

Development of an Algorithm with Artificial Vision using Convolutional Neural Networks for Blueberry Quality Control

Ryan Abraham León León, PhD¹, Martín Antonio Rentería Davila, Engineer²
^{1,2}Universidad Privada del Norte, Perú, ryan.leon@upn.edu.pe, n00268483@upn.pe

ABSTRACT: *The paper presents the development of a computer vision algorithm using convolutional neural networks for blueberry quality control. The YOLOv8 network was used for the detection and classification of blueberries according to their quality. A total of 840 blueberry images were collected and labeled using the Roboflow platform. After training and evaluating the model, an accuracy between 91% and 98%, and F1-Scores between 90% and 97% were obtained in classifying blueberries as good or bad in seven different production areas. The results demonstrate the effectiveness of the YOLOv8-based machine vision system for accurate detection of blueberry quality, optimizing the sorting process and reducing human intervention.*

Keywords: *Artificial vision, convolutional neural networks, YOLOv8, quality control, blueberries, agricultural product classification.*

*Digital Object Identifier: (only for full papers, inserted by LEIRD).
ISSN, ISBN: (to be inserted by LEIRD).
DO NOT REMOVE*

Desarrollo de un Algoritmo con Visión Artificial utilizando Redes Neuronales Convolucionales para el Control de la Calidad del Arándano

Ryan Abraham León León, Doctor¹, Martín Antonio Rentería Dávila, Ingeniero²
^{1,2}Universidad Privada del Norte, Perú, ryan.leon@upn.edu.pe, n00268483@upn.pe

Resumen: El artículo presenta el desarrollo de un algoritmo con visión artificial utilizando redes neuronales convolucionales para el control de la calidad del arándano. Se empleó la red YOLOv8 para la detección y clasificación de los arándanos según su calidad. Se recolectaron y etiquetaron 840 imágenes de arándanos utilizando la plataforma Roboflow. Después de entrenar y evaluar el modelo, se obtuvo una precisión entre el 91% y el 98%, y F1-Scores entre el 90% y el 97% en la clasificación de arándanos como buenos o malos en siete zonas de producción distintas. Los resultados demuestran la efectividad del sistema de visión artificial basado en YOLOv8 para la detección precisa de la calidad de los arándanos, optimizando el proceso de selección y reduciendo la intervención humana.

Palabras Claves: Visión artificial, redes neuronales convolucionales, YOLOv8, control de calidad, arándanos, clasificación de productos agrícolas.

I. INTRODUCCIÓN

En la industria de los arándanos, la evaluación manual de la calidad enfrenta desafíos debido a la subjetividad en la interpretación de características como tamaño, color y madurez. Esta subjetividad provoca inconsistencias en los estándares de calidad, lo que afecta la selección de productos y, en consecuencia, la satisfacción del cliente. La detección manual de defectos como magulladuras y pudrición es ineficiente, lo que puede llevar a que productos de menor calidad se mezclen con lotes de mayor calidad. Además, la clasificación manual de grandes volúmenes de arándanos consume tiempo y recursos, lo que destaca la necesidad de análisis en tiempo real para mejorar la eficiencia y mantener estándares consistentes.

Por ello, el desarrollo de un algoritmo basado en visión artificial con una red neuronal convolucional es crucial para el control de calidad de los arándanos. Este enfoque permite una evaluación objetiva y consistente, eliminando la subjetividad y la variabilidad humana. Al identificar y separar de manera eficiente los productos de menor calidad, se asegura que solo los mejores arándanos lleguen al mercado, mejorando la satisfacción del cliente y fortaleciendo la reputación de la marca. Además, este método optimiza la eficiencia operativa al reducir la dependencia de la mano de obra manual, permitiendo la clasificación en tiempo real de grandes volúmenes de productos, lo que incrementa la productividad y maximiza el uso de los recursos disponibles.

El uso de algoritmos de visión artificial está revolucionando la industria agrícola al integrar tecnología avanzada con las necesidades prácticas de los productores y

consumidores [1]. Estos algoritmos permiten detectar y clasificar rápidamente los arándanos según su calidad, mejorando la precisión y eficiencia del control de calidad [2,3]. Además, pueden diferenciar entre distintas etapas de maduración utilizando modelos de aprendizaje automático que reconocen características visuales específicas de cada fase [4]. También son capaces de identificar daños, manchas o deformidades, asegurando que solo los frutos de alta calidad sean aptos para exportación [5,6].

Entre los algoritmos de aprendizaje profundo, las redes neuronales convolucionales (CNN) se destacan por su capacidad de percibir detalles locales y compartir parámetros [7]. Estas redes han demostrado ser efectivas en la identificación de características clave como color, forma, tamaño y estado de maduración de los arándanos [8,9,10]. Procesan imágenes capturadas por cámaras especializadas, extrayendo información relevante para el control de calidad, y permiten distinguir rápidamente entre frutos de alta calidad y aquellos con defectos como magulladuras o pudrición [11,12].

YOLO (You Only Look Once) es una de las CNN más utilizadas en aplicaciones de visión por computadora [13]. Estudios han abordado los desafíos de la recolección manual de arándanos en entornos naturales utilizando YOLOv5 para un control de calidad óptimo [14]. También se han entrenado variantes de YOLO para reconocer diferentes niveles de madurez, utilizando conjuntos de datos extensos. Por ejemplo, la red YOLOv3 ha demostrado un rendimiento excelente en la identificación de etapas de madurez en el campo [15]. Otro estudio implementó YOLOv8 para estimar con precisión los rendimientos de arándanos, información vital para optimizar la gestión de cultivos y mejorar la rentabilidad [16].

