



Balanceo de Conjuntos de Datos Basado en Redes Generativas Aplicado a Imágenes del Sector Agrícola

Generative Network-Based Data Balancing Applied to Agricultural Image Datasets

Autores

* **Luis Jesús Montesdeoca Espinoza**

✉ luis.montesdeoca.41@espam.edu.ec

Stalin Joel Zambrano Rojas

✉ stalin.zambrano@espam.edu.ec

Victor Joel Pinargote Bravo

✉ vpinargote@espam.edu.ec

Luis Cristobal Cedeño Valarezo

✉ lcedeno@espam.edu.ec

Escuela Superior Politécnica
Agropecuaria de Manabí Manuel Félix
López, Calceta, Manabí, Ecuador.

*Autor para correspondencia

Comó citar el artículo:

Montesdeoca Espinoza, L.J., Zambrano Rojas, S.J., Pinargote Bravo, V.J. & Cedeño Valarezo, L.C. (2025). Balanceo de Conjuntos de Datos Basado en Redes Generativas Aplicado a Imágenes del Sector Agrícola. *Informática y Sistemas* 9(2), pp. 164-176. <https://doi.org/10.33936/iys.v9i2.7782>

Enviado: 11/08/2025

Aceptado: 12/09/2025

Publicado: 12/09/2025

Resumen

El desbalance de clases en conjuntos de datos agrícolas limita el rendimiento de los modelos de clasificación basados en redes neuronales convolucionales, al dificultar la identificación precisa de clases minoritarias. Con el objetivo de mitigar este problema, se adaptó la metodología CRISP-DM para generar datos sintéticos mediante redes generativas adversarias Wasserstein (WGAN-GP), utilizando defectos segmentados de enfermedades del aguacate (Scab y Anthracnose) extraídos con técnicas de visión por computador. Estas anomalías fueron integradas en imágenes de frutos sanos para construir un conjunto de datos equilibrado. Posteriormente, se entrenó un modelo de clasificación utilizando la arquitectura InceptionV3 con aprendizaje por transferencia, evaluando su desempeño tanto con el conjunto desequilibrado como con el conjunto balanceado. Los resultados evidenciaron mejoras notables en la precisión, especialmente en la clasificación de enfermedades, cuando se utilizó el conjunto equilibrado, alcanzando una precisión de validación del 97.74%. Este estudio demuestra que el uso de datos sintéticos puede ser una solución efectiva para mejorar la capacidad predictiva de modelos en contextos donde la recopilación de datos reales es limitada o costosa.

Palabras clave: Aprendizaje profundo; Balanceo de clases; Datos sintéticos; Redes convolucionales; Segmentación de imágenes

Abstract

The imbalance of classes in agricultural datasets limits the performance of classification models based on convolutional neural networks, as it makes it difficult to accurately identify minority classes. To mitigate this problem, the CRISP-DM methodology was adapted to generate synthetic data using Wasserstein generative adversarial networks (WGAN-GP), using segmented defects from avocado diseases (Scab and Anthracnose) extracted with computer vision techniques. These anomalies were integrated into images of healthy fruits to construct a balanced dataset. The InceptionV3 architecture with transfer learning was then used to train a classification model, and its performance was assessed using both the balanced and unbalanced datasets. Using the balanced dataset resulted in significant accuracy gains, particularly in disease classification, with a validation accuracy of 97.74%. This study shows that in situations where real data collection is expensive or limited, using synthetic data can be a useful way to increase the predictive power of models.

Keywords: Deep learning; Class balancing; Convolutional networks; Synthetic data; Image segmentation





1. Introducción

La agricultura es un pilar esencial para la seguridad alimentaria y el desarrollo económico global (Thayer et al., 2020). Sin embargo, enfrenta desafíos crecientes como el cambio climático, la degradación del suelo y la propagación de enfermedades en cultivos, que demandan soluciones tecnológicas innovadoras. La Organización de las Naciones Unidas para la Alimentación y la Agricultura (FAO, 2021) ha resaltado la necesidad de integrar herramientas digitales para lograr sistemas agrícolas más resilientes y eficientes.

El aprendizaje automático, una de las tecnologías emergentes, ha demostrado tener un alto potencial para optimizar los procesos productivos, ayudar a la toma de decisiones fundamentadas en datos y mejorar la detección oportuna de enfermedades (Soria Olivas et al., 2023; Shalev-Shwartz & Ben-David, 2014). En el tratamiento de imágenes agrícolas, se han empleado en gran medida las arquitecturas codificador-decodificador, U-Net y las redes neuronales convolucionales (CNN), con resultados sobresalientes en la clasificación y segmentación de hojas y frutos (Lei et al., 2024). Sin embargo, su desempeño tiende a estar limitado por la disponibilidad de datos equilibrados y representativos.

Investigaciones anteriores han indagado sobre la aplicación de las redes generativas adversarias (GAN) para incrementar las clases minoritarias en fotografías agrícolas, alcanzando avances en sensibilidad y precisión (Lu et al., 2022; Carvajal Chávez, 2023). No obstante, estos estudios tienen limitaciones típicas: el riesgo de sobreajuste al volver a utilizar imágenes de validación en el proceso de generación, la escasa evaluación de la variabilidad morfológica que se produce y la falta de control sobre cómo se integran los datos sintéticos. Asimismo, la mayor parte se ha enfocado en enfermedades o cultivos concretos sin especificar la posibilidad de reproducir sus métodos.

En los años recientes (2023-2025), la literatura sobre visión por computadora en el sector agrícola ha investigado tácticas de rebalanceo de clases que combinan técnicas de oversampling con datos sintéticos. En concreto, numerosos estudios indican que los modelos generativos (GANs y sus variantes) tienen la capacidad de incrementar la diversidad morfológica de las lesiones y optimizar la identificación de clases minoritarias, lo cual tiene efectos positivos en el recall/precisión y F1 en cultivos y enfermedades foliares ampliamente investigadas como el tomate (Wang et al., 2025).

Este conjunto de pruebas más recientes en el ámbito agrícola amplía y complementa las revisiones anteriores acerca de GANs en la agroalimentación, destacando la importancia del empleo de métricas de calidad de síntesis (como FID e IS) para verificar la utilidad que tienen los datos producidos en tareas posteriores a la clasificación (Wang et al., 2025).

A diferencia de estudios previos que aplican GANs de manera genérica para aumentar imágenes agrícolas, el aporte central de este trabajo consiste en entrenar una WGAN-GP

exclusivamente con defectos segmentados de enfermedades reales de aguacate (Scab y Anthracnose) y posteriormente insertarlos de forma controlada en frutos sanos. Esta estrategia garantiza que las imágenes sintéticas conserven la textura, iluminación y fondo natural del fruto, al tiempo que diversifican la morfología de las lesiones. Con ello se logra un balanceo más realista y contextualizado del conjunto de datos, evitando las duplicaciones artificiales propias de técnicas clásicas como oversampling o SMOTE. Además, al definir la partición de los datos antes de la generación, se controla la fuga de información entre conjuntos y se asegura la reproducibilidad del pipeline. Este enfoque, complementado con la adaptación explícita de la metodología CRISP-DM y la evaluación con métricas robustas por clase, constituye la principal innovación metodológica frente a propuestas anteriores en el área.

