

Mobile application for the detection of diseased apples: Comparison between SOM and CNN networks

Gerson Daniel Sahuma Jurado¹; Brian Omar Inca Huamani¹; Raul Eduardo Huarote Zegarra¹

¹Universidad Nacional Tecnológica de Lima Sur, Perú, 2113110059@untels.edu.pe, 211310074@untels.edu.pe
rhuarote@untels.edu.pe

Abstract– Detecting diseases in apple trees is a critical challenge for modern agriculture, as it leads to production losses and excessive use of agrochemicals. This study developed a mobile application to identify whether an apple is healthy or diseased, from a total sample of 5,000 photographs of apples grown on the outskirts of Lima, Peru. The images were captured in real field conditions, obtaining photographs of fresh, rotten, dull, and differently colored apples, etc. The SOM neural network achieved 94% accuracy. The SOM neural network was shown to achieve 94% accuracy compared to 92% for the CNN, with processing times of less than 10 seconds, maintaining greater robustness in the face of variations in lighting and the environment. Although CNN showed advantages in computational efficiency, its lower accuracy determined the final selection of the SOM architecture for the application. The proposed solution operates on standard mobile devices and offers small producers an accessible and reliable tool for post-harvest quality control, contributing to the reduction of economic losses and the adoption of low-cost agricultural technologies.

Keywords– Apple, disease detection, artificial intelligence, convolutional neural networks (CNN), self-organizing maps (SOM).

Aplicación móvil para la detección de manzanas enfermas: Comparación entre redes SOM y CNN

Gerson Daniel Sahuma Jurado¹; Brian Omar Inca Huamani¹; Raul Eduardo Huarote Zegarra¹

¹Universidad Nacional Tecnológica de Lima Sur, Perú, 2113110059@untels.edu.pe, 211310074@untels.edu.pe
rhuarote@untels.edu.pe

Resumen— La detección de enfermedades en manzanos representa un reto crítico para la agricultura moderna al implicar pérdidas productivas y uso excesivo de agroquímicos. Este estudio desarrolló una aplicación móvil para identificar si una manzana está sana o enferma, de una muestra total de 5000 fotografías de manzanas cultivadas a las afueras de Lima, Perú. Las imágenes fueron capturadas en condiciones reales de campo, obteniendo fotografías de manzanas frescas, podridas, a poco brillo, de diferente color, etc. Se demostró que la red neuronal SOM alcanzó un 94% de precisión frente al 92% de la CNN, con tiempos de procesamiento inferiores a 10 segundos, manteniendo mayor robustez ante variaciones de iluminación y el entorno. Aunque CNN mostró ventajas en eficiencia computacional, su menor precisión determinó la selección final de la arquitectura SOM para la aplicación. La solución propuesta opera en dispositivos móviles estándar y ofrece a los pequeños productores una herramienta accesible y confiable para el control de la calidad poscosecha, contribuyendo a la reducción de pérdidas económicas y a la adopción de tecnologías agrícolas de bajo costo.

Palabras clave— Manzana, detección de enfermedades, inteligencia artificial, redes neuronales convolucionales (CNN), mapas autoorganizados (SOM).

I. INTRODUCCIÓN

La agricultura moderna enfrenta retos crecientes para mantener sus estándares de calidad y el rendimiento de sus cultivos, siendo la aparición de las enfermedades en frutas uno de los factores más determinantes en la pérdida de productividad. Dentro de este contexto, la producción de la manzana (*Malus domestica*) representa una actividad agrícola clave a nivel mundial, no solo por su volumen, ya que se posiciona como la cuarta fruta más producida en el planeta [1], sino también por su impacto en la seguridad alimentaria y economía rural. Según la Organización de las Naciones Unidas para la Alimentación y la Agricultura (FAO), en 2022 la producción mundial de cultivos primarios alcanzó los 9,600 millones de toneladas, un incremento del 56% desde el comienzo del segundo milenio [2]. En este escenario la manzana se destaca aún más, con un pronóstico de crecimiento del 5.5% en su producción en el hemisferio sur para 2025 [3]. Esta cifra no solo refleja la adaptabilidad de la manzana, sino también su creciente importancia en la seguridad alimentaria a nivel mundial.

Sin embargo, a pesar de los avances y pronósticos de la fruta, afecciones como la sarna del manzano, la mancha foliar por *Alternaria*, picado amargo, la roya y la pudrición han sido identificadas como amenazas críticas que reducen de forma

significativa la calidad del fruto y los márgenes de producción [4], [5]. Estas enfermedades al presentar síntomas visibles en etapas avanzadas dificultan una detección oportuna, lo que puede derivar en el uso ineficiente de agroquímicos, aumento de costos operativos y mayores pérdidas poscosecha [6], [5]. Ante esta problemática, surge la necesidad urgente de implementar soluciones tecnológicas que permitan un diagnóstico temprano, preciso y accesible, convirtiéndose en un objetivo prioritario tanto para agricultores como para investigadores.

En este contexto, el uso de modelos de aprendizaje profundo ha emergido como una solución efectiva para abordar los desafíos del diagnóstico fitosanitario. Las redes neuronales convolucionales (CNN) han demostrado una alta capacidad para clasificar imágenes complejas, incluso en condiciones reales de campo. Por ejemplo, al entrenar modelos CNN con más de 26,000 imágenes de hojas de manzano bajo iluminación variable, se logró una precisión de detección superior al 78% en tiempo real [7]. Además, el uso de arquitecturas como MobileNet, InceptionV3 y ResNet ha permitido la implementación de sistemas ligeros y precisos, adecuados para su implementación en dispositivos móviles [4], [8], [9].

