were used within the general framework of

CNNs. However, some studies combined different types of neural networks, such as KNN and RNN, to achieve higher accuracy.[4,5,6]

Regarding the architectures used, a variety of approaches was observed. Classic CNN architectures were used, such as CNN LSTM or AlexNet, while others used more modern and deep architectures, such as VGG (Visual Geometry Group) or ResNet. In addition to the architectures, different optimizers were used in some studies. For example, the first article addressed the classification of mangosteen fruits using an optimized CNN model using the Adam optimizer, achieving improvements in the accuracy of the results.[7] Other studies also explored different optimizers as shown in table 2.

Table 2. Optimizers Used in some articles

	Sumari et al.[7]	Vasumathi et al.[8]	Nasir et al.[9]	Palakodati et al.[10]	Siddiqi[11]
Neural Networks Used	CNN optimized	CNN LSTM and CNN LSTM. HYBRID	CNN with VGG19 and PHOG architecture	CNN with a Softmax layer	(CNN) IndusNet, adjusted VGG16 and Mobilenet
Optimizer	Adam	Dragonfly	mRMR	Adam	AdaBoostbased

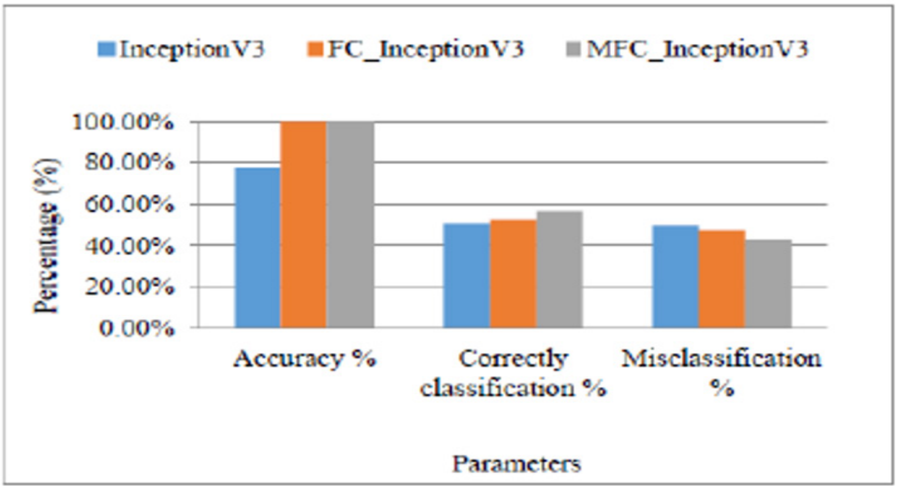

Figure 2. Optimizers Used in some articles

On the other hand, when analyzing the articles, it was observed that some focused on the classification of a single specific fruit, such as dragon fruits.[4] However, there were also studies that addressed the classification of various types of fruits, as shown in table 3.

In addition to the variety of fruits, architectures and approaches used, the quality of the images used was found to be an important factor for the performance and accuracy of the classification techniques. Some studies highlighted the importance of evaluating imaging performance in poor environmental conditions such as fog, while others used thermal imaging to obtain better results.[18] Likewise, some studies focus on the classification of diseases in fruits, using images with diseases or in poor condition.[8,9,10]

In the case of the fourth article analyzed, 6520 images were used, divided into 80 % for training the neural network and 20 % for testing.[8] Specifically, in table 4 you can find the quantities of images used in each study, which varied in each investigation.