Se propone como hipótesis que la utilización de datos sintéticos producidos por WGAN-GP para el balanceo de clases podría incrementar la habilidad de generalización de modelos que clasifican imágenes agrícolas, en lugar del entrenamiento con conjuntos desbalanceados. El propósito general es crear y analizar un flujo de trabajo que pueda repetirse, incluyendo la extracción de defectos, la producción controlada de imágenes sintéticas y la evaluación del impacto que tiene en el rendimiento de un modelo basado en InceptionV3.

2. Materiales y Métodos

La investigación se desarrolló entre abril y julio de 2025 en la Unidad de Docencia Investigación y Vinculación (UDIV), perteneciente a la Escuela Superior Politécnica Agropecuaria de Manabí “Manuel Félix López” (ESPA MFL), Ecuador. El laboratorio cuenta con estaciones de trabajo equipadas con GPU NVIDIA RTX 3060 (12 GB) y entornos configurados con Python 3.10, TensorFlow 2.x, OpenCV y JupyterLab. El tiempo total de cómputo fue de aproximadamente 120 horas, distribuidas entre preparación de datos, segmentación de defectos, entrenamiento de redes generativas adversarias (WGAN-GP) y entrenamiento de modelos clasificadores mediante aprendizaje por transferencia.

2.1 Conjunto de Datos y Caracterización

El conjunto de datos utilizado corresponde a imágenes de aguacates capturadas en un estudio fotográfico controlado en Morelos, México, durante un único ciclo productivo, disponible en el repositorio de Kaggle llamado “Clasificación de enfermedades del aguacatero” del usuario camposfe1, este es un repositorio abierto a todo público con autorización para que se haga el uso adecuado de esa colección de imágenes. Las imágenes están clasificadas en tres condiciones sanitarias: Healthy (sano), Scab y Anthracnose, con un total de 3983 registros distribuidos en

1764, 1197 y 1022 imágenes respectivamente. Esta distribución presenta un desbalance moderado (proporciones aproximadas del 44%, 30% y 26%), que podría afectar el rendimiento de los modelos de clasificación, particularmente en la identificación de clases minoritarias. También en este mismo repositorio existe una carpeta test, esta carpeta no está clasificada en clases. Esta carpeta se utilizará para comprobar la efectividad del modelo con 100 imágenes (estos datos no se usarán para el entrenamiento será exclusivamente para la validación del modelo).

En el repositorio menciona que las condiciones de captura fueron estandarizadas para minimizar la variabilidad en iluminación, enfoque y encuadre. La caja de estudio utilizada tiene dimensiones de $22 \times 24 \times 23$ cm, con una intensidad lumínica de 450 lm, fondo blanco y colocación central del fruto.

Así mismo también menciona que las imágenes se capturaron con cámara digital en formato .jpg con las siguientes especificaciones: resolución de 4496×3000 píxeles, relación de aspecto 4:3, apertura f/3.8, velocidad de obturación 1/100 s, valor ISO 200, distancia focal 24 mm, compensación de exposición 0 paso y sin uso de flash. La iluminación estable y el fondo uniforme permitieron obtener imágenes de alta calidad.

Para evitar fuga de información hacia el conjunto de validación, la generación de imágenes sintéticas se realizó únicamente a partir de los ejemplos incluidos en el conjunto de entrenamiento, definiendo las particiones antes del entrenamiento de la WGAN-GP. Las imágenes están organizadas según el archivo labels.csv como se muestra en la Tabla 1, que estructura el conjunto de datos en formato multiclase mediante los siguientes atributos: *Identification* (identificador único de cada imagen), *Condition* (etiqueta de clase principal), y tres columnas binarias *Healthy*, *Scab*, *Anthraco* que indican la pertenencia a cada clase.

Tabla 1. Tabla representativa del archivo labels.csv.
Fuente: Los Autores.

Identification	Condition	Healthy	Scab	Anthraco
Avocado 001	Scab	0	1	0
Avocado 002	Scab	0	1	0
Avocado 003	Healthy	1	0	0
Avocado 004	Scab	0	1	0

La enfermedad Scab es causada por el hongo Sphaceloma perseae y se manifiesta mediante lesiones oscuras y corchosas en la superficie del fruto, mientras que la Anthracnose, provocada por Colletotrichum gloeosporioides, se presenta con manchas oscuras que evolucionan a lesiones necróticas profundas, afectando significativamente la calidad del fruto (Bustamante et al., 2022).

2.2 Segmentación con OpenCV

Se implementaron técnicas de segmentación con la biblioteca OpenCV para entrenar una red generativa con el fin de simular efectos visuales realistas de patologías. Para obtener las zonas de las enfermedades presentes en imágenes de aguacate con Scab y Anthracnose, se usaron máscaras morfológicas, algoritmos de umbralización y transformaciones cromáticas dentro de los espacios Lab y HSV.

Se ha comprobado que estas técnicas son particularmente eficaces para dividir lesiones foliares y zonas enfermas a través de umbralización utilizando el canal H del espacio HSV, lo cual mejora la precisión sin necesidad de modelos pesados (Hai et al., 2025). De igual forma, herramientas de procesamiento parecidas se han empleado con éxito para automatizar la cuantificación de los síntomas de patologías y luchar contra las fluctuaciones en la iluminación a través de segmentación binaria adaptativa (Paauw et al., 2024).

2.3 Arquitectura del Modelo Generativo (WGAN-GP)

La WGAN-GP no realiza una tarea supervisada ni se evalúa con métricas de clasificación; su finalidad es aprender la distribución morfológica de los defectos segmentados (máscaras binarias de Scab/Anthracnose) y sintetizar nuevas instancias de defecto. Por esta razón, no se definió un particionado train/val/test para la GAN: su objetivo no es “generalizar” a datos no vistos, sino producir parches de lesión plausibles que luego se insertan sobre frutos sanos.

La arquitectura de la red generativa utilizada fue una WGAN como se muestra en la Figura 1 con penalización de gradiente (WGAN-GP), compuesta por:

- **Generador:** red secuencial con capas Dense, BatchNormalization, LeakyReLU y Conv2DTranspose, finalizando con activación tanh.
- **Discriminador (Crítico):** arquitectura espejo inversa con capas Conv2D, LeakyReLU y Dropout, sin activación final (por propiedad de la función de pérdida Wasserstein).

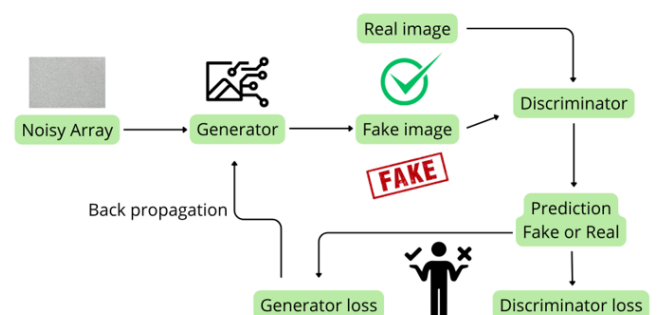


Figura 1. Diagrama de la arquitectura WGAN-GP mostrando el flujo de datos y retroalimentación.
Fuente: Los Autores.

Se emplearon las máscaras obtenidas con OpenCV como fundamento para entrenar una red generativa adversarial de Wasserstein (WGAN) que se enfoca únicamente en la creación de imperfecciones (necrosis, manchas, escoriaciones). La red, después de ser entrenada, posibilitó la síntesis de nuevos defectos, que se superpusieron posteriormente en imágenes aleatorias de aguacates sanos. De este modo, se generó una representación sintética de frutas enfermas. Esta estrategia posibilitó que la representación de las clases minoritarias crezca de forma controlada, sin modificar la distribución de la clase saludable (Salem et al., 2020).

