

ARTÍCULO DE INVESTIGACIÓN

Detección de sarna y oídio en manzanos con YOLOv12

Automated Detection of Scab and Powdery Mildew in Apple Crops using YOLOv12

DOI: <https://doi.org/10.61154/metanoia.v12i1.4199>

Manolo Sebastián Muñoz-Espinoza¹

Ángel Santiago Cruz Velasco²

Walter Eduardo Moreno Castillo³

¹**E-mail:** mmunoz@uta.edu.ec **Afiliación:** Universidad Técnica de Ambato. Tungurahua, Ambato, Ecuador. **ORCID:** <https://orcid.org/0000-0001-5897-1211>

² **E-mail:** as.cruz@uta.edu.ec **Afiliación:** Universidad Técnica de Ambato. Tungurahua, Ambato, Ecuador. **ORCID:** <https://orcid.org/0009-0002-0474-216X>

³ **E-mail:** we.moreno@uta.edu.ec **Afiliación:** Universidad Técnica de Ambato. Tungurahua, Ambato, Ecuador. **ORCID:** <https://orcid.org/0000-0003-3907-371X>

Recibido: 20/08/2025

Revisado: 30/08/2025

Aprobado: 20/09/2025

Publicado: 01/01/2026

RESUMEN

Este estudio presenta un sistema de detección automática de sarna (*Venturia inaequalis*), oídio (*Podosphaera leucotricha*) y hojas sanas en manzanos (*Malus domestica*) mediante el uso de un modelo de detección de objetos basado en redes neuronales convolucionales, específicamente una versión optimizada de YOLOv12. La investigación se enmarca en un diseño metodológico experimental y cuantitativo, con un enfoque aplicado y nivel descriptivo-predicativo, cuyo objetivo es evaluar la eficacia de la inteligencia artificial en el diagnóstico temprano de enfermedades foliares en frutales. Se recolectaron 300 imágenes de hojas de manzano en condiciones de campo, 100 para cada clase, las cuales fueron preprocesadas, aumentadas y etiquetadas mediante la plataforma Roboflow, obteniendo un conjunto de datos final de 720 imágenes. El modelo fue entrenado y validado utilizando un conjunto de datos dividido en entrenamiento (92%), validación (4%) y prueba (4%). Los resultados mostraron una precisión del 99.2 %, un recall del 89.5 % y un mAP50 del 92.5 %, tras 100 épocas de entrenamiento. La matriz de confusión reveló un bajo índice de falsos negativos, lo que indica una alta sensibilidad del modelo. Estos hallazgos demuestran que YOLOv12 puede ser una herramienta efectiva para la detección temprana de enfermedades en sistemas de agricultura de precisión, reduciendo la dependencia del diagnóstico visual por expertos. Este trabajo contribuye al desarrollo de tecnologías inteligentes para la sanidad vegetal en entornos agrícolas reales.

Descriptores: Inteligencia artificial, enfermedad de las plantas, agricultura.

ABSTRACT

This study introduces an automated detection system for scab (*Venturia inaequalis*), powdery mildew (*Podosphaera leucotricha*), and healthy leaves in apple trees (*Malus domestica*) utilizing an object detection model based on convolutional neural networks, specifically an optimized version of YOLOv12. The research adopts an experimental and quantitative methodological design, characterized by an applied approach and a descriptive-predictive level, aimed at assessing the effectiveness of artificial intelligence for the early diagnosis of foliar diseases in fruit trees. A total of 300 images of apple leaves were gathered in field conditions, with 100 images per class. These images were preprocessed, augmented, and labeled using the Roboflow platform, leading to a final dataset comprising 720 images. The model was trained

and validated on a dataset split into training (92%), validation (4%), and testing (4%) subsets. The results demonstrated a precision of 99.2%, a recall of 89.5%, and an mAP50 of 92.5% after 100 training epochs. The confusion matrix indicated a low false negative rate, showcasing the model's high sensitivity. These findings suggest that YOLOv12 is a potent tool for the early detection of diseases in precision agriculture, thereby decreasing the reliance on expert visual diagnosis. This research contributes to the advancement of smart technologies for enhancing plant health in practical agricultural contexts.

