



Detección automática de enfermedades foliares del cafeto mediante reconocimiento de patrones

Automatic detection of coffee leaf diseases through pattern recognition

✉ Santa-María, John C.^{1,2*}

✉ Rodríguez, Ciro¹

¹Universidad Nacional Mayor de San Marcos, Lima, Perú

²Universidad Nacional de San Martín, Tarapoto, Perú

Recibido: 15 May. 2025 | Aceptado: 22 Dic. 2025 | Publicado: 20 Ene. 2026

Autor de correspondencia*: john.santamaria@unmsm.edu.pe

Cómo citar este artículo: Santa-María, J. C., & Rodríguez, C. (2026). Detección automática de enfermedades foliares del cafeto mediante reconocimiento de patrones. *Revista Científica de Sistemas e Informática*, 6(1), e1349. <https://doi.org/10.51252/rcsi.v6i1.1349>

RESUMEN

La detección temprana de enfermedades foliares del cafeto es clave para reducir pérdidas productivas; sin embargo, el diagnóstico visual en campo presenta limitaciones asociadas a la subjetividad y la variabilidad ambiental. El objetivo de este estudio fue diseñar y evaluar un modelo híbrido de reconocimiento de patrones para clasificar hojas de cafeto sanas, con roya y con ojo de gallo a partir de imágenes capturadas en condiciones reales en Saposoa (San Martín, Perú). Se empleó un conjunto de datos propio de 1500 imágenes validadas por especialista (500 por clase), ampliado mediante aumento de datos controlado hasta 6000 imágenes balanceadas. ResNet18 fue utilizado como extractor de características por transferencia de aprendizaje y se compararon tres clasificadores supervisados: SVM, Random Forest y XGBoost. La evaluación se realizó mediante validación cruzada estratificada de 10 pliegues y un conjunto de prueba independiente (20%). El modelo ResNet18 + SVM obtuvo el mejor desempeño, con una accuracy de 0.9742, F1-macro de 0.9730 y AUC-macro de 0.9968, superando a Random Forest (accuracy = 0.9367) y XGBoost (accuracy = 0.9583). El análisis inferencial mediante ANOVA y la prueba de Tukey HSD confirmó diferencias estadísticamente significativas entre modelos ($p < 0.001$). Los resultados evidencian la robustez y viabilidad del enfoque propuesto para apoyar el diagnóstico fitosanitario del cafeto en condiciones reales de campo.

Palabras clave: aprendizaje profundo; diagnóstico fitosanitario; reconocimiento de patrones; transferencia de aprendizaje; visión artificial

ABSTRACT

Early detection of coffee leaf diseases is essential to reduce production losses; however, visual field diagnosis presents limitations associated with subjectivity and environmental variability. The objective of this study was to design and evaluate a hybrid pattern-recognition model to classify healthy coffee leaves and those affected by coffee leaf rust and brown eye spot using images captured under real field conditions in Saposoa (San Martín, Peru). A proprietary dataset of 1500 images validated by a specialist (500 per class) was used and expanded through controlled data augmentation to 6000 balanced images. ResNet18 was employed as a feature extractor using transfer learning, and three supervised classifiers were compared: SVM, Random Forest, and XGBoost. Model performance was evaluated using 10-fold stratified cross-validation and an independent test set (20%). The ResNet18 + SVM model achieved the best performance, with an accuracy of 0.9742, a macro F1-score of 0.9730, and a macro-AUC of 0.9968, outperforming Random Forest (accuracy = 0.9367) and XGBoost (accuracy = 0.9583). Inferential analysis using ANOVA and Tukey's HSD test confirmed statistically significant differences among models ($p < 0.001$). The results demonstrate the robustness and feasibility of the proposed approach to support coffee phytosanitary diagnosis under real-world field conditions.

Keywords: computer vision; deep learning; pattern recognition; phytosanitary diagnosis; transfer learning



1. INTRODUCCIÓN

La detección temprana y precisa de enfermedades foliares del cafeto constituye un factor crítico para la sostenibilidad productiva y económica de los pequeños productores en América Latina. Entre las patologías más severas destacan la roya del café (*Hemileia vastatrix*) y el ojo de gallo (*Mycena citricolor*), responsables de pérdidas significativas de rendimiento y calidad, así como de impactos directos en la seguridad alimentaria y la estabilidad financiera de los caficultores (Avelino et al., 2015; Poma-Angamarca et al., 2024).

A pesar de su relevancia agronómica, el diagnóstico de estas enfermedades en campo continúa basándose predominantemente en la inspección visual realizada por agricultores o técnicos. Este método presenta limitaciones importantes asociadas a la subjetividad del evaluador, la variabilidad en la manifestación de síntomas y la influencia de condiciones ambientales como iluminación, humedad y estado fenológico de la planta, lo que incrementa la probabilidad de errores diagnósticos (Abade et al., 2020; Abdullah et al., 2023).

Los diagnósticos imprecisos generan decisiones agronómicas inadecuadas, tales como aplicaciones tardías o innecesarias de fungicidas, incrementando los costos de producción y reduciendo la eficacia del manejo fitosanitario. Esta problemática resulta especialmente crítica para los pequeños productores, quienes cuentan con recursos limitados y acceso restringido a asistencia técnica especializada, lo que amplifica el impacto económico de estas enfermedades (Julca-Otiniano et al., 2024; Saavedra-Ramírez, 2023).

En los últimos años, la visión artificial y el aprendizaje profundo se consolidan como herramientas prometedoras para la detección automática de enfermedades en plantas. Diversos estudios demuestran que las redes neuronales convolucionales y los enfoques de transferencia de aprendizaje permiten clasificar enfermedades foliares con altos niveles de precisión, incluyendo patologías del cafeto como la roya y la cercospora (Archana & Jeevaraj, 2024; Mansouri et al., 2024; Martinez et al., 2022).

Adicionalmente, se proponen modelos híbridos que combinan la extracción automática de características mediante deep learning con clasificadores supervisados tradicionales, como SVM o Random Forest, logrando mejoras en robustez y desempeño frente a variaciones de iluminación y ruido visual (Abuhayi & Mossa, 2023; Ayikpa et al., 2022). No obstante, persisten desafíos relacionados con la generalización de los modelos y su confiabilidad en escenarios reales de campo.

Una limitación recurrente en la literatura es que muchos modelos se entrena con conjunto de datos internacionales o imágenes capturadas en condiciones controladas, con hojas aisladas y fondos uniformes. Estas condiciones difieren significativamente de los entornos reales de cultivo, donde las imágenes presentan sombras, occlusiones, múltiples estados de infección y alta variabilidad ambiental, lo que afecta el desempeño del diagnóstico automático (Chavarro et al., 2023).

En el contexto peruano, y particularmente en la región San Martín, la caficultura representa un pilar económico y social fundamental. Investigaciones recientes reportan la presencia de nuevas razas de *Hemileia vastatrix* y una elevada incidencia del ojo de gallo en sistemas agroforestales húmedos, incrementando la complejidad del diagnóstico visual y la necesidad de herramientas tecnológicas adaptadas a las condiciones locales (Ehrenbergerová et al., 2018; Julca-Otiniano et al., 2024).

A pesar de los avances tecnológicos, se identifica un vacío científico relevante: la ausencia de modelos automáticos robustos entrenados con imágenes reales de cafetales peruanos que permitan identificar de manera diferenciada la roya y el ojo de gallo en condiciones propias de la región San Martín. Esta brecha limita la adopción efectiva de sistemas inteligentes para el manejo fitosanitario local.

