

Desarrollo de un Sistema de Visión Artificial con la Red Neuronal Convolutiva (YOLO v8) para Clasificar el Arándano por su Grado de Madurez

Ryan A LEÓN LEÓN

La Libertad-Universidad Privada del Norte
Trujillo, 13011, Perú

Carol A. BARRANTES VARGAS

La Libertad- UPN, Universidad Privada del Norte
Trujillo, 13011/ Perú

Fernando D. BACILIO DE LA CRUZ

La Libertad- UPN, Universidad Privada del Norte
Trujillo, 13011/ Perú

RESUMEN: En el contexto de la agricultura de precisión, el uso de sistemas de visión artificial ha cobrado importancia para optimizar la clasificación de productos agrícolas, especialmente los arándanos. Este estudio se centra en la implementación de un sistema basado en la red neuronal convolutiva YOLO v8 para clasificar arándanos según su grado de madurez. La precisión en la clasificación de la madurez de los arándanos es fundamental para garantizar su calidad y optimizar su valor comercial. La metodología incluye la recolección y etiquetado de 666 imágenes de arándanos en tres etapas de madurez: inmaduro, pintón y maduro, utilizando la plataforma Roboflow. Se aplicaron técnicas de aumentación de datos para mejorar la variabilidad del conjunto de datos. El modelo YOLOv8, entrenado en Google Colab con una GPU Tesla T4, mostró alta precisión y eficiencia en la clasificación, demostrando una precisión superior al 95%. La arquitectura YOLOv8 permite la detección simultánea de objetos y su clasificación mediante convoluciones y capas de anclaje, optimizando la evaluación del estado de madurez de los arándanos. A pesar de los resultados prometedores de YOLOv8, se necesita más investigación para ampliar su aplicación. Es crucial aumentar la diversidad del conjunto de datos y mejorar el entrenamiento para mantener la precisión en condiciones adversas. Integrar YOLOv8 con otros sistemas de IA, adaptarlo a otros cultivos, desarrollar interfaces amigables y realizar estudios de costo-beneficio son esenciales. Estas mejoras pueden aumentar la eficiencia y precisión del sistema, promoviendo prácticas agrícolas más sostenibles y rentables.

Palabras Claves: Agricultura de precisión, visión artificial, YOLO v8, Red neuronal convolutiva, clasificación de madurez, arándanos.

Development of an artificial vision system with the convolutional neural network (YOLO v8) to classify blueberry by its degree of maturity

ABSTRACT: In the context of precision agriculture, the use of computer vision systems has gained importance in optimizing the classification of agricultural products, especially blueberries. This study focuses on the implementation of a system based on the YOLOv8 convolutional neural network to classify blueberries according to their maturity level. The accuracy in classifying blueberry maturity is crucial to ensure their quality and optimize their commercial value. The methodology includes the collection and labeling of 666 images of blueberries in three stages of maturity: unripe, ripening, and ripe, using the Roboflow platform. Data augmentation techniques were applied to improve the variability of the dataset. The YOLOv8 model, trained in Google Colab with a Tesla T4 GPU, demonstrated high accuracy

and efficiency in classification, achieving an accuracy of over 95%. The YOLOv8 architecture allows for simultaneous object detection and classification through convolutions and anchor layers, optimizing the evaluation of blueberry maturity. Despite the promising results of YOLOv8, further research is needed to expand its application. It is crucial to increase the diversity of the dataset and improve training to maintain accuracy under adverse conditions. Integrating YOLOv8 with other AI systems, adapting it to other crops, developing user-friendly interfaces, and conducting cost-benefit studies are essential. These improvements can enhance the system's efficiency and accuracy, promoting more sustainable and profitable agricultural practices.

Keywords: Precision agriculture, computer vision, YOLO v8, Convulsive Regional Network, maturity classification, blueberries.

1. INTRODUCCIÓN

En el contexto actual de la agricultura de precisión, el uso de sistemas de visión artificial se ha vuelto esencial para optimizar la clasificación de productos agrícolas, especialmente los arándanos. La problemática asociada a la clasificación manual de estos frutos destaca la necesidad de incorporar tecnologías de la industria 4.0 en el sector agrícola. Es por ello que, la integración de sistemas de visión artificial mejora la precisión y eficiencia mientras impulsa la digitalización y la toma de decisiones basada en datos en tiempo real. Esto enfatiza la necesidad imperativa de integrar sistemas de automatización.

Este estudio se enfoca en la implementación de un sistema de visión artificial basado en la red neuronal convolutiva YOLO v8, con el propósito específico de clasificar el arándano según su grado de madurez. La capacidad de este sistema para identificar y distinguir niveles de madurez en los arándanos no solo promete mejorar la eficiencia de la cadena de producción, sino también garantizar una calidad óptima del producto final, satisfaciendo así las demandas del mercado y las necesidades de los productores agrícolas. Asimismo, hoy en día, los arándanos son reconocidos a nivel mundial por su contenido vitamínico, cualidades antioxidantes y ventajas para la salud cardiovascular, convirtiéndose en uno de los cultivos con mayor proyección productiva y comercial [1,2]. La clasificación de los arándanos según su nivel de madurez es una etapa crucial en sector agroindustrial, ya que ayuda a determinar el mejor estado y calidad de los frutos para su comercialización [3]. La precisión en esta tarea permite garantizar que los arándanos lleguen al mercado en su punto óptimo, maximizando su valor y