Furthermore, by analyzing each item to gain a better understanding and focus, and determine its efficiency compared to traditional techniques. Most studies do not directly compare traditional classification; however, they agree that approaches using traditional manual extraction of fruit characteristics may be deficient as they may lead to inappropriate and late identification, grading and sorting. of fruits in the agricultural field.[4,7,11,13,14,15,16,17] In the commercial setting, this procedure often leads to errors, as sellers must correctly recognize each type of vegetable and fruit, which can be a significant challenge even for highly trained employees.[12]

Finally, the results obtained in each study were analyzed, comparing the methods used, the proposed methods, the method with the best result and the levels of precision achieved, as shown in table 5.

https://doi.org/10.56294/dm2023175

Table 3. Types of Fruits Used

		Studies														
		Sumari et al.[7]	Trieu et al.[4]	Katarzyna et al.[12]	Vasumathi et al.[8]	Nasir et al.[9]	Gill et al.[5]	Palakodati et al.[10]	Gulzar[13]	Gill et al.[6]	Siddiqi[11]	Fu et al.[14]	Ibrahim et al.[15]	Azadnia et al.[16]	Meshram et al.[17]	Mohd Ali et al.[18]
Fruits	Apple			x		x	x	x		x		x			x	
	Banana					x		x	x	x		x			x	
	Manngis	x														
	Dragon fruit		x									x				
	Pomegranate				x										x	
	Peach					x										
	Cherry					x										
	Oranges						x	x		x		x			x	
	Chestnuts								x							
	Kiwi								x			x				
	Starfruit								x							
	Hawthorn													x		
	Lime														x	
	Pear											x				
	Dates										x		x			
	Watermelon															
	Pineapple															x

https://doi.org/10.56294/dm2023175

Table 4. Images used and distribution

Article	Number of images	Distribution
Sumari et al.[7]	1000 images	80 % training and 20 % as a test
Trieu et al.[4]	6500 images	90 % training and 10 % proof
Katarzyna et al.[12]	6161 images	70 % training, 15 % validation and 15 % testing
Vasumathi et al.[8]	6520 images	80 % training and 20 % proof
Nasir et al.[9]	65429 images	48 905 Images for training and 16421 test images
Palakodati et al.[10]	5989 images	3596 for training 596 for validation Palakodati and 1797 for testing
Siddiqi[11]	3640 images	2800 for training, 280 validation and 560 testing
Fu et al.[14]	4000 images	90 % for training and 10 % for validation
Ibrahim et al.[15]	628 images	80 % for training and 20 % for testing
Azadnia et al.[16]	2400 images	80 % for training and 20 % for testing
Meshram et al.[17]	12000 images	70 % training, 15 % validation and 15 % testing
Mohd Ali et al.[18]	3240 thermal images	80 % for training, 10 % for validation and 10 % for testing

Table 5. Results obtained

	Comparative Methods	Proposed Methods	Method with best results	Precision Results
Sumari et al.[7]	Xception, VGG16 and ResNet50 that are optimized	Optimized CNN using 3 types of layers (1) convolutional layers, 2) pooling layers and 3) fully connected (FC) layers	Proposed	94,99 %
Trieu &Thinh [4]	Image Processing + RNA, other types of CNN	K-Nearest Neighbor (KNN) model, CNN, ANN	Proposed	97,38 %
Katarzyna & Paweł [12]	Two ways for classification, the object detection method (YOLO V3) and the full frame classifier and ROIs	Simplified nine-layer CNN architecture, Yolo V3	Simplified nine-layer CNN	99,78 %
Vasumathi & Kamarasan[8]	CNN LSTM and CNN LSTM.HYBRID With Dragonfly optimizer	CNN LSTM and CNN LSTM. HYBRID	LSTM.HYBRID	97,10 %
Nasir et al.[9]	CNN with cubic SVM algorithm and CNN with SVM with quadratic algorithm	CNN with VGG19 and PHOG architecture	CNN with cubic SVM algorithm	99,60 %

DISCUSSION

The analyzed studies reveal a consensus in the predominant use of convolutional neural networks (CNN) as the main approach for fruit classification. These investigations used a wide range of architectures such as CNN LSTM or AlexNet and VGG or ResNet. Furthermore, to improve the classification accuracy, the combination of different types of neural networks such as KNN and RNN is used.

Likewise, the diversity of fruits studied, and the quality of the images used emerge as critical aspects in the performance of the classification models. The importance of evaluating image quality in adverse environmental conditions, such as fog or smog, has been noted, as well as the inclusion of thermal images to improve accuracy. Also, some studies focused on classifying diseases in fruits, using images that represent fruits in poor condition or with conditions. This comprehensive and diversified approach to representing actual fruit conditions significantly influences the neural networks' ability for accurate classification.