En este contexto, el propósito de la presente investigación es diseñar y evaluar un modelo híbrido de análisis automático basado en reconocimiento de patrones, que integre técnicas de visión artificial y aprendizaje profundo para identificar la presencia de roya y ojo de gallo en hojas de cafeto a partir de imágenes capturadas en campo. El modelo se entrena con un conjunto de datos propio representativo de

las condiciones agroecológicas de San Martín, con el fin de ofrecer un diagnóstico objetivo, reproducible y aplicable que complemente la evaluación visual tradicional y fortalezca la toma de decisiones en el manejo integrado de enfermedades.

2. MATERIALES Y MÉTODOS

La investigación se desarrolló como un estudio de enfoque aplicado, orientado al diseño y evaluación de un modelo automático de reconocimiento de patrones para la clasificación de enfermedades foliares del cafeto a partir de imágenes digitales capturadas en condiciones reales de campo. El nivel de la investigación fue explicativo, dado que se analizó cómo la extracción de características mediante transferencia de aprendizaje, combinada con clasificadores supervisados, permitió discriminar entre hojas sanas, afectadas por roya y por ojo de gallo. El diseño fue no experimental y transversal, ya que no se manipuló la condición fitosanitaria de las hojas y los datos fueron recolectados en un periodo determinado. Asimismo, el estudio incorporó un enfoque tecnológico, al proponer e implementar un modelo computacional con aplicabilidad práctica en el diagnóstico fitosanitario del cultivo de café.

Los materiales utilizados para la realización de esta investigación se dividieron en software y hardware, los cuales se describieron en la Tabla 1.

Tabla 1. Software y hardware empleados

Componente	Detalle
Software	Google Colab (entrenamiento y evaluación), Google Drive (almacenamiento), Python (scripts del pipeline), OpenCV y scikit-image (preprocesamiento), torch y torchvision (extractor CNN), scikit-learn (SVM/Random Forest y métricas), XGBoost (clasificador boosting), pandas y matplotlib (análisis y visualización).
Hardware	Teléfono móvil Samsung Galaxy A54 5G (captura de imágenes en campo).

En la Figura 1 se representó la metodología general, estructurada como un sistema de visión artificial para clasificar hojas de cafeto en tres clases fitosanitarias. La metodología siguió las siguientes etapas: 1) zona y diseño del estudio, 2) recolección y control de calidad, 3) construcción del conjunto de datos, 4) desarrollo del modelo híbrido, 5) evaluación y comparación estadística.

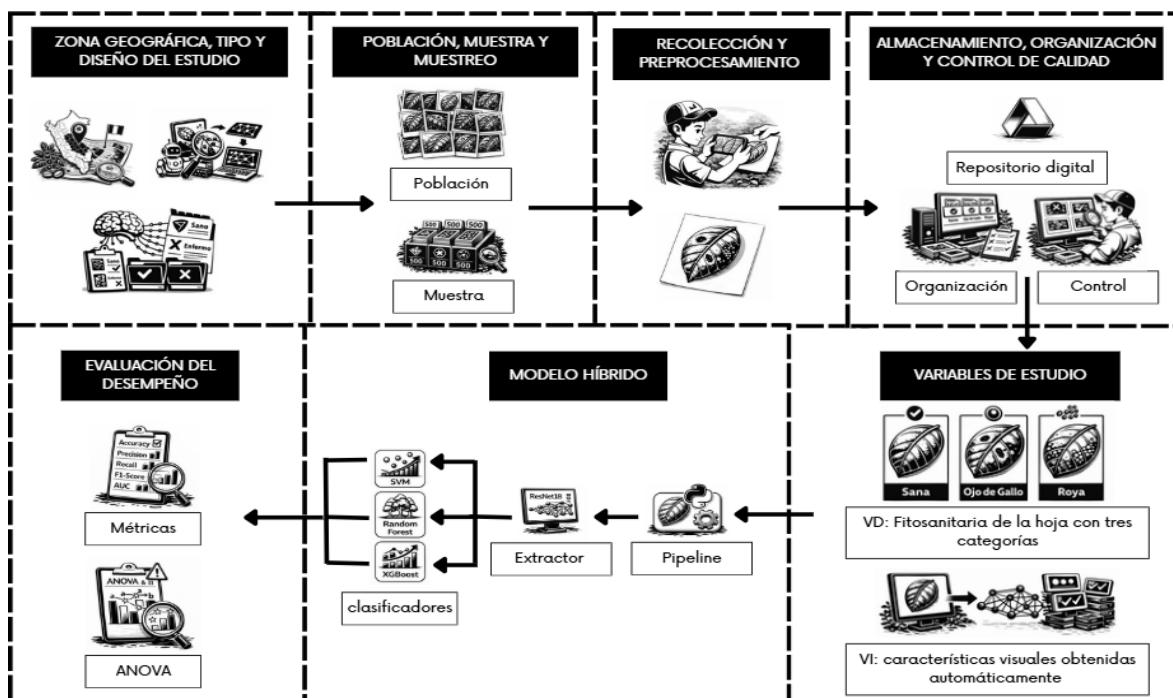


Figura 1. Arquitectura metodológica del modelo híbrido de reconocimiento de patrones para la clasificación de enfermedades foliares del cafeto

1ra etapa: zona geográfica, tipo y diseño del estudio

El estudio se ejecutó en plantaciones de café del distrito de Saposoa, región San Martín, Perú, bajo condiciones reales de campo. La investigación se desarrolló con enfoque tecnológico-aplicado orientado al diseño y evaluación de un modelo de reconocimiento de patrones. El nivel fue explicativo, debido a que se examinó cómo la extracción de características por transferencia de conocimiento, combinada con clasificadores supervisados, permitió discriminar clases de enfermedad. El diseño fue no experimental y de corte transversal, debido a que no se manipuló la condición fitosanitaria y la captura se realizó en un periodo definido.

2da etapa: población, muestra y muestreo

La población estuvo conformada por hojas de cafeto sanas y enfermas presentes en cultivos de Saposoa, registrables mediante imágenes digitales y con diagnóstico inicial de especialistas en sanidad vegetal. Se trabajó con una muestra de 1 500 imágenes, distribuidas en tres estratos: 500 hojas sanas, 500 con ojo de gallo y 500 con roya. La selección se realizó mediante muestreo no probabilístico por conveniencia, con estratificación por condición fitosanitaria; dentro de cada estrato se priorizaron imágenes con diversidad visual y calidad técnica adecuada.

Criterios de inclusión: (i) procedencia Saposoa; (ii) captura con protocolo uniforme (distancia, iluminación y enfoque); (iii) clasificación confirmada por especialista (sana, ojo de gallo, roya); (iv) calidad suficiente (buena resolución, enfoque correcto y sin sombras excesivas).

Criterios de exclusión: (i) imágenes borrosas, incompletas, sobre/subexpuestas; (ii) hojas ocluidas por objetos externos; (iii) ausencia de confirmación por especialista; (iv) estados intermedios sin diferenciación clara de enfermedad.

3ra etapa: recolección y preprocesamiento de imágenes

La captura de imágenes se realizó directamente en campo con apoyo de un ingeniero agrónomo, quien evaluó cada hoja antes del registro fotográfico. Las visitas se efectuaron aproximadamente entre las 10:00 a. m. y 1:00 p. m., periodo con iluminación natural relativamente uniforme. Para la toma, cada hoja se colocó sobre un fondo blanco y se fotografió manteniendo una distancia constante, asegurando encuadre completo y adecuada nitidez de bordes y lesiones. Las imágenes se almacenaron en formato JPG, conservando la resolución nativa del dispositivo y evitando procesos adicionales de compresión.

Previo al entrenamiento de los modelos, las imágenes fueron sometidas a un proceso de preprocesamiento estandarizado. Todas las muestras se redimensionaron a 224 × 224 píxeles, acorde a los requerimientos de entrada de la arquitectura ResNet18, y se normalizaron utilizando los parámetros de media y desviación estándar del conjunto ImageNet, con el fin de garantizar compatibilidad con el modelo preentrenado y estabilidad numérica durante el entrenamiento.

