

Evaluación de Arquitecturas de Redes Neuronales Convolucionales para la Detección de Enfermedades en las Hojas de la Papa

Evaluation of Convolutional Neural Network Architectures for Potato Leaf Disease Detection



Reyna Isabel Chipa Dávalos

Universidad Nacional Micaela Bastidas de Apurímac
<https://orcid.org/0009-0007-7475-9801>
172143@unamba.edu.pe

Sindy Malu Huilca Elguera

Universidad Nacional Micaela Bastidas de Apurímac
<https://orcid.org/0009-0008-4812-7322>
172153@unamba.edu.pe

Mario Aquino Cruz

Universidad Nacional Micaela Bastidas de Apurímac
<https://orcid.org/0000-0002-2552-5669>
maquino@unamba.edu.pe

Universidad Nacional Micaela Bastidas de Apurímac – Perú
 Riqchary, revista de investigación en ciencia y tecnología
[ISSN: 2810-8124 \(en línea\) / ISSN: 2706-543x](https://doi.org/10.57166/riqchary/v8.n1.2026)
 Vol. 8 Núm. 1 (2026) - Publicado: 24/08/2026
doi.org/10.57166/riqchary/v8.n1.2026
 Páginas: 89- 98
 Recibido 17/06/2026 ; Aceptado 03/08/2026
<https://doi.org/10.57166/riqchary/v8.n1.2026.11>

Resumen— La papa, uno de los principales cultivos alimentarios básicos a nivel mundial, es altamente susceptible a diversas enfermedades foliares que afectan significativamente su productividad y representan una seria amenaza para la seguridad alimentaria, contribuyendo así a pérdidas económicas y afectando los ingresos de los agricultores. Por lo tanto, la detección temprana y precisa es esencial. Los métodos de detección convencionales se basan principalmente en la observación manual, que consume mucho tiempo y requiere personal especializado. En este estudio, se evaluaron cinco arquitecturas de redes neuronales convolucionales (CNN) para la clasificación automática de enfermedades foliares de la papa, incluyendo modelos preentrenados (ResNet50, MobileNet y VGG16) y modelos entrenados desde cero (AlexNet y LeNet-5). El conjunto de datos se construyó integrando y seleccionando imágenes de repositorios públicos disponibles en Kaggle alcanzando un total de 6691 imágenes, que comprenden cinco clases: tizón temprano, tizón tardío, virus del enrollamiento de la hoja de la papa (PLRV), virus del mosaico (PVY) y hojas sanas. Se realizaron múltiples experimentos variando hiperparámetros como el tamaño del lote, los optimizadores y el número de épocas de entrenamiento. Los resultados muestran que VGG16 obtuvo el mejor rendimiento, con una precisión del 99,87 %, superando a las demás arquitecturas. Además, se desarrolló una aplicación móvil basada en el modelo óptimo para la detección en tiempo real. Estos hallazgos demuestran el potencial del aprendizaje profundo para sistemas de diagnóstico agrícola inteligentes y escalables.

Palabras clave: Cultivos de papa, Detección de enfermedades de las plantas, Procesamiento de imágenes, Redes neuronales convolucionales (CNN)

Abstract— Potato, one of the world's most important staple food crops, is highly susceptible to various foliar diseases that significantly affect its productivity and pose a serious threat to food security, thereby contributing to economic losses and impacting farmers' income. Therefore, early and accurate detection is essential. Conventional detection methods rely primarily on manual observation, which is time-consuming and requires specialized personnel. In this study, five convolutional neural network (CNN) architectures were evaluated for the automatic classification of potato leaf diseases, including pretrained models (ResNet50, MobileNet, and VGG16) and models trained from scratch (AlexNet and LeNet-5). The dataset was constructed by integrating and selecting images from publicly available Kaggle repositories, resulting in a total of 6,691 images distributed across five classes: early blight, late blight, potato leafroll virus (PLRV), mosaic virus (PVY), and healthy leaves. Multiple experiments were conducted by varying hyperparameters such as batch size, optimizers, and the number of training epochs. The results show that VGG16 achieved the best performance, with an accuracy of 99.87%, outperforming the other architectures. Additionally, a mobile application based on the optimal model was developed for real-time detection. These findings demonstrate the potential of deep learning for intelligent and scalable agricultural diagnostic systems.

Keywords: Convolutional Neural Networks (CNN), Image processing, Plant disease detection, Potato crops

1 INTRODUCCIÓN

Las enfermedades foliares en los cultivos representan una de las principales amenazas para la producción agrícola mundial, planteando serios desafíos para la seguridad alimentaria y la estabilidad económica. En este contexto, el cultivo de papa destaca como uno de los cultivos más importantes a nivel mundial, no solo por su alto valor nutricional, sino también por su significativo impacto socioeconómico, ya que constituye una fuente primaria de ingresos para millones de agricultores. Sin embargo, los cultivos de papa están continuamente expuestos a una amplia gama de enfermedades foliares que reducen sustancialmente tanto el rendimiento como la calidad [1].

Según la Organización de las Naciones Unidas para la Alimentación y la Agricultura (FAO), hasta el 40% de los cultivos alimentarios mundiales se pierden anualmente debido a plagas y enfermedades de las plantas. Estas pérdidas tienen efectos devastadores en la productividad y los ingresos, particularmente en las comunidades vulnerables cuyos medios de subsistencia dependen en gran medida de la agricultura [2]. Informes del norte de la India indican que, durante la temporada 2023-2024, las enfermedades foliares dañaron entre el 8% y el 10% de los cultivos de papa, y algunas regiones experimentaron pérdidas locales que superaron el 30% de la producción esperada [3]. De manera similar, en algunas zonas de España, como Algaida (Islas Baleares), la producción de papa se redujo en aproximadamente un 90 % durante un otoño reciente debido a brotes severos de enfermedades foliares, lo que representa un caso extremo de impacto económico y productivo [4].

Este problema también representa una seria preocupación en países latinoamericanos como Perú. Un estudio realizado por MIDAGRI informa que la brecha nacional de rendimiento de papa entre 1997 y 2023 alcanzó las 12,9 toneladas por hectárea (t/ha), lo que corresponde al 45 % del rendimiento alcanzable para los productores agrícolas. En términos prácticos, los agricultores peruanos producen, en promedio, solo el 55 % de su rendimiento potencial en condiciones óptimas de cultivo, lo que implica que se pierde casi la mitad de la producción potencial [5]. Estas cifras resaltan la necesidad urgente de implementar estrategias preventivas que permitan la detección temprana de los síntomas de enfermedades foliares y una mitigación más eficiente de las pérdidas.

Para abordar este problema, los recientes avances en inteligencia artificial, en particular las redes neuronales convolucionales (CNN), han surgido como una solución prometedora e innovadora para la detección automática de enfermedades en las hojas de papa. Las CNN se componen de capas convolucionales, de agrupación y totalmente conectadas, lo que les permite aprender automáticamente características discriminativas directamente a partir de datos de imagen. Gracias a su capacidad para capturar patrones espaciales complejos, las CNN se han adaptado ampliamente en tareas de análisis de imágenes y visión artificial [6].

