

Redes Convolucionales para Agricultura: MobileNetV3 vs. InceptionV3 en el Reconocimiento de Enfermedades en Fresas

RESUMEN: La detección temprana de enfermedades en plantas, como el tizón foliar y cenicilla, es esencial para minimizar pérdidas económicas y preservar la calidad de los cultivos. Para ello, resulta indispensable contar con métodos de diagnóstico que sean precisos, accesibles y eficientes. En este contexto, las redes neuronales profundas han demostrado un gran potencial en el ámbito agrícola, particularmente en el análisis de imágenes foliares para identificar síntomas en etapas iniciales. El presente estudio propone el desarrollo de un sistema de clasificación de enfermedades en hojas de fresa, basado en modelos preentrenados de redes neuronales convolucionales (CNN), específicamente MobileNetV3 e InceptionV3. La metodología incluyó la recolección y preprocesamiento de imágenes, seguidos del entrenamiento, ajuste y validación de ambos modelos. El desempeño de las redes se evaluó mediante métricas estándar como precisión, sensibilidad y especificidad. Los resultados obtenidos permiten una comparación detallada entre MobileNetV3 e InceptionV3, contribuyendo al desarrollo de herramientas inteligentes para la detección eficiente de enfermedades en hojas de fresa. Este avance abre la posibilidad de implementar soluciones portátiles para el diagnóstico en campo, fortaleciendo así el enfoque de agricultura inteligente.

PALABRAS CLAVE: agricultura inteligente, redes neuronales convolucionales (CNN), clasificación de imágenes, enfermedades de la fresa, MobileNetV3-Small, InceptionV3.



Colaboración

Ever Essaú Rodríguez Sandoval; Raquel Ochoa Ornelas; Ansel Yoan Rodríguez González, Tecnológico Nacional de México / Ciudad Guzmán

Fecha de recepción: 10 de septiembre de 2025

Fecha de aceptación: 14 de septiembre de 2025

ABSTRACT: Early detection of plant diseases, such as Leaf Scorch and Powdery Mildew, is essential to minimize economic losses and preserve crop quality. In this context, deep neural networks have shown great potential in the agricultural field, particularly in the analysis of leaf images to identify symptoms in early stages. The present study proposes the development of a strawberry leaf disease classification system based on pretrained convolutional neural network (CNN) models, specifically MobileNetV3 and InceptionV3. The methodology included image collection and preprocessing, followed by training, tuning, and validation of both models. The performance of the networks was evaluated using standard metrics such as accuracy, sensitivity and specificity. The results obtained allow a detailed comparison between MobileNetV3 and InceptionV3, contributing to the development of intelligent tools for efficient disease detection in strawberry leaves. This breakthrough opens the possibility of implementing portable solutions for in-field diagnostics, thus strengthening the smart agriculture approach.

KEYWORDS: intelligent agriculture; convolutional neural networks (CNN); image classification; strawberry diseases; MobileNetV3-Small; InceptionV3.

INTRODUCCIÓN

Las fresas representan un cultivo de gran relevancia para la agricultura mexicana y desempeñan un papel clave en la eco-

nomía global, debido a su alta demanda en los mercados internacionales. México se ha posicionado como uno de los tres principales exportadores de esta fruta, con ingresos cercanos a los 400 millones de pesos en 2020, impulsados principalmente por la producción en estados como Michoacán, Baja California, Jalisco y Guanajuato [1].

Sin embargo, este crecimiento sostenido se ve amenazado por la presencia de enfermedades y plagas de alto impacto, como la cenicilla, *Diplocarpon earlianum*, *Tetranychus urticae* y *Heliothis virescens*. Estos agentes no solo comprometen la calidad y el volumen de la producción, sino que, en algunos casos, también representan riesgos para la salud humana [2].

Frente a este panorama, resulta imprescindible el desarrollo de sistemas eficaces para la detección temprana y el monitoreo continuo de problemas fitosanitarios. En este contexto, la Agricultura de Precisión (AP) ha experimentado un notable auge en los últimos años, integrando tecnologías como la robótica, la teledetección, el análisis de datos y el procesamiento de imágenes para optimizar la gestión de los cultivos [3]. Particularmente, el procesamiento de imágenes y el aprendizaje automático se han consolidado como herramientas eficaces para la detección temprana de plagas y enfermedades, permitiendo análisis rápidos, precisos y rentables a partir de imágenes de hojas, tallos, flores y frutos. Estos métodos, además, posibilitan estimar la severidad de una enfermedad comparando las áreas afectadas con la totalidad de la planta [4].