Adicionalmente, se aplicó un aumento de datos offline controlado para incrementar la variabilidad visual del conjunto y mejorar la capacidad de generalización de los modelos frente a condiciones reales de campo. Para cada imagen original se generaron exactamente tres imágenes aumentadas, empleando transformaciones complementarias: (i) aumentos geométricos (rotaciones aleatorias de $\pm 50^\circ$, escalado y traslación, con volteos horizontales/verticales), (ii) variaciones de color e iluminación en el espacio HSV (ajustes de tono, saturación y brillo), y (iii) recorte aleatorio con redimensionamiento y adición de ruido gaussiano. El aumento se realizó de forma offline para mantener trazabilidad entre imágenes originales y aumentadas, evitar fugas de información durante la validación y asegurar un balance homogéneo entre clases.

4ta etapa: almacenamiento, organización y control de calidad

Las imágenes se transfirieron a un repositorio digital y se organizaron en tres directorios: sanas, ojo de gallo y roya, según el diagnóstico del especialista. El control de calidad se aplicó en dos momentos: (i) revisión inmediata durante la captura para repetir fotografías con desenfoque, sombras fuertes o encuadres incompletos; y (ii) revisión posterior en el entorno de almacenamiento para descartar defectos técnicos o dudas de clasificación. Cuando existió incertidumbre diagnóstica, la etiqueta se verificó nuevamente con el especialista, garantizando consistencia del conjunto final.

5ta etapa: variables de estudio

La variable dependiente fue la clase fitosanitaria de la hoja con tres categorías: sana, ojo de gallo y roya. La variable independiente correspondió a la representación de características visuales obtenidas automáticamente desde la imagen mediante un extracto convolucional basado en transferencia de conocimiento, las cuales alimentaron a los clasificadores supervisados del modelo híbrido.

6ta etapa: desarrollo del modelo híbrido (reconocimiento de patrones)

El pipeline se implementó en Python y se estructuró en: (i) ingreso y normalización básica de la imagen, (ii) extracción de características mediante ResNet18 preentrenada usada como extracto (transfer learning), y (iii) clasificación supervisada. Se evaluaron tres combinaciones híbridas: ResNet18 + SVM, ResNet18 + Random Forest y ResNet18 + XGBoost, manteniendo el mismo conjunto de datos y el mismo esquema de validación para asegurar comparabilidad (Figura 2).

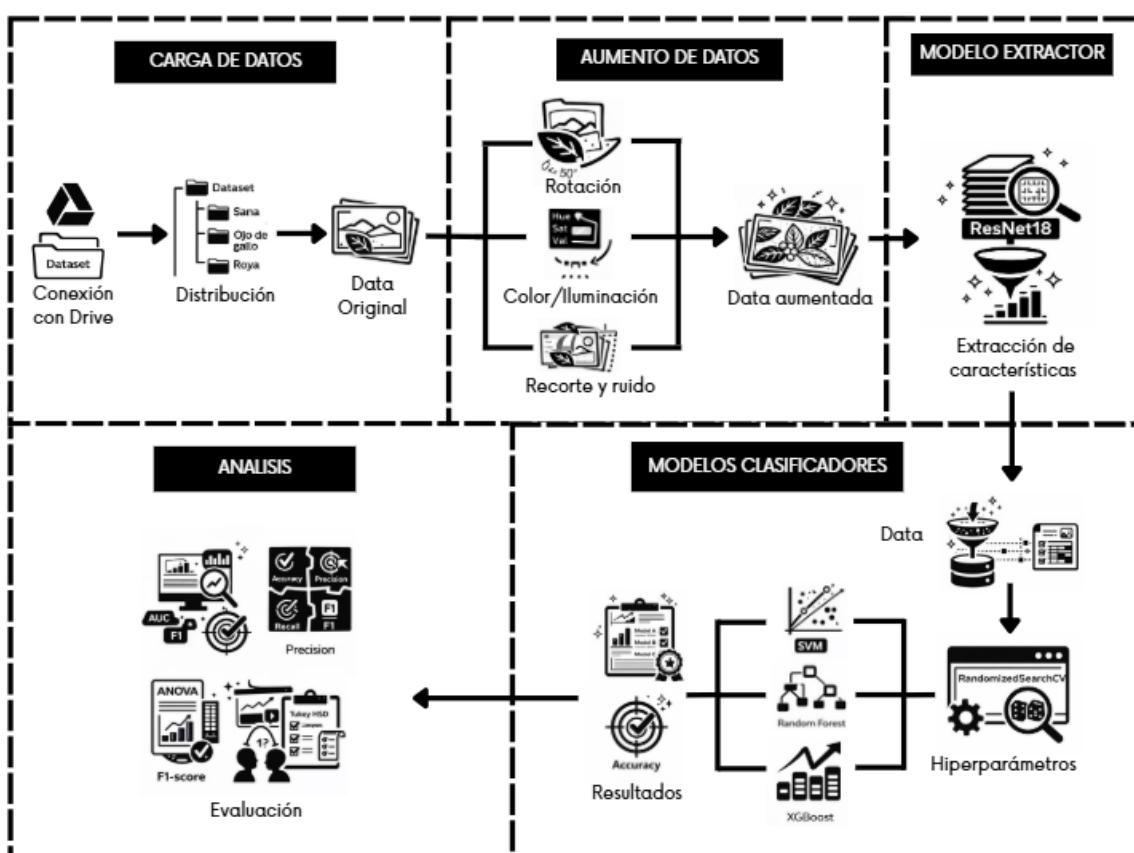


Figura 2. Flujograma del pipeline metodológico del modelo híbrido de reconocimiento de patrones

Para asegurar una comparación equitativa entre los tres modelos híbridos (ResNet18+SVM, ResNet18+Random Forest y ResNet18+XGBoost), los clasificadores no se entrenaron con parámetros por defecto. Los hiperparámetros se ajustaron mediante búsqueda aleatoria (RandomizedSearchCV) sobre el conjunto de entrenamiento (80%), manteniendo el conjunto de prueba (20%) completamente

independiente para la evaluación final. En cada clasificador se exploraron 30 configuraciones y se utilizó validación interna estratificada, preservando proporciones similares de las tres clases. La métrica objetivo empleada para seleccionar la mejor configuración fue la accuracy, por corresponder a la métrica principal utilizada posteriormente en la comparación estadística (val_acc). Para SVM se empleó un pipeline con estandarización (StandardScaler) y (opcionalmente) reducción de dimensionalidad (PCA) antes del clasificador; en Random Forest y XGBoost se evaluaron combinaciones de hiperparámetros asociadas a complejidad del modelo y regularización, manteniendo el mismo número de iteraciones y el mismo esquema de validación para garantizar condiciones homogéneas entre modelos.

7ma etapa: evaluación del desempeño y tratamiento estadístico

La evaluación se realizó mediante validación cruzada estratificada de 10 pliegues, preservando proporciones similares de las tres clases en cada fold. En cada pliegue se calcularon métricas estándar: accuracy, precision, recall, F1-score y AUC, obteniéndose 10 observaciones por modelo. Para comparar el rendimiento entre los tres modelos híbridos se aplicó un ANOVA de un factor con bloqueo por fold; cuando se detectó significancia estadística, se ejecutó la prueba post-hoc Tukey HSD para identificar diferencias por pares. Con base en consistencia y desempeño global, ResNet18 + SVM se seleccionó como modelo final para la clasificación automática de hojas sanas, con ojo de gallo y con roya.

3. RESULTADOS Y DISCUSIÓN

La investigación generó como principal aporte un conjunto de datos propio de hojas de cafeto capturadas en campo (Saposoa, San Martín) y la evaluación comparativa de tres modelos híbridos basados en extracción de características con ResNet18 y clasificación supervisada (SVM, Random Forest y XGBoost). Los resultados mostraron un desempeño alto en la identificación de hojas sanas, ojo de gallo y roya, así como diferencias estadísticamente significativas entre modelos en validación cruzada.

