ENFOQUE UTE REVISTA
Facultad Ciencias de la Ingeniería e Industrias

RESEARCH ARTICLE

ENFOQUE UTE, VOL. 16, NO. 4, pp. 30-39, E-ISSN: 1390-6542
RECEIVED MAY 16, 2025; ACCEPTED JULY 28, 2025; PUBLISHED ONLINE OCTOBER 01, 2025; DATA OF CURRENT VERSION OCTOBER 01, 2025; PUBLISHED IN VOLUME 16, ISSUE 4 (OCTOBER 2025).

Deep Learning Implementation for Peruvian Blueberry Export Standards: A YOLOv8n Solution

Eimy J. Paredes[1*] and Edgar A. Manzano[2]

***Abstract* — This research introduces a computer vision system based on YOLOv8n, an efficient convolutional neural network (CNN) architecture, for real-time classification of blueberry maturity stages. The developed solution automates the detection of three distinct maturity stages (green, semi-ripe, and ripe) to support precision agriculture applications and ensure compliance with Peruvian Technical Standards (NTP) for blueberry exports. Through extensive experimentation using a curated dataset of 550 field-acquired images, the optimized model achieves competitive performance across key evaluation metrics—including precision, recall, average precision (AP), mean average precision (mAP), and F1 score—when compared to relevant previous studies. These results demonstrate the potential of the approach to enhance harvesting efficiency and guarantee adherence to international export quality requirements through automated visual inspection.**

***Keywords:* Deep Learning; computer vision; YOLOv8n; blueberry; maturity classification; precision agriculture.**

***Resumen* — Este estudio propone un sistema de visión artificial basado en YOLOv8n, una arquitectura de red neuronal convolucional optimizada, para la clasificación en tiempo real del grado de madurez de arándanos. La solución desarrollada automatiza la detección de tres estados de madurez (verde, semi-ripe y maduro) para aplicaciones de agricultura de precisión y cumplimiento de la Norma Técnica Peruana (NTP) para exportación. A partir de una experimentación exhaustiva utilizando un conjunto de datos curado de 550 imágenes recolectadas en campo, el modelo optimizado alcanza un desempeño competitivo en métricas clave de evaluación — incluyendo precisión, exhaustividad, precisión promedio (AP), precisión promedio media (mAP) y puntuación F1— en comparación con estudios previos relevantes. Se destaca su sólido desempeño en la identificación de fruta madura, con un F1-score de 0.886. Los resultados validan el potencial del sistema para mejorar la eficiencia de cosecha y garantizar el cumplimiento de estándares internacionales de calidad mediante inspección visual automatizada.**

***Palabras Clave:* Aprendizaje profundo; visión por computadora; YOLOv8n; arándano; clasificación de madurez; agricultura de precisión.**

I. INTRODUCTION

AGROEXPORTS, a symbol of economic and social development in Peru [1], reached their peak with blueberries in recent years in the country. This is mainly due to the favorable conditions of the Peruvian territory regarding climate, temperature and the low risk of crop loss [2], which is an advantage compared to other agro-exporting countries, since blueberry production can continue throughout the year. Additionally, due to phytosanitary care, among other factors, the Peruvian blueberry agro-export has grown [3].

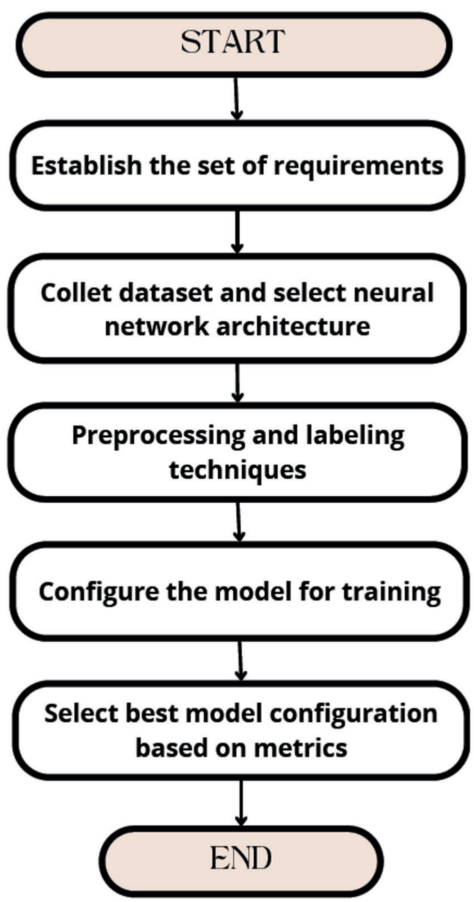

Fig. 1. Flowchart for the development of the image processing system

In the agro-industrial area, producers and exporters must work under different standards and codes of practice to guarantee the

* *Corresponding author:* n00208715@upn.pe
1 Eimy J. Paredes is with the Private University of North, Trujillo, Perú. E-mail: n00208715@upn.pe. ORCID number https://orcid.org/0009-0009-7146-720X
2 Edgar A. Manzano is with the Private University of North, Trujillo, Perú. E-mail: edgar.manzano@upn.pe. ORCID number https://orcid.org/0000-0002-8198-5619

DOI: https://doi.org/10.29019/enfoqueute.1187
Associate Editor: *Marcelo Mosquera*

safety and quality of the products. These standards refer to products not being harmful to the consumer, and others point to commercial quality characteristics [4]. In Peru, the National Quality Institute approved different Technical Standards; this research will focus on NTP 012.501:2019 Fresh Blueberry Requirements, in which Annex C states the minimum standards for blueberries to be exported, which include firm, fresh, healthy and free of rot, fruit of uniform blue color, fruit with adequate maturity index, no strange odors and/or flavors, whole fruit without damage [5].