No obstante, a pesar de los avances en sensores hiperespectrales, su elevado costo limita la adopción por parte de pequeños y medianos productores. En este sentido, el uso de dispositivos móviles y cámaras convencionales en espectro visible se plantea como una alternativa viable y económica para democratizar el acceso a herramientas de diagnóstico, especialmente en regiones rurales con recursos limitados [5].

En la actualidad, la mayoría de los estudios se enfocan en diferentes cultivos o emplean arquitecturas costosas en términos computacionales [6], lo cual dificulta su implementación en dispositivos móviles. No obstante, trabajos recientes muestran la viabilidad de modelos ligeros en condiciones reales de campo [7]. En este sentido, este trabajo se diferencia al comparar el desempeño de MobileNetV3 e InceptionV3 exclusivamente en hojas de fresa, con el propósito de detectar enfermedades en una etapa temprana y determinar cuál arquitectura resulta más adecuada para su despliegue en entornos agrícolas reales.

MATERIAL Y MÉTODOS

Se detalla un planteamiento sistemático de varias etapas clave para llevar a cabo el estudio, que se muestra en términos generales en la Figura 1.

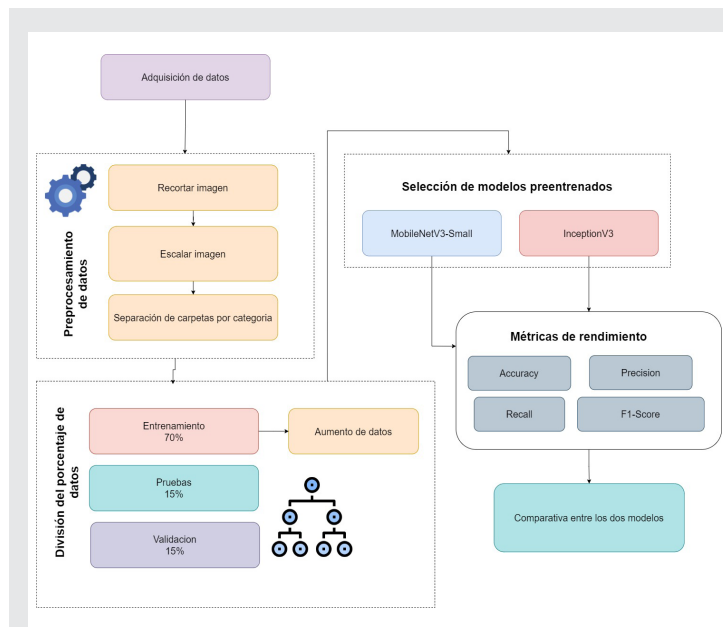


Figura 1: Método propuesto para la clasificación de enfermedades de las hojas de fresa.

Fuente: Elaboración propia.

Conjunto de datos

El conjunto de datos utilizado en este estudio fue recopilado a partir de múltiples fuentes, incluyendo invernaderos locales, campos de cultivo y repositorios públicos como Kaggle [8]. Las imágenes provinieron principalmente de cámaras digitales convencionales y teléfonos móviles, en condiciones variadas de iluminación natural y artificial, así como con diferencias en ángulos de captura. Estas variaciones se mantuvieron deliberadamente para simular escenarios reales de trabajo en campo.

Tabla 1. Cantidad de imágenes obtenidas por fuente.

Fuente	Cantidad de imágenes
Kaggle	1400
Invernaderos	150
Campos de cultivo	168

Fuente: Elaboración propia.

Para garantizar diversidad real, se aplicó el algoritmo de hashing perceptual, el cual consiste en transformar cada imagen en una representación numérica reducida que permite comparar similitudes. Se estableció un umbral de distancia de Hamming de 5 bits para determinar duplicados; es decir, si dos imágenes difieren en menos de cinco posiciones, una de ellas se elimina. Esto permitió depurar imágenes redundantes y asegurar mayor representatividad del conjunto final. Tras la consolidación, el dataset quedó conformado por 4,276 imágenes de hojas de fresa, clasificadas en cinco categorías: saludable, tizón foliar, cenicilla, blossom bli-

ght y tizón angular, con resolución original de 419 x 419 píxeles. Con el fin de balancear las clases, se empleó la librería Augmentor en Python para realizar aumentos mediante rotaciones, escalados, giros y variaciones de brillo. La Figura 2 muestra ejemplos representativos de cada clase. Con el objetivo de mejorar la variabilidad y balancear las clases, se empleó la biblioteca Augmentor de Python. A partir de un conjunto inicial de 1,808 imágenes donde: 700 saludables, 380 con leaf scorch, 329 con cenicilla, 125 con blossom y 274 con angular leaf scorch; fueron aumentadas mediante transformaciones estocásticas como rotaciones, escalados, giros y cambios de brillo, equilibrando así cada clase a 700 imágenes. Estas técnicas ayudan a simular condiciones del mundo real y a mejorar la capacidad de generalización de los modelos durante el entrenamiento.