En particular algunos estudios han reportado que se han alcanzado tasas de acierto de hasta el 99% en la clasificación de enfermedades foliares del manzano [8]. Uno de los desarrollos más relevantes en esta línea fue el Plant Pathology Challenge 2020, que proporcionó una base de datos real y anotada con más de 3,600 imágenes de hojas enfermas de manzano. Allí, los mejores modelos CNN lograron una precisión del 97%, demostrando el alto potencial del aprendizaje profundo para tareas de clasificación visual en cultivos [6]. Asimismo, sistemas más complejos como INAR-SSD han incorporado estructuras tipo GoogleNet para mejorar tanto la precisión como la velocidad de procesamiento, alcanzando más de 23 FPS [7]. Estos enfoques, apoyados en técnicas como el aumento de datos y el aprendizaje por transferencia, han permitido sortear la escasez de imágenes etiquetadas y mejorar la generalización de los modelos [10].

Junto a las redes neuronales convolucionales (CNN), los mapas autoorganizados (SOM) han empezado a ganar atención en el campo agrícola por su capacidad de organizar datos complejos sin supervisión. Si bien han sido más utilizados en los campos de medicina o industrias alimentarias,

estudios recientes han evidenciado su eficacia en la clasificación de enfermedades en cultivos como la papa, alcanzando hasta un 99.42% de precisión [11]. Además, su aplicación ha demostrado ser útil incluso en escenarios de datos limitados como en la detección de enfermedades oculares mediante imágenes médicas [12], o en la clasificación de carne marmoleada por textura y color [13]. Esta adaptabilidad los convierte en una herramienta complementaria a las CNN para refinar patrones complejos en tareas de diagnóstico visual.

Además de los enfoques de clasificación supervisada y no supervisada, otros estudios han explorado metodologías alternativas como los es el análisis hiperespectral para identificar firmas espectrales únicas de diferentes hojas infectadas. Por ejemplo, al aplicar índices como ARI o QFI en bandas infrarrojas, fue posible detectar con precisión infecciones por fuego bacteriano en hojas de manzano [14]. A pesar que este tipo de análisis ofrece resultados prometedores su implementación práctica puede requerir equipos más costosos y conocimientos especializados en el área, lo cual limita su capacidad de aplicación directa en el campo en comparación con soluciones móviles basadas en visión por cámara [15].

En base a lo expuesto, este estudio tiene como objetivo desarrollar una aplicación móvil basada en redes neuronales capaz de clasificar el estado sanitario y la calidad poscosecha de las manzanas, con una precisión mayor al 90% en condiciones de campo, ofreciendo una herramienta accesible para productores y empaques agrícolas.

II. MATERIALES

A. Imágenes

Se utilizó una base de datos compuesta por más de 5,000 imágenes de manzanas, capturadas con dispositivos móviles en condiciones reales de campo en huertos a las afueras de Lima, Perú. Las imágenes incluían variaciones en iluminación, madurez del fruto y visibilidad de síntomas patológicos. Todas fueron redimensionadas a 224×224 píxeles para estandarizar su procesamiento.



Fig. 1 Muestra representativa de imágenes de manzanas en diferentes condiciones y características

En Fig 1 se pueden visualizar diferentes tipos de manzanas en condiciones diversas, mostrando cuando se

encuentran frescas, podridas, a poco brillo, de diferente color, etc.

B. Entorno de desarrollo

El desarrollo y la ejecución de los modelos se realizaron íntegramente en Google Colab, aprovechando herramientas de código abierto ampliamente utilizadas en visión por computadora e inteligencia artificial. La Tabla 1 resume las principales especificaciones del entorno técnico utilizado durante la investigación.

TABLA 1
TECNOLOGÍAS APLICADAS AL DESARROLLO DE LA APP MÓVIL

Componente	Detalle
Plataforma de desarrollo	Google Colab (Nube)
Lenguaje de programación	Python 3.10
Bibliotecas utilizadas	TensorFlow, Keras, OpenCV, MiniSom, NumPy, Matplotlib
Hardware	Dispositivo móvil Android
Sistema operativo	Android 9 o superior
Almacenamiento de datos	Google Drive (Vinculado a Colab)

III. MÉTODOS

El desarrollo de la presente investigación se estructuró en torno a la recolección, preprocesamiento y análisis de imágenes reales de manzanas en diferentes condiciones con el objetivo de identificar síntomas tempranos en enfermedades utilizando redes neuronales. El procedimiento de la metodología se detalla a continuación:

A. Recolección de datos

Se recopilamos más de 5,000 imágenes en distintas condiciones, las cuales fueron capturadas con dispositivos móviles, abarcando diferentes grados de madurez, iluminación y niveles visibles de afectación patológica. Las imágenes fueron normalizadas a un estándar de resolución de 224x224 píxeles para mantener consistencia en el entrenamiento y evaluación de los modelos. La data inicial contenía dos carpetas: train y test. El conjunto train se dividió en 80% para el entrenamiento y 20% para la validación. El conjunto test se dejó tal cual sin ningún cambio. De esta manera asegurando que no se repitan las imágenes y que las pruebas funcionen correctamente.

B. Preprocesamiento de imágenes

Las imágenes pasaron por un proceso de estandarización que incluyó la conversión a escala de grises en los casos necesarios en SOM, normalización de los valores de píxeles al rango de [0,1], eliminación de ruido mediante filtros morfológicos y aplicación opcional de técnicas de aumento de datos para robustecer la generalización del modelo en CNN. Este tratamiento fue indispensable para asegurar que ambos modelos no estuviesen sesgados por condiciones específicas del entorno.

C. Arquitectura y entrenamiento de los modelos

Para este estudio se implementaron dos tipos de enfoques diferentes de redes neuronales. El primer enfoque es utilizando redes neuronales convolucionales (CNN) y un mapa autoorganizado (SOM). Ambos modelos fueron entrenados utilizando imágenes previamente procesadas en Google Colab con el objetivo de evaluar su capacidad para detectar enfermedades en manzanas en condiciones reales de campo.