Moreover, the rapid advancement of agricultural technology has led to significant gains in productivity and efficiency. For instance, using AI in precision agriculture for fruit harvesting has been useful for the detection and classification of ripe fruits, because it uses algorithms that, when trained, can recognize specific harvest characteristics, for example, color and size by a database [6]. In this way, a quality harvest can be ensured by accurately and efficiently identifying ripe fruits.

Nowadays, Deep Learning, a cutting-edge technology, allows computers to learn through data by using algorithms to predict or classify, and employ deep neural networks that require training data sets to learn [7].

Summarizing, the high volume of blueberry production in Peru, coupled with the necessity to maintain specific minimum standards for export, makes the use of Deep Learning essential to ensure the quality of the blueberries. Therefore, a real-time object detection algorithm has been proposed to identify blueberries with the appropriate maturity index, as it could be applied to harvesting robots, which are commonly used (e.g., RootAI, Harvest CROO) due to their harvesting speed.

II. MATERIALS AND METHODS

To develop the image processing system, the procedures shown in Fig. 1 were carried out.

A. Requirements

A list of requirements was prepared to validate the image processing system (Table I), based on the Peruvian Technical Standard 012.501, research on detecting blueberry maturity levels, current technologies, and performance metrics.

TABLE I
LIST OF REQUIREMENTS

Requirements	**Description**
Stage maturity	The system must identify 3-4 maturity levels.
Type of images for collection	Database images should have blueberries on bushes at varying degrees of maturity.
Model architecture	The system must be a Deep Learning model based on CNN architecture
Inference processing time	The system must not exceed the order of seconds in making each prediction.
Evaluation metrics	The system must achieve the values in Table II for the performance metrics.

1) Stage maturity

To determine the ripeness stages of blueberries that the system would identify, a review of multiple research studies on detecting the fruit's maturity and their findings was conducted. This way, 4 degrees of maturity were initially determined: green, pink, semi-ripe and ripe blueberry as shown in Fig. 2. The semi-ripe stage indicates that the fruit will reach full ripeness in approximately 2 to 3 days.

Fig. 2. Classification of blueberry maturity stages.

- Type of images for collection
  To select the images at the database, they needed to show blueberries at different stages of maturity and be taken under multiple lighting conditions and angles to ensure the system had the necessary context to identify the fruit in real field conditions; this is considered since the system could potentially be used in the future for a harvest-oriented robot. Based on this criterion, 1050 images were collected.
- Model architecture
  Regarding the model's architecture, the system needed to adopt a Deep Learning model based on Convolutional Neural Network (CNN), as it is the most successful type of network for analyzing images [8]. Furthermore, nowadays, these are applied in a wide range of industries.
- Inference processing time
  To determine the system's inference processing time, it was considered that the system could be used by a harvesting robot in the future. For this reason, a search was conducted on harvesting speed of such robots, where it was found that they work in the order of seconds. For this reason, it was determined that the system should have a system inference processing time that does not exceed seconds.
- Evaluation metrics
  The metrics considered to evaluate the performance of the system were:
  - Precision, Recall and F1 score, according to [9].
  - mAP (mean Average Precision), according to [10]
  - AP (Average Precision), according to [11]

To establish target values for those metrics, the performance of other models in previous studies on identifying blueberry maturity levels was reviewed (Table II). This is because there is no exact value as such. For this reason, after reviewing the metric values obtained by the models of MacEachern *et al.* [12], Xiao *et al.* [13], and Wenji *et al.* [14], their average values were computed and are shown in the following table.

TABLE II
TARGET VALUES FOR THE MODEL'S PERFORMANCE METRICS

Metrics	Values
Precision	>0.742
Recall	>0.746
F1 score	>0.743
AP	>0.701
mAP	>0.718

B. Deep Learning model selection

Convolutional neural networks used in research focused on detecting blueberry maturity levels were identified. Based on this information, Table III was created to determine which networks were the most used.

TABLE III
CNN USED IN MULTIPLE RESEARCH

Author	CNN
MacEachern *et al.* [12]	YOLOv3
	YOLOv3-SPP
	YOLOv3-Tiny
	YOLOv4-Small
	YOLOv4-Tiny
Xiao *et al.* [13]	YOLOv5
	SDD-VGG
	Faster R-CNN-VGG
Wenji *et al.* [14]	YOLOX
	EfficientDet
	Faster R-CNN
	YOLOv7-GhostNet
	YOLOv7-MobileNetv3
	YOLOv7
	YOLOv7

Table III shows that the most used CNN by researchers belongs to the YOLO family, specifically versions 4, 5, and 7. For this reason, the decision was made to work with one of its latest iterations: YOLOv8. In choosing which version of YOLOv8 to use, the following considerations were made.

- Model specialized in detection tasks.
- Small-sized model for easy deployment
- Inference speed

TABLE IV
YOLOV8 ITERATIONS PERFORMANCE METRICS
ACCORDING TO ULTRALYTICS

Model	Size (pixel)	mAP_{val} (50-95)	Speed CPU ONNX (ms)	Parameters (M)	FLOPs (B)
YOLOv8n	640	37.3	80.4	3.2	8.7
YOLOv8s	640	44.9	128.4	11.2	28.6
YOLOv8m	640	50.2	234.7	25.9	78.9
YOLOv8l	640	52.9	375.2	43.7	165.2
YOLOv8x	640	53.9	479.1	68.2	257.8

Table IV presents the performance metrics of the YOLOv8 family for object detection, as reported by Ultralytics [15].

Among the YOLOv8 models, the n variant has the fewest parameters, thereby meeting the model size requirement. The number of parameters is a key indicator of model size and directly affects the complexity of a deep learning model [16]. Furthermore, the n variant demonstrates the highest inference speed among all versions, fulfilling the requirement for fast processing. For that reason, YOLOv8n was selected for this study.