Este estudio tiene como objetivo identificar la arquitectura de red neuronal convolucional (CNN) más precisa para la detección de enfermedades de la papa, centrándose en las arquitecturas AlexNet [7], MobileNet [8], VGG16 [9], ResNet50 [10] y LeNet-5 [11]. Para su evaluación, los modelos se clasificaron en dos grupos: modelos preentrenados (MobileNet, VGG16 y ResNet50) y modelos entrenados desde cero (AlexNet y LeNet-5). Las arquitecturas seleccionadas representan una amplia gama de diseños de CNN, desde modelos relativamente simples y computacionalmente eficientes hasta arquitecturas más sofisticadas y especializadas.

Como toda investigación científica, este estudio presenta ciertas limitaciones. Aunque los modelos evaluados demostraron un rendimiento sólido en condiciones experimentales controladas, su eficacia puede verse afectada por factores externos como variaciones en la iluminación, ruido visual, complejidad del fondo y cambios en la orientación o el estado de las hojas. Estos factores pueden reducir la precisión de la clasificación en entornos agrícolas reales. Además, la disponibilidad de recursos computacionales constituyó una limitación importante durante el proceso de entrenamiento, ya que el manejo de grandes volúmenes de datos y la obtención de tiempos de entrenamiento razonables requirieron el uso de una cuenta de Google Colab Pro+.

Mediante esta evaluación, el estudio busca contribuir al desarrollo de soluciones tecnológicas mediante un análisis comparativo de cinco arquitecturas CNN para identificar sus fortalezas y limitaciones en la detección de enfermedades foliares en la papa. Como contribución práctica, este trabajo también propone el desarrollo de una aplicación móvil que integre la arquitectura CNN de mejor rendimiento, lo que permitirá a los agricultores acceder a una herramienta eficiente y fácil de usar para la detección temprana de enfermedades.

1.1 Trabajos Relacionados

Estudios recientes han explorado ampliamente el uso de la inteligencia artificial para la detección y prevención de enfermedades en las hojas de la papa. Tales como el estudio de Angel y Bryan donde desarrollaron un prototipo de software basado en técnicas de inteligencia artificial, proponiendo un novedoso modelo de red neuronal convolucional (CNN) de 54 capas. El proceso de entrenamiento se configuró utilizando el optimizador Adam y se llevó a cabo durante 50 épocas, con un conjunto de datos de 2501 imágenes distribuidas en tizón tardío, tizón temprano, hoja con pulguilla y hojas sanas obtenidas de diversas fuentes, como el conjunto de datos de PlantVillage del repositorio Kaggle y un conjunto de datos propio. Además, los autores compararon su modelo propuesto con cinco arquitecturas basadas en transfer learning como: ResNet50, VGG19, InceptionV3, EfficientNetB0 y MobileNet. El modelo implementado demostró un rendimiento competitivo, alcanzando una precisión del 95,34 %, y se integró con éxito en un prototipo de aplicación móvil [12].

Gisell Natalia Cristiano Muñoz diseñó un sistema ba-

sado en redes neuronales convolucionales (CNN) preentrenadas y propuso una metodología de tres fases para el análisis y la clasificación de enfermedades de la papa a partir imágenes recopiladas en escenarios reales. En la primera fase, se emplea transfer learning y fine-tuning para entrenar tres arquitecturas CNN preentrenadas (VGG16, ResNet50 e InceptionV3) utilizando subimágenes de tamaño 224x224 píxeles, en la segunda fase, se utilizó la CNN con mejor desempeño para generar mapas de calor que resaltan las regiones de interés relacionados con las enfermedades en las hojas. Finalmente, en la tercera fase, se aplicaron nuevamente transfer learning y fine-tuning para reentrenar los tres modelos CNN utilizando los mapas de calor generados como entrada, obteniendo como el mejor modelo a InceptionV3 con una precisión del 99,67%. [13].

Hritwik y colegas [14] evaluaron el rendimiento de modelos de redes neuronales convolucionales (CNN como VGG19, DenseNet121 y ResNet50, para la identificación y clasificación precisa de las enfermedades de las hojas de papa. El estudio realizó un análisis experimental en el que se evaluaron las métricas de rendimiento de cada modelo, para lo cual se utilizó un conjunto de datos recopilado directamente del campo, que contenía diversas muestras de hojas sanas y hojas afectadas con tizón temprano y tizón tardío. Los resultados demostraron que los tres modelos CNN lograron un alto rendimiento, siendo VGG19 con el mejor desempeño, obteniendo una precisión del 98,77%.

William y sus colegas [15] propusieron un nuevo modelo de red neuronal convolucional (CNN) el cual consta de 11 capas, diseñado para procesar imágenes de entrada de 227×227 píxeles e inspirado en la arquitectura secuencial de VGG16. El objetivo del estudio fue analizar el rendimiento de la arquitectura propuesta, que se evaluó mediante métricas de rendimiento estándar y se comparó con arquitecturas conocidas como AlexNet, ZFNet, VGG16 y VGG19. Para ello, los autores emplearon el conjunto de datos PlantVillage, obtenido de Kaggle. El modelo propuesto alcanzó un rendimiento cercano a los modelos de mejor desempeño en las pruebas, logrando una precisión del 91%, mientras que VGG16 y AlexNet lograron precisiones del 96% y 94%, respectivamente.

2 MÉTODO

2.1 Tipo y nivel de investigación

Este trabajo corresponde a una investigación aplicada con un diseño experimental. Se evaluó el efecto de diferentes algoritmos de aprendizaje automático y diversas configuraciones de hiperparámetros sobre el desempeño del modelo de clasificación.

2.2 Conjunto de datos

El conjunto de datos utilizado en este estudio se construyó mediante la integración de imágenes obtenidas de repositorios públicos en Kaggle [16] [17] [18] [19] [20], con el objetivo de crear un conjunto de datos robusto para la

clasificación de enfermedades de las hojas de papa. Debido a la heterogeneidad de las fuentes, se llevó a cabo un exhaustivo proceso de curación de datos. Las imágenes conservadas mantuvieron las etiquetas originales proporcionadas por los autores de cada conjunto de datos, por lo que no fue necesario realizar una validación de contenido mediante juicio de expertos. Dado que las imágenes provienen de repositorios públicos, no están asociadas a una ubicación geográfica específica. No obstante, el proceso de integración, curación y entrenamiento de los modelos se llevó a cabo en la ciudad de Abancay, departamento de Apurímac, Perú, durante el año 2025.

Para la construcción del conjunto de datos, se establecieron criterios de inclusión destinados para garantizar la calidad y uniformidad de las imágenes utilizadas en el entrenamiento de los modelos. Se seleccionaron imágenes digitales en formato JPG correspondientes a hojas de papa, asegurando una resolución adecuada para conservar características discriminativas clave, como la textura foliar, patrones de coloración y manifestaciones visuales asociadas a enfermedades. Posteriormente, se aplicó un proceso de limpieza y depuración en el que se descartaron imágenes borrosas, iluminación inadecuada, ruido excesivo, etiquetas incorrectas, duplicados o aquellas que no correspondían a hojas de papa. Finalmente, mediante una revisión manual, se verificó la coherencia visual y semántica de cada categoría, conformando un conjunto de datos final estructurado en cinco clases: tizón temprano, tizón tardío, virus del enrollamiento de hoja (PLRV), virus del mosaico (PVY) y hojas sanas., como se muestra en la figura 1.

