



Universidad Nacional Mayor de San Marcos

Universidad del Perú. Decana de América

Dirección General de Estudios de Posgrado
Facultad de Ingeniería de Sistemas e Informática
Unidad de Posgrado

Modelo de análisis automático mediante reconocimiento de patrones para determinar enfermedades en hojas de cafeto

TESIS

Para optar el Grado Académico de Magíster en Ingeniería de
Sistemas e Informática con mención en Gestión de Tecnología de
Información y Comunicaciones

AUTOR

John Clark SANTA MARIA PINEDO

ASESOR

Dr. Ciro RODRIGUEZ RODRIGUEZ

Lima, Perú

2026



Reconocimiento - No Comercial - Compartir Igual - Sin restricciones adicionales

<https://creativecommons.org/licenses/by-nc-sa/4.0/>

Usted puede distribuir, remezclar, retocar, y crear a partir del documento original de modo no comercial, siempre y cuando se dé crédito al autor del documento y se licencien las nuevas creaciones bajo las mismas condiciones. No se permite aplicar términos legales o medidas tecnológicas que restrinjan legalmente a otros a hacer cualquier cosa que permita esta licencia.

Referencia bibliográfica

Santa Maria, J. (2026). *Modelo de análisis automático mediante reconocimiento de patrones para determinar enfermedades en hojas de café*. [Tesis de maestría, Universidad Nacional Mayor de San Marcos, Facultad de Ingeniería de Sistemas e Informática, Unidad de Posgrado]. Repositorio institucional Cybertesis UNMSM.

Metadatos complementarios

Datos de autor	
Nombres y apellidos	John Clark Santa Maria Pinedo.
Tipo de documento de identidad	DNI
Número de documento de identidad	41720526
URL de ORCID	https://orcid.org/0000-0002-8594-4865
Datos de asesor	
Nombres y apellidos	Ciro Rodriguez Rodriguez
Tipo de documento de identidad	DNI
Número de documento de identidad	06020241
URL de ORCID	https://orcid.org/0000-0003-2112-1349
Datos del jurado	
Presidente del jurado	
Nombres y apellidos	William Martín Enriquez Maguiña
Tipo de documento	DNI
Número de documento de identidad	06179457
Miembro del jurado 1	
Nombres y apellidos	César Augusto Angulo Calderón
Tipo de documento	DNI
Número de documento de identidad	32907109
Miembro del jurado 2	
Nombres y apellidos	Carlos Ernesto Chavez Herrera
Tipo de documento	DNI
Número de documento de identidad	40788722
Datos de investigación	
Línea de investigación	C.0.3.17. Reconocimiento de Patrones

Grupo de investigación	GI04 Inteligencia Artificial
Agencia de financiamiento	Sin financiamiento
Ubicación geográfica de la investigación	País: Perú Departamento: San Martín Provincia: Huallaga Distrito: Saposoa Centro poblado: Caserío Santa Rosa Latitud: -6.972967 Longitud: -76.858876
Año o rango de años en que se realizó la investigación	2021 - 2025
URL de disciplinas OCDE	Ciencias de la computación https://purl.org/pe-repo/ocde/ford#1.02.01 Ingeniería de sistemas y comunicaciones https://purl.org/pe-repo/ocde/ford#2.02.04



UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS

Universidad del Perú. Decana de América
Facultad de Ingeniería de Sistemas e Informática
Vicedecanato de Investigación y Posgrado
Unidad de Posgrado

**ACTA DE SUSTENTACIÓN DE TESIS PARA OPTAR EL GRADO ACADÉMICO DE
MAGÍSTER EN INGENIERÍA DE SISTEMAS E INFORMÁTICA CON MENCIÓN EN
GESTIÓN DE TECNOLOGÍA DE INFORMACIÓN Y COMUNICACIONES**

A los veinticinco (25) días del mes de junio de 2026, siendo las 11:00 am., se reunieron en el Auditorio Alfredo Celso Alva Bravo, el Jurado de Tesis conformado por los siguientes docentes:

Mg. William Martin Enriquez Maguiña (Presidente)

Mg. César Augusto Angulo Calderón (Miembro)

Mg. Carlos Ernesto Chavez Herrera (Miembro)

Dr. Ciro Rodriguez Rodriguez (Miembro Asesor)

Se inició la Sustentación invitando al candidato a Magíster **JOHN CLARK SANTA MARIA PINEDO**, para que realice la exposición oral de la tesis para optar el Grado Académico de Magíster en Ingeniería de Sistemas e Informática con mención en Gestión de Tecnología de Información y Comunicaciones., siendo la Tesis intitulada:

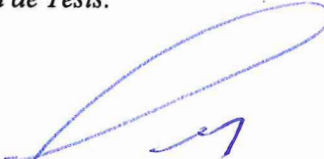
**"MODELO DE ANÁLISIS AUTOMÁTICO MEDIANTE RECONOCIMIENTO DE PATRONES
PARA DETERMINAR ENFERMEDADES EN HOJAS DE CAFETO"**

Concluida la exposición, los miembros del Jurado de Tesis procedieron a formular sus preguntas que fueron absueltas por el graduando; acto seguido se procedió a la evaluación correspondiente, habiendo obtenido la siguiente calificación:

VEINTE (20) EXCELENTE

Por tanto, el presidente del Jurado, de acuerdo con el Reglamento General de Estudios de Posgrado, otorga al Bachiller **JOHN CLARK SANTA MARIA PINEDO** el Grado Académico de Magíster en Ingeniería de Sistemas e Informática con mención en Gestión de Tecnología de Información y Comunicaciones,

Siendo las 2.00 horas, el presidente del Jurado de Tesis, da por concluido el acto académico de Sustentación de Tesis.


Mg. William Martin Enriquez Maguiña
(Presidente)


Mg. César Augusto Angulo Calderón
(Miembro)


Mg. Carlos Ernesto Chavez Herrera
(Miembro)


Dr. Ciro Rodriguez Rodriguez
(Miembro Asesor)



Anexo 1

CERTIFICADO DE SIMILITUD

Yo, **Ciro Rodriguez Rodriguez** en mi condición de asesor acreditado mediante Dictamen N°000638-2025-UPG-VDIP-FISI/UNMSM, cuyo título es, **“Modelo de análisis automático mediante reconocimiento de patrones para determinar enfermedades en hojas de cafeto”**, presentado por el bachiller **John Clark Santa Maria Pinedo** para optar el Grado Académico de Magíster en Ingeniería de Sistemas e Informática con mención en Gestión de Tecnología de Información y Comunicaciones, CERTIFICO que se ha cumplido con lo establecido en la Directiva de Originalidad y de Similitud de Trabajos Académicos, de Investigación y Producción Intelectual. Según la revisión, análisis y evaluación mediante la plataforma PIDS, el documento evaluado cuenta con el 4% de similitud, nivel PERMITIDO para continuar con los trámites correspondientes y para su **publicación en el repositorio institucional**.

Se emite el presente certificado en cumplimiento de lo establecido en las normas vigentes, como uno de los requisitos para la obtención del Grado Académico de Magíster en Ingeniería de Sistemas e Informática con mención en Gestión de Tecnología de Información y Comunicaciones.

Lima, 14 de abril de 2026

Firma del Asesor

Nombres y Apellidos del asesor: **Ciro Rodriguez Rodriguez**

DNI: 06020241

* Según corresponda: bachiller, magíster, doctor; título profesional o título de segunda especialidad

La vigencia del presente certificado corresponde a un período de noventa (90) días desde la fecha de su emisión; siempre que no haya observaciones en el texto durante el proceso de revisión por el jurado informante o surja alguna contingencia.

DEDICATORIA

Con profundo amor y gratitud, dedico este trabajo a mi madre, por su ejemplo y entrega; a mi esposa, por su apoyo inquebrantable y paciencia; y a mis hijos, por ser la inspiración más noble de mi vida.

AGRADECIMIENTO

Un especial agradecimiento a mi asesor de tesis, por su valiosa orientación, compromiso constante y permanente apoyo académico, los cuales fueron fundamentales para el desarrollo y culminación satisfactoria del presente trabajo de investigación.

Índice general

Página de aceptación o veredicto de la Tesis por los miembros del Jurado Examinador	Error!
Bookmark not defined.	
Página de dedicatoria y agradecimiento (Opcional)	3
Índice general	5
Lista de tablas	7
Lista de figuras	8
Resumen	9
Abstract	10
CAPÍTULO 1. INTRODUCCIÓN	1
1.1 Situación Problemática	1
1.2 Formulación del Problema	2
1.2.1 Problema general	2
1.2.2 Problemas específicos	2
1.3 Hipótesis	3
1.3.1 Hipótesis general.....	3
1.3.2 Hipótesis específicas	3
1.4 Justificación teórica	3
1.5 Justificación práctica	4
1.6 Objetivos	6
1.6.1 Objetivo general.....	6
1.6.2 Objetivos específicos	6
CAPÍTULO 2. MARCO TEÓRICO	7
2.1 Antecedentes de investigación	7
2.2 Bases Teóricas	9
2.2.1 Modelo de análisis automático mediante reconocimiento de patrones	9
2.2.2 Enfermedades en hojas de cafeto	10
2.3 Estado del Arte	11
2.3.1 Introducción.....	11
2.3.2 Revisión epistemológica	14
2.3.3 Metodología.....	15
2.3.4 Resultados y discusión	17
2.3.5 Conclusiones.....	21
CAPÍTULO 3. METODOLOGÍA	23
3.1 Variables	23
3.2 Operacionalización de las variables.....	23
3.3 Matriz de consistencia	1
3.4 Tipo, nivel y diseño de la investigación.....	1
3.5 Objeto de estudio, población y muestra	1

3.5.1	Tamaño de muestra	2
3.5.2	Selección de muestra.....	2
3.6	Técnicas de recolección de datos.....	3
3.6.1	Proceso de obtención de imágenes.....	3
3.6.2	Almacenamiento y organización de imágenes.....	3
3.6.3	Control de calidad	4
3.7	Instrumentos de recolección de datos	4
3.8	Tratamiento estadístico de los datos	5
3.8.1	Análisis descriptivo.....	5
3.8.2	Validación cruzada estratificada	5
3.8.3	Comparación de modelos mediante ANOVA	5
CAPÍTULO 4. DESARROLLO DEL MODELO DE ANÁLISIS AUTOMÁTICO MEDIANTE RECONOCIMIENTO DE PATRONES PARA DETERMINAR ENFERMEDADES EN HOJAS DE CAFETO		7
4.1	OE 1 Construcción del conjunto de datos de imágenes digitales de hojas de cafeto.....	7
4.1.1	Comprensión del negocio.....	7
4.1.2	Comprensión de los datos.	8
4.2	OE 2 Identificación de un modelo híbrido de análisis automático basado en reconocimiento de patrones, que integre técnicas de transferencia de conocimiento para la extracción de características y algoritmos de clasificación supervisada para identificar enfermedades en las hojas del cafeto	10
4.2.1	Preparación de los datos.....	11
4.2.2	Modelo híbrido 1: ResNet18 + SVM.....	17
4.2.3	Modelo híbrido 2: ResNet18 + Random Forest	24
4.2.4	Modelo híbrido 3: ResNet18 + XGBoost	29
4.3	OE 3 Evaluación del desempeño del modelo híbrido propuesto analizando métricas de rendimiento en la identificación de enfermedades en las hojas del cafeto	35
4.3.1	Visualización de la distribución de resultados por métrica y modelo.....	36
4.3.2	Resultados del análisis ANOVA	38
4.3.3	Aplicación de la prueba POST HOC	38
4.3.4	Interpretación y selección del modelo óptimo.....	39
CAPÍTULO 5. RESULTADOS Y DISCUSIÓN.....		40
5.1	Análisis, interpretación y discusión de resultados	40
5.2	Pruebas de hipótesis.....	40
5.2.1	Hipótesis específica H1	40
5.2.2	Hipótesis específica H2	41
5.2.3	Hipótesis específica H3	43
5.3	Discusión.....	44
CONCLUSIONES		47
RECOMENDACIONES		48
REFERENCIAS BIBLIOGRÁFICAS		49

Lista de tablas

Tabla 1. Clasificación de arquitecturas utilizadas en el reconocimiento de enfermedades en hojas de café	17
Tabla 2. Tipo de tarea predominante en los estudios	18
Tabla 3. Resumen de exactitud (“accuracy” extraída del mejor resultado cuando fue comparable) por categoría de arquitectura	19
Tabla 4. Datasets identificados con uso recurrente	20
Tabla 5. Accuracy de validación (val_acc) por fold en TRAIN (10-fold CV)	37
Tabla 6. ANOVA para val_acc en CV (TRAIN)	38
Tabla 7. Comparaciones múltiples (Tukey HSD) para val_acc en CV (TRAIN)	38
Tabla 8. Desempeño en TEST (20% hold-out)	39
Tabla 9. Conjunto de datos a comparar	41
Tabla 10. Valores de métricas evaluadas de cada conjunto de datos	41
Tabla 11. Resumen de métricas superiores por conjunto de datos	41
Tabla 12. Comparación de diagnósticos del modelo híbrido y del experto	43
Tabla 13. Escala de interpretación del coeficiente de la Kappa de Cohen (K)	43
Tabla 14. Resultado del análisis de concordancia entre el modelo híbrido y el experto	43

Lista de figuras

Figura 1. Proceso de construcción del dataset	8
Figura 2. Recolección de hojas	9
Figura 3. Ejemplares del dataset	10
Figura 4. Esquema del proceso de data augmentation aplicado al conjunto de imágenes	12
Figura 5. Ejemplares de aumento de datos aplicado por clase	13
Figura 6. Distribución por clases (antes y después)	14
Figura 7. Matriz de confusión del modelo híbrido ResNet18 + SVM en el conjunto de entrenamiento (80%)	20
Figura 8. Matriz de confusión del modelo híbrido ResNet18 + SVM en el conjunto de prueba (20%)	21
Figura 9. Curvas ROC OVR del modelo híbrido ResNet18 + SVM en el conjunto de entrenamiento (Macro-AUC=1.000)	22
Figura 10. Curvas ROC OVR del modelo híbrido ResNet18 + SVM en el conjunto de prueba (Macro-AUC=0.997)	23
Figura 11. Evolución de Accuracy y F1-macro por fold (10-fold CV) en el conjunto de entrenamiento	23
Figura 12. Matriz de confusión del modelo ResNet18 + RandomForest en el conjunto de entrenamiento	27
Figura 13. Matriz de confusión del modelo ResNet18 + RandomForest en el conjunto de prueba	27
Figura 14. Evolución de la accuracy y el F1-macro	29
Figura 15. Matriz de confusión del modelo ResNet18 + XGBoost en el conjunto de entrenamiento	32
Figura 16. Matriz de confusión del modelo ResNet18 + XGBoost en el conjunto de prueba	33
Figura 17. Curvas ROC OVR del modelo ResNet18 + XGBoost en entrenamiento	34
Figura 18. Curvas ROC OVR del modelo ResNet18 + XGBoost en prueba	34
Figura 19. Evolución de la accuracy y el F1-macro	35
Figura 20. Distribución de val_acc en CV por modelo	36
Figura 21. Gráfico de medias de val_acc en CV (TRAIN) por modelo	37
Figura 22. F1-macro en TEST (20%) por modelo	37

Resumen

La finalidad de esta investigación fue diseñar y evaluar un modelo de análisis automático basado en reconocimiento de patrones para la identificación de enfermedades foliares del cafeto, con el propósito de apoyar el diagnóstico fitosanitario en condiciones reales de campo en el distrito de Saposoa, región San Martín, Perú. El objetivo general fue desarrollar un modelo de análisis automático basado en reconocimiento de patrones para identificar y clasificar enfermedades en hojas de cafeto. Se empleó un enfoque cuantitativo para el desarrollo del estudio, de nivel explicativo y alcance tecnológico, siguiendo una estructura metodológica basada en el estándar CRISP-DM. Como procedimiento principal, se construyó un dataset propio compuesto por imágenes digitales capturadas en campo y validadas por especialistas en sanidad vegetal, al cual se aplicó un aumento de datos controlado para mejorar la capacidad de generalización. Se empleó ResNet18 como extractor de características mediante transferencia de conocimiento, combinándolo con tres clasificadores supervisados: Support Vector Machine (SVM), Random Forest y XGBoost. El rendimiento de los modelos fue evaluado mediante validación cruzada estratificada de 10 pliegues y un conjunto de prueba independiente, utilizando métricas de exactitud, F1-macro y AUC-ROC. Los resultados obtenidos pusieron de manifiesto diferencias relevantes a nivel estadístico entre los modelos evaluados (ANOVA, $p < 0.05$), destacando el modelo ResNet18 + SVM como el de mejor desempeño global, con alta estabilidad y capacidad discriminativa. En conclusión, la investigación demuestra que los modelos híbridos basados en transferencia de conocimiento constituyen una estrategia eficaz y viable para la clasificación automatizada de enfermedades foliares del cafeto, ofreciendo un recurso replicable y con potencial de implementación en herramientas informáticas diseñadas para respaldar la toma de decisiones agrícolas en contextos con recursos limitados.

Palabras clave: aprendizaje automático, clasificación de imágenes, enfermedades foliares, transferencia de aprendizaje, visión artificial

Abstract

The purpose of this research was to design and evaluate an automatic analysis model based on pattern recognition for the identification of foliar diseases in coffee plants, with the aim of supporting phytosanitary diagnosis under real field conditions in the district of Saposoa, San Martín region, Peru. The general objective was to develop an automatic pattern-recognition-based model to identify and classify diseases in coffee leaves. A quantitative approach was adopted for the development of the study, at an explanatory level and with a technological scope, following a methodological framework based on the CRISP-DM standard. As the main procedure, a proprietary dataset was constructed, consisting of digital images captured in the field and validated by plant health specialists. Controlled data augmentation was applied to enhance the generalization capability of the models. ResNet18 was employed as a feature extractor through transfer learning and combined with three supervised classifiers: Support Vector Machine (SVM), Random Forest, and XGBoost. Model performance was evaluated using stratified 10-fold cross-validation and an independent test set, applying accuracy, macro F1-score, and AUC-ROC as evaluation metrics. The results revealed statistically significant differences among the evaluated models (ANOVA, $p < 0.05$), highlighting the ResNet18 + SVM model as the best overall performer, demonstrating high stability and discriminative capability. In conclusion, the research demonstrates that hybrid models based on transfer learning constitute an effective and viable strategy for the automated classification of coffee foliar diseases, providing a replicable approach with strong potential for implementation in computational tools designed to support agricultural decision-making in resource-limited contexts.

Keywords: computer vision, image classification, machine learning, plant disease detection, transfer learning

CAPÍTULO 1. INTRODUCCIÓN.

1.1 Situación Problemática

La detección temprana y precisa de enfermedades foliares del cafeto es fundamental para preservar la productividad y la persistencia económica de los pequeños productores. Sin embargo, en la caficultura de América Latina persisten deficiencias significativas en el diagnóstico oportuno de patologías críticas como la roya del café (*Hemileia vastatrix*) y el ojo de gallo (*Mycena citricolor*), las cuales representan las principales amenazas para la sostenibilidad del cultivo. Diversas investigaciones documentan cómo la roya ha generado crisis recurrentes con impactos severos en la producción, la estabilidad financiera y la seguridad alimentaria de los pequeños caficultores (Avelino et al., 2015a). Asimismo, el ojo de gallo, ampliamente extendido en zonas húmedas, causa defoliación severa y pérdidas considerables de rendimiento, especialmente en plantaciones bajo sombra y condiciones de alta humedad (Poma-Angamarca et al., 2024).

Los métodos tradicionales de diagnóstico principalmente la inspección visual realizada por técnicos y agricultores continúan siendo el mecanismo predominante para identificar estas enfermedades en campo. Sin embargo, este procedimiento presenta desafíos importantes: variabilidad en la interpretación, dependencia del nivel de experiencia del evaluador y condiciones ambientales que pueden modificar la apariencia visual de los síntomas (Abade et al., 2020). Tanto las clasificaciones erróneas por exceso como aquellas por omisión pueden generar decisiones agronómicas erróneas, incrementando los costos de manejo, retrasando intervenciones y reduciendo la eficacia de los programas de control (Abdullah et al., 2023). Esta situación es especialmente problemática en pequeños productores, quienes deben asumir los costos económicos derivados de diagnósticos incorrectos, afectando su capacidad financiera y la estabilidad del sistema productivo.

En este contexto, el diagnóstico de enfermedades foliares del cafeto se ve condicionado no solo por la diversidad de patógenos presentes, sino también por la alta variabilidad en la expresión visual de los síntomas a lo largo del ciclo del cultivo. Factores como las condiciones microclimáticas locales, la estructura del sistema agroforestal y el estado fenológico de la planta pueden modificar la apariencia, intensidad y distribución de los síntomas, lo que complica su identificación temprana y confiable en campo (Sera et al., 2022). Asimismo, se ha observado que distintas enfermedades y estreses abióticos pueden generar manifestaciones visuales similares, lo que aumenta el riesgo de confusión diagnóstica cuando se emplean métodos exclusivamente visuales (Paulos & Woldeyohannis, 2022; Siddiqua et al., 2022).

En la región San Martín (Perú), la caficultura constituye uno de los pilares económicos y sociales más importantes para los pequeños productores. Estudios recientes han identificado la presencia de nuevas razas de *H. vastatrix* en campos peruanos, lo que incrementa la complejidad del manejo de la enfermedad y la probabilidad de fallas en los programas de control (Julca-Otiniano et al.,

2024a). Además, investigaciones realizadas en plantaciones del país muestran que la incidencia de estas enfermedades está influenciada por factores como la sombra, la altitud y la estructura del sistema agroforestal, lo que aumenta la variabilidad en la manifestación de síntomas y dificulta aún más el diagnóstico visual (Ehrenbergerová et al., 2018). A esto se suma que, en zonas como San Martín, el número de técnicos disponibles suele ser menor que la demanda real, lo que genera sobrecarga operativa, diagnósticos tardíos y limitada capacidad de monitoreo fitosanitario.

Adicionalmente, aunque la investigación internacional muestra avances significativos en modelos de clasificación de enfermedades en café incluyendo enfoques basados en EfficientNet, MobileNetV2 o YOLO (Adelaja & Pranggono, 2025; Fragoso et al., 2025; Shetty, 2024) la mayoría de estos desarrollos se han centrado en enfermedades como roya, cercospora o antracnosis, mientras que la clasificación del ojo de gallo aparece de manera menos frecuente o completamente ausente en algunos datasets, a pesar de su relevancia agronómica (Novtahaning et al., 2022). Asimismo, muchos modelos trabajan con hojas recortadas y sobre fondos uniformes, condiciones que distan de la realidad de campo donde las imágenes presentan oclusiones, sombras, variaciones climáticas y múltiples estados intermedios de infección (Chavarro et al., 2023).

Esta revisión evidencia que, aunque existen múltiples esfuerzos por automatizar el diagnóstico de enfermedades del cafeto, persiste una brecha crítica: la ausencia de un modelo robusto, contextualizado y entrenado con imágenes reales de cafetales peruanos, capaz de identificar de manera precisa y diferenciada la roya y el ojo de gallo en condiciones propias de la región San Martín. Esta carencia limita la habilidad de los productores para realizar decisiones en momentos críticos, incrementa el riesgo de pérdidas económicas y compromete la sostenibilidad de la caficultura local.

Por lo tanto, resulta necesario desarrollar un modelo de análisis automático mediante reconocimiento de patrones que permita determinar con alta precisión la presencia de roya y ojo de gallo a partir de imágenes digitales de hojas de cafeto capturadas en campo. Dicho modelo debe integrar técnicas de visión artificial y aprendizaje profundo, trabajar con un dataset propio y representar adecuadamente las condiciones agroecológicas de San Martín, a fin de ofrecer un diagnóstico objetivo, reproducible y aplicable, que complemente el análisis visual tradicional y contribuya a mejorar el proceso de decisión vinculados al manejo integral de enfermedades.

1.2 Formulación del Problema

1.2.1 Problema general

- ¿Cómo un modelo de análisis automático mediante reconocimiento de patrones identifica y clasifica enfermedades en hojas de cafeto?

1.2.2 Problemas específicos

- ¿Cómo estructurar un conjunto de imágenes de hojas de cafeto sanas y enfermas que permita entrenar adecuadamente un modelo automático de clasificación?

- ¿Qué técnicas de transferencia de conocimiento y algoritmos supervisados permiten construir un modelo híbrido capaz de extraer características relevantes y clasificar enfermedades en las hojas de café?
- ¿Cuál es el desempeño del modelo híbrido propuesto para la identificación y clasificación de enfermedades en hojas de café?

1.3 Hipótesis

1.3.1 Hipótesis general

El desarrollo de un modelo de análisis automático mediante reconocimiento de patrones identifica y clasifica enfermedades en hojas de café en el distrito de Saposa.

Hipótesis nula (H_0): El desarrollo de un modelo de análisis automático mediante reconocimiento de patrones no identifica ni clasifica enfermedades en hojas de café en el distrito de Saposa.

Hipótesis alternativa (H_1): El desarrollo de un modelo de análisis automático mejora significativamente la identificación y clasificación de enfermedades en hojas de café en el distrito de Saposa.

1.3.2 Hipótesis específicas

- La estructuración de un conjunto de imágenes de hojas de café sanas y enfermas permite entrenar adecuadamente un modelo automático de clasificación.
- La aplicación de técnicas de transferencia de conocimiento y algoritmos supervisados permite construir un modelo híbrido capaz de extraer características relevantes y clasificar enfermedades en hojas de café.
- El modelo híbrido propuesto presenta un desempeño adecuado para la identificación y clasificación de enfermedades en hojas de café.