La implementación de un sistema de visión artificial utilizando CNN para la gestión de calidad de arándanos se basa en la necesidad de automatizar y mejorar el proceso de evaluación de estos frutos en la industria agrícola. La evaluación manual es propensa a errores, lo que afecta la precisión en la selección de productos para el mercado. La visión artificial ofrece una solución innovadora al permitir la detección automática y precisa de características clave como color, tamaño, integridad y madurez, optimizando el control de calidad y reduciendo los costos operativos [17].

El objetivo es desarrollar un algoritmo basado en YOLOv8 para mejorar la eficiencia y precisión en la selección de productos agrícolas, entrenando la red con imágenes de arándanos para que reconozca y clasifique características relevantes de calidad. Se utilizarán funciones de activación avanzadas, como Leaky ReLU, para mejorar el rendimiento en tareas de detección y clasificación. A diferencia de ReLU, Leaky ReLU permite el paso de pequeñas cantidades de valores negativos, lo que ayuda a evitar el problema de los gradientes muertos y mejora la capacidad de la red para aprender patrones de datos complejos. Esta función proporciona una mayor flexibilidad en comparación con ReLU, lo que puede ser beneficioso en la clasificación de imágenes con variaciones sutiles en la calidad de los arándanos.

II. METODOLOGÍA

La etapa inicial del proyecto se centró en la recolección y etiquetado de imágenes de arándanos con diversas calidades, utilizando la plataforma Roboflow. Se capturaron 840 imágenes de arándanos con una cámara de alta resolución, asegurando la diversidad del conjunto de datos mediante distintas condiciones de iluminación y ángulos. La cantidad de imágenes empleadas se basó en un estudio anterior que utilizó 620 imágenes [18]. Con el pasar de los años, las redes neuronales han avanzado notablemente, mejorando su capacidad para aprender rápidamente. Por esta razón, se decidió entrenar el modelo con imágenes de alta calidad y nitidez para obtener resultados más eficientes y favorables [19]. Durante la selección de imágenes, se eliminaron aquellas que no cumplían con los estándares de calidad o que estaban duplicadas. En el proceso de etiquetado, se identificaron gráficamente los arándanos según su calidad utilizando la herramienta Roboflow (Fig.1).



Fig. 1 Etiquetado de arándanos mediante en software Roboflow.

El color y la textura de los arándanos son esenciales para la clasificación, ya que el software detecta y marca los arándanos en las imágenes, asignando etiquetas basadas en su calidad. Esta clasificación visual permite que el modelo aprenda a distinguir entre las diferentes calidades de los arándanos de manera precisa, tal como indican estudios recientes [19,20]. Además, se utilizaron técnicas de aumento de datos como rotación, ajuste

de brillo y contraste para aumentar la variabilidad y robustez del conjunto de datos. Las características recopiladas se basan en criterios visuales y morfológicos, tales como el color, tamaño y textura de los arándanos. Una de las características más notables para reconocer un arándano de alta calidad es su color uniforme y su textura suave, mientras que los arándanos de menor calidad pueden tener colores más claros, texturas más firmes o presentar manchas y una textura blanda [20]. Esta información es fundamental para el entrenamiento de la red neuronal.

YOLO (You Only Look Once) es una serie de modelos de redes neuronales convolucionales (CNN) utilizados para la detección de objetos en imágenes y videos en tiempo real [21]. En este proyecto, se empleó la versión YOLOv8, la iteración más reciente de esta tecnología, reconocida por su alta precisión y eficiencia. YOLOv8 se basa en una arquitectura de red neuronal convolucional (Fig. 2) que permite la detección de objetos con una sola pasada a través de la red, a diferencia de otros enfoques que requieren múltiples etapas de procesamiento. La decisión de utilizar YOLOv8 en comparación con YOLOv5 y EfficientDet fue motivada por sus mejoras significativas en precisión y eficiencia. YOLOv8 supera a YOLOv5 al ofrecer un aumento del 3-5% en mAP y mantener velocidades similares, lo que lo hace ideal para tareas en tiempo real. Frente a EfficientDet, aunque este último destaca por su alta precisión, YOLOv8 logra un mejor equilibrio entre precisión y velocidad, con una arquitectura optimizada para ser más ligera y rápida, sin sacrificar rendimiento.

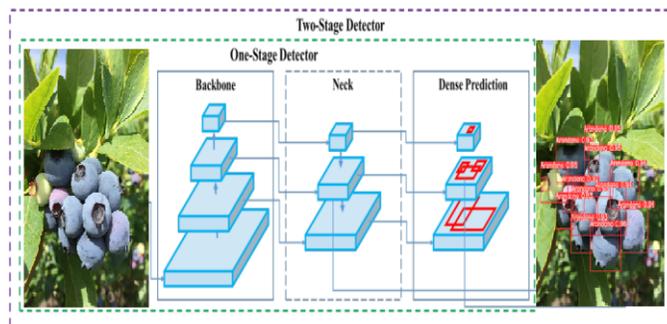


Fig. 2 Arquitectura YOLOv8.

Además, YOLOv8 es fácil de ajustar, lo que lo hace adecuado para aplicaciones personalizadas. [21]. Esta arquitectura se distingue por:

Convoluciones: Estas capas extraen características importantes de las imágenes mediante el uso de filtros que detectan bordes, texturas y patrones.