Descriptors: Artificial intelligence, plant disease, agriculture.

INTRODUCCIÓN

El uso de inteligencia artificial en la agricultura ha cobrado relevancia en los últimos años, especialmente en la detección temprana de enfermedades en cultivos. Entre las tecnologías empleadas destacan Big Data, Machine Learning y Deep Learning (Rodríguez Torres et al., 2023), cuyo impacto ha sido notable en el desarrollo de sistemas de diagnóstico automático basados en imágenes.

En el ámbito de los granos almacenados, (Analuisa-Aroca et al., 2023) utilizaron una red neuronal convolucional tipo ResNet-50 para identificar la presencia del gorgojo del maíz (*Sitophilus zeamais*) en imágenes de granos obtenidos de mercados locales en Ecuador, con un conjunto de 1380 imágenes etiquetadas, el modelo alcanzó niveles altos de precisión (>0.94) en métricas como precisión, F-score, AUC e índice de calidad.

Por otro lado, (Facuy-Delgado et al., 2022) exploraron el uso de algoritmos basados en inteligencia artificial para detectar enfermedades fúngicas en arroz (*Oryza sativa*), específicamente *pyriculariosis* y *helminthosporiosis*, utilizando técnicas como redes neuronales convolucionales (CNN), máquinas de soporte vectorial (SVM) y lógica difusa, lograron identificaciones con precisiones superiores al 95% a partir de imágenes de hojas afectadas.

En el caso del frijol, (Saldaña Valenzuela, 2023) desarrolló un sistema basado en Inteligencia Artificial para la detección de enfermedades como la Mancha angular y la Roya del frijol, aplicando técnicas de Deep Learning y Computación Visual, los resultados muestran que los modelos predictivos pueden ofrecer alternativas efectivas a los métodos tradicionales de diagnóstico, optimizando el uso de insumos y reduciendo pérdidas en la producción.

Varios estudios han demostrado la eficacia de las redes neuronales convolucionales (CNN) en la identificación precisa de enfermedades foliares, (Mora et al., 2021) destacan que

arquitecturas como VGG16 y Xception son herramientas poderosas en este ámbito, especialmente cuando su implementación se realiza en dispositivos económicos y accesibles como Raspberry Pi.

Un ejemplo particularmente destacable es el estudio de (Huanca Namuche et al., 2025), quienes desarrollaron un modelo CNN para la clasificación de la mancha negra en cítricos. Este modelo alcanzó una precisión del 100%, respaldada por un área bajo la curva ROC (AUC) igual a 1.0, (Li et al., 2024) presentaron un método innovador para la detección de enfermedades en hojas de manzano, integrando tres componentes clave: DSCon, BiFormer e IF-CIoU. Este modelo alcanzó un mAP50 (promedio de precisión a 50% de solapamiento) del 93.81% y un mAP50-95 del 95.69%, lo cual representa una mejora significativa respecto a versiones anteriores de YOLO y otros modelos existentes.

Además, (Wang et al., 2022) propusieron MGA-YOLO, un modelo optimizado basado en YOLOv5s, que obtiene un mAP máximo del 89.3% en el conjunto de prueba ALDOD (Apple Leaf Disease Object Detection), superando a modelos avanzados como YOLOX-L. Este modelo también destaca por su eficiencia computacional, alcanzando velocidades de hasta 84.13 FPS (cuadros por segundo), lo que lo hace adecuado para aplicaciones en tiempo real.

En este contexto, el presente artículo propone el uso del modelo YOLOv12 para la detección precisa y rápida de enfermedades mediante visión por computador, con el objetivo de proporcionar una herramienta eficiente y escalable para su aplicación en condiciones reales de campo.