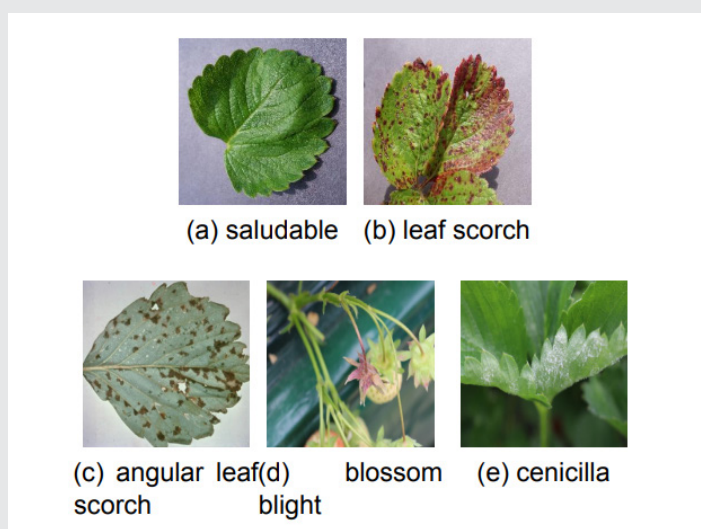


Figura 2: Muestras del conjunto de datos separados por categoría.

Fuente: Elaboración propia.

Preprocesamiento de datos

La etapa de preprocesamiento incluyó la organización sistemática del conjunto de datos, cargando imágenes desde directorios etiquetados según la clase correspondiente. Las imágenes, originalmente en formato RGB, fueron redimensionadas a 224 x 224 píxeles para ajustarse a los requerimientos de entrada de las arquitecturas MobileNetV3 e InceptionV3. Se empleó la clase ImageDataGenerator de TensorFlow para realizar el aumento de datos y cargar las imágenes durante el entrenamiento. Entre las transformaciones aplicadas se incluyeron rotación, traslación, ajustes de brillo y contraste, con el objetivo de incrementar la variabilidad y reducir posibles sesgos en la clasificación. El conjunto fue dividido con barajado y estratificación en subconjuntos de entrenamiento (70%), validación (15%) y prueba (15%), asegurando una distribución representativa de las clases. Este enfoque permitió construir un conjunto de datos balanceado y representativo, adecuado para evaluar la generalización de los modelos en condiciones realistas.

MobileNetV3-SMALL

En otras investigaciones [9] se presenta una modificación de MobileNetV3-small como un modelo de clasificación capaz de ofrecer un alto rendimiento en dispositivos con recursos limitados, como teléfonos móviles o sistemas embebidos [9].

En particular, MobileNetV3-Small se desarrolló con el objetivo de equilibrar precisión y eficiencia computacional, lo que la convierte en una excelente opción para tareas de clasificación de imágenes en tiempo real en entornos con capacidad de cómputo restringida. La Figura 3 muestra de forma esquemática su arquitectura. Una comparación entre los modelos MobileNetV3-Large y EfficientNet-B0 para la clasificación de enfermedades en hojas de fresa para su implementación en dispositivos con baja capacidad de procesamiento y almacenamiento, como sistemas embebidos o móviles [10].