HiperParámetros de la WGAN-GP

Para garantizar la reproducibilidad del modelo generativo, en la Tabla 2 se resumen los principales hiper parámetros utilizados en el entrenamiento de la WGAN-GP. Estos valores fueron seleccionados siguiendo las configuraciones recomendadas en la literatura y ajustados de forma preliminar mediante pruebas exploratorias.

Tabla 2. HiperParámetros utilizados en la WGAN-GP.
Fuente: Los Autores.

Parámetro	Valor
Dimensión del vector de ruido (z)	100
Optimizador	Adam ($\beta_1=0.5$, $\beta_2=0.9$)
Tasa de Aprendizaje	0.0001
Batch size	32
Épocas	5000
Penalización de gradiente (λ)	10
Número de actualizaciones del crítico (n_critic)	5
Función de activación (generador)	tanh
Función de activación (crítico)	LeakyReLU (sin capa final)
Generador	Dense($4 \times 4 \times 256$) $\rightarrow$ BN $\rightarrow$ LeakyReLU $\rightarrow$ Conv2DTrans(128, k=4, s=2) $\rightarrow$ BN $\rightarrow$ LeakyReLU $\rightarrow$ Conv2DTrans(64, k=4, s=2) $\rightarrow$ BN $\rightarrow$ LeakyReLU $\rightarrow$ Conv2DTrans(3, k=3, s=1, act=tanh) (ajusta canales/tamaño real)
Crítico	Conv2D(64, k=4, s=2) $\rightarrow$ LeakyReLU $\rightarrow$ Conv2D(128, k=4, s=2) $\rightarrow$ LeakyReLU $\rightarrow$ Conv2D(256, k=4, s=2) $\rightarrow$ Flatten $\rightarrow$ Dense(1)
Selección de checkpoint	Se guardó cada 200 épocas; y se eligió el mejor modelo manualmente (el modelo que fue más regular generando los defectos sintéticos).

Política de Partición de Datos

En el caso particular del entrenamiento de la WGAN-GP, no se implementó una política de partición tradicional en conjuntos de entrenamiento, validación y prueba, como suele hacerse en tareas de aprendizaje supervisado. El objetivo de este modelo generativo no es medir su rendimiento en datos no vistos, sino aprender la distribución estadística de los defectos segmentados de aguacate para poder sintetizar nuevas instancias realistas. En consecuencia, todas las imágenes disponibles de defectos fueron utilizadas para el entrenamiento, aplicando previamente técnicas de aumento de datos (rotaciones, variación de brillo y espejado).

De esta manera, la “fuga de información” entendida como el uso indebido de datos de validación o prueba en el entrenamiento no representa un riesgo en esta etapa, dado que la WGAN-GP no se evalúa mediante métricas de clasificación convencionales. La validación del enfoque se llevó a cabo posteriormente, en la fase de clasificación con Inception V3, donde sí se aplicó una división explícita de los datos en entrenamiento, validación y prueba para garantizar la independencia entre los conjuntos y asegurar la reproducibilidad del experimento.

A su vez se logró controlar qué frutos del mismo origen no aparezcan en conjuntos distintos realizando una partición por espécime controlando los datos mediante carpetas separadas.

2.4 Clasificación con Red Convolutacional InceptionV3

Arquitectura del modelo

Para analizar el efecto de equilibrar clases a través de imágenes sintéticas generadas, se utilizó una red neuronal convolutacional con la metodología de transfer learning, que tiene como fundamento la arquitectura InceptionV3 y fue entrenada previamente con el conjunto de datos ImageNet. Esta arquitectura profunda ha demostrado ser muy eficaz en la clasificación de imágenes naturales, lo que posibilita aprovechar el conocimiento previamente adquirido y agilizar el proceso de entrenamiento, además de optimizar el rendimiento (Hossen et al., 2025).

Para prevenir el sobreajuste, se implementó la técnica EarlyStopping con patience de 5 y restore_best_weights=True en la configuración. Esto posibilitó parar el entrenamiento antes de tiempo. En la situación con datos desbalanceados, esta estrategia fue crucial porque el entrenamiento se detuvo sin intervención humana en la época 24. Esto indica que el modelo llegó a su mejor rendimiento antes de finalizar las 50 épocas programadas, lo cual optimizó el tiempo de cómputo y mantuvo la capacidad de generalización.

La arquitectura del modelo base de InceptionV3 se muestra en la Figura 2.

Fase 1 – Transfer Learning:

- Modelo base: InceptionV3 preentrenado en ImageNet.
- Congelación de capas: todas las capas convolucionales del modelo base fueron congeladas.
- Optimizador: Adam (tasa de aprendizaje por defecto).
- Función de pérdida: *categorical_crossentropy*.
- *EarlyStopping*: habilitado con *patience=5* y *restore_best_weights=True*.

Fase 2 – Fine-Tuning:

- Se descongelaron las últimas 50 capas del modelo base.
- Épocas: 50 (Programadas).
- *EarlyStopping* mantenido.

Una capa final de salida *Dense* (3, *activation='softmax'*) para realizar la clasificación multiclase en tres categorías: *Healthy*, *Scab* y *Anthracoze*.

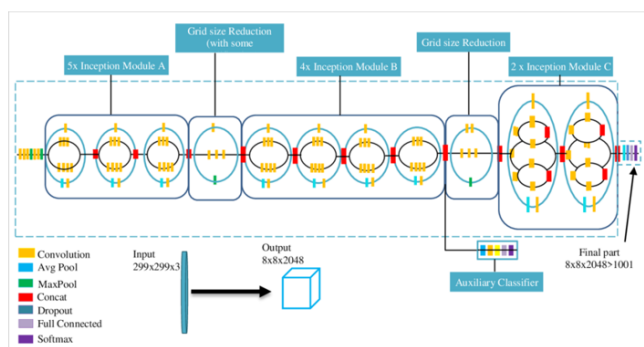


Figura 2. Estructura completa de la red InceptionV3 con sus bloques de inception y clasificador auxiliar.

Fuente: (Iparraguirre-Villanueva et al., 2022)

Configuración de entrenamiento

El modelo final fue compilado con los siguientes parámetros:

- Optimizador: Adam.
- Función de pérdida: *categorical_crossentropy*, adecuada para clasificación multiclase.
- Métrica: Precisión (*accuracy*).

El entrenamiento se realizó durante 50 épocas programadas

con un tamaño de *batch* de 32, utilizando imágenes de tamaño 224x224 píxeles y aceleración mediante GPU.

Preprocesamiento de los datos

Se utilizó la clase *ImageDataGenerator* para normalizar las imágenes (*rescale=1./255*) y dividir automáticamente el conjunto de datos en:

- Entrenamiento: 80% de las imágenes.
- Validación: 20% de las imágenes.

Las imágenes utilizadas provenían de la carpeta datos-balanceados y datos-desbalanceados, la cuales contienen tres subdirectorios representando las clases correspondientes. Se observó el comportamiento del modelo y, durante el entrenamiento, se visualizaron las curvas de pérdida y precisión para las dos fases (entrenamiento y validación), lo cual permitió identificar potenciales problemas como el sobreajuste.

Finalmente, los modelos entrenados fueron almacenados para futuras evaluaciones o integraciones en sistemas de diagnóstico automatizado.