METODOLOGÍA

Diseño metodológico

Esta investigación se desarrolló bajo un diseño experimental con enfoque cuantitativo y aplicado, con el fin de desarrollar y evaluar un modelo de detección de objetos para identificar sarna y oídio en imágenes de hojas de manzano. El estudio se clasifica como de nivel descriptivo-predicativo, ya que no solo describe las características visuales de las enfermedades, sino que también predice su presencia y ubicación dentro de las imágenes. El procedimiento siguió un flujo estructurado compuesto por adquisición de datos, preprocesamiento, entrenamiento del modelo y evaluación cuantitativa de su desempeño.

Adquisición de imágenes

Las imágenes utilizadas en este estudio fueron obtenidas de la plataforma Kaggle (*Plant*

Pathology 2021 - FGVC8, 2021), las cuales consisten en fotografías capturadas directamente en condiciones de campo. El conjunto de datos seleccionado incluye un total de 300 imágenes, distribuidas equitativamente en tres clases representativas de estado fitosanitario, sarna: 100 imágenes, oídio: 100 imágenes y hojas sanas: 100 imágenes

Este dataset fue utilizado como base para el entrenamiento del modelo de detección.

Preparación y Preprocesamiento de las Imágenes

Con el objetivo de mejorar la robustez del modelo frente a variaciones en iluminación, ángulos de captura y otras condiciones ambientales, se aplicaron técnicas de data augmentation mediante la plataforma Roboflow (*Roboflow: Computer Vision Tools for Developers and Enterprises*, 2025).

Las transformaciones empleadas incluyeron: ajuste de resolución de 640 x 640 píxeles, flip: horizontal y vertical, rotación: 90° en sentido horario, antihorario y hacia abajo, desenfoque: hasta 2.1 píxeles, ruido: inserción en hasta el 3.78% de los píxeles, rotación de bounding boxes: entre -15° y +15°, exposición: Variación entre -10% y +10%.

Como resultado del proceso de aumento de datos, el tamaño del conjunto de entrenamiento aumentó a 780 imágenes.

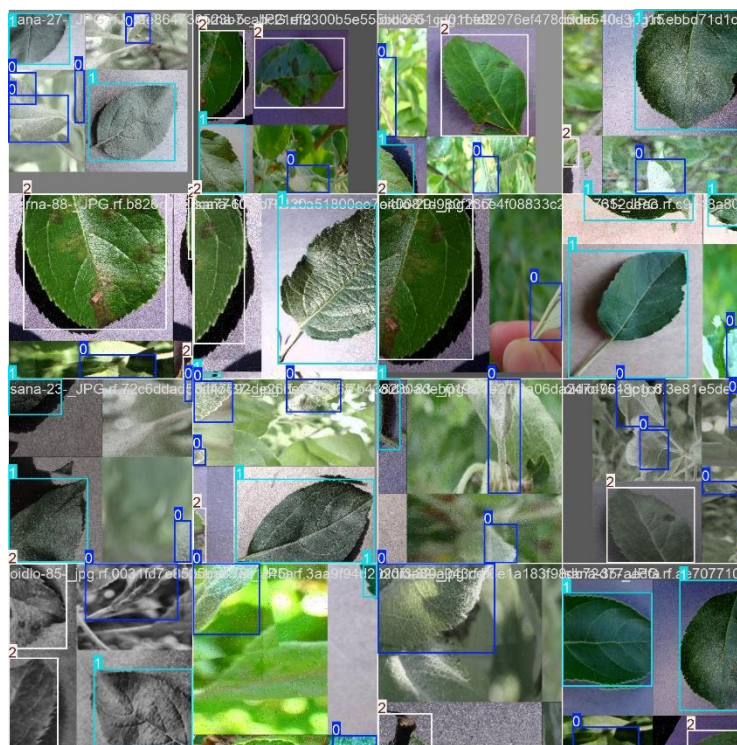
Finalmente, el dataset completo fue dividido en tres subconjuntos: Entrenamiento: 780 imágenes, Validación: 30 imágenes y de Prueba: 30 imágenes