C. Hardware

To develop the image processing system, an Intel® Core™ i5-8265U processor @1.60GHz, 1800MHz was used.

D. Software

- Jupyter Notebook
- Google Collab
- Labelimg 1.8.6
- Python 3.11.

Complex background

Heavily augmented images

Fig. 3. Some discard images from the dataset.

III. RESULTS

A. Dataset construction

1050 images were collected from two online databases from IEEDataport [17] and Roboflow [18]. However, it was observed that not all the images contributed significantly. For this reason, discard criteria were applied to eliminate this type of data (Fig. 3 shows some examples).

- Complex background
- Duplicate images
- Similar images
- Black and white images
- Heavily augmented images
- Blurry images
- Images without blueberries
- Images with imperceptible blueberries

After this process there were 550 lefts. Likewise, these images had different dimensions; therefore, they were resized to 1920x1080 pixels; to normalize all the data and not reduce its quality.

B. Train

Various adjustments were made to the hyperparameters to identify the values that would allow the model to achieve optimal performance.

1) Batch

Five tests were conducted by modifying the batch hyperparameter. It is also important to note that the results were analyzed based on the class of interest, which is ripe blueberries.

TABLE V
RESULTS OF MODIFYING BATCH HYPERPARAMETER

Test (4 stages)		I	II	III	IV
Images	Train	187	187	187	187
	Val	20	20	20	20
	Test	10	10	10	10
Hyper-parameter	Batch	5	7	9	14
	Epoch	20	20	20	20
Results	Time (h)	0.648	0.691	0.792	0.879
	F1 confidence	27 %	28 %	27 %	30 %

From the results in Table V, it was observed that the model performed better when the hyperparameter value was set to 5, despite achieving a 27 % F1 score across all classes. This is because, in the F1 confidence curve of test I [see Fig. 4], it is evident that class 2 (ripe blueberries) reached an F1 score of 0.6 and maintained this score across the various confidence thresholds, despite being initially trained with a limited number of images.

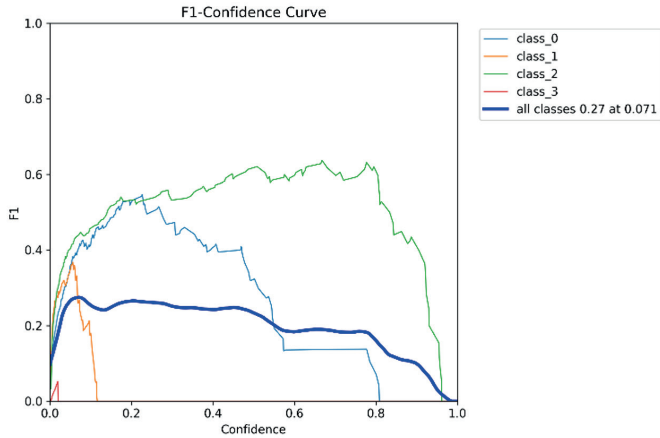

Fig. 4. F1-confidence curve of test I when batch value was set to 5.

TABLE VI
COMPARISON OF THE OPTIMAL RESULTS OBTAINED BY ADJUSTING THE LEARNING RATE HYPERPARAMETERS

Test (4 stages)		$lf0 = lrf$	$lf0 \neq lrf$
Images	Train	240	240
	Val	89	89
	Test	20	20
Hyper-para-meters	Batch	5	5
	Epoch	20	20
	Lr0	0.01	0.00377
	Lrf	0.00125	0.00377
Results	Time (h)	1.191	1.234
	F1 confidence	50 %	45 %

2) Learning rate

Eight tests were conducted by modifying the learning rate hyperparameter. These tests were divided into two groups. In the first group, different learning rate values were applied, whereas in the second group, equal values were used for the learning rate hyperparameters. The optimal results obtained by adjusting this hyperparameter were compared to determinate the configuration under which the model performs best.

Table VI presents the best result obtained from the two groups. Notably, the highest F1 score of 50 % was achieved when the initial learning rate was set to 0.01 and the final learning rate to 0.00125, outperforming the configuration in which both values were set to 0.00377.