1.4 Justificación teórica

Esta investigación se basa en algunas ideas importantes. Una de ellas es el reconocimiento de patrones. Esto significa que un sistema puede encontrar cosas que se repiten y estructuras importantes en datos complejos. De esta manera, se puede clasificar información visual de forma objetiva y sin errores. El reconocimiento de patrones es útil cuando se trabaja con imágenes digitales. Permite convertir la información visual en números que un ordenador puede entender (Bishop, 2006). Esto facilita que el ordenador identifique automáticamente clases que ya han sido definidas. La visión artificial y el aprendizaje automático también son fundamentales en esta investigación. Todos estos conceptos son parte de la inteligencia artificial y se utilizan para analizar imágenes de manera efectiva.

Diversos estudios señalan que las redes neuronales convolucionales (CNN) representan el marco teórico más sólido para el análisis automático de imágenes, debido a su capacidad para aprender jerarquías de características visuales tales como bordes, texturas y patrones complejos directamente a partir de los datos, sin requerir una definición manual de descriptores (LeCun et

al., 2015a). Las redes neuronales convolucionales son muy buenas para clasificar imágenes y se usan mucho en la agricultura y para cuidar las plantas (Ferentinos, 2018a).

Sin embargo, hay algunos problemas con los modelos de aprendizaje profundo. No funcionan bien cuando no hay muchos datos o cuando las condiciones son muy diferentes, como sucede en los campos de cultivo. En estos casos, una buena estrategia es el aprendizaje por transferencia. Esto significa usar lo que los modelos han aprendido de grandes colecciones de imágenes para mejorar la capacidad de generalizar y reducir la necesidad de datos y recursos computacionales (Pan & Yang, 2010a; Too et al., 2019a).

Estudios recientes muestran que la combinación de modelos híbridos es una buena forma de lograr un equilibrio entre la capacidad de representar datos, la estabilidad y la eficiencia computacional. Esto se consigue uniendo extractores de características basados en redes neuronales convolucionales con clasificadores tradicionales como SVM, Random Forest o XGBoost (Abuhayi & Mossa, 2023a; Ayikpa, Ayikpa, et al., 2022). Este enfoque permite separar la forma en que se representan las imágenes de la forma en que se toman las decisiones. Esto facilita analizar cómo funciona cada modelo, compararlos entre sí y adaptarlos a diferentes situaciones. La inteligencia artificial es muy útil en la agricultura. La visión por computadora es una herramienta importante que ayuda a monitorear los cultivos de manera sistemática. Esto permite tomar decisiones informadas con datos objetivos (Barbedo, 2016a; Karunathilake et al., 2023a). Por ejemplo, se pueden usar modelos automáticos para identificar enfermedades en las hojas de las plantas. Esto reduce la subjetividad del diagnóstico visual tradicional y establece criterios de evaluación estándar. Esto es importante para mejorar la confiabilidad del monitoreo fitosanitario, como se menciona en varios estudios (Mohanty et al., 2016a; Siddiqua et al., 2022).

En este sentido, la presente investigación se justifica teóricamente porque integra y aplica de manera coherente los fundamentos del reconocimiento de patrones, el aprendizaje profundo, la transferencia de conocimiento y los modelos híbridos de clasificación, aportando evidencia empírica sobre su desempeño en un dominio agrícola específico. Asimismo, contribuye al cuerpo teórico existente al analizar estas estrategias bajo condiciones reales de campo y con un conjunto de datos contextualizado, fortaleciendo la comprensión de los alcances y limitaciones de los modelos de inteligencia artificial en la identificación automática de enfermedades foliares del cafeto.

1.5 Justificación práctica

La investigación surge de una problemática concreta y muy cercana a la realidad agrícola de la Región San Martín. Las enfermedades del cafeto comprometen el rendimiento y la calidad del cultivo, afectando la estabilidad productiva de los agricultores (Avelino et al., 2015b). En el contexto de las familias cafetaleras, dicha afectación puede repercutir en la estabilidad económica, debido a la dependencia de sus ingresos respecto de la producción y comercialización del café (Gichuru et al., 2021). En estudios recientes, se reconoce además la necesidad de fortalecer

estrategias tecnológicas de diagnóstico y monitoreo para responder oportunamente a los problemas fitosanitarios del cultivo (Koutouleas et al., 2024).

El problema se agrava debido a que la identificación de enfermedades en campo suele depender exclusivamente de la percepción visual del agricultor, un proceso inherentemente subjetivo y susceptible a errores, especialmente cuando las lesiones son incipientes o cuando no se cuenta con personal técnico capacitado en todas las zonas de cultivo (Julca-Otiniano et al., 2024b). Esta incertidumbre retrasa el proceso de decisión y, en diversas situaciones, conduce a la aplicación tardía o inadecuada de fungicidas, elevando los costos de producción y reduciendo significativamente el rendimiento del cultivo (Salazar-Navarro et al., 2024a).

En este contexto, propongo el desarrollo de un modelo híbrido basado en visión artificial y aprendizaje automático, capaz de analizar imágenes digitales de hojas de café y clasificar automáticamente si una planta se encuentra sana o afectada por roya u ojo de gallo (Abuhayi & Mossa, 2023b; Martínez et al., 2022a). Esta propuesta no busca simplemente automatizar un proceso manual, sino reducir la subjetividad diagnóstica, estandarizar la evaluación, y ofrecer a los productores un sistema que traduzca la imagen de una hoja en un diagnóstico objetivo, cuantificado y reproducible (Ayikpa, Mamadou, et al., 2022). Estudios recientes han demostrado que los modelos híbridos basados en transferencia de conocimiento, combinados con clasificadores como SVM, Random Forest o XGBoost, pueden alcanzar precisiones muy elevadas para detectar enfermedades agrícolas (Novtahaning et al., 2022; Pham et al., 2023).

En términos metodológicos, el estudio representa una innovación porque integra un enfoque moderno de transfer learning con el uso de clasificadores tradicionales optimizados, permitiendo aprovechar lo mejor del deep learning sin requerir datasets masivos como los utilizados en investigaciones internacionales (Karunathilake et al., 2023b). Esta estrategia permitirá emular, mediante patrones visuales cuantificados, el proceso de diagnóstico que realiza un especialista en sanidad vegetal, pero con acceso directo para productores y técnicos de la región (Mendes et al., 2020). Para ello se trabajará con imágenes reales recolectadas directamente en las hectáreas de cultivo de café ubicadas en el distrito de Saposoa, asegurando representatividad y pertinencia territorial, como recomiendan diversas investigaciones en el marco de los enfoques de agricultura de precisión (Siddiqua et al., 2022).

El impacto esperado es significativo. En el plano productivo, el modelo permitirá que los agricultores tomen decisiones oportunas sobre el manejo fitosanitario, evitando pérdidas y aumentando la eficiencia del uso de insumos. En el plano económico, una detección más temprana de enfermedades contribuye a mantener la calidad del café orgánico. En el plano social, una mejora en el rendimiento de los cultivos se traduce en mayor estabilidad financiera para las familias cafetaleras, fortaleciendo las cadenas de valor locales y promoviendo el bienestar comunitario.

Finalmente, el estudio no solo generará un modelo automatizado de diagnóstico, sino también un

dataset digital estructurado de hojas sanas y enfermas, que podrá servir como base para futuras investigaciones en inteligencia artificial aplicada al sector agropecuario peruano (Ristaino et al., 2021). Esta contribución se alinea con la tendencia global hacia el uso de tecnologías inteligentes en agricultura, cuyo objetivo es aumentar la resiliencia productiva frente a enfermedades que, como la roya, siguen constituyendo un riesgo persistente para la seguridad alimentaria y económica de miles de agricultores (Zheng & Kamruzzaman, 2024).

1.6 Objetivos

1.6.1 Objetivo general

- Desarrollar un modelo de análisis automático basado en reconocimiento de patrones para identificar y clasificar enfermedades en hojas de cafeto.

1.6.2 Objetivos específicos

- Estructurar un conjunto de imágenes digitales de hojas de cafeto sanas y enfermas para entrenar adecuadamente un modelo automático de clasificación.
- Determinar las técnicas de transferencia de conocimiento y algoritmos supervisados para construir un modelo híbrido capaz de extraer características relevantes y clasificar enfermedades en las hojas de cafeto.
- Evaluar el desempeño del modelo híbrido propuesto para la identificación y clasificación de enfermedades en hojas de cafeto.

CAPÍTULO 2. MARCO TEÓRICO

2.1 Antecedentes de investigación

En el estudio de (Abuhayi & Mossa, 2023a) señalan que la identificación precoz de enfermedades del café es primordial para garantizar la calidad y productividad del cultivo. Para abordar este problema, propusieron un modelo de aprendizaje profundo dividido en tres fases: preprocesamiento, extracción de características y clasificación. Utilizaron técnicas de suavizado gaussiano junto con técnicas de ampliación del conjunto de datos para optimizar la fidelidad visual del material de imagen y la robustez del modelo. Posteriormente, combinaron arquitecturas CNN basadas en GoogLeNet y ResNet para extraer características de alto nivel, las cuales fueron clasificadas mediante perceptrones multicapa y clasificadores de conjunto. Sus resultados demostraron una precisión del 99,08 %, superando los modelos tradicionales, y concluyen que la combinación de CNN y técnicas adecuadas de preprocesamiento constituye un método altamente eficaz para la identificación automática de enfermedades del café.

Así también, (Mansouri et al., 2024; Martinez et al., 2022b) destacan que muchas enfermedades del café son altamente contagiosas y requieren una detección temprana para evitar pérdidas. El estudio plantea un modelo de aprendizaje profundo fundamentado en transfer learning, utilizando 195 arquitecturas preentrenadas —desde modelos ligeros como MobileNet hasta modelos densos como EfficientNet y ResNet— para la detección de cuatro enfermedades foliares. El equipo evaluó el rendimiento de cada modelo ajustando los parámetros mediante técnicas de entrenamiento de pocos disparos. Los resultados indican que EfficientNet-B0 logró la mayor precisión con 99,99 %, posicionándolo como la arquitectura más adecuada para construir aplicaciones móviles o plataformas de apoyo a caficultores. Concluyen que el aprendizaje por transferencia facilita el desarrollo de soluciones de diagnóstico precisas y eficientes, incluso con recursos computacionales limitados.

Por su parte, (Novtahaning et al., 2022) analizan la necesidad de sistemas de reconocimiento robustos para las enfermedades más comunes de las hojas de café. Su estudio clasificó cuatro patologías principales (phoma, minador, roya y cercospora) mediante un ensamble de modelos de deep learning basado en CNN preentrenadas. Para ello, extrajeron características profundas utilizando múltiples arquitecturas, seleccionaron las tres de mayor rendimiento y concatenaron sus vectores para formar una capa de conjunto. Además, se incorporaron estrategias de aumento de datos con el fin de incrementar la capacidad de generalización del sistema. Su enfoque obtuvo una exactitud de 97,31 % en validación, superando a las redes individuales, y concluyen que los modelos ensamblados mejoran la estabilidad y reducen los errores en escenarios con variaciones de iluminación y estados intermedios de infección.

En otro estudio, (Fragoso et al., 2025) abordaron la detección de enfermedades y plagas del café utilizando modelos de la familia YOLO, reconocidos por su capacidad de detección en tiempo real. Emplearon las versiones YOLO v8, YOLO v9, YOLO v10 y YOLO v11, junto con sus

variantes, utilizando el conjunto de datos BRACOL anotado por expertos. Evaluaron el rendimiento de cada modelo mediante métricas como mAP, F1-Score y recall, incorporando análisis cuantitativos y cualitativos. Identificaron que YOLO v8s ofreció el mejor desempeño, con un mAP de 54,5 % y un tiempo de inferencia de 11,4 ms, lo cual lo convierte en una alternativa ideal para aplicaciones en campo que requieren rapidez y precisión simultánea. El estudio concluye que los modelos de detección resultan especialmente útiles para la localización temprana y precisa de lesiones en hojas de café.

Asimismo, (Adelaja & Pranggono, 2025) abordan las dificultades de aplicar métodos tradicionales de identificación de enfermedades, que requieren mano de obra especializada y no permiten análisis en tiempo real. Proponen un sistema fundamentado en métodos basados en redes neuronales profundas orientadas a la identificación instantánea de enfermedades en hojas de café usando diferentes modelos preentrenados: ResNet 101, Xception, VGG16 y su modelo propuesto, CoffNet. Utilizaron un conjunto diverso de imágenes tomadas con distintos dispositivos y evaluaron el rendimiento en términos de precisión y velocidad. Los resultados muestran que CoffNet alcanzó la mayor precisión con 99,89 % y una velocidad de 125,93 fps, lo que lo hace adecuado para aplicaciones portátiles y sistemas móviles de monitoreo agrícola. Concluyen que este tipo de modelos puede reducir significativamente la carga laboral de los agricultores y mejorar la capacidad de respuesta ante brotes de enfermedades.

Finalmente, (Atila et al., 2021) señalan que, aunque la mayoría de las enfermedades de las plantas presentan síntomas visibles en las hojas, el diagnóstico convencional sigue dependiendo de la observación óptica realizada por fitopatólogos experimentados, un proceso lento, subjetivo y estrechamente ligado a la pericia del especialista. Para superar estas limitaciones, proponen el uso de la arquitectura de aprendizaje profundo EfficientNet para la clasificación automática de enfermedades en hojas de plantas y comparan su desempeño con otros modelos de deep learning de vanguardia. Utilizan el conjunto de datos PlantVillage, que abarca 39 categorías asociadas a distintos tipos de enfermedades y estados foliares, entrenando los modelos tanto con el conjunto original (55 448 imágenes) como con un conjunto aumentado (61 486 imágenes), sin requerir una extracción manual de características ni un preprocesamiento complejo. Todos los modelos se entrenan bajo un esquema de aprendizaje por transferencia, configurando todas las capas como entrenables para ajustar finamente los parámetros a la tarea específica. Los resultados obtenidos en el conjunto de prueba muestran que los modelos EfficientNet-B5 y EfficientNet-B4 alcanzan los valores más altos frente al resto de arquitecturas evaluadas, logrando precisiones de 99,91 % y 99,97 %, y exactitudes de 98,42 % y 99,39 %, respectivamente. Los autores dicen que EfficientNet funciona mejor para clasificar enfermedades de las hojas. Los resultados sugieren que la arquitectura propuesta presenta un desempeño prometedor para su aplicación en sistemas de diagnóstico asistido por computadora en cultivos como el café.

2.2 Bases Teóricas

2.2.1 Modelo de análisis automático mediante reconocimiento de patrones

El análisis automático mediante reconocimiento de patrones se ha convertido en una herramienta clave dentro de la inteligencia artificial, especialmente cuando se trabaja con datos complejos como imágenes digitales. En términos generales, este enfoque busca identificar regularidades que permitan diferenciar una clase de otra a partir de características observables. En el procesamiento de imágenes, esto implica que el sistema no se basa en reglas rígidas previamente programadas, sino que aprende a reconocer estructuras visuales que se repiten en los datos, reduciendo así la dependencia de evaluaciones subjetivas (LeCun et al., 2015b).

En este contexto, la visión artificial aporta los métodos necesarios para capturar y procesar las imágenes antes de su análisis. Más que limitarse a “leer” una imagen, el sistema debe convertirla en información numérica que pueda ser interpretada computacionalmente. Este paso es fundamental, ya que transforma una representación visual en un conjunto de valores que permiten realizar comparaciones, mediciones y clasificaciones de manera objetiva (M. & Zahraddeen Yakuu, 2024).

Por su parte, el aprendizaje profundo ha cobrado especial relevancia en los últimos años debido a su capacidad para trabajar directamente con grandes volúmenes de datos visuales. A diferencia de los enfoques tradicionales, en los que era necesario definir manualmente qué características extraer, las redes profundas aprenden estas representaciones de forma automática a partir de los datos de entrenamiento. Esta característica ha permitido mejorar significativamente el desempeño en tareas como clasificación, detección y segmentación de imágenes (Ferentinos, 2018; LeCun et al., 2015).

Dentro de estas arquitecturas, las redes neuronales convolucionales (CNN) han demostrado ser particularmente adecuadas para el análisis de imágenes, ya que están diseñadas para captar patrones espaciales y texturales. Su estructura les permite adaptarse a variaciones comunes en las imágenes, como cambios de iluminación o de escala, lo que explica su amplio uso en problemas de reconocimiento visual (Atila et al., 2021; Too et al., 2019).

No obstante, es importante reconocer que estos modelos también presentan limitaciones. El entrenamiento de redes profundas suele requerir recursos computacionales elevados y conjuntos de datos amplios para alcanzar un buen nivel de generalización. Estas exigencias han impulsado la búsqueda de estrategias que permitan optimizar su rendimiento o combinarlos con otros métodos dentro del reconocimiento de patrones, especialmente en escenarios donde los datos o los recursos son limitados (Barbedo, 2016b).

En conjunto, los modelos de análisis automático mediante reconocimiento de patrones basados en inteligencia artificial constituyen una base teórica sólida para el desarrollo de sistemas de clasificación visual, permitiendo abordar problemas complejos de identificación y categorización de información a partir de imágenes digitales de manera objetiva, reproducible y escalable.

2.2.2 Enfermedades en hojas de cafeto

El cafeto (*Coffea arabica* L.) es un cultivo perenne de gran importancia económica y social en las regiones tropicales y subtropicales, particularmente en América Latina, donde constituye una de las principales fuentes de ingreso para pequeños productores. Las plantaciones de cafeto se desarrollan comúnmente bajo sistemas agroforestales o de sombra, los cuales influyen directamente en el microclima del cultivo y condicionan la incidencia, severidad y distribución de enfermedades fitosanitarias que afectan a la planta (Avelino et al., 2015a; Ehrenbergerová et al., 2018).

Las hojas del cafeto cumplen un rol fisiológico fundamental en los procesos de fotosíntesis, respiración y transpiración, por lo que cualquier alteración en su estructura o funcionalidad repercute directamente en el crecimiento vegetativo, el llenado del fruto y el rendimiento final del cultivo. Debido a su constante exposición a factores ambientales y biológicos, las hojas constituyen uno de los órganos más susceptibles al ataque de patógenos, especialmente hongos fitopatógenos, los cuales pueden provocar enfermedades foliares de alta incidencia y difícil control (Avelino et al., 2015a; Salazar-Navarro et al., 2024b).

Entre las principales enfermedades foliares del cafeto destaca la roya del café, causada por el hongo *Hemileia vastatrix*, considerada una de las patologías más destructivas del cultivo a nivel mundial. Esta enfermedad se manifiesta mediante la aparición de pústulas anaranjadas en el envés de las hojas, acompañadas de clorosis, necrosis y defoliación prematura, lo que reduce significativamente la capacidad fotosintética de la planta y afecta de manera directa la productividad del cultivo. Estudios recientes señalan que la roya del café continúa representando una amenaza constante para las plantaciones debido a su elevada capacidad de dispersión, adaptación y a la aparición de nuevas razas del patógeno, lo que incrementa la complejidad de su manejo fitosanitario (Salazar-Navarro et al., 2024b).

Otra enfermedad foliar de relevancia agronómica es el ojo de gallo, causado por el hongo *Mycena citricolor*. Esta patología se caracteriza por la presencia de manchas circulares bien delimitadas, con centros cloróticos o necróticos, que pueden confluir y provocar una defoliación severa en estados avanzados de infección. El ojo de gallo se presenta con mayor frecuencia en condiciones de alta humedad, sombra intensa y precipitaciones constantes, condiciones habituales en muchas zonas cafetaleras tropicales, donde puede generar pérdidas productivas significativas al reducir el área foliar funcional de la planta (Poma-Angamarca et al., 2024; SENASICA, 2014).

En el contexto de las plantaciones cafetaleras, diversos estudios han evidenciado que la incidencia y severidad de las enfermedades foliares están estrechamente relacionadas con factores ambientales, el manejo agronómico y el estado fenológico de la planta. Investigaciones realizadas en cafetales han demostrado que existe una correlación significativa entre la incidencia y la severidad de la roya del café, lo que resalta la importancia de una identificación adecuada y oportuna de los síntomas para la toma de decisiones eficaces en el manejo fitosanitario del cultivo

(Acosta Trinidad et al., 2025).

Las enfermedades foliares del cafeto, en particular la roya del café y el ojo de gallo, representan una de las principales limitantes para la sostenibilidad y productividad del cultivo. La correcta identificación de estas patologías resulta esencial para la implementación de estrategias de manejo oportunas y eficaces, orientadas a reducir pérdidas económicas, optimizar el uso de insumos agrícolas y garantizar la estabilidad de los sistemas productivos cafetaleros.

2.3 Estado del Arte

2.3.1 Introducción

La sanidad foliar del cafeto constituye un elemento decisivo para la productividad y sostenibilidad del cultivo, especialmente en sistemas de pequeña agricultura donde las pérdidas por enfermedades pueden comprometer seriamente la estabilidad económica de los productores (Alvarado-Huamán et al., 2020). Entre las patologías más relevantes destacan la roya del café (*Hemileia vastatrix*) y el ojo de gallo (*Mycena citricolor*), debido a su alta incidencia en regiones tropicales húmedas y a su capacidad de provocar defoliación, reducción del rendimiento y deterioro de la calidad del grano (Avelino et al., 2015a). La identificación oportuna de estas enfermedades resulta esencial para implementar estrategias de manejo oportunas; sin embargo, en la práctica agrícola el diagnóstico continúa dependiendo mayoritariamente de la inspección visual, un proceso sujeto a la experiencia del evaluador, a la variabilidad ambiental y a la similitud visual entre distintas afecciones foliares (Atila et al., 2021; Barbedo, 2016b).

En los últimos años, diversos estudios han explorado la incorporación de métodos de inteligencia artificial y visión computacional orientados a la automatización del diagnóstico de enfermedades foliares del cafeto (Abuhayi & Mossa, 2023a; Mansouri et al., 2024), con el objetivo de optimizar la precisión, rapidez y objetividad del diagnóstico fitosanitario integral. En este contexto, las redes neuronales convolucionales (CNN) han sido ampliamente empleadas para la obtención de descriptores visuales y la clasificación de hojas sanas y afectadas, mientras que enfoques basados en transfer learning han permitido aprovechar modelos preentrenados para compensar la limitada disponibilidad de datos (Adelaja & Pranggono, 2025; Kumar et al., 2020; Nawaz et al., 2024). Más recientemente, se han incorporado arquitecturas avanzadas y modelos híbridos que combinan distintas estrategias de aprendizaje orientadas a mejorar el desempeño en escenarios complejos (Mehta et al., 2025; Thakur et al., 2024). Sin embargo, a pesar de los resultados promisorios reportados, persisten limitaciones metodológicas relevantes: muchos modelos dependen de conjuntos de datos reducidos o capturados en condiciones controladas, presentan problemas de sobreajuste y muestran una capacidad limitada de generalización cuando se aplican a imágenes adquiridas en condiciones reales de campo, caracterizadas por variaciones de iluminación, fondo y estado fenológico (Ferentinos, 2018b; Mohanty et al., 2016b; Too et al., 2019b). En este contexto, surge una interrogante clave: ¿Qué arquitecturas y enfoques de IA se han utilizado para la clasificación de enfermedades foliares del cafeto, y cuáles presentan mayores limitaciones

metodológicas en condiciones reales de campo?

Además de estas limitaciones técnicas, se observa una preocupación conceptual relevante en la literatura sobre clasificación automática de enfermedades en hojas de café: (Rocha-Calderón et al., 2023a) presenta los resultados de la clasificación visual basada en imágenes como si constituyeran una confirmación fitosanitaria concluyente. En muchos casos, los modelos se limitan a identificar patrones visuales asociados a determinadas enfermedades, sin considerar que la confirmación fitosanitaria en el manejo agrícola requiere la integración de información adicional, como el estado fenológico del cultivo, las condiciones ambientales, la progresión de los síntomas a lo largo del tiempo y, en ciertos casos, evaluaciones técnicas complementarias realizadas por especialistas (Novtahaning et al., 2022).

Desde el punto de vista agronómico, la identificación visual de síntomas foliares representa únicamente una etapa preliminar del proceso de diagnóstico fitosanitario. La presencia de manchas, decoloraciones o necrosis en las hojas puede estar asociada a múltiples causas, incluyendo enfermedades fúngicas, estrés hídrico, deficiencias nutricionales o daños mecánicos, las cuales presentan similitudes morfológicas significativas (Rocha-Calderón et al., 2023b). Por ello, la clasificación automática basada exclusivamente en imágenes no puede considerarse, por sí sola, equivalente a un diagnóstico definitivo, sino más bien como una herramienta de apoyo al monitoreo y a la toma de decisiones iniciales (Arif et al., 2025).

Este uso impreciso de los términos no es trivial, ya que interpretar una predicción visual como diagnóstico concluyente puede generar decisiones agronómicas inadecuadas. Por ejemplo, una clasificación errónea puede inducir la aplicación innecesaria de fungicidas, incrementar los costos de producción, favorecer la aparición de resistencia en patógenos o, en el caso contrario, retrasar intervenciones necesarias cuando se producen falsos negativos. Estas decisiones afectan de manera directa la eficiencia del manejo integrado de enfermedades y la sostenibilidad económica de los pequeños productores, especialmente en sistemas productivos con recursos limitados (Montalbo, 2022a).

Asimismo, la falta de rigor conceptual en la presentación de los resultados puede generar expectativas poco realistas sobre las capacidades reales de los sistemas de inteligencia artificial, tanto en productores como en técnicos agrícolas. La confianza excesiva en modelos que no han sido validados en condiciones reales de campo ni contrastados con criterios agronómicos sólidos puede derivar en una dependencia tecnológica mal fundamentada, debilitando los procesos tradicionales de evaluación técnica y afectando la toma de decisiones a nivel del sistema productivo (Arif et al., 2025; Rocha-Calderón et al., 2023b).

En este contexto, resulta fundamental distinguir claramente entre la clasificación visual automática de síntomas foliares y el diagnóstico fitosanitario integral del cultivo al desarrollar y reportar modelos de inteligencia artificial aplicados a la caficultura. El uso impreciso de estos conceptos compromete la validez metodológica de los estudios y puede tener implicancias

directas en la toma de decisiones agronómicas (Martínez S. et al., 2022a). Por ello, surge la siguiente interrogante: ¿Qué riesgos metodológicos y de toma de decisiones surgen cuando los estudios presentan la clasificación visual automática de hojas de cafeto como un diagnóstico fitosanitario concluyente, sin considerar procesos de validación agronómica ni las limitaciones inherentes a los modelos de visión artificial?