Etiquetado de las Imágenes

El etiquetado de las imágenes se realizó mediante la herramienta Roboflow (*Roboflow: Computer Vision Tools for Developers and Enterprises*, 2025), definiendo bounding boxes precisos alrededor de las áreas afectadas por cada enfermedad. Además, las hojas sanas fueron identificadas como una clase independiente. El etiquetado manual se realizó siguiendo el formato estándar de YOLO, con archivos de texto asociados a cada imagen que contienen las coordenadas normalizadas de los recuadros y sus respectivas clases:

- Clase 0: "oidio"
- Clase 1: "sana"
- Clase 2: "sarna"

Figura 1. Etiquetado de hoja de manzana.



Fuente: Los autores

Arquitectura del Modelo

Para el desarrollo del sistema de detección automática de enfermedades en plantas, se utilizó la arquitectura YOLOv12 (Tian et al., 2025), última versión de la familia YOLO (Redmon et al., 2016), implementada dentro del entorno Ultralytics YOLO. Este modelo fue elegido por su capacidad de realizar detección y clasificación simultánea, además de su eficiencia computacional y alto rendimiento en dispositivos con recursos limitados.

El entrenamiento se llevó a cabo en la plataforma Google Colab, utilizando una GPU para acelerar el procesamiento. La configuración del modelo permitió el aprendizaje basado en detalles visuales complejos y su adaptación a diferentes condiciones de campo.

Entrenamiento del Modelo

Se utilizaron los siguientes parámetros durante el entrenamiento del modelo: Epochs: 100, Batch size: 16, Learning rate: Inicializado en 0.01, ajustado dinámicamente durante el entrenamiento, Optimizador: AdamW, Función de pérdida: Combinación de pérdidas de caja (box loss), clasificación (cls loss) y distribución focal (dfl loss)

El entorno de desarrollo estuvo basado en Python, con el uso de bibliotecas como PyTorch y Ultralytics YOLO.

Evaluación del Modelo

Durante el entrenamiento, se monitorearon métricas clave como:

- Precisión (Precision): Proporción de predicciones correctas sobre todas las predicciones realizadas.
- Recall (Recall): Proporción de predicciones correctas sobre todas las instancias reales.
- mAP (Mean Average Precision): Media ponderada de la precisión a diferentes umbrales de confianza.

El rendimiento del modelo fue evaluado utilizando el conjunto de prueba (30 imágenes).

Además, se generaron gráficos de pérdida (loss) tanto para el conjunto de entrenamiento como de validación, así como curvas de precisión y recall para analizar el desempeño del modelo a lo largo de los epochs.

RESULTADOS

El modelo YOLOv12 alcanzó los siguientes resultados después de 100 epochs de entrenamiento:

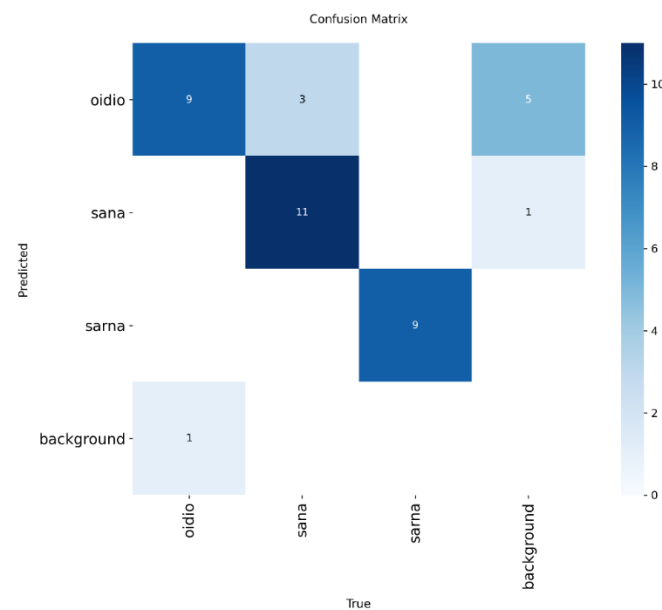
- Precisión (Precision): Máximo de 0.9923.
- Recall (Recall): Máximo de 0.89524.
- mAP50 (Mean Average Precision at IoU=0.5): Máximo de 0.92512.