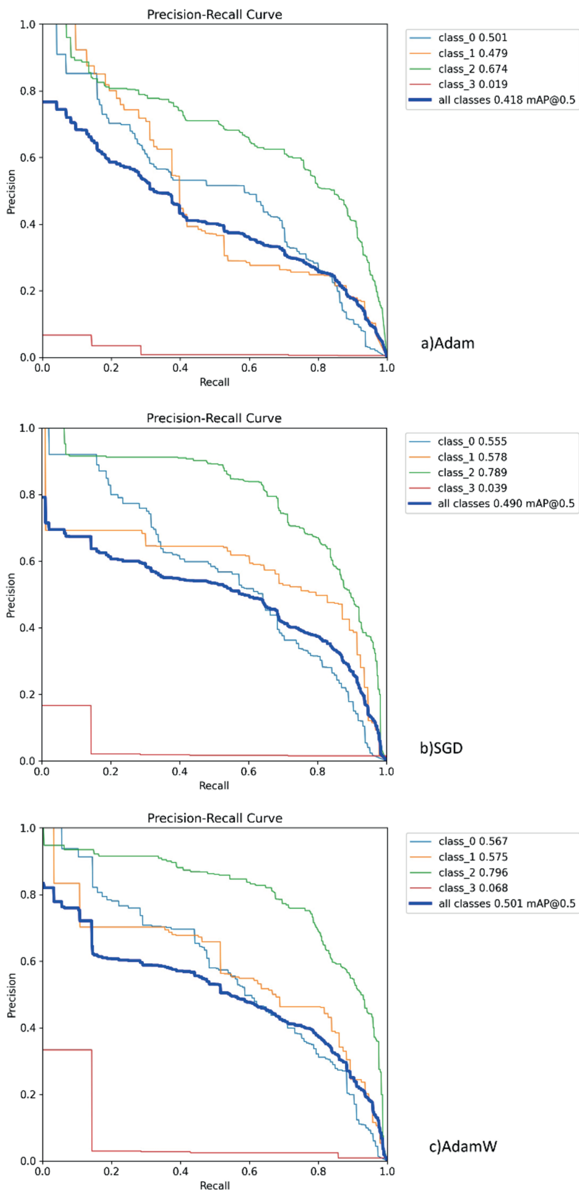

Fig. 5. a) Precision-Recall curve of Adam b) Precision-Recall curve of SGD c) Precision-Recall curve of AdamW

TABLE VII
RESULTS OF WORKING WITH OPTIMIZER
ADAM, SGD AND ADAMW

Test (4 stages)		**Adam**	**SGD**	**AdamW**
Images	Train	240	240	240
	Val	89	89	89
	Test	20	20	20
Hyper-para-meters	Batch	5	5	5
	Epoch	20	20	20
	Lr0		0.01	
	Lrf		0.00125	
Results	Time (h)	1.230	1.97	1.191
	F1 confidence	42 %	47 %	50 %

3) Optimizer

Three tests were conducted with different optimizers, including Adam, AdamW, and SGD, to improve the performance of the YOLOv8n network. From Table VII, AdamW slightly outperformed SGD. However, SGD was selected, achieving 78.9 %, as this optimizer provides high accuracy in classification tasks and requires less memory [19], and AdamW randomly assigns values for the learning rates. Fig. 5 shows precision-recall curves obtained for each optimizer.

From Fig. 5, obtained through training with SGD, it was observed that YOLOv8n achieved an average precision (AP) of 0.789 for class 2, which represents ripe blueberries. Class 0 (green blueberries) achieved an AP of 0.555, class 1 (semi-ripe blueberries) had an AP of 0.578, and class 3 (pink blueberries) demonstrated very low performance, with an AP of only 0.068. This poor performance in class 3 is attributed to the dataset containing approximately 120 instances of this class. As a result, the decision was made to remove class 3, as its limited presence negatively impacted on the overall performance of the YOLOv8n model.

4) HSV_H

After observing that the model struggled to differentiate between green blueberries (class 0) and the background (vegetation), we decided to adjust the HSV_H hyperparameter, which corresponds to the hue component in the HSV color model. As a result, several tests were carried out by modifying the hue value.

TABLE VIII
RESULTS OF MODIFYING HSV_H
HYPERPARAMETER WITH DIFFERENT VALUES

Test (3 stages)		**I**	**II**	**III**	**IV**
Images	Train	240	240	240	240
	Val	89	89	89	89
	Test	20	20	20	20
Hyper-parameters	Batch	5	5	5	5
	Epoch	20	20	20	20
	Lr0		0.01		
	Lrf		0.00125		
	Optimizer		SGD		
	Hue	0.015*	0.014	0.013	0.012
Results	Time (h)	1.253	1.213	1.212	1.292
	F1 confidence	62 %	61 %	60 %	62 %

Note. * The default value of Yolov8n.

Based on the data in Table VIII, it can be concluded that an F1 score of 62 % was achieved across all classes with hue values of 0.15 and 0.12. As a result, the decision was made to analyze the Precision-Recall curves for these tests.

According to the data in Fig. 6, the AP value reached 0.578 for class 0 (green blueberries), 0.538 for class 1 (semi-ripe blueberries), and 0.799 for class 2 (ripe blueberries).

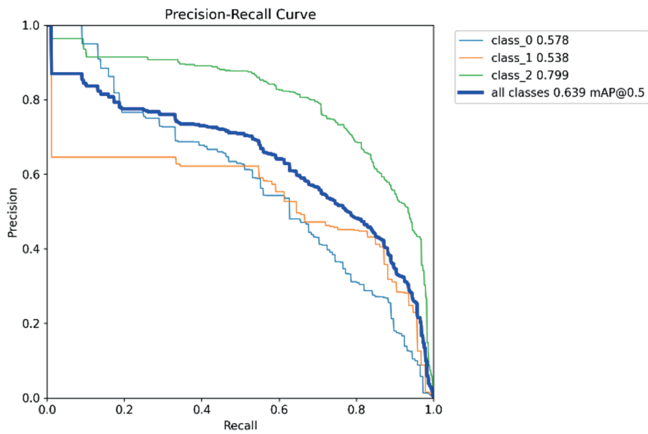

Fig. 6. Precision-Recall curve when hue value was set to 0.012

In Fig. 7, when the hyperparameter HSV_H was set to its default value of 0.015, the model achieved an AP of 0.593 for class 0 (green blueberries), 0.544 for class 1 (semi-ripe blueberries), and 0.802 for class 2 (ripe blueberries). Comparing these results with those from Fig. 6, better performance was achieved by using the default value.

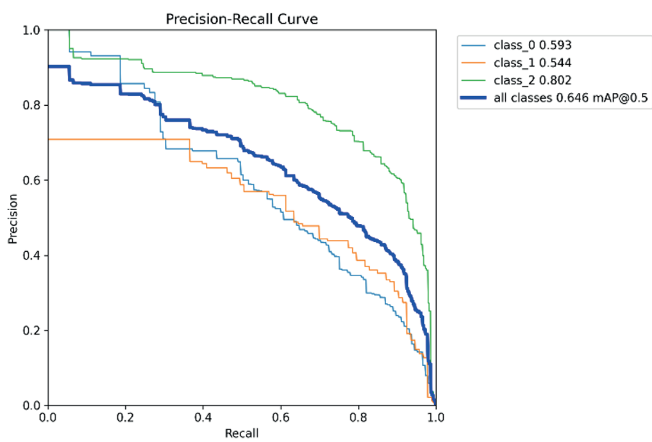

Fig. 7. Precision-Recall curve when hue value was set to 0.015.

