

LOCALIZACIÓN DE HOJAS Y FRUTOS EN PLANTAS DE JITOMATE MEDIANTE YOLO8VN

LOCATION OF LEAVES AND FRUITS IN TOMATO PLANTS USING YOLO8VN

Carranza Flores José Luis¹, Magadán Salazar Andrea¹, Cristóbal Alejo Jairo², Fuentes-Pacheco Jorge¹

¹Tecnológico Nacional de México/ Centro Nacional de Investigación y Desarrollo Tecnológico (CENIDET), d23ce167@cenidet.tecnm.mx, <https://orcid.org/0000-0002-0235-8685>.

¹Tecnológico Nacional de México/ Centro Nacional de Investigación y Desarrollo Tecnológico (CENIDET), andrea.ms@cenidet.tecnm.mx, <https://orcid.org/0000-0001-5474-4150>.

²Tecnológico Nacional de México/CONKAL, jairo.ca@conkal.tecnm.mx, <https://orcid.org/0000-0001-9354-1129>.

¹Tecnológico Nacional de México/ Centro Nacional de Investigación y Desarrollo Tecnológico (CENIDET), jorge.fp@cenidet.tecnm.mx, <https://orcid.org/0000-0002-9060-8276>.

<https://doi.org/10.61117/ipsumtec.v8i2.383>

Resumen -- La localización de objetos de interés con aprendizaje profundo es un área de la visión por computadora que se enfoca en identificar y delimitar la ubicación de objetos específicos dentro de una imagen. En este artículo se presenta una propuesta de localizar el rectángulo mínimo de hojas y frutos de jitomate de manera automática, mediante el análisis de imágenes con Yolo8vn. Los resultados obtenidos muestran un buen desempeño aún frente a factores como la presencia de cambios en la perspectiva, escala, iluminación, presencia de polvo, patologías, etc.

Palabras Clave: deep learning, CNN, YOLO8vn, hojas de tomate, fruto del tomate.

Abstract -- Locating objects of interest with deep learning is an area of computer vision that focuses on identifying and delineating the location of specific objects within an image. This article presents a proposal to locate the minimum rectangle of tomato leaves and fruits automatically, through image analysis with Yolo8vn. The results obtained show good performance even in the face of factors such as the presence of changes in perspective, scale, lighting, presence of dust, pathologies, etc.

Key words – deep learning, CNN, YOLO8vn, tomato leaf, tomato fruit.

INTRODUCCIÓN

De acuerdo con la Organización de las Naciones Unidas, para el 2050 habrá 9,700 millones de personas habitando el planeta [1], esto significa que la demanda de alimentos aumentará drásticamente, por lo que, el sector agrícola debe hacer uso de tecnologías como apoyo para tomar decisiones en el cuidado de los cultivos agrícolas, surgiendo una nueva área denominada agricultura de precisión [2]. La agricultura de precisión se apoya del avance de la inteligencia artificial en las áreas de visión artificial, aprendizaje de máquinas y aprendizaje

profundo, a través de las cuales se lleva a cabo el diseño y desarrollo de sistemas inteligentes que automatizan procesos que, para el ser humano implican mucho tiempo. El jitomate es la hortaliza más consumida en todo el mundo. En México constituye una de las hortalizas más importantes debido a la cantidad de empleos directos e indirectos que genera su cultivo y al número de divisas que ingresan al país por concepto de su comercialización [3]. Además de su importancia económica, el jitomate es también fuente de vitaminas, minerales y antioxidantes, los cuales son fundamentales para la nutrición y la salud humana. La presencia de enfermedades en su cultivo provoca grandes pérdidas o malformaciones de los frutos. Es por ello por lo que, se propone agilizar el proceso de reconocimiento de las patologías, desde etapas tempranas para apoyar a los agricultores en la toma de decisiones para evitar su propagación y el uso indiscriminado de pesticidas y fungicidas.