satisfaciendo las expectativas de los consumidores en términos de sabor, textura y apariencia [4]. Para dicha etapa, hacer uso exclusivamente de la clasificación manual puede resultar en ineficiencias operativas, costos elevados, baja calidad del producto y limitaciones en la capacidad de tomar decisiones estratégicas en la industria agrícola [5]. Es por ello que, actualmente se emplea la inteligencia artificial, representando un avance significativo en esta industria. La implementación de un sistema de clasificación por visión artificial permite una evaluación más precisa y exhaustiva del nivel de maduración de los arándanos, optimizando el cumplimiento de estándares de calidad y uniformidad de los lotes destinados a exportación [6,7]. El uso de algoritmos en los últimos años, ha sido ampliamente explorado por diversos investigadores quienes han desarrollado numerosos algoritmos para clasificar frutas mediante visión por computadora, siendo el color, la textura y la morfología las características más comúnmente utilizadas para identificar el estado de madurez [8,9]. Las investigaciones actuales han identificado YOLO como una de las redes neuronales convolucionales más prometedoras para la clasificación precisa del grado de madurez de los arándanos. Esta red neuronal, junto con sus variantes y versiones, ha sido entrenada con conjuntos de datos específicos que contienen imágenes de arándanos en diferentes estados de madurez, además también se aplicó la técnica de aumentación de datos. [10,14]. Además, estudios previos también han explorado diversas arquitecturas de redes neuronales para detectar y clasificar el estado de madurez de arándanos, utilizando técnicas avanzadas de aprendizaje profundo y visión por computadora. La aplicación de técnicas de visión artificial en la industria agrícola ha experimentado progresos significativos, especialmente en la clasificación automatizada de frutas según su estado de madurez. [11,12]. Por ejemplo, Zhu et al. [13] propusieron un algoritmo basado en R-CNN para la detección y el reconocimiento de frutos de arándanos, que fue capaz de identificar los arándanos y distinguir su madurez con mayor precisión y rapidez en condiciones soleadas, teniendo un excelente efecto de reconocimiento [13]. MacEachern et al. [14] aplicaron YOLOv4 a la detección de la madurez de los arándanos y demostraron que el algoritmo tiene una alta precisión para la detección de su madurez. Yang et al. [15] propusieron un algoritmo ligero para la detección de frutos de arándanos que utiliza un innovador mecanismo de atención de bloques convolucionales. Al combinar este enfoque con la red neuronal YOLOv5, lograron mejorar la precisión en la detección de arándanos. Asimismo, un estudio reciente empleó YOLOv5 mejorado para reconocer de manera precisa y rápida las etapas de madurez de los arándanos [16]. En este sentido, los algoritmos y las técnicas de visión por computadora en conjunto están revolucionando el manejo y procesamiento de los arándanos, brindando soluciones innovadoras y eficientes para la evaluación y selección de estos productos agrícolas [17].

El cultivo de arándanos ha experimentado un crecimiento significativo en las últimas décadas debido a la creciente demanda mundial, impulsada por sus beneficios para la salud. Según datos de la FAO, la producción mundial de arándanos ha aumentado en más del 50% en los últimos diez años, situándose en aproximadamente 1.5 millones de toneladas anuales. Los principales países productores incluyen Estados Unidos, Canadá, Polonia, Chile y Perú, con este último emergiendo como un importante exportador debido a sus condiciones climáticas favorables y su capacidad para producir arándanos fuera de temporada para el hemisferio norte. Este crecimiento no solo beneficia a los grandes productores y exportadores, sino también a pequeñas y medianas explotaciones agrícolas que dependen de

la calidad y clasificación precisa de sus productos para acceder a mercados internacionales y obtener precios competitivos. Por lo tanto, este proyecto se justifica por la necesidad de optimizar la clasificación de arándanos mediante la implementación de un sistema de visión artificial basado en la red neuronal YOLOv8, alineándose con los estándares actuales de la Industria 4.0. En este contexto, la automatización promete uniformidad y fiabilidad en el mercado, destacando la integración de tecnologías de IA y visión por computadora para mejorar la eficiencia y calidad en la producción y clasificación de arándanos.

Esta investigación, tiene como objetivo desarrollar un sistema de visión artificial con la red neuronal convolucional YOLOv8 para clasificar los arándanos según su grado de madurez. Además, se han definido tres objetivos específicos: implementar DATASET para la recopilación de imágenes de arándanos en diversos grados de madurez, el entrenamiento de la red neuronal YOLOv8 con los datos obtenidos, y testear el algoritmo desarrollado.

2. METODOLOGÍA

El proyecto consistió en la recolección y etiquetado de imágenes de arándanos según su grado de madurez, utilizando la plataforma Roboflow. El objetivo de la recolección de imágenes es para observar la forma, dimensión y estructura del arándano y detectar posibles defectos, ya que las fotografías se han tomado en un ambiente no controlado. Se han elegido tres etapas de madurez específicas: maduro, pintón e inmaduro, basándose en motivos de exportación y comercialización, ya que estas categorías son las más relevantes para determinar la calidad y el estado de los arándanos en los mercados internacionales. Los arándanos inmaduros son importantes para identificar cuándo no están listos para la cosecha, los pintones son cruciales para prever el momento óptimo de recolección y los maduros son esenciales para asegurar la calidad del producto en su punto óptimo de consumo. Esta clasificación permite a las empresas tomar decisiones precisas sobre el momento de la cosecha y los estándares de exportación. Para construir un conjunto de datos robusto y diverso, se capturaron 666 imágenes de arándanos en distintos estados de madurez utilizando una cámara de alta resolución Canon EOS 5D Mark IV con un lente macro EF 100mm f/2.8L. Las imágenes se tomaron en el campo de cultivo de la empresa Camposol, ubicado en el distrito de Chao, provincia de Virú, en el departamento de La Libertad. Para asegurar la diversidad del conjunto de datos, las fotografías se capturaron bajo diferentes condiciones de iluminación (natural y artificial) y desde varios ángulos (incluyendo vista superior, lateral y en ángulo inclinado). El equipo fotográfico se configuró para diferentes ajustes de apertura y velocidad de obturación, lo que permitió capturar imágenes con variaciones en la profundidad de campo y nitidez. Además, se utilizó un trípode y un control remoto para minimizar cualquier movimiento de la cámara y asegurar la consistencia en la calidad de las imágenes. Un grupo de expertos en horticultura inspeccionó cada arándano y los clasificó manualmente según su madurez. Se utilizó un software de anotación de imágenes, LabelImg, para marcar las regiones de interés en cada fotografía, asegurando que cada arándano estuviera correctamente etiquetado. Este proceso incluyó una doble verificación para minimizar errores humanos. La cantidad de imágenes utilizadas se tomó como referencia de otro estudio similar que empleó 620 imágenes [18]. En los últimos años, las redes neuronales han evolucionado significativamente, 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 precisos y eficientes [19], también se aplicó aumentación de datos en las imágenes. Durante la selección de imágenes, se descartaron aquellas que no cumplieran con los estándares de calidad o que estaban duplicadas. Durante el etiquetado, se identificaron gráficamente los arándanos según su estado de madurez utilizando la herramienta Roboflow (Fig. 1). El color y la textura de los arándanos son cruciales para la clasificación, ya que el software detecta y marca los arándanos en las imágenes, asignando etiquetas como "maduro u óptimo", "inmaduro" y "sobremaduro". Esta clasificación visual permite que el modelo aprenda a distinguir entre los diferentes estados de madurez de manera precisa, tal como lo señalan trabajos recientes [20,21]. Adicionalmente, se aplicaron técnicas de aumentación de datos como rotación, cambio de brillo y contraste para incrementar la variabilidad y robustez del conjunto de datos. Las características recopiladas se basan en criterios visuales y morfológicos como el color, tamaño y textura de los arándanos. Una de las características más evidentes para reconocer un arándano maduro es su color uniforme y su textura suave (C), mientras que los arándanos inmaduros suelen tener un color más claro y una textura más firme (A), y los arándanos pintones pueden presentar manchas verdosas y moradas (B). Esta información sienta las bases fundamentales para el entrenamiento de la red neuronal.