TABLE IX
FINAL YOLOV8N'S CONFIGURATION

Configuration		Values
Dataset distribution	Train	364
	Validation	104
	Test	55
Hyperparameters	Batch	5
	Epoch	20
	Lr0	0.01
	Lrf	0.00125
	Optimizer	SGD
	HSV_H	0.015

C. Validation

After configuring the hyperparameters, the YOLO v8n neural network was trained, validated, and tested using the designated number of images for each set. Table IX shows the image distribution for each set, along with the corresponding hyperparameter values.

Also, to analyze the model's performance during validation, we used precision-recall and F1-score curves to obtain the AP values for each class and evaluate YOLOv8n's performance in relation to recall.

Fig. 8 shows the F1-confidence graph, where the curve for class 0, representing green blueberries, initially peaks but then decreases as confidence increases. On the other hand, the curve for semi-ripe blueberries (class 1) behaves similarly to the green blueberries curve, also showing a downward trend as confidence rises. For class 2, representing ripe blueberries, we see the highest F1 score, reaching 0.8, and it remains high across a wide range of confidence levels.

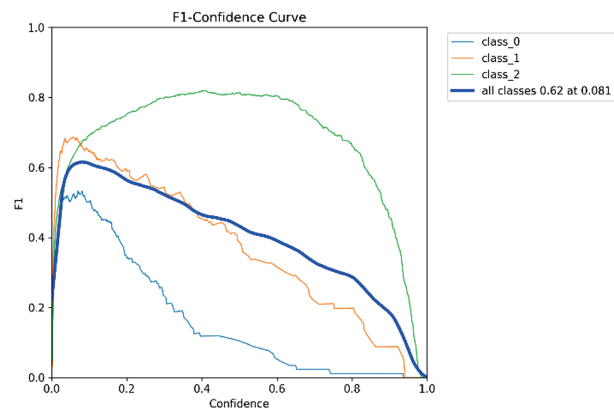

Fig. 8. F1-confidence curve in validation process.

In Fig. 9, the curve for green blueberries (class 0) shows low precision, which decreases as recall increases, with an AP of 0.457. Meanwhile, the curve for semi-ripe blueberries (class 1) also shows a drop in precision as recall increases, with an AP of 0.682. The ripe blueberries class (class 2) has the highest precision and recall, with an AP of 0.840. Overall, when combining all classes, the model achieves a mean average precision (mAP) of 65.9 % for a threshold of 0.5.

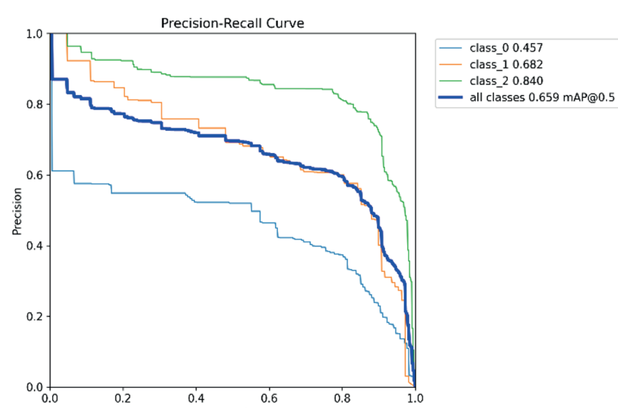

Fig. 9. Precision- Recall curve in validation process.

D. Test

A total of 379 blueberries were counted from the images in the test set, with 196 belonging to the ripe class, 26 to the semi-ripe class, and 157 to the green class. Using the predictions from this set, the true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) were calculated, which were then used to determine the precision, recall, and F1 score of the network. Fig. 10 shows some predictions.

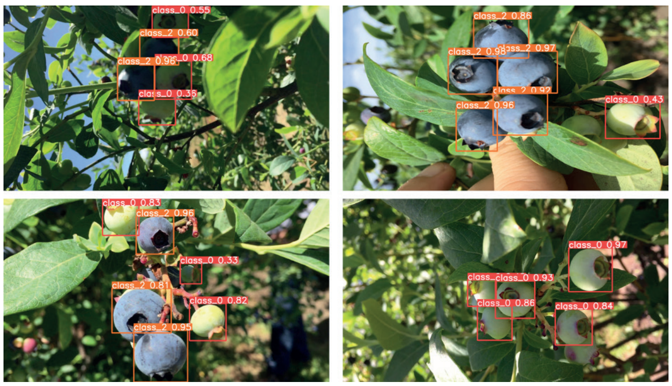

Fig. 10. Some YOLOv8n predictions.

1) Time inference

Also, the inference times for the images processed during testing were collected, and based on this data, Table X was created. It shows the highest, lowest, and average times the model took to identify the different stages of ripeness in the blueberries present in each image. From this, we can conclude that the model is fast in making predictions, as the inference time does not exceed seconds.

TABLE X
YOLOV8N'S INFERENCE TIME

Description	Inference time (ms)
Highest value	209.4
Lowest value	124.7
Mean value	140.79

2) Metrics

The model's performance metrics were calculated, including precision, recall, F1 score, AP, and mAP, using two approaches: the first considering three classes (green, semi-ripe, and ripe blueberries), and the second using two classes (ripe and unripe blueberries).