En la literatura del área, se observa que hay un interés especial por el cultivo de jitomate debido a su importancia en el mercado mundial. Las actividades principales que se llevan a cabo son la identificación de los síntomas en la hoja para determinar la enfermedad y/o el patógeno causante (virus, bacteria, hongos) como en [4-13]. Es notorio que pocos trabajos consideran el reconocimiento de patologías en el fruto como [14-17]. Esto porque los virus, bacterias, hongos y plagas afectan primeramente las hojas de las plantas y una vez que han ocasionados daños se presentan también en tallos y frutos.

También se observa que, los sistemas reportados realizan la detección de diversas enfermedades en los cultivos agrícolas, mejorando su rendimiento mediante el uso de arquitecturas de aprendizaje profundo: AlexNet [18], GoogleNet [19-20], ResNet50 [21], ResNet18 [22], Xception [23], VGG16 [24-27]. En la mayoría de ellos

lleva a cabo la detección de agentes patógenos mediante el análisis de imágenes en condiciones controladas con nivel de daño avanzado. Es decir, son imágenes de una hoja recortada (segmentada), alineada y que contrasta con respecto al fondo como en [22], [24], [28-32]. Sin embargo, estas condiciones no siempre se presentan en un ambiente real, donde lo importante es realizar la identificación de los síntomas en una etapa temprana, lo que permitirá se tomen medidas en la eliminación de los patógenos. Para el reconocimiento de la presencia de los primeros síntomas en las plantas, lo primero que se necesita es contar con un sistema que realice la localización o segmentación de las hojas y frutos de las plantas de jitomate en condiciones reales.

El presente trabajo describe la metodología para localizar las hojas y el fruto del jitomate; considerando que la imagen adquirida en condiciones no controladas puede contener la presencia de varias hojas al mismo tiempo con diferentes perspectivas y tamaños, oclusión entre ellas, sombras, presencia de polvo, agua; además de cambios en la intensidad luminosa, presencia de cambios en la tonalidad de las hojas por los síntomas de patologías, etc. Factores que aumentan la complejidad del proceso de segmentación y que por ello pocas veces son considerados. Para su desarrollo, el sistema propuesto utiliza el enfoque de aprendizaje profundo, mediante la implementación, entrenamiento y optimización de la red YOLO8vn [34].

El artículo está organizado de la siguiente forma: se presenta el marco teórico de la arquitectura YOLO8vn y las métricas utilizadas para evaluar el rendimiento del sistema. Se detalla la metodología seguida para el desarrollo el sistema propuesto. Además, se presenta la experimentación realizada y se discuten los resultados y, finalmente se dan las conclusiones del trabajo.

ARQUITECTURA YOLO

Desde su creación en 2015, la variante YOLO (*You Only Look Once*) de detectores de objetos ha crecido rápidamente, YOLO8vn fue lanzado en enero de 2023. Las variantes YOLO se basan en el principio de rendimiento en tiempo real y de alta clasificación, basado en parámetros computacionales limitados pero eficientes [35]. Este principio se ha encontrado en el ADN de todas las versiones de YOLO con creciente intensidad, a medida que las variantes evolucionan para abordar los requisitos de la inspección de calidad automatizada dentro del dominio de detección de defectos de superficies industriales, como la necesidad de una detección rápida, alta precisión y despliegue en áreas restringidas.

En la literatura se observa que la familia de YOLO reporta buenos resultados en la localización de objetos

diversos. Sin embargo, dichas arquitecturas se han considerado para la localización de hojas en distintos cultivos como se puede ver en [6], [11], [27], [33], [34]. Sin embargo, es notorio que, la mayoría de los artículos señalan llevar a cabo la localización mediante el análisis de imágenes en condiciones controladas. Por ejemplo, en los trabajos [4-26], reportan un desempeño satisfactorio en condiciones controladas; es decir, se requiere contar con imágenes cuyo objeto principal (hoja o fruto) se encuentre recortado, alineado, localizado e incluso el fondo sea contrastante. YOLO8vn [36] es un modelo de visión artificial de última generación creado por *ultralytics*. Diseñado específicamente para ser ligera y eficiente, ideal para dispositivos con recursos limitados como teléfonos móviles, dispositivos de IoT, y otras aplicaciones embebidas. Además, esta optimizado para ser mucho más pequeño en comparación con sus predecesores. El modelo contiene soporte listo para usar en tareas de detección, clasificación y segmentación de objetos. YOLO8vn viene con mejoras tanto arquitectónicas como de experiencia para desarrolladores. En comparación con las versiones anteriores a YOLO8vn viene con:

1. Nuevo sistema de detección de anclajes
2. Cambios en los bloques convoluciones utilizados en el modelo
3. Aumento de mosaico aplicado durante el entrenamiento, desactivado antes de las ultimas 10 épocas

Además, YOLO8vn viene con cambios para mejorar la experiencia del desarrollador con el modelo, y el modelo se encuentra empaquetado como una biblioteca que puede instalar en código Python.