Capas de Anclaje: YOLOv8 utiliza cajas de anclaje predefinidas para predecir la ubicación y el tamaño de los

objetos en la imagen [21]. Las coordenadas predichas para una caja de anclaje (bx,by,bw,bh) se ajustan de la siguiente manera:

$$bx = \sigma(tx) + cx \quad (1)$$

$$by = \sigma(ty) + cy \quad (2)$$

$$bw = pw e^{\{tw\}} \quad (3)$$

$$bh = ph e^{\{th\}} \quad (4)$$

Donde:

cx,cy: es la ubicación de la celda de la cuadrícula

pw,ph: son las dimensiones de la caja de anclaje predeterminada.

tx,ty,tw,th: son las predicciones de la red.

Predicción Simultánea: A diferencia de otros métodos que primero generan propuestas de regiones y luego clasifican, YOLOv8 realiza la predicción de las clases y las ubicaciones de los objetos simultáneamente, lo que mejora la velocidad y la eficiencia [21].

Función de Pérdida: La función de pérdida de YOLOv8 combina la precisión de la clasificación y la precisión de la localización en un único valor escalar, optimizando así ambos aspectos durante el entrenamiento. La pérdida total L se define como:

$$L = Lloc + Lconf + Lcls \quad (5)$$

Lloc es la pérdida de localización, que mide el error en las predicciones de las coordenadas de las cajas. Lconf es la pérdida de confianza, que mide el error en las predicciones de la confianza de que una caja contiene un objeto y Lcls es la pérdida de clasificación, que mide el error en las predicciones de las clases de los objetos. Durante el entrenamiento del modelo YOLOv8, la función de pérdida se implementa y ajusta de manera cuidadosa para equilibrar los componentes de localización, confianza y clasificación. Se utiliza la pérdida GIoU para mejorar la precisión de las coordenadas de las cajas, la pérdida de entropía cruzada binaria para evaluar la confianza en la presencia de objetos y la pérdida de entropía cruzada categórica para clasificar correctamente los objetos.

El uso del optimizador Adam en el proyecto de detección y clasificación de arándanos con YOLOv8 implica la actualización adaptativa de los pesos del modelo para minimizar la función de pérdida durante el entrenamiento. El proceso comienza con la evaluación de la pérdida, que mide la diferencia entre las predicciones del modelo y las etiquetas reales (arándanos buenos o malos). Luego, mediante backpropagation, se calculan los gradientes de la pérdida con respecto a los pesos. Adam utiliza estos gradientes junto con estimaciones de primer y segundo momento para ajustar los pesos del modelo en cada iteración, reduciendo gradualmente la pérdida. Este ajuste continuo es visualizado a través de la

disminución de la función de pérdida a lo largo del tiempo, como se muestra en el gráfico adjunto. Hemos utilizado Adam debido a su capacidad para adaptarse a los cambios en la tasa de aprendizaje, lo que acelera la convergencia y mejora la precisión del modelo en comparación con otros optimizadores.

La convolución es un proceso matemático que fusiona dos conjuntos de datos. En el ámbito de las CNNs, se trata de aplicar un filtro (o núcleo) sobre una imagen para identificar características significativas [22]. Este filtro se desplaza por la imagen de entrada y, en cada posición, calcula un valor de salida basado en una suma ponderada de los valores de la imagen y los del filtro. Durante el entrenamiento de YOLOv8, las capas convolucionales se utilizan para extraer características visuales cruciales de las imágenes de arándanos [22]. Estas características pueden incluir bordes, texturas, colores y formas que reflejan la calidad de los arándanos.

La activación del kernel es un aspecto esencial en este proceso. Cada vez que el filtro se desplaza sobre la imagen, se calcula una función de activación que determina la salida de la neurona en esa posición específica. Esta activación permite que la red neuronal distinga entre las características importantes y las irrelevantes, mejorando la precisión en la detección de patrones relevantes en los arándanos.

La activación comúnmente utilizada en las CNNs es la función Leaky ReLU (Rectified Linear Unit con pendiente pequeña para valores negativos), que introduce no linealidad en el modelo, permitiendo que aprenda representaciones más complejas. A diferencia de ReLU, Leaky ReLU permite el paso de valores negativos pequeños, evitando que las neuronas queden inactivas durante el entrenamiento. Esto es especialmente útil en redes profundas, donde el problema de los gradientes muertos puede afectar negativamente el rendimiento. Con Leaky ReLU, el modelo conserva cierta sensibilidad a las entradas negativas, lo que mejora su capacidad de aprender patrones más ricos y detallados.

En el primer bloque, denominado Kernel D-Conv (Fig.3), se utilizan tres filtros de tamaño 3x3 ($K1, K2, K3$) que aplican operaciones convolucionales sobre las entradas ($I1, I2, I3$), generando mapas de características intermedias ($L1, L2, L3$). Estas características son luego procesadas en el segundo bloque, Kernel P-Conv, que utiliza filtros de convolución 1x1 ($K1, K2, K3$) para combinar y refinar la información obtenida en la etapa anterior. Este diseño permite una eficiente extracción y combinación de características, resultando en salidas ($O1, O2, O3$) que representan la imagen procesada, mejorando la precisión en la clasificación basada en las características visuales capturadas por la red convolucional profunda.