3.1. Construcción y balance del conjunto de datos

El conjunto inicial estuvo conformado por 1 500 imágenes (500 por clase) los cuales son observables en la Figura 3. Luego, mediante aumento de datos offline controlado (tres transformaciones por imagen), el conjunto de datos alcanzó 6 000 imágenes que se puede evidenciar en la Figura 4, manteniendo un balance perfecto de 2 000 imágenes por clase. Este incremento amplió la variabilidad visual bajo condiciones similares a campo, favoreciendo la estabilidad del entrenamiento y la comparación justa entre modelos.

Se muestra un fragmento del código utilizado para la visualización de las hojas iniciales:

```
for cls in classes:
    path = os.path.join(segmentado_path, cls)
    imgs = [f for f in os.listdir(path)
            if f.lower().endswith(valid_ext) and "_aug" not in f.lower()][:-3]
    fig, axes = plt.subplots(1, len(imgs), figsize=(12, 4))
    axes = axes if isinstance(axes, (list, np.ndarray)) else [axes]
    for ax, name in zip(axes, imgs):
        img = cv2.cvtColor(cv2.imread(os.path.join(path, name)), cv2.COLOR_BGR2RGB)
        ax.imshow(img)
        ax.axis("off")
    plt.show()
```



Figura 3. Muestras representativas del conjunto de datos de hojas de cafeto por clase fitosanitaria

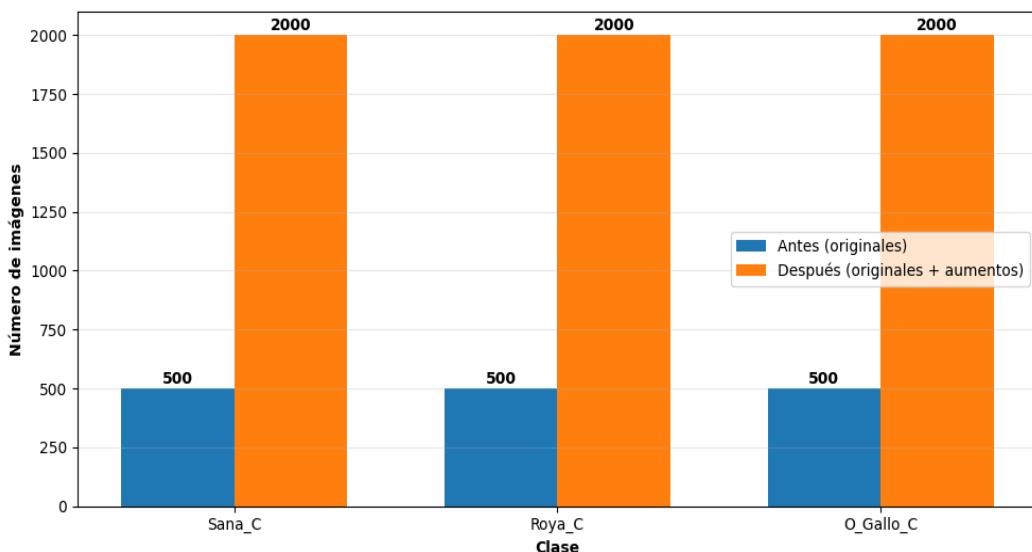


Figura 4. Distribución por clases (antes y después)

Antes de presentar los resultados comparativos, en la Tabla 2 se resume el tamaño final del conjunto de datos por clase.

Tabla 2. Distribución del conjunto de datos antes y después del aumento

Clase	Antes del aumento	Después del aumento	Porcentaje de incremento
Sana	500	2000	150%
Roya	500	2000	150%
Ojo de gallo	500	2000	150%
Total	1500	6000	150%

Es importante precisar que el aumento de datos no introdujo riesgo de sobreajuste ni sesgos metodológicos en la evaluación de los modelos. El proceso de data augmentation se realizó de manera offline y controlada, previo a la partición del conjunto de datos, y todas las imágenes aumentadas fueron agrupadas junto con su imagen original durante la separación entrenamiento-prueba. De este modo, se garantizó que ninguna versión aumentada de una misma hoja apareciera simultáneamente en ambos subconjuntos, evitando fugas de información (data leakage). Esta estrategia permitió incrementar la variabilidad visual del conjunto de entrenamiento sin comprometer la independencia del conjunto de prueba, fortaleciendo la validez de los resultados y la capacidad de generalización de los modelos evaluados.

3.2. Desempeño de los modelos híbridos en el conjunto de prueba

Los tres enfoques híbridos alcanzaron un desempeño alto en datos no vistos (20% hold-out), con superioridad consistente del modelo ResNet18 + SVM. En la Tabla 3 se presentan las métricas globales obtenidas en el conjunto de prueba.

Tabla 3. Desempeño de los modelos en TEST (20% hold-out)

Modelo	Accuracy	F1-macro	AUC-macro	n_train	n_test
ResNet18 + SVM	0.974167	0.973033	0.996806	4800	1200
ResNet18 + Random Forest	0.936667	0.934056	0.987143	4800	1200
ResNet18 + XGBoost	0.958333	0.956459	0.993550	4800	1200

En términos comparativos, el orden de desempeño en TEST fue SVM > XGBoost > Random Forest, coherente con la tendencia observada durante la validación cruzada. En particular, ResNet18 + SVM obtuvo los mejores valores de accuracy (0.9742), F1-macro (0.9730) y AUC-macro (0.9968), lo que sugiere una combinación más eficaz entre los embeddings extraídos por ResNet18 y un clasificador capaz de construir fronteras de decisión no lineales robustas. Este comportamiento es esperable, ya que SVM tiende a generalizar mejor en espacios de características densos y de alta dimensionalidad, mientras que los enfoques basados en árboles (Random Forest y XGBoost) pueden presentar mayor sensibilidad al ruido y a variaciones locales cuando operan sobre vectores de embeddings.

Complementariamente, el análisis de las matrices de confusión permitió interpretar los tipos de error más frecuentes. En los tres modelos, la mayor proporción de confusiones ocurrió entre roya y sano, lo que indica que ciertos casos de roya —especialmente en estadios tempranos o con baja severidad— comparten rasgos visuales con hojas sanas bajo variabilidad de iluminación y textura propia de imágenes de campo. En ResNet18 + SVM, la clase ojo de gallo mostró un reconocimiento particularmente estable, mientras que la roya concentró la mayoría de errores, principalmente como Roya → Sana y, en menor medida, Sana → Roya. En Random Forest y XGBoost se mantuvo el mismo patrón de confusión, pero con mayor magnitud relativa en Random Forest, lo que explicó su menor desempeño global.

Desde una perspectiva operativa, la confusión Roya → Sana constituye el error más crítico, pues implica falsos negativos que podrían retrasar intervenciones fitosanitarias; en contraste, Sana → Roya puede inducir falsos positivos y decisiones preventivas innecesarias. Por ello, aunque el AUC-macro se mantuvo alto en los tres casos y evidenció buena capacidad discriminativa multiclasé, los resultados sugieren que la mejora futura debe priorizar la separación entre roya y sano, incorporando mayor diversidad de casos leves, condiciones de iluminación más extremas y variación de fondo/occlusiones para reducir el error con mayor impacto práctico.