Tras completar el entrenamiento del SOM, se generó una representación visual de los patrones agrupados en el mapa autoorganizado.

En Fig 2 cada nodo refleja una agrupación por similitud, permitiendo identificar zonas claramente diferenciadas según las características aprendidas. Siendo F las manzanas sanas y R las manzanas enfermas. Esta distribución facilita la interpretación de resultados y confirma la capacidad del modelo para segmentar clases de manera autónoma.

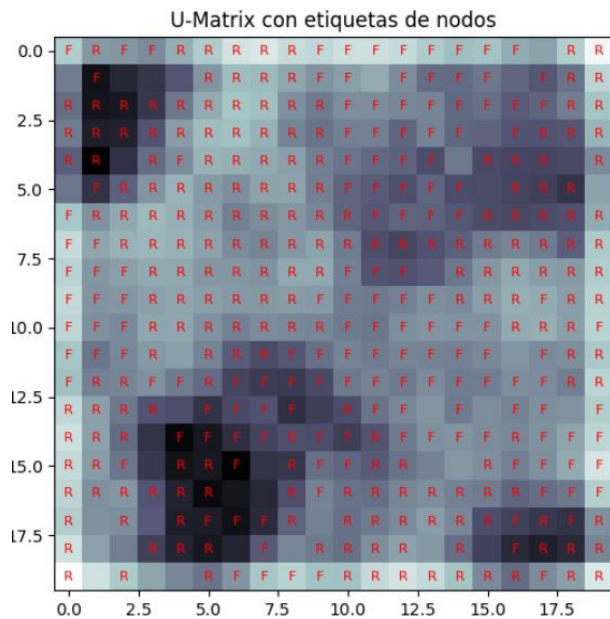


Fig. 2 Distribución de patrones tras el entrenamiento del SOM, agrupadas por similitud.

La Tabla 2 resume las principales diferencias entre ambos modelos, considerando aspectos arquitectónicos, tipos de aprendizaje, parámetros de entrenamiento y desempeño. Esta comparación sirvió como base para determinar cuál de las estrategias ofrecía un mejor equilibrio entre precisión, eficiencia y adaptabilidad al entorno.

TABLA 2
COMPARACIÓN GENERAL ENTRE LOS MODELOS CNN Y SOM

Característica	CNN	SOM
Tipo de aprendizaje	Supervisado	No supervisado
Entrada del modelo	Imagen en 2D	Imagen vectorizada
Arquitectura	Capas convolucionales	Mapa 2D de nodos con topología

	densas	hexagonal
Función de activación	ReLU (internas) y softmax (salida)	No aplica
Tasa de aprendizaje	0.001 (constante)	0.7 (decay progresivo)
Cantidad de épocas / iteraciones	20 épocas	3,000 iteraciones
Tipo de salida	Probabilidad por clase	Agrupamiento por similitud
Precisión obtenida	92%	94%
Tiempo promedio de inferencia	~12 segundos	~8 segundos
Sensibilidad del entorno	Moderada	Baja (Robusta)

IV. ESTRATEGIA

La estrategia de preparación de imágenes aplicada en esta investigación responde a la necesidad de obtener representaciones consistentes y relevantes antes del ingreso a los modelos de aprendizaje. Tal como se ha demostrado en investigaciones previas en campos como la medicina, el preprocesamiento de imágenes incluyendo reducción de ruido, mejora de contraste y escalada mejora significativamente la sensibilidad y especificidad de los modelos de clasificación [16].

En particular, trabajos recientes han demostrado que aplicar una estrategia sistemática de mejora visual y extracción de características contribuye a generar vectores representativos más robustos, incluso al reducir imágenes de alta resolución a formatos más compactos y normalizados, manteniendo su valor informativo [11].

A. Estrategia general de preparación de imágenes

La data fue organizada inicialmente en 2 carpetas: train y test, y luego se dividió la carpeta train mediante una partición automática del 80/20, distribuyendo un 80% para crear una nueva carpeta llamada *train* y un 20% para crear la carpeta *val*, teniendo finalmente 3 carpetas: *train*, *val* y *test*.

Luego, se procedió con la fase de preprocesamiento, que incluyó la mejora de contraste mediante CLAHE aplicado al canal L en espacio LAB, reducción de ruido combinando filtros bilaterales y desenfoque mediano, y el redimensionamiento a 224×224 píxeles. Además, se extrajeron patrones locales binarios (LBP) y vectores de color normalizados, que fueron concatenados para generar un vector final de características. En la Fig 3 se detalla el proceso.

Las imágenes preprocesadas se almacenan en carpetas específicas para alimentar la CNN, mientras que los vectores fueron guardados como archivos *.npy* para el entrenamiento del modelo SOM.