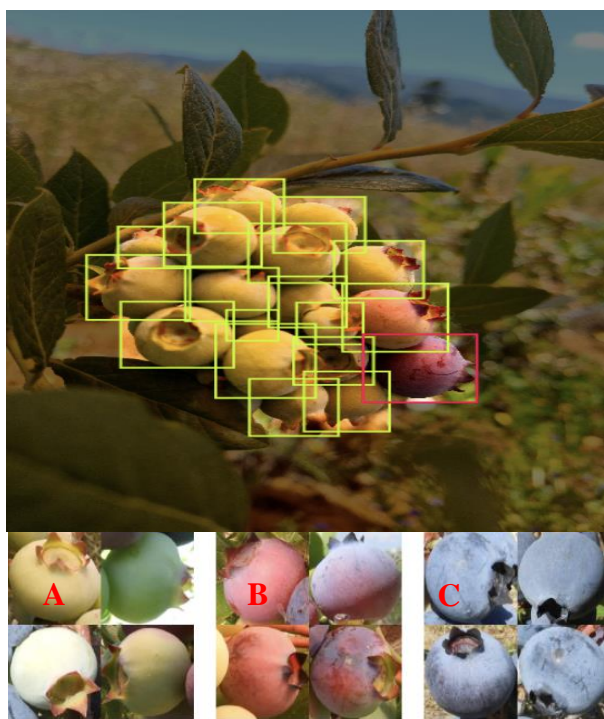


Figura 1. Etiquetado de arándanos mediante el software Roboflow.

YOLO (You Only Look Once) es una familia de modelos de redes neuronales convolucionales (CNN) utilizada para la detección de objetos en imágenes y videos en tiempo real [21]. En este proyecto, se utilizó la versión YOLOv8, que representa la última iteración de esta tecnología, conocida por su alta precisión y eficiencia. YOLOv8 se basa en una arquitectura de red neuronal convolucional que permite la detección de objetos con una sola pasada a través de la red, en lugar de realizar múltiples etapas de procesamiento como en otros enfoques. Esta elección se fundamenta en una investigación reciente donde se utilizó YOLOv8, demostrando su superioridad en términos de precisión y velocidad en comparación con versiones anteriores y

otros modelos de detección de objetos. Dicha arquitectura se caracteriza 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. Las coordenadas predichas para una caja de anclaje (bx, by, bw, bh) se ajustan de la siguiente manera Ec. (1).

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

En la ecuación 1 se calcula la coordenada x del centro de la caja de anclaje ajustada. Se aplica una función de activación sigmoide (σ) a la predicción tx de la red neuronal, y luego se suma la coordenada x de la celda de la cuadrícula (cx). Esto ayuda a localizar el centro horizontal de la caja de anclaje en relación con la celda de la cuadrícula Ec. (2).

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

En la ecuación 2 se calcula la coordenada y del centro de la caja de anclaje ajustada. Se aplica la función de activación sigmoide (σ) a la predicción ty de la red neuronal, y luego se suma la coordenada y de la celda de la cuadrícula (cy). Esto ayuda a localizar el centro vertical de la caja de anclaje en relación con la celda de la cuadrícula Ec. (3).

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

En la ecuación 3 se calcula el ancho de la caja de anclaje ajustada. Se utiliza la función de activación exponencial (e^{tw}) aplicada a la predicción tw de la red neuronal, y luego se multiplica por el ancho predeterminado de la caja de anclaje (pw). Esto ajusta dinámicamente el ancho de la caja de anclaje basado en la predicción de la red Ec. (4).

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

En la ecuación 4 se calcula el alto de la caja de anclaje ajustada. Se utiliza la función de activación exponencial (e^{th}) aplicada a la predicción th de la red neuronal, y luego se multiplica por la altura predeterminada de la caja de anclaje (ph). Esto ajusta dinámicamente la altura de la caja de anclaje basado en la predicción 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. 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 Ec. (5).

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

$Lloc$ 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.