3.3. Resultados por validación cruzada y comparación inferencial

Con el objetivo de evaluar la estabilidad y robustez de los modelos frente a variaciones del conjunto de entrenamiento, se aplicó una validación cruzada estratificada de 10 pliegues sobre el conjunto de entrenamiento. Este procedimiento permitió analizar el comportamiento de cada modelo bajo diferentes particiones de los datos, simulando escenarios de variabilidad inherentes a imágenes capturadas en

condiciones reales de campo, donde factores como iluminación, textura foliar y severidad de la enfermedad pueden cambiar entre muestras. Al preservar la proporción de las tres clases en cada pliegue, la validación cruzada garantizó una evaluación equilibrada y redujo la dependencia de una única partición de entrenamiento, fortaleciendo la evidencia de capacidad de generalización del enfoque propuesto. Con el fin de transparentar la base empírica utilizada en el análisis inferencial, en la Tabla 4 se reportan los valores de accuracy obtenidos en cada fold (val_acc) para los tres modelos híbridos evaluados.

Tabla 4. Accuracy por fold en TRAIN (10-fold CV)

Fold	ResNet18+SVM	ResNet18+RF	ResNet18+XGBoost
1	0.9875	0.9563	0.9625
2	0.9771	0.9396	0.9542
3	0.9875	0.9313	0.9604
4	0.9646	0.9208	0.9396
5	0.9792	0.9396	0.9604
6	0.9750	0.9167	0.9646
7	0.9771	0.9354	0.9604
8	0.9771	0.9375	0.9563
9	0.9667	0.9208	0.9437
10	0.9563	0.9167	0.9396

A partir de estos valores se aplicó un ANOVA con bloqueo por fold, evidenciando diferencias significativas en el desempeño entre modelos. En la Tabla 5 se reportan los resultados del ANOVA.

Tabla 5. ANOVA para val_acc en CV (TRAIN) con bloqueo por fold

Fuente	Suma de cuadrados	gl	MC	F	p-value
Modelo	0.009396	2	0.004698	133.025489	1.647730e-11
Fold	0.002532	9	0.000281	7.965407	1.068262e-04
Residual	0.000636	18	0.000035	—	—

El contraste indicó efecto significativo del factor “modelo” ($p < 0.05$). Las medias de val_acc en CV fueron: SVM = 0.974792, XGBoost = 0.954167 y RF = 0.931458, manteniendo el mismo orden observado en TEST.

Dado el resultado significativo, se aplicó Tukey HSD para identificar diferencias por pares. La Tabla 6 resume las comparaciones múltiples.

Tabla 6. Tukey HSD para val_acc (TRAIN)

Modelo A	Modelo B	Diferencia de medias	p-valor	¿Diferencia significativa?
ResNet18+RF	ResNet18+SVM	0.0433	0.0000	Sí
ResNet18+RF	ResNet18+XGBoost	0.0227	0.0002	Sí
ResNet18+SVM	ResNet18+XGBoost	-0.0206	0.0006	Sí

Con el propósito de complementar el análisis estadístico inferencial y facilitar la interpretación visual de las diferencias entre modelos, se incorporaron representaciones gráficas post hoc basadas en los valores de accuracy obtenidos durante la validación cruzada. En la Figura 5 se presenta un diagrama de cajas (boxplot) del val_acc por modelo, el cual permite analizar la dispersión, la estabilidad y el grado de solapamiento del desempeño a lo largo de los 10 pliegues. En esta figura se observa que el modelo ResNet18 + SVM concentra los valores más altos de accuracy y presenta una menor variabilidad inter-fold, evidenciando un comportamiento más estable frente a variaciones en el conjunto de entrenamiento. Por el contrario, ResNet18 + Random Forest muestra una mayor dispersión y valores centrales más bajos, mientras que ResNet18 + XGBoost exhibe un desempeño intermedio.

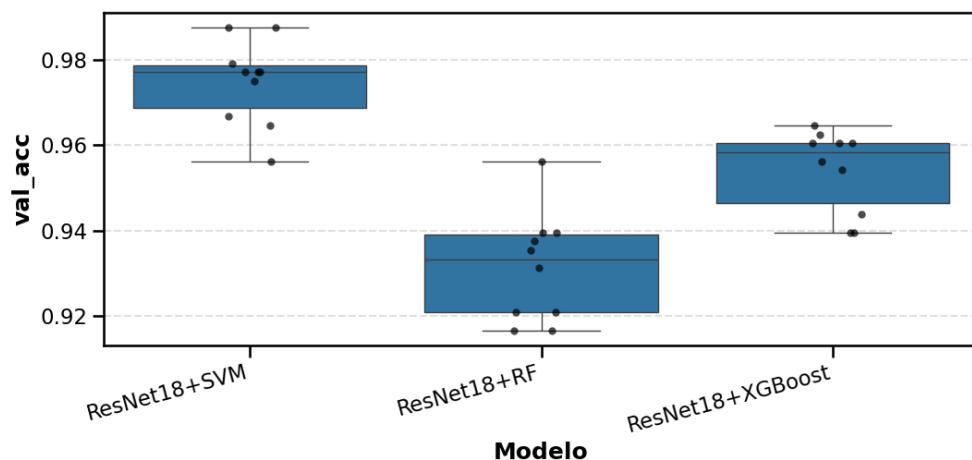


Figura 5. Distribución de val_acc en CV por modelo

De manera complementaria, la Figura 6 muestra el gráfico de medias de val_acc en validación cruzada, reforzando visualmente el orden de desempeño identificado por el ANOVA y la prueba post hoc de Tukey HSD (ResNet18 + SVM > ResNet18 + XGBoost > ResNet18 + Random Forest). Estas visualizaciones confirman que las diferencias estadísticas observadas no solo son significativas desde el punto de vista inferencial, sino también consistentes en términos de estabilidad y generalización, aspecto especialmente relevante considerando que las imágenes fueron capturadas en condiciones reales de campo.

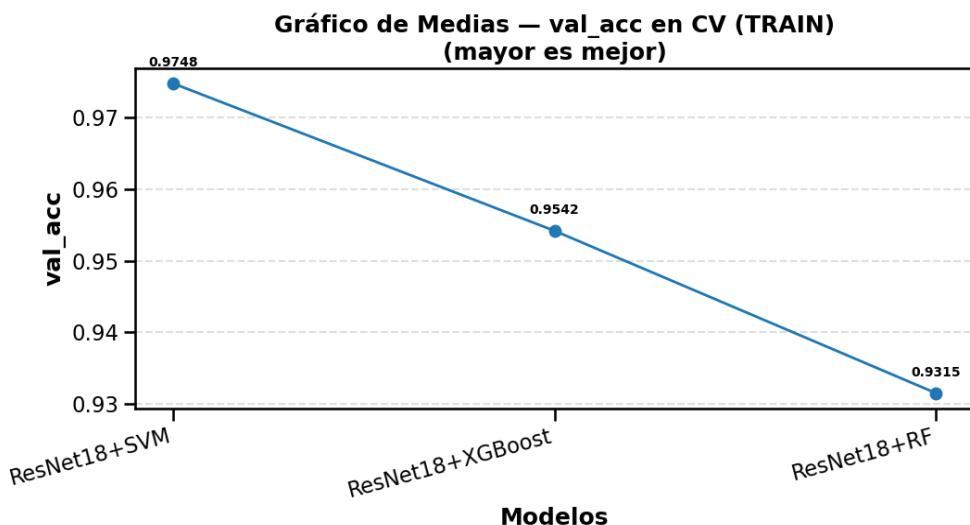


Figura 6. Gráfico de medias de val_acc en validación cruzada

3.4. Análisis detallado del modelo seleccionado (ResNet18 + SVM)

Dado que el modelo ResNet18 + SVM mostró el mejor desempeño global y la mayor consistencia tanto en validación cruzada como en el conjunto de prueba, se presenta a continuación un análisis detallado de su comportamiento clasificadorio, con el fin de profundizar en la interpretación de los resultados y evaluar su aplicabilidad práctica en escenarios reales.