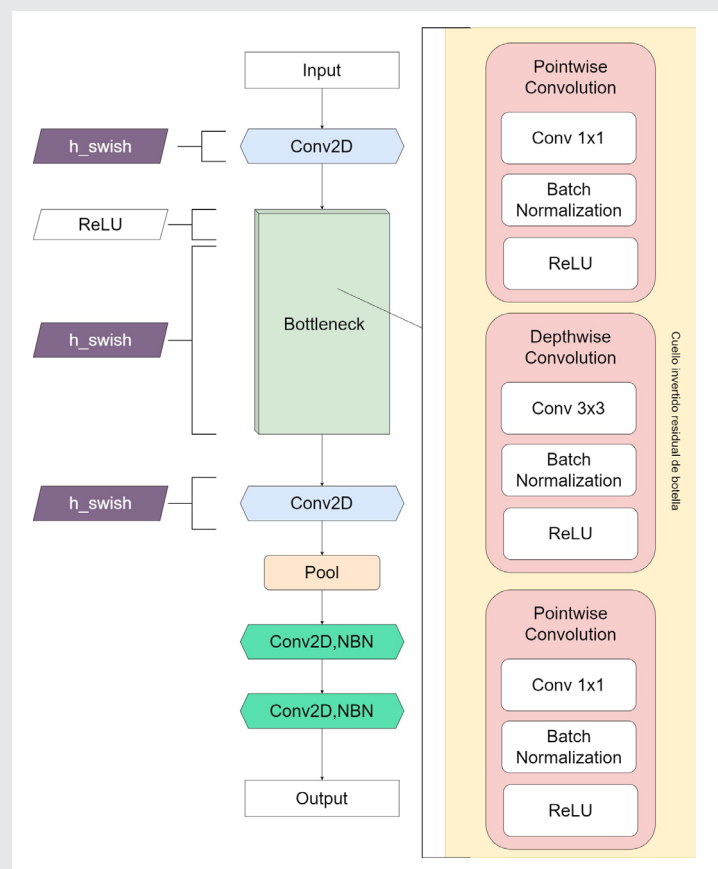


Figura 3: Arquitectura de MobileNetV3-Small.

Fuente: Elaboración propia.

InceptionV3

El modelo InceptionV3 es una red neuronal convolucional profunda ampliamente utilizada en tareas de clasificación de imágenes debido a su alta precisión y capacidad para aprender representaciones complejas. Sin embargo, su entrenamiento desde cero requiere recursos computacionales considerables, lo cual limita su uso en entornos con capacidades restringidas.

Para superar esta limitación, se aplica el aprendizaje por transferencia [11,12], utilizando una versión preentrenada sobre el conjunto de datos ImageNet. En este enfoque, se conservan los pesos de las capas convolucionales y se sustituye la capa densa final, la cual es reentrenada con imágenes de hojas de fresa clasificadas en cinco categorías: saludable, leaf scorch, cenicilla, blossom y angular. La Figura 4 ilustra la arquitectura de InceptionV3 utilizada en este estudio.

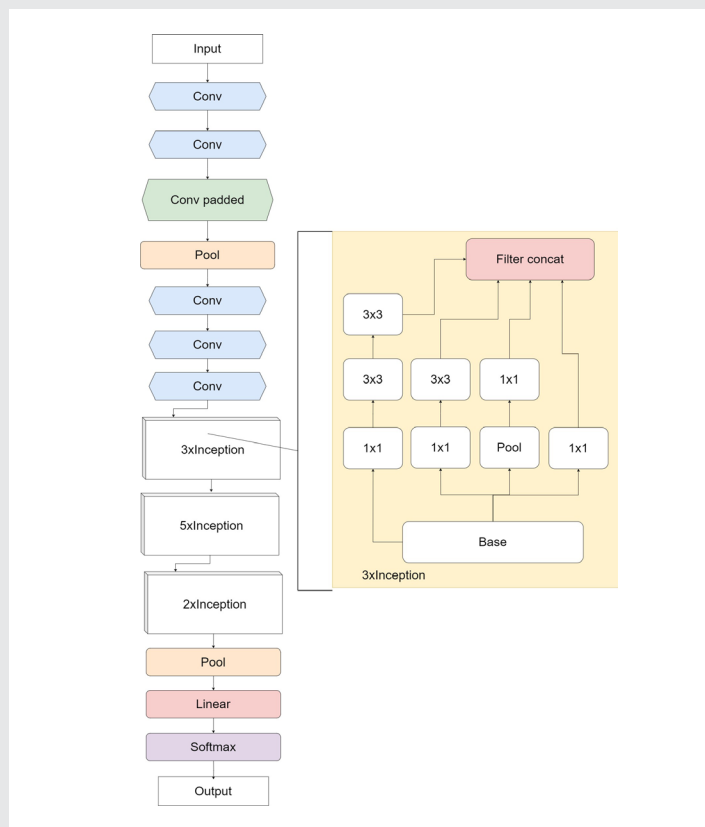


Figura 4: Arquitectura de InceptionV3.

Fuente: Elaboración propia.