La disponibilidad de datasets públicos ha sido un factor clave para la creación de modelos basados en inteligencia artificial para la identificación de enfermedades foliares del cafeto, ya que ha permitido entrenar y comparar arquitecturas de aprendizaje profundo de manera estandarizada. Bases de datos ampliamente utilizadas contienen imágenes etiquetadas de hojas sanas y enfermas, capturadas bajo condiciones controladas, lo que ha facilitado la evaluación inicial del desempeño de los modelos y la reproducción de resultados en distintos estudios (Aufar & Kaloka, 2022; Ayikpa, Ayikpa, et al., 2022).

Adicionalmente, la mayoría de estos datasets públicos están compuestos por imágenes de hojas aisladas, generalmente recortadas y fotografiadas sobre fondos homogéneos, con iluminación controlada y sin interferencias visuales. Esta configuración dista considerablemente de las condiciones reales de campo, donde las hojas se encuentran parcialmente ocluidas, presentan variaciones de iluminación natural, fondos complejos, distintos estados fenológicos y múltiples niveles de severidad de la enfermedad. En consecuencia, los modelos entrenados exclusivamente con este tipo de datos pueden no representar adecuadamente la complejidad visual presente en cafetales reales (Ayikpa, Ayikpa, et al., 2022).

Adicionalmente, la variabilidad inherente a los datasets externos —asociada a la región geográfica, al cultivar, al protocolo de captura de imágenes y a las condiciones agroecológicas— limita la capacidad de los modelos para generalizar a nuevos contextos productivos. En muchos casos, los datasets utilizados provienen de regiones distintas a aquellas donde se pretende aplicar el modelo, sin considerar factores locales como clima, sombra, altitud o prácticas de manejo, lo que afecta la validez de los resultados cuando los sistemas se despliegan en escenarios reales (Martínez S. et al., 2022b).

Estas limitaciones técnicas pueden traducirse en una disminución significativa del desempeño del modelo al enfrentarse a imágenes no vistas durante el entrenamiento, incrementando la tasa de errores y reduciendo la confiabilidad de las predicciones. La falta de datasets representativos y contextualizados no solo afecta la precisión de la clasificación, sino que también dificulta la evaluación objetiva del verdadero potencial de los modelos de IA para apoyar el monitoreo fitosanitario y la toma de decisiones agronómicas (Aufar & Kaloka, 2022; Martínez S. et al., 2022b).

En este contexto, resulta fundamental analizar críticamente el papel que desempeñan los datasets públicos y externos en el desarrollo de modelos de inteligencia artificial para la caficultura, así como las implicancias que su variabilidad tiene sobre la validez y aplicabilidad de los resultados.

Por ello, surge la siguiente interrogante: ¿De qué manera influye el uso de datasets públicos o externos (y su variabilidad como la iluminación, fondo, región, cultivar y protocolo de captura) en la precisión y validez de los modelos para roya y ojo de gallo, especialmente al aplicarlos en cafetales de San Martín?

Dada esta problemática, resulta necesario llevar a cabo una revisión sistemática y crítica de la literatura existente que aplica inteligencia artificial en la clasificación de enfermedades foliares del cafeto. Esta revisión permitirá identificar los tipos de datasets más utilizados, analizar su grado de representatividad, evaluar cómo influyen en el desempeño de los modelos propuestos y reconocer las principales limitaciones técnicas asociadas a su uso. Asimismo, se busca sentar una base sólida para el desarrollo de enfoques más robustos y contextualizados, capaces de responder a las condiciones reales de producción en cafetales de la región San Martín.

2.3.2 Revisión epistemológica

El cultivo del café y la sanidad vegetal

El cultivo del café constituye uno de los sistemas agrícolas más relevantes en regiones tropicales, no solo por su aporte económico, sino también por su impacto social y ambiental, especialmente en contextos de pequeña agricultura (Montalbo, 2022b). La sanidad vegetal representa un eje central en la sostenibilidad del cultivo, ya que la presencia de enfermedades afecta directamente el rendimiento, la calidad del grano y la estabilidad productiva de los cafetales. Desde una perspectiva agronómica, el mantenimiento de un adecuado estado sanitario permite optimizar el uso de insumos, reducir pérdidas y asegurar la viabilidad del sistema productivo a largo plazo.

La sanidad vegetal se concibe como un proceso integral que involucra la prevención, detección, evaluación y manejo de enfermedades, considerando la interacción entre el cultivo, el patógeno y el ambiente. En el caso del cafeto, esta interacción es particularmente compleja debido a la influencia de factores agroecológicos como la altitud, la humedad, la sombra y las prácticas de manejo, los cuales condicionan la aparición y severidad de las enfermedades foliares. En este contexto, la detección temprana y precisa de alteraciones en las hojas es un componente clave para el manejo integrado del cultivo.

Enfermedades foliares del cafeto

Las enfermedades foliares del cafeto representan una de las mayores amenazas para la productividad del cultivo, al afectar directamente el área fotosintética y provocar defoliación, debilitamiento de la planta y reducción del rendimiento. Entre las patologías más relevantes destacan la roya del café (*Hemileia vastatrix*) y el ojo de gallo (*Mycena citricolor*), debido a su amplia presencia en zonas tropicales húmedas y a su impacto recurrente en los sistemas productivos (Ayikpa, Ayikpa, et al., 2022).

La roya del café se manifiesta principalmente mediante la aparición de pústulas anaranjadas en el envés de las hojas, mientras que el ojo de gallo se caracteriza por lesiones circulares con centros necróticos. No obstante, la expresión visual de estas enfermedades puede variar

considerablemente según el estado fenológico de la planta, las condiciones ambientales y el nivel de severidad, lo que dificulta su identificación precisa mediante inspección visual. Además, otras afecciones foliares y factores de estrés pueden generar síntomas similares, incrementando la probabilidad de errores en el reconocimiento de la enfermedad (Novtahaning et al., 2022).

Inteligencia artificial aplicada a la agricultura

La inteligencia artificial ha surgido como una herramienta clave en la agricultura moderna, particularmente en el contexto de la agricultura de precisión y el monitoreo fitosanitario. Mediante la aplicación de métodos de aprendizaje automático y aprendizaje profundo, la inteligencia artificial posibilita el análisis de grandes volúmenes de información agrícola, la detección de patrones complejos y el respaldo a la toma de decisiones de forma objetiva y consistente (Aufar & Kaloka, 2022).

En aplicaciones agrícolas basadas en imágenes, la visión por computadora ha demostrado un alto potencial para la identificación y clasificación de enfermedades en plantas. Estos enfoques se apoyan en modelos capaces de aprender características visuales relevantes directamente a partir de los datos, reduciendo la dependencia de reglas manuales y aumentando la escalabilidad de las soluciones. En el caso del cafeto, la IA se ha propuesto como una alternativa para complementar el monitoreo tradicional, ofreciendo herramientas que faciliten la identificación temprana de enfermedades y el seguimiento del estado sanitario del cultivo.

Clasificación automática vs diagnóstico agronómico

Desde una perspectiva epistemológica, es fundamental distinguir entre la clasificación automática de imágenes y el diagnóstico agronómico integral. La clasificación automática se apoya en la identificación de patrones visuales aprendidos a partir de conjuntos de datos, permitiendo asignar una imagen a una categoría predefinida, como hoja sana o enferma. Sin embargo, este proceso no equivale a un diagnóstico fitosanitario completo, el cual implica la integración de múltiples fuentes de información, incluyendo observaciones de campo, condiciones ambientales, estado fenológico del cultivo y conocimiento experto (Ayikpa, Ayikpa, et al., 2022).

El diagnóstico agronómico constituye un proceso interpretativo más amplio, orientado a comprender las causas subyacentes de los síntomas observados y a definir estrategias de manejo adecuadas. En este sentido, los sistemas de inteligencia artificial deben entenderse como instrumentos de apoyo en los procesos de decisión, capaces de proporcionar información preliminar o indicios tempranos, pero no como sustitutos del criterio técnico ni de los procedimientos de validación agronómica. Reconocer esta diferencia es esencial para evitar interpretaciones erróneas sobre el alcance de los modelos propuestos y para promover un uso responsable de la inteligencia artificial en el manejo fitosanitario del cafeto (Montalbo, 2022b).

2.3.3 Metodología

Para llevar a cabo esta revisión, se realizó una búsqueda sistemática de literatura científica en la base de datos Scopus, elegida por su amplia cobertura multidisciplinaria y el rigor de sus criterios

de indexación académica. La búsqueda se realizó en abril de 2025 mediante una estrategia inicial amplia, utilizando la cadena: TITLE (coffee leaf) AND PUBYEAR > 2020 AND PUBYEAR < 2027 AND LIMIT-TO (DOCTYPE, "ar"), con el fin de identificar investigaciones recientes relacionadas con el estudio de hojas de café. Esta búsqueda arrojó un total de 313 artículos.

A partir del conjunto inicial de artículos recuperados, se aplicaron criterios de inclusión y exclusión previamente definidos con el fin de garantizar la pertinencia y relevancia de los estudios analizados en el estado del arte. Se incluyeron únicamente artículos originales que abordaran el uso de métodos basados en inteligencia artificial y visión por computadora aplicados a la identificación y categorización de enfermedades foliares del cafeto, empleando enfoques como redes neuronales convolucionales (CNN), transfer learning, modelos híbridos u otras arquitecturas de aprendizaje automático, así como aquellos que discutieran aspectos metodológicos relacionados con el uso de datasets, validación en condiciones reales o limitaciones técnicas. Por otro lado, se excluyeron estudios duplicados, trabajos de enfoque exclusivamente agronómico sin integración de métodos computacionales, investigaciones centradas en otros cultivos sin aplicabilidad directa al cafeto y artículos cuyo texto completo no estuvo disponible. La aplicación de estos criterios permitió conformar la muestra final de artículos analizados, alineada con las preguntas de investigación planteadas.

Para sistematizar la información extraída de los artículos seleccionados, se diseñó una plantilla de registro en Microsoft Excel, que permitió organizar los datos de manera estructurada y homogénea. En dicha plantilla se consideraron elementos clave como la identificación del artículo (autor, título, año de publicación y referencia), el problema abordado, los objetivos planteados, el tipo de enfoque metodológico empleado y las técnicas de inteligencia artificial utilizadas. Asimismo, se registraron las arquitecturas de los modelos propuestos, los tipos de datasets empleados (públicos, externos o propios), las condiciones bajo las cuales se adquirieron las imágenes, los indicadores de evaluación empleados, los aportes más relevantes identificados, los principales hallazgos y las líneas de investigación futura propuestas.

El análisis de la información se orientó a responder a las interrogantes de investigación planteadas en el estudio, relacionadas con las arquitecturas y enfoques metodológicos utilizados en la clasificación automática de enfermedades foliares del cafeto, los riesgos metodológicos y de toma de decisiones asociados a la interpretación de la clasificación visual como diagnóstico fitosanitario, y la influencia del uso de datasets públicos o externos en la precisión y validez de los modelos desarrollados. Para ello, los datos extraídos se organizaron en tablas temáticas, lo que permitió identificar patrones recurrentes, limitaciones técnicas comunes y avances metodológicos destacados. Estas tablas facilitaron la comparación entre distintas arquitecturas (como CNN, transfer learning y modelos híbridos), así como la evaluación de las propiedades y limitaciones de los datasets empleados en los estudios revisados.

Finalmente, se elaboró una síntesis descriptiva de los resultados obtenidos, integrando los

hallazgos más relevantes dentro del marco teórico y del análisis crítico desarrollado en el estado del arte. Este enfoque permitió destacar las principales contribuciones de la literatura existente, identificar brechas metodológicas y conceptuales, y establecer una base sólida para proponer futuras líneas de investigación orientadas a la creación de modelos de inteligencia artificial con mayor capacidad de generalización, contextualizados y aplicables a las condiciones reales de los cafetales de la región San Martín.

2.3.4 Resultados y discusión

Arquitecturas de inteligencia artificial usadas en hojas de café

El análisis de los 38 estudios consolidados en la matriz evidencia que la investigación en reconocimiento de enfermedades en hojas de café ha transitado desde enfoques clásicos (índices de color, segmentación básica y clasificadores tradicionales) hacia modelos profundos basados en CNN, y más recientemente hacia propuestas híbridas (fusión de características, combinación CNN + clasificadores tradicionales o CNN + atención/transformers). En general, la selección arquitectónica responde a tres factores recurrentes: (i) tipo de imagen (controlada vs. campo), (ii) tamaño/variabilidad del dataset, y (iii) objetivo final (clasificación, severidad, detección o despliegue en aplicaciones).

Los resultados del análisis evidencian que la mayoría de los estudios se orientan a tareas de clasificación, incluyendo variantes que incorporan estimación de severidad de la enfermedad o generación de recomendaciones. En segundo lugar, se identifican enfoques que integran etapas de segmentación o definición de regiones de interés (ROI), cuyo propósito principal es reducir el ruido de fondo y aislar la hoja o las zonas afectadas antes de la clasificación. En menor proporción, se reportan modelos basados en detección, enfocados en la localización explícita de lesiones o regiones relevantes dentro de la imagen.

Tabla 1.

Clasificación de arquitecturas utilizadas en el reconocimiento de enfermedades en hojas de café

Categoría	Modelos/Arquitecturas	Propósito
CNN (transfer learning / backbones clásicos)	DenseNet, ResNet, MobileNet/MobileNetV2/V3, EfficientNet, VGG, Inception, Xception	Clasificación multiclase o binaria; extracción jerárquica de características visuales.
Modelos híbridos	Ensamblados (p.ej., combinación de backbones), fusión de características (CNN + CNN), CNN + ML (DT/RF/SVM), CNN + handcrafted features	Mejorar generalización; balancear precisión vs. interpretabilidad/costo; robustez ante variación visual.
Transformers / atención (minoría, ViT / Swin (y fusiones con CNN) pero emergente)		Captura de relaciones globales y dependencia a larga distancia; útil cuando el fondo o el contexto confunden.
Segmentación / extracción de ROI	Esquemas con sustracción de fondo, segmentación previa, pipelines de “semi-segmentación”	Reducir sesgo por fondo; aislar hoja/lesión antes de extraer características.

Categoría	Modelos/Arquitecturas	Propósito
ML tradicional / modelos interpretables	SVM, Decision Tree, Random Forest, KNN, Naïve Bayes, lógica difusa	Clasificación basada en características (color/textura o deep features) con mayor interpretabilidad relativa.
Enfoques por índices de color / pixel-level	EG, GCC, reglas difusas (Mamdani)	Modelos más interpretables para escenarios de bajo cómputo, típicamente centrados en roya.

El análisis de la matriz de resultados evidencia que las redes neuronales convolucionales (CNN) constituyen el enfoque predominante en la literatura aplicada al análisis de hojas de café. No obstante, se observa un crecimiento sostenido de arquitecturas híbridas, lo cual responde de manera directa a dos limitaciones recurrentes identificadas en los estudios revisados: (i) la disminución del desempeño cuando los modelos entrenados en condiciones controladas son evaluados en escenarios de campo, y (ii) la necesidad de desarrollar modelos más ligeros y desplegables en entornos móviles o web, sin una pérdida significativa de precisión.

Distribución de enfoques y tipo de tarea reportada

En cuanto a la distribución de los enfoques según el tipo de tarea, los resultados muestran que la mayoría de los trabajos se concentra en clasificación, mientras que una proporción menor incorpora estimación de severidad (por ejemplo, grado o área afectada) o implementaciones funcionales orientadas a sistemas web o aplicaciones móviles. Esta distribución es relevante, ya que las métricas reportadas tienden a ser más elevadas en escenarios de clasificación con imágenes capturadas bajo condiciones controladas, mientras que los valores de desempeño se moderan cuando el problema incluye fondos complejos, variabilidad en la iluminación y adquisición de imágenes en condiciones reales de campo.

Tabla 2.

Tipo de tarea predominante en los estudios

Tipo de tarea (agrupada)	n (de 38)	Interpretación
Clasificación	23	Predomina el enfoque “predecir clase” (sana/enferma o multiclase).
Clasificación + implementación (app/web)	6	Priorizan arquitectura ligera y respuesta rápida (p.ej., MobileNet).
Clasificación + severidad	5	Incorporan reglas por área/porcentaje o etapas del daño.
Detección	3	Se orientan a localizar regiones/lesiones; no siempre reportan “accuracy”.
Segmentación	1	Menos frecuente, pero relevante para generalización en

campo.

El análisis de los estudios que incluyen implementaciones web o móviles muestra una tendencia consistente hacia la selección de modelos compactos, como MobileNet y sus variantes. Aunque estos modelos no siempre alcanzan el mayor valor de precisión (“top-1”) en evaluaciones comparativas, su adopción se justifica por las restricciones propias del despliegue en entornos reales, donde el tamaño del modelo, el consumo de memoria y el tiempo de inferencia penalizan arquitecturas más profundas.

Desempeño reportado y brecha entre condiciones controladas y campo

Al sintetizar los resultados reportados en la matriz de análisis, se identifica una brecha clara entre el desempeño obtenido en condiciones controladas y aquel observado en escenarios de campo. Las métricas más elevadas —frecuentemente cercanas al 99–100% de exactitud— se concentran en estudios que emplean datasets de gran tamaño, particiones estándar de entrenamiento/validación y condiciones de captura relativamente homogéneas. En contraste, cuando los modelos son evaluados con imágenes adquiridas en condiciones reales, caracterizadas por fondos complejos, variaciones de iluminación y capturas mediante dispositivos móviles en campo, el rendimiento tiende a disminuir o requiere la incorporación de estrategias adicionales para mantener niveles aceptables de precisión.

Un indicador relevante de esta brecha se observa en los estudios que utilizan datasets real-world, como RoCoLe, donde el desempeño se mantiene en rangos más moderados incluso al emplear arquitecturas avanzadas. Este comportamiento es consistente con los problemas de generalización reportados en la literatura, evidenciando que los modelos entrenados en entornos controlados no siempre transfieren su rendimiento de forma directa a escenarios operativos reales.

Tabla 3.
Resumen de exactitud (“accuracy” extraída del mejor resultado cuando fue comparable)
por categoría de arquitectura

Categoría	Estudios (n)	Mediana de accuracy (%)	Máximo reportado (%)	Lectura práctica
CNN (incluye TL)	19	99.49	100.00	Muy alto en escenarios favorables; sensible a fondo/iluminación si no se controla.
Híbrido	13	97.30	99.49	Mejora robustez; combina fortalezas (deep + interpretabilidad o reducción dimensional).
ML tradicional / clasificador	3	98.60	98.65	Competitivo cuando las características están bien diseñadas/seleccionadas.
Detección de objetos	3	73.30	97.80	No siempre comparable por métricas distintas (mAP/IoU); suele ser

Categoría	Estudios (n)	Mediana de accuracy (%)	Máximo reportado (%)	Lectura práctica más exigente.
-----------	--------------	-------------------------	----------------------	--------------------------------

Nota: en detección/segmentación muchas veces no corresponde “accuracy”, por eso se interpreta con cautela

En conjunto, los resultados sugieren que valores elevados de exactitud no deben interpretarse automáticamente como evidencia de aplicabilidad en campo. El desempeño tiende a sostenerse de manera más estable cuando los estudios incorporan al menos una de las siguientes estrategias metodológicas: (i) segmentación o sustracción de fondo para reducir ruido visual, (ii) aumento de datos con transformaciones realistas, (iii) fusión de características o esquemas de ensamble, y/o (iv) selección de arquitecturas considerando explícitamente criterios de despliegue, como tamaño del modelo y tiempo de inferencia.

Uso de datasets y su impacto (frecuencia y contexto)

La distribución de los datasets analizados evidencia una fuerte dependencia de **conjuntos de datos públicos ampliamente utilizados**, junto con una menor proporción de datasets recolectados en condiciones reales. Entre los más recurrentes se encuentra **BRACOL**, que aparece como uno de los datasets predominantes en la literatura revisada, debido a su disponibilidad y a la claridad de sus clases. De manera similar, **JMuBEN** es utilizado con frecuencia, especialmente en estudios orientados a escenarios multiclase.

Tabla 4. Datasets identificados con uso recurrente

Dataset (o tipo)	Veces observado
BRACOL	13
JMuBEN	8
RoCoLe (real-world)	4
Dataset propio/privado	2
Dataset Rwanda	1
Arabica	
PlantVillage	1

Por su parte, **RoCoLe** aparece con menor frecuencia, pero adquiere un rol relevante al representar condiciones reales de captura, con mayor complejidad de fondo y variabilidad ambiental. Asimismo, se identifican **datasets propios o privados**, asociados a contextos locales específicos —como recolecciones en Rwanda u otras regiones productoras—, cuya contribución resulta valiosa, aunque su comparabilidad entre estudios depende en gran medida de la estandarización de los protocolos de captura y anotación.

Los resultados muestran que los valores más altos de desempeño suelen concentrarse en datasets con **condiciones controladas** o patrones visuales más consistentes, mientras que los conjuntos de datos real-world ponen en evidencia la necesidad de **mayor robustez metodológica**,

favoreciendo enfoques basados en segmentación, fusión de información o arquitecturas híbridas.

Condiciones de captura y validez práctica

Al analizar las condiciones de captura reportadas en los estudios revisados, se observa que la mayoría de los trabajos se basa en imágenes provenientes de repositorios o escenarios controlados. En particular, se identifican 26 estudios que emplean imágenes capturadas en condiciones controladas, frente a 10 estudios que utilizan imágenes adquiridas directamente en campo y 2 estudios con esquemas mixtos.

Esta distribución explica en gran medida la elevada frecuencia de métricas de desempeño altas reportadas en la literatura, ya que la mayoría de los modelos es evaluada bajo condiciones con menor variabilidad que la observada en plantaciones reales. En consecuencia, los resultados refuerzan la necesidad de interpretar con cautela los valores de exactitud reportados y de promover evaluaciones más representativas del entorno operativo final, especialmente para aplicaciones orientadas al uso práctico en agricultura de precisión.

2.3.5 Conclusiones

En la presente revisión se identificaron tanto avances importantes como limitaciones que aún persisten en el uso de técnicas de inteligencia artificial para la detección y clasificación de enfermedades en hojas de café. Estos hallazgos se analizaron considerando aspectos metodológicos, técnicos y también su aplicabilidad en contextos agrícolas reales, lo cual resulta clave para su adopción práctica. A partir de la revisión de los estudios seleccionados, se pueden destacar tres ideas principales.

En primer lugar, se observa que las redes neuronales convolucionales (CNN) continúan siendo el enfoque más utilizado en la literatura, especialmente en tareas de clasificación de enfermedades foliares. Sin embargo, en los últimos años ha ido ganando terreno el uso de enfoques híbridos. Estos combinan las CNN con otras técnicas, como métodos de segmentación, extracción de características o incluso clasificadores tradicionales, buscando superar limitaciones relacionadas con la generalización de los modelos. En varios casos, estas propuestas logran un mejor equilibrio entre precisión y robustez, sobre todo cuando se enfrentan a imágenes más complejas. Aun así, su desempeño sigue dependiendo en gran medida de la calidad de los datos con los que son entrenados.

En segundo lugar, se evidencia una diferencia clara entre los resultados obtenidos en entornos controlados y aquellos logrados en condiciones reales de campo. Muchas investigaciones reportan niveles de precisión bastante altos, incluso superiores al 95%, pero estos suelen provenir de datasets públicos con imágenes bien definidas, buena iluminación y fondos simples. Por el contrario, cuando los modelos se evalúan con imágenes capturadas directamente en plantaciones —donde existen variaciones en iluminación, fondos y dispositivos—, el rendimiento tiende a ser más moderado. En estos casos, además, es frecuente que se requieran técnicas adicionales de preprocesamiento o estrategias más complejas para mantener la estabilidad del modelo. Esto

sugiere que una alta precisión en laboratorio no necesariamente garantiza un buen desempeño en escenarios reales.

Finalmente, el uso de datasets públicos ha sido fundamental para el desarrollo de este campo, destacando algunos conjuntos ampliamente utilizados como BRACOL y JMuBEN. No obstante, estos datasets suelen presentar limitaciones en cuanto a la diversidad de condiciones de captura, lo que puede afectar la capacidad de generalización de los modelos. Por otro lado, datasets que incluyen imágenes tomadas en campo, como RoCoLe o aquellos recolectados de forma local, aunque menos utilizados, resultan especialmente valiosos para evaluar el comportamiento real de las propuestas. En este sentido, se hace evidente la necesidad de desarrollar bases de datos más variadas, que incluyan diferentes condiciones de cultivo y que cuenten con criterios de anotación más estandarizados.

En conjunto, los resultados de esta revisión indican que el progreso en la aplicación de inteligencia artificial a la sanidad del cultivo de café depende no solo del desarrollo de arquitecturas más precisas, sino también de una evaluación más realista del desempeño, del uso responsable de métricas y de la consolidación de datasets que reflejen la complejidad del entorno agrícola. Estos elementos son clave para avanzar hacia soluciones verdaderamente transferibles y sostenibles en el marco de la agricultura de precisión.

CAPÍTULO 3. METODOLOGÍA

3.1 Variables

Independiente: Modelo de análisis automático basado en reconocimiento de patrones.

Dependiente: Enfermedades en hojas de cafeto.