La división en entrenamiento (80%) y validación (20%) se realizó mediante el parámetro *validation_split* de la clase *ImageDataGenerator* en Keras/TensorFlow, lo cual implica que la separación fue aleatoria a nivel global y no estratificada por clase. En consecuencia, las proporciones exactas de cada clase en entrenamiento y validación pueden variar ligeramente respecto a la distribución original del conjunto. Esta estrategia fue seleccionada debido a que:

- Permite mantener un flujo automatizado y reproducible en Keras sin necesidad de particiones manuales.
- Evita posibles sesgos introducidos por una estratificación rígida, dado que el objetivo principal del estudio era evaluar el impacto global del balanceo con imágenes sintéticas.
- Se complementa con un conjunto de prueba completamente independiente (100 imágenes de la carpeta test del repositorio Kaggle), lo que asegura la evaluación fuera del dominio de entrenamiento y validación.

2.5 Diseño Experimental y Comparación

Para evaluar el impacto del balanceo conjunto de datos mediante imágenes sintéticas, se diseñó un experimento comparativo empleando una red neuronal convolucional con arquitectura InceptionV3, entrenada bajo dos condiciones distintas:

- Conjunto de datos original desbalanceado.
- Conjunto de datos balanceado mediante imágenes generadas con WGAN.

Ambos modelos fueron entrenados utilizando los mismos parámetros de arquitectura, optimización y preprocesamiento, lo que garantizó condiciones experimentales equivalentes y permitió una comparación objetiva del desempeño.

Métrica de evaluación

En esta investigación, la métrica principal utilizada para evaluar la eficiencia del modelo de clasificación fue la precisión (accuracy), dado que el objetivo principal consistía en comparar el rendimiento entre un conjunto de datos desbalanceado y otro balanceado con imágenes sintéticas. Adicionalmente, para reforzar la validez del enfoque y confirmar la capacidad de generalización, se consideró la evaluación en un conjunto de prueba independiente no visto durante el entrenamiento ni la validación, reportando métricas de desempeño por clase (precisión, recall y F1-score).

2.5 Arquitectura General del Flujo Experimental

El flujo general del sistema se estructuró en las siguientes etapas que se puede ver en la Figura 3. Este diseño garantiza la replicabilidad del experimento y puede adaptarse a otros cultivos o enfermedades agrícolas.

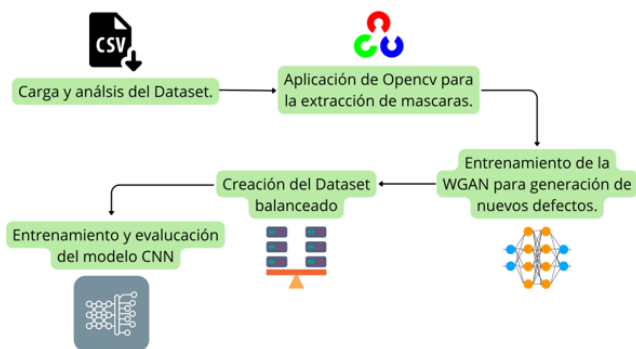


Figura 3. Flujo de trabajo completo del sistema: desde carga de datos hasta entrenamiento de modelos.

Fuente: Los Autores.

El código, las configuraciones y los archivos de división (CSV) se encuentran disponibles en <https://github.com/JesusMe001/Balanceo-de-Conjuntos-de-Datos-Basados-en-Redes-Generativas-Aplicado-a-Im-genes-del-Sector-Agr-cola.git>, incluyendo ejemplos representativos de máscaras y sintéticos. Se documentan versiones de librerías y modelos ya generados listos para usarse.

3. Resultados y Discusión

3.1 Preparación del Conjunto de Datos y Segmentación de Defectos

El procedimiento inicial consistió en clasificar el conjunto de datos original de imágenes de aguacates en tres carpetas, cada una de las cuales correspondía a sus respectivas clases: Scab, Anthracnose y Healthy. Esta estructura facilita un flujo automatizado para cargar y procesar imágenes, ya sea para la clasificación o para las labores de segmentación.

Para aislar solamente las zonas afectadas por enfermedades en las imágenes que se han marcado como Scab y Anthracnose, se llevó a cabo un proceso personalizado de extracción de defectos, empleando la biblioteca OpenCV. Este procedimiento fue esencial para crear un conjunto de datos realista y enfocado, el cual se usará como entrada para entrenar el modelo generativo WGAN.

El procedimiento incluyó los siguientes pasos:

Conversión a espacio de color HSV: para facilitar la segmentación según tonalidades específicas del fruto y su fondo.

Definición de máscaras de exclusión:

- Se creó una máscara para eliminar el fondo blanco (valores altos en el canal V).
- Se definió otra máscara para excluir el verde sano del aguacate (tonos entre 35° y 90° en el canal H).

Combinación de máscaras: se unieron ambas zonas a excluir y se invirtió la máscara, conservando únicamente los elementos que no son ni blancos ni verdes, es decir, las áreas enfermas.

Limpieza morfológica: se aplicó una operación de apertura para reducir ruido y mejorar la definición del área segmentada.

Extracción final del defecto: se aplicó la máscara obtenida sobre la imagen original, obteniendo una representación aislada del defecto, lista para ser utilizada en el entrenamiento de la red generativa.

El resultado de este procedimiento fue un conjunto de máscaras binarias limpias que destacan únicamente las zonas de enfermedad, y una serie de imágenes que contienen sólo los efectos patológicos extraídos del aguacate cabe recalcar que por cada aguacate enfermo se extrajo una máscara como se destaca en la Figura 4 y Figura 5.

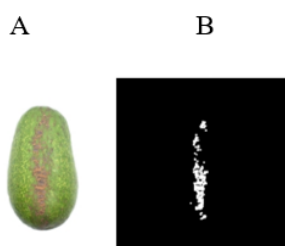


Figura 4. Resultado del algoritmo OpenCV Scab para detección de Scab: (a) Fruto original; (b) Región segmentada.
Fuente: Los Autores.



Figura 5. Resultado del algoritmo OpenCV Scab para detección de Anthracnose: (a) Fruto original; (b) Región segmentada.
Fuente: Los Autores.

Esta perspectiva permitió, además de mantener las propiedades naturales del aguacate saludable, registrar con exactitud las anomalías inherentes a las enfermedades foliares. Posteriormente, estos efectos aislados se usaron como base de datos para entrenar la red generativa, con el objetivo de replicar esos patrones en imágenes sintéticas nuevas.

3.2 Generación de Defectos Sintéticos con WGAN

Después de conseguir las máscaras que reflejan únicamente los efectos patológicos de las enfermedades Anthracnose y Scab, se entrenó una red generativa adversaria con el propósito de crear nuevos defectos que replican visualmente las lesiones vistas en los frutos reales. A este fin, se utilizó una Wasserstein GAN con penalización de gradiente (WGAN-GP) porque tiene una habilidad superior para estabilizar el proceso de capacitación y producir imágenes más verosímiles.

Aumento de datos previo al entrenamiento

Las imágenes de los defectos extraídos se sometieron a una etapa de aumento de datos (data augmentation) para ampliar la diversidad y solidez del conjunto de entrenamiento previo. Este procedimiento comprendió:

- Rotaciones en 0°, 90°, 180° y 270°.
- Variación aleatoria de brillo.
- Espejado horizontal.

Cada imagen resultante fue redimensionada a 128×128 píxeles y normalizada al rango [-1, 1], lo que resultó en un conjunto suficientemente amplio para entrenar el generador.