Estos resultados indican que el modelo tiene una alta capacidad para detectar correctamente las enfermedades de sarna y oídio, así como identificar correctamente las hojas sanas.

Análisis de la matriz de confusión

La matriz de confusión muestra el desempeño del modelo en términos de clasificación:

Figura 2. Matriz de confusión



Fuente: Los autores

Oídio: El modelo clasificó correctamente 9 imágenes de oídio, pero hubo 3 falsos positivos (clasificados como sana), 5 falsos positivos (clasificados como sarna) y un falso negativo (clasificado como background)

Sana: El modelo clasificó correctamente 11 imágenes de hojas sanas, pero hubo 3 falsos positivos (clasificados como oídio).

Sarna: El modelo clasificó correctamente 9 imágenes de sarna, sin falsos positivos ni falsos negativos significativos.

Background: Hubo un caso donde el modelo clasificó incorrectamente una imagen como background.

Estos resultados sugieren que el modelo tiene una buena capacidad para distinguir entre las clases, aunque existen algunas confusiones entre oídio y hojas sanas.

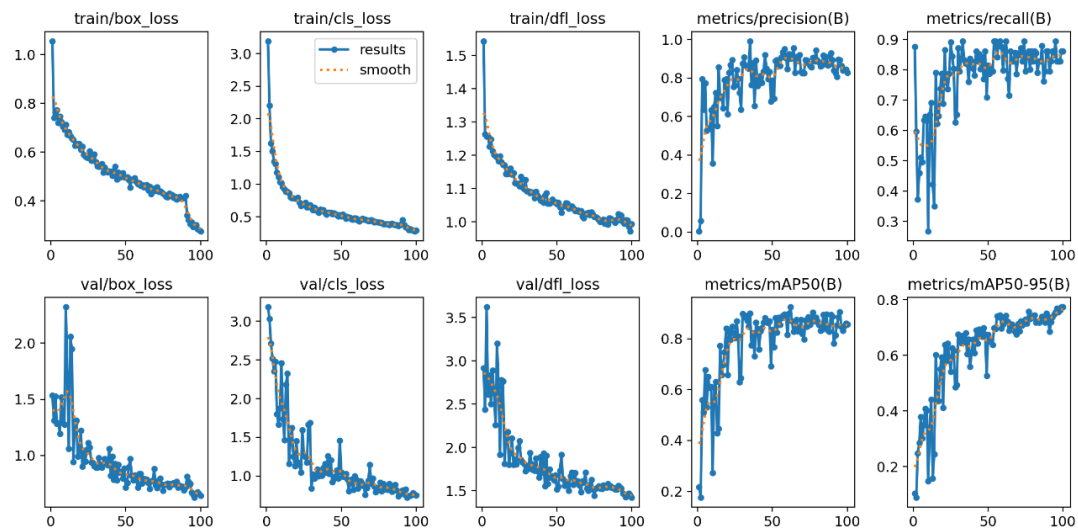
Gráficos de pérdida y métricas

Los gráficos de pérdida muestran una tendencia descendente en las pérdidas de caja (box loss), clasificación (cls loss) y distribución focal (dfl loss) tanto en el conjunto de entrenamiento como de validación. Esto indica que el modelo converge adecuadamente y mejora su rendimiento a medida que avanza el entrenamiento.

Las curvas de precisión y recall también muestran un comportamiento estable y alto desempeño, especialmente en el rango final de epochs. El mAP50 alcanzó valores cercanos a 0.92, lo cual

es muy satisfactorio para este tipo de tareas de detección.