3.2 Operacionalización de las variables

Variable	Dimensiones	Indicadores	Instrumento / Técnica	Escala de medición
Modelo de análisis automático basado en reconocimiento de patrones (Variable independiente)	Procesamiento de imágenes	- Número de imágenes de hojas de cafeto procesadas correctamente - Tasa de error en la lectura de imágenes - Número de características extraídas por cada modelo de transferencia de conocimiento	Registro automatizado del sistema / Log del modelo	Cuantitativa discreta
	Extracción de características	Estabilidad de los vectores de características	Reporte técnico del modelo / Scripts de extracción	Cuantitativa continua
	Implementación del modelo híbrido	- Número de predicciones generadas por el modelo - Tiempo promedio de procesamiento por imagen	Bitácora experimental / Cronómetro del sistema	Cuantitativa continua
Enfermedades en hojas de cafeto (Variable dependiente)	Precisión diagnóstica	- Porcentaje de clasificación correcta (Accuracy) - Sensibilidad y especificidad del modelo - AUC-ROC - Nivel de concordancia entre el modelo y el especialista (Kappa de Cohen)	Matriz de confusión, AUC-ROC	Cuantitativa continua
	Confiabilidad diagnóstica	Repetibilidad de los resultados	Análisis de concordancia estadística	Cuantitativa continua
	Eficiencia operativa	Tiempo promedio de diagnóstico por caso	Cronometrado en pruebas reales	Cuantitativa continua

3.3 Matriz de consistencia

Problemas	Objetivos	Hipótesis	Variables	Método de Investigación
¿Cómo un modelo de análisis automático mediante reconocimiento de patrones identifica y clasifica enfermedades en hojas de cafeto?	Desarrollar un modelo de análisis automático basado en reconocimiento de patrones para identificar y clasificar enfermedades en hojas de cafeto.	El desarrollo de un modelo de análisis automático mediante reconocimiento de patrones identifica y clasifica enfermedades en hojas de cafeto en el distrito de Saposoa.	Indep: Modelo híbrido de análisis automático basado en reconocimiento de patrones Dep: Enfermedades foliares en hojas de cafeto	Revisión bibliográfica de métodos de detección de enfermedades en café. Diseño experimental, de tipo aplicada y nivel explicativo. Desarrollo del modelo híbrido y evaluación mediante experimentación computacional.
¿Cómo estructurar un conjunto de imágenes de hojas de cafeto sanas y enfermas que permita entrenar adecuadamente un modelo automático de clasificación?	Estructurar un conjunto de imágenes digitales de hojas de cafeto sanas y enfermas para entrenar adecuadamente un modelo automático de clasificación.	La estructuración de un conjunto de imágenes digitales de hojas de cafeto, compuesto por clases sana y enfermas, balanceado y etiquetado por tipo de enfermedad, con control de calidad y aumento de datos, permite entrenar adecuadamente un modelo automático de clasificación para el distrito de Saposoa.		
¿Qué técnicas de transferencia de conocimiento y algoritmos supervisados permiten construir un modelo híbrido capaz de extraer características relevantes y clasificar enfermedades en las hojas de cafeto?	Determinar las técnicas de transferencia de conocimiento y algoritmos supervisados para construir un modelo híbrido capaz de extraer características relevantes y clasificar enfermedades en las hojas de cafeto.	La aplicación de técnicas de transferencia de conocimiento combinadas con algoritmos supervisados en un enfoque híbrido, permite construir un modelo capaz de extraer características relevantes y clasificar enfermedades en hojas de cafeto con mayor precisión en comparación con modelos convencionales.		
¿Cuál es el desempeño del modelo híbrido propuesto para la identificación y clasificación de enfermedades en hojas de cafeto?	Evaluar el desempeño del modelo híbrido propuesto para la identificación y clasificación de enfermedades en hojas de cafeto.	El modelo híbrido desarrollado mejora significativamente la precisión, confiabilidad y eficiencia en la identificación y clasificación de enfermedades en hojas de cafeto respecto a los enfoques tradicionales de diagnóstico visual.		

3.4 Tipo, nivel y diseño de la investigación

El presente enfoque de investigación se enmarca dentro de un paradigma cuantitativo de tipo aplicada, orientado a comprender las particularidades visuales de las imágenes digitales de hojas de cafeto afectadas por enfermedades foliares en el distrito de Saposoa, región San Martín, Perú. El estudio se desarrolla mediante la implementación de un modelo de análisis automático basado en reconocimiento de patrones, con el propósito de determinar y clasificar de manera precisa las enfermedades presentes en dichas hojas.

Este estudio se sitúa en un nivel explicativo, ya que no sólo busca describir las características visuales que diferencian una hoja sana de una hoja enferma, sino también explicar cómo un modelo híbrido de aprendizaje automático, apoyado en técnicas de transferencia de conocimiento, puede identificar patrones relevantes para la detección de enfermedades foliares en el cafeto (Shmueli, 2010); (Pan & Yang, 2010)).

El diseño de la investigación es experimental, sustentado en la comparación sistemática de algoritmos de clasificación y en la aplicación de pruebas estadísticas para evaluar diferencias significativas en su desempeño (Dietterich, 1998).

Además, el proyecto incorpora elementos de investigación correlacional y tecnológica:

- **Correlacional**, porque analiza la relación entre las clasificaciones realizadas por especialistas agrícolas y los resultados generados por el modelo híbrido de IA, evaluando la concordancia entre ambas metodologías y la consistencia del modelo frente al diagnóstico experto.
- **Tecnológico**, ya que implica el diseño, desarrollo e implementación de un modelo computacional basado en transferencia de conocimiento y clasificación supervisada, cuya finalidad no es únicamente validar principios teóricos sobre reconocimiento de patrones, sino proponer una solución innovadora con aplicabilidad real en el ámbito agrícola, especialmente en el monitoreo fitosanitario del cultivo de café.

3.5 Objeto de estudio, población y muestra

La población de estudio está conformada por las hojas de cafeto sanas y enfermas presentes en los cultivos en el distrito de Saposoa. Estas hojas pueden ser documentadas mediante imágenes digitales y cuentan con la evaluación inicial realizada por especialistas o técnicos en sanidad vegetal.

A partir de esta población, se trabajó con un conjunto de imágenes digitales de hojas de cafeto seleccionadas mediante un muestreo no probabilístico por conveniencia y estratificado según la condición fitosanitaria (sanas, afectadas por ojo de gallo y afectadas por roya), priorizando aquellas muestras con síntomas claramente identificables y diagnóstico confirmado por especialistas, de modo que se garantice su utilidad para el entrenamiento y evaluación del modelo automático.

Criterios de Inclusión:

- Hojas de cafeto pertenecientes al distrito de Saposoa.
- Imágenes capturadas siguiendo un protocolo uniforme de distancia, iluminación y enfoque.
- Hojas clasificadas previamente por especialistas como:
 - Sanas
 - Afectadas por ojo de gallo
 - Afectadas por roya
- Imágenes con calidad suficiente para análisis (alta resolución, enfoque correcto, sin sombras excesivas).

Criterios de Exclusión:

- a. Imágenes borrosas, incompletas, sobreexpuestas o subexpuestas.
- b. Hojas cubiertas parcial o totalmente por objetos externos (ramas, manos, tierra, etc.).
- c. Hojas cuya clasificación no haya sido confirmada por un especialista.
- d. Hojas en estados intermedios donde no se puede diferenciar claramente la enfermedad.

3.5.1 Tamaño de muestra

La muestra estuvo constituida por 1 500 imágenes digitales de hojas de cafeto, distribuidas en tres grupos: 500 imágenes de hojas sanas, 500 imágenes de hojas afectadas por ojo de gallo y 500 imágenes de hojas afectadas por roya. Las imágenes fueron seleccionadas de manera estratificada según la condición fitosanitaria de la hoja (sana, ojo de gallo, roya) y, dentro de cada estrato, se realizó una selección aleatoria de las muestras disponibles, con el fin de asegurar la diversidad y representatividad de las diferentes condiciones presentes en los cultivos del distrito de Saposoa.

3.5.2 Selección de muestra

Para la selección de la muestra se siguió un procedimiento sistemático y estratificado, con el fin de garantizar la diversidad y representatividad de la población de estudio. La muestra total estuvo constituida por 1 500 imágenes digitales de hojas de cafeto, distribuidas en tres estratos definidos según su condición fitosanitaria: 500 imágenes de hojas sanas, 500 imágenes de hojas afectadas por ojo de gallo y 500 imágenes de hojas afectadas por roya.

En una primera etapa, se realizó la recolección y registro fotográfico de hojas de cafeto directamente en los cultivos en el distrito de Saposoa, asegurando que se incluyan hojas en diferentes estados de desarrollo y con diversos niveles de afectación visible. Posteriormente, todas las hojas serán evaluadas y clasificadas por especialistas o técnicos en sanidad vegetal, quienes determinaron si la hoja está sana o presenta síntomas compatibles con ojo de gallo o roya del cafeto.

A partir de este universo de imágenes registradas, se procedió a la selección estratificada de la muestra, definiendo tres estratos:

- a. Estrato 1: hojas sanas.
- b. Estrato 2: hojas con ojo de gallo.

c. Estrato 3: hojas con roya.

Dentro de cada estrato, se seleccionaron las imágenes mediante un muestreo no probabilístico por conveniencia, priorizando aquellas que presenten síntomas claramente identificables en el caso de hojas enfermas y buen estado visual en el caso de hojas sanas (enfoque adecuado, contraste suficiente y ausencia de obstrucciones). De este modo, se garantiza que las imágenes seleccionadas sean útiles para el entrenamiento y evaluación del modelo de análisis automático. Finalmente, todas las imágenes incluidas en la muestra fueron validadas en cuanto a su clasificación y calidad por el equipo de especialistas, con el propósito de confirmar que cada etiqueta (sana, ojo de gallo, roya) sea correcta y que las condiciones técnicas de las imágenes sean compatibles con los requerimientos del sistema de visión artificial. Este procedimiento riguroso de selección de la muestra contribuyó a que los resultados obtenidos sean confiables, representativos y aplicables a la población objetivo de los cultivos en el distrito de Saposoa.

3.6 Técnicas de recolección de datos

Se definieron técnicas estructuradas que aseguraron la calidad, trazabilidad y control del conjunto de datos obtenidos.

3.6.1 Proceso de obtención de imágenes

El proceso de recolección de las imágenes se realizó directamente en las plantaciones de café de la cooperativa, iniciando con la identificación y selección de las plantas de cafeto en las diferentes parcelas de cultivo. Con el apoyo de un ingeniero agrónomo, se realizó la evaluación visual de las hojas para determinar su condición fitosanitaria, clasificando en hojas sanas, hojas con síntomas característicos de ojo de gallo y hojas con síntomas característicos de roya del cafeto. La participación del especialista permitió asegurar la veracidad del diagnóstico de cada hoja antes de su registro fotográfico.

Las visitas de campo se llevaron a cabo en un horario aproximado de 10:00 a. m. a 1:00 p. m., periodo en el cual se disponía de una iluminación natural suficiente y relativamente uniforme para la revisión visual. Una vez identificadas, las hojas seleccionadas eran cuidadosamente manipuladas para evitar daños, dobleces o sombras que pudieran comprometer la calidad visual. Para la captura de las imágenes se utilizó un teléfono móvil con cámara de alta resolución, siguiendo un protocolo estandarizado que consistió en colocar cada hoja sobre una superficie de papel blanco y mantener una distancia adecuada entre la cámara y la hoja, de modo que la hoja aparecerá completa dentro del encuadre y en foco. El enfoque se centró en la hoja, procurando que los bordes y las posibles lesiones se aprecian con la mayor nitidez posible.

Las imágenes obtenidas se registraron en formato JPG, conservando la resolución nativa del dispositivo y evitando procesos de compresión adicionales que pudieran degradar la calidad de la información visual necesaria para el posterior análisis mediante visión artificial.

3.6.2 Almacenamiento y organización de imágenes

Una vez capturadas, las imágenes fueron transferidas a un medio de almacenamiento digital y organizadas en una estructura sencilla de carpetas, de acuerdo con la condición fitosanitaria de las hojas. Para ello, se crearon tres directorios principales: Carpeta 1: hojas sanas, Carpeta 2: hojas con ojo de gallo y Carpeta 3: hojas con roya. Cada imagen fue almacenada en la carpeta correspondiente según la clasificación realizada en campo con el apoyo del ingeniero agrónomo.

3.6.3 Control de calidad

El control de calidad de las imágenes se llevó a cabo de manera continua durante y después del proceso de captura. En una primera instancia, durante las visitas a campo, se revisaba de inmediato cada fotografía tomada, verificando que la hoja se encontrará completamente visible, bien enfocada y con una iluminación adecuada. En los casos en los que se observaban problemas evidentes, como desenfoque, sombras intensas o encuadres incompletos, la imagen se repetía hasta obtener un resultado aceptable.

Posteriormente, ya en el entorno de almacenamiento digital, se realizó una segunda revisión visual de las imágenes organizadas en las tres carpetas. En esta etapa se descartaron aquellas fotografías que presentaban defectos técnicos (borrosas, sobreexpuestas, subexpuestas o con elementos que distrajeran la atención de la hoja) o dudas respecto a su condición fitosanitaria. Cuando surgían discrepancias o incertidumbres sobre la clasificación de una hoja, se procedía a contrastar nuevamente la información con el ingeniero agrónomo, con la finalidad de asegurar que las imágenes incluidas en el conjunto final mantuvieran una calidad visual adecuada y una clasificación confiable.

Este proceso de revisión sistemática permitió depurar el conjunto de imágenes, garantizando que los datos utilizados en la etapa de modelado fueran consistentes y pertinentes para los objetivos de la investigación.

3.7 Instrumentos de recolección de datos

Los instrumentos utilizados fueron tanto físicos como digitales, incluyendo:

- **Dispositivo móvil Samsung Galaxy A54 5G**, utilizado para la captura de imágenes en campo bajo un protocolo estandarizado de distancia, enfoque e iluminación
- **Imágenes digitales en formato JPG**, conservando la resolución nativa del dispositivo para mantener la información visual necesaria para el análisis.
- **Google Drive** como entorno de almacenamiento, procesamiento y organización de datos.
- **Scripts en Python** desarrollados en Google Colab para entrenamiento, evaluación y despliegue de modelos.
- **Bibliotecas especializadas** en visión computacional y aprendizaje profundo, tales como:
 - OpenCV y scikit-image para preprocesamiento y normalización de imágenes.
 - torch y torchvision para la implementación del extractor de características basado en redes neuronales convolucionales.

- scikit-learn para el entrenamiento de clasificadores tradicionales (SVM y Random Forest) y cálculo de métricas.
- XGBoost para el entrenamiento del clasificador basado en boosting.
- matplotlib, seaborn y pandas para análisis visual, estructuración y presentación de resultados.
- **Modelos entrenados:** ResNet18 + SVM, ResNet18 + Random Forest y ResNet18 + XGBoost, evaluados mediante validación cruzada estratificada.

Estos instrumentos fueron empleados para trabajar de forma integrada a lo largo del pipeline de procesamiento, desde la imagen capturada en campo hasta la inferencia automatizada de la condición fitosanitaria (sana, ojo de gallo o roya).

3.8 Tratamiento estadístico de los datos

El análisis estadístico incluyó técnicas descriptivas e inferenciales, orientadas a evaluar el desempeño de los modelos híbridos propuestos para la identificación y clasificación de enfermedades foliares del cafeto (sana, ojo de gallo y roya) en el distrito de Saposa.

3.8.1 Análisis descriptivo

Se calcularon métricas de desempeño estándar para cada modelo:

- **Accuracy (exactitud)**
- **Precision (precisión)**
- **Recall (sensibilidad)**
- **F1-score**
- **AUC (área bajo la curva ROC)**

Estas métricas fueron calculadas en cada pliegue de validación cruzada, permitiendo evaluar la estabilidad de cada modelo frente a la variabilidad del conjunto de datos.

3.8.2 Validación cruzada estratificada

Se aplicó validación cruzada de 10 pliegues estratificada, asegurando que en cada subconjunto se mantuvieran proporciones similares de las tres clases (sana, ojo de gallo y roya). Esta estrategia permitió estimar el rendimiento generalizado de los modelos bajo distintas particiones de entrenamiento y prueba, reduciendo el sesgo asociado a una única división del dataset.

3.8.3 Comparación de modelos mediante ANOVA

Se realizó un análisis de varianza (ANOVA) de un factor, incorporando el bloqueo por fold, para comparar el desempeño entre los tres modelos híbridos entrenados (ResNet18 + SVM, ResNet18 + Random Forest y ResNet18 + XGBoost). El análisis se ejecutó sobre los resultados obtenidos en una validación cruzada estratificada de 10 pliegues, generando 10 observaciones por modelo (una por cada fold), las cuales sirvieron como base para la prueba estadística.

La hipótesis nula (H_0) consideró que no existían diferencias significativas en el rendimiento de los modelos híbridos, mientras que la hipótesis alternativa (H_1) planteó que al menos uno de ellos presentaba un desempeño significativamente distinto. Tras obtener un resultado significativo, se

realizó una prueba post hoc de Tukey HSD con el objetivo de determinar las diferencias existentes entre los distintos pares de modelos.

El modelo ResNet18 + SVM resultó ser estadísticamente superior en consistencia y desempeño general, motivo por el cual fue seleccionado como el modelo final para la identificación y clasificación automática de hojas sanas, con ojo de gallo y con roya en el distrito de Saposoa.

CAPÍTULO 4. DESARROLLO DEL MODELO DE ANÁLISIS AUTOMÁTICO MEDIANTE RECONOCIMIENTO DE PATRONES PARA DETERMINAR ENFERMEDADES EN HOJAS DE CAFETO

Este capítulo se presenta el desarrollo integral del modelo híbrido de clasificación automática de enfermedades foliares del cafeto, el cual representa el núcleo fundamental de este estudio. Se describe la construcción del conjunto de datos estructurado, el diseño e implementación de los modelos de clasificación basados en visión artificial y aprendizaje automático, así como los procedimientos de evaluación orientados al diagnóstico fitosanitario automatizado de hojas de cafeto sanas, afectadas por ojo de gallo y afectadas por roya.

Con el propósito de organizar de manera lógica y sistemática las distintas fases del desarrollo del modelo, se adoptó la estructura metodológica del estándar CRISP-DM (Cross Industry Standard Process for Data Mining). Este enfoque permitió representar de forma clara el flujo completo del proceso, desde la comprensión del problema agronómico hasta la evaluación final del sistema automatizado de clasificación, asegurando coherencia metodológica y trazabilidad en cada etapa. A lo largo de este capítulo se desarrollan de manera secuencial las fases de comprensión del problema, comprensión y preparación de los datos, modelado, evaluación y selección del modelo óptimo, alineadas tanto a la lógica agronómica del diagnóstico fitosanitario convencional como a los principios computacionales de la minería de datos y la visión por computador.

El presente apartado describe el proceso orientado a la construcción de un dataset estructurado de imágenes digitales de hojas de cafeto, debidamente clasificadas por especialistas en sanidad vegetal según su condición fitosanitaria (sana, ojo de gallo y roya). El desarrollo se organizó conforme a las primeras etapas del modelo CRISP-DM, garantizando la adecuada preparación de los datos, la coherencia en la asignación de etiquetas y la compatibilidad del conjunto de imágenes con los requerimientos de los modelos híbridos de clasificación empleados en esta investigación.

4.1 Estructuración del conjunto de datos de imágenes digitales de hojas de cafeto

4.1.1 Comprensión del negocio

El diagnóstico fitosanitario de enfermedades foliares en el cultivo de café constituye una actividad crítica para la productividad y sostenibilidad del sector cafetalero en el distrito de Saposa; sin embargo, en la práctica, la evaluación visual en campo puede verse afectada por la variabilidad entre evaluadores, la similitud de síntomas en etapas tempranas, las condiciones cambiantes de iluminación y el tiempo limitado para inspeccionar grandes áreas de cultivo, lo que reduce la oportunidad y precisión del diagnóstico. Bajo esta premisa, la investigación se orientó al diseño, construcción y evaluación comparativa de modelos híbridos de clasificación basados en imágenes digitales de hojas de cafeto, con el propósito de determinar cuál configuración ofrece el mejor desempeño para distinguir entre hojas sanas, afectadas por ojo de gallo y afectadas por roya; para

ello, se desarrollaron y compararon las combinaciones ResNet18 + SVM, ResNet18 + Random Forest y ResNet18 + XGBoost, evaluando su rendimiento y estabilidad mediante validación cruzada. En este contexto, la creación de un dataset estructurado de imágenes etiquetadas por especialistas constituyó un paso clave, ya que aseguró la calidad de los datos y la base necesaria para entrenar y evaluar objetivamente los modelos propuestos.

4.1.2 Comprensión de los datos

El proceso de construcción del dataset siguió un flujo estructurado que abarcó tres etapas principales: pre recolección, recolección y almacenamiento de imágenes, tal como se muestra en la **Figura 1**.



Figura 1. Proceso de construcción del dataset

Pre recolección de imágenes

La fase de pre recolección comprendió la selección del cultivo de interés y la identificación de hojas representativas de distintas condiciones fitosanitarias en plantaciones ubicadas en el distrito de Saposoa, región San Martín, Perú, como refleja la **Figura 2**. Se consideraron tres categorías relevantes para el estudio: hojas sanas, hojas afectadas por ojo de gallo (*Mycena citricolor*) y hojas afectadas por roya del café (*Hemileia vastatrix*).

Cada muestra fue evaluada y confirmada por un especialista en sanidad vegetal (ingeniero agrónomo), quien estableció la etiqueta de referencia (ground truth) utilizada posteriormente para el entrenamiento y validación de los modelos híbridos de clasificación. Asimismo, en esta etapa se definieron los parámetros estandarizados de captura (distancia, iluminación, fondo y encuadre), con el fin de asegurar uniformidad en la adquisición de las imágenes.



Figura 2. Recolección de hojas

Recolección de imágenes

La digitalización de las muestras se realizó mediante la captura fotográfica de las hojas de cafeto previamente seleccionadas, con el objetivo de transformar las muestras físicas en imágenes digitales aptas para su procesamiento mediante técnicas de visión artificial y aprendizaje automático.

Para este proceso se utilizó un dispositivo móvil Samsung Galaxy A54 5G, equipado con una cámara de alta resolución. La toma de imágenes se llevó a cabo siguiendo un protocolo estandarizado, orientado a preservar los detalles visuales relevantes de la superficie foliar, tales como textura, coloración, bordes y presencia de lesiones asociadas a ojo de gallo y roya del cafeto. Durante la captura, cada hoja fue colocada sobre un fondo blanco con el fin de reducir interferencias visuales del entorno y facilitar la posterior segmentación. Se mantuvo una distancia aproximada de 30 a 40 cm entre la cámara y la hoja, asegurando que la muestra se encontrara completamente visible dentro del encuadre y correctamente enfocada. Asimismo, se procuró una iluminación natural uniforme, evitando sombras pronunciadas, reflejos o variaciones bruscas de luz que pudieran afectar la calidad de la imagen.

Almacenamiento

Las imágenes fueron registradas en formato JPG, conservando la resolución nativa del dispositivo y evitando procesos de compresión adicionales. Cada imagen digital fue asociada a información

básica que incluyó un identificador único, la condición fitosanitaria asignada (sana, ojo de gallo o roya) y la validación del especialista, conformando así un conjunto de datos estructurado el cual alimentó al pipeline de entrenamiento y evaluación de los modelos híbridos de clasificación, como se muestra en **Figura 3**.

Este proceso de digitalización resultó esencial para garantizar la consistencia visual y del material de entrada, condiciones necesarias para el correcto desempeño de los modelos basados en identificación de patrones utilizados en la presente investigación.



Figura 3. Ejemplares del dataset

4.1.1 Construcción de modelos híbridos de análisis automático basados en reconocimiento de patrones, que integren técnicas de transferencia de conocimiento para la extracción de características y algoritmos de clasificación supervisada para identificar enfermedades en imágenes de hojas del cafeto

4.2.1 Preparación de los datos

A fin de estructurar adecuadamente el conjunto de imágenes digitales de hojas de cafeto para su uso en modelos híbridos de reconocimiento de patrones y aprendizaje automático, se implementó un pipeline completo de preparación de datos, diseñado para garantizar consistencia visual, trazabilidad computacional y un balance representativo entre clases. Este flujo metodológico contempló el montaje del entorno de trabajo, la indexación del dataset, la aplicación de un aumento de datos controlado, el análisis del balance por clases antes y después del aumento, y un proceso de separación y validación que evita fugas de información entre los conjuntos de entrenamiento y prueba.

Montaje y carga de datos

La ejecución del proceso de preparación de datos se ejecutó íntegramente en la plataforma Google Colab, con integración directa a Google Drive para el almacenamiento estructurado del conjunto de imágenes. Las imágenes fueron organizadas previamente en carpetas según su condición fitosanitaria: hojas sanas, hojas con roya y hojas con ojo de gallo.

Se verificó la correcta carga de las imágenes desde cada directorio y se construyó un índice tabular que registra la ruta de cada archivo y su respectiva etiqueta de clase, asegurando la coherencia entre la estructura física del dataset y su representación computacional.

```

DATASET_DIR =
Path('/content/drive/MyDrive/Tesis_Hojas/tesishojas')
CLASSES = ['Sana_C', 'Roya_C', 'O_Gallo_C']

for c in CLASSES:
    imgs = list((DATASET_DIR / c).glob('*.jpg'))
    print(f"{c}: {len(imgs)} imágenes")

```

Aumento de datos controlado

La **Figura 4** presenta el esquema general del proceso de aumento aplicado al conjunto original por clase, ilustrando las transformaciones geométricas, cromáticas y de encuadre/textura empleadas, así como el resultado final de las versiones aumentadas generadas.