La convolución es una operación matemática que combina dos conjuntos de información. En el contexto de las CNNs, se refiere a la aplicación de un filtro (o núcleo) sobre una imagen para extraer características importantes [22]. Este filtro se desplaza sobre la imagen de entrada y, en cada posición, se calcula un valor de salida que es el resultado de una suma ponderada de los valores de la imagen y los valores del filtro. Durante el entrenamiento de YOLOv8, las capas convolucionales se utilizan para extraer características visuales importantes de las imágenes de arándanos. Estas características pueden incluir bordes, texturas, colores y formas que son indicativos de los diferentes estados de madurez (maduro, inmaduro, pintón). Los filtros convolucionales (Fig. 2) son matrices pequeñas que se aplican sobre las imágenes para procesarlas, entrenados para identificar características específicas. En el caso de los arándanos, los filtros

convolucionales de YOLOv8 son esenciales para la clasificación por madurez. Este proceso comienza con la primera capa de convolución, que detecta características básicas como bordes, texturas y áreas con color uniforme. Las capas intermedias combinan y refinan estas características, enfocándose en patrones más complejos como texturas específicas y variaciones de color. Finalmente, las capas profundas combinan todas las características extraídas para formar una representación global de la imagen, permitiendo distinguir entre las diferentes etapas de madurez de los arándanos. En la aplicación práctica, la capa de salida de YOLOv8 utiliza las características abstractas para identificar y clasificar los arándanos en la imagen, determinando la ubicación y el tipo de cada uno (maduro, verde, pintón). Los filtros en diferentes capas detectan el color azul oscuro y la textura uniforme de los arándanos maduros, los tonos verdes y las texturas más firmes de los arándanos verdes, y las variaciones de color y manchas de los arándanos pintones.

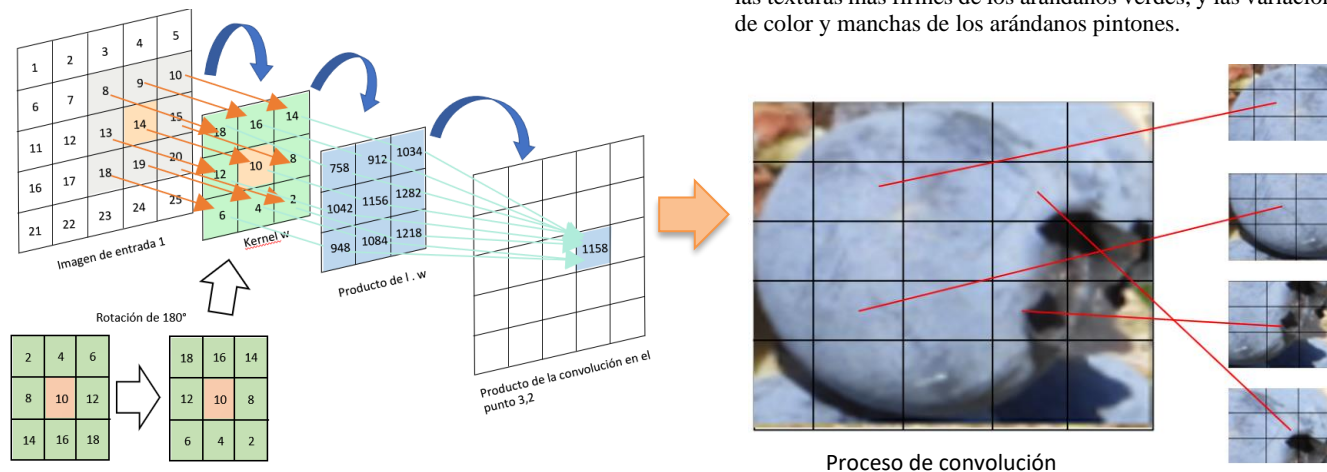


Figura 2. Filtros convolucionales de YOLOv8 extrayendo características críticas de las imágenes de arándanos.

La arquitectura de YOLOv8 (Fig. 3) lleva a cabo un proceso para clasificar la madurez de los arándanos en maduro, verde y pintón con alta precisión y eficiencia. Comienza con el Backbone, una sección inicial de la red que analiza la imagen de los arándanos para extraer características visuales clave, como la textura y color. En este paso, los filtros convolucionales del Backbone identifican patrones específicos relacionados con la madurez de los arándanos, como tonos azules oscuros para los maduros y tonos más verdes para los verdes. Estas características se refinan aún más en el Neck, donde se fusionan y combinan para formar representaciones más complejas y detalladas de la madurez.

El Backbone utiliza combinaciones de capas conocidas como C2f (Conv2d + BatchNorm2d + LeakyReLU) y C2l (otra configuración de Conv2d + BatchNorm2d + LeakyReLU) para extraer características de diferentes niveles de abstracción. Conv2d es una capa de convolución 2D que aplica filtros convolucionales a la imagen para extraer características, mientras que BatchNorm2d normaliza las activaciones de la capa anterior, ayudando a acelerar el entrenamiento y proporcionar estabilidad. LeakyReLU introduce no linealidad en el modelo, permitiendo que las neuronas transmitan información incluso cuando la activación es negativa. En el Neck, se utilizan operaciones de submuestreo (upsample) para aumentar la resolución espacial de las características extraídas, generalmente utilizando interpolación bilineal o vecinos más cercanos. Además, la operación de concatenación (concat) se emplea para combinar

características de diferentes capas o niveles de la red, formando representaciones más ricas y detalladas.

Durante la etapa de Predicción, estas representaciones fusionadas se utilizan para generar resultados específicos sobre la madurez de los arándanos en la imagen. La red examina las características extraídas y busca señales distintivas de madurez. Estos criterios son aprendidos y codificados por los filtros de la red a lo largo del entrenamiento. Como resultado, YOLOv8 puede identificar con precisión los arándanos maduros en la imagen, proporcionando información detallada sobre su ubicación y clasificación. Este enfoque exhaustivo y detallado garantiza una clasificación precisa y confiable de la madurez de los arándanos, lo que puede ser invaluable en aplicaciones agrícolas para optimizar la selección y manejo de la fruta.

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.