3.4.1. Matrices de confusión en entrenamiento y prueba

La Figura 7 muestra la matriz de confusión correspondiente al conjunto de entrenamiento (80%). Se observa una diagonal claramente dominante, lo que evidencia una separación casi perfecta entre las clases evaluadas (O_Gallo_C, Roya_C y Sana_C), confirmando la capacidad del modelo para aprender patrones visuales discriminativos a partir de los embeddings extraídos por ResNet18.

Por su parte, la Figura 8 presenta la matriz de confusión obtenida en el conjunto de prueba (20% hold-out), compuesto por imágenes nunca vistas durante el entrenamiento. En este escenario, el modelo clasificó correctamente la gran mayoría de las instancias, registrándose un número reducido de errores concentrados principalmente en confusiones entre las clases Roya_C y Sana_C. Este comportamiento resulta esperable desde una perspectiva agronómica, debido a la similitud visual que pueden presentar hojas sanas y hojas con roya en estadios tempranos de infección, especialmente bajo variabilidad de iluminación y textura propia de condiciones de campo.

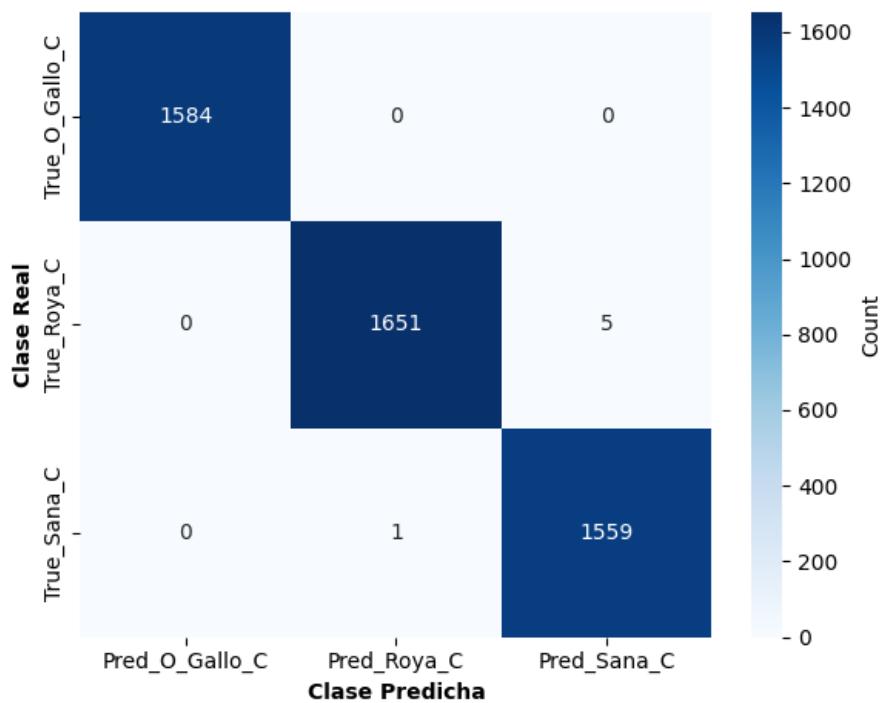


Figura 7. Matriz de confusión del modelo ResNet18 + SVM en el conjunto de entrenamiento (80%)

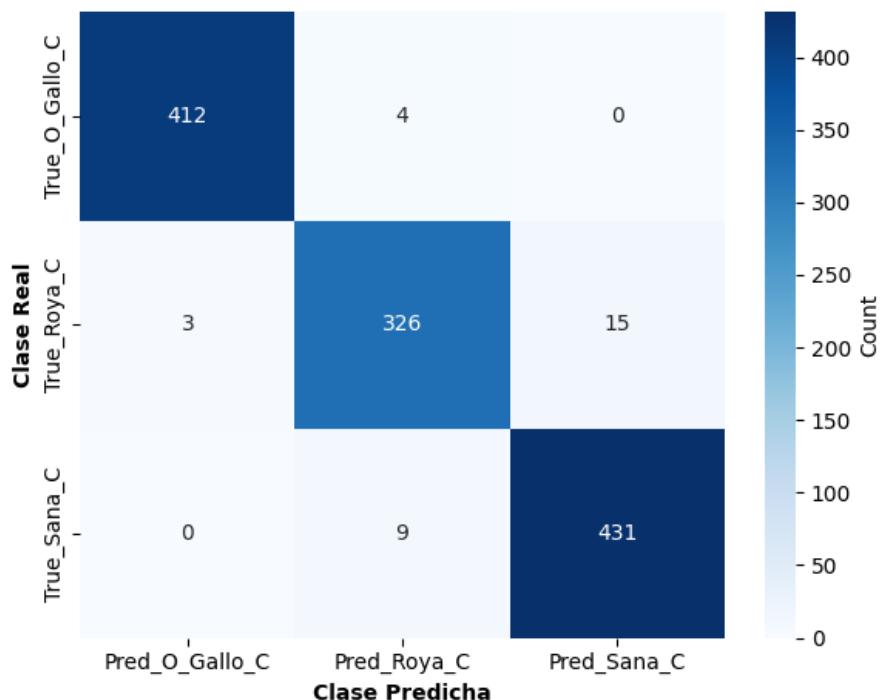


Figura 8. Matriz de confusión del modelo ResNet18 + SVM en el conjunto de prueba (20% hold-out)

3.4.2. Curvas ROC y capacidad discriminativa

La capacidad discriminativa del modelo se evaluó mediante curvas ROC multiclas bajo el esquema One-vs-Rest. La Figura 9 muestra las curvas ROC correspondientes al conjunto de prueba, donde el modelo alcanzó un Macro-AUC de 0.997, con valores individuales superiores a 0.99 para las tres clases analizadas. Estos resultados confirman una separación robusta entre categorías y una alta confiabilidad del sistema para distinguir entre hojas sanas, hojas con roya y hojas con ojo de gallo, incluso en presencia de variabilidad propia del entorno agrícola.

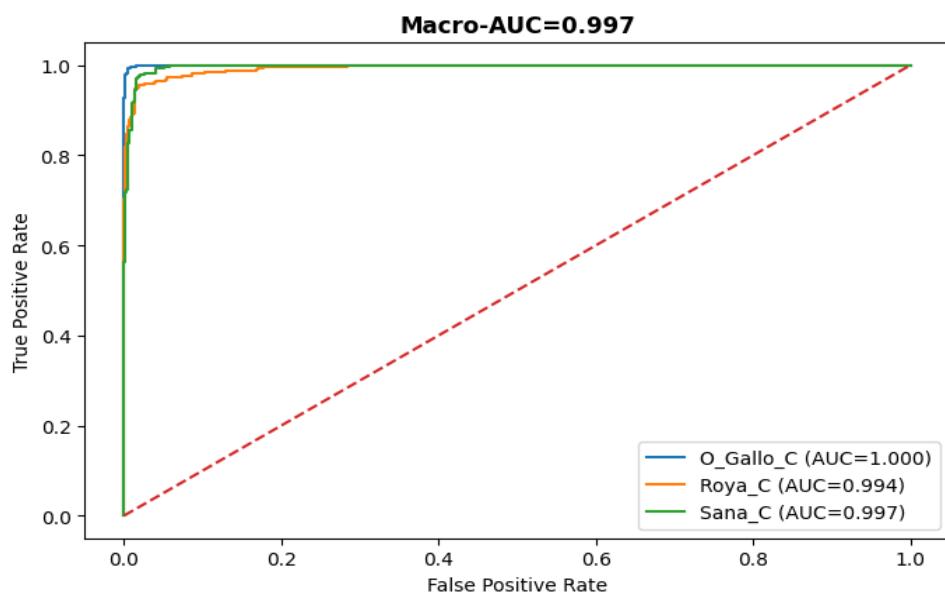


Figura 9. Curvas ROC multiclas (One-vs-Rest) del modelo ResNet18 + SVM en el conjunto de prueba (Macro-AUC = 0.997)

3.4.3. Estabilidad del modelo mediante validación cruzada

La estabilidad del modelo seleccionado se analizó a través de la evolución de las métricas accuracy y F1-macro a lo largo de los 10 pliegues de la validación cruzada estratificada. La Figura 10 evidencia que ambas métricas se mantienen en valores elevados y con variaciones controladas entre folds, sin caídas abruptas de desempeño. Este comportamiento confirma que el modelo ResNet18 + SVM es robusto frente a cambios en la partición del conjunto de entrenamiento, reforzando su capacidad de generalización sobre imágenes capturadas en condiciones reales de campo.