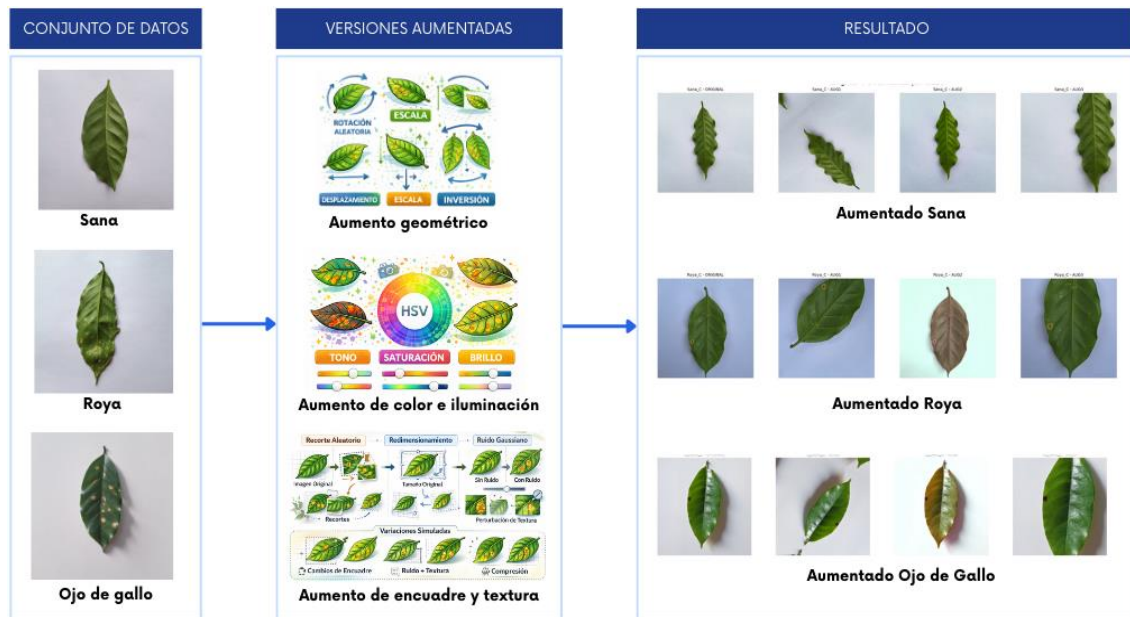


Figura 4. Esquema del proceso de data augmentation aplicado al conjunto de imágenes

Dado que el conjunto inicial estuvo compuesto por 500 imágenes por clase, se aplicó un aumento de datos (data augmentation) offline con el objetivo de incrementar la variabilidad visual y optimizar la capacidad de generalización de los modelos híbridos frente a condiciones reales de campo (cambios de orientación, iluminación y encuadre). Para cada imagen original se generaron exactamente 3 nuevas imágenes aumentadas, empleando tres “recetas” de transformación: (i) aumentos geométricos fuertes, (ii) aumentos de color/iluminación en el espacio HSV y (iii) aumentos por recorte con redimensionamiento y ruido gaussiano. La **Figura 5** presenta ejemplos representativos del proceso de aumento aplicado a cada clase, mostrando la imagen original y las tres transformaciones generadas.

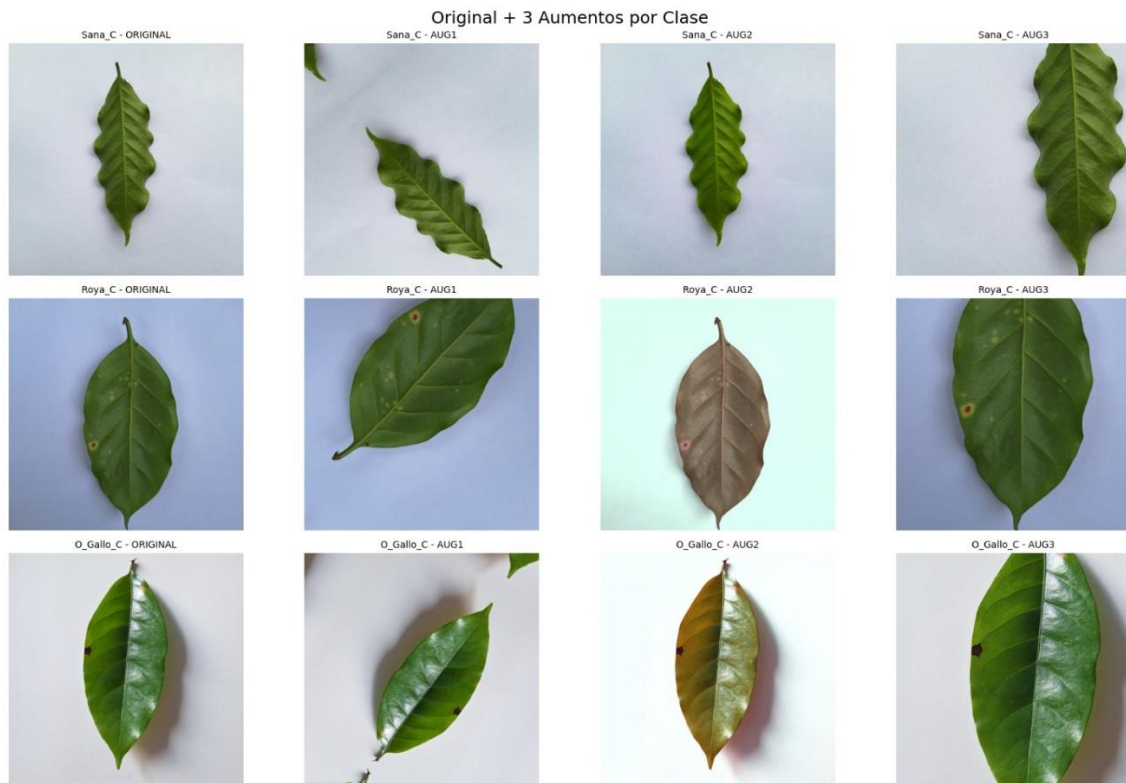


Figura 5. Ejemplares de aumento de datos aplicado por clase

Estas transformaciones fueron aplicadas a todas las clases y los archivos resultantes se guardaron con el sufijo _aug1, _aug2 y _aug3, manteniendo trazabilidad del origen y evitando duplicación no controlada.

NUM_AUMENTOS = 3

```
def aug_geom(img):
    """Aumento 1: geometría fuerte (rotación+escala+traslación)
    + flip aleatorio."""
    out = random_affine(out, angle_range=(-50, 50),
scale_range=(0.65, 1.35), translate_frac=0.25)
    if np.random.rand() < 0.8:
        out = random_flip(out)
    return out

def aug_color(img):
    """Aumento 2: color/iluminación fuerte en HSV."""
    return random_color_hsv_strong(img, h_shift_range=(-30, 30),
s_scale_range=(0.5, 1.6), v_scale_range=(0.5, 1.6))

def aug_crop_noise(img):
    """Aumento 3: recorte + redimensionado + ruido."""
    out = random_crop_resize(img, scale_range=(0.5, 0.8))
    out = random_gaussian_noise(out, sigma_range=(10, 30))
    return out

AUG_FUNCS = [aug_geom, aug_color, aug_crop_noise]

for k, func in enumerate(AUG_FUNCS, start=1):
```

```
img_aug = func(img)
nuevo_nombre = f"{nombre_base}_aug{k}{ext}"
cv2.imwrite(ruta_nueva, img_aug)
```

Balance final por clases

La **Figura 6** muestra la distribución del número de imágenes por clase antes y después del aumento de datos. Inicialmente, el dataset estuvo conformado por 1 500 imágenes, distribuidas equitativamente en 500 imágenes por clase. Tras la aplicación del aumento de datos, el conjunto final alcanzó un total de 6 000 imágenes, con 2 000 imágenes por clase, manteniendo un balance perfecto entre categorías.

Distribución por clases:

- Antes del aumento:
 - Sana: 500
 - Roya: 500
 - Ojo de gallo: 500
- Después del aumento:
 - Sana: 2 000
 - Roya: 2 000
 - Ojo de gallo: 2 000

Este incremento controlado permitió mejorar la representatividad del conjunto de entrenamiento y reducir el riesgo de sobreajuste, especialmente relevante en modelos basados en aprendizaje profundo y enfoques híbridos.

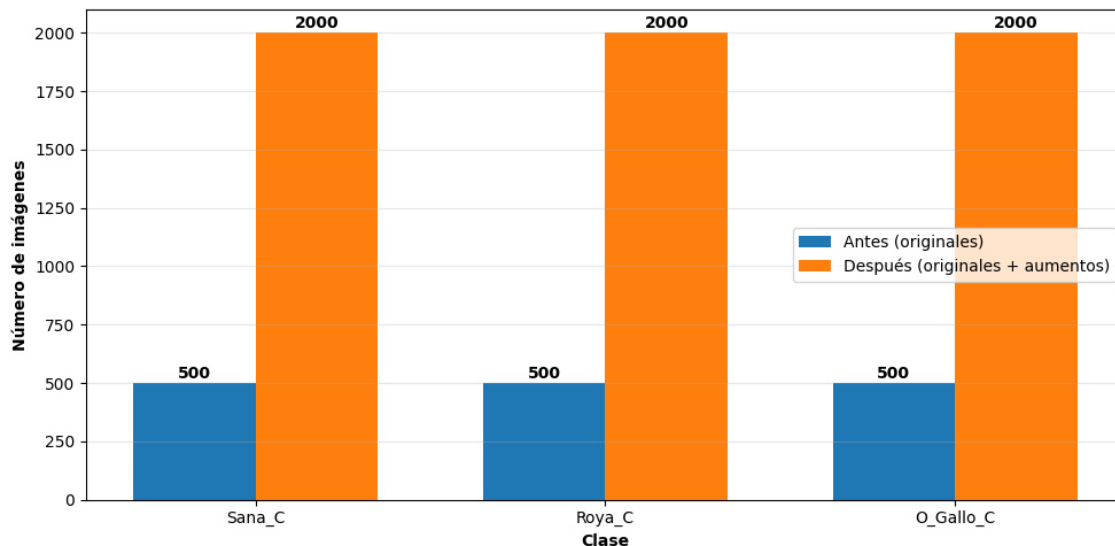


Figura 6. Distribución por clases (antes y después)

La figura evidencia que el proceso de aumento no alteró el balance entre clases, sino que amplió el volumen total de datos de forma homogénea, condición favorable para la comparación objetiva del desempeño de los modelos.

Separación de datos por grupos y validación cruzada

Con el objetivo de garantizar una evaluación rigurosa del desempeño de los modelos y evitar fugas de información (data leakage) entre los conjuntos de entrenamiento y prueba, se aplicó un esquema de partición de los datos fundamentado en agrupamiento por imagen original, considerando que cada hoja de cafeto generó múltiples versiones aumentadas durante el proceso de data augmentation.

En una primera etapa, el conjunto de datos final fue dividido mediante una partición hold-out 80/20, destinando el 80 % de las imágenes al conjunto de entrenamiento y el 20 % al conjunto de prueba. Esta separación se realizó utilizando un esquema por grupos, donde cada imagen original y todas sus imágenes aumentadas asociadas fueron consideradas como una misma unidad lógica. De esta forma, se aseguró que ninguna versión aumentada de una hoja aparezca simultáneamente en entrenamiento y prueba, preservando la independencia real entre ambos conjuntos.

```
from sklearn.model_selection import GroupShuffleSplit

def base_id_from_path(p):
    return Path(p).stem.split("_aug")[0]

df["group_id"] = df["path"].apply(base_id_from_path)

gss = GroupShuffleSplit(n_splits=1, test_size=0.20,
                        random_state=42)
train_idx, test_idx = next(gss.split(
    df, df["label_idx"], groups=df["group_id"]
))

df_train = df.iloc[train_idx]
df_test = df.iloc[test_idx]
```

Posteriormente, dentro del conjunto de entrenamiento, se aplicó una validación cruzada estratificada de k pliegues, manteniendo proporciones similares de las tres clases (hojas sanas, con ojo de gallo y con roya) en cada pliegue. Esta estrategia permitió evaluar la estabilidad y consistencia del rendimiento de los modelos híbridos bajo distintas particiones de los datos, reduciendo la dependencia de una única división del conjunto de entrenamiento.

Durante este proceso, el agrupamiento por imagen original se mantuvo, asegurando que las imágenes aumentadas derivadas de una misma hoja no se distribuyeran entre diferentes pliegues de validación. De este modo, se preservó la coherencia del proceso de evaluación y se evitó una sobreestimación del desempeño de los modelos.

```
from sklearn.model_selection import StratifiedKFold

skf = StratifiedKFold(n_splits=10, shuffle=True,
                      random_state=42)

for fold, (tr_idx, val_idx) in enumerate(
    skf.split(X_train, y_train), start=1):
    # entrenamiento y validación por fold
```

pass

Esta combinación de partición hold-out 80/20 por grupos y validación cruzada estratificada sobre el conjunto de entrenamiento permitió obtener una evaluación robusta del modelo, diferenciando claramente entre el rendimiento estimado durante la validación cruzada y el desempeño final medido sobre datos no vistos (conjunto de prueba).

Nota final de la preparación de los datos:

El conjunto final de 6 000 imágenes digitales de hojas de cafeto, balanceado entre clases y con un índice estructurado y trazable, constituyó la base oficial de datos utilizada en todas las etapas del desarrollo del modelo híbrido, incluyendo:

- Exploración inicial del dataset
- Entrenamiento de los modelos
- Validación cruzada
- Comparación estadística del desempeño
- Evaluación final en el conjunto de prueba

La integridad, balance y trazabilidad del conjunto de datos fueron controladas rigurosamente para garantizar la replicabilidad, robustez estadística y validez metodológica del estudio.

A partir del conjunto de imágenes digitales de hojas de cafeto descrito en la sección de comprensión y preparación de los datos, se implementaron modelos híbridos de clasificación orientados a la identificación automática de la condición fitosanitaria en el distrito de Saposoa. Estos modelos combinan una red neuronal convolucional preentrenada como extractor de características (feature extractor) con clasificadores supervisados tradicionales, con el propósito de aprovechar la capacidad de representación visual del aprendizaje profundo y, a la vez, la robustez de algoritmos clásicos para el aprendizaje sobre vectores de características.

En la presente investigación, el extractor de características se basó en ResNet18 preentrenada (transfer learning). La elección de esta arquitectura respondió a criterios de eficiencia computacional, debido a que ResNet18 es una versión más ligera en comparación con variantes más profundas (p. ej., ResNet 50/ResNet 101), lo que reduce el costo de cómputo y memoria durante inferencia. Esta característica resulta pertinente considerando escenarios de uso futuro en dispositivos móviles o entornos con recursos limitados, donde se requiere un modelo menos demandante sin sacrificar capacidad de representación. En consecuencia, su uso se alinea con una posible implementación práctica orientada a aplicaciones accesibles para productores y técnicos agrícolas en campo.

A partir de ResNet18 se obtuvieron embeddings representativos de patrones visuales relevantes (textura, coloración, bordes y lesiones). Posteriormente, dichos vectores fueron utilizados como entrada para tres clasificadores supervisados: SVM, Random Forest y XGBoost, conformando los siguientes modelos híbridos: ResNet18 + SVM, ResNet18 + Random Forest y ResNet18 + XGBoost.

Los modelos se construyeron sobre imágenes estandarizadas a 224×224 píxeles, correspondientes a tres clases: hojas sanas, hojas con ojo de gallo y hojas con roya. El pipeline fue desarrollado en Google Colab bajo entorno Python 3, con soporte GPU y gestión de archivos desde Google Drive. En los apartados siguientes se describe el proceso de extracción de características, entrenamiento y evaluación comparativa aplicado a cada modelo híbrido.

4.2.2 Modelo híbrido 1: ResNet18 + SVM

Justificación del modelo híbrido

Se diseñó un modelo híbrido que combina ResNet18 como extractor de características visuales con un clasificador SVM. ResNet18 fue seleccionada por ser una arquitectura ligera y eficiente en comparación con versiones más profundas (p. ej., ResNet50), lo cual reduce el consumo de memoria y cómputo durante la inferencia; esto resulta pertinente para una implementación futura en dispositivos móviles, facilitando su adopción en campo por productores y técnicos agrícolas. La SVM, por su parte, es un clasificador robusto para datos tabulares, especialmente eficaz cuando se trabaja con vectores de embeddings extraídos por una red convolucional, permitiendo una separación no lineal mediante kernel RBF.

Configuración inicial y carga del índice de datos

Las imágenes se organizaron en tres clases (Sana_C, Roya_C, O_Gallo_C) y se indexaron en un DataFrame con la ruta del archivo y su etiqueta. Se empleó LabelEncoder para codificar las clases y se verificó el total de imágenes por categoría (2 000 por clase; 6 000 en total).

```
rows = []
for label in CLASSES:
    class_dir = DATASET_DIR / label
    for p in class_dir.iterdir():
        if p.is_file() and p.suffix.lower() in IMG_EXTS:
            rows.append({'path': str(p), 'label': label})
```

```
df = pd.DataFrame(rows)
df["label_idx"] = le.fit_transform(df["label"])
```

Partición 80/20 por grupos (control de leakage por augmentation)

Dado que el dataset contiene imágenes aumentadas con sufijo _augk, se aplicó una separación hold-out 80/20 usando GroupShuffleSplit, agrupando por un identificador base (group_id) derivado del nombre del archivo. Con ello se aseguró que una imagen original y sus aumentos no se mezclen entre entrenamiento y prueba, evitando fugas de información.

```
def base_id_from_path(p: str) -> str:
    stem = Path(p).stem
    stem = re.sub(r"_aug\d+$", "", stem)
    return stem
```

```
df["group_id"] = df["path"].apply(base_id_from_path)
```

```
gss = GroupShuffleSplit(n_splits=1, test_size=0.20,
random_state=SEED)
train_idx, test_idx = next(gss.split(df, df["label_idx"],
groups=df["group_id"]))
df_train = df.iloc[train_idx].reset_index(drop=True)
df_test  = df.iloc[test_idx].reset_index(drop=True)
```

Fine-tuning ligero del extractor ResNet18 (solo cabeza + early stopping)

Antes de extraer embeddings, se realizó un fine-tuning ligero de ResNet18, entrenando únicamente la capa final (head) y congelando el resto de parámetros. Se aplicó early stopping para conservar el mejor rendimiento interno. En tu ejecución, se alcanzó un mejor val_acc interna de 0.8427 durante el ajuste preliminar del backbone.

```
model =
models.resnet18(weights=models.ResNet18_Weights.IMAGENET1K_V1)
freeze_all_except_head(model)
# entrenamiento con early stopping...
```

Extracción de embeddings (features) con ResNet18

Luego del fine-tuning, se reemplazó la capa final por Identity() para usar ResNet18 como extractor. Cada imagen se redimensionó a 224×224 y se normalizó con parámetros ImageNet. El resultado fue un vector de 512 características por imagen. En tu experimento se extrajeron:

- X_train: (4800, 512) para el 80%
- X_test: (1200, 512) para el 20%

```
model.fc = nn.Identity() # extractor
X_train, y_train, _ = extract_features(model, df_train, test_tf,
DEVICE, batch_size=64)
X_test, y_test, _ = extract_features(model, df_test, test_tf,
DEVICE, batch_size=64)
```

Configuración del clasificador SVM + PCA y optimización (HPO)

Sobre los embeddings se implementó un pipeline con StandardScaler + PCA + SVM (kernel RBF). Para ajustar hiperparámetros se aplicó RandomizedSearchCV con 30 candidatos y un split estratificado equivalente a validación del 10% (1/(folds=10)). El mejor modelo encontrado incluyó: pca__n_components=512, kernel=rbf, gamma=scale, C=1.5849.

```
pipe_svm = Pipeline([
    ('scaler', StandardScaler()),
    ('pca', PCA(random_state=SEED)),
```

```

('clf', SVC(probability=True, kernel='rbf'))
])
search = RandomizedSearchCV(pipe_svm, param_dist_svm, n_iter=30,
cv=cv, scoring='accuracy')

```

Resultados de clasificación (TRAIN y TEST)

Finalmente, el valor de macro promedio en F1-score (0.926) y AUC (0.986) confirma que el modelo mantiene un rendimiento equilibrado entre clases, sin favorecer desproporcionadamente a las clases mayoritarias. Esto evidencia la efectividad del enfoque híbrido propuesto, especialmente en un contexto clínico donde el balance entre sensibilidad y especificidad es crucial para la toma de decisiones diagnósticas.

Las métricas evidencian un rendimiento alto y consistente. En TRAIN (80%) se obtuvo:

- Accuracy = 0.9988, F1-macro = 0.9988, AUC-macro = 0.9997.

En TEST (20% nunca visto) se obtuvo:

- Accuracy = 0.9742, F1-macro = 0.9730, AUC-macro = 0.9968.

El classification report en TEST confirma desempeño equilibrado por clase, destacando O_Gallo_C con precisión/recall ≈ 0.99 , y un rendimiento ligeramente menor en Roya_C (recall ≈ 0.95), coherente con los errores observados en la matriz de confusión.

Visualización de resultados

El análisis visual de los resultados obtenidos por el modelo híbrido ResNet18 + SVM permite complementar las métricas cuantitativas y facilita la interpretación del desempeño clasificatorio en las tres clases fitosanitarias consideradas: hojas sanas (Sana_C), hojas afectadas por roya (Roya_C) y hojas afectadas por ojo de gallo (O_Gallo_C). A través de matrices de confusión, curvas ROC y el análisis de la evolución por pliegues, se evalúa la capacidad discriminativa, estabilidad y comportamiento general del modelo.

• Matriz de confusión multiclase

La **Figura 7** muestra la matriz de confusión obtenida para el modelo híbrido ResNet18 + SVM en el conjunto de entrenamiento (80%). Se observa una diagonal claramente dominante, lo que pone de manifiesto una elevada concordancia entre las etiquetas verdaderas y las salidas generadas por el modelo. En términos cuantitativos, la clase O_Gallo_C fue identificada sin errores, mientras que la clase Roya_C presentó únicamente 5 instancias confundidas con Sana_C, y la clase Sana_C registró 1 confusión hacia Roya_C, confirmando un comportamiento altamente preciso en entrenamiento.

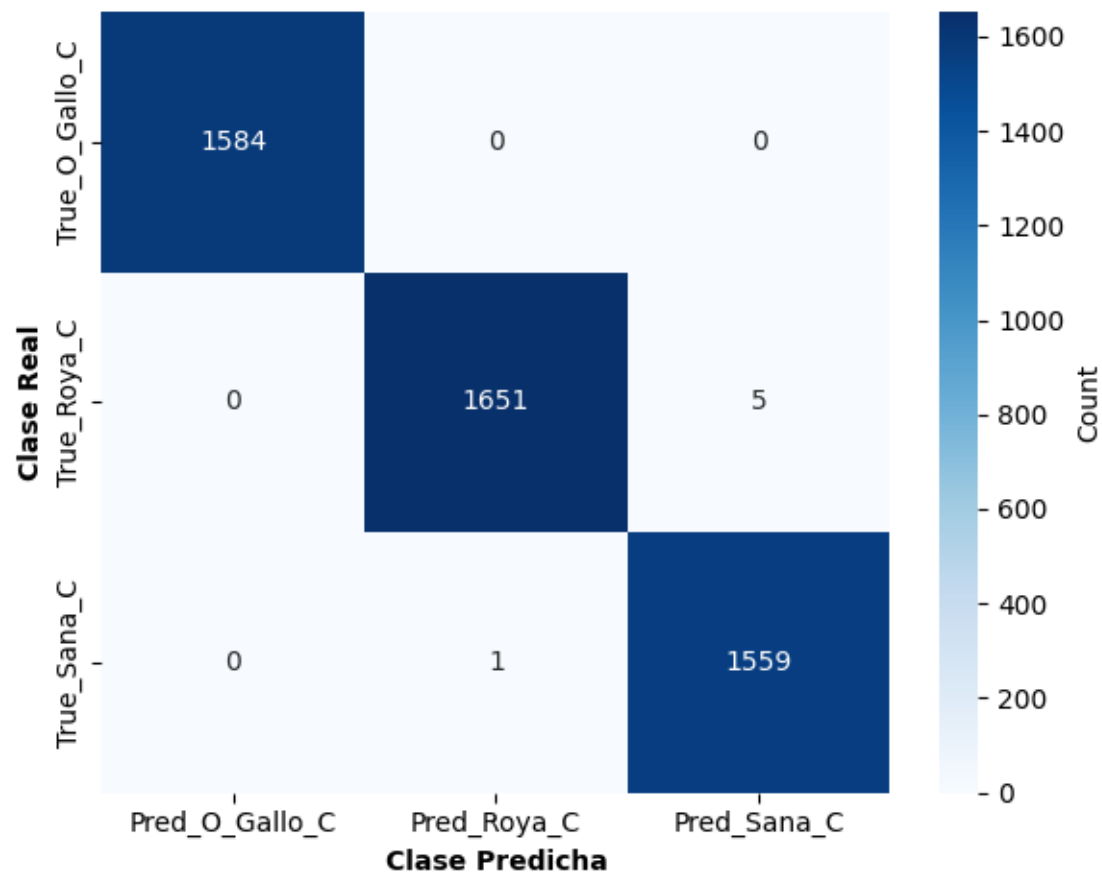


Figura 7. Matriz de confusión del modelo híbrido ResNet18 + SVM en el conjunto de entrenamiento (80%)

La **Figura 8** presenta la matriz de confusión correspondiente al conjunto de prueba (20%), compuesto por imágenes nunca vistas durante el entrenamiento. Se clasificaron correctamente 412 casos de O_Gallo_C, 326 de Roya_C y 431 de Sana_C. Las principales confusiones se concentraron entre $Roya_C \rightarrow Sana_C$ (15 casos) y $Sana_C \rightarrow Roya_C$ (9 casos), además de un número reducido de casos $O_Gallo_C \rightarrow Roya_C$ (4 casos). Este patrón resulta esperable debido a la variabilidad de iluminación y al grado de severidad de síntomas, donde ciertas manifestaciones tempranas pueden presentar características visuales similares.