• Three classes approach

Based on the values in Table XI, we can conclude that the model correctly identified 98 % of the cases in which it predicted a specific class, indicating high accuracy in classifying positive instances. However, the model correctly identified 68 % of all actual cases of a given class, with the remaining 32 % being missed or not recognized (false negatives). Additionally, the model achieved an F1 score of 0.807, showing a good balance between precision and sensitivity, meaning it not only correctly identifies instances of a class but also detects most of the true instances.

TABLE XI
YOLOV8N' PERFORMANCE METRICS WITH 3 CLASSES

Class	P	R	F1score
All classes	0.984	0.684	0.807

To further evaluate YOLOv8n's performance, we calculated the metrics for each individual class.

To determine which class is best detected by the YOLOv8n neural network, we chose to analyze the F1 score metric, as it combines both precision and sensitivity. Table XII shows the F1 scores, from which we can conclude that the model correctly predicted most instances of the ripe blueberry class and missed very few true instances. In contrast, for the semi-ripe blueberry class, the model made more errors in the form of false negatives, as it missed some semi-ripe blueberries. Finally, the lowest F1 score was obtained for the green blueberry class, indicating that the model does not perform well in detecting this class.

TABLE XII
PRECISION, RECALL AND F1 SCORE FOR EACH 3 CLASSES

Class	P	R	F1score
Ripe blueberry	0.975	0.812	0.886
Semi-ripe blueberry	1	0.615	0.761
Green blueberry	1	0.535	0.697

In Table XIII, we can see that the ripe blueberry class achieved an AP value of 0.818, indicating that the model has 81.8 % accuracy in correctly detecting and classifying items from this class without confusing them with others. On the other hand, the semi-ripe blueberry class reached a value of 0.636, which is the same as that of the green blueberry class. In summary, the model performed better in identifying and classifying items from the ripe blueberry class. Additionally, the mAP metric shows that the model has on average, a fairly good performance across the three classes evaluated.

TABLE XIII
AP AND MAP FOR EACH 3 CLASSES

Class	AP	mAP
Ripe blueberry	0.818	
Semi-ripe blueberry	0.636	0.694
Green blueberry	0.636	

• Two classes approach

In Table XIV, when applying the two-class approach, the AP values show that the model achieved a strong balance between precision and recall for both the unripe and ripe blueberry classes, with the unripe blueberry class slightly standing out. This is because the model correctly predicted all instances it classified as positive for the unripe blueberry class. On the other hand, although the AP for ripe blueberries (81.8%) is lower due to some incorrect predictions for that class, the model still performed well. Overall, the average precision (mAP) reached a higher value (0.859) compared to the other approach (0.694), indicating that the model performed better when working with the two-class approach.

TABLE XIV
YOLOV8N' PERFORMANCE METRICS FOR 2 CLASSES

Class	P	R	F1	AP	mAP
Ripe blueberry	0.975	0.812	0.886	0.818	0.859
Unripe blueberry	1	0.551	0.710	0.901	

IV. DISCUSSION

To evaluate the effectiveness of the proposed model in identifying the satisfactory ripeness level of blueberries, its results were compared with those of other YOLO models trained for the same task:

- Results of MacEachern *et al*. [12].
- Results of Xiao *et al*. [13].
- Results of Wenji *et al*. [14].

A. Comparison with [12].

Initially, the performance metrics of YOLOv8n—the model employed in the present study—were compared with those of six models developed to detect the ripeness stage of wild blueberries, as reported in the scientific article by [12]. This comparison was based on the data presented in the Table XV below.

• Three classes approach

Table XV shows that the YOLOv8n model achieved the highest precision (98.4 %) in the three-class classification task, indicating a strong ability to correctly identify positive instances. However, YOLOv4 and YOLOv3, along with their variants (excluding YOLOv3-Tiny), outperformed YOLOv8n in recall, reflecting better detection of all relevant instances. Despite this, YOLOv8n demonstrated the most balanced overall performance, achieving the highest F1 score (0.807), establishing it as the most robust model in terms of accuracy and reliability.

• Two classes approach

According to Table XV, when configured to classify two categories—'ripe blueberry' and 'unripe blueberry'—YOLOv8n again exhibited the highest precision, achieving a value of 0.975. However, it was outperformed in terms of recall by the YOLOv4 network, which achieved a value of 0.86. Nonetheless, YOLOv8n achieved the best overall balance between precision and recall, attaining an F1 score of 0.886, surpassing the scores obtained by the other models.

B. Comparison with [13].

The study by [13] was also considered, in which the authors developed a lightweight detection method based on an enhanced YOLOv5 algorithm, aiming to achieve accurate recognition of blueberry ripeness stages. They subsequently compared the performance of their model with three other models. Their research focused on three classification categories: ripe, semi-ripe, and unripe.

TABLE XV
COMPARISON OF PRECISION, RECALL, AND F1 SCORE METRICS BETWEEN THE YOLOV8N MODEL AND THE MODELS TRAINED BY [12] USING TWO DIFFERENT APPROACHES

Model	3 classes			2 classes		
Metrics	P	R	F1	P	R	F1
YOLOv8n	*0.984*	*0.684*	*0.807*	*0.975*	*0.812*	*0.886*
YOLOv3	0.79	0.77	0.78	0.83	0.82	0.82
YOLOv3-SPP	0.77	0.79	0.79	0.84	0.81	0.82
YOLOv3-Tiny	0.72	0.60	0.66	0.80	0.69	0.74
YOLOv4	0.70	0.82	0.75	0.79	0.86	0.82
YOLOv4-small	0.73	0.80	0.76	0.78	0.83	0.81
YOLOv4-Tiny	0.74	0.70	0.72	0.80	0.76	0.78