Métricas

Las métricas de evaluación permiten medir el rendimiento de un modelo predictivo. Es decir, su objetivo es proporcionar de manera objetiva cómo el modelo está funcionando. Las métricas clásicas para evaluar el desempeño del modelo son:

- a) *Recall* (Sensibilidad o tasa de verdaderos positivos); es la proporción de verdaderos positivos entre todos los casos que realmente son positivos. Mide la proporción de instancias positivas que fueron correctamente identificadas por el modelo, ver la expresión 1.

$$Recall = \frac{TP}{TP+FN} \quad \text{Ec. (1)}$$

- b) *Precision*; Es la proporción de verdaderos positivos entre todas las instancias que han sido predichas como positivas. La expresión 2 muestra el cálculo de *precision*.

$$Precision = \frac{TP}{TP+FP} \quad \text{Ec. (2)}$$

- c) *F1 Score*: es la media armónica de la *precisión* y *recall*. Proporciona un balance entre ambas métricas y es útil cuando hay un desbalance en las clases. La expresión 3 muestra el cálculo de la métrica antes mencionada.

$$F1\ Score = 2 * \frac{Precision * Recall}{Precision+Recall} \quad \text{Ec. (3)}$$

DESARROLLO

Metodología

De manera general, la propuesta para llevar a cabo la localización de hojas y frutos, en imágenes de plantas de jitomate adquiridas en condiciones reales, mediante la red YOLO8vn, se integra de los siguientes pasos que son detallados, posteriormente, en las siguientes subsecciones, ver la figura 1.

- Se adquiere y define el conjunto de imágenes que se utilizan para llevar a cabo el entrenamiento y validación del sistema.
- Las imágenes se normalizan a un formato y tamaño.
- Para el entrenamiento se requiere contar con el *ground truth*, que se crea mediante un proceso de etiquetado.
- Se definen y aplican las operaciones para realizar un aumento de datos.
- Se distribuyen de manera aleatoria las imágenes en 80:10:10 para los subconjuntos a utilizarse en el entrenamiento, validación y pruebas.
- En el entrenamiento de la red, se optimizan los parámetros y como resultado se obtienen las predicciones de cuadros delimitadores en diferentes escalas y posiciones. Posteriormente, se aplica un umbral de confianza para refinar los cuadros delimitadores. Para ello, el modelo es entrenado con los parámetros que se describen en la sección de entrenamiento, ver tabla 4.
- Finalmente, la red proporciona los cuadros delimitadores finales con etiquetas y las puntuaciones, ver la sección de experimentación.



Figura 1. Esquema del proceso de localización.

DataSet

Existen diversos conjuntos de datos públicos que se integran de imágenes de hojas sanas y enfermas de diversos cultivos. En especial, aquellos conjuntos que contienen hojas de jitomate y se encuentran entre los más utilizados son: Plantvillage [37], ImageNet [38] y Kaggle [39]. Es importante destacar que, la mayoría de las imágenes de los dos primeros conjuntos, se encuentran en condiciones controladas; es decir, el objeto de interés se muestra en un fondo blanco o en un color uniforme de tal manera que no hay ruido que obstruya los objetos de estudio. Además, dos de los datasets antes mencionados solo contienen imágenes de hojas de diferentes cultivos y no presentan frutos

El conjunto tomato leaf disease dataset de Kaggle contiene algunas imágenes en fondo blanco, pero la mayoría de las fotografías fueron capturadas en un ambiente no controlado, como se observa en la figura 2. Por esta razón, se optó por usar el conjunto de Kaggle y se seleccionaron las imágenes necesarias para el desarrollo del proyecto. El conjunto contiene 10000 imágenes de diversos tamaños, que presentan diferentes tipos de enfermedades de las plantas, incluidas infecciones bacterianas, enfermedades fúngicas, infestaciones de plagas e infecciones virales.