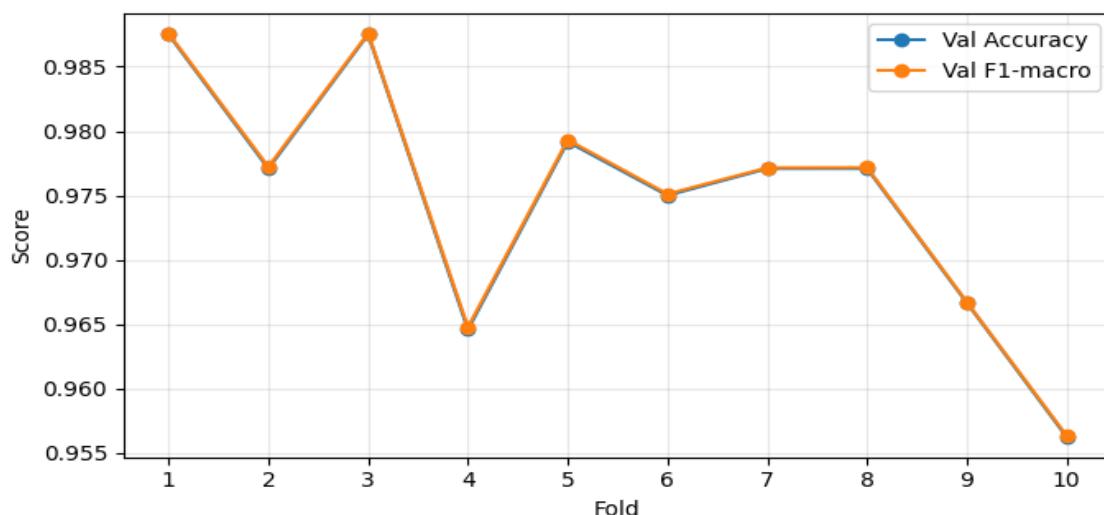


Figura 10. Evolución de accuracy y F1-macro por fold en la validación cruzada (10-fold CV) del modelo ResNet18 + SVM

3.5. Discusión

Los resultados obtenidos evidenciaron que el enfoque híbrido basado en transfer learning y clasificadores supervisados permitió una identificación precisa y estable de hojas de cafeto sanas, con roya y con ojo de gallo, destacando el modelo ResNet18 + SVM como la alternativa de mejor desempeño global. Este hallazgo fue consistente con lo reportado por Abuhayi & Mossa (2023) y Novtahaning et al. (2022), quienes demostraron que la combinación de extracción profunda de características y técnicas de clasificación mejora la robustez del diagnóstico automático frente a variaciones visuales. Asimismo, estudios recientes han señalado que el uso de arquitecturas profundas como extractores de características, combinadas con clasificadores tradicionales, favorece un mejor equilibrio entre discriminación y estabilidad del modelo en tareas de clasificación de enfermedades del cafeto (Pham et al., 2023). No obstante, investigaciones ampliamente citadas como la de (Atila et al., 2021), desarrollada sobre el conjunto de datos PlantVillage, reportaron precisiones cercanas al 99 % bajo condiciones controladas de laboratorio, caracterizadas por fondos uniformes, iluminación homogénea y hojas aisladas. Si bien dichos resultados evidencian el alto potencial del aprendizaje profundo, también ponen de manifiesto una limitación recurrente asociada a la generalización del modelo en escenarios reales de campo. En contraste, el presente estudio se basó en imágenes capturadas directamente en plantaciones de café bajo condiciones no controladas, incorporando variabilidad real en iluminación, textura foliar y severidad de la enfermedad, lo que refuerza el valor aplicado del enfoque propuesto y su mayor pertinencia para contextos agrícolas operativos.

Si bien investigaciones como las de Atila et al. (2021) y Mansouri et al. (2024) reportaron precisiones cercanas al 100 % empleando arquitecturas profundas como EfficientNet y conjunto de datos estandarizados, los resultados de la presente investigación se obtuvieron a partir de imágenes capturadas en campo, bajo condiciones reales de iluminación y variabilidad ambiental. En este contexto, el uso de ResNet18, una arquitectura más ligera, permitió mantener un alto desempeño con menores requerimientos computacionales, favoreciendo su viabilidad para aplicaciones prácticas en entornos agrícolas. Esta ventaja coincide con estudios que resaltan el potencial de modelos livianos como MobileNetV2 y EfficientNet-B0 para escenarios con recursos limitados y despliegue en dispositivos de bajo consumo (Arif et al., 2025; Aufar & Kaloka, 2022). No obstante, estos resultados deben interpretarse considerando ciertas limitaciones del estudio, entre ellas el tamaño del conjunto de datos y su procedencia de una única región productora, así como la dependencia de condiciones ambientales propias del momento de captura y la ausencia de una validación externa en otras zonas cafetaleras, aspectos que deberán abordarse en investigaciones futuras para evaluar la generalización espacial del modelo.

Por otro lado, investigaciones orientadas a la detección en tiempo real mediante modelos YOLO priorizaron la localización de lesiones con alta velocidad de inferencia (Adelaja & Pranggono, 2025; Fragoso et al., 2025). En contraste, el enfoque propuesto se centró en la clasificación fitosanitaria multiclasé, logrando un balance adecuado entre precisión, estabilidad y generalización, como se evidenció en la validación cruzada y el bajo gap entre entrenamiento y prueba. Esta estrategia resulta especialmente relevante, considerando que diversas aplicaciones móviles y sistemas de diagnóstico automatizado aún presentan limitaciones en generalización cuando se enfrentan a condiciones no controladas de captura de imágenes (Siddiqua et al., 2022).

CONCLUSIONES

Se diseñó y evaluó un modelo de análisis automático basado en reconocimiento de patrones para la clasificación de hojas de cafeto en tres condiciones fitosanitarias (sana, roya y ojo de gallo), confirmándose que el enfoque híbrido, que integra a ResNet18 como extractor de características y un clasificador supervisado, alcanzó un desempeño alto, estable y consistente; en la comparación entre modelos, ResNet18 + SVM presentó el mejor equilibrio entre precisión, consistencia y capacidad discriminativa, por lo que se seleccionó como la alternativa óptima. Asimismo, se construyó un conjunto de datos propio, estructurado

y balanceado, compuesto por imágenes capturadas en condiciones reales de campo y validadas por especialista, donde el aumento de datos controlado permitió incrementar el volumen sin alterar el balance entre clases, fortaleciendo el entrenamiento y la evaluación. La validación cruzada estratificada y el análisis inferencial mediante ANOVA y la prueba post hoc de Tukey HSD evidencian diferencias estadísticamente significativas entre los modelos, respaldando de forma objetiva la selección del enfoque final y su capacidad de generalización en el contexto del distrito de Saposoa (San Martín). Desde una perspectiva aplicada, los resultados evidencian el potencial del enfoque propuesto para su integración en sistemas móviles de apoyo al diagnóstico fitosanitario y herramientas de monitoreo agrícola en campo, orientadas a productores y técnicos, así como su posible adaptación a la detección de otras enfermedades foliares y a distintos cultivos, previo ajuste del conjunto de datos; en este sentido, se recomienda como trabajo futuro ampliar la recolección a otras regiones y campañas agrícolas, incorporar escenarios con mayor variabilidad ambiental y evaluar su despliegue en prototipos móviles para medir el rendimiento operativo en condiciones reales de uso.