Arquitectura del modelo

El modelo generativo fue compuesto por:

- Un generador, diseñado con capas Dense, Reshape, y múltiples Conv2DTranspose con activación tanh, el cual recibe vectores de ruido de dimensión 100 como entrada.
- Un crítico (discriminador) que actúa como evaluador del realismo, construido con capas Conv2D, LeakyReLU y Flatten, sin activación final para cumplir con la función de pérdida Wasserstein.

Ambos modelos fueron optimizados con Adam (learning rate 0.0001), configurado con los parámetros recomendados para WGAN-GP (beta_1=0.5, beta_2=0.9), y se implementó la penalización de gradiente con $\lambda = 10$ para reforzar la estabilidad en la retropropagación.

Proceso de entrenamiento

El entrenamiento se llevó a cabo durante 5000 épocas, y cada 200 épocas se guardó el modelo con el fin de conservar aquel que ofreciera los mejores resultados, utilizando un tamaño de batch de 32 imágenes. En cada iteración:

- El crítico fue entrenado 5 veces por cada actualización del generador, como sugiere el algoritmo WGAN-GP.
- Se calculó la penalización de gradiente entre muestras reales y generadas, lo que aseguró una convergencia más estable.

A lo largo del entrenamiento se registraron pérdidas en ambos modelos como se muestra en la Figura 6.

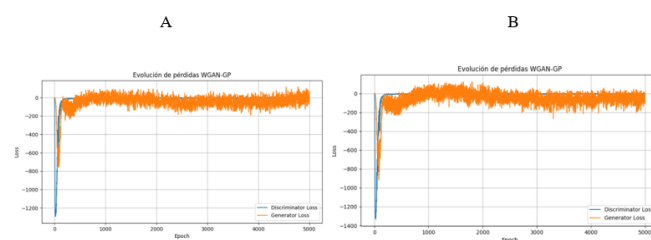


Figura 6. Evolución de pérdidas WGAN-GP. (a) Evolución Scab; (b) Evolución Anthracnose.
Fuente: Los Autores.

Muestras generadas

Cada cierto número de épocas, se generaron y guardaron imágenes sintéticas de prueba utilizando vectores aleatorios (noise) como entrada al generador entrenado. Estas imágenes fueron visualmente inspeccionadas para verificar la calidad de los defectos generados como se muestra en la Figura 7.

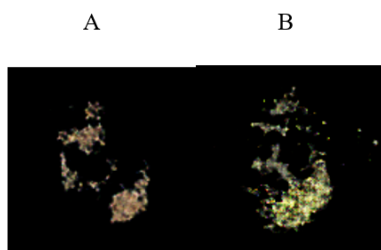


Figura 7. Muestras sintéticas de defectos generados por la WGAN-GP. (a) Defecto Scab; (b) Defecto Anthracnose.
Fuente: Los Autores.

Después de entrenar la WGAN, se produjeron 567 defectos sintéticos nuevos para Scab y 742 para Anthracnose. Se nota en las muestras una coherencia visual con los defectos auténticos obtenidos de las imágenes originales. Esto confirma la habilidad del generador para aprender y reproducir los atributos cromáticos y morfológicos característicos de las enfermedades Scab y Anthracnose.

3.3 Síntesis de Imágenes Balanceadas e Inserción de Defectos

Una vez completado el entrenamiento del modelo generativo WGAN-GP, y tras validar la calidad visual de los defectos sintetizados, se procedió a integrarlos dentro de imágenes de aguacates sanos con el fin de construir un conjunto balanceado de clases.

El procedimiento consistió en superponer los efectos patológicos causados en frutas sanas, de modo que se imita la apariencia verdadera de las frutas infectadas por *Scab* y *Anthracnose*. La inserción se llevó a cabo de un modo visualmente realista, teniendo en cuenta la orientación, la proporción y la colorimetría del fruto base. Este método posibilitó incrementar de manera artificial la clase minoritaria sin duplicar imágenes originales ni modificar la clase *Healthy*.

Balanceo del Conjunto de Datos

Con el fin de igualar la cantidad de muestras entre clases, se seleccionó una cantidad de 742 imágenes sintéticas para complementar la clase *Anthracnose-E1* y 547 imágenes sintéticas *Scab-E2* respectivamente las cuales se encontraban originalmente subrepresentadas. Las nuevas imágenes fueron distribuidas en sus carpetas correspondientes a su clase *data/Anthracnose-E1* y *data/Scab-E2* para que puedan ser procesadas como parte del flujo estándar de entrenamiento de modelos.

Esta estrategia de inserción permitió transformar un conjunto de

datos originalmente desbalanceado en uno equilibrado en las tres clases, lo cual es esencial para evitar el sesgo de predicción hacia clases mayoritarias durante el entrenamiento del clasificador, en las Figura 8 y Figura 9 se ilustra el proceso de construcción de una muestra sintética:

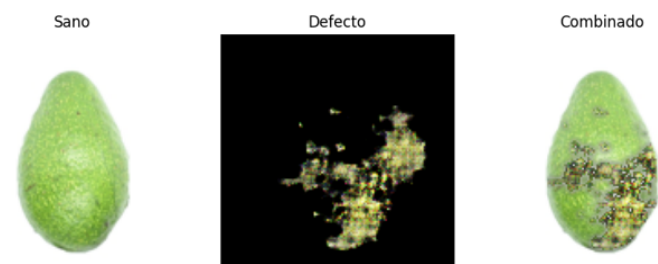


Figura 8. Generación de muestras sintéticas de Anthracnose: fruto original, defecto extraído y resultado combinado.
Fuente: Los Autores.

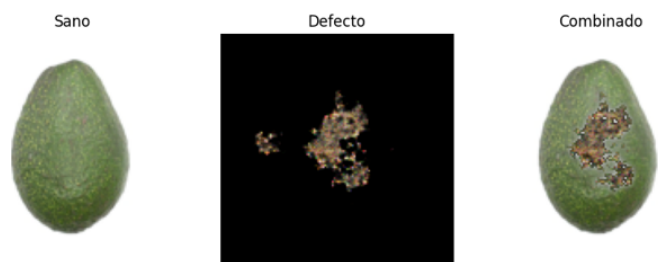


Figura 9. Generación de muestras sintéticas de Scab: fruto original, defecto extraído y resultado combinado.
Fuente: Los Autores.

Este procedimiento no solo logró un equilibrio numérico en el conjunto de datos, sino que además amplió la diversidad morfológica de los ejemplos, lo cual le proporcionó al clasificador final una mayor capacidad para generalizar cuando se enfrenta a nuevas muestras.

3.4 Comparación de Resultados: Conjunto de Datos Desbalanceados y Balanceados con WGAN

Se entrenaron modelos de clasificación con la misma estructura (InceptionV3 + top model personalizado) en dos contextos diferentes: uno que empleaba el conjunto de datos original desequilibrado y otro que lo hacía con el conjunto equilibrado por medio de la inserción de fallas sintéticas. Esto se hizo para verificar el efecto de balancear los datos conjuntos a través de imágenes sintéticas creadas por la WGAN.

Cabe señalar que los tamaños reportados (3188/795 en el escenario original y 4236/1056 en el balanceado) provienen de aplicar la política de división 80/20 descrita en la sección de Métodos. Esta división fue realizada de manera aleatoria a nivel global, por lo que los conteos específicos por clase pueden variar ligeramente entre las particiones. En la Tabla 3 se resumen los totales de imágenes utilizados en cada escenario, incluyendo el conjunto de prueba independiente de 100 imágenes

Tabla 3. Distribución global de imágenes por partición en los escenarios original y balanceado.
Fuente: Los Autores.