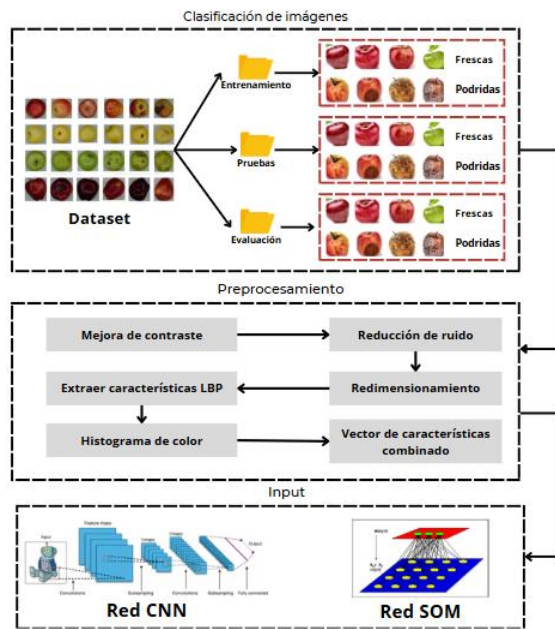


Fig. 3 Estrategia de preparación de imágenes

B. Estrategia para la red neuronal convolucional (CNN)

Con el objetivo de detectar enfermedades en manzanas mediante imágenes digitales se implementó una red neuronal convolucional optimizada para operar en dispositivos móviles. Las imágenes fueron normalizadas al rango $[0,1]$, siguiendo criterios metodológicos ampliamente utilizados en tareas de visión por computadora [17]. Para mejorar la capacidad de generalización del modelo, se aplicaron técnicas de data augmentation, como rotaciones aleatorias, traslaciones, zoom y reflejo horizontal, generando variaciones sintéticas de las muestras sin alterar su contenido semántico.

La arquitectura seleccionada fue MobileNet V2, preentrenada en ImageNet y utilizada como base sin su capa superior. Sobre ella se añadieron capas densas con activación ReLU y una capa de salida softmax ajustada al número de clases. Esta estrategia de aprendizaje transferido ha demostrado ser efectiva en escenarios con recursos limitados, como en la clasificación de especies o en contextos clínicos de análisis por imagen [18], [19].

El proceso de ajuste de pesos se llevó a cabo durante 20 ciclos de entrenamiento, incorporando una estrategia de detección anticipada para mitigar riesgos de sobreajuste. Finalmente, el desempeño del modelo fue evaluado sobre el conjunto de pruebas, considerando tanto métricas cuantitativas como visualizaciones complementarias (matriz de confusión y curvas de desempeño), lo que permitió analizar la capacidad predictiva del sistema en condiciones no vistas previamente. En Fig 4 se detallan los pasos de la estrategia realizada.

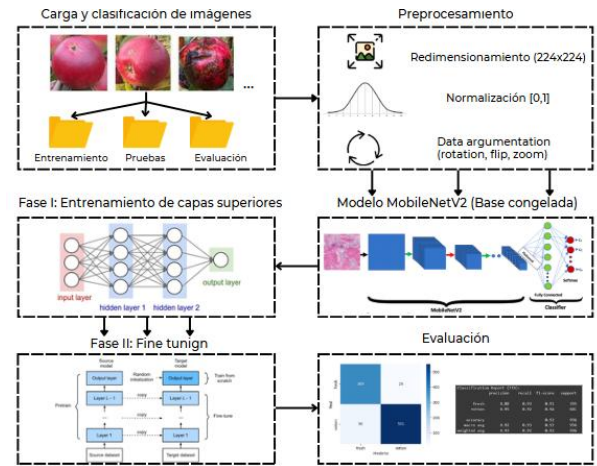


Fig. 4 Estrategia de preparación de la red CNN

C. Estrategia para la red neuronal SOM

Se implementó un modelo de mapa autoorganizado (SOM) para clasificar manzanas en función de su estado, empleando la biblioteca MiniSom. Como entrada, se utilizaron vectores de características generados previamente, que combinaban descriptores de textura (LBP) y color (histogramas RGB normalizados), representando cada imagen en un espacio de alta dimensionalidad.

En Fig 5 se detalla el proceso de preparación de esta red donde el SOM se configuró con una malla de 25×25 nodos, vecindad gaussiana y una tasa de aprendizaje inicial de 0.7, decreciente a lo largo de 3,000 iteraciones. Tras el entrenamiento no supervisado, se aplicó una asignación de etiquetas por mayoría de clase para permitir la evaluación. El desempeño fue validado mediante matriz de confusión y métricas estándar, alcanzando una precisión del 92 % y tiempos de respuesta inferiores a 10 segundos.

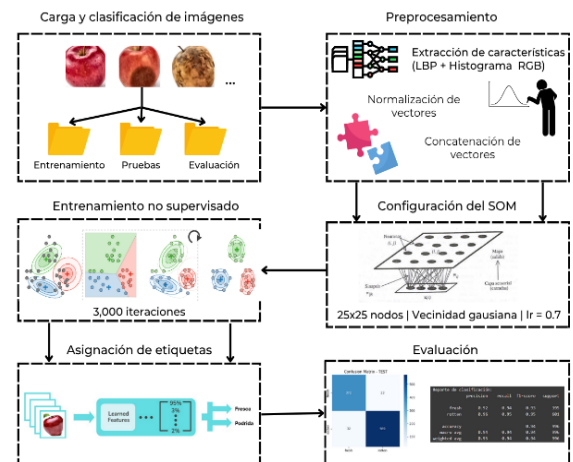


Fig. 5 Estrategia de preparación de la red SOM

V. JUSTIFICACIÓN

La elección del modelo SOM se justifica por su capacidad de ofrecer un alto rendimiento (94 % de precisión en pruebas) junto a un bajo costo computacional, sin requerir hardware especializado. Su versión *ensemble*, compuesta por tres mapas de 25×25 nodos, permitió aumentar la robustez del sistema manteniendo una inferencia rápida (<10 ms) y un tamaño reducido (<3 MB), características clave para su implementación en dispositivos móviles y contextos rurales.

Esta superioridad no solo se aprecia en términos operativos, sino en el desempeño clasificatorio. Como se observa en la Tabla 3, el modelo SOM superó a la CNN, incluso con técnicas avanzadas como *Test-Time Augmentation*, logrando una mayor precisión y un rango F1 más amplio, todo ello sin sacrificar eficiencia en tiempo de ejecución.