FINANCIAMIENTO

Los autores no recibieron ningún patrocinio para llevar a cabo este estudio-artículo.

CONFLICTO DE INTERESES

No existe ningún tipo de conflicto de interés relacionado con la materia del trabajo.

CONTRIBUCIÓN DE AUTORES

Conceptualización; Curación de datos; Metodología; Administración del proyecto; Redacción – borrador original; Redacción – revisión y edición: Santa-María, J. C.; Análisis formal; Investigación; Supervisión: Santa María, J. C. y Rodríguez, C. La validación fue realizada por: Rodríguez, C.

DISPONIBILIDAD DE DATOS DEPOSITADOS

El conjunto de datos utilizado para este estudio está disponible abiertamente en Mendeley Data en <https://doi.org/10.17632/mfpwg4y65r.1>.

REFERENCIAS

- Abade, A. S., Ferreira, P. A., & Vidal, F. de B. (2020). *Plant Diseases recognition on images using Convolutional Neural Networks: A Systematic Review*. Cornell University.
<https://doi.org/https://doi.org/10.48550/arXiv.2009.04365>
- Abdullah, H. M., Mohana, N. T., Khan, B. M., Ahmed, S. M., Hossain, M., Islam, K. S., Redoy, M. H., Ferdush, J., Bhuiyan, M. A. H. B., Hossain, M. M., & Ahamed, T. (2023). Present and future scopes and challenges of plant pest and disease (P&D) monitoring: Remote sensing, image processing, and artificial intelligence perspectives. *Remote Sensing Applications: Society and Environment*, 32, 100996.
<https://doi.org/10.1016/j.rsase.2023.100996>
- Abuhayi, B. M., & Mossa, A. A. (2023). Coffee disease classification using Convolutional Neural Network based on feature concatenation. *Informatics in Medicine Unlocked*, 39, 101245.
<https://doi.org/10.1016/j.imu.2023.101245>
- Adelaja, O., & Pranggono, B. (2025). Leveraging Deep Learning for Real-Time Coffee Leaf Disease Identification. *AgriEngineering*, 7(1), 13. <https://doi.org/10.3390/agriengineering7010013>
- Archana, R., & Jeevaraj, P. S. E. (2024). Deep learning models for digital image processing: a review.

Artificial Intelligence Review, 57(1), 11. <https://doi.org/10.1007/s10462-023-10631-z>

- Arif, A., Putrawansyah, F., & Jangcik, I. (2025). Detection of Coffee Leaf Diseases Using Lightweight Deep Learning: A Comparative Study of EfficientNet-B0 and Vision Transformer. *Ingénierie Des Systèmes d'Information*, 30(9), 2393–2404. <https://doi.org/10.18280/isi.300915>
- Atila, Ü., Uçar, M., Akyol, K., & Uçar, E. (2021). Plant leaf disease classification using EfficientNet deep learning model. *Ecological Informatics*, 61, 101182. <https://doi.org/10.1016/j.ecoinf.2020.101182>
- Aufar, Y., & Kaloka, T. P. (2022). Robusta coffee leaf diseases detection based on MobileNetV2 model. *International Journal of Electrical and Computer Engineering (IJECE)*, 12(6), 6675. <https://doi.org/10.11591/ijece.v12i6.pp6675-6683>
- Avelino, J., Cristancho, M., Georgiou, S., Imbach, P., Aguilar, L., Bornemann, G., Läderach, P., Anzueto, F., Hruska, A. J., & Morales, C. (2015). The coffee rust crises in Colombia and Central America (2008–2013): impacts, plausible causes and proposed solutions. *Food Security*, 7(2), 303–321. <https://doi.org/10.1007/s12571-015-0446-9>
- Ayikpa, K. J., Ayikpa, K. J., Ayikpa, K. J., Mamadou, D., Gouton, P., & Adou, K. J. (2022). Experimental Evaluation of Coffee Leaf Disease Classification and Recognition Based on Machine Learning and Deep Learning Algorithms. *Journal of Computer Science*, 18(12), 1201–1212. <https://doi.org/10.3844/jcssp.2022.1201.1212>
- Chavarro, A. F., Renza, D., & Ballesteros, D. M. (2023). Influence of Hyperparameters in Deep Learning Models for Coffee Rust Detection. *Applied Sciences*, 13(7), 4565. <https://doi.org/10.3390/app13074565>
- Ehrenbergerová, L., Kučera, A., Cienciala, E., Trochta, J., & Volařík, D. (2018). Identifying key factors affecting coffee leaf rust incidence in agroforestry plantations in Peru. *Agroforestry Systems*, 92(6), 1551–1565. <https://doi.org/10.1007/s10457-017-0101-x>
- Fragoso, J., Silva, C., Paixão, T., Alvarez, A. B., Júnior, O. C., Florez, R., Palomino-Quispe, F., Savian, L. G., & Trazzi, P. A. (2025). Coffee-Leaf Diseases and Pests Detection Based on YOLO Models. *Applied Sciences*, 15(9), 5040. <https://doi.org/10.3390/app15095040>
- Julca-Otiniano, A., Alvarado-Huamán, L., Castro-Cepero, V., Borjas-Ventura, R., Gómez-Pando, L., Pereira, A. P., Nielen, S., Ingelbrecht, I., Silva, M. do C., & Várzea, V. (2024). New Races of Hemileia vastatrix Detected in Peruvian Coffee Fields. *Agronomy*, 14(8), 1811. <https://doi.org/10.3390/agronomy14081811>
- Mansouri, N., Guessmi, H., & Alkhailil, A. (2024). A deep learning model for detection and classification of coffee-leaf diseases using the transfer-learning technique. *International Journal of Advances in Intelligent Informatics*, 10(3), 379. <https://doi.org/10.26555/ijain.v10i3.1573>
- Martinez, F., Montiel, H., & Martinez, F. (2022). A Machine Learning Model for the Diagnosis of Coffee Diseases. *International Journal of Advanced Computer Science and Applications*, 13(4). <https://doi.org/10.14569/IJACSA.2022.01304110>
- Novtahaning, D., Shah, H. A., & Kang, J.-M. (2022). Deep Learning Ensemble-Based Automated and High-Performing Recognition of Coffee Leaf Disease. *Agriculture*, 12(11), 1909. <https://doi.org/10.3390/agriculture12111909>
- Pham, T. C., Nguyen, V. D., Le, C. H., Packianather, M., & Hoang, V.-D. (2023). Artificial intelligence-based solutions for coffee leaf disease classification. *IOP Conference Series: Earth and Environmental Science*, 1278(1), 012004. <https://doi.org/10.1088/1755-1315/1278/1/012004>
- Poma-Angamarca, R. A., Rojas, J. R., Sánchez-Rodríguez, A., & Ruiz-González, M. X. (2024). Diversity of Leaf

Fungal Endophytes from Two Coffea arabica Varieties and Antagonism towards Coffee Leaf Rust.
Plants, 13(6), 814. <https://doi.org/10.3390/plants13060814>

Saavedra-Ramírez, J. (2023). Impacto social de los Proyectos de Extensión Agraria durante el 2005 al 2010 en San Martín y Amazonas, Perú. *Revista Amazónica de Ciencias Económicas*, 2(1), e433. <https://doi.org/10.51252/race.v2i1.433>

Siddiqua, A., Kabir, M. A., Ferdous, T., Ali, I. B., & Weston, L. A. (2022). Evaluating Plant Disease Detection Mobile Applications: Quality and Limitations. *Agronomy*, 12(8), 1869. <https://doi.org/10.3390/agronomy12081869>