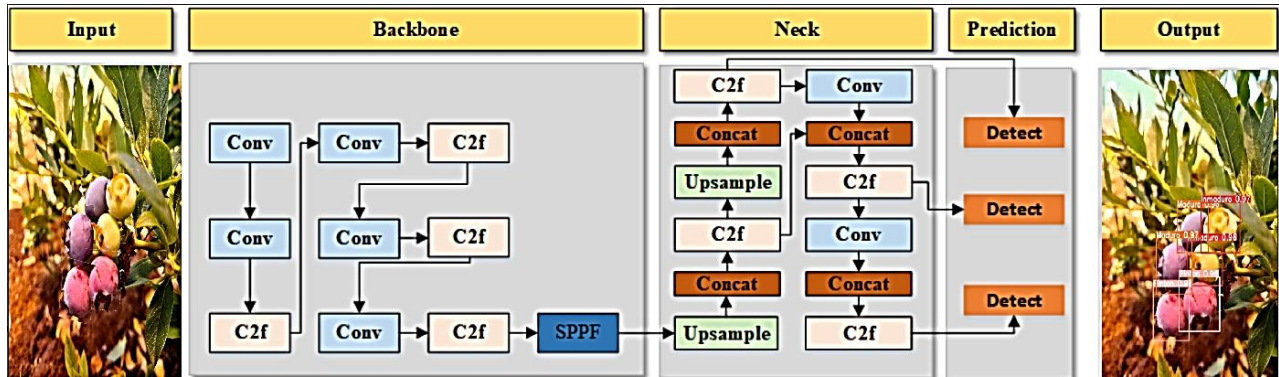


Figura 3. Arquitectura YOLOv8, un sistema diseñado para clasificar la madurez de los arándanos en tres etapas clave: Backbone, Neck y Prediction.

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 que permite realizar operaciones algebraicas de manera eficiente en la GPU. 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 [23-25]. 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 150. 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. Conjuntamente, se muestra el entrenamiento exitoso del modelo YOLOv8, donde las pérdidas de entrenamiento y validación disminuyen consistentemente y las métricas de precisión, recuperación y mAP alcanzan valores altos y estables. Esto indica que el modelo está mejorando continuamente en la detección y clasificación de objetos, generalizando bien a datos no vistos y proporcionando datos precisos y fiables. 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 garantizó la robustez y precisión del modelo en la detección de la madurez de los arándanos.

Para entrenar el sistema de visión artificial con redes neuronales, se sigue un procedimiento detallado que comienza con la recolección de fotografías de arándanos y termina con la red neuronal siendo capaz de identificar el grado de madurez en el que se encuentra. A continuación, se describe paso a paso cómo desarrollar este proyecto. El algoritmo YOLO detecta objetos en imágenes dividiendo la imagen en una cuadrícula de $S \times S$ celdas. Cada celda predice varias cajas delimitadoras y las probabilidades de clases de objetos. La predicción incluye las coordenadas de las cajas, la confianza de que la caja contiene un objeto y las probabilidades de clase. La red neuronal se entrena

minimizando una función de pérdida que penaliza los errores en las coordenadas de las cajas, la confianza y la clasificación de los objetos. El proceso de entrenamiento implica pasar imágenes etiquetadas a través de la red, calcular la pérdida, y ajustar los pesos de la red mediante retro propagación para mejorar la precisión Tabla (1).

Tabla 1. Algoritmo principal del entrenamiento YOLOv8.

Línea de código	Explicación
<code>from yolov8.models.yolo import Model</code>	Importa el modelo YOLOv8.
<code>class Dataset</code>	Define un dataset
<code>loss_fn = ComputeLoss(model)</code>	Define la función de pérdida.
<code>for epoch in range(epochs)</code>	Inicia el bucle de entrenamiento.
<code>model.train()</code>	Pone el modelo en modo de entrenamiento.
<code>imgs = imgs.to('cuda')</code>	Envía las imágenes a la GPU.
<code>labels = labels.to('cuda')</code>	Envía las etiquetas a la GPU.
<code>optimizer.step()</code>	Actualiza los parámetros del modelo.
<code>torch.save(model.state_dict(), 'yolov8_trained_model.pth')</code>	Guarda los pesos del modelo entrenado.

Después de realizar las predicciones, se aplica la supresión de no máximos (NMS) para eliminar las cajas delimitadoras redundantes y conservar solo aquellas con la mayor confianza. Este enfoque permite a YOLO realizar detecciones rápidas y precisas en tiempo real. Durante el entrenamiento, el modelo se ajusta iterativamente usando conjuntos de datos de entrenamiento y validación para optimizar los pesos y evitar el sobreajuste, lo que resulta en un modelo robusto capaz de detectar objetos con alta precisión.

Para implementar el algoritmo de detección de la madurez de los arándanos, se utilizó Python en el entorno de desarrollo PyCharm. Se configuró PyCharm con un entorno virtual y se instalaron las dependencias necesarias, incluyendo 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, incluyendo la captura de imágenes desde una cámara o la carga desde el sistema de archivos. Utilizando el modelo YOLOv8, se detectaron y clasificaron los arándanos en

las imágenes según su madurez, mostrando los resultados mediante cuadros delimitadores y etiquetas sobre las imágenes originales Tabla (2).

Tabla 2. Descripción detallada de los pasos del algoritmo principal de YOLOv8.

Paso	Descripción
Detección y Seguimiento de Arándanos	El modelo YOLOv8 se utiliza para detectar y seguir arándanos en cada fotograma. Se adapta para identificar arándanos y clasificarlos según su madurez. La función model.track() realiza esta tarea.
Visualización de los Resultados	Los resultados de la detección y seguimiento se muestran visualmente en el fotograma original, con cuadros delimitadores alrededor de los arándanos detectados y etiquetados que indican su madurez. La visualización se realiza mediante la función plot().
Visualización en Tiempo Real	La función cv2.imshow() muestra el fotograma procesado en una ventana titulada "fame", actualizándose en tiempo real con cada nuevo fotograma para visualizar la detección, seguimiento y clasificación de arándanos por madurez. El bucle while sigue ejecutándose hasta que se presiona la tecla "q", lo que rompe el bucle y cierra la ventana.