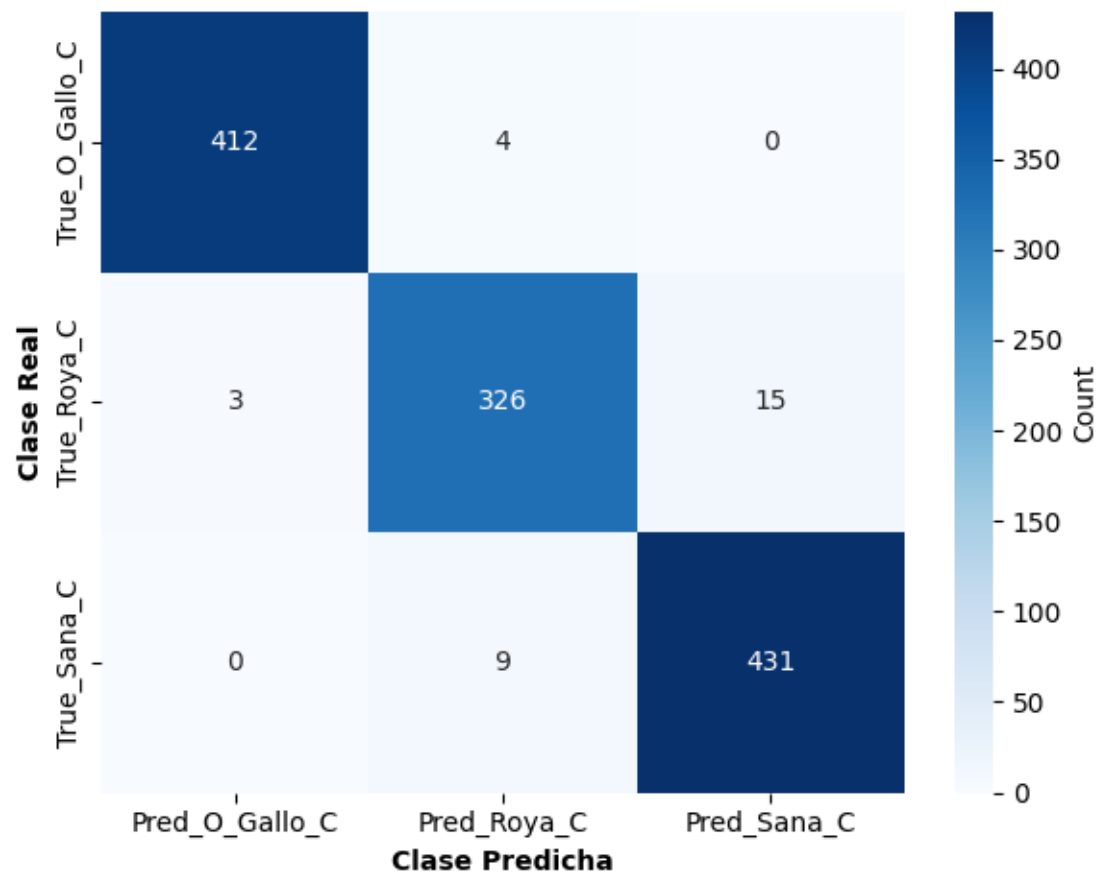


Figura 8. Matriz de confusión del modelo híbrido ResNet18 + SVM en el conjunto de prueba (20%)

- **Curvas ROC y AUC por clase**

Las curvas ROC (Receiver Operating Characteristic) se emplean para evaluar la habilidad del modelo en distinguir correctamente cada una de las clases. En la **Figura 9** se presentan las curvas ROC correspondientes al conjunto de entrenamiento, evidenciando un comportamiento casi ideal: el modelo alcanzó Macro-AUC = 1.000, con valores por clase próximos a la unidad, lo que indica una separación casi perfecta entre las categorías.

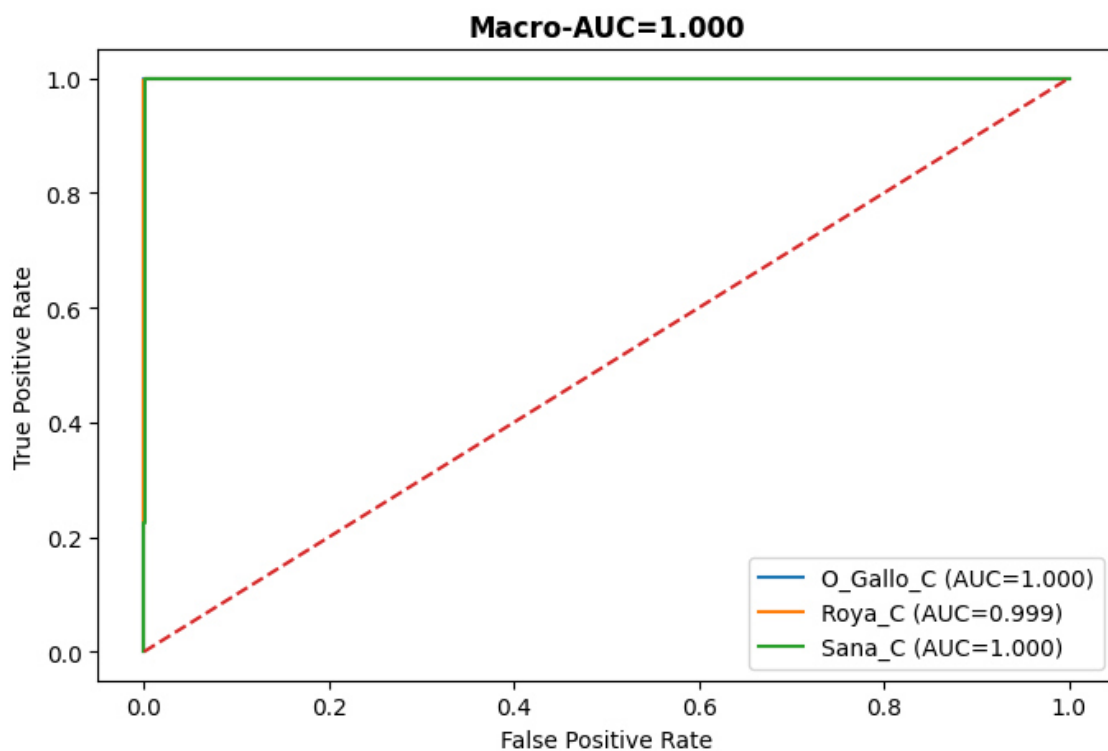


Figura 9. Curvas ROC OVR del modelo híbrido ResNet18 + SVM en el conjunto de entrenamiento (Macro-AUC=1.000)

En la partición de prueba, la **Figura 10** muestra que el modelo conserva una excelente capacidad de generalización, alcanzando Macro-AUC = 0.997. De forma individual, la clase O_Gallo_C presentó AUC = 1.000, mientras que Sana_C y Roya_C mantuvieron valores muy altos (≈ 0.997 y ≈ 0.994 , respectivamente), confirmando una discriminación sólida aun en datos no vistos.

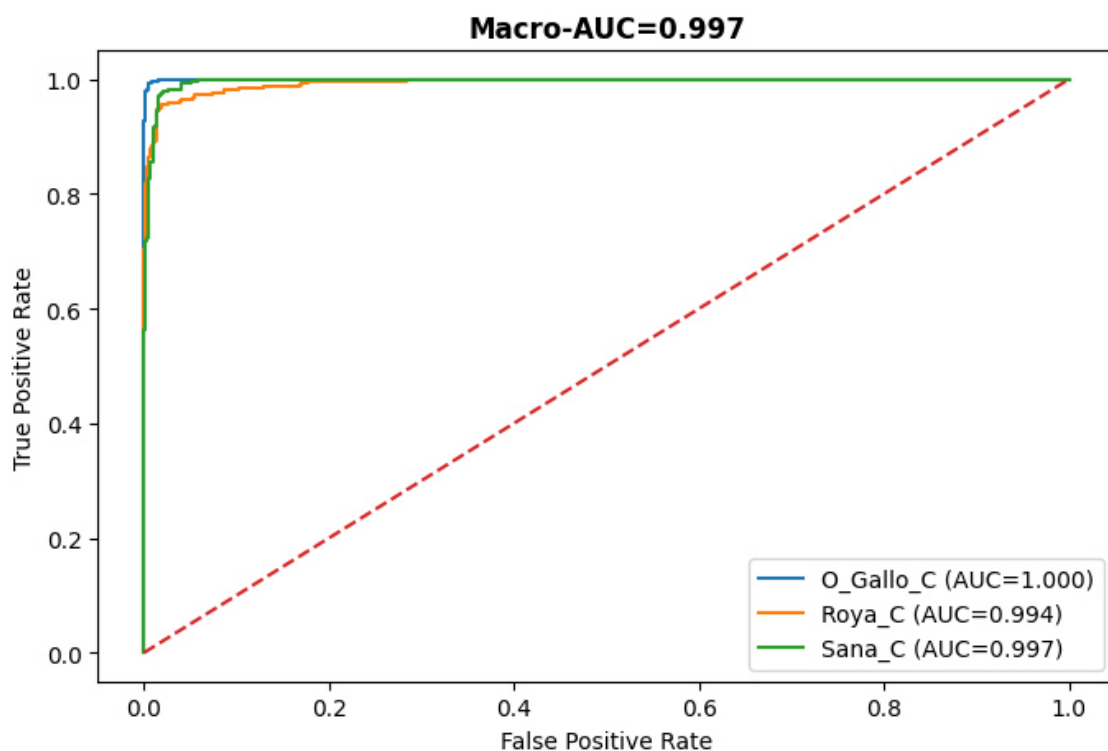


Figura 10. Curvas ROC OVR del modelo híbrido ResNet18 + SVM en el conjunto de prueba (Macro-AUC=0.997)

• Estabilidad del modelo y evolución por pliegues

La consistencia del modelo se examinó mediante un procedimiento de validación cruzada estratificada con 10 folds sobre los datos de entrenamiento. La **Figura 11** presenta la evolución de *accuracy* y *F1-macro* por fold, mostrando un comportamiento estable, sin variaciones abruptas, con valores que oscilaron aproximadamente entre 0.956 y 0.988. Esta consistencia respalda la confiabilidad del modelo y su independencia respecto a una partición específica de los datos.

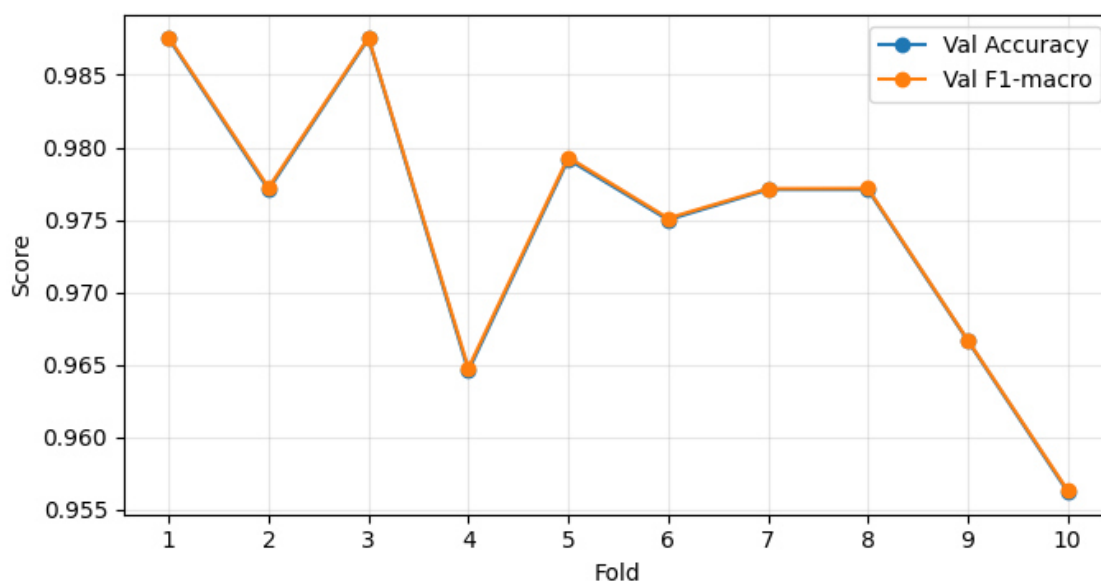


Figura 11. Evolución de Accuracy y F1-macro por fold (10-fold CV) en el conjunto de entrenamiento

• Interpretación agronómica y operativa

Desde una perspectiva agronómica, los resultados obtenidos consolidan el potencial del modelo híbrido ResNet18 + SVM como herramienta de apoyo al diagnóstico fitosanitario en el cultivo de café. La alta capacidad de identificación de ojo de gallo y roya, enfermedades de impacto económico, permite una detección oportuna en campo. Asimismo, el buen desempeño en hojas sanas reduce falsas alarmas y optimiza decisiones de intervención.

Finalmente, el uso de ResNet18 como extractor por su eficiencia y menor demanda computacional refuerza la viabilidad operativa del enfoque para una futura integración en aplicaciones móviles o sistemas de agricultura de precisión, orientados a productores y técnicos en contextos con recursos limitados.

Conclusión

El modelo híbrido basado en ResNet18 como extractor de características y SVM como clasificador demostró un rendimiento robusto, estable y generalizable en la clasificación automática de hojas de cafeto en tres condiciones fitosanitarias (Sana_C, Roya_C y O_Gallo_C),

evidenciando una alta capacidad discriminativa tanto en entrenamiento como en prueba, además de consistencia a lo largo de la validación cruzada. El enfoque híbrido permitió aprovechar la transferencia de conocimiento de ResNet18 para capturar patrones visuales relevantes (textura, coloración y lesiones) y la eficacia de la SVM para separar los embeddings con precisión, logrando un desempeño elevado bajo variabilidad propia del entorno agrícola. Asimismo, su diseño modular facilita la reutilización de representaciones y reduce la complejidad computacional frente a modelos monolíticos end-to-end, lo que respalda su viabilidad técnica y operativa como alternativa eficiente para el diagnóstico fitosanitario automatizado y su proyección a una futura integración en herramientas móviles o sistemas de apoyo a la decisión en campo.

4.2.3 Modelo híbrido 2: ResNet18 + Random Forest

Justificación del modelo híbrido

Para este estudio también se planteó un segundo modelo híbrido, en el que se utiliza ResNet18 como extractor de características visuales y Random Forest como clasificador. La elección de ResNet18 se debe a que es una arquitectura relativamente ligera dentro de la familia ResNet, ya que tiene menos capas y parámetros que otras versiones más profundas. Esto la hace adecuada para escenarios donde no se dispone de gran capacidad de cómputo, como aplicaciones móviles o herramientas que se usan directamente en campo. En contextos agrícolas, este aspecto es especialmente importante, ya que se busca un balance entre eficiencia y buen desempeño en la identificación de enfermedades.

En cuanto al clasificador, se optó por Random Forest debido a su buen comportamiento en distintos tipos de problemas. Este algoritmo, basado en múltiples árboles de decisión, suele ser bastante estable frente al sobreajuste y tiene la capacidad de capturar relaciones complejas entre las variables. Además, funciona bien incluso cuando se trabaja con muchas características, lo cual es útil considerando que las representaciones generadas por la red neuronal pueden ser de alta dimensionalidad. La combinación de embeddings visuales extraídos por ResNet18 con RandomForest permite evaluar un enfoque alternativo al SVM, orientado a maximizar estabilidad y generalización del modelo híbrido.

Configuración inicial y carga del índice de datos

El conjunto de datos fue indexado automáticamente a partir de la estructura de carpetas correspondiente a las tres clases diagnósticas: Sana_C, Roya_C y O_Gallo_C, totalizando 6,000 imágenes balanceadas tras el proceso de aumento de datos. Cada imagen fue registrada en un DataFrame que incluye la ruta absoluta del archivo, la etiqueta de clase y un identificador de grupo (group_id) derivado del nombre base de la imagen original, lo cual permite agrupar correctamente las imágenes aumentadas y evitar fugas de información.

La separación del dataset se realizó mediante un esquema 80 % entrenamiento / 20 % prueba, empleando GroupShuffleSplit, garantizando que todas las imágenes aumentadas derivadas de una misma imagen original permanezcan exclusivamente en uno de los subconjuntos. Este procedimiento asegura una evaluación honesta del desempeño del modelo sobre datos nunca vistos.

Fine-tuning ligero del extractor ResNet18

Se utilizó una instancia de ResNet18 preentrenada sobre ImageNet, sobre la cual se realizó un fine-tuning ligero, entrenando únicamente la capa final (fully connected) y dejando inalterados los restantes pesos de la arquitectura. Este enfoque permite adaptar el extractor a las características específicas de las hojas de cacao sin incurrir en altos costos computacionales ni riesgo de sobreajuste.

El fine-tuning se ejecutó durante un máximo de 4 épocas, incorporando early stopping basado en la precisión de validación interna. El mejor modelo alcanzó una $val_accuracy = 0.8427$, valor que fue utilizado como referencia para la extracción definitiva de embeddings.

Extracción de embeddings visuales

Una vez ajustada la red, ResNet18 fue configurada como extractor puro de características, reemplazando la capa de clasificación por una capa identidad. Cada imagen fue procesada para generar un vector de características de 512 dimensiones, el cual representa una descripción compacta y discriminativa de la información visual contenida en la hoja.

Se extrajeron embeddings para:

- 4,800 imágenes del conjunto de entrenamiento (80 %)
- 1,200 imágenes del conjunto de prueba (20 %)

Estos vectores fueron almacenados en archivos comprimidos .npz para su posterior uso por el clasificador RandomForest.

Configuración del clasificador RandomForest

El clasificador RandomForestClassifier fue implementado dentro de un pipeline que incluye opcionalmente una etapa de reducción de dimensionalidad mediante PCA. La optimización de hiperparámetros se realizó utilizando RandomizedSearchCV, explorando 30 combinaciones posibles con un split estratificado interno del 10 % del conjunto de entrenamiento.

Los hiperparámetros óptimos seleccionados incluyen:

- $n_estimators = 400$
- $max_depth = 30$
- $min_samples_split = 5$
- $min_samples_leaf = 2$
- $max_features = sqrt$
- $bootstrap = False$

- `pca__n_components = 512`

Este proceso permitió maximizar el desempeño del clasificador manteniendo estabilidad y control del sobreajuste.

Validación cruzada y evaluación del modelo

El modelo optimizado fue evaluado bajo tres escenarios complementarios:

- Rendimiento en entrenamiento (80 %)
 - Accuracy = 1.000
 - F1-macro = 1.000
 - ROC-AUC = 1.000
- Evaluación en conjunto de prueba (20 % nunca visto)
 - Accuracy = 0.9367
 - F1-macro = 0.9341
 - ROC-AUC = 0.9871

Validación cruzada estratificada de 10 folds sobre el conjunto de entrenamiento, obteniendo valores de F1-macro entre 0.9169 y 0.9565, lo que evidencia un desempeño consistente y estable del modelo.

Resultados de clasificación

El desempeño del modelo híbrido ResNet18 + RandomForest demuestra una capacidad sólida para discriminar entre las tres clases evaluadas. En el conjunto de prueba, la clase O_Gallo_C presentó los mejores resultados (F1 = 0.97), seguida por Sana_C (F1 = 0.94). La clase Roya_C, aunque ligeramente inferior (F1 = 0.89), mantiene un desempeño competitivo considerando la similitud visual entre lesiones.

El valor de F1-macro = 0.9341 confirma que el modelo mantiene un equilibrio adecuado entre clases, sin favorecer desproporcionadamente a ninguna de ellas.

Visualización de resultados

La exploración visual de los resultados permite realizar el análisis cuantitativo y facilitar la interpretación del comportamiento del modelo híbrido ResNet18 + RandomForest en cada una de las clases diagnósticas.

Matrices de Confusión

La **Figura 12** y **Figura 13** **Error! Reference source not found.** presentan las matrices de confusión correspondientes a los conjuntos de entrenamiento y prueba, respectivamente. En ambas se observa una diagonal dominante, lo que refleja un alto grado de coincidencia entre las clases verdaderas y las predicciones obtenidas.

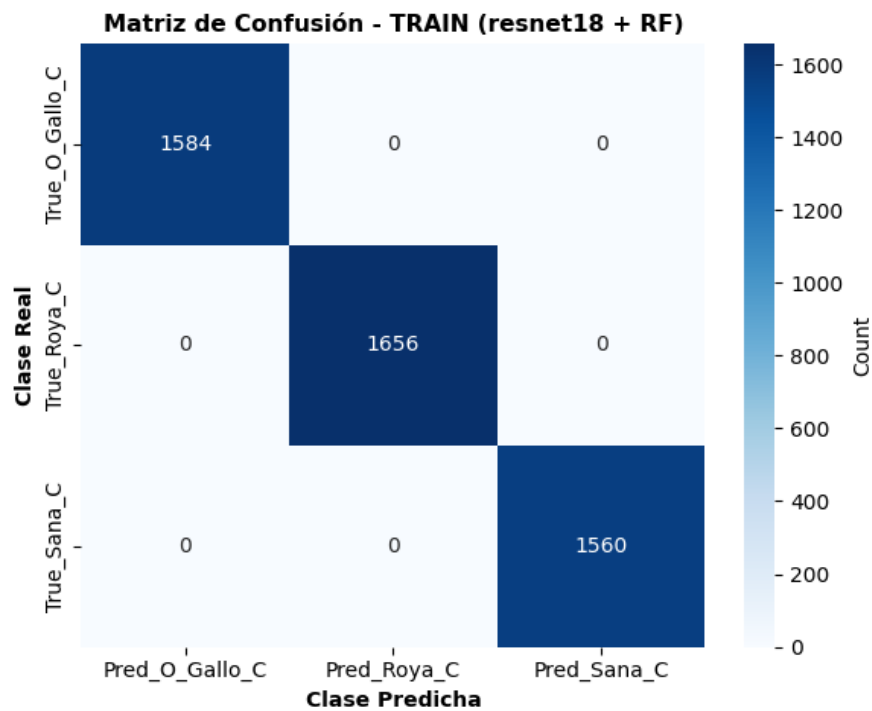


Figura 12. Matriz de confusión del modelo ResNet18 + RandomForest en el conjunto de entrenamiento

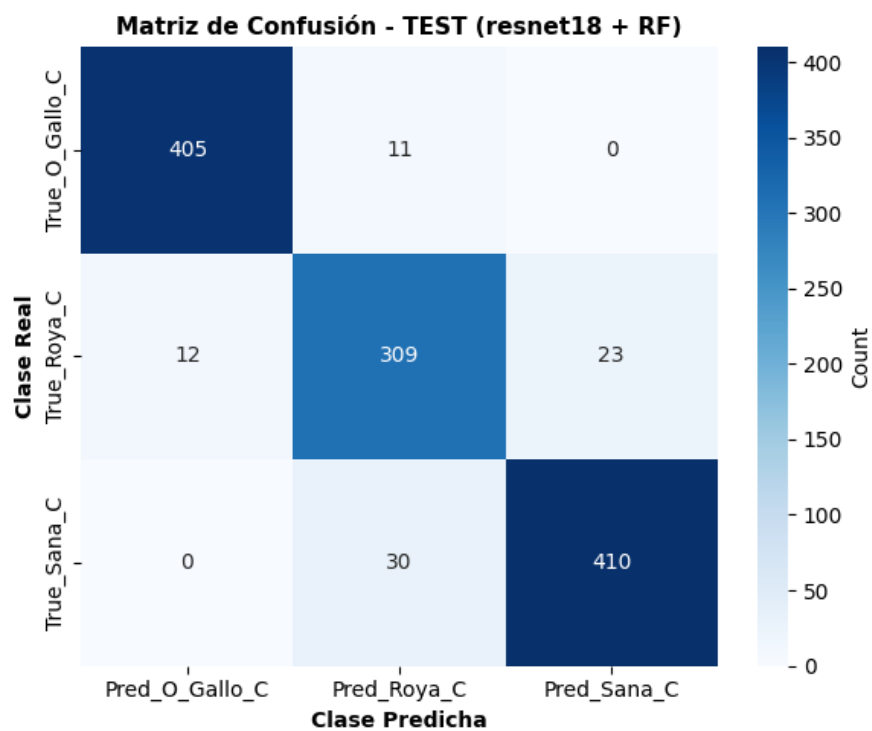


Figura 13. Matriz de confusión del modelo ResNet18 + RandomForest en el conjunto de prueba

Los principales errores de clasificación se concentran entre Roy_C y Sana_C, lo cual es coherente con la proximidad visual de estas condiciones en etapas tempranas de la enfermedad.

Curvas ROC y AUC

Las Figura 1 y Figura 2 muestran las curvas ROC multiclase (One-vs-Rest) para entrenamiento y prueba. El modelo alcanza un AUC macro = 0.987 en test, con valores individuales elevados para todas las clases, confirmando una alta capacidad discriminativa.