El conjunto de datos unificado recopilado a partir del repositorio público de Kaggle, comprende un total de 6691 imágenes distribuidas de la siguiente manera: 1500 imágenes de hojas sanas, 1500 imágenes de tizón tardío, 1500 imágenes de tizón temprano, 1054 imágenes del virus del enrollamiento de la hoja y 1137 imágenes del virus del mosaico. Dado que se utilizaron todas las imágenes disponibles del conjunto de datos, la población y la muestra del análisis incluyeron las 6691 imágenes.

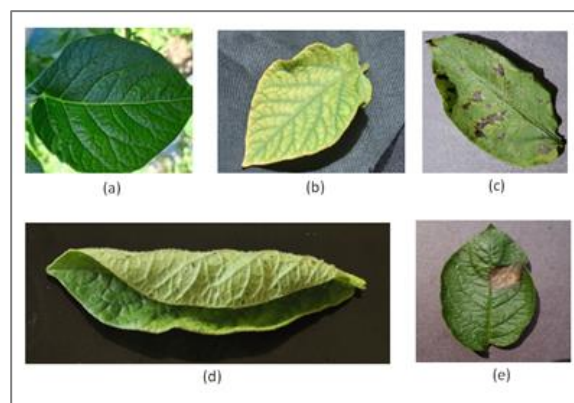


Fig. 1. Imágenes de muestra de las cinco clases en el conjunto de datos: (a) hoja sana, (b) virus del mosaico, (c) tizón temprano, (d) virus del enrollamiento de la hoja y (e) tizón tardío.

2.3 Procedimiento

El estudio comenzó con el preprocesamiento de un conjunto de datos que contiene imágenes de hojas de papa, clasificadas en hojas sanas y diversas clases de enfermedades, como se muestra en el diagrama metodológico en la figura 2. Todas las imágenes se redimensionaron a una resolución uniforme de 224×224 píxeles y se normalizaron para facilitar el aprendizaje del modelo. A continuación, el conjunto de datos se dividió automáticamente en subconjuntos de entrenamiento, validación y prueba mediante una estrategia de división estratificada. Específicamente, se empleó la función “train_test_split” de la biblioteca Sci-kit-learn para asignar el 70 % de las imágenes al entrenamiento, el 15 % a la validación y el 15 % a las pruebas. Esta división se seleccionó porque proporciona un equilibrio adecuado entre el aprendizaje y la evaluación del modelo.

Además, no se aplicó validación cruzada, ya que la partición independiente permitió ajustar y evaluar los modelos sin fuga de información. Para mejorar el rendimiento del modelo y mitigar el sobreajuste durante la fase de entrenamiento, se aplicaron técnicas de aumento de datos, como rotaciones, volteo horizontal, zoom y traslaciones, exclusivamente al conjunto de entrenamiento. Estas técnicas tienen como objetivo mejorar la capacidad de generalización de los modelos y reducir el sobreajuste.

Posteriormente, las arquitecturas CNN MobileNet, ResNet50, VGG16, AlexNet y LeNet se entrenaron utilizando dos enfoques de entrenamiento distintos y varias configuraciones de hiperparámetros, como se presenta en las Tablas 1 y 2, con el fin de realizar un análisis comparativo exhaustivo.

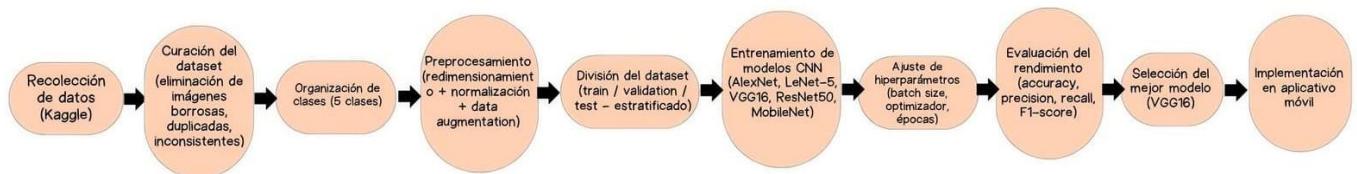


Fig. 2. Diagrama de flujo metodológico.

2.3.1 Modelos entrenados desde cero

Las arquitecturas AlexNet y LeNet se entrenaron inicialmente desde cero, lo que significa que no se incorporó ningún conocimiento previo y todos los pesos de la red se inicializaron aleatoriamente. Este enfoque permitió evaluar la capacidad de cada modelo para aprender representaciones discriminativas exclusivamente a partir del conjunto de datos disponible.

2.3.2 Modelos preentrenados

Además, las arquitecturas VGG16, ResNet50 y MobileNet se entrenaron utilizando pesos preentrenados en el conjunto de datos ImageNet. Por el contrario, AlexNet y LeNet-5 se excluyeron de este enfoque debido a la ausencia de pesos preentrenados estándar e incompatibilidades arquitectónicas. El proceso de aprendizaje por transferencia (transfer learning) se llevó a cabo en dos fases:

a) Fase 1 (Feature Extractor)

Las capas convolucionales del modelo base se mantuvieron congeladas y solo se entrenaron las capas de clasificación finales.

b) Fase 2 (Fine-Tuning)

Las últimas capas convolucionales se descongelaron selectivamente de forma controlada, y el modelo se ajustó con precisión utilizando una tasa de aprendizaje más baja para refinar las características aprendidas hacia el dominio específico de las enfermedades de las hojas de la papa.

Todos los modelos se entrenaron utilizando los optimizadores Adam y RMSprop, con funciones de pérdida adecuadas para la clasificación multiclase. En concreto, se aplicó la entropía cruzada categórica a los modelos entre-

nados desde cero, mientras que se utilizó la entropía cruzada categórica dispersa para los modelos preentrenados. Además, para mitigar el impacto del desequilibrio de clases, los pesos de las clases se calcularon automáticamente a partir del conjunto de datos de entrenamiento y se incorporaron durante el proceso de entrenamiento.