3. RESULTADOS

Se desarrolló un sistema de visión artificial para clasificar los arándanos según su grado de madurez utilizando la red neuronal convolucional YOLOv8. Para evaluar la eficiencia del sistema, se realizaron una serie de 150 tomas de arándanos en diferentes estados de madurez. Las pruebas se llevaron a cabo con los arándanos en diversas posiciones y cantidades, permitiendo determinar la precisión y eficacia de la red neuronal en la clasificación de sus tres estados de madurez (maduro, pintón e inmaduro). Para evaluar la efectividad del sistema de visión artificial para clasificar arándanos según su grado de madurez, se realizaron pruebas con arándanos de diferentes lotes.

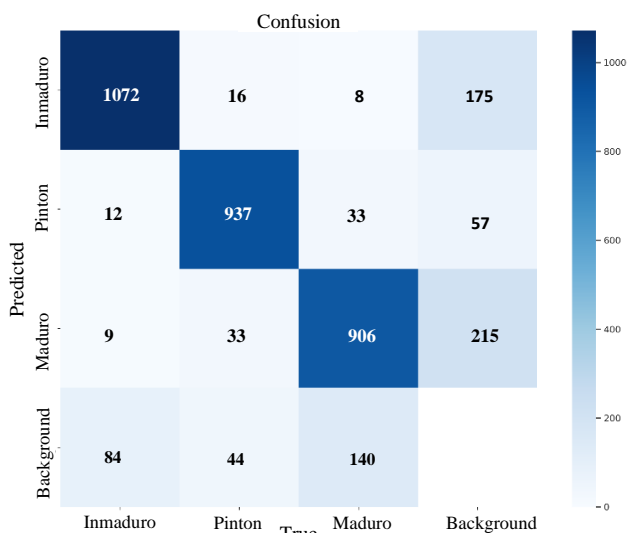


Figura 4. Matriz de confusión con los resultados del modelo de clasificación de arándanos.

Para la categoría inmaduro, hubo 1072 verdaderos positivos

(Inmaduro predicho como Inmaduro). Los falsos negativos incluyeron 16 predichos como Pintón, 8 como Maduro y 175 como Background. En la categoría pintón, se registraron 937 verdaderos positivos (Pintón predicho como Pintón). Los falsos negativos consistieron en 12 predichos como Inmaduro, 33 como Maduro y 57 como Background. Para la categoría maduro, hubo 906 verdaderos positivos (Maduro predicho como Maduro). Los falsos negativos incluyeron 9 predichos como Inmaduro, 33 como Pintón y 215 como Background. En la categoría background, se obtuvieron 140 verdaderos positivos (Background predicho como Background). Los falsos negativos fueron 84 predichos como Inmaduro, 44 como Pintón y 140 como Maduro (Fig. 4).

Tabla 3. Análisis de precisión y F1-score por lote de arándanos clasificados por madurez.

Lotes	Análisis de Precisión y F1- score
Lote1	Maduro: Se clasificaron 50 arándanos maduros con una precisión del 95% y un F1-score de 0.96. Inmaduro: Se evaluaron 40 arándanos inmaduros, alcanzando una precisión del 92% y un F1-score de 0.93. Pintón: Se clasificaron 35 arándanos pintones, logrando una precisión del 89% y un F1-score de 0.91.
Lote 2	Maduro: Se evaluaron 45 arándanos maduros con una precisión del 96% y un F1-score de 0.97. Inmaduro: Se clasificaron 42 arándanos inmaduros con una precisión del 93% y un F1-score de 0.94. Pintón: Se analizaron 38 arándanos pintones, obteniendo una precisión del 90% y un F1-score de 0.92.
Lote 3	Maduro: Se clasificaron 48 arándanos maduros con una precisión del 94% y un F1-score de 0.95. Inmaduro: Se evaluaron 39 arándanos inmaduros con una precisión del 91% y un F1-score de 0.92. Pintón: Se clasificaron 38 arándanos pintones, logrando una precisión del 88% y un F1-score de 0.90.
Lote 4	Maduro: Se evaluaron 49 arándanos maduros con una precisión del 97% y un F1-score de 0.98. Inmaduro: Se clasificaron 40 arándanos inmaduros con una precisión del 94% y un F1-score de 0.95. Pintón: Se analizaron 36 arándanos pintones, obteniendo una precisión del 91% y un F1-score de 0.93.

Los resultados indican que el sistema de visión artificial basado en YOLOv8 es altamente efectivo para clasificar arándanos por su grado de madurez. Las precisiones y F1-scores obtenidos fueron consistentemente altos a lo largo de los diferentes lotes. La precisión varió entre el 88% y el 97%, demostrando una alta capacidad del modelo para realizar clasificaciones correctas. Los F1-scores, que balancean la precisión y el recall, estuvieron en un rango de 0.90 a 0.98, reflejando un buen rendimiento general del sistema. Estos resultados confirman la viabilidad y efectividad del sistema propuesto para la clasificación automática de arándanos en aplicaciones agrícolas Tabla (3).

Se eligió la función de activación LeakyReLU para YOLOv8 debido a su capacidad para mitigar el problema de desvanecimiento del gradiente, mejorar la convergencia del entrenamiento al mantener activaciones no saturadas, y reducir el riesgo de "neurons dying" al permitir un pequeño gradiente negativo. Durante el desarrollo del programa, se han realizado ajustes significativos, incluyendo la optimización de hiperparámetros como el tamaño del lote y la tasa de aprendizaje, la implementación de técnicas de aumento de datos para mejorar la robustez del modelo, ajustes en la arquitectura para mejorar la

extracción y fusión de características, y el uso de técnicas de regularización para prevenir el sobreajuste, asegurando así una clasificación precisa y eficiente de la madurez de los arándanos.