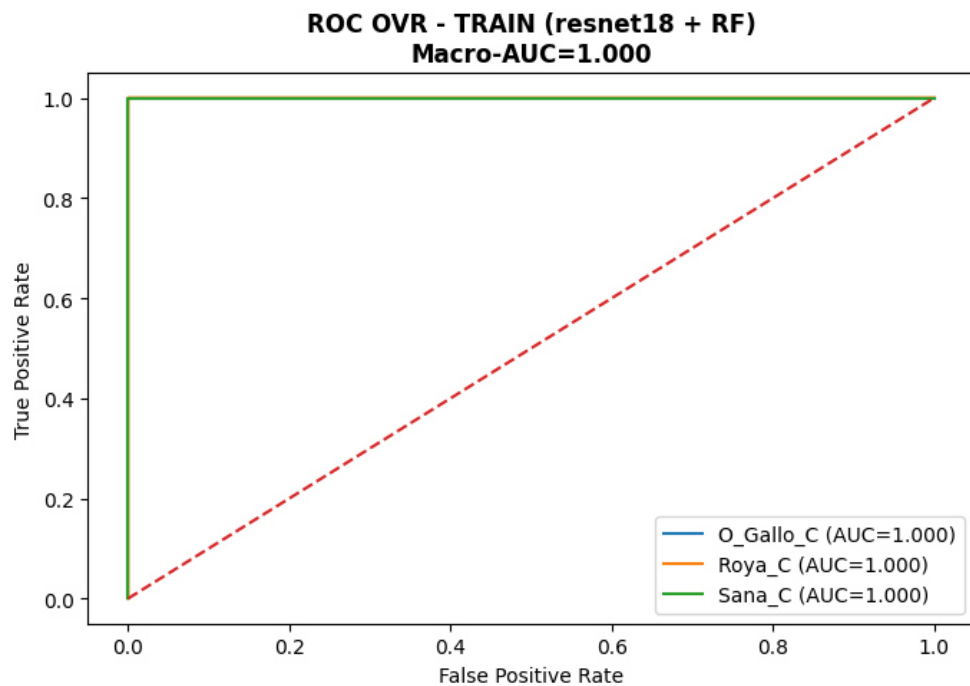


Figura 1. Curvas ROC OVR del modelo ResNet18 + RandomForest en entrenamiento

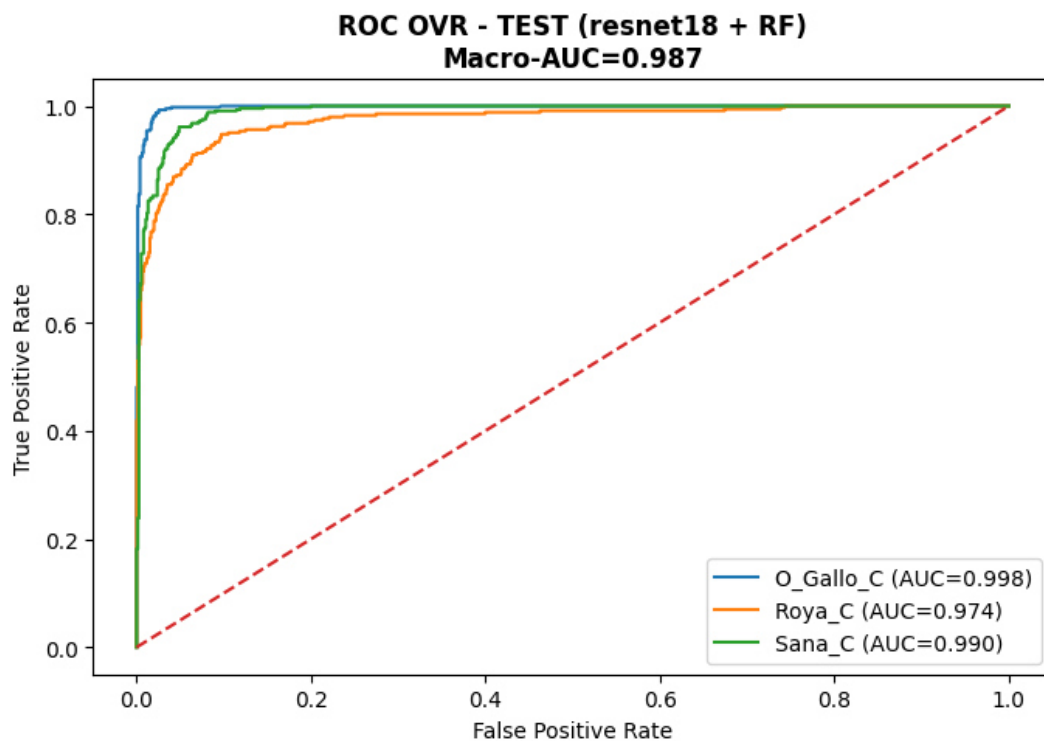


Figura 2. Curvas ROC OVR del modelo ResNet18 + RandomForest en prueba

Evolución por fold

La **Figura 14** ilustra la evolución de la accuracy y el F1-macro a lo largo de los 10 folds de validación cruzada, evidenciando estabilidad y ausencia de colapsos en el rendimiento.

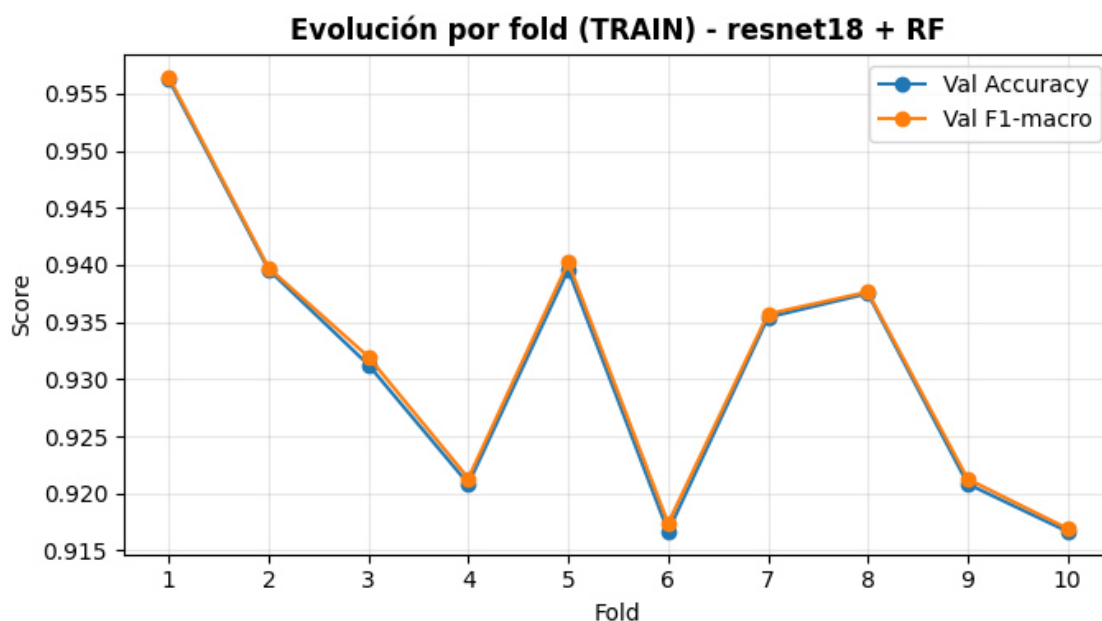


Figura 14. Evolución de la accuracy y el F1-macro

Conclusión

El modelo híbrido ResNet18 + RandomForest demostró un desempeño sólido, estable y generalizable en la clasificación automática de hojas de cacao, alcanzando métricas elevadas tanto en validación cruzada como en el conjunto de prueba independiente. Aunque su rendimiento es ligeramente inferior al obtenido con el clasificador SVM, este enfoque ofrece ventajas relevantes en términos de interpretabilidad, robustez y tolerancia al ruido. La combinación de un extractor profundo ligero con un clasificador de ensamble confirma la viabilidad de este modelo como una alternativa complementaria dentro del sistema propuesto, especialmente en escenarios donde se prioriza estabilidad y simplicidad computacional para futuras implementaciones en dispositivos de bajo consumo orientados a agricultores y técnicos de campo.

4.2.4 Modelo híbrido 3: ResNet18 + XGBoost

Justificación del modelo híbrido

Se diseñó un modelo híbrido que combina la arquitectura ResNet18 como extractor de características visuales con un clasificador XGBoost para la clasificación multiclase de hojas de cacao en las categorías Sana_C, Roya_C y O_Gallo_C.

ResNet18 fue seleccionada por ser una arquitectura convolucional ligera y eficiente, con un número reducido de parámetros en comparación con versiones más profundas (ResNet34, ResNet50), lo que evidencia su potencial de uso con restricciones computacionales con potencial para una implementación futura en dispositivos móviles utilizados por agricultores y técnicos de campo.

Por su parte, XGBoost es un algoritmo de gradient boosting altamente optimizado que destaca por su capacidad para modelar relaciones no lineales complejas, su robustez frente al sobreajuste y su excelente desempeño en espacios de características de alta dimensionalidad, como los embeddings generados por redes convolucionales profundas. La combinación de ambos enfoques permite aprovechar la riqueza representacional del aprendizaje profundo junto con la potencia discriminativa de los modelos de ensamble basados en árboles.

Configuración inicial y carga de datos

El pipeline se desarrolló en un entorno Google Colab, utilizando Python 3 y aceleración por GPU (CUDA). El dataset estructurado se organizó en tres clases balanceadas, con 2 000 imágenes por clase, alcanzando un total de 6 000 imágenes tras el proceso de aumento de datos.

Se realizó una partición 80/20, donde el 80 % de las imágenes se utilizó para entrenamiento y el 20 % para prueba, garantizando que el conjunto de test permaneciera completamente no visto durante el ajuste del modelo. Adicionalmente, para la validación interna y el ajuste de hiperparámetros se empleó validación cruzada estratificada de 10 folds.

Fine-tuning ligero del extractor ResNet18

Se cargó una instancia preentrenada de ResNet18 desde torchvision.models, con pesos iniciales ajustados sobre ImageNet. Con el objetivo de preservar las características generales aprendidas y evitar sobreajuste, se aplicó un fine-tuning ligero, entrenando únicamente la capa final (cabeza de clasificación), mientras que el resto de la red permaneció congelado.

El entrenamiento se realizó durante un máximo de cuatro épocas, empleando early stopping basado en la precisión de validación. Como resultado, se obtuvo una accuracy de validación interna máxima de 0.8427, valor suficiente para proceder a la extracción de características profundas.

Extracción de embeddings visuales

Una vez finalizado el fine-tuning, ResNet18 se utilizó como extractor fijo de características, generando vectores de 512 dimensiones para cada imagen.

Se extrajeron embeddings para:

- 4 800 imágenes del conjunto de entrenamiento (80 %)
- 1 200 imágenes del conjunto de prueba (20 %)

Estos embeddings fueron almacenados en memoria y posteriormente utilizados como entrada directa para el clasificador XGBoost, separando claramente la etapa de representación visual del proceso de clasificación.

Configuración del clasificador XGBoost

El clasificador XGBoost fue entrenado sobre los embeddings generados por ResNet18. Para optimizar su desempeño, se aplicó un proceso de búsqueda aleatoria de hiperparámetros

(RandomizedSearchCV), utilizando un split estratificado interno (10 % del conjunto de entrenamiento) y validación cruzada de 10 folds.

Entre los hiperparámetros optimizados se incluyeron:

- número de estimadores (n_estimators)
- profundidad máxima del árbol (max_depth)
- tasa de aprendizaje (learning_rate)
- regularización L1 y L2 (reg_alpha, reg_lambda)
- fracción de muestras y características (subsample, colsample_bytree)
- reducción de dimensionalidad mediante PCA

La mejor configuración alcanzó un F1-macro = 1.000 en el split de validación utilizado para la optimización, evidenciando una separación altamente efectiva entre clases.

Validación cruzada estratificada (10 folds)

Con los hiperparámetros óptimos, se aplicó un esquema de validación cruzada estratificada en diez particiones sobre el conjunto de entrenamiento. En cada fold se calcularon las métricas de accuracy, F1-macro y ROC-AUC macro, obteniéndose valores consistentes y estables.

Los resultados promedio mostraron accuracies cercanas al 96 %, con variaciones controladas entre folds, lo que confirma la robustez del modelo y su baja dependencia de particiones específicas del dataset.

Resultados de clasificación

El modelo híbrido ResNet18 + XGBoost alcanzó los siguientes resultados:

- Entrenamiento (80 %)
 - Accuracy: 1.0000
 - F1-macro: 1.0000
 - ROC-AUC: 1.0000
- Prueba (20 % – datos no vistos)
 - Accuracy: 0.9583
 - F1-macro: 0.9565
 - ROC-AUC: 0.9936

El reporte de clasificación en el conjunto de prueba evidencia un desempeño equilibrado entre clases, destacando O_Gallo_C como la clase mejor identificada, seguida de Sana_C y Roya_C, esta última con métricas ligeramente inferiores debido a la similitud visual con hojas sanas en ciertos estadios tempranos.

Visualización de resultados

La representación gráfica de los resultados facilita analizar de manera cualitativa y cuantitativa el comportamiento del modelo híbrido ResNet18 + XGBoost.

Matrices de confusión

Las matrices de confusión del conjunto de entrenamiento muestran una clasificación perfecta, sin errores entre clases (**Figura 15** y **Figura 16**). En el conjunto de prueba, se observa una diagonal dominante, con errores mínimos concentrados principalmente entre *Roya_C* y *Sana_C*, lo cual resulta coherente desde el punto de vista agronómico debido a la similitud morfológica en etapas iniciales de infección.

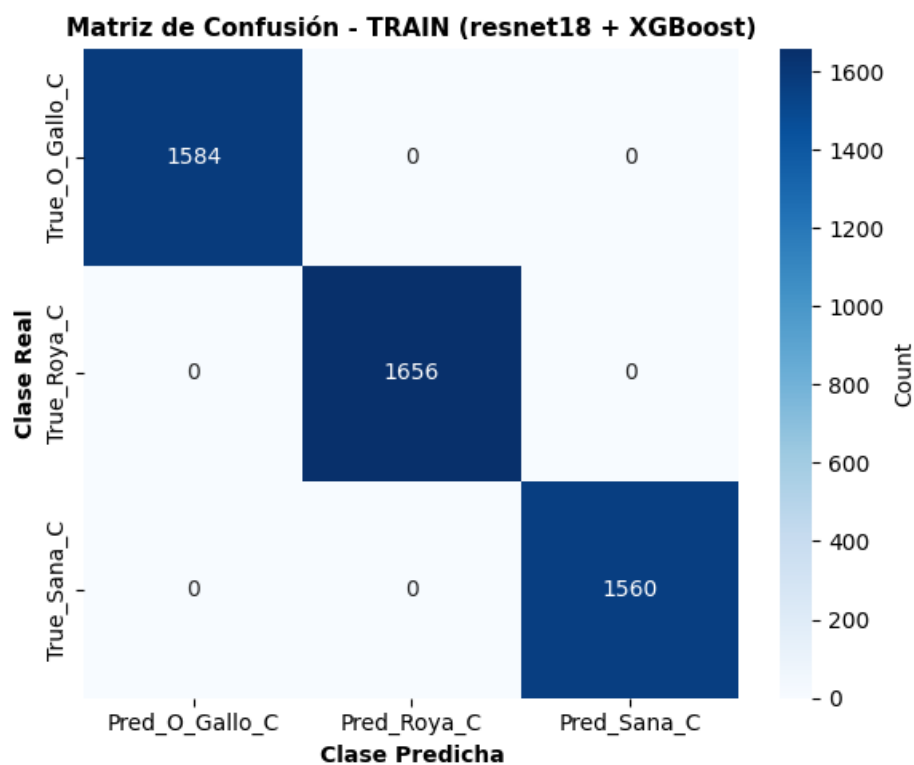


Figura 15. Matriz de confusión del modelo ResNet18 + XGBoost en el conjunto de entrenamiento

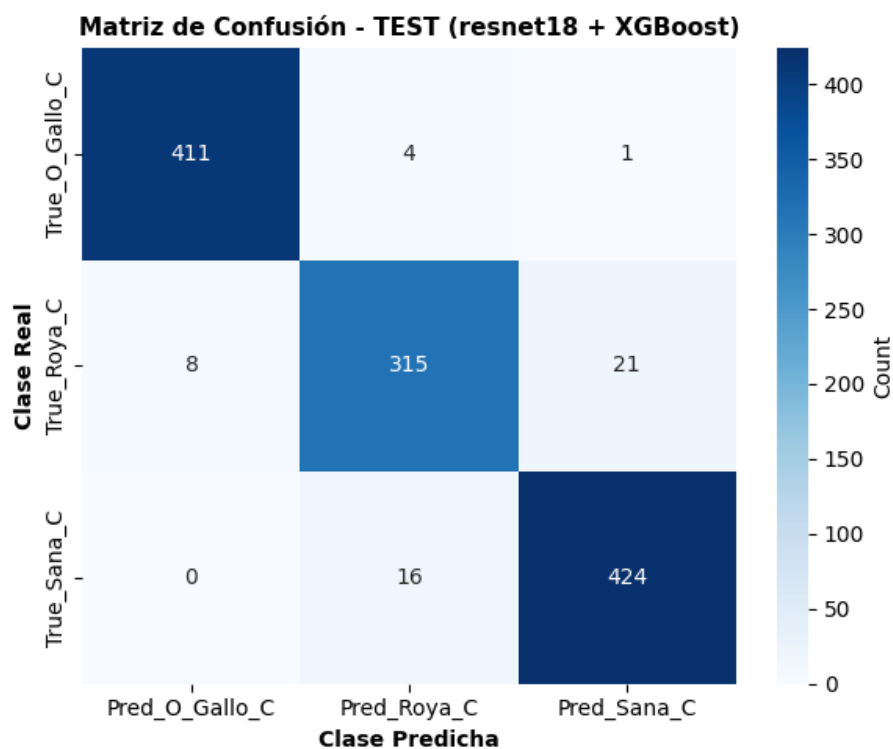


Figura 16. Matriz de confusión del modelo ResNet18 + XGBoost en el conjunto de prueba

Curvas ROC y AUC

Las curvas ROC muestran una excelente capacidad discriminativa para todas las clases (**Figura 17, Figura 18 y Figura 19**), con evidenciando una alta capacidad de separabilidad, reflejada en AUC superiores a 0.99 en gran parte de los escenarios analizados, confirmando una separación clara entre categorías.

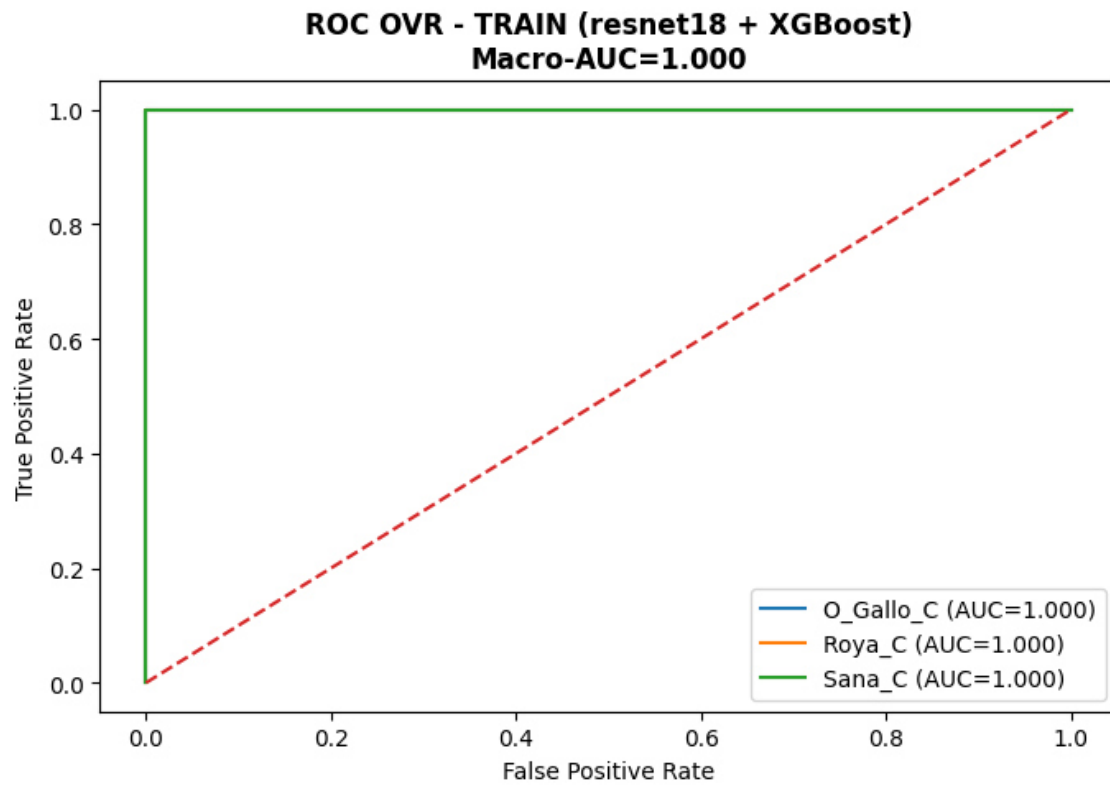


Figura 17. Curvas ROC OVR del modelo ResNet18 + XGBoost en entrenamiento

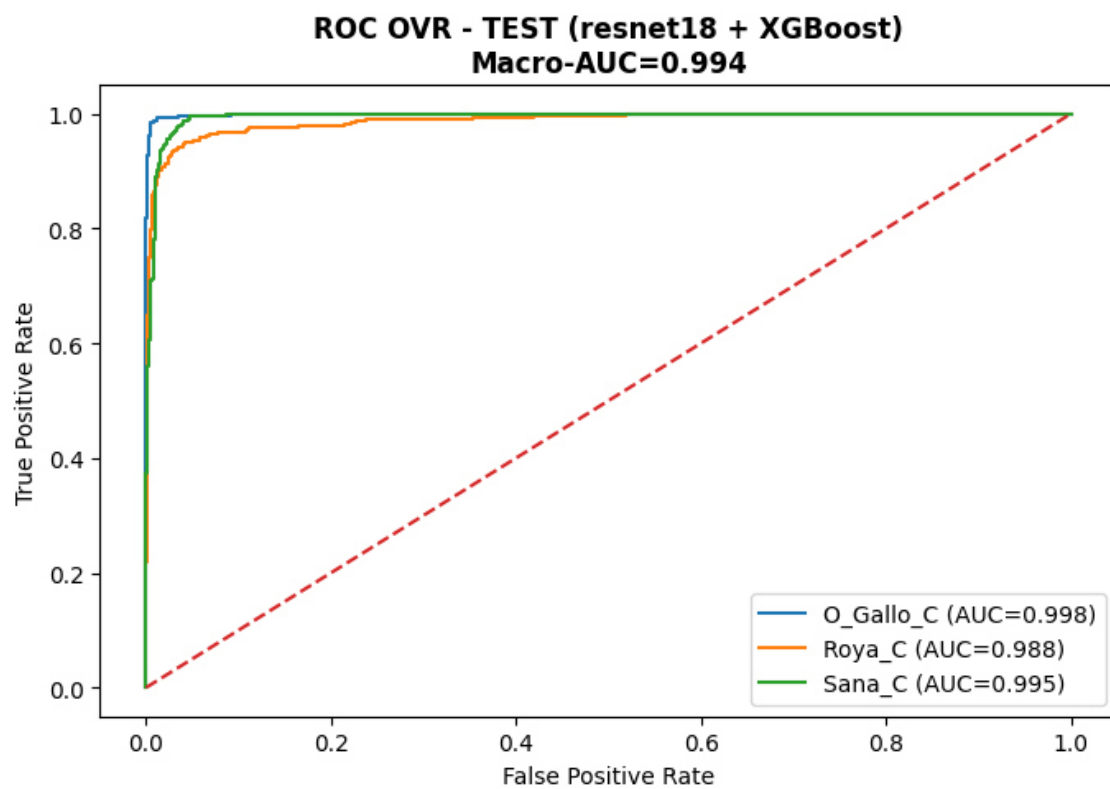


Figura 18. Curvas ROC OVR del modelo ResNet18 + XGBoost en prueba
Evolución por fold

La evolución de accuracy y F1-macro por fold evidencia estabilidad del modelo durante la validación cruzada, sin fluctuaciones abruptas.



Figura 19. Evolución de la accuracy y el F1-macro

Desde una perspectiva operativa, estos resultados confirman que el modelo es adecuado para escenarios reales de monitoreo agrícola, donde se requiere precisión, robustez y eficiencia computacional.

Conclusión

El modelo híbrido ResNet18 + XGBoost demostró un rendimiento elevado y consistente en la clasificación automática de hojas de cacao en las categorías Sana_C, Roya_C y O_Gallo_C, alcanzando métricas sobresalientes tanto en entrenamiento como en evaluación sobre datos no vistos. La combinación de un extractor convolucional ligero con un clasificador de gradient boosting permitió obtener una solución robusta, eficiente y altamente discriminativa, con un balance adecuado entre precisión y generalización. Su diseño modular, bajo requerimiento computacional y estabilidad frente a la variabilidad del conjunto de datos lo posicionan como una alternativa viable para su futura implementación en sistemas móviles o plataformas orientados a facilitar los procesos de decisión en contextos agrícolas reales.

4.2 Evaluación del desempeño del modelo híbrido propuesto analizando métricas de rendimiento en la identificación de enfermedades en las hojas del café

La evaluación comparativa de los tres modelos híbridos propuestos (ResNet18+SVM, ResNet18+Random Forest y ResNet18+XGBoost) se desarrolló mediante un enfoque estadístico inferencial basado en ANOVA, complementado con una prueba post hoc de Tukey HSD, utilizando como evidencia empírica los resultados obtenidos en los 10 folds de validación cruzada

(métrica principal: val_acc). Adicionalmente, se realizó un análisis descriptivo del desempeño en TEST (hold-out 20%), incorporando métricas globales (Accuracy, F1-macro y AUC macro) y una verificación de generalización mediante el gap entre el promedio de CV y el rendimiento en TEST. Todas las pruebas se ejecutaron con un nivel de significancia $\alpha = 0.05$.

4.3.1 Visualización de la distribución de resultados por métrica y modelo.

Con el propósito de comparar la dispersión y consistencia del desempeño entre modelos a lo largo de los 10 folds, se elaboraron gráficos de distribución y gráficos de medias basados en val_acc (mayor es mejor) (**Figura 20**, **Figura 21** y **Figura 22**).