TABLA 1
Resumen de las configuraciones de los modelos CNN preentrenado

Parámetro	VGG16	ResNet50	MobilNet
Tipo de modelo	CNN preentrenada	CNN preentrenada	CNN preentrenada
Pesos iniciales	ImagenNet	ImagenNet	ImagenNet
Tamaño de imagen	$224 \times 224 \times 3$	$224 \times 224 \times 3$	$224 \times 224 \times 3$
Image Batch	16, 32	16, 32	16, 32
Estrategia de entrenamiento	Transfer Learning	Transfer Learning	Transfer Learning
Fase 1	Feature extractor	Feature extractor	Feature extractor
Capas entrenable (fase 1)	Solo clasificador	Solo clasificador	Solo clasificador
Épocas (fase 1)	10, 15	10, 15	10, 15
Learning rate (fase 1)	$1e-4$	$1e-4$	$1e-4$
Fase 2	Fine-tuning	Fine-tuning	Fine-tuning
Capas ajustadas (fase 2)	Últimas 30 capas	conv5	Últimas 25 capas

Parámetro	VGG16	ResNet50	MobilNet
Batch Normalization	Congelado	Congelado	Congelado
Épocas (Fase 2)	10, 15	10, 15	10, 15
Learning rate (fase 2)	1e-5	1e-5	1e-5
Optimizador	Adam, RMSprop	Adam, RMSprop	Adam, RMSprop
Métrica función de pérdida	Sparse Categorical Crossentropy	Sparse Categorical Crossentropy	Sparse Categorical Crossentropy

TABLA 2
Resumen de las configuraciones de los modelos CNN entrenados desde cero

Parámetro	LeNet	AlexNet
Tipo de modelo	Desde cero	Desde cero
Pesos iniciales	Aleatorios	Aleatorios
Tamaño de imagen	224 × 224 × 3	224 × 224 × 3
Image Batch	32	32
Épocas	30	30
Learning rate (fase 1)	1e-3, 1e-4	1e-3, 1e-4
Optimizador	Adam, RMSprop	Adam, RMSprop
Métrica función de pérdida	Categorical Crossentropy	Categorical Crossentropy

Además, se implementó la técnica de Early Stopping para detener el entrenamiento cuando la pérdida de validación no mejoró durante cinco épocas consecutivas. Asimismo, se utilizó ModelCheckpoint para guardar automáticamente el mejor modelo según el rendimiento en validación. El entrenamiento y la evaluación del modelo se realizaron utilizando un acelerador GPU A100 disponible en Google Colab Pro+, lo que permitió una reducción significativa del tiempo de cálculo. Sin esta infraestructura computacional, el entrenamiento de arquitecturas profundas como VGG16 y ResNet50 habría requerido una ejecución continua durante varios días. Así mismo se utilizó el framework de TensorFlow 19 y el API de Keras 3.9.

Finalmente, el rendimiento de cada arquitectura se evaluó en el conjunto de prueba utilizando métricas estándar, como exactitud (accuracy), precisión (precision), exhaustividad (recall), puntuación F1 (F1-score) promedio macro. Este enfoque permitió una comparación justa y reproducible entre los diferentes modelos analizados. Los modelos entrenados se guardaron en formato .keras y .json para facilitar su posterior evaluación e implementación.

3 RESULTADOS

A continuación, se presentan los resultados obtenidos

a partir de la evaluación de las arquitecturas VGG16, AlexNet, ResNet50, MobileNet y LeNet para la detección de enfermedades foliares de la papa. Los resultados se organizan en las siguientes secciones:

3.1 Curvas de entrenamiento y validación

La Figura 3 y 4 presenta las curvas de precisión y pérdida respectivamente para los modelos VGG16, MobileNet y ResNet50, utilizando las mejores configuraciones de hiperparámetros identificadas en la Tabla 1. Los tres modelos muestran una rápida convergencia, caracterizada por un aumento progresivo de la precisión y una disminución constante de la función de pérdida durante el entrenamiento. Este comportamiento resalta la eficacia de aprovechar los pesos preentrenados de ImageNet para la clasificación de enfermedades en las hojas de papa.

Las curvas de entrenamiento y validación se mantienen estrechamente alineadas, lo que indica una buena capacidad de generalización y la ausencia de sobreajuste significativo. MobileNet demuestra una convergencia más rápida, alcanzando altos niveles de precisión en las primeras etapas del entrenamiento. En contraste, VGG16 y ResNet50 muestran un proceso de aprendizaje más gradual, alcanzando finalmente niveles de rendimiento estables y comparables al final del entrenamiento.

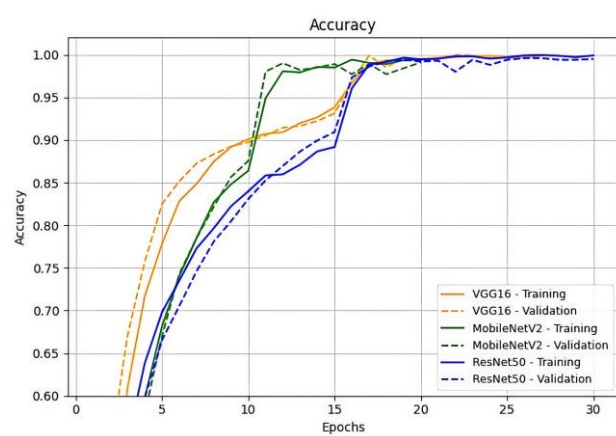


Fig. 3. Curvas de precisión durante el entrenamiento de los modelos preentrenados para la detección de enfermedades en las hojas de la papa.

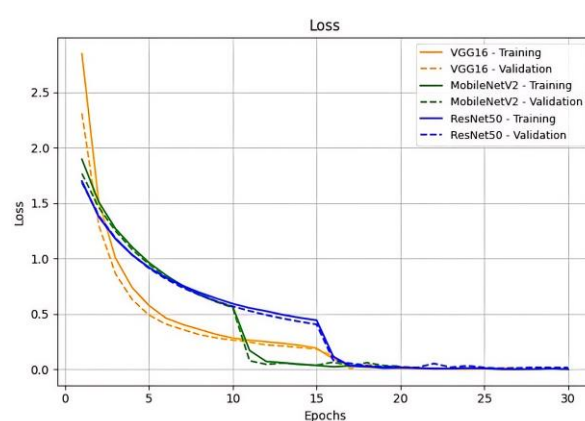


Fig. 4. Curvas de pérdida durante el entrenamiento de los modelos preentrenados para la detección de enfermedades en las hojas de la papa.

La Figura 5 y 6 presenta las curvas de precisión y pérdida de los modelos LeNet y AlexNet entrenados. LeNet muestra un aprendizaje gradual, con una mejora constante en la precisión y una reducción progresiva de la pérdida, lo que evidencia estabilidad entre las curvas de entrenamiento y validación. Este comportamiento indica un aprendizaje consistente y la ausencia de sobreajuste significativo, aunque los niveles de rendimiento finales se mantienen moderados. Por el contrario, AlexNet demuestra una convergencia más rápida, alcanzando valores de precisión más altos y niveles de pérdida más bajos. La comparación entre ambos modelos revela que, si bien LeNet ofrece un rendimiento aceptable, AlexNet logra una generalización superior debido a su mayor profundidad y capacidad mejorada para extraer características relevantes.

Estos hallazgos respaldan la idea de que entrenar redes neuronales convolucionales (CNN) desde cero es menos efectivo cuando se dispone de conjuntos de datos limitados. También justifican el uso de modelos preentrenados como una alternativa más robusta para aplicaciones de detección de enfermedades agrícolas.