4. DISCUSIÓN

Los resultados obtenidos de las pruebas con el sistema de visión artificial basado en la red neuronal convolucional YOLOv8 para clasificar arándanos según su grado de madurez muestran un rendimiento altamente satisfactorio. En comparación con estudios previos que han utilizado tecnologías similares para la clasificación de arándanos nuestro sistema logró alcanzar precisiones similares en la clasificación de arándanos maduros, inmaduros y pintones. Específicamente, en una investigación basada en redes neuronales convolucionales [13], donde el objetivo fue el reconocimiento de arándanos en diferentes niveles de madurez, se reportó una precisión de reconocimiento del 94.62%. Por otro lado, en un estudio previo [16], cuyo objetivo fue desarrollar un método de detección ligero basado en un algoritmo mejorado de YOLOv5 para lograr un reconocimiento preciso y rápido de las etapas de madurez de los arándanos, solo se alcanzó una precisión del 92%. En contraste, nuestro sistema YOLOv8 demostró una precisión superior del 95%, destacándose significativamente frente a los resultados anteriores y subrayando la eficacia y mejora que nuestro método aporta al reconocimiento de las etapas de madurez de los arándanos. Estos resultados sugieren que YOLOv8 es competitivo y efectivo en la clasificación de arándanos por su grado de madurez, proporcionando una base sólida para aplicaciones agrícolas donde la precisión y la eficiencia en la clasificación automática son fundamentales. La ligera variabilidad observada entre lotes podría atribuirse a diferencias inherentes en las características visuales de los arándanos entre los lotes. Esta variabilidad podría ser un área para futuras investigaciones, enfocándose en mejorar la capacidad del modelo para poder adaptarse a diferentes condiciones de producción de arándanos.

Con respecto a los cuatro lotes evaluados, mostraron variaciones significativas en sus resultados de clasificación de arándanos por grado de madurez. El Lote 1 destacó por su alta precisión en arándanos maduros e inmaduros, con un F1-score de hasta 0.96, aunque la categoría pintón mostró la precisión más baja (89%). En contraste, el Lote 2 exhibió una mejora general, alcanzando una precisión del 90% para arándanos pintones. El Lote 3, mantuvo una precisión del 94% para arándanos maduros y un F1-score de 0.95. Finalmente, el Lote 4 destacó con los mejores resultados generales, logrando una precisión máxima del 97% para arándanos maduros y un destacado F1-score de 0.98, indicando condiciones particularmente favorables para la clasificación en este lote.

La precisión y el F1-score son métricas cruciales para evaluar el rendimiento de un sistema de clasificación. La alta precisión indica que el sistema tiene una baja tasa de falsos positivos, mientras que los altos F1-scores indican un buen equilibrio entre precisión y recall. Esto es fundamental en aplicaciones agrícolas donde la precisión en la clasificación puede impactar directamente en la calidad y el valor comercial del producto.

La automatización de la clasificación puede acelerar el proceso de selección de arándanos, reduciendo la dependencia en el trabajo manual y mejorando la eficiencia operativa. Dado el éxito con los arándanos, el mismo enfoque podría adaptarse para clasificar otras frutas y vegetales, ampliando el alcance de las aplicaciones de visión artificial en la agricultura. A pesar de los resultados prometedores, hay algunas limitaciones que deben ser

abordadas en futuras investigaciones como, evaluar el sistema bajo una mayor variedad de condiciones ambientales y de iluminación para asegurarse de que el rendimiento se mantenga consistente. Entrenar el modelo con un conjunto de datos más amplio y diverso podría mejorar aún más su capacidad de difusión y precisión. La implementación y evaluación del sistema en un entorno de producción en tiempo real sería un siguiente paso para validar su viabilidad práctica y eficiencia operativa.

5. CONCLUSIÓN

El desarrollo e implementación de un sistema de visión artificial utilizando la red neuronal convolucional YOLOv8 para clasificar arándanos por su grado de madurez ha demostrado ser una solución efectiva y precisa. El sistema logró una precisión superior del 95% y un F1-score entre 0.90 y 0.98, indicando una alta fiabilidad en la clasificación de arándanos maduros, inmaduros y pintones. Los resultados consistentes a lo largo de diferentes lotes sugieren que el modelo YOLOv8 es capaz de distinguir, adaptándose a variaciones naturales en los arándanos. La implementación de este sistema puede optimizar significativamente el proceso de selección de arándanos, reduciendo la necesidad de clasificación manual y mejorando la eficiencia operativa. La clasificación precisa por grado de madurez puede mejorar la calidad del producto final ofrecido a los consumidores, asegurando que los arándanos sean comercializados en su estado óptimo. Además, el éxito demostrado con los arándanos sugiere que el enfoque basado en YOLO v8 puede ser adaptado para la clasificación de otras frutas y vegetales, ampliando el alcance de la aplicación de la visión artificial en la agricultura. Basándonos en los resultados del estudio, para mejorar aún más el sistema de visión artificial basado en YOLOv8 para la clasificación de arándanos por su grado de madurez, se sugieren varias líneas de investigación futuras. Primero, es crucial evaluar el rendimiento del sistema bajo una variedad más amplia de condiciones ambientales y de iluminación para garantizar su consistencia y fiabilidad en diferentes entornos agrícolas. Además, ampliar el conjunto de datos de entrenamiento para incluir una mayor diversidad de imágenes podría mejorar la precisión y la capacidad de generalización del modelo, asegurando una clasificación más precisa y robusta en diversas situaciones.

REFERENCIAS BIBLIOGRÁFICAS

- [1] Silva, Sara, et al. Health promoting properties of blueberries: A review. *Critical reviews in food science and nutrition*, 2020; 60(2); 181-200. <https://doi.org/10.1080/10408398.2018.1518895>.
- [2] Rashidinejad, A. Blueberries. In *Nutritional composition and antioxidant properties of fruits and vegetables*. Academic Press, 2020. p. 467-482. <https://doi.org/10.1016/B978-0-12-812780-3.00029-5>
- [3] Gonzalez, Sebastian; Arellano, Claudia; Tapia, Juan E. Deepblueberry: Quantification of blueberries in the wild using instance segmentation. *Ieee Access*, 2019, vol. 7, p. 105776-105788. <https://doi.org/10.1016/j.biosystemseng.2022.10.005>
- [4] Vargas, Q y Best, S. Introducción a la Agricultura de Precisión 4.0 en huertos de arándanos. 2021. Disponible en: <https://bibliotecadigital.ciren.cl/items/8d7bd362-1679-4549-91eb-494d59ea390b>
- [5] Wang D, Zhang M, Mujumdar AS y Yu D. Advanced Detection Techniques Using Artificial Intelligence in Processing of Berries. *Food Eng Rev*. 2022; 14: 176-199.