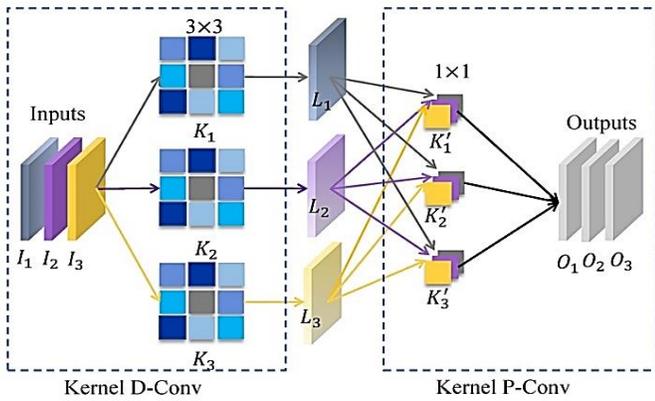


Fig. 3 Kernel convolucional Yolov8 denominado Kernel D-Conv

Para mejorar aún más la eficiencia, YOLOv8 incorpora convoluciones separables en profundidad (depthwise separable convolutions), donde una convolución profunda es seguida por una convolución puntual (pointwise convolution). Esta técnica reduce significativamente el número de cálculos necesarios y los parámetros del modelo, sin comprometer la precisión.

Finalmente, YOLOv8 utiliza un bloque de salida específico diseñado para realizar la detección de objetos. Este bloque consiste en múltiples capas totalmente conectadas que generan las predicciones finales, incluyendo las coordenadas del cuadro delimitador (bounding box), la clase del objeto y la confianza en la predicción. La combinación de estas técnicas avanzadas y la arquitectura optimizada permite a YOLOv8 ofrecer una detección de alta precisión y velocidad, haciendo de esta versión una herramienta poderosa para aplicaciones de visión por computadora en tiempo real.

Los filtros convolucionales son pequeñas matrices aplicadas a las imágenes. Cada filtro se entrena para reconocer características específicas (Fig. 4). Estudios previos han demostrado su eficacia en la clasificación y detección de diversas características en frutas, subrayando su utilidad en el análisis de productos agrícolas [23]. En el caso de los arándanos, algunos filtros pueden aprender a detectar el color uniforme y la textura suave de un arándano de alta calidad, mientras que otros pueden identificar texturas más firmes o manchas que indican una calidad inferior.

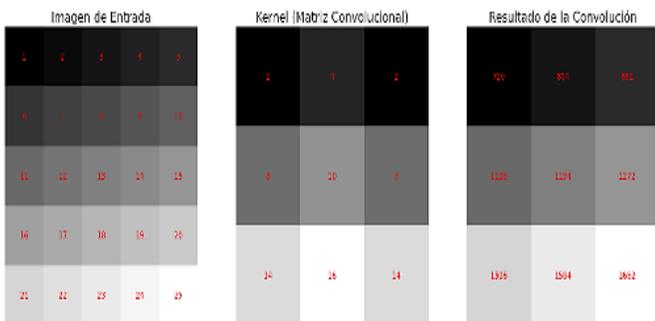


Fig. 4 Filtros convolucionales.

Para el entrenamiento del modelo YOLOv8, se aprovechó la potencia computacional de Google Colab, que ofrece recursos especializados para el entrenamiento de redes neuronales, incluyendo GPUs. En particular, se utilizó una GPU Tesla T4 con 15 GB de memoria, compatible con CUDA, lo que permitió acelerar significativamente el proceso de entrenamiento. La capacidad de procesamiento de esta GPU aseguró una rápida convergencia del modelo y tiempos de entrenamiento eficientes.

Una vez configurado el entorno en Google Colab, se importaron las bibliotecas necesarias como torch y yolov8 para el desarrollo y entrenamiento del modelo. PyTorch, importado como torch, es una biblioteca de código abierto muy popular para el aprendizaje profundo. Es especialmente conocida por su facilidad de uso y flexibilidad, lo que la hace ideal para la investigación y el desarrollo de modelos de redes neuronales. Su eficacia y versatilidad han sido respaldadas por numerosas revisiones [24,25]. Además, se montó Google Drive para facilitar el almacenamiento y acceso a los conjuntos de datos y modelos.

El proceso de entrenamiento se llevó a cabo en Google Colab, donde se descargaron los datos etiquetados desde Roboflow y se cargaron en el entorno de trabajo. Posteriormente, se configuró el modelo YOLOv8 ajustando parámetros como el tamaño del lote, la tasa de aprendizaje y el número de épocas, fijando este último en un total de 250. Esta configuración se fundamentó en indagaciones anteriores que demostraron la eficacia de este número de épocas para optimizar el rendimiento del modelo. Durante el entrenamiento, se monitorearon métricas como la pérdida de entrenamiento y la precisión en el conjunto de validación, permitiendo ajustes en los hiperparámetros según fuera necesario para mejorar el rendimiento del modelo.

La figura 5 muestra una serie de gráficos de pérdida y métricas de evaluación correspondientes al entrenamiento de un modelo YOLOv8. En la fila superior, los gráficos de la izquierda muestran la disminución de la pérdida de caja delimitadora (train/box_loss), la pérdida de clasificación (train/cls_loss) y la pérdida de objeto (train/obj_loss) durante el entrenamiento, indicando una convergencia efectiva del modelo. Los gráficos de la derecha presentan las métricas de precisión (metrics/precision[B]) y recuperación (metrics/recall[B]), las cuales aumentan consistentemente, reflejando una mejora en la capacidad del modelo para identificar correctamente los objetos. En la fila inferior, los gráficos de pérdida de validación (val/box_loss, val/cls_loss, val/obj_loss) también muestran una tendencia decreciente, lo que sugiere que el modelo no está sobreajustando. Finalmente, los gráficos de las métricas de mAP (metrics/mAP50(B), metrics/mAP50-95(B)) evidencian un incremento sostenido, confirmando que el modelo mantiene un buen desempeño

generalizado en la tarea de detección de objetos a lo largo de las iteraciones del entrenamiento.