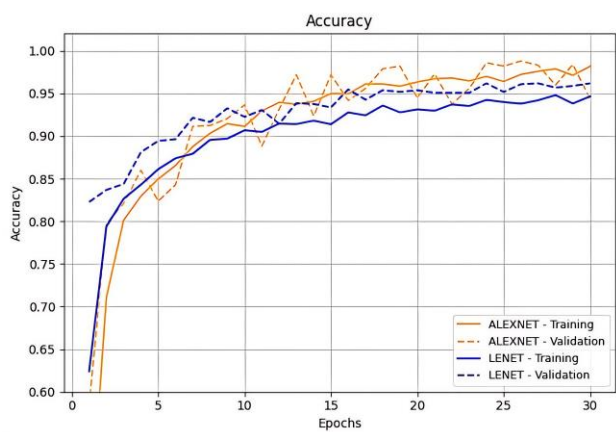


Fig. 5. Curvas de precisión durante el entrenamiento de los modelos desde cero para la detección de enfermedades en las hojas de la papa.

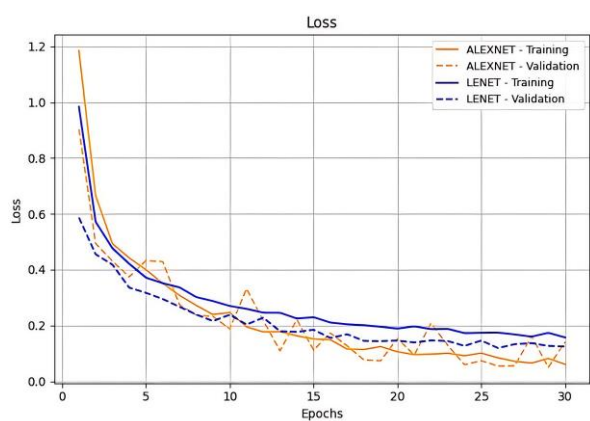


Fig. 6. Curvas de pérdida durante el entrenamiento de los modelos desde cero para la detección de enfermedades en las hojas de la papa.

3.2 Presición en la validación, el entrenamiento y las pruebas

La comparación de la precisión alcanzada en las etapas de entrenamiento, validación y prueba de los modelos preentrenados VGG16, MobileNet y ResNet50 se muestra en la figura 7; para ello se toma como referencia la configuración óptima de hiperparámetros que fue determinada experimentalmente. Los resultados demuestran que las arquitecturas evaluadas son muy efectivas para detectar enfermedades en hojas de papa; sobresale VGG16 por su exactitud y MobileNet por su eficacia en términos computacionales.

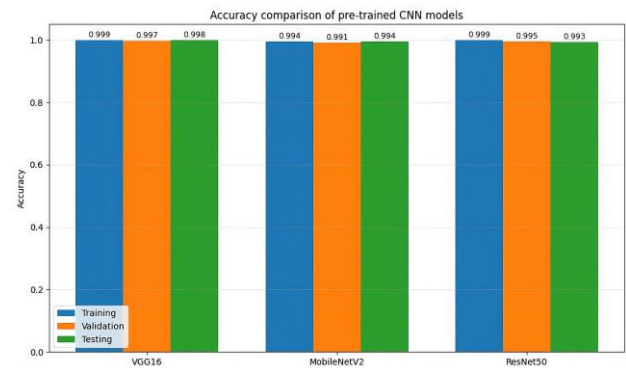


Fig. 7. Rendimiento comparativo de las arquitecturas CNN preentrenadas en las fases de entrenamiento, validación y prueba para las hojas de papa (sanas y enfermas).

La Figura 8 presenta el desempeño de los modelos LeNet y AlexNet bajo distintos optimizadores, evaluados en los conjuntos de entrenamiento, validación y prueba. En los dos modelos, las diferencias entre validación y prueba son mínimas, lo que señala una buena capacidad de generalización. En total, AlexNet tiene un desempeño superior al de LeNet.

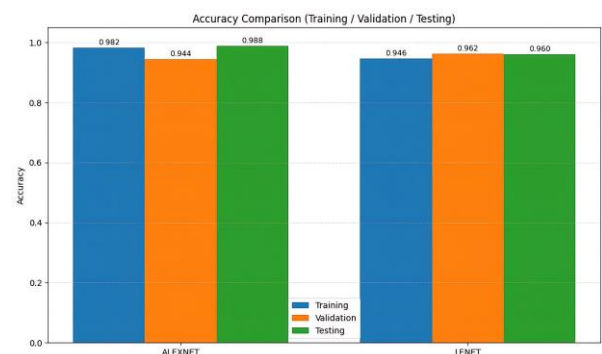


Fig. 8. Rendimiento comparativo de las arquitecturas CNN desde cero en las fases de entrenamiento, validación y prueba para las hojas de papa (sanas y enfermas).

3.3 Matriz de confusión

Las matrices de confusión de los modelos con mejor rendimiento tras la optimización de hiperparámetros revelan una alta precisión en la clasificación de las arquitecturas VGG16, MobileNet y ResNet50, como se ilustra en las figuras 9, 10 y 11, respectivamente. VGG16 logra el mejor rendimiento general, con excelentes resultados de clasificación y una mínima confusión entre las clases de virus

del mosaico y hoja sana. MobileNet también alcanza una alta precisión, con pequeñas clasificaciones erróneas, principalmente entre tizón temprano y tizón tardío. ResNet50 presenta un nivel de confusión ligeramente superior, sobre todo entre tizón temprano y tizón tardío; sin embargo, su rendimiento general sigue siendo competitivo.

En general, estos resultados confirman la discriminación eficaz de las enfermedades de las hojas de papa, destacando a VGG16 como el modelo más preciso y a MobileNet como una alternativa eficiente que ofrece un equilibrio favorable entre precisión y complejidad computacional.

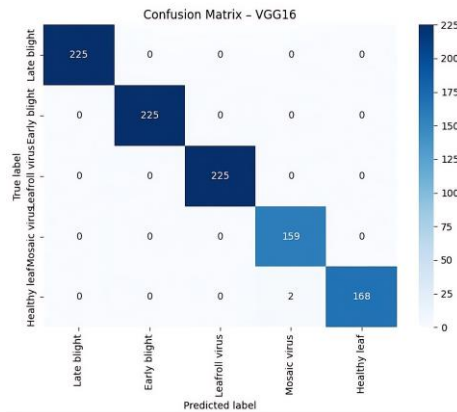


Fig. 9. Matriz de confusión de la arquitectura VGG16 para la clasificación de hojas de papa (sanas y enfermas).

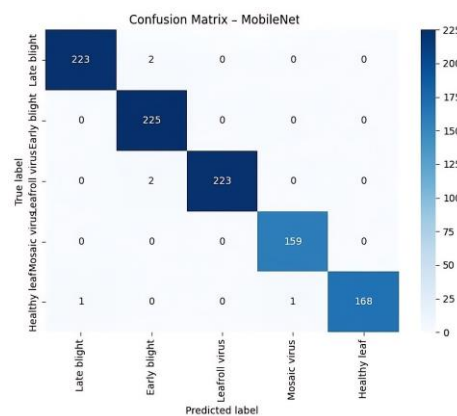


Fig. 10. Matriz de confusión de la arquitectura MobileNet para la clasificación de hojas de papa (sanas y enfermas).