Figura 2. Imágenes del conjunto de datos de Kaggle [4].

Con el objetivo de obtener un conjunto de imágenes más amplio y diverso, se llevó a cabo una visita a invernaderos de jitomate en el Estado de Guerrero, donde se capturaron fotografías (como se muestra en la Figura 3), utilizando tres dispositivos diferentes.



Figura 3. Ejemplo de imágenes adquiridas en la sierra de Guerrero, México.

Finalmente, se complementó el conjunto de datos con imágenes descargadas de internet, tanto de frutos sanos como de aquellos que presentan enfermedades. Algunos ejemplos se presentan en la Figura 4.



Figura 4. Imágenes descargadas de internet.

En la tabla 1 se muestra la cantidad de imágenes capturadas en invernaderos de Guerrero, del conjunto de datos kaggle y las descargadas de internet.

Tabla 1. Cantidad de imágenes del dataset creado.

Propio		Kaggle		Internet	
Hoja	Fruto	Hoja	Fruto	Hoja	Fruto
400	400	800	-	100	100
Total: 1800 [1300 de hojas y 500 de frutos]					

Preprocesamiento

a) Normalización:

Debido a que se tienen imágenes de diversas fuentes, es necesario normalizar las imágenes. Para ello, el formato de la imagen establecido es jpg, y se modificó el tamaño de la imagen de entrada. Algunas de ellas tenían el tamaño de 1266 x 1390, otras de 4224 x 5632 y otras de 3468 x 4624. Todas ellas fueron redimensionadas a un tamaño de 640 x 640 píxeles.

b) Etiquetado:

El proceso de etiquetado consiste localizar el recuadro mínimo que contenga el objeto de interés, es decir, asignar la posición del punto superior izquierdo y del punto inferior derecho para que la red aprenda a identificar los objetos. Este proceso, es indispensable para obtener resultados que aseguren una mayor precisión en la localización de los objetos. Hay diferentes formas de hacerlo. El etiquetado manual implica que los humanos observen las imágenes y asignen las posiciones manualmente.

El etiquetado semiautomático implica el uso de algoritmos para sugerir las posiciones, que luego son revisadas y corregidas por humanos. El etiquetado automático utiliza algoritmos de procesamiento de imágenes para localizar los objetos

Para el proceso de etiquetado se utilizó Roboflow [40] y de manera manual se hicieron precisiones en algunas imágenes donde le rectángulo mínimo no estaba correcto. Además, se asignan etiquetas a las imágenes de las hojas y frutos, en este caso en particular de la planta de jitomate como se observa en la figura 5. Durante el proceso de asignación se crearon dos categorías: la clase 0 que hace referencia a los frutos y la clase 1 a las hojas. Posteriormente, se generan de manera aleatoria los tres subconjuntos: entrenamiento, validación y prueba.

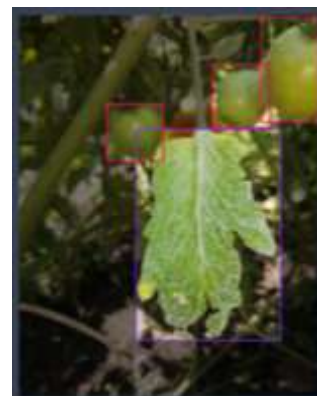


Figura 5. Imagen etiquetada en RoboFlow.

- El primer conjunto proporciona los datos de entrada para el entrenamiento de la red YOLO8vn.
- El segundo grupo de imágenes permite ajustar los parámetros durante cada época del entrenamiento del algoritmo, de tal forma que se optimice la precisión del mismo.
- El tercer grupo permite evaluar el modelo, una vez finalizado el entrenamiento. Con este conjunto se obtienen los resultados finales de la detección automática reportados.

Los datos se exportan para el entrenamiento en tres carpetas que contienen las imágenes y las etiquetas.