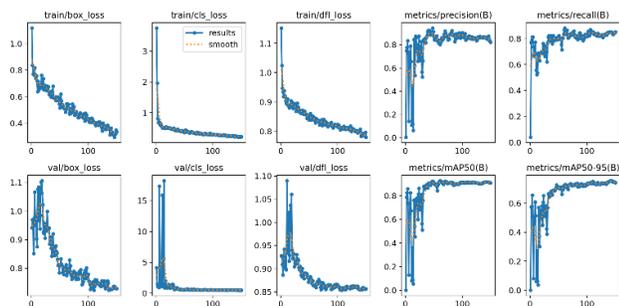


Fig. 5 Resultados del entrenamiento.

En los resultados obtenidos, se pueden observar las métricas de desempeño tanto del conjunto de datos de entrenamiento como del de validación (Fig.6), lo cual es fundamental para evaluar la efectividad del modelo. La primera imagen ilustra la fase de entrenamiento, donde se muestran las detecciones y clasificaciones realizadas por el modelo en el conjunto de datos de entrenamiento, evidenciadas por los cuadros delimitadores y las etiquetas de confianza. La segunda imagen presenta los resultados de la fase de validación, demostrando la capacidad del modelo para generalizar sus predicciones en datos no vistos previamente, con similares representaciones visuales de las detecciones y sus respectivas puntuaciones de confianza. Estos resultados permiten una evaluación exhaustiva del rendimiento del modelo, facilitando la identificación de áreas de mejora y la optimización de los parámetros de entrenamiento.

garantizó la robustez y precisión del modelo en la detección de la madurez de los arándanos.

Para implementar el algoritmo de detección de la calidad de los arándanos, se empleó Python en el entorno de desarrollo PyCharm. Se configuró PyCharm con un entorno virtual y se instalaron las dependencias necesarias, como opencv-python y pytorch [26]. Se cargó el modelo entrenado de YOLOv8 y se desarrolló un script en Python para procesar nuevas imágenes de arándanos, ya sea capturando imágenes desde una cámara o cargándolas desde el sistema de archivos. Utilizando el modelo YOLOv8, se detectaron y clasificaron los arándanos en las imágenes según su calidad, mostrando los resultados mediante cuadros delimitadores y etiquetas sobre las imágenes originales (Fig. 7).

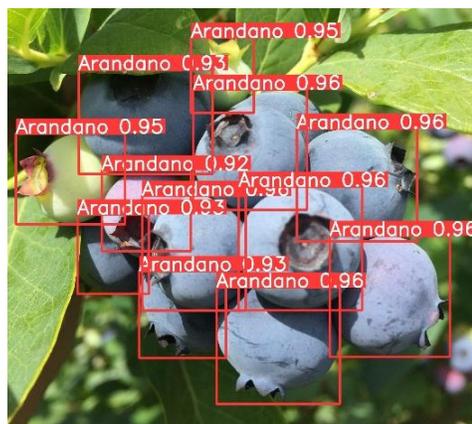


Fig. 7 Prueba de la red neuronal YOLOv8 para la detección de arándanos.



Fig. 6 Resultados de la red neuronal para la clasificación de arándanos según su calidad

Al culminar el proceso de entrenamiento, se evaluó el modelo utilizando un conjunto de datos de prueba separado, asegurando que no hubiera sobreajuste y validando su capacidad para generalizar a datos no vistos. Este enfoque

El proceso integral de entrenamiento y aplicación del modelo destaca cada etapa crucial que contribuye a su alta eficacia en la identificación y clasificación automática de objetos. En la fase de entrenamiento, el modelo se nutre de un vasto conjunto de datos etiquetados que le permiten aprender y reconocer patrones relevantes. Este aprendizaje se ajusta mediante técnicas avanzadas de machine learning, optimizando los parámetros para maximizar la precisión y eficiencia del modelo. Una vez completado el entrenamiento, el modelo se implementa en el entorno de aplicación, donde se enfrenta a datos no vistos anteriormente. Aquí, su capacidad para identificar y clasificar objetos se pone a prueba, demostrando su habilidad para realizar estas tareas de manera rápida y precisa. Los criterios de calidad previamente definidos juegan un papel crucial, ya que el modelo utiliza estos estándares para evaluar y categorizar cada objeto de manera consistente y fiable.

III. RESULTADOS

Los resultados indican que el sistema de visión artificial basado en YOLOv8 es altamente efectivo para detectar la

calidad de los arándanos, clasificándolos como buenos o malos. La precisión varió entre el 91% y el 98%, y los F1-scores entre 92% y 99%. Estos resultados subrayan la capacidad robusta del modelo para clasificar arándanos de distintas calidades de manera consistente a través de diversas zonas, cada una representando diferentes entornos o condiciones de prueba. La elección de siete zonas fue suficiente para cubrir una variedad razonable de condiciones sin crear una carga excesiva en términos de recolección de datos o complejidad en la evaluación del sistema.