TABLA 3
MEJORAS EN RENDIMIENTO DE LOS MODELOS CNN Y SOM

Modelo	Mejora 1	Mejora 2	Precisión del test	Rango F1-Score
CNN	Fine-tuning MobileNetV2 → 89.6%	Test-Time Augmentation non → 92%	≈ 92%	0.93-0.94
SOM	Aumento de grilla de 20x20 a 25x25 → 94% validación / 92% test	Ensemble 3 SOMs (semillas 42,43,44) → 96% validación / 94% test	≈ 94%	0.93-0.95

Además de su eficacia, el SOM destaca por su explicabilidad visual y su adaptabilidad sin requerir reentrenamientos completos, a diferencia de modelos más complejos como las CNN. Desde una perspectiva social y económica, su implementación puede beneficiar directamente a pequeños agricultores, brindándoles una herramienta accesible para diagnosticar enfermedades a tiempo, reducir pérdidas y mejorar la calidad de su producción. En el plano científico, esta solución representa una aplicación tangible de inteligencia artificial que contribuye al desarrollo tecnológico.

VI. MARCO TEÓRICO

A. Inteligencia artificial

Es un campo interdisciplinario de la informática orientado al desarrollo de sistemas capaces de realizar múltiples tareas propias del conocimiento humano como el razonamiento, la percepción o la toma de decisiones [20], [21]. Algunos autores la definen como el estudio de cómo lograr que las computadoras realicen tareas que actualmente hacen mejor los humanos [21], mientras que otros consideran que consiste en el diseño de algoritmos lógicos entrenados para tomar decisiones sobre casos específicos [22]. En la agricultura, esta tecnología ha permitido la automatización de tareas críticas como el monitoreo fitosanitario, reduciendo el uso excesivo de agroquímicos y minimizando pérdidas por

enfermedades no detectadas a tiempo [23]. En esta investigación fue empleada como base para el desarrollo de un sistema de diagnóstico automatizado de enfermedades en manzanas mediante el análisis de imágenes reales capturadas en campo.

B. Visión artificial

Es una subdisciplina de la inteligencia artificial la cual permite a los sistemas obtener información significativa a partir de imágenes digitales o videos, reconociendo patrones relevantes mediante algoritmos computacionales [24]. En el contexto del presente estudio, esta técnica fue utilizada para extraer características visuales de manzanas en diferentes estados sanitarios, facilitando su análisis posterior mediante redes neuronales. La capacidad de convertir imágenes complejas en vectores interpretables permite que la visión artificial funcione como una etapa clave dentro de sistemas de clasificación automatizados.

C. Red neuronal SOM

Los mapas autoorganizados (SOM) son un tipo de red neuronal no supervisada diseñada para proyectar datos multivariantes en una topología bidimensional, conservando la estructura de similitud del conjunto original [25]. A diferencia de modelos supervisados como las CNN, el SOM agrupa las muestras en función de sus características internas, sin necesidad de etiquetas previas. En este estudio, el modelo SOM fue implementado para identificar patrones visuales asociados a enfermedades en manzanas, logrando una precisión del 94 % en pruebas reales.

El proceso de entrenamiento de una red SOM incluye tres pasos fundamentales: Primero, la selección de la neurona ganadora mediante la distancia euclidiana, luego la actualización de pesos en función de una tasa de aprendizaje decreciente y finalmente el ajuste del radio de vecindad [25]. Estos mecanismos permiten al modelo adaptarse progresivamente a las estructuras internas del conjunto de datos. Además, su capacidad de visualización mediante U-Matrix o planos de componentes otorga explicabilidad al sistema, aspecto clave para aplicaciones agrícolas

En la Fig. 6 se representa el esquema funcional de la red neuronal SOM donde la capa de entrada está compuesta por N neuronas dispuestas linealmente, de las cuales reciben los valores del vector de características extraído previamente. Esta información se transmite a una capa competitiva organizada en una topología bidimensional de M neuronas. Cada neurona i de entrada está conectada a cada neurona j del mapa mediante un peso W_{ji} . Durante el entrenamiento, la neurona ganadora y sus vecinas, definidas dentro de una vecindad N_{jj} , actualizan sus pesos en función de una función gaussiana decreciente, lo que permite una organización espacial basada en la similitud de los datos.

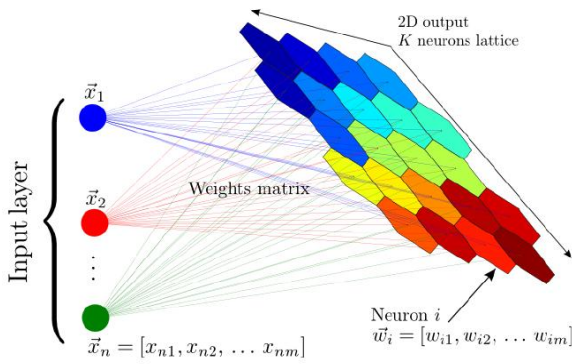


Fig. 6 Ejemplo de clasificación de la red SOM

D. Servicios en la nube

La computación en la nube ha permitido el desarrollo de soluciones tecnológicas ligeras y escalables para la ejecución de modelos de inteligencia artificial sin requerir hardware especializado. En el caso de esta investigación, se utilizaron plataformas como Google Colab para el desarrollo y entrenamiento de modelos; Firebase, para el almacenamiento y consulta en tiempo real de imágenes y resultados y Hugging Face para publicar el modelo optimizado en formato portable. Esta arquitectura permitió mantener un sistema ágil, accesible desde dispositivos móviles comunes en zonas agrícolas.