Aumento de Datos.

Considerando que las redes neuronales convolucionales necesitan grandes cantidades de imágenes, se decidió aplicar algunas transformaciones para tener una mayor cantidad de información y evitar el sobreajuste, la tabla 2 muestra las técnicas implementadas. Finalmente, se obtuvieron 3119 imágenes.

Tabla 2. Técnicas aplicadas para el aumento de datos.

Técnica	Valor
Rotación	90° upside down
Saturación	34%
Ruido Gaussiano	1.96%
Redimensión	640 x 640

Después del aumento de datos. El dataset se distribuye con 80% para entrenamiento, 10% para validación y 10% para las pruebas como se observa en la tabla 3.

Tabla 3. Distribución del conjunto de imágenes.

Cantidad de imágenes	Entrenamiento	Validación	Pruebas
	80%	10%	10%
3119	2730	150	150

Entrenamiento

La implementación se llevó a cabo utilizando la herramienta Google Colab para el entrenamiento de la red y Pycharm 2024.2.3 para las pruebas. El proceso de entrenamiento consiste en el aprendizaje y evaluación iterativo, esto debido a que el objetivo del modelo es localizar el objeto de interés dentro de la imagen. Para iniciar con el entrenamiento es necesario configurar algunos parámetros básicos. Después de una serie de entrenamientos se observó que el modelo tenía un buen rendimiento con los siguientes valores, ver tabla 4.

Tabla 4. Hiperparámetros utilizados.

Hiperparámetros	valor
Batch size	64
Ephocs	20
Width shape y height shape	640 x 640

DISCUSIÓN Y ANÁLISIS DE RESULTADOS

El objetivo del experimento fue evaluar el rendimiento del modelo para localizar de manera automática el recuadro mínimo que contenga al objeto de interés, en este caso, clasificando si se trata de una hoja o de un fruto, ya que posteriormente requieren un análisis diferente. La figura 6 muestra de manera visual los resultados obtenidos al encerrar en un *bounding box* los objetos de interés; a pesar de que las imágenes al ser adquiridas en ambientes no controlados presentan hojas y frutos sanos y enfermos. Incluso, se observa en dichas imágenes, que el sistema es capaz de segmentar a las hojas y frutos a pesar de la escala, perspectiva, edad de las hojas, traslape entre ellas, aún ante la presencia de polvo, y problemas de tonalidad y forma.



Figura 6. Resultados obtenidos.

El modelo muestra un buen desempeño en la fase de validación, alcanzando una precisión de 97.6% (véase Figura 7), un recall de 97% (véase Figura 8) y un F1 de 84.28% (véase Figura 9). Con respecto al valor de precisión significa que, de todos los frutos y hojas que el modelo detecta, efectivamente son verdaderos.

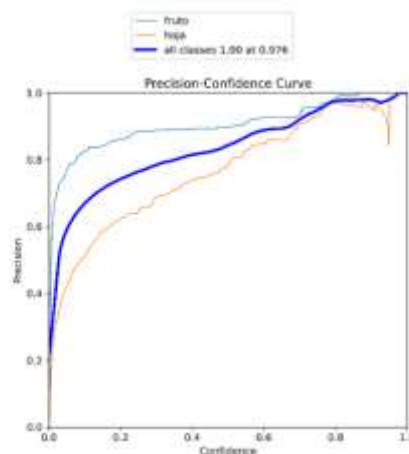


Figura 7. Gráfica de la curva de confianza en la precisión.

Los resultados con recall nos dice que, de todos los frutos y hojas existentes, el modelo identifica la mayor parte de instancias en la imagen.

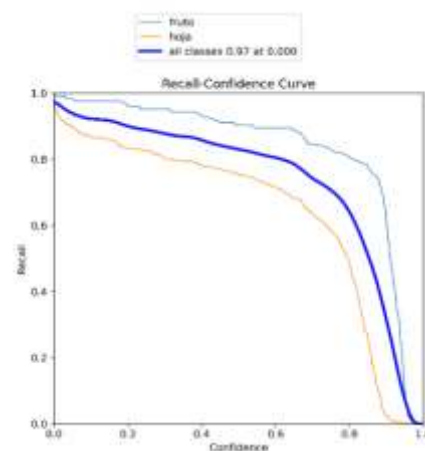


Figura 8. Gráfica de la curva de confianza de recall.