Figura 3. Gráficos de pérdida y métricas

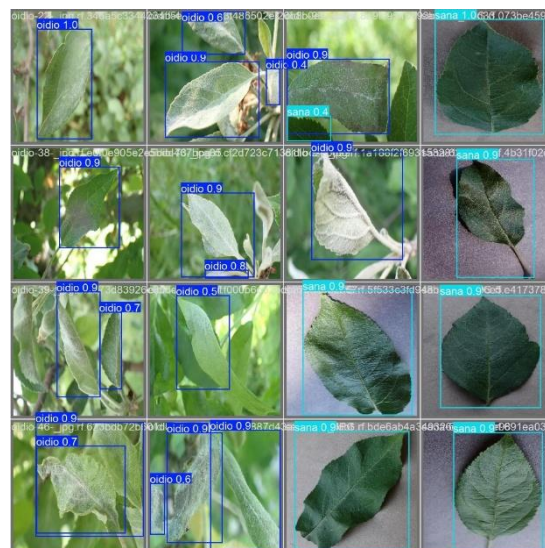


Fuente: Los autores

Imágenes de los resultados

En la Figura 4 se muestran algunas imágenes representativas del conjunto de validación, donde se puede observar el desempeño del modelo al momento de detectar y segmentar visualmente las diferentes condiciones foliares. En dichas imágenes, las cajas delimitadoras generadas por el modelo se superponen sobre las hojas clasificadas como sanas o enfermas, lo cual permite validar tanto la precisión espacial como la capacidad de reconocimiento semántico del sistema desarrollado.

Figura 4. Imágenes de validación del modelo.



Fuente: Los autores

DISCUSIÓN

El uso de modelos basados en inteligencia artificial, como YOLOv12, ha demostrado ser una herramienta poderosa para la detección de enfermedades en cultivos. En particular, el modelo desarrollado en este estudio logró una alta precisión, Recall y mAP50 en la identificación de hojas sanas y afectadas por oídio y sarna. Estos resultados son consistentes con estudios previos que han utilizado modelos YOLO para la detección de enfermedades de las hojas del manzano (Wang et al., 2022)(Xu & Wang, 2023)

Uno de los aspectos clave del presente trabajo es la importancia de la calidad de los datos de entrenamiento. La correcta anotación de las imágenes utilizando Roboflow garantizó que el modelo recibiera información precisa y bien estructurada, lo cual fue fundamental para obtener resultados confiables. Sin embargo, es necesario destacar que la calidad del modelo depende directamente de la cantidad y diversidad de los datos disponibles. Futuros estudios podrían beneficiarse de conjuntos de datos más grandes y diversificados, incluyendo imágenes capturadas bajo diferentes condiciones ambientales y en diferentes etapas de desarrollo de las enfermedades.

Además, la alta Recall y mAP50 observadas en el caso de la sarna (1.0) sugieren que el modelo tiene un excelente rendimiento en la detección de esta enfermedad. Sin embargo, la precisión ligeramente menor en el caso del oídio (0.75) indica que aún existen oportunidades para mejorar la capacidad del modelo en la diferenciación de esta enfermedad, especialmente en casos donde los síntomas sean menos evidentes.

CONCLUSIONES

Este trabajo presenta un sistema de detección automática de enfermedades de sarna y oídio en hojas de plantas, basado en el modelo YOLOv12. El uso de técnicas de data augmentation y un enfoque riguroso en la preparación del dataset permitió obtener resultados altamente prometedores:

- Precisión máxima de 0.9923.
- Recall máximo de 0.89524.
- mAP50 de 0.92512.

Estos resultados demuestran que el modelo es capaz de detectar con alta precisión las

enfermedades en imágenes de hojas, lo que podría tener importantes aplicaciones prácticas en la agricultura, facilitando la inspección temprana y la toma de decisiones para el control de plagas y enfermedades.

No obstante, se identificaron algunas limitaciones, como la confusión entre oídio y hojas sanas, lo que sugiere que el modelo podría beneficiarse de un dataset más diverso y de la incorporación de características adicionales para mejorar su discriminación.