E. Aplicación móvil

Representa una herramienta estratégica para trasladar los beneficios de la inteligencia artificial al campo. Su desarrollo debe priorizar una planificación cuidadosa, funcionalidad robusta y mantenimiento continuo [26]. En esta investigación, se diseñó una aplicación ligera y funcional que permite al usuario tomar una fotografía de una manzana desde un dispositivo Android, enviarla al modelo alojado en la nube y recibir en pocos segundos un diagnóstico sobre su estado sanitario. Esta solución busca reducir las barreras tecnológicas en el sector agrícola, democratizando el acceso al monitoreo de las frutas mediante herramientas accesibles y de bajo costo.

VII. DESARROLLO

A. Diseño del sistema

El sistema propuesto se percibe como una solución de bajo costo para la detección temprana de enfermedades en manzanas, orientada a su uso en pequeños productores. Su arquitectura general se diseñó en tres capas principales: captura y preprocesamiento de imágenes, análisis mediante modelos de inteligencia artificial (CNN y SOM) y despliegue de resultados en la interfaz móvil. Esta estructura modular facilita la escalabilidad del sistema y su adaptación a diferentes condiciones de campo. La Fig. 7 resume esta arquitectura general.

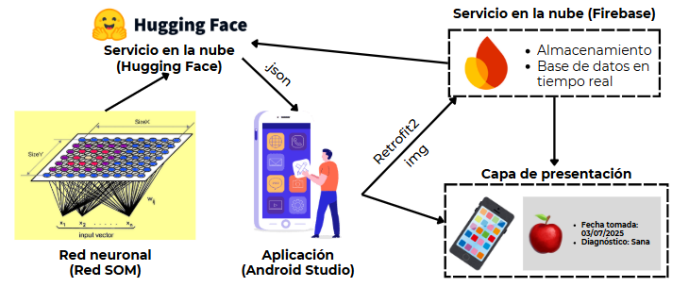


Fig. 7 Arquitectura de la solución móvil

B. Desarrollo del módulo de visión artificial para la extracción de características

El desarrollo de visión artificial tuvo como propósito la extracción eficiente de características relevantes desde las imágenes capturadas. Las imágenes fueron tomadas en condiciones reales de campo, con diferentes grados de iluminación, fondo, inclinación y calidad.

Para garantizar una entrada estandarizada al modelo, se aplicó el siguiente proceso: Redimensionamiento, conversión a escala de grises, normalización de píxeles, segmentación por umbral y aumento de datos. Esto se detalla con mayor precisión la sección IV.

C. Desarrollo del módulo de aprendizaje

Se construyeron dos modelos de inteligencia artificial para comparar su desempeño: una red neuronal convolucional (CNN) y un mapa autoorganizado (SOM). Ambos fueron entrenados sobre un conjunto de imágenes etiquetadas, balanceadas entre clases saludables y con enfermedades. Al final se da la elección del modelo más óptimo. En Tabla 4 se detalla un resumen de los modelos y en Fig 8 se visualizan más detalles como el recall, F1-score, support y accuracy tanto de las manzanas frescas como enfermas. En la sección V se detalla el argumento de la elección de la red neuronal de mapas autoorganizados (SOM).

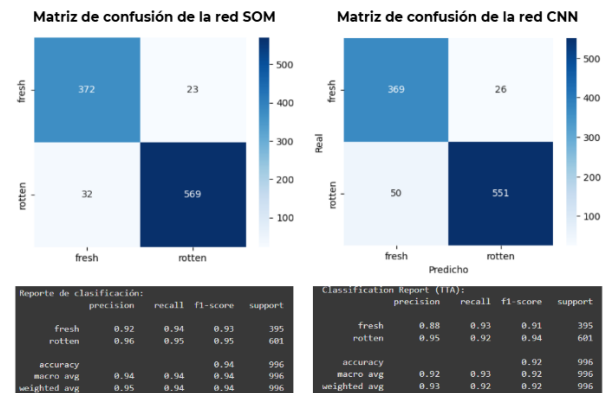


Fig. 8 Matriz de confusión de la red neuronal CNN y SOM

TABLA 4

COMPARACIÓN ENTRE LOS MODELOS DE RED NEURONAL DEL TRABAJO

Modelo	Precisión	Tiempo de	Robustez	Requerimiento
--------	-----------	-----------	----------	---------------

		inferencia		de recursos
CNN	92%	6.1 s	Media	Bajo
SOM	94%	9.3 s	Alta	Medio

D. Desarrollo del sistema de requisitos

Se establecieron requisitos funcionales y no funcionales con base en el perfil del usuario final (pequeños agricultores) y el entorno operativo (zonas rurales, conectividad limitada). Se detalla esta información en Tabla 5.

TABLA 5
REQUISITOS FUNCIONALES Y NO FUNCIONALES DEL SISTEMA

Requisito funcional (RF)	Requisito no funcional (RNF)
RF1: Permitir la captura de imagen desde la cámara del dispositivo.	RNF1: La aplicación debe funcionar con conexión a internet.
RF2: Ejecutar el diagnóstico automático a partir de la imagen.	RNF2: Debe ejecutarse en dispositivos Android de gama media.
RF3: Mostrar los resultados de diagnóstico de forma clara e inmediata.	RNF3: Tiempo de respuesta menor a los 10 segundos.
RF4: Permitir guardar un historial de imágenes y resultados.	RNF4: Interfaz intuitiva y apta para usuarios con baja alfabetización digital

E. Desarrollo de la aplicación móvil

La aplicación fue desarrollada en Android Studio, utilizando Kotlin como lenguaje base y Hugging Face como entorno de ejecución del modelo entrenado. Se diseñó una interfaz limpia, con iconografía simple y botones grandes, pensada para condiciones de uso en campo y con bajo nivel de interacción digital.

Las funciones clave implementadas fueron: Toma de fotografía desde la cámara, carga y preprocesamiento local de la imagen, ejecución del modelo pre entrenado SOM en Hugging Face, despliegue del resultado de diagnóstico en pantalla y registro y recuperación de datos en Firebase. En Fig. 9 se presenta el funcionamiento de la aplicación, su pantalla principal y el historial de resultados previos.