In the same way, the number and distribution of data sets used in these studies vary considerably, which could influence the generalization ability and accuracy of the trained models. The variation in the number of images used, as well as the distribution between training, validation and test sets, poses challenges and highlights the importance of establishing optimal criteria for effective training.

Unlike traditional methods that employ manual feature extraction, studies suggest that these approaches may be inadequate for accurate and timely classification of fruits. Neural networks prove to be more efficient and accurate in identification and classification, which is crucial in both agricultural and commercial environments where accuracy in fruit identification is essential.

CONCLUSIONS

The superiority of neural networks over traditional methods for fruit sorting, especially in agricultural and commercial settings, suggests significant potential for optimizing identification and sorting processes in the agricultural industry and fruit supply chain.

The variety of fruits studied in these analyzes highlights the ability of neural networks to adapt to multiple types of agricultural products. This versatility in sorting is essential in agricultural and commercial environments, where accurate identification of different fruits is crucial. Then the variation in the number of images used and the distribution between training, validation and testing sets highlights the need to standardize the data sets for more accurate benchmarking. Furthermore, it highlights the challenge of establishing appropriate criteria for data collection and distribution in future studies.

Finally, the consideration of adverse environmental conditions, such as fog or smog, as well as the focus on fruits with diseases or in poor condition, shows a holistic approach to classification. This suggests that neural networks can be trained to identify not only fruits in ideal conditions, but also in less favorable situations, which improves their applicability in variable environments.

REFERENCES

1. Muñoz Villalobos IA, Bolt A. Diseño y desarrollo de aplicación móvil para la clasificación de flora nativa chilena utilizando redes neuronales convolucionales. RIAI - Revista Iberoamericana de Automatica e Informatica Industrial [Internet]. 2017 Jan 1; 14(1):104–14. Available from: https://revistas.ufpr.br/atoz/article/view/81419

2. Leelavathy B, Sri Datta YSS, Rachana YS. Quality Assessment of Orange Fruit Images Using Convolutional Neural Networks. Springer eBooks [Internet]. 2020; 56:403–12. Available from: https://www.scinapse.io/papers/3115286542

3. Coa YMF, Crisostomo NWF, Díaz-Barriga GE. Desarrollo económico sostenible bajo un régimen social sin preceptos éticos y morales: auditoría forense en contraposición de la corrupción. Revista Científica Empresarial Debe-Haber 2023;1:48-62

4. Trieu NM, Thinh NT. A Study of Combining KNN and ANN for Classifying Dragon Fruits Automatically. Journal of Image and Graphics(United Kingdom) [Internet]. 2022 Mar 1; 10(1):28–35. Available from: https://www.researchgate.net/publication/358940623_A_Study_of_Combining_KNN_and_ANN_for_Classifying_Dragon_Fruits_Automatically

5. Gonzalez-Argote J. Analyzing the Trends and Impact of Health Policy Research: A Bibliometric Study. Health Leadership and Quality of Life 2023;2:28-28. https://doi.org/10.56294/hl202328

6. Gutiérrez VF. La estructura organizacional del Gobierno Regional de Moquegua y su eficiencia funcional. Sincretismo 2021;2

7. Sumari P, Wan Ahmad WMA, Hadi F, Mazlan M, Liyana NA, Bello RW, et al. A Precision Agricultural Application: Manggis Fruit Classification Using Hybrid Deep Learning. Revue d'Intelligence Artificielle [Internet]. 2021 Oct 1; 35(5):375–81. Available from: https://www.iieta.org/journals/ria/paper/10.18280/ria.350503

8. Gonzalez-Argote J. Patterns in Leadership and Management Research: A Bibliometric Review. Health Leadership and Quality of Life 2022;1:10-10. https://doi.org/10.56294/hl202210

9. Nasir IM, Bibi A, Shah JH, Khan MA, Sharif M, Iqbal K, et al. Deep Learning-Based Classification of Fruit Diseases: An Application for Precision Agriculture. Computers, Materials & Continua [Internet]. 2020 Nov 26; 66(2):1949–62. Available from: https://www.techscience.com/cmc/v66n2/40656/html