En resumen, este estudio demuestra la viabilidad de utilizar modelos de aprendizaje profundo, específicamente YOLOv12, para la detección automatizada de enfermedades agrícolas. Se recomienda continuar investigando para optimizar aún más el desempeño del modelo y explorar su integración en sistemas de monitoreo en tiempo real, como drones o cámaras inteligentes.

REFERENCIAS BIBLIOGRÁFICAS

- Analuisa-Aroca, I., Vergara-Romero, A., Betzaida, I., & Almeida, P. (2023). Convolutional neural networks ResNet-50 for weevil detection in corn kernels. *Redes neuronales convolucionales ResNet-50 para la detección de gorgojo en granos de maíz. Scientia Agropecuaria*, 14(3), 385–394. <https://doi.org/10.17268/sci.agropecu.2023.034>
- Facuy-Delgado, J., Molina-Oleas, W., & Ortega-Ponce, L. (2022). Revisión sistemática de los algoritmos para detección de enfermedades fúngicas en el arroz. *MQRInvestigar*, 6(4), 608–620. <https://doi.org/10.56048/MQR20225.6.4.2022.608-620>
- Huanca Namuche, A. A., Terry Alvarado, B. S., & García Estrella, C. (2025). Redes neuronales convolucionales para la clasificación de la mancha negra en los cítricos. *Sapiens International Multidisciplinary Journal*, 2, 1–20. <https://doi.org/https://doi.org/10.71068/4x9sw526>
- Li, T., Zhang, L., & Lin, J. (2024). Precision agriculture with YOLO-Leaf: advanced methods for detecting apple leaf diseases. *Frontiers in Plant Science*, 15, 1452502. <https://doi.org/10.3389/FPLS.2024.1452502/BIBTEX>
- Mora, E. A. H., Huitrón, V. G., Rangel, H. R., & Sosa, L. E. A. (2021). Detección de enfermedades foliares con arquitecturas de redes neuronales convolucionales. *RINDERESU*, 5(1), 18–40. <http://rinderesu.com/index.php/rinderesu/article/view/46>
- Plant Pathology* 2021 - FGVC8. (2021, March 15). <https://www.kaggle.com/datasets/itselif/comprehensive-plant-disease-dataset>
- Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You Only Look Once: Unified, Real-

Time Object Detection. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. <http://pjreddie.com/yolo/>

Roboflow: Computer vision tools for developers and enterprises. (2025). <https://roboflow.com/>

Rodríguez Torres, D. F., Garzón Alfonso, W., & Santiago Cely, C. P. (2023). Machine Learning en la detección de enfermedades en cultivos. *Revista de La Escuela Colombiana de Ingeniería*, 133, 23–31. <http://www.escuelaing.edu.co/revista.htm>

Saldaña Valenzuela, S. (2023). Inteligencia Artificial: Detección de enfermedades en el cultivo de frijol. *Revista Naturaleza, Sociedad Y Ambiente*, 10, 53–58. <https://revistacunsurori.com/index.php/revista/article/view/87/115>

Tian, Y., Ye, Q., & Doermann, D. (2025). *YOLOv12: Attention-Centric Real-Time Object Detectors*. <https://doi.org/10.0>

Wang, Y., Wang, Y., & Zhao, J. (2022). MGA-YOLO: A lightweight one-stage network for apple leaf disease detection. *Frontiers in Plant Science*, 13, 927424. <https://doi.org/10.3389/FPLS.2022.927424/BIBTEX>

Xu, W., & Wang, R. (2023). ALAD-YOLO:an lightweight and accurate detector for apple leaf diseases. *Frontiers in Plant Science*, 14, 1204569. <https://doi.org/10.3389/FPLS.2023.1204569/BIBTEX>

©2026 por los autores. Este artículo es de acceso abierto y distribuido según los términos y condiciones de la licencia Creative Commons Atribución-NoComercial-CompartirIgual 4.0 Internacional (CC BY-NC-SA 4.0) (<https://creativecommons.org/licenses/by-nc-sa/4.0/>)