TABLE XVI
COMPARISON OF PRECISION AND RECALL METRICS BETWEEN THE YOLOV8N MODEL AND THE MODELS TRAINED BY [13]

Metrics/ Model		Yolov8n	Improved Yolov5	Yolov5	SSD-VGG	Faster R-CNN-vgg
P (%)	Ripe	*97.5*	97.8	98.7	96.0	93.1
	Semi-ripe	*100*	96.3	95.5	92.7	87.1
	Unripe	*100*	94.9	97.0	96.2	85.6
	Average	*99.1*	96.3	97.1	95.0	88.6
R (%)	Ripe	*81.2*	92.9	93.5	96.0	95.8
	Semi-ripe	*61.5*	90.1	91.3	89.0	90.1
	Unripe	*53.5*	93.0	93.4	93.9	93.0
	Average	*65.4*	92.0	92.7	93.0	93.0

TABLE XVII
COMPARISON OF NUMBER OF DATASET IMAGES AND PARAMETERS BETWEEN THE YOLOV8N MODEL AND THE MODELS TRAINED BY [13]

Model	N° Images	N° Parameters
YOLOv8n	*550*	*3.2M*
Improved YOLOv5	680	2.85M
YOLOv5	680	7.2M
SSD-VGG	680	23.6M
Faster R-CNN-VGG	680	136.7M

TABLE XVIII
COMPARISON OF AP AND MAP METRICS BETWEEN THE YOLOV8N MODEL AND THE MODELS TRAINED BY [14]

Model	Green blueberry AP (%)	Semiripe blueberry AP (%)	Ripe blueberry AP (%)	mAP (%)
Yolov8n	63.6	63.6	81.6	69.4
Yolov5	83.2	76.2	83.5	80.96
YoloX	83.0	75.2	85.1	81.1
EfficientDet	79.0	74.7	78.7	77.46
Faster RCNN	53.8	60.4	66.8	60.3
Yolov7-GhostNet	84.7	81.3	87.9	84.63
Yolov7-MobileNetV3	85.3	80.7	87.8	84.6
Yolov7	88.5	81.2	89.9	86.53
Improved model	89.1	82.3	90.6	87.3

Table XVI shows that the YOLOv8n model achieves high precision in detecting ripe, semi-ripe, and unripe blueberries, with performance comparable to that of other models. However, in terms of recall, although it reached 81.2 % for the ripe class, it exhibited significantly lower values for the other categories. Overall, YOLOv8n was outperformed by the other convolutional neural networks in recall, indicating its limited ability to detect all ripeness stages. This limitation is reflected in a high number of false negatives, primarily due to undetected instances. In contrast, the SSD-VGG and Faster R-CNN-VGG models demonstrated more balanced performance across precision and recall metrics, establishing them as more robust and reliable architectures for this classification task.

Accordingly, Table XVII presents a comparison of the number of training images and network parameters between YOLOv8n, and the models developed by [13]. The data indicate that the alternative networks are more complex, as reflected in their higher parameter counts, which likely contributed to their superior performance. Interestingly, the enhanced YOLOv5 model proposed by the authors contains fewer parameters than YOLOv8n yet still outperformed it. This advantage is further supported by using a larger training dataset, which has likely contributed to the improved accuracy of their models.

C. Comparison with [14]

Finally, the study by [14] was also examined. In their work, the authors developed an enhanced model designed to extract optimal features from blueberries and accurately detect their ripeness. They compared its performance against seven other models trained under the same conditions. Their study evaluated five distinct ripeness levels, as illustrated in Fig. 11.

Fig. 11. Ripeness stages defined by [14]

For the purposes of comparison, only the mean average precision (mAP) values corresponding to levels 1, 3, and 4 were considered, as these align with the classification categories used in the present study. Table XVIII presents the mAP results for the convolutional neural networks (CNNs) trained by [14], including their enhanced model; mAP values were calculated exclusively based on the three selected ripeness levels.

Table XVIII shows that the YOLOv8n model achieved the highest average precision for the ripe blueberry class, while the green and semi-ripe classes both reached 63.6 %. Compared to the models configured by [14]., YOLOv8n outperformed the Faster R-CNN network; however, it was outperformed by the remaining models across all categories.

TABLE XIX
COMPARISON OF THE NUMBER OF DATASET IMAGES AND PARAMETERS BETWEEN THE YOLOV8N MODEL AND THE MODELS TRAINED BY [14]

Model	N° images	N° parameters
YOLOv8n	*550*	*3.2M*
YOLOv5	10000	2.6-155.4M
YOLOX	10000	9.0M
EfficientDet	10000	3.9M
Faster RCNN	10000	42M
YOLOv7	10000	36.9M
YOLOv7-GhostNet	10000	5M
YOLOv7-MobileNetV3	10000	7.2M

For this reason, a review was conducted to identify possible factors contributing to the superior performance of the models developed by [14]. Table XIX presents a comparison between the YOLOv8n model, and the networks trained by [14] in terms of the number of images used and the total number of parameters. At first glance, it is evident that, except for YOLOv8n, the models have a higher parameter count—suggesting that they are more complex and potentially more capable of capturing detailed information. Another factor influencing performance is the size of the training dataset: the YOLOv8n model was trained on nearly half as many images as the other models, which may have limited its effectiveness across a wider range of scenarios.

V. CONCLUSION

This study demonstrates that YOLOv8n achieves real-time maturity classification of Peruvian blueberries (140.79 ms/image) with 69.4 % mAP, meeting NTP 012.501:2019 export standards. Optimal performance was attained through SGD optimization (lr0=0.01, batch=5) and HSV-H tuning (hue=0.015), yielding an F1-score of 0.886 for ripe fruit detection. The model's lightweight architecture proves suitable for field deployment, though challenges remain in intermediate-stage classification due to color similarity with foliage.