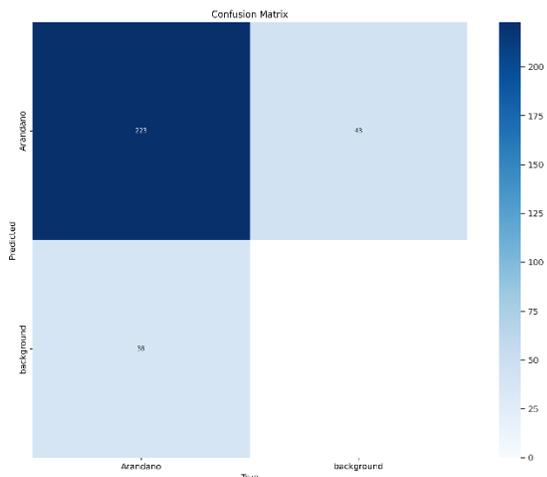


Fig. 8 Matriz de confusión

Los valores en la diagonal principal indican el número de predicciones correctas, con 223 instancias de "Arándano" correctamente clasificadas y 38 instancias de "Background" correctamente identificadas. Sin embargo, la matriz también revela errores en la clasificación: 43 instancias de "Arándano" fueron erróneamente etiquetadas como "Background". Este análisis sugiere que, aunque el modelo tiene una alta precisión en la identificación de "Arándano", hay un margen de mejora en la precisión para la clase "Background", evidenciado por la confusión existente entre las clases (Fig. 8).

De las siete zonas de producción de arándanos, cada una estuvo dividida en categorías de calidad ("Bueno" y "Malo") (Fig. 9).

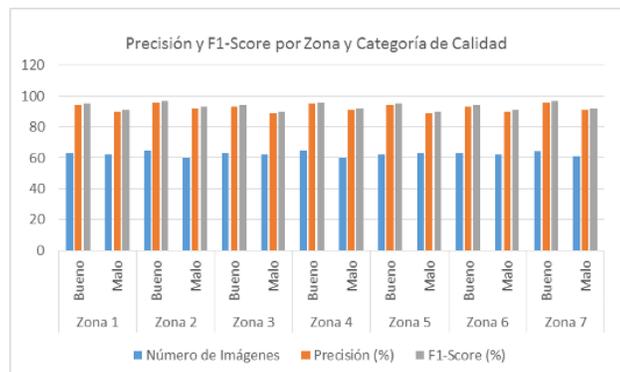


Fig. 9 Precisión y F1-Score por Zona y Categoría de Calidad

IV. DISCUSIONES

Se evaluaron siete zonas de producción de arándanos, clasificando las imágenes en "Bueno" y "Malo", y se observaron variaciones en los resultados de precisión entre las diferentes zonas estudiadas. En este estudio, los resultados destacan que el sistema de visión artificial basado en YOLOv8 fue altamente efectivo para la clasificación de arándanos en ambas categorías. La precisión promedio para arándanos "Bueno" fue del 98 % y para arándanos "Malo" fue del 91%. Estos hallazgos no solo proporcionan una visión detallada de la calidad del producto en distintas regiones, sino que también permiten comparaciones significativas con investigaciones previas en el campo de la visión artificial aplicada a la agricultura. Comparando los resultados obtenidos en nuestro estudio con la literatura existente sobre sistemas de visión artificial para la clasificación de arándanos, se observa una coherencia en la tendencia hacia altos niveles de precisión. Por ejemplo, en una investigación anterior se propuso una identificación automática de arándanos utilizando el algoritmo de aprendizaje automático YOLOv5, alcanzando una precisión del 93%. No obstante, nuestro proyecto logró una precisión del 98%, gracias a un proceso exhaustivo de evaluación y ajuste iterativo, lo que garantizó que el modelo no solo fuera preciso sino también consistente en su desempeño [14]. Asimismo, otro estudio evaluó el rendimiento de la visión artificial para la clasificación de arándanos silvestres en los arbustos antes de la cosecha utilizando YOLOv3, obteniendo una precisión del 91% [15]. Sin embargo, nuestro proyecto fue superior porque utilizamos YOLOv8, una versión más avanzada y optimizada de la arquitectura YOLO. Esta versión incluye mejoras en la detección de objetos y el procesamiento de imágenes, lo que resulta en una mayor precisión y velocidad. Además, empleamos funciones de activación avanzadas como Leaky ReLU, permitiendo a nuestra red aprender y representar características complejas de manera más efectiva. Leaky ReLU mejora el aprendizaje al evitar que las neuronas queden inactivas frente a entradas negativas, manteniendo una pequeña pendiente para esos valores. Esto proporciona una mayor capacidad de representación y un mejor rendimiento en modelos profundos, especialmente en escenarios donde es crucial no perder información de los datos, como en la detección y clasificación de objetos complejos.

Además, el desempeño consistente a través de diferentes zonas subraya la robustez del modelo en diversas condiciones agroindustriales, lo cual es crucial para su aplicabilidad en escenarios del mundo real. Sin embargo, es importante considerar que la efectividad de cualquier sistema de visión artificial puede verse afectada por factores como la iluminación, la calidad de las imágenes capturadas y la variabilidad inherente en los productos agrícolas. Por lo tanto, futuras investigaciones podrían centrarse en la optimización continua de los algoritmos y la integración de tecnologías avanzadas para mejorar aún más la precisión y la confiabilidad en la clasificación automatizada de arándanos.