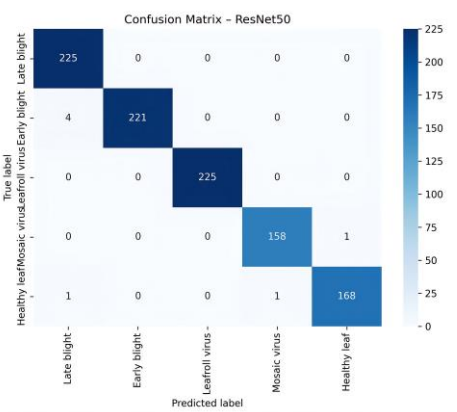


Fig. 11. Matriz de confusión de la arquitectura ResNet50 para la clasificación de hojas de papa (sanas y enfermas).

La matriz de confusión indica que el modelo AlexNet logra un rendimiento de clasificación muy preciso, como se muestra en la Figura 12, con la mayoría de las muestras correctamente ubicadas a diagonal principal. Las clasificaciones erróneas son limitadas y ocurren principalmente entre las clases de tizón temprano y tizón tardío, lo que probablemente se deba a la similitud visual de los síntomas en las hojas de papa. Además, se observan algunos errores entre las clases de virus del mosaico y hoja sana, aunque estos ocurren con una frecuencia relativamente baja. En general, AlexNet demuestra una gran capacidad para distinguir entre hojas enfermas y sanas, manteniendo una baja tasa de error entre clases.

En contraste, el modelo LeNet-5 presenta un mayor grado de confusión de clases. Se observan clasificaciones erróneas significativas entre el virus del mosaico y las hojas sanas, así como errores más frecuentes entre el virus del enrollamiento de la hoja, tizón temprano y tizón tardío, lo que indica limitaciones para capturar patrones visuales sutiles y síntomas estrechamente relacionados, como se muestra en la Figura 13. Si bien LeNet-5 logra un rendimiento satisfactorio, el análisis comparativo indica que AlexNet proporciona una discriminación de clases y robustez superiores, lo que lo convierte en el modelo más eficaz para esta tarea de clasificación.

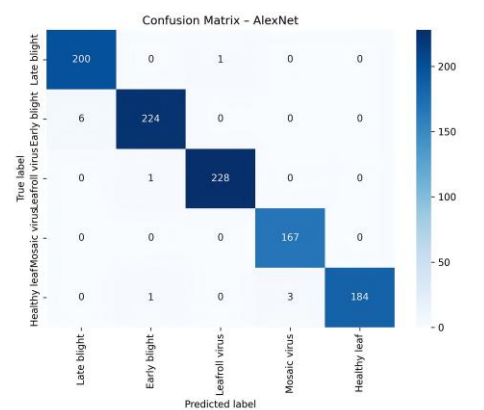


Fig. 12. Matriz de confusión de la arquitectura AlexNet para la clasificación de hojas de papa (sanas y enfermas).

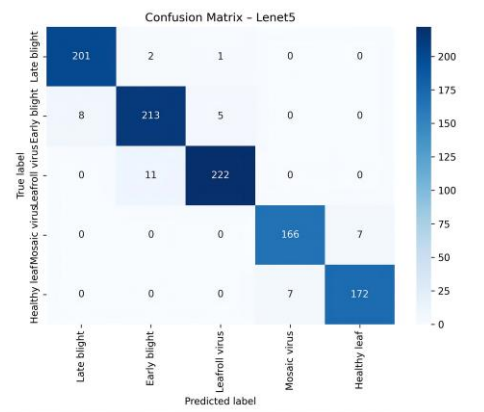


Fig. 13. Matriz de confusión de la arquitectura LeNet5 para la clasificación de hojas de papa (sanas y enfermas).

4 MÉTRICAS DE CLASIFICACIÓN

En las tablas 3 y 4 se muestran el resumen del rendimiento de clasificación de cada arquitectura evaluada para

la detección de enfermedades de las hojas de papa, distinguiendo entre modelos preentrenados y modelos entrenados desde cero.

TABLA 3
Resumen de las métricas de clasificación del modelo preentrenado

Modelo	Image Batch	Epocas (Fase 1 y 2)	Optimizador	Accuracy	Precision (macro)	Recall (macro)	F1-score (macro)	Entrenamiento	Validación	Prueba
MobileNet	32	20	Adam	0.991	0.9911	0.9911	0.991	0.9908	0.992	0.991
VGG16	32	20	Adam	0.995	0.9946	0.9942	0.9944	0.9977	0.995	0.995
ResNet50	32	20	Adam	0.988	0.9886	0.988	0.9882	0.9983	0.995	0.988
MobileNet	16	30	RMSprop	0.994	0.994	0.9937	0.9939	0.9932	0.989	0.994
VGG16	16	30	RMSprop	0.994	0.9937	0.9934	0.9935	0.9985	0.996	0.994
ResNet50	16	30	RMSprop	0.988	0.9881	0.9865	0.9873	0.9983	0.992	0.988
MobileNet	16	20	Adam	0.994	0.9944	0.9941	0.9942	0.9942	0.991	0.994
VGG16	16	20	Adam	0.994	0.9937	0.9934	0.9935	0.9972	0.997	0.994
ResNet50	16	20	Adam	0.988	0.9886	0.988	0.9882	0.9972	0.997	0.993
MobileNet	32	30	Adam	0.989	0.9897	0.9893	0.9894	0.9908	0.991	0.989
VGG16	32	30	Adam	0.998	0.9975	0.9976	0.9976	0.9987	0.997	0.998
ResNet50	32	30	Adam	0.989	0.9894	0.9889	0.9891	0.9991	0.995	0.989

TABLA 4
Resumen de las métricas de clasificación de los modelos entrenados desde cero

Model0	Learning rate	Epocas	Optimizador	Accuracy	Precision (macro)	Recall (Macro)	F1-score (macro)	Entrenamiento	Validación	Prueba
Alex Net	1e-3	30	Adam	0.9803	0.981	0.9789	0.9798	0.9697	0.9879	0.9803
LeNet5	1e-3	30	Adam	0.9005	0.9045	0.9022	0.9021	0.8446	0.9012	0.9005
Alex Net	1e-4	30	Adam	0.9882	0.988	0.9887	0.9882	0.9816	0.9879	0.9882
LeNet5	1e-4	30	Adam	0.9596	0.9597	0.9602	0.9599	0.9479	0.9617	0.9596
Alex Net	1e-3	30	RMSprop	0.9783	0.9787	0.9788	0.9786	0.9626	0.9839	0.9783
LeNet5	1e-3	30	RMSprop	0.8916	0.8946	0.892	0.8923	0.8435	0.9214	0.8916

5 INTERPRETABILIDAD DEL MODELO MEDIANTE GRAD-CAM

Para analizar el comportamiento interno del modelo y comprender qué regiones de la imagen son determinantes en sus predicciones, se empleó la técnica de Mapeo de Activación de Clase Ponderado por Gradiente (Grad-CAM). Este método genera mapas de calor que resaltan las regiones más relevantes consideradas por la red neuronal durante el proceso de clasificación. La Figura 14 presenta ejemplos representativos de Grad-CAM aplicados a imágenes de hojas de papa afectadas por diferentes enfermedades.