Escenario	Entrenamiento	Validación	Total (train+val)	Prueba
Original	3188	795	3983	-
Balanceado	4236	1056	5292	100

Desempeño con Conjunto de Datos Desbalanceado

El entrenamiento con el conjunto original, compuesto por 3188 imágenes de entrenamiento y 795 de validación, mostró un incremento progresivo en la precisión hasta aproximadamente la época 24, momento en el que se alcanzó el mejor valor de validación (83,43 %). A partir de este punto, el modelo se estabilizó y se aplicó EarlyStopping para evitar sobreajuste. La fluctuación en las curvas de pérdida y precisión indica que la distribución desigual de clases influyó en la capacidad de generalización del modelo. Este comportamiento puede observarse en la Figura 10.

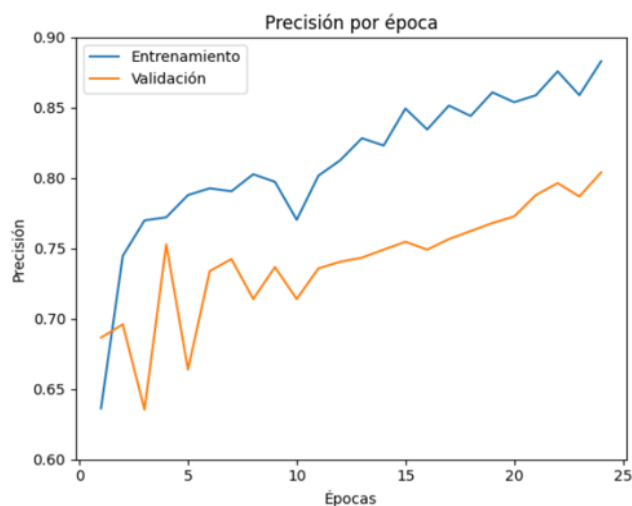


Figura 10. Desempeño con el Conjunto de Datos Desbalanceado.
Fuente: Los Autores.

Desempeño con Conjunto de Datos Balanceado (WGAN)

El modelo que fue entrenado con datos sintéticos compuesto por 4236 imágenes de entrenamiento y 1056 de validación producidos a través de WGAN-GP logró un 97,74 % de precisión

en la validación durante la época 50, en comparación con el 83,43% adquirido en el escenario desequilibrado. Esta mejora de más de 14 puntos porcentuales se debe especialmente a dos factores:

- Disminución de la tendencia hacia la clase mayoritaria: El equilibrio numérico entre las tres clases redujo el sesgo de predicción que se notaba en la situación inicial, lo que permitió al modelo adquirir patrones más representativos de las clases minoritarias.
- Mayor diversidad morfológica: Los fallos producidos por la WGAN-GP aumentaron la gama de formas, texturas y ubicaciones de las lesiones, lo cual mejoró la habilidad del modelo para generalizar en el grupo de validación.

A pesar de que la curva de entrenamiento demostró una baja constante de la pérdida y niveles altos de precisión en validación, el aumento notado debía ser analizado con precaución, pues podría indicar un ajuste a las particularidades propias de los datos sintéticos. Se llevó a cabo un análisis adicional sobre un conjunto de prueba independiente de 100 imágenes que no se habían examinado antes, ni en la validación ni durante el entrenamiento, para abordar esta inquietud.

En este examen, el modelo demostró una notable capacidad de generalización, logrando un macro F1-score de 0.90 y una precisión global del 91%. La matriz de confusión mostró un desempeño ideal en la categoría Healthy (con una precisión y un recall del 100 %), un alto recall en Anthracnose (0,94) y un rendimiento algo inferior en Scab (con 0,92 de precisión y 0,76 de recall), debido a ciertas confusiones con Anthracnose tal como se muestra en la Tabla 4.

Tabla 4. Resultados de clasificación en conjunto de prueba independiente (100 imágenes).
Fuente: Los Autores.

Clase	Precisión	Recall	F1-Score	Soporte
Scab	0.92	0.76	0.83	29
Anthracnose	0.83	0.94	0.88	36
Healthy	1.00	1.00	1.00	35
Promedio Macro	0.92	0.90	0.90	100
Promedio Ponderado	0.91	0.91	0.91	100
Accuracy Global	-	-	0.91	100

Estos hallazgos demuestran la efectividad de la estrategia de balanceo por medio de WGAN-GP, como se ilustra en la Figura 11; no obstante, también indican áreas para mejorar, en particular



en lo que respecta a distinguir entre clases con patrones visuales parecidos. Para disminuir la confusión y examinar métodos de clasificación fina (fine-grained classification) que posibiliten diferenciar entre características sutiles de enfermedades que se ven de manera parecida, se aconseja mejorar la producción de defectos en investigaciones futuras.

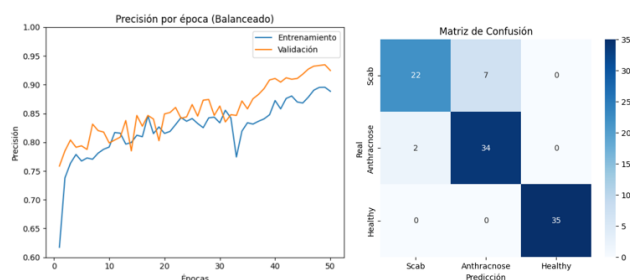


Figura 11. Evaluación del modelo con dataset balanceado: curvas de precisión por época y matriz de confusión final.

Fuente: Los Autores.

Mejores Épocas de Cada Escenario

Tabla 5. Épocas en cada escenario.

Fuente: Los Autores.

Conjunto de Datos	Mejor Precisión Validación	Época	Pérdida Validación	Tamaño de validación
Desbalanceado	83.43%	24	0.4216	795
Balanceado (WGAN)	97.74%	50	0.0704	1056

Validación de la hipótesis mediante bootstrapping

Para evaluar la significancia estadística de la diferencia observada, se aplicó un procedimiento de remuestreo bootstrapping (10 000 iteraciones) sobre las predicciones de validación de ambos escenarios. El análisis arrojó una diferencia media en accuracy de 0,1402 (14,02 puntos porcentuales), con un intervalo de confianza del 95 % de [0,1134, 0,1679] como se muestra en la Tabla 6.

Dado que el límite inferior del intervalo es mayor que cero, se concluye que la mejora es estadísticamente significativa y no atribuible al azar como se observa en la Figura 12.

Si bien este resultado respalda la hipótesis de que el balanceo de clases mediante WGAN-GP mejora el rendimiento del modelo,

Tabla 6. Resultados del bootstrapping sobre la exactitud del modelo en el conjunto de prueba.

Fuente: Los Autores.

Parámetro	Época
Total de imágenes	100
Accuracy observado	0.91
Iteraciones (bootstrap)	10000
Semilla	2025
IC 95 % (percentil 2.5–97.5)	[0.85, 0.96]

se reconoce que el bootstrapping no evalúa posibles sesgos derivados de la generación de datos. Por ello, en trabajos futuros se recomienda complementar este análisis con pruebas como McNemar, validación cruzada repetida o evaluación en bases de datos externas, para confirmar la robustez del modelo y descartar efectos de sobreajuste.