En cuanto a implicaciones prácticas, es crucial continuar optimizando los algoritmos para mejorar la precisión y confiabilidad en la clasificación automatizada de arándanos. Además, explorar la integración de tecnologías avanzadas podría mitigar desafíos como la variabilidad de iluminación y calidad de imagen. La adaptación de estos sistemas para otros cultivos también podría ampliar significativamente su utilidad en la agricultura de precisión, promoviendo mejoras sustanciales en la implementación y aplicación práctica de la visión artificial en la industria agrícola.

V. CONCLUSIONES

Esta investigación ha desarrollado con éxito un algoritmo basado en visión artificial utilizando la red neuronal convolucional YOLOv8 para el control automatizado de la calidad de arándanos, mejorando la eficiencia y precisión en la selección de productos agrícolas. La red fue entrenada utilizando datos de imágenes de arándanos, lo que le permite reconocer y clasificar características relevantes de calidad. Además, se emplearon funciones de activación avanzadas, como Leaky ReLU, permitiendo a la red aprender y representar características complejas en las imágenes de manera más efectiva. A diferencia de la ReLU tradicional, Leaky ReLU permite el flujo de valores negativos pequeños, evitando que las neuronas queden inactivas, lo que es crucial para mantener un aprendizaje robusto en redes profundas. Esto optimiza el desempeño en tareas de detección y clasificación, ya que la red puede captar detalles sutiles y características más diversas en las imágenes.

El sistema de visión artificial basado en YOLOv8 demostró ser altamente efectivo para la detección de la calidad de los arándanos, con una precisión del 98%. Su implementación para el control de calidad del arándano representa un avance significativo en el ámbito de la agricultura y la tecnología de alimentos. Este desarrollo no solo optimiza el proceso de selección y clasificación de los frutos, asegurando que los productos que llegan al consumidor cumplan con los más altos estándares de calidad, sino que también reduce la intervención humana, minimizando errores y mejorando la eficiencia operativa. La aplicación de las CNN permite una identificación precisa y rápida de defectos y características deseables en los arándanos, destacando el potencial de la inteligencia artificial para revolucionar prácticas agrícolas y asegurar productos de mejor calidad en el mercado. Para futuros trabajos, se recomienda adaptar el modelo para operar en tiempo real durante la recolección, así como integrarlo con sistemas de gestión agrícola y explorar su aplicación en otros cultivos. Estas mejoras tienen el potencial de ampliar considerablemente su impacto en la mejora de la calidad y eficiencia.

El aporte científico de esta investigación radica en la aplicación innovadora de la red neuronal YOLOv8 en el ámbito agrícola, contribuyendo al avance del conocimiento en la integración de inteligencia artificial y agricultura de precisión. Este estudio ofrece una metodología robusta para el entrenamiento y validación de modelos de visión artificial en tareas de clasificación de calidad, proporcionando una base científica para futuras investigaciones en la automatización de procesos agrícolas. Asimismo, los resultados obtenidos validan el uso de técnicas avanzadas de procesamiento de imágenes y aprendizaje profundo en entornos reales, abriendo nuevas oportunidades para el desarrollo de tecnologías que mejoren la productividad y sostenibilidad del sector agrícola.

REFERENCIAS

- [1] Rematales J, & Hancock J. 2018. Blueberries, 2nd ed. Wallingford, UK,: CABI Publishing. <https://doi.org/10.1017/S001447971200049X>
- [2] Pathan M, et al. Artificial cognition for applications in smart agriculture: A comprehensive review. *Artificial Intelligence in Agriculture*, 2020, vol. 4, p. 81-95. <https://doi.org/10.1016/j.aiia.2020.06.001>
- [3] Bhargava, A; Bansal, A. Fruits and vegetables quality evaluation using computer vision: A review. *Journal of King Saud University-Computer and Information Sciences*, 2021, vol. 33, no 3, p. 243-257. <https://doi.org/10.1016/j.jksuci.2018.06.002>
- [4] Horna W, Manzano A. Sistema de visión artificial basado en redes neuronales convolucionales para la selección de arándanos según estándares de exportación. *Campus*. 2021; 26(32). <https://doi.org/10.24265/campus.2021.v26n31.0>
- [5] Bao Y, Sun Y. Fruit Quality Inspection and Classification Based on Computer Vision: A Review. *Journal of Food Engineering*. 2020; 259: 21-30. [https://doi.org/10.1016/S0260-8774\(03\)00183-3](https://doi.org/10.1016/S0260-8774(03)00183-3)
- [6] Ribeiro A, Mayer F, Barbedo J. Review of Computer Vision Tools in the Agro-Food Sector. *Computers and Electronics in Agriculture*. 2017; 143: 112-122. <https://www.sciencedirect.com/science/article/abs/pii/S1537511020300301>
- [7] Momeny, M., Jahanbakhshi, A., Jafarnejad, K., and Zhang, Y. (2020). Accurate classification of cherry fruit using deep CNN based on hybrid pooling approach. *Postharvest Biol. Technol.* 166:111204. doi: 10.1016/j.postharvbio.2020.111204
- [8] Zhang M, et al. Redes totalmente convolucionales para hematomas de arándanos y segmentación del cáliz mediante imágenes de transmitancia hiperspectral. *Ingeniería de Biosistemas*, 2020, vol. 192, pág. 159-175. <https://www.sciencedirect.com/science/article/pii/S2772375522000648>
- [9] Maceachern, Craig B., et al. Detection of fruit maturity stage and yield estimation in wild blueberry using deep learning convolutional neural networks. *Smart Agricultural Technology*, 2023, vol. 3, p. 100099. <https://doi.org/10.1016/j.atech.2022.100099>
- [10] Ukwuoma, Chiagoziem C., et al. Avances recientes en la detección y clasificación de frutas mediante técnicas de aprendizaje profundo. *Problemas matemáticos en ingeniería*, 2022, vol. 2022, pág. 1-29. <https://doi.org/10.1155/2022/9210947>
- [11] NI, Xueping, et al. A deep learning-based web application for segmentation and quantification of blueberry internal bruising. *Computers and Electronics in Agriculture*, 2022, vol. 201, p. 107200. <https://doi.org/10.1016/j.compag.2022.107200>
- [12] QIAO, Shicheng, et al. Detección y clasificación de la descomposición temprana en arándanos basada en una red neuronal convolucional 3D residual profunda mejorada en imágenes hiperspectrales. *Programación científica*, 2020, vol. 2020. <https://doi.org/10.1155/2020/8895875>
- [13] WANG, Chenglin y col. Aplicación de métodos de detección basados en redes neuronales convolucionales en la producción de fruta fresca: una