En general, el modelo concentró su atención en regiones asociadas a síntomas característicos, como lesiones, áreas necróticas y alteraciones de color, evidenciando la capacidad de aprender características discriminativas relevantes para la clasificación. Sin embargo, se observó variabilidad en la localización de las activaciones, presentándose patrones tanto concentrados como difusos según la complejidad de la imagen y la severidad de los síntomas. Estos resultados aportan interpretabilidad al modelo y respaldan su fiabilidad para la detección de enfermedades de las hojas de papa.

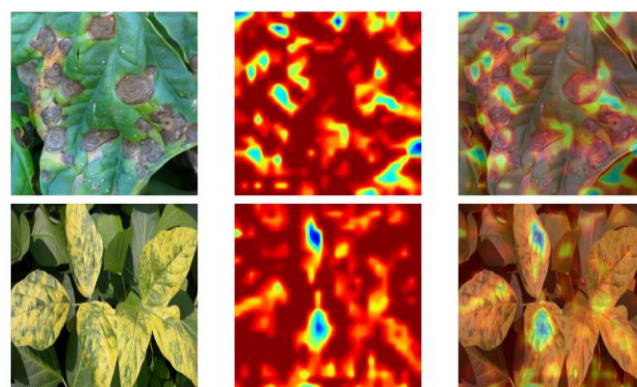


Fig. 14. Visualización del método Grad-CAM aplicado a una hoja de papa con síntomas de enfermedad. De izquierda a derecha, se muestran la imagen original, el mapa de calor generado por el modelo y la superposición de Grad-CAM.

6 DISCUSIÓN

Los resultados obtenidos en este estudio demuestran que los modelos preentrenados de redes neuronales convolucionales superan significativamente a los modelos entrenados desde cero en la clasificación de enfermedades foliares de la papa. En particular, el modelo preentrenado

VGG16 alcanzó el mejor rendimiento general, obteniendo valores de precisión de 0.9987 durante el entrenamiento, 0.997 en la validación y 0.998 en la prueba. Este comportamiento puede atribuirse a la capacidad de los modelos preentrenados para aprovechar representaciones previamente aprendidas a partir de grandes volúmenes de datos, favoreciendo una extracción más eficiente de características discriminatorias presentes en las imágenes foliares. En contraste, el modelo AlexNet entrenado desde cero mostró un rendimiento competitivo, aunque inferior al de los modelos basados en aprendizaje por transferencia (transfer learning), alcanzando valores de 0.9816 en entrenamiento, 0.9879 en validación y 0.9882 en prueba.

Estos hallazgos concuerdan con el trabajo de Hritwik et al. (2023) [14], quienes evaluaron arquitecturas como VGG19, DenseNet121 y ResNet50 para la clasificación de enfermedades de hojas de papa e informaron que VGG alcanzó la mayor precisión (98,77 %). De manera similar, en el presente estudio, VGG16 demostró una capacidad superior para extraer características relevantes de imágenes foliares, incluso en comparación con arquitecturas entrenadas desde cero, lo que refuerza su eficacia en tareas de clasificación agrícola.

En comparación con el estudio de William et al. [15], quienes propusieron una CNN de 11 capas inspirada en VGG16 y la evaluaron en el conjunto de datos PlantVillage, su modelo alcanzó una precisión del 91 %, mientras que AlexNet y VGG16 lograron precisiones del 94 % y 96 %, respectivamente. En contraste, el presente estudio reporta una precisión de prueba de 99,8 % utilizando VGG16 preentrenada y 98,82 % con AlexNet, lo que evidencia una mejora sustancial en el rendimiento de clasificación de enfermedades foliares de la papa. Estas diferencias podrían estar asociadas a variaciones en las configuraciones experimentales, estrategias de entrenamiento, selección de hiperparámetros o composición del conjunto de datos empleado. Por otro lado, los resultados reportados por Cristiano Muñoz [13] evidencian una precisión del 99,67 % mediante transfer learning y fine-tuning con InceptionV3. Estos hallazgos respaldan la premisa de que los modelos preentrenados pueden alcanzar niveles elevados de precisión en tareas de clasificación agrícola. En este contexto, VGG16 obtuvo un rendimiento excepcional mediante procedimientos estándar de entrenamiento y validación, confirmando la efectividad de los enfoques basados en transfer learning para la detección automática de enfermedades en las hojas de papa.

En síntesis, los resultados confirman que el modelo VGG16 preentrenado constituye la arquitectura con mayor precisión y robustez para la clasificación de enfermedades en las hojas de la papa dentro del conjunto de modelos evaluados, superando tanto a los modelos entrenados desde cero como a diversos enfoques reportados en la literatura.

7 CONCLUSIONES

En este estudio se evaluaron y compararon arquitecturas de redes neuronales convolucionales preentrenadas y entrenadas desde cero para la clasificación automática de enfermedades foliares de la papa. Los resultados demostraron que los modelos preentrenados presentan un desempeño superior en precisión, estabilidad y capacidad de generalización frente a los modelos entrenados desde cero.

Entre las arquitecturas evaluadas, VGG16 obtuvo el mejor rendimiento global, alcanzando una precisión de prueba del 99,8 %, seguida de MobileNet y ResNet50, con precisiones de prueba entre 98,8 % y 99,4 %.". Además, las reducidas diferencias entre entrenamiento, validación y prueba evidenciaron una adecuada generalización y ausencia de sobreajuste significativo. En los modelos entrenados desde cero, AlexNet alcanzó el mejor desempeño (98,8 %), mientras que LeNet-5 logró resultados satisfactorios (95,9 %), destacando por su simplicidad y menor costo computacional. Los hallazgos respaldan la eficacia del aprendizaje por transferencia (transfer learning) para la detección automática de enfermedades en hojas de papa.

La principal contribución de este estudio radica en la evaluación comparativa de arquitecturas preentrenadas y entrenadas desde cero sobre un conjunto de datos integrado a partir de múltiples repositorios públicos, incorporando cinco categorías relevantes de enfermedades de las hojas de papa. Además, el desarrollo de una aplicación móvil basada en el modelo óptimo demuestra la viabilidad de trasladar los resultados experimentales hacia aplicaciones agrícolas en tiempo real. La aplicación, denominada RaphiAI, se implementó utilizando React Native para la interfaz de usuario y FastAPI para la gestión del backend y el despliegue del modelo entrenado, como se ilustra en la figura 15. El código fuente de la aplicación móvil y el modelo entrenado están disponibles públicamente en un repositorio [GitHub](#), además proporciona la aplicación descargable para pruebas y uso práctico.