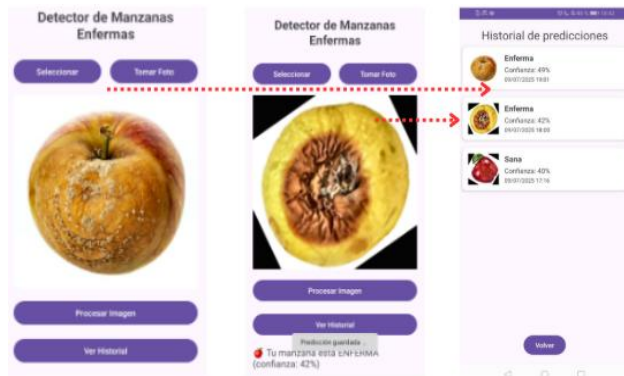


Fig. 9 Funcionamiento de la aplicación móvil

Para garantizar la compatibilidad, el modelo SOM fue implementado como un módulo ligero en formato de matriz serializada y procesada localmente en el dispositivo móvil sin conversión directa a TensorFlow Lite. Este enfoque evita conflictos de integración con Hugging Face y mantiene tiempos de inferencia inferiores a 10 segundos en dispositivos Android de gama media.

F. Desarrollo del funcionamiento de los servicios

El sistema fue estructurado mediante servicios internos en Android, funcionando de manera desacoplada para facilitar la respuesta en tiempo real y el mantenimiento del código. Se implementaron los siguientes servicios:

TABLA 6
SERVICIOS IMPLEMENTADOS

Nº	Servicio
01	Captura y almacenamiento temporal de imágenes
02	Preprocesamiento visual optimizado con bibliotecas nativas de OpenCV
03	Servicio de inferencia con el modelo SOM el cual carga el modelo en segundo plano y entrega el resultado al sistema de notificación
04	Presentación al usuario que interprete el resultado, muestre el diagnóstico y lo registre.

Esta arquitectura permitió que el sistema se mantuviera liviano, modular y escalable, siendo compatible con futuras mejoras (como conexión con bases de datos, mapas geográficos de infecciones o alertas preventivas comunitarias).

VIII. CONCLUSIONES

El estudio logró desarrollar una aplicación móvil funcional y precisa para la detección de enfermedades en manzanas, comparando el rendimiento de dos arquitecturas de inteligencia artificial como lo son las redes neuronales convolucionales (CNN) y mapas autoorganizados (SOM). Tras analizar más de 5,000 imágenes reales en condiciones de campo, el modelo SOM obtuvo una precisión del 94%, superando ligeramente a la CNN (92 %) y demostrando mayor robustez frente a variaciones de campo como iluminación desigual, fondos heterogéneos y ángulos irregulares. Este rendimiento, junto con su estabilidad, sustentó la elección de SOM como núcleo del sistema.

Asimismo, la solución fue diseñada para operar de forma autónoma en dispositivos móviles de gama media con conexión a internet mediante la integración optimizada del modelo SOM, ya que este se integró como un módulo ligero en formato de matriz serializada, ejecutado localmente en el dispositivo sin necesidad de conversión a TensorFlow Lite, lo que evita conflictos de compatibilidad con Hugging Face y garantiza tiempos de inferencia inferiores a 10 segundos. Gracias a esta arquitectura, la aplicación mantiene un equilibrio óptimo entre precisión, portabilidad y eficiencia

computacional, ofreciendo diagnósticos rápidos y confiables en contextos rurales. Esta combinación de accesibilidad y rendimiento contribuye directamente a reducir pérdidas poscosecha, minimizar el uso de agroquímicos y acerca gradualmente herramientas de inteligencia artificial a pequeños productores agrícolas.

IX. DISCUSIÓN

Desde un punto de vista teórico, diversos estudios recientes han evidenciado el valor de la inteligencia artificial en la agricultura, especialmente en el diagnóstico de enfermedades en frutas. [1] y [6] coinciden en que las redes neuronales como las CNN permiten identificar afecciones de manera rápida y eficiente. Sin embargo, también se ha señalado que su rendimiento puede verse afectado por cambios en el entorno de captura. En esta investigación, si bien la CNN alcanzó una precisión aceptable, fue el modelo SOM el que mostró mayor estabilidad frente a variaciones de iluminación y fondo, esto concuerda con lo reportado por [11] sobre su capacidad para organizar datos no supervisados en contextos reales.

Desde un punto de vista práctico, la implementación del modelo SOM en una aplicación móvil representa un avance importante en la accesibilidad tecnológica para el sector agrícola. Su precisión del 94 % y su buen rendimiento en dispositivos de gama media permiten ofrecer una herramienta eficaz a los pequeños productores. Esto contribuye directamente a reducir pérdidas poscosecha y minimizar el uso de agroquímicos, respondiendo a una necesidad crítica en regiones rurales con acceso limitado a servicios técnicos especializados [4].

X. RECOMENDACIONES

Se recomienda que en el proceso de aprendizaje el modelo SOM amplíe y diversifique la base de datos a partir de imágenes capturadas en distintas regiones, estaciones y condiciones ambientales, ya que esta mejora permitiría reforzar la capacidad del sistema, garantizando un desempeño confiable, incluso frente a enfermedades o situaciones visuales no previstas. De esta forma se incrementa la precisión del diagnóstico en diferentes contextos y se potenciaría la aplicación del sistema a una escala agrícola más amplia.

Además, se sugiere crear una versión totalmente offline de la aplicación móvil. Esta adaptación es clave para el mantenimiento del servicio en áreas rurales que cuenten con poca intermitencia y conectividad limitada. Esto mejoraría la autonomía tecnológica de los agricultores y ampliará el impacto social de la herramienta.