Configuración de entorno y entrenamiento

Los modelos fueron entrenados en un entorno Anaconda, utilizando Python 3.9.21 en un equipo computacional con 32 GB de RAM, 5 GB de memoria de intercambio y 5 GB de almacenamiento libre. Aunque no se realizaron modificaciones arquitectónicas, el desempeño de los modelos depende en gran medida de contar con imágenes bien etiquetadas y preprocesadas, lo que resalta la importancia de procesos robustos de curación de datos. Se empleó un tamaño de lote de 16 y una configuración de early stopping con una paciencia de tres épocas para evitar sobreajuste, permitiendo detener el entrenamiento cuando la pérdida de validación no mejoraba. El número máximo de épocas fue de 30, y se utilizaron generadores de datos para alimentar los conjuntos de entrenamiento y validación de forma eficiente. Los modelos fueron optimizados con Adam (Adaptive Moment Estimation), configurando una tasa de aprendizaje de 0.0001. La función de pérdida utilizada fue la sparse categorical crossentropy, adecuada

para tareas de clasificación multiclase con etiquetas enteras. En la ecuación 1 se presenta dicha función:

$$L = -\frac{1}{N} \sum_{n=1}^N \log(\hat{p}_{y^{(n)}}^{(n)}), \quad \text{Ec. (1)}$$

donde L es la pérdida promedio de la Sparse Categorical Crossentropy, N es el número total de muestras en el conjunto de datos, $y^{(n)}$ es la etiqueta verdadera (entera) de la muestra n , con $y^{(n)} \in \{0, 1, \dots, C - 1\}$, $\hat{p}_{y^{(n)}}^{(n)}$ es la probabilidad predicha por modelo para la clase correcta $y^{(n)}$ en la muestra n , y C el número total de clases posibles.

RESULTADOS

Los resultados preliminares muestran diferencias relevantes entre ambas arquitecturas. Es importante señalar que la calidad y condiciones de captura de las imágenes influyen parcialmente en el desempeño de los modelos, particularmente en clases con síntomas leves o condiciones de iluminación desiguales. En la Figura 5(a) y 5(b) se presentan las curvas de exactitud y pérdida correspondientes a las fases de entrenamiento y validación del modelo MobileNetV3. De forma similar, la Figura 6(a) y 6(b) muestra las mismas métricas para el modelo InceptionV3. En cuanto al aumento de datos, se realizaron experimentos de validación que confirmaron que no se produjo un sobreajuste severo. El uso de early stopping y la división estratificada de los conjuntos (70% entrenamiento, 15% validación, 15% prueba) contribuyó a mitigar este riesgo. Sin embargo, debe considerarse que, en algunas clases, la proporción de imágenes sintéticas fue elevada, lo que podría afectar la capacidad de generalización en escenarios totalmente nuevos.

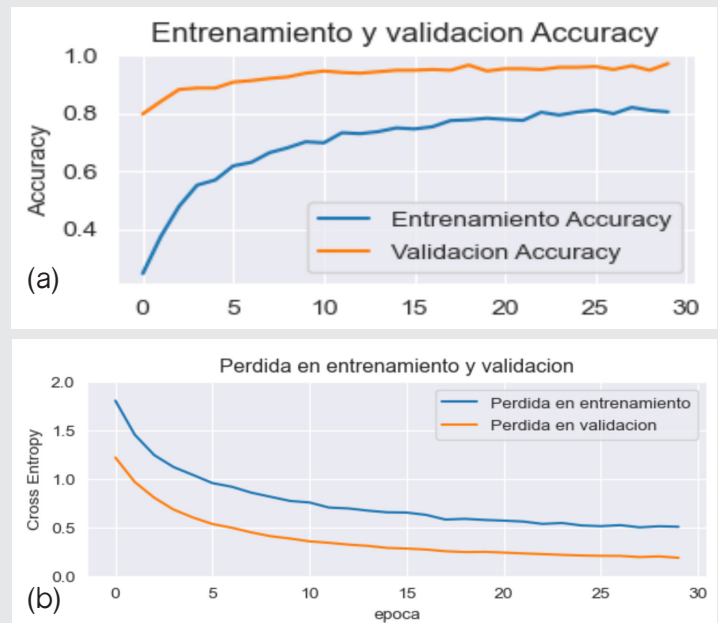


Figura 5. Curvas de exactitud y pérdidas a lo largo de épocas de MobileNetV3. (a) Precisión de entrenamiento y validación, (b) Pérdida de entrenamiento y validación.

Fuente: Elaboración propia.