Los resultados de F1 muestran que el modelo clasifica correctamente, pero hay algunos objetos (hojas y frutos) no localizados esto debido a que la mayoría de las imágenes fueron capturadas en ambientes no controlados y algunas de ellas tienen algunas oclusiones.

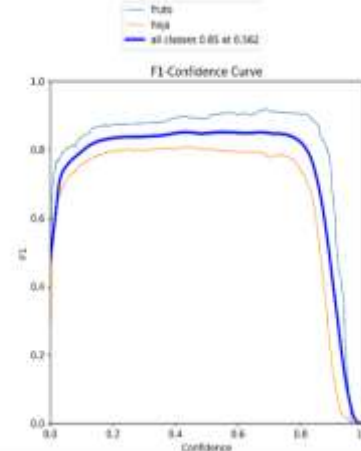


Figura 9. Gráfica de la curva de confianza de F1.

Como se observa en el aprendizaje de la red se tienen algunos problemas de localización en las imágenes que contienen hojas y frutos en el fondo de la imagen. Además, algunas imágenes del conjunto de prueba estaban desenfocadas.

CONCLUSIONES

De acuerdo con los puntajes obtenidos, se concluye que el modelo obtenido con YOLO8vn permite que la localización sea precisa en imágenes capturas en ambientes no controlados. Además, que el uso de herramientas para el etiquetado es eficaz en este caso en particular debido a que el *bounding box* encierra completamente el objeto de interés. La experimentación muestra que la red YOLO8vn es una buena opción para llevar a cabo la localización de las hojas y frutos. Además, el aplicar un aumento de datos, permite contar con una mayor cantidad y diversidad de imágenes lo que ayudará a obtener un mejor rendimiento.

En trabajo ha presentado resultados iniciales que demuestran la factibilidad de emplear técnicas de visión por computadora y aprendizaje profundo para la clasificación de frutos y hojas. Sin embargo, los hallazgos obtenidos representan tan solo el primer paso hacia un sistema más completo de reconocimiento de patologías en plantas. Por ello, se propone como trabajo futuro:

1. Ampliar el Conjunto de Datos: Incluir un mayor número de imágenes que abarquen: condiciones de luz y escenarios de campo, así como diversas etapas de desarrollo y niveles de infección.
2. Incorporar Patologías Específicas: Extender la metodología para abordar el diagnóstico de enfermedades comunes en las plantas analizadas, de modo que el modelo pueda diferenciar entre hojas sanas y hojas con patologías específicas.

Estas propuestas buscan no solo mejorar la exactitud y eficiencia del modelo presentado, sino también avanzar hacia aplicaciones prácticas que contribuyan al control y la prevención de patologías en plantas, beneficiando al sector agrícola y la sostenibilidad de los cultivos.