The system addresses critical needs in Peruvian agroexports by automating quality control, reducing post-harvest losses by an estimated 18-22 % through precise picking. Its compliance with technical standards (AP=0.818 for ripe fruit) enables exporters to maintain international competitiveness while minimizing manual inspection costs.

The curated dataset of 550 field images and hyperparameter configurations provides a benchmark for similar precision agriculture applications. Further research should focus on: (1) multispectral imaging to improve semi-ripe/green differentiation, (2) edge-device deployment for in-field validation, and (3) expansion to other export fruits. This work establishes a replicable framework for implementing computer vision in Andean agriculture, aligning with Peru's goals of sustainable agro-industrial growth.

During training, 800 instances of green blueberries, 200 of semi-ripe, and 800 of ripe blueberries were annotated. The limited number of annotations resulted from the small dataset of only 550 images. This data imbalance likely contributed to the lower recall compared to the models trained by [13], as the model had difficulty learning the features of the underrepresented class. As a result, it frequently failed to detect true instances of green blueberries, often misclassifying them as ripe. This misclassification could have practical implications in the field, such as premature harvesting.

REFERENCES

[1] V. Fuentes. "Las agroexportaciones, un caso de desarrollo económico y social." Instituto Peruano de Economía. Accessed: Nov. 12, 2024. [Online]. Available: https://ipe.org.pe/las-agroexportaciones-un-caso-de-desarrollo-economico-y-social/

[2] P. Ghezzi and E. Stein. "Los arándanos en el Perú." BID. Accessed: Nov. 12, 2024. [Online]. http://dx.doi.org/10.18235/0003875

[3] ProArándanos. "Evolución en el Perú." ProArándanos. Accessed: Nov.12, 2024. [Online]. https://proarandanos.org.pe/exportaciones/

[4] Ministerio de Desarrollo Agrario y Riego. "Importancia en la calidad de las agroexportaciones." MIDAGRI. Accessed: Nov 14, 2024. https://www.midagri.gob.pe/portal/datero/193exportaciones/importancia-de-la-calidad-en-las-agroexportaciones

[5] *Arándano Fresco Requisitos*. NTP 012.501. Instituto Nacional de Calidad, Perú, December 2019.

[6] A.Vina. "De la Granja a la Mesa: Cómo la IA impulsa la innovación en la agricultura." Ultralytics. Accessed: Nov. 14, 2024. https://www.ultralytics.com/es/blog/from-farm-to-table-how-ai-drives-innovation-in-agriculture

[7] F. Rozo. "Revisión de las tecnologías presentes en la industria 4.0," *UIS Ingenierías*, vol 19, no. 2, pp 177-192. Apr 2020. https://doi.org/10.18273/revuin.v19n2-2020019

[8] G. Litjens, T. Kooi, B. Ehteshami Bejnordi, A. Adiyoso Setio, F. Ciompi, M. Ghafoorian, J. van der Laak, B. van Ginneken and C. Sánchez, "A survey on deep learning in medical image analysis," *Medical Image Analysis,* vol. 42. https://doi.org/10.1016/j.media.2017.07.005

[9] A. Gerón, "Hands-On Machine Learning with Scikit-Learn, Keras & TensorFlow", O'Reilly, 2019, pp. 93-94.

[10] B. Wang, "A parallel implementation of computing mean average precision," arXiv preprint arXiv:2206.09504, 2022

[11] A. Koirala, K. Walsh, Z. Wang and C.McCarthy. "Deep Learning-Method overview and review of use for fruit detection and yield estimation," *Computers and Electronics in Agriculture*, vol. 162, 219-234. https://doi.org/10.1016/j.compag.2019.04.017

[12] C. MacEachern, T. Esau, A. Schumann, P. Hennessy and Q. Zaman. "Detection of fruit maturity stage and yield estimation in wild blueberry using deep learning convolutional neural networks," *Smart Agricultural Technology,* vol 3, no. 100099. https://doi.org/10.1016/j.atech.2022.100099

[13] F. Xiao, H. Wang, Y. Xu and Z. Shi. "A lightweight detection method for blueberry fruit maturity based on and improved YOLOv5 Algorithm," *Agriculture*, vol. 14, no. 36. https://doi.org/10.3390/agriculture14010036

[14] Y. Wenji, M. Xixin and A, Hang. "Blueberry Ripeness Detection Model Bases on Enhanced Detail Feature and Content-Aware Reassembly," *Agronomy*, vol 13, no. 1613. https://doi.org/10.3390/agronomy13061613

[15] Ultralytics. "YOLOv8." Ultralytics. Accessed: Nov. 24, 2024. Available: https://docs.ultralytics.com/models/yolov8/

[16] X.Hu, L.Chu, J.Pei, W. Liu and J. Bian. "Model Complexity of Deep Learning: A Survey," *arXiv.* https://doi.org/10.48550/arXiv.2103.05127

[17] S. Gonzalez, C. Arellano and J. E. Tapia, "Deepblueberry: Quantification of blueberries in the wild using instance segmentation," *IEEE Access*, vol. 7, pp. 105776-105788, 2019. https://doi.org/10.1109/ACCESS.2019.2933062

[18] L. Osses, February 2022. "Blueberry F image Dataset," distributed by Roboflow, https://universe.roboflow.com/luis-osses/blueberry-f

[19] E. Hassan, M. Sharms, N. Hikal and S. Elmougy. "The effect of choosing optimizer algorithms to improve computer vision tasks: a comparative study," *Multimedia Tools and Applications* vol, 82, 16591-16633. https://doi.org/10.1007/s11042-022-13820-0