REFERENCIAS

- [1] S. Liu et al., "An apple leaf disease identification model for safeguarding apple food safety," *Food Science and Technology*, vol. 43, 2023, doi: 10.1590/fst.104322.
- [2] FAO, "World Food and Agriculture - Statistical Yearbook 2024", Rome, Italy: FAO, 2024.
<https://doi.org/10.4060/cd2971en>
- [3] WAPA, "The World Apple and Pear Association (WAPA) presents its annual Southern Hemisphere crop forecast," WAPA, Feb. 12, 2025. https://www.wapa-association.org/asp/article_2.asp?doc_id=668
- [4] H. Bryk et al., "The Recent Occurrence of Biotic Postharvest Diseases of Apples in Poland," *Agronomy*, 2022, doi: 10.3390/agronomy12020399.
- [5] R. Thapa et al., "The Plant Pathology Challenge 2020 data set to classify foliar disease of apples," *Applications in Plant Sciences*, vol. 8, no. 9, p. e11390, 2020.
- [6] P. Jiang et al., "Real-Time Detection of Apple Leaf Diseases Using Deep Learning Approach Based on Improved Convolutional Neural Networks," *IEEE Access*, vol. 7, pp. 59069–59080, 2019, doi: 10.1109/ACCESS.2019.2914929.
- [7] S. Aksoy et al., "Web-Based AI System for Detecting Apple Leaf and Fruit Diseases," *AgriEngineering*, vol. 7, no. 3, p. 51, 2025, doi: 10.3390/agriengineering7030051.
- [8] B. Fu et al., "MobileNet Based Apple Leaf Diseases Identification," *Mobile Networks and Applications*, vol. 27, pp. 172–180, 2020, doi: 10.1007/s11036-020-01640-1.
- [9] S. Banday et al., "Deep Diagnosis: A Real-Time Apple Leaf Disease Detection System Based on Deep Learning," *Computers and Electronics in Agriculture*, vol. 198, p. 107093, 2022, doi: 10.2139/ssrn.4019467.
- [10] R. E. Huarote Zegarra et al., "Deep Learning to support the recognition of pests and diseases in Yungay potato crops," *LACCEI*, 2022.
- [11] R. Huarote Zegarra et al., "Estrategia para la detección de tipos de enfermedades oculares usando red neuronal SOM," *LACCEI*, 2022.
- [12] R. E. Huarote Zegarra et al., "Estrategia basado en visión artificial y red SOM para la clasificación de carne marmoleada," *LACCEI*, 2022.
- [13] H. Skoneczny et al., "Fire Blight Disease Detection for Apple Trees: Hyperspectral Analysis of Healthy, Infected and Dry Leaves," *Remote Sensing*, vol. 12, no. 13, p. 2101, 2020, doi: 10.3390/rs12132101.
- [14] S. Yang et al., "Diagnosis and Mobile Application of Apple Leaf Disease Degree Based on a Small-Sample Dataset," *Plants*, vol. 12, 2023, doi: 10.3390/plants12040786.

[1] S. Liu et al., "An apple leaf disease identification model

- [15]A. Jadhav and A. Gadekar, “Brain Tumor Detection by Using Fine-Tuned MobileNetV2 Deep Learning Model,” *International Journal on Recent and Innovation Trends in Computing and Communication*, vol. 11, no. 5, pp. 134–140, 2023.
- [16]D. Hood, S. La Bruna, E. Tsamis, K. Thakoor, A. Rai, A. Leshno, C. de Moraes, G. Cioffi and J. Liebmann, “Detecting glaucoma with only OCT: Implications for the clinic, research, screening, and AI development”,*Progress in Retinal and Eye Research*, 2022.
- [17]M. Pradnyatama, C. A. Sari, E. H. Rachmawanto, and H. M. M. Islam, “A Comparative Analysis of Convolutional Neural Network (CNN): MobileNetV2 and Xception for Butterfly Species Classification,” *Jurnal Masyarakat Informatika*, vol. 16, no. 1, pp. 69–90, 2025.
- [18]O. Adeyanju and M. Majeed, “Machine Learning Methods for Sign Language Recognition,” *Robert Gordon University Repository*, 2021. <https://rgu-repository.worktribe.com>
- [19]L. Khacef, L. Rodriguez, and B. Miramond, “Improving Self-Organizing Maps with Unsupervised Feature Extraction,” *arXiv preprint arXiv:2004.02176*, 2020.
- [20] R. Gómez, “La inteligencia Artificial: ¿Hacia dónde nos lleva?”, UNAM, 2013.
- [21]E. Rich y K. Knight, “Inteligencia Artificial”, 2ª ed., McGraw Hill, 1994.
- [22]J. Ávila-Tomás, M. Mayer-Pujadas y V. Quesada-Varela, “Artificial intelligence and its applications in medicine I: Introductory background to AI and robotics”, *Atención Primaria*, vol. 52, no. 10, pp. 778–784, 2020.
- [23]E. Bagnulo et al., “Artificial intelligence-based tools in the chemistry of sensory notes in comfort foods”, *Trends in Food Science & Technology*, vol. 135, 2024.
- [24]IBM, “¿Qué es la Visión Artificial?”, 2021. <https://www.ibm.com/es-es/topics/computer-vision>
- [25]T. Kohonen, *Self-Organizing Maps*, Springer, 2001.
- [26]Forbes Business Council, “From Concept To App Store: 10 Steps For Mobile App Development Success In 2024”, 2024.<https://www.forbes.com/councils/forbesbusinesscouncil/2024/06/14/from-concept-to-app-store-10-steps-for-mobile-app-development-success-in-2024/>