BIBLIOGRAFÍA

- [1] Organización de las Naciones Unidas, 2022. [Consultado 08 de agosto de 2024] Disponible en: <https://www.un.org/es/global-issues/population>.
- [2] Hernández, R. R. (2021). La agricultura de precisión. Una necesidad actual. *Revista Ingeniería Agrícola*, 11(1), 67-74.
- [3] Instituto de Investigación y capacitación agropecuaria, acuícola y forestal ICAMEX, (2023). Cultivo de Jitomate.
- [4] Jia, W., Xu, Y., Lu, Y., Yin, X., Pan, N., Jiang, R., & Ge, X. (2023). An accurate green fruits detection method based on optimized YOLOX-m. *Frontiers in Plant Science*, 14, 1187734.
- [5] Singh, A. K., Sreenivasu, S. V. N., Mahalaxmi, U. S. B. K., Sharma, H., Patil, D. D., & Asenso, E. (2022). Hybrid feature-based disease detection in plant leaf using convolutional neural network, bayesian optimized SVM, and random forest classifier. *Journal of Food Quality*, 2022, 1-16.
- [6] Jia, W., Xu, Y., Lu, Y., Yin, X., Pan, N., Jiang, R., & Ge, X. (2023). An accurate green fruits detection method based on optimized YOLOX-m. *Frontiers in Plant Science*, 14, 1187734.
- [7] Peng, D., Li, W., Zhao, H., Zhou, G., & Cai, C. (2023). Recognition of tomato leaf diseases based on DIMPCNET. *Agronomy*, 13(7), 1812.
- [8] Liang, J., & Jiang, W. (2023). A ResNet50-DPA model for tomato leaf disease identification. *Frontiers in Plant Science*, 14, 1258658.
- [9] Chen, H., Wang, Y., Jiang, P., Zhang, R., & Peng, J. (2023). LBFNet: A Tomato Leaf Disease Identification Model Based on Three-Channel Attention Mechanism and Quantitative Pruning. *Applied Sciences*, 13(9), 5589.
- [10] Debnath, A., Hasan, M. M., Raihan, M., Samrat, N., Alsulami, M. M., Masud, M., & Bairagi, A. K. (2023). A Smartphone-Based Detection System for Tomato Leaf Disease Using EfficientNetV2B2 and Its Explainability with Artificial Intelligence (AI). *Sensors*, 23(21), 8685.
- [11] Umar, M., Altaf, S., Sattar, K., Somroo, M. W., & Sivakumar, S. (2023). Multi-Disease Recognition in Tomato Plants: Evaluating the Performance of CNN and Improved YOLOv7 Models for Accurate Detection and Classification.
- [12] Flores Colorado, O. E., Cervantes Canales, J., García-Lamont, F. y Ruiz Castilla, J. S. (2023). Identificación de las principales enfermedades de la planta del café (*Coffea arabica*) a través de visión artificial. *CIENCIA ergo-sum*, 30(3).
- [13] Fadhillah, M., & Suryani, D. (2023). Android Application for Tomato Leaf Disease Prediction Based on MobileNet Fine-tuning. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, 7(6), 1260-1267.
- [14] Islam, M. M., Talukder, M. A., Sarker, M. R. A., Uddin, M. A., Akhter, A., Sharmin, S., ... & Debnath, S. K. (2023). A deep learning model for cotton disease prediction using fine-tuning with smart web application in agriculture. *Intelligent Systems with Applications*, 20, 200278.
- [15] Khatoon, S., Hasan, M. M., Asif, A., Alshamari, M., & Yap, Y. (2021). Image-based automatic diagnostic system for tomato plants using deep learning. *Comput. Mater. Contin*, 67(1), 595-612.