Descripción del procedimiento:

- Se crean vectores binarios que representan aciertos y errores en cada escenario.
- Se realiza un remuestreo aleatorio con reemplazo (bootstrapping) para cada conjunto.
- En cada iteración se calcula la diferencia en accuracy.
- Se construye un intervalo de confianza al 95 % a partir de las diferencias obtenidas.
- Si el límite inferior del intervalo es mayor que cero, la mejora se considera estadísticamente significativa

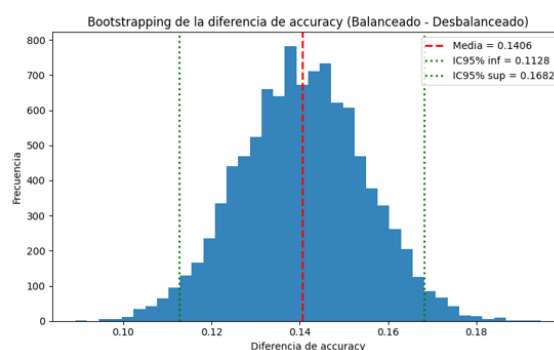


Figura 12. Análisis estadístico bootstrapping: distribución de diferencias de accuracy entre datasets balanceado y desbalanceado.

Fuente: Los Autores.

Discusión

Los resultados obtenidos respaldan la hipótesis de que el uso de datos sintéticos generados mediante WGAN-GP puede mejorar el rendimiento de modelos de clasificación de imágenes agrícolas. Para evaluar la significancia estadística de la diferencia observada entre el modelo entrenado con el conjunto original y el modelo entrenado con el conjunto balanceado, se aplicó un procedimiento de remuestreo bootstrapping (10 000 iteraciones) implementado en Python. Este análisis arrojó una diferencia media en accuracy de 0,1402 (14,02 puntos porcentuales), con un intervalo de confianza del 95 % de [0,1134, 0,1679], lo que indica que la mejora es estadísticamente significativa y no atribuible al azar.

En el escenario inicial desequilibrado, el modelo llegó a una precisión de validación del 83,43 % en la época 24. Esto ocurrió debido a que se implementó EarlyStopping con una paciencia de 5 épocas, ya que las métricas se estabilizaron. Las fluctuaciones que se aprecian en la curva de validación indican problemas de generalización vinculados con el desbalance. Esto es consistente con lo que Bi et al. (2020) informaron: detectaron aumentos significativos de precisión al emplear WGAN-GP para equilibrar conjuntos de datos agrícolas.

Por otro lado, el modelo que fue entrenado con el conjunto balanceado a través del WGAN-GP logró un 97,74 % de precisión en la validación al llegar a la época número 50, además de una pérdida baja (0,0704). La disminución del sesgo hacia la clase mayoritaria y la inclusión de una diversidad morfológica más amplia en las clases minoritarias pueden ser responsables de este incremento. No obstante, es importante interpretar con precaución las mejoras de esta magnitud, pues podrían ser el resultado de un ajuste a los rasgos particulares de los datos sintéticos. Por esta razón, se aconseja verificar en conjuntos externos e independientes para corroborar la habilidad de generalización.

Estos resultados son coherentes con lo que se ha documentado en la literatura: desde 2017, las GAN han ayudado a incrementar el desempeño de modelos de visión por computadora en labores agrícolas, entre ellas la detección poscosecha, el fenotipado y la clasificación de enfermedades (Carvajal Chávez, 2023; Lu et al., 2022). Investigaciones como las de Bi et al. (2020) han demostrado que la estabilidad y el realismo de las imágenes generadas se incrementan con el uso de variantes como WGAN-GP, que se combinan con técnicas de regularización como cGAN-LSR, logrando así un aumento en la precisión sin menoscabar la robustez.

En resumen, los hallazgos corroboran que WGAN-GP tiene el potencial de ser un instrumento útil para optimizar la actuación de modelos clasificadores en situaciones con clases minoritarias. Sin embargo, para fortalecer esta evidencia, se aconseja que investigaciones posteriores incorporen pruebas de McNemar, validaciones cruzadas que se realicen varias veces y métricas de similitud perceptual (IS, FID). Esto se hace con el objetivo de

evitar sobreajuste y asegurar que el modelo sea generalizable.

Comparación con métodos clásicos de balanceo

Frente a estrategias tradicionales como el oversampling aleatorio o SMOTE, que duplican o interpolan ejemplos en el espacio de características y pueden inducir sobreajuste o generar instancias poco plausibles a nivel visual, el enfoque propuesto actúa directamente sobre la representación visual del fenómeno patológico. La inserción controlada de defectos segmentados/generados sobre frutos sanos preserva el contexto de iluminación, textura y fondo, a la vez que incrementa la variabilidad morfológica de las lesiones. En contraste, class weights únicamente pondera/ajusta la función de pérdida sin corregir el sesgo en la distribución de muestras. Por su parte, técnicas de mezcla como CutMix, MixUp o Copy-Paste combinan regiones de forma genérica; nuestro pipeline, en cambio, restringe los parches a regiones patológicas plausibles y regula su escala, orientación y colorimetría para mantener coherencia semántica y fotométrica. Como trabajo futuro, una comparación ablation directa contra estas alternativas permitiría cuantificar el beneficio marginal de cada componente del pipeline.

Calidad y diversidad de los datos sintéticos

Aunque la validación principal se centra en el impacto sobre el clasificador, resulta pertinente auditar los sintéticos con métricas perceptuales como FID (Fréchet Inception Distance) e IS (Inception Score), a nivel global y por clase. El uso de FID/IS ayudaría a detectar colapso de modos, texturas repetitivas o artefactos, y a establecer umbrales mínimos de aceptación antes de incorporar nuevas tandas al entrenamiento.

Aplicabilidad en condiciones reales de campo

Dado que el conjunto base proviene de un entorno controlado, la transferencia a campo donde cambian iluminación, fondo y oclusiones requiere medidas adicionales: (i) aplicar aumentos fotométricos y geométricos más agresivos durante la síntesis e inserción (variaciones de temperatura de color, sombras, desenfoque, rotaciones y escalas); (ii) introducir domain randomization al posicionar los defectos (rango de tamaños, múltiples lesiones por fruto, ligeras perturbaciones de tono/saturación para compatibilizar con el fondo); y (iii) realizar evaluación externa con imágenes de otras cámaras y locaciones, seguida de un ajuste fino del clasificador con un subconjunto anotado en campo. La confusión observada entre *Scab* y *Anthraco* sugiere explorar enfoques de fine-grained classification y mecanismos de atención local centrados en bordes y texturas de la lesión para mejorar la discriminación en escenarios reales.

4. Conclusiones

Esta investigación muestra cómo las redes generativas adversarias, en particular una WGAN que penaliza el gradiente, tienen el potencial de transformarse en un instrumento fundamental para

encarar un desafío común pero complejo: la desproporción entre clases dentro de conjuntos de datos agrícolas. Su puesta en práctica no solamente permitió que la distribución de datos se equilibrara, sino que además los modelos de clasificación basados en aprendizaje profundo mejoraron significativamente su rendimiento. Se mantuvo la coherencia visual global al producir imágenes sintéticas que reflejan con realismo los defectos típicos de enfermedades como el Scab y Anthracnose, lo que resultó en un incremento constante de la estabilidad y precisión del modelo, a su vez un proceso de bootstrapping permitió una firme validación estadística, la cual apoyó este progreso.

La aportación primordial de esta investigación es proporcionar una estrategia que se puede replicar, automatizar y adaptar para enriquecer datos en situaciones donde no hay suficientes muestras disponibles. Este método, a diferencia de las técnicas tradicionales de aumento, no solo homologa la cantidad de ejemplos por clase, sino que también diversifica la representación morfológica, lo cual posibilita que el modelo aprenda patrones más sólidos y transferibles.