- Available in: <https://doi.org/10.1007/s12393-021-09298-5>.
- [6] Rodríguez Sánchez, E. G., & Zárata Cáceres, M. E. Diseño e implementación de un sistema de clasificación por visión artificial basado en redes neuronales y servo posicionamiento lineal industrial. [Tesis de pregrado]. Arequipa: Universidad Católica Santa María; 2019. Recuperado de: <https://repositorio.ucsm.edu.pe/items/7ecb270a-0f1e-4830-89bc-211769993399>
- [7] Kim, S. et al. Recent advances of artificial intelligence in manufacturing industrial sectors: A review. *International Journal of Precision Engineering and Manufacturing*, 2022, p. 1-19. <https://doi.org/10.1007/s12541-021-00600-3>
- [8] Seema, A & Gill G, "Automatic Fruit Grading and Classification System Using Computer Vision: A Review," *Second International Conference on Advances in Computing and Communication Engineering*. 2015: 598-603, doi: 10.1109/ICACCE.2015.15.
- [9] Huang W, Wang X, Zhang J, Xia J, Zhang X. Improvement of blueberry freshness prediction based on machine learning and multi-source sensing in the cold chain logistics. *Food Control* [Internet]. 2023;145(109496):109496. <http://dx.doi.org/10.1016/j.foodcont.2022.109496>.
- [10] Liu Y, Zheng H, Zhang Y, Zhang Q, Chen H, Xu X, et al. "Is this blueberry ripe?": a blueberry ripeness detection algorithm for use on picking robots. *Front Plant Sci* [Internet]. 2023;14. Available in: <http://dx.doi.org/10.3389/fpls.2023.1198650>
- [11] Tan K, Lee WS, Gan H, Wang S. Recognising blueberry fruit of different maturity using histogram oriented gradients and colour features in outdoor scenes. *Biosyst Eng* [Internet]. 2018;176:59–72. Disponible en: <http://dx.doi.org/10.1016/j.biosystemseng.2018.08.011>
- [12] Li Z, Chang Hong M, and KangXuan J. "Prediction of blueberry maturity based on deep learning." 2023; 51(5): 232-236. doi:10.3969/j.issn.0517-6611.2023.05.053
- [13] Zhu, X., Ma, H., Ji, J. T., Jin, X., Zhao, K. X., Zhang, K. Detecting and identifying blueberry canopy fruits based on faster r-CNN. *J. South. Agric.* 2020; 51 (6): 1493–1501. doi: 10.3969/j.issn.2095-1191.2020.06.032
- [14] MacEachern CB, Esau TJ, Schumann AW, Hennessy PJ, Zaman QU. Detection of fruit maturity stage and yield estimation in wild blueberry using deep learning convolutional neural networks. *Smart Agricultural Technology* [Internet]. 2023;3(100099):100099. Available in: <http://dx.doi.org/10.1016/j.atech.2022.100099>
- [15] Yang W, Ma X, Hu W, Tang P. Lightweight blueberry fruit recognition based on multi-scale and attention fusion NCBAM. *Agronomy*. 2022; 12 (10): 2354. doi: 10.3390/agronomy12102354
- [16] Xiao, F.; Wang, H.; Xu, Y.; Shi, Z. Un método de detección ligero para la madurez de la fruta de arándanos basado en un algoritmo YOLOv5 mejorado. *Agricultura* 2024, 14, 36. <https://doi.org/10.3390/agriculture14010036>
- [17] Rizzo M, Marcuzzo M, Zangari A, Gasparetto A, Albarellia A. Fruit ripeness classification: A survey. *Artificial Intelligence in Agriculture*, 2023; 7: 44-57. <https://doi.org/10.1016/j.aiaa.2023.02.004>
- [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] Das, Anup Kumar, et al. Machine vision system for real-time debris detection on mechanical wild blueberry harvesters. *Smart Agricultural Technology*, 2023, vol. 4, p. 100166. <https://doi.org/10.1016/j.atech.2022.100166>
- [20] S. Espinoza, C. Aguilera, L. Rojas y PG Campos, "Análisis de imágenes de frutas con aprendizaje profundo: una revisión sistemática de la literatura y direcciones futuras", en *IEEE Access*, vol. 12, págs. 3837-3859, 2024, doi: 10.1109/ACCESS.2023.3345789
- [21] 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
- [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] Bongulwar, DM; Talbar, SN. BONGULWAR, D. M.; TALBAR, S. N. Robust Convolutional Neural Network Model For Recognition of Fruits. *Indian Journal of Science and Technology*, 2021, vol. 14, no 45, p. 3318-3334. <https://doi.org/10.17485/IJST/v14i45.1493>
- [24] Lemsalu, Madis, et al. Real-Time CNN-based Computer Vision System for Open-Field Strawberry Harvesting Robot. *IFAC-PapersOnLine*, 2022, vol. 55, no 32, p. 24-29. <https://doi.org/10.1016/j.ifacol.2022.11.109>
- [25] Hussain, Nazar, et al. Application of deep learning to detect Lamb's quarters (*Chenopodium album* L.) in potato fields of Atlantic Canada. *Computers and Electronics in Agriculture*, 2021, vol. 182, p. 106040. <https://doi.org/10.1155/2022/4194874>
- [26] Wang, Chenglin, et al. YOLO-BLBE: A Novel Model for Identifying Blueberry Fruits with Different Maturities Using the I-MSRCR Method. *Agronomy*, 2024, vol. 14, no 4, p. 658. <https://doi.org/10.3390/agronomy14040658>