- [16] Khattak, A., Asghar, M. U., Batool, U., Asghar, M. Z., Ullah, H., Al-Rakhami, M., & Gumaei, A. (2021). Automatic detection of citrus fruit and leaves diseases using deep neural network model. *IEEE access*, 9, 112942-112954.
- [17] Zeng, Q., Sun, J., & Wang, S. (2023). DIC-Transformer: interpretation of plant disease classification results using image caption generation technology. *Frontiers in Plant Science*, 14.
- [18] Da Silva Abade, A., de Almeida, A. P. G., & de Barros Vidal, F. (2019). Plant Diseases Recognition from Digital Images using Multichannel Convolutional Neural Networks. In *VISIGRAPP (5: VISAPP)* (pp. 450-458).
- [19] López, J. A. M., De la Torre Gutiérrez, H., & López, F. J. H. Detección de antespacios urbanos usando YOLO: Caso de estudio Mexicali.
- [20] Mohanty, S. P., Hughes, D. P., & Salathé, M. (2016). Using deep learning for image-based plant disease detection. *Frontiers in plant science*, 7, 1419.
- [21] Sagar, A., & Dheeba, J. (2020). On using transfer learning for plant disease detection. *BioRxiv*, 2020-05.
- [22] Zhao, S., Peng, Y., Liu, J., & Wu, S. (2021). Tomato leaf disease diagnosis based on improved convolution neural network by attention module. *Agriculture*, 11(7), 651.
- [23] Chipuli, J. P., & Luwemba, G. W. (2023). Tomato Plant Leaf Disease Detection Using Image Recognition: A Case Study of Mlali in Morogoro Region, Tanzania. *European Journal of Information Technologies and Computer Science*, 3(4), 18-25.
- [24] Rangarajan, A. K., Purushothaman, R., & Ramesh, A. (2018). Tomato crop disease classification using pre-trained deep learning algorithm. *Procedia computer science*, 133, 1040-1047.
- [25] Latif, G., Abdelhamid, S. E., Mallouhy, R. E., Alghazo, J., & Kazimi, Z. A. (2022). Deep learning utilization in agriculture: Detection of rice plant diseases using an improved CNN model. *Plants*, 11(17), 2230.
- [26] Jung, M., Song, J. S., Shin, A. Y., Choi, B., Go, S., Kwon, S. Y., ... & Kim, Y. M. (2023). Construction of deep learning-based disease detection model in plants. *Scientific Reports*, 13(1), 7331.
- [27] Liu, Y., & Yu, Q. (2024). Real-time and lightweight detection of grape diseases based on Fusion Transformer YOLO. *Frontiers in Plant Science*, 15, 1269423.
- [28] Fuentes, A., Yoon, S., Kim, S. C., & Park, D. S. (2017). A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition. *Sensors*, 17(9), 2022.
- [29] Fuentes, A. F., Yoon, S., Lee, J., & Park, D. S. (2018). High-performance deep neural network-based tomato plant diseases and pests diagnosis system with refinement filter bank. *Frontiers in plant science*, 9, 1162.
- [30] Gonzalez-Huitron, V., León-Borges, J. A., Rodríguez-Mata, A. E., Amabilis-Sosa, L. E., Ramírez-Pereda, B., & Rodríguez, H. (2021). Disease detection in tomato leaves via CNN with lightweight architectures implemented in Raspberry Pi 4. *Computers and Electronics in Agriculture* 181, 105951.
- [31] Bhujel, A., Kim, N. E., Arulmozhi, E., Basak, J. K., & Kim, H. T. (2022). A lightweight Attention-based convolutional neural networks for tomato leaf disease classification. *Agriculture*, 12(2), 228.
- [32] Ullah, Z., Alsubaie, N., Jamjoom, M., Alajmani, S. H., & Saleem, F. (2023). EffiMob-Net: A Deep Learning-Based Hybrid Model for Detection and Identification of Tomato Diseases Using Leaf Images. *Agriculture*, 13(3), 737.
- [33] Mbouembe, P. L. T., Liu, G., Park, S., & Kim, J. H. (2023). Accurate and fast detection of tomatoes based on improved YOLOv5s in natural environments. *Frontiers in Plant Science*, 14.
- [34] Bochkovskiy, A., Wang, C. Y., & Liao, H. Y. M. (2020). Yolov4: Optimal speed and accuracy of object detection. *arXiv preprint arXiv:2004.10934*.
- [35] Yolo. [Consultado el 12 de mayo de 2024] Disponible en: docs.ultralytics.com/es/models/yolov8/
- [36] Yolo8vn. [Consultado el 12 de mayo de 2024] <https://github.com/ultralytics/ultralytics?tab=readme-ov-file>
- [37] Plantvillage. [Consultado el 12 de octubre de 2024] Disponible en: <https://plantvillage.psu.edu/>
- [38] ImageNet. [Consultado el 12 de octubre de 2024] Disponible en: <https://www.image-net.org/>
- [39] Kaggle. [Consultado el 12 de mayo de 2024] Disponible en: <https://www.kaggle.com/datasets>
- [40] Roboflow. [Consultado el 12 de mayo de 2024] Disponible en: <https://roboflow.com/>

Tabla de Roles de Contribución

Rol	Autor (es)
Conceptualización	Jairo Cristóbal Alejo
Metodología	Andrea Magadán Salazar
Administración del proyecto	Andrea Magadan Salazar, Jorge Fuentes-Pacheco

Software	Jose Luis Carranza Flores
Validación	Jorge Fuentes-Pacheco
Redacción	Jose Luis Carranza Flores, Andrea Magadán Salazar



Esta obra está bajo
una licencia internacional
Creative Commons Atribución 4.0.