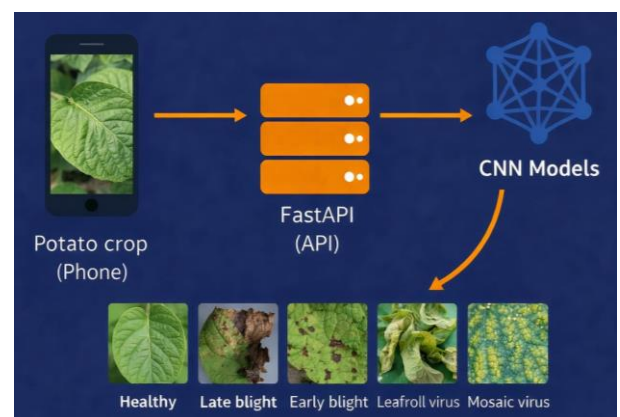


Fig. 15. Arquitectura del sistema propuesto.

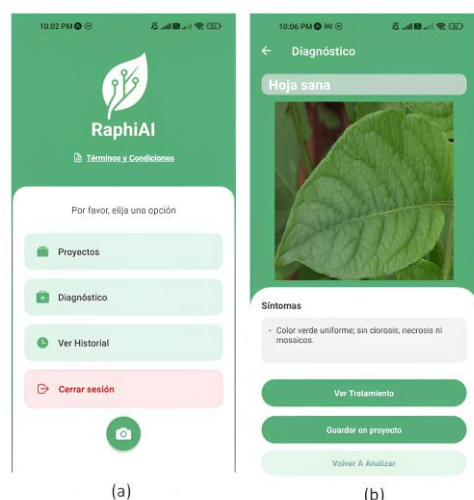


Fig. 16. Interfaz de la aplicación: (a) menú principal, (b) vista del resultado de la predicción.

REFERENCIAS

- [1] T. Utkarsh Yashwant, D. A. Shobanadevi and D. A. Shanthini, "Potato Leaf Disease Classification using Deep Learning: A Convolutional Neural Network Approach," 04 11 2023, <https://doi.org/10.48550/arXiv.2311.02338>
- [2] R. E. Ferreira Candia, "Food and Agriculture Organization of the United Nations," 10 12 2024. [Online]. Available: <https://www.fao.org/paraguay/noticias/detail-events/en/c/1730017/>.
- [3] H. Bajwa, "The New Indian Express," 04 01 2024. [Online]. Available: https://www.newindianexpress.com/nation/2024/Jan/04/late-blight-disease-10-potato-crops-damaged-in-punjab-prices-fall-to-rs-5-per-kg-2647839.html?utm_source.
- [4] Fresh Plaza, "Fresh Plaza," 10 12 2024. [Online]. Available: <https://www.freshplaza.com/north-america/article/2105776/spain-90-of-potato-crop-lost-to-disease/?>.
- [5] Ministerio de desarrollo agrario y riego, "Análisis de brechas de rendimiento en el cultivo de papa en el Perú," 2023. [Online]. Available: <https://cdn.www.gob.pe/uploads/document/file/6305660/5543457-analisis-brecha-rendimiento-en-el-cultivo-de-papa.pdf?>.
- [6] M. Krichen, "Convolutional Neural Networks: A Survey," MDPI, 07 07 2023, <https://doi.org/10.3390/computers12080151>
- [7] D. Renza and D. M. Ballesteros, "Fundamentos de visión por computador utilizando aprendizaje profundo," Cali, Redipe, 2023, <https://editorial.redipe.org/index.php/1/catalog/download/156/277/5544?inline=1>
- [8] K. S. Gamez Abril, "Deep Learning en agricultura: conceptos y aplicaciones en la identificación de cultivos sobre imágenes satelitales," Bogotá, Colombia, 2024, <https://hdl.handle.net/1992/73839>
- [9] F. J. Martín Ameneiro, "Análisis Automático de imágenes para detección de melanoma," Madrid, 2020, https://repositorio.uam.es/bitstream/handle/10486/692735/martin_ameneiro_francisco-javier_tfg.pdf?sequence=1
- [10] L. M. K. Sheikh, A. Shaikh, A. Sandupatla, R. Pudale, A. Bakare, P. M. Chavan and I. j. f. j. G, "Transfer Learning Based Plant Diseases Detection Using Res-Net5," ijrasnet, 20 04 2024, <https://doi.org/10.22214/ijrasnet.2024.60677>
- [11] T. D. Salka, M. B. Hanafi, S. M. S. A. A. Rahman, D. B. M. Zulperi and Z. Omar, "Plant leaf disease detection and classification using convolution neural networks model: a review," Artificial Intelligence Review, vol. 58, no. 322, 2025, <https://doi.org/10.1007/s10462-025-11234-6>
- [12] A. R. Gavilanez Guanoluiza and B. A. Saragosin Ghuamushig, "Desarrollo de un prototipo para la identificación automática de plagas y enfermedades en el cultivo de papa utilizando técnicas de inteligencia artificial en la ciudad de Latacunga," Universidad Técnica de Cotopaxi, Ecuador, 2024, <http://repositorio.utc.edu.ec/handle/27000/12029>
- [13] G. N. Cristiano Munuz, "Clasificación de enfermedades en cultivos de papa empleando redes neuronales convolucionales," La pontificia universidad Javerina, Bogotá, 2023, <https://repository.javeriana.edu.co/handle/10554/65257>
- [14] H. Ghosh, I. Sadiq Rahat, K. Shaik, S. Khasim and M. Yesubabu, "Potato Leaf Disease Recognition and Prediction using Convolutional Neural Networks," EAI Endorsed Transactions on Scalable Information Systems, vol. 10, no. 6, p. 8, 2023, <https://doi.org/10.4108/eetsis.3937>
- [15] W. A. Lozada Portilla, M. J. Suarez Baron and E. Aven-
daño Fernadez, "Aplicación de Redes Neuronales Convolucionales para la Detección del Tizón Tradío Phytophthora Infestans en papa Solanum Tuberosum," Revista Facultad Nacional de Agronomía Medellín, vol. 74, no. 2, p. 18, 06 12 2021, <https://doi.org/10.31910/rudca.v24.n2.2021.1917>
- [16] N. Sankalana, "Kaggle," 08 2024. [Online]. Available: <https://www.kaggle.com/datasets/nirmal-sankalana/potato-viral-disease-dataset/code>.
- [17] H. Nouman, "Kaggle," 08 2024. [Online]. Available: <https://www.kaggle.com/datasets/hafiznouman786/potato-plant-diseases-data?select=Potato-Plants>.
- [18] Sadik, "Kaggle," 08 2025. [Online]. Available: <https://www.kaggle.com/code/renjiabarai/potato-leaf-disease-prediction/input>.
- [19] B. Olimov, "Kaggle," 08 2025. [Online]. Available: <https://www.kaggle.com/code/kill92/f1-0-98-potato-disease-classifier-pytorch/input>.
- [20] A. Wagih Ibrahim, "Kaggle," 08 2023. [Online]. Available: <https://www.kaggle.com/datasets/abdallahalidev/plantvillage-dataset>.