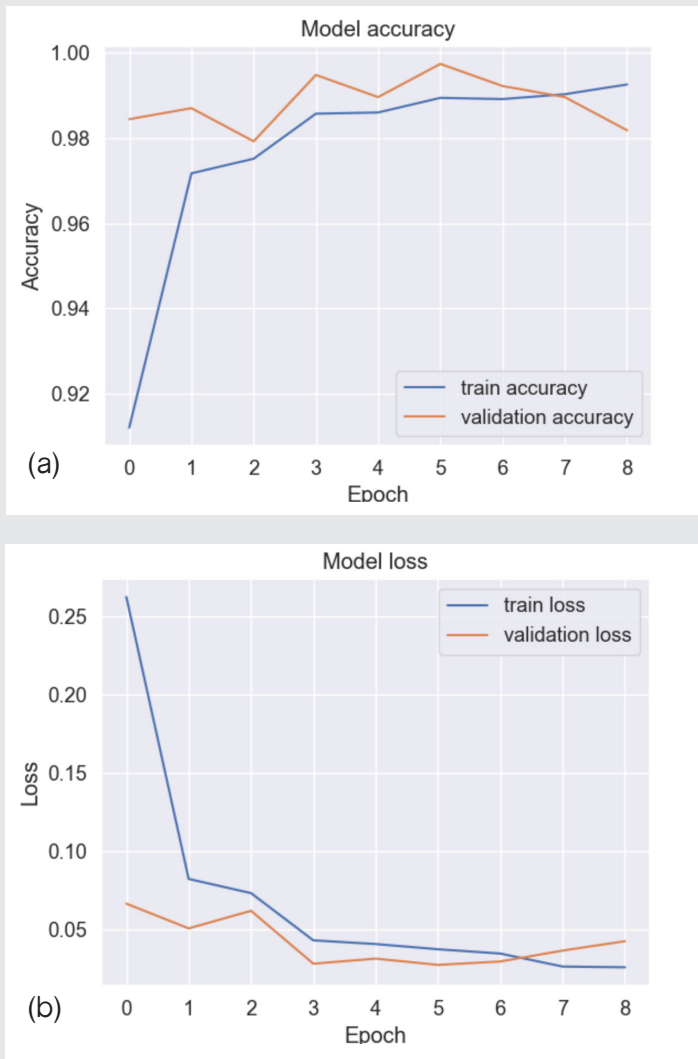


Figura 6: Curvas de precisión y pérdidas a lo largo de épocas de InceptionV3. (a) Precisión de entrenamiento y validación (b) Pérdida de entrenamiento y validación. Fuente: Elaboración propia.

InceptionV3

La evaluación final del modelo InceptionV3 se llevó a cabo utilizando el conjunto de validación, obteniendo una precisión (ecuación 2) del 98.18%, lo que indica un alto grado de exactitud en sus predicciones. Esto demuestra una excelente capacidad de generalización y un mínimo sobreajuste por parte del modelo. Asimismo, las métricas de recuperación (ecuación 3) y F1-Score (ecuación 4) oscilaron entre el 96% y 98% para las distintas clases, lo cual confirma la eficacia del modelo en la clasificación de enfermedades. La precisión evalúa qué proporción de las predicciones positivas fue realmente correcta; la recuperación mide cuántos de los casos positivos reales fueron correctamente identificados por el modelo; y el F1-score combina ambas métricas en un solo valor que representa el equilibrio entre precisión y recuperación, especialmente útil cuando se presentan clases desbalanceadas. En la Figura 7 se muestra la matriz de confusión, una herramienta que

permite visualizar el desempeño del modelo en términos de verdaderos positivos, falsos negativos, falsos positivos y verdaderos negativos. El modelo InceptionV3 demostró un desempeño sobresaliente, clasificando correctamente los casos con errores mínimos.

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

$$\text{Recall} = \frac{TP}{TP + FN} \quad \text{Ec. (3)}$$

$$\text{F1-Score} = \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \quad \text{Ec. (4)}$$

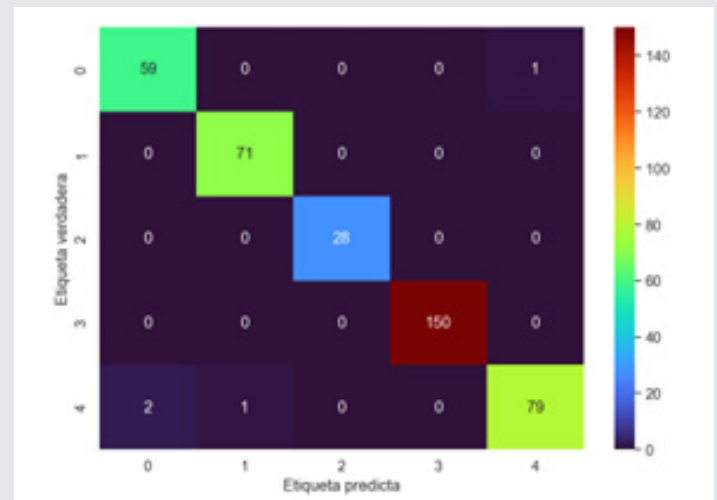


Figura 7: Matriz de confusión del InceptionV3, en donde 0: Angular leaf scorch, 1: Cenicilla, 2: Blossom Blight, 3: Saludable, 4: leaf scorch