Sin embargo, el estudio tiene limitaciones que deben tenerse en cuenta:

- La variabilidad de las condiciones ambientales y de cultivo se ve limitada porque el conjunto de datos proviene de un solo ciclo productivo y una sola localización geográfica, que es México.
- La precisión del clasificador final fue el foco de la evaluación de la WGAN, y no se incluyeron indicadores específicos de similitud perceptual (por ejemplo, IS o FID) que posibiliten medir la diversidad y calidad de las imágenes sintéticas.

Estos descubrimientos son significativos para los desarrolladores e investigadores que se dedican a las imágenes en el ámbito agrícola, particularmente en contextos de agricultura de precisión donde adquirir datos puede ser caro o imposible. Investigaciones futuras podrían ampliar esta metodología a otras enfermedades o cultivos, incluir evaluaciones con conjuntos de datos independientes y emplear métodos avanzados para validar las imágenes producidas. Además, podrían investigar técnicas de regularización que disminuyan el peligro de sobreajuste y arquitecturas generativas más novedosas.

Agradecimientos

Agradecemos a la ESPAM MFL por su apoyo técnico y académico a través del proyecto institucional Aprendizaje automático y su comportamiento con distintos tipos de conjuntos de datos agropecuarios CUP:91880000.0000.388091 SENPLADES, Ecuador, así como a los responsables del conjunto de datos

utilizados, cuya disponibilidad fue clave para el desarrollo de esta investigación.

Contribución de los autores

Luis Jesús Montesdeoca Espinoza: Conceptualización, Curación de datos, Análisis formal, Metodología, Software, Investigación, Visualización, Redacción – borrador original del artículo, Redacción – revisión y edición del artículo. **Stalin Joel Zambrano Rojas:** Conceptualización, Curación de datos, Análisis formal, Metodología, Software, Investigación, Visualización, Redacción – borrador original del artículo, Redacción – revisión y edición del artículo. **Víctor Joel Pinargote Bravo:** Supervisión, Administración del proyecto, Validación, Redacción – revisión y edición del artículo. **Luis Cristobal Cedeño Valarezo:** Supervisión, Administración del proyecto, Validación, Redacción – revisión y edición del artículo.

Conflicto de interés

Los autores no tienen conflictos de intereses.

Referencias bibliográficas

- Bi, W., Wang, X., Xu, W., & Wang, C. (2020). Generation of synthetic training data for deep learning-based plant disease detection system. *Frontiers in Plant Science*, 11, 583438. <https://doi.org/10.3389/fpls.2020.583438>
- Bustamante, M. I., Osorio-Navarro, C., Fernández, Y., Bourret, T. B., Zamorano, A., & Henríquez-Sáez, J. L. (2022). First record of Colletotrichum anthracnose causing anthracnose on avocado fruits in Chile. *Pathogens*, 11(10), 1204. <https://doi.org/10.3390/pathogens11101204>
- Carvajal Chávez, C. A. (2023). Uso de técnicas como la regresión y redes neuronales para anticipar el rendimiento del maíz. *Recimundo*, 8(4), 126–135. [https://doi.org/10.26820/recimundo/8.\(4\).diciembre.2024.126-135](https://doi.org/10.26820/recimundo/8.(4).diciembre.2024.126-135)
- Food and Agriculture Organization of the United Nations. (2021). The state of food and agriculture 2021: Making agrifood systems more resilient to shocks and stresses. FAO. <https://doi.org/10.4060/cb4476en>
- Hai, T., Shao, Y., Zhang, X., Yuan, G., Jia, R., Fu, Z., Wu, X., Ge, X., Song, Y., Dong, M., & Yan, S. (2025). An efficient model for leafy vegetable disease detection and segmentation based on few-shot learning framework and prototype attention mechanism. *Plants*, 14(5), 760.

- <https://doi.org/10.3390/plants14050760>
- Hossen, M. I., Awrangjeb, M., Pan, S., & Al Mamun, A. (2025). Transfer learning in agriculture: A review. *Artificial Intelligence Review*, 58, Article 97. <https://doi.org/10.1007/s10462-024-11081-x>
- Iparraguirre-Villanueva, O., Guevara-Ponce, V., Paredes, O., Sierra-Liñan, F., Zapata-Paulini, J., & Cabanillas-Carbonell, M. (2022). Convolutional neural networks with transfer learning for pneumonia detection. *International Journal of Advanced Computer Science and Applications*, 13(9), 592–599. <https://doi.org/10.14569/IJACSA.2022.0130963>
- Lei, L., Yang, Q., Yang, L., Shen, T., Wang, R., & Fu, C. (2024). Deep learning implementation of image segmentation in agricultural applications: A comprehensive review. *Artificial Intelligence Review*, 57, Article 149. <https://doi.org/10.1007/s10462-024-10775-6>
- Lu, Y., Chen, D., Olaniyi, E., & Huang, Y. (2022). Redes generativas antagónicas (GAN) para el aumento de imágenes en la agricultura: Una revisión sistemática. *Computers and Electronics in Agriculture*, 200, 107208. <https://doi.org/10.1016/j.compag.2022.107208>
- Paauw, M., Hardeman, G., Taks, N. W., Lambalk, L., Berg, J. A., Pfeilmeier, S., & van den Burg, H. A. (2024). ScAnalyzer: An image processing tool to monitor plant disease symptoms and pathogen spread in Arabidopsis thaliana leaves. *Plant Methods*, 20, Article 80. <https://doi.org/10.1186/s13007-024-01213-3>
- Salem, H. M., Fan, D. P., & Shao, L. (2020). Data augmentation using GAN for improving skin disease classification. *Computer Methods and Programs in Biomedicine*, 198, 105769. <https://doi.org/10.1016/j.cmpb.2020.105769>
- Shalev-Shwartz, S., & Ben-David, S. (2014). Understanding machine learning: From theory to algorithms. Cambridge University Press.
- Soria Olivas, E., Sánchez-Montañés Isla, A., Gamero Cruz, R., Castillo Caballero, B., & Cano Michalena, P. (2023). Sistemas de aprendizaje automático. RA-MA Editorial.
- Tassi, A., McGough, A. S., & Armitage, D. W. (2022). Applications of generative adversarial networks in agriculture: A review. arXiv. <https://doi.org/10.48550/arXiv.2204.04707>
- Thayer, A. W., Vargas, A., Castellanos, A. A., Lafon, C. W., McCarl, B. A., Roelke, D. L., Winemiller, K. O., & Lacher, T. E. (2020). Integrating agriculture and ecosystems to find suitable adaptations to climate change. *Climate*, 8(1), 10. <https://doi.org/10.3390/cli8010010>
- Wang, C., Xia, Y., Xia, L. et al. Dual discriminator GAN-based synthetic crop disease image generation for precise crop disease identification. *Plant Methods* 21, 46 (2025). <https://doi.org/10.1186/s13007-025-01361-0>
- Wang, S.; Xu, D.; Liang, H.; Bai, Y.; Li, X.; Zhou, J.; Su, C.; Wei, W. Advances in Deep Learning Applications for Plant Disease and Pest Detection: A Review. *Remote Sens.* 2025, 17, 698. <https://doi.org/10.3390/rs17040698>
- Zhou, Y., Liu, X., & Zhang, J. (2023). A hybrid autoencoder-GAN model for small-scale image synthesis in imbalanced datasets. *Applied Sciences*, 13(2), 905. <https://doi.org/10.3390/app13020905>