- revisión exhaustiva. *Fronteras en la ciencia vegetal*, 2022, vol. 13, pág. 868745. <https://doi.org/10.3389/fpls.2022.868745>
- [14] Zhai, Y et al. Método de detección de arándanos basado en aprendizaje automático por CIE-YOLOv5. En Conferencia Internacional sobre Gráficos por Computadora, Inteligencia Artificial y Procesamiento de Datos (ICCAID 2023). SPIE, 2024. pág. 462-467. <https://doi.org/10.1117/12.3026626>
- [15] Schumann, Arnold W., et al. Detection of three fruit maturity stages in wild blueberry fields using deep learning artificial neural networks. En 2019 ASABE Annual International Meeting. American Society of Agricultural and Biological Engineers, 2019. p. 1. doi:10.13031/aim.201900533
- [16] Zhengkun L, Changying L; Muñoz Patricio. Estimación del rendimiento de arándanos mediante imágenes multivista con detección de objetos YOLOv8. En 2023 Reunión Internacional Anual de ASABE. Sociedad Estadounidense de Ingenieros Agrícolas y Biológicos, 2023. p. 1. doi:10.13031/aim.202300883
- [17] Kamilaris, A; Prenafeta F. Francesc X. A review of the use of convolutional neural networks in agriculture. *The Journal of Agricultural Science*, 2018, vol. 156, no 3, p. 312-322. <https://doi.org/10.1017/S0021859618000436>
- [18] Ayan P, et al. Smart solutions for capsicum Harvesting: Unleashing the power of YOLO for Detection, Segmentation, growth stage Classification, Counting, and real-time mobile identification. *Computers and Electronics in Agriculture*, 2024, vol. 219, p. 108832. <https://doi.org/10.3390/agriculture14030331>
- [19] HENNESSY, Patrick J., et al. Evaluation of cameras and image distance for CNN-based weed detection in wild blueberry. *Smart Agricultural Technology*, 2022, vol. 2, p. 100030. <https://doi.org/10.1016/j.atech.2021.100030>
- [20] LI, Zhengkun; LI, Changying; Muñoz, Patricio. Blueberry Yield Estimation Through Multi-View Imagery with YOLOv8 Object Detection. En 2023 ASABE Annual International Meeting. American Society of Agricultural and Biological Engineers, 2023. p. 1. doi:10.13031/aim.202300883
- [21] Xiuyan, G. A. O.; zhang, Yanmin. Detection of Fruit using YOLOv8-based Single Stage Detectors. *International Journal of Advanced Computer Science & Applications*, 2023, vol. 14, no 12. doi: 10.14569/ijacsa.2023.0141208
- [22] LIU, Yao; PU, Hongbin; SUN, Da-Wen. Efficient extraction of deep image features using convolutional neural network (CNN) for applications in detecting and analysing complex food matrices. *Trends in Food Science & Technology*, 2021, vol. 113, p. 193-204. <https://doi.org/10.1016/j.tifs.2021.04.042>
- [23] Quiroz, Ignacio A.; Alférez, Germán H. Image recognition of Legacy blueberries in a Chilean smart farm through deep learning. *Computers and Electronics in Agriculture*, 2020, vol. 168, p. 105044. <https://doi.org/10.1016/j.compag.2019.105044>
- [24] M. Yi, C. Zhao, F. Liao and W. Yao, "Classification of Blueberry Varieties Based on Improved EfficientNet," 2022 4th International Academic Exchange Conference on Science and Technology Innovation (IAECST), Guangzhou, China, 2022, pp. 411-415, doi: 10.1109/IAECST57965.2022.10062152.
- [25] P. Brayam Cubas Muñoz, E. Martín Fiestas Sorogastua and S. Ricardo Prado Gardini, "Detection and Classification of ventura-blueberries in five levels of ripeness from images taken during pre-harvest stage using Deep Learning techniques," 2022 IEEE ANDESCON, Barranquilla, Colombia, 2022, pp. 1-6, doi: 10.1109/ANDESCON56260.2022.9989578
- [26] Xiao, Feng, et al. A Lightweight Detection Method for Blueberry Fruit Maturity Based on an Improved YOLOv5 Algorithm. *Agriculture*, 2023, vol. 14, no 1, p. 36. <https://doi.org/10.3390/agriculture14010036>