Fuente: Elaboración propia

La Tabla 2 muestra que el modelo InceptionV3 obtuvo un rendimiento consistentemente alto en todas las clases evaluadas, destacando blossom blight y saludable, que alcanzaron métricas perfectas. Otras clases como cenicilla y leaf scorch lograron F1-scores cercanos al 97%, mientras que angular leaf scorch obtuvo la puntuación más baja con un respetable 97%.

Tabla 2. Resultados por clase para InceptionV3.

Clase	Precisión	Recall	F1-Score
Angular leaf scorch	96%	98%	97%
Cenicilla	98%	100%	99%
Blossom blight	100%	100%	100%
Saludable	100%	100%	100%
Leaf scorch	98%	96%	97%

Fuente: Elaboración propia.

MobileNetV3-Small

La evaluación final de MobileNetV3- Small se realizó con el conjunto de validación y arrojó una exactitud del 82% durante el entrenamiento. Este modelo arrojó porcentajes inferiores para puntuación F1: 86%, recuperación: 94% y precisión: 91%. En la Figura 8 se demuestra un mayor error de clasificación especialmente en la clase de leaf scorch, seguido de angular leaf scorch.

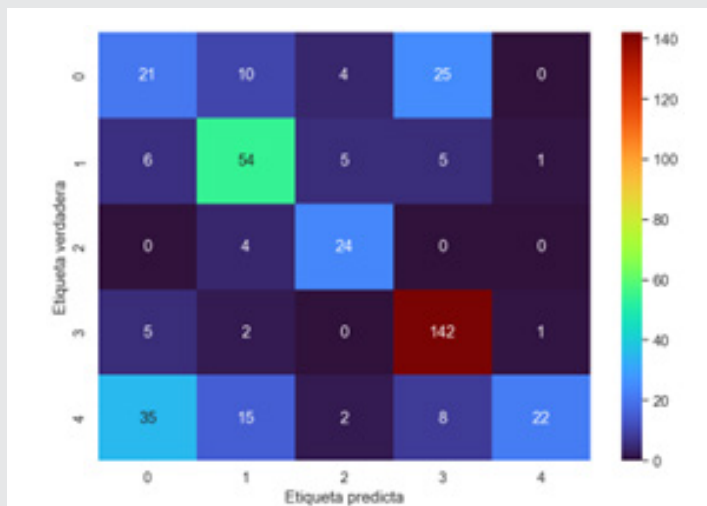


Figura 8: Matriz de confusión del MobileNetV3-Small, en donde 0: Angular leaf scorch, 1: Cenicilla, 2: Blossom Blight, 3: Saludable, 4: leaf scorch.

Fuente: Elaboración propia.

La Tabla 3 revela que MobileNetV3-Small presentó una mayor variabilidad en el rendimiento entre clases. La clase Saludable mostró una disminución notable en comparación con InceptionV3, y angular leaf scorch obtuvo el F1-score más bajo de todos los casos evaluados, con 84%. Esta caída podría explicarse por la similitud visual con otras enfermedades y la menor capacidad de representación del modelo.

Tabla 3. Resultados por clase para MobileNetV3-Small.

Clase	Precisión	Recall	F1-Score
Angular leaf scorch	90%	93%	85%
Cenicilla	92%	94%	87%
Blossom blight	93%	96%	88%
Saludable	89%	95%	86%
Leaf scorch	91%	92%	84%

Fuente: Elaboración propia.

Estos resultados preliminares reflejan las características de cada arquitectura: InceptionV3, más profunda y compleja, logra una mejor generalización, mientras que MobileNetV3-Small, optimizado para eficiencia, sacrifica precisión en clases visualmente similares. La Tabla

4 resume los promedios generales de precisión, recuperación y F1-score, permitiendo comparar globalmente el desempeño de ambos modelos.

Tabla 4. Resultados globales entre InceptionV3 y MobileNetV3-Small.

Modelo	Exactitud	Precisión	Recall	F1-Score
InceptionV3	98.18%	98%	98%	98.7%
MobileNetV3-Small	82%	91%	94%	86%

Fuente: Elaboración propia.