10. Castillo-Gonzalez W. Charting the Field of Human Factors and Ergonomics: A Bibliometric Exploration. Health Leadership and Quality of Life 2022;1:6-6. https://doi.org/10.56294/hl20226

11. Siddiqi R. Fruit-classification model resilience under adversarial attack. SN Appl Sci [Internet]. 2022 Jan 1; 4(1):1–22. Available from: https://link.springer.com/article/10.1007/s42452-021-04917-6

12. Katarzyna R, Paweł M. A Vision-Based Method Utilizing Deep Convolutional Neural Networks for Fruit Variety Classification in Uncertainty Conditions of Retail Sales. Applied Sciences 2019 [Internet]. 2019 Sep 22; 9(19):3971–88. Available from: https://www.mdpi.com/2076-3417/9/19/3971/htm

13. Auza-Santiváñez JC, Díaz JAC, Cruz OAV, Robles-Nina SM, Escalante CS, Huanca BA. mHealth in health systems: barriers to implementation. Health Leadership and Quality of Life 2022;1:7-7. https://doi.org/10.56294/hl20227

14. Fu Y, Nguyen M, Yan WQ. Grading Methods for Fruit Freshness Based on Deep Learning. SN Comput Sci [Internet]. 2022 Jul 1; 3(4):1–

13. Available from: https://link.springer.com/article/10.1007/s42979-022-01152-7

15. Murillo-Ticona TA, Berneso-Soto ML. Los Entornos Virtuales de Aprendizaje al rescate del servicio educativo. Sincretismo 2020;1

16. Azadnia R, Fouladi S, Jahanbakhshi A. Intelligent detection and waste control of hawthorn fruit based on ripening level using machine vision system and deep learning techniques. Results in Engineering [Internet]. 2023 Mar 1; 17(100891):1–13. Available from: https://www.sciencedirect.com/science/article/pii/S259012302300018X

17. Uman JMM, Arias LVC, Romero-Carazas R. Factores que dificultan la graduación: El caso de la carrera profesional de contabilidad en las universidades peruanas. Revista Científica Empresarial Debe-Haber 2023;1:58-74

18. Mohd Ali M, Hashim N, Abd Aziz S, Lasekan O. Utilisation of Deep Learning with Multimodal Data Fusion for Determination of Pineapple Quality Using Thermal Imaging. Agronomy [Internet]. 2023 Jan 30; 13(2):401–14. Available from: https://www.mdpi.com/2073-4395/13/2/401/htm

FINANCING
The authors did not receive financing for the development of this research.

CONFLICT OF INTEREST
None.

AUTHORSHIP CONTRIBUTION
Conceptualization: Franklin Moza-Villalobos; Juan Natividad-Villanueva.
Data curation: Franklin Moza-Villalobos; Juan Natividad-Villanueva; Brian Meneses-Claudio.
Formal analysis: Franklin Moza-Villalobos; Juan Natividad-Villanueva.
Acquisition of funds: Franklin Moza-Villalobos; Juan Natividad-Villanueva.
Research: Franklin Moza-Villalobos; Juan Natividad-Villanueva; Brian Meneses-Claudio.
Methodology: Franklin Moza-Villalobos; Juan Natividad-Villanueva; Brian Meneses-Claudio.
Project management: Franklin Moza-Villalobos; Juan Natividad-Villanueva.
Resources: Franklin Moza-Villalobos; Juan Natividad-Villanueva.
Supervision: Franklin Moza-Villalobos; Juan Natividad-Villanueva.
Validation: Franklin Moza-Villalobos; Juan Natividad-Villanueva; Brian Meneses-Claudio.
Display: Franklin Moza-Villalobos; Juan Natividad-Villanueva; Brian Meneses-Claudio.
Drafting - original draft: Franklin Moza-Villalobos; Juan Natividad-Villanueva.
Writing - proofreading and editing: Brian Meneses-Claudio.