CONCLUSIONES

El presente estudio demuestra que el uso de aprendizaje por transferencia con arquitecturas como InceptionV3 y MobileNetV3-Small es una estrategia eficaz para la clasificación automática de enfermedades en hojas de fresa. Los resultados obtenidos evidencian que es posible desarrollar modelos precisos, eficientes y adaptables, adecuados para diferentes contextos agrícolas según las restricciones técnicas del entorno. En este trabajo, las imágenes provinieron deliberadamente de dispositivos móviles, con variaciones en ángulos de captura y en condiciones de iluminación natural y artificial, con el objetivo de simular un entorno real de uso. Un aspecto relevante a explorar en trabajos posteriores es el impacto específico de estos parámetros en los resultados, ya que conocer esta influencia permitiría anticipar la factibilidad de integrar los modelos en dispositivos concretos y bajo condiciones determinadas de operación.

No obstante, el estudio presenta algunas limitaciones: el conjunto de datos, aunque representativo, sigue siendo relativamente pequeño; la dependencia de la calidad y el ángulo de captura puede afectar la robustez en escenarios no controlados; y el aumento intensivo de datos en ciertas clases plantea la necesidad de validar la generalización en conjuntos externos no aumentados.

Como trabajo futuro, se plantea la implementación de estos modelos en dispositivos móviles, aplicando técnicas de cuantización y poda de parámetros para reducir su tamaño. Asimismo, se propone ampliar el dataset con nuevas imágenes de enfermedades capturadas en condiciones diversas y emplear técnicas de aumento avanzadas, como modificaciones cromáticas o simulaciones de ruido de campo, para mejorar la capacidad de generalización.

AGRADECIMIENTOS

Esta investigación fue financiada por el Tecnológico Nacional de México (TecNM) bajo el número de proyecto [22298.25-P].

BIBLIOGRAFÍA

[1] Rocha-Ibarra, J.E., Mireles-Arriaga, A.I., Ruiz-Nieto, J. E., & Maki-Díaz, G. (2024). Production and ex-

port of berries in Mexico's agricultural development: A study of competitive advantage. *Agrociencia*, 1–14.

[2] Rehman, S., Attique Khan, M., Alhaisoni, M., Armghan, A., Alenezi, F., Alqahtani, A., Vesal, K., & Nam, Y. (2023). Fruit leaf diseases classification: A hierarchical deep learning framework. *Computers, Materials & Continua*, 75(1), 1179–1194.

[3] Ngugi, L. C., Abelwahab, M., & Abo-Zahhad, M. (2021). Recent advances in image processing techniques for automated leaf pest and disease recognition – A review. *Information Processing in Agriculture*, 8(1), 27–51.

[4] Biswas, M., Ray, M., Jahan, I., Khan, S., Ahmad Saad, S., & Bharman, P. (2022). Deep learning in agriculture: A review. *Asian Journal of Research in Computer Science*, 28–47.

[5] Galván García, A. C., Sánchez, M. G., & Rodríguez González, A. J. (2024). Desarrollo e implementación de un sistema de monitoreo agrícola basado en micromódulos: Un enfoque integrado para la gestión de plantaciones. *Ingeniantes*, 3(1), 104–112.

[6] Afzaal, U., Bhattarai, B., Pandeya, Y. R., & Lee, J. (2021). An instance segmentation model for strawberry diseases based on Mask R-CNN. *Sensors*, 21(19), 6565.

[7] Cao, Z., Li, J., Fang, L., Li, Z., Yang, H., & Dong, G. (2025). Research on efficient classification algorithm for coal and gangue based on improved MobileNetV3-small. *International Journal of Coal Preparation and Utilization*, 45(2), 437–462.

[8] Pramudhita, D. A., Azzahra, F., Arfat, I. K., Magdalena, R., & Saidah, S. (2023). Strawberry plant diseases classification using CNN based on MobileNetV3-Large and EfficientNet-B0 architecture. *Jurnal Ilmiah Teknik Elektro Komputer dan Informatika*, 9(3), 522–534.

[9] Ochoa-Ornelas, R., Gudiño-Ochoa, A., García-Rodríguez, J. A., & Uribe-Toscano, S. (2024). Lung and colon cancer detection with InceptionResNetV2: A transfer learning approach. *Journal of Research and Development*, 10(25).

[10] Li, L., Zhang, S., & Wang, B. (2021). Plant disease detection and classification by deep learning—A review. *IEEE Access*, 9, 56683–56698.