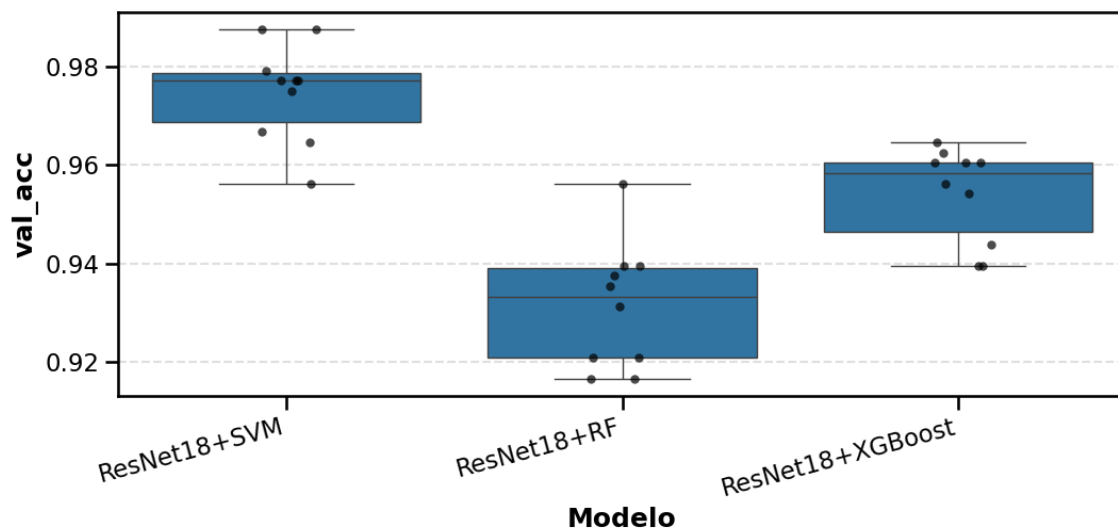


Figura 20. Distribución de val_acc en CV por modelo

- El diagrama de caja muestra que ResNet18+SVM concentra los valores más altos de val_acc y mantiene una dispersión relativamente baja, reflejando un comportamiento estable entre folds.
- ResNet18+XGBoost presenta un rendimiento intermedio, con variabilidad moderada y valores consistentemente por debajo de SVM.
- ResNet18+Random Forest exhibe la media más baja y una mayor tendencia a folds con desempeño inferior, lo cual sugiere menor robustez comparativa en validación.

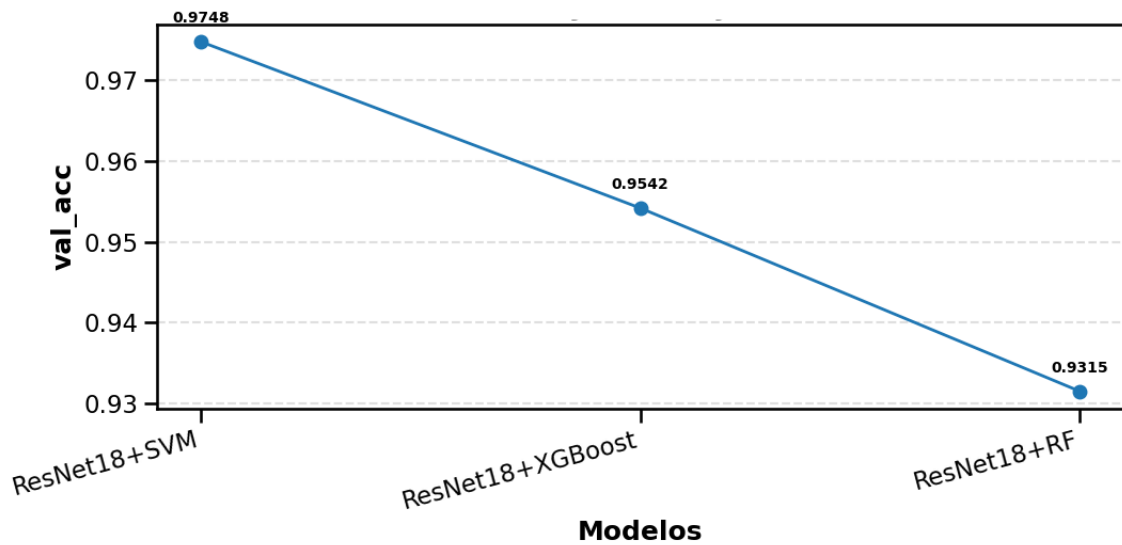


Figura 21. Gráfico de medias de val_acc en CV (TRAIN) por modelo

- Las medias obtenidas en validación cruzada confirman el orden de desempeño: ResNet18+SVM > ResNet18+XGBoost > ResNet18+RF, coherente con la separación observada en los boxplots.

De manera complementaria, se incluyó la comparación del rendimiento en el conjunto de prueba:

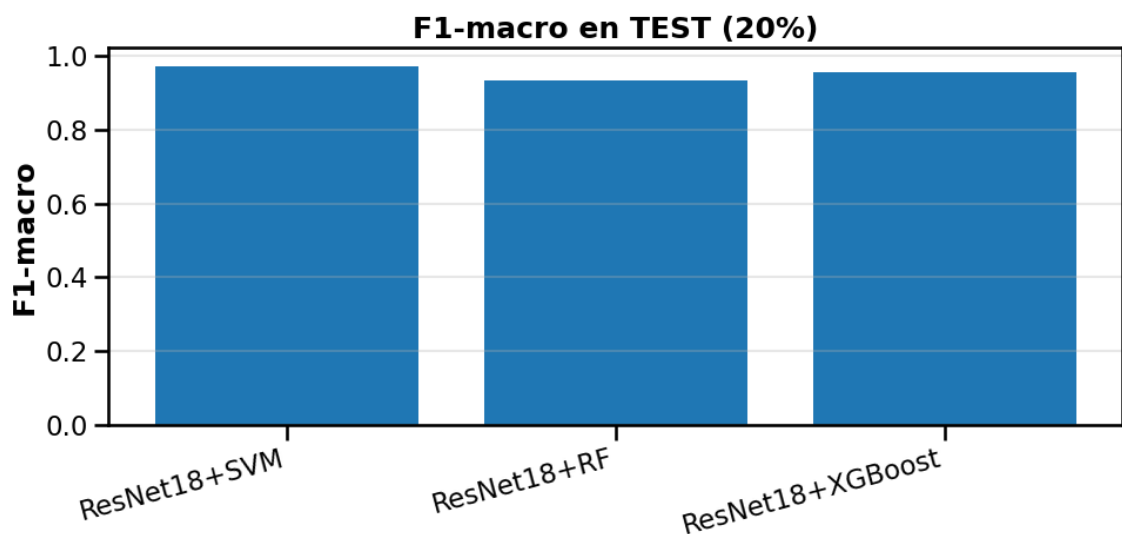


Figura 22. F1-macro en TEST (20%) por modelo

- El gráfico evidencia que ResNet18+SVM mantiene el mayor F1-macro en TEST, seguido por ResNet18+XGBoost, mientras que ResNet18+RF registra el valor más bajo.

Tablas de métricas por fold de los modelos híbridos evaluados

Con el fin de garantizar transparencia y trazabilidad del análisis inferencial, se presenta la tabla completa con val_acc por fold (10 ejecuciones) para cada modelo híbrido.

Tabla 5. Accuracy de validación (val_acc) 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

Estas métricas constituyen la base empírica utilizada en el ANOVA, permitiendo evaluar diferencias entre modelos controlando la variabilidad entre folds (bloqueo por fold).

4.3.2 Resultados del análisis ANOVA

Dado que los modelos fueron evaluados sobre la misma partición de folds, se aplicó un ANOVA con bloqueo por fold, considerando los factores modelo y fold. Los análisis realizados evidencian la existencia de diferencias significativas en el desempeño de los modelos evaluados.

Tabla 6. ANOVA para val_acc en CV (TRAIN)

Fuente	Suma de cuadrados	gl	MC	F	Sig.
C(model)	0.009396	2	0.004698	133.025489	1.647730e-11
C(fold)	0.002532	9	0.000281	7.965407	1.068262e-04
Residual	0.000636	18	0.000035	—	—

- Se obtuvo un p-valor (modelo) < 0.05 , por lo que se rechaza H_0 : existen diferencias significativas entre los modelos.
- El tamaño de efecto fue $\eta^2 = 0.937$ (grande), lo que indica que la diferencia entre modelos no solo es significativa, sino también sustantiva en magnitud.
- La prueba de Levene no evidenció problemas de homogeneidad ($p > 0.05$), apoyando la validez del contraste.

En términos descriptivos, las medias de **val_acc** en CV fueron:

- **ResNet18+SVM:** 0.974792
- **ResNet18+XGBoost:** 0.954167
- **ResNet18+RF:** 0.931458

4.3.3 Aplicación de la prueba POST HOC

Dado que el ANOVA resultó significativo, se aplicó la prueba Tukey HSD para identificar qué pares de modelos difieren entre sí.

Tabla 7. Comparaciones múltiples (Tukey HSD) para val_acc en CV (TRAIN)

Modelo A	Modelo B	Diferencia de medias	Sig.	IC 95% Inferior	IC 95% Superior	Rechaza H_0
ResNet18+RF	ResNet18+SVM	0.0433	0.0000	0.0313	0.0553	Sí
ResNet18+RF	ResNet18+XGBoost	0.0227	0.0002	0.0107	0.0347	Sí

ResNet18+SVM	ResNet18+XGBoost	-0.0206	0.0006	-	-0.0086	Sí
				0.0326		

Los resultados muestran que todos los pares difieren significativamente, confirmando el orden de rendimiento observado en los gráficos: SVM > XGBoost > RF.

Tabla 8. Desempeño en TEST (20% hold-out)

Modelo	Accuracy_test	F1macro_test	AUCmacro_test	n_train	n_test
ResNet18+SVM	0.974167	0.973033	0.996806	4800	1200
ResNet18+RF	0.936667	0.934056	0.987143	4800	1200
ResNet18+XGBoost	0.958333	0.956459	0.993550	4800	1200

Finalmente, el gap entre CV y TEST fue menor a 2% en todos los modelos, indicando buena generalización (sin caídas fuertes al pasar a datos no vistos).

4.3.4 Interpretación y selección del modelo óptimo

En conjunto, el ANOVA y la aplicación de la prueba post hoc de Tukey permitió corroborar la existencia de diferencias significativas entre los tres enfoques híbridos, siendo ResNet18+SVM el modelo con mejor desempeño global y mayor consistencia en validación cruzada (val_acc promedio = 0.974792), además de mantener el mejor rendimiento en TEST (Accuracy = 0.974167; F1-macro = 0.973033; AUC macro = 0.996806) con un gap mínimo entre CV y TEST, por lo que se selecciona ResNet18+SVM como el modelo óptimo para la identificación y clasificación de hojas de cafeto sanas, con roya y con ojo de gallo en el contexto del distrito de Saposo.

CAPÍTULO 5. RESULTADOS Y DISCUSIÓN

5.1 Análisis, interpretación y discusión de resultados

En este apartado se presenta un análisis general de los resultados obtenidos a partir de la evaluación del modelo de análisis automático basado en reconocimiento de patrones para la identificación de enfermedades en hojas de café. Los resultados presentados en el capítulo anterior reflejan el desempeño de los modelos híbridos evaluados a través de métricas como exactitud, F1-macro y AUC-ROC. Estas fueron calculadas utilizando un esquema de validación cruzada estratificada, así como un conjunto de prueba independiente. Este enfoque permite tener una visión general del comportamiento de los modelos antes de realizar la contratación de las hipótesis planteadas.

Antes de aplicar pruebas estadísticas inferenciales, se revisó la distribución de los datos para comprobar si se cumplía el supuesto de normalidad. Para ello, se utilizó la prueba de Shapiro-Wilk con un nivel de significancia de $\alpha = 0.05$. Los resultados obtenidos no evidenciaron desviaciones importantes respecto a una distribución normal, ya que los valores de significancia se mantuvieron por encima del umbral establecido. En base a esto, se consideró adecuado emplear pruebas estadísticas de tipo paramétrico en el análisis posterior.

En cuanto al análisis descriptivo, se pueden notar algunas diferencias en el desempeño de los modelos evaluados. En términos generales, la combinación ResNet18 + SVM muestra resultados más altos y consistentes en comparación con ResNet18 + Random Forest y ResNet18 + XGBoost. Este comportamiento se mantiene tanto en la validación cruzada como en el conjunto de prueba, lo que sugiere una mayor estabilidad en su rendimiento. Asimismo, la consistencia de los resultados a lo largo de los distintos pliegues de validación sugiere una adecuada capacidad de generalización del modelo propuesto frente a la variabilidad presente en las imágenes capturadas en condiciones reales de campo.

Estos resultados permiten establecer una base analítica sólida para la contratación de las hipótesis de investigación. La interpretación detallada de los hallazgos, su comparación con estudios previos y el análisis de sus implicancias técnicas y agronómicas se desarrollan de manera específica en el apartado de discusión, donde se profundiza en el alcance y las limitaciones del modelo propuesto.

5.2 Pruebas de hipótesis

5.2.1 Hipótesis específica H1

Para la hipótesis específica 1: “El conjunto de imágenes digitales de hojas de café recolectado y etiquetado presenta una estructura y calidad técnica significativamente

superior a los conjuntos de datos disponibles para el análisis de enfermedades foliares”, se planteó lo siguiente:

H₀: No superioridad del conjunto de datos propio frente a otros.

H_a: Superioridad del conjunto de datos propio frente a otros.

Tabla 9. Conjunto de datos a comparar

Conjunto de datos	Cantidad de imágenes	Cita
Propio	1500	(Santa Maria Pinedo, 2026)
RoCoLe	1560	(Parraga-Alava et al., 2019)
coffee_dataset	3312	(Chelangat et al., 2025)

Tabla 10. Valores de métricas evaluadas de cada conjunto de datos

Métrica	Propio	RoCoLe	coffee_dataset
Megapíxeles promedio	9.79	3.37	0.07
Nitidez (varianza Laplaciano)	218.03	2016.50	145.17
Brillo medio	189.50	116.34	143.50
Contraste medio	49.31	46.81	34.35
Saturación negra (ratio)	0.000001	0.000894	0.000634
Saturación blanca (ratio)	0.000089	0.004811	0.013137
Imágenes borrosas (%)	0.73 %	0.00 %	61.52 %
Imágenes subexpuestas (%)	0.00 %	0.96 %	0.16 %
Imágenes sobreexpuestas (%)	0.07 %	5.45 %	17.08 %

Tabla 11. Resumen de métricas superiores por conjunto de datos

Comparación	Total de métricas evaluadas	Métricas donde MI_DATASET es superior	Métricas donde el otro dataset es superior
Propio vs RoCoLe	9	7	2
Propio vs coffee_dataset	9	9	0

De acuerdo con los resultados obtenidos, el conjunto de datos propio presenta valores superiores en la mayoría de las métricas de calidad técnica evaluadas en comparación con los conjuntos de datos RoCoLe y coffee_dataset. En particular, el dataset propuesto supera a RoCoLe en 7 de 9 métricas y al coffee_dataset en la totalidad de las métricas evaluadas (9 de 9). En consecuencia, considerando el criterio de decisión basado en la cantidad de métricas superiores, se rechaza la hipótesis nula (H_0) y se acepta la hipótesis alternativa (H_a), concluyendo que el conjunto de imágenes digitales de hojas de café recolectado y etiquetado presenta una estructura y calidad técnica superior para el análisis de enfermedades foliares.

5.2.2 Hipótesis específica H2

Para la hipótesis específica 2: “Existe al menos un modelo híbrido basado en transferencia de conocimiento que muestra una capacidad significativamente mayor para extraer características relevantes y clasificar enfermedades en hojas de café en comparación

con modelos convencionales”, se planteó lo siguiente:

H₀: No existen diferencias significativas en el desempeño de los modelos evaluados para la clasificación de enfermedades en hojas de cafeto.

H_a: Existen diferencias significativas en el desempeño de los modelos evaluados para la clasificación de enfermedades en hojas de cafeto.

Para la contrastación de esta hipótesis se tomaron como referencia los resultados del análisis estadístico presentados en el Capítulo 4. En particular, se consideraron los resultados del análisis ANOVA con bloqueo por fold aplicado a la métrica val_acc en validación cruzada (**Tabla 6**), así como las comparaciones múltiples realizadas mediante la prueba post hoc de Tukey HSD (**Tabla 7**). Asimismo, se analizaron los resultados de desempeño obtenidos en el conjunto de prueba independiente (20%), presentados en la **Tabla 8**.

Criterio de decisión:

Si el valor de significancia obtenido en el análisis ANOVA para el factor modelo es inferior a 0.05, se rechaza la hipótesis nula. Posteriormente, la prueba post hoc de Tukey permite identificar los pares de modelos que presentan diferencias significativas y establecer el orden de desempeño correspondiente.

Resultados y veredictos:

De acuerdo con los resultados del análisis ANOVA, se obtuvo un valor de significancia $p < 0.05$ para el factor modelo, lo que evidencia la existencia de diferencias estadísticamente significativas en el desempeño de los modelos evaluados. El tamaño de efecto obtenido ($\eta^2 = 0.937$) indica que dichas diferencias no solo son significativas, sino también sustantivas en magnitud. La prueba post hoc de Tukey mostró que todos los pares de modelos difieren significativamente entre sí, confirmando el orden de rendimiento ResNet18 + SVM > ResNet18 + XGBoost > ResNet18 + Random Forest.

Este comportamiento se ve reforzado por los resultados obtenidos en el conjunto de prueba independiente, donde el modelo ResNet18 + SVM alcanzó los valores más altos en las métricas de exactitud, F1-macro y AUC-ROC, manteniendo además un gap inferior al 2% entre los resultados de validación cruzada y prueba, lo que evidencia una adecuada capacidad de generalización.

En consecuencia, considerando el criterio de decisión establecido y la evidencia estadística presentada, se rechaza la hipótesis nula (H₀) y se acepta la hipótesis alternativa (H_a), concluyendo que existe al menos un modelo híbrido basado en transferencia de conocimiento que presenta una capacidad superior para la extracción de características y

clasificación de enfermedades en hojas de cafeto.

5.2.3 Hipótesis específica H3

Para la hipótesis específica 3: “El modelo híbrido desarrollado mejora significativamente la precisión, confiabilidad y eficiencia en la identificación y clasificación de enfermedades en hojas de cafeto respecto a los enfoques tradicionales de diagnóstico visual”, se planteó lo siguiente:

H₀: No existe concordancia significativa entre la clasificación realizada por el modelo híbrido y la evaluación del experto.

H_a: Existe concordancia significativa entre la clasificación realizada por el modelo híbrido y la evaluación del experto.

Criterio de decisión:

Si $\kappa \leq 0.80$, se acepta la hipótesis nula.

Si $\kappa > 0.80$, se rechaza la hipótesis nula y se acepta la hipótesis alternativa.

Tabla 12. Comparación de diagnósticos del modelo híbrido y del experto

Nº Imágenes	Predicción Modelo	Respuesta del experto
1	Sana	Sana
2	Sana	Sana
3	Sana	Sana
4	Sana	Sana
5	Sana	Sana
6	Roya	Roya
7	Roya	Roya
8	Sana	Sana
9	Roya	Roya
10	Roya	Roya
11	Ojo de gallo	Ojo de gallo
12	Roya	Ojo de gallo
13	Ojo de gallo	Ojo de gallo
14	Ojo de gallo	Ojo de gallo
15	Ojo de gallo	Ojo de gallo

Tabla 13. Escala de interpretación del coeficiente de la Kappa de Cohen (K)

κ	Interpretación
< 0.20	Concordancia pobre
0.21 – 0.40	Débil
0.41 – 0.60	Moderada
0.61 – 0.80	Sustancial
0.81 – 1.00	Casi perfecta

Tabla 14. Resultado del análisis de concordancia entre el modelo híbrido y el experto

Comparación	Nº de imágenes	Coeficiente κ	Interpretación
Modelo híbrido vs Experto	15	0.899	Concordancia casi perfecta

Los resultados del análisis de concordancia entre la predicción del modelo híbrido y la evaluación

del experto, evaluados mediante el coeficiente de Cohen's Kappa, evidencian un valor de $\kappa = 0.899$, el cual, según la escala de interpretación adoptada, indica una concordancia casi perfecta. En función del criterio de decisión definido, se rechaza la hipótesis nula (H_0) y se acepta la hipótesis alternativa (H_a), confirmando que el modelo híbrido desarrollado presenta una alta concordancia y confiabilidad en la identificación y clasificación de enfermedades foliares en hojas de cafeto respecto al diagnóstico visual tradicional.

5.3 Discusión

Los resultados obtenidos en la presente investigación confirman que los modelos híbridos basados en reconocimiento de patrones, que integran redes neuronales convolucionales preentrenadas como extractores de características y clasificadores supervisados tradicionales, constituyen una estrategia eficaz y robusta para la identificación automática de enfermedades foliares del cafeto en condiciones reales de campo. En particular, el modelo ResNet18 + SVM mostró un desempeño superior y estadísticamente significativo frente a las combinaciones ResNet18 + Random Forest y ResNet18 + XGBoost, tanto en validación cruzada como en el conjunto de prueba independiente, lo que respalda la hipótesis general planteada en el estudio.

Estos hallazgos son coherentes con lo reportado por (Abuhayi & Mossa, 2023a), quienes demostraron que la combinación de CNN con técnicas adecuadas de extracción y clasificación permite alcanzar altos niveles de precisión en la identificación de enfermedades del café. Sin embargo, mientras dicho estudio se apoya en arquitecturas más profundas y en esquemas de preprocesamiento intensivo, los resultados del presente trabajo evidencian que un extractor más ligero como ResNet18, combinado con un clasificador SVM optimizado, puede lograr un desempeño comparable, manteniendo una menor complejidad computacional. Este aspecto resulta relevante para escenarios de aplicación práctica, especialmente en contextos agrícolas con recursos limitados.

Asimismo, los resultados concuerdan con lo planteado por (Mansouri et al., 2024), quienes resaltan el papel del aprendizaje por transferencia para desarrollar modelos precisos aun cuando el tamaño del conjunto de datos es limitado. No obstante, a diferencia de dicho estudio, que evalúa un gran número de arquitecturas preentrenadas bajo condiciones relativamente homogéneas, la presente investigación demuestra que la selección cuidadosa de una arquitectura ligera y su integración con clasificadores tradicionales puede ofrecer una solución más balanceada entre precisión, estabilidad y viabilidad operativa, sin depender de infraestructuras computacionales avanzadas.

En relación con los enfoques de ensamble reportados por (Novtahaning et al., 2022),

quienes evidencian mejoras en la estabilidad y reducción de errores mediante la combinación de múltiples CNN, los resultados de este estudio muestran que una arquitectura híbrida más simple también puede alcanzar niveles elevados de desempeño cuando se trabaja con un dataset contextualizado y capturado en condiciones reales. Esto sugiere que la robustez del modelo no depende únicamente de la complejidad arquitectónica, sino también de la representatividad del conjunto de datos y del control riguroso de la evaluación, aspectos que fueron cuidadosamente abordados en este trabajo. Por otro lado, los modelos de detección en tiempo real basados en YOLO, como los evaluados por (Fragoso et al., 2025), ofrecen ventajas claras en términos de localización de lesiones y velocidad de inferencia. Sin embargo, dichos enfoques se orientan principalmente a tareas de detección y no necesariamente a la clasificación detallada del estado fitosanitario completo de la hoja. En contraste, el enfoque propuesto en esta investigación se centra en la clasificación multiclase de hojas completas (sanas, con roya y con ojo de gallo), lo que resulta más adecuado como herramienta inicial de apoyo al diagnóstico y al monitoreo fitosanitario en campo, especialmente cuando el objetivo es facilitar decisiones tempranas de manejo.

De manera similar, los resultados obtenidos guardan correspondencia con los reportados por (Adelaja & Pranggono, 2025), quienes destacan la utilidad de modelos profundos ligeros para aplicaciones móviles. En este sentido, el desempeño alcanzado por ResNet18 + SVM refuerza la viabilidad de desarrollar herramientas portátiles basadas en visión artificial que puedan ser utilizadas directamente por productores o técnicos, sin sacrificar precisión diagnóstica ni estabilidad del sistema.

Un aspecto particularmente relevante de esta investigación es que confirma varias de las limitaciones metodológicas identificadas en el estado del arte. Tal como señalan (Atila et al., 2021) y (Barbedo, 2016), los modelos entrenados en datasets capturados bajo condiciones controladas tienden a sobreestimar su desempeño cuando se aplican en escenarios reales de campo. En este estudio, el uso de un dataset propio, capturado en plantaciones reales del distrito de Saposoa y validado por especialistas, permitió evaluar el comportamiento del modelo frente a variaciones de iluminación, fondo y severidad de síntomas, reduciendo el riesgo de sobreajuste y proporcionando una estimación más realista de su capacidad de generalización.

Asimismo, los resultados respaldan la postura crítica expuesta en el estado del arte respecto a la distinción entre clasificación visual automática y diagnóstico fitosanitario integral. Si bien el modelo desarrollado muestra una alta capacidad para identificar

patrones visuales asociados a roya y ojo de gallo, sus predicciones deben interpretarse como un apoyo al proceso de decisión y no como un reemplazo del diagnóstico agronómico completo, en concordancia con lo señalado por (Rocha-Calderón et al., 2023) y (Arif et al., 2025). En este sentido, la presente investigación contribuye a un uso más responsable y técnicamente fundamentado de la inteligencia artificial en el contexto agrícola.

Finalmente, el análisis estadístico inferencial mediante ANOVA y prueba post hoc de Tukey permitió demostrar que las diferencias observadas entre los modelos híbridos evaluados no solo son descriptivas, sino también estadísticamente significativas, lo que fortalece la validez de los resultados obtenidos. La superioridad del modelo ResNet18 + SVM, tanto en métricas de rendimiento como en estabilidad entre folds y generalización en el conjunto de prueba, confirma que este enfoque representa una alternativa óptima dentro del marco metodológico planteado.

En conjunto, los hallazgos de esta investigación no solo corroboran resultados previamente reportados en la literatura, sino que también aportan evidencia empírica en un contexto geográfico y productivo poco explorado, contribuyendo a cerrar la brecha entre desarrollos experimentales de laboratorio y aplicaciones reales en cafetales peruanos. Esto posiciona al modelo propuesto como una base sólida para futuras investigaciones y desarrollos tecnológicos orientados al monitoreo fitosanitario inteligente del cultivo de café.

CONCLUSIONES

1. Se logró diseñar y evaluar un modelo de análisis automático basado en reconocimiento de patrones, obteniendo resultados satisfactorios. En particular, el enfoque híbrido —que combina un extractor profundo con un clasificador supervisado— demostró ser efectivo para identificar y clasificar hojas de cafeto sanas, con roya y con ojo de gallo. Entre las configuraciones evaluadas, el modelo ResNet18 + SVM destacó por presentar el mejor desempeño y una mayor consistencia en los resultados.
2. Asimismo, se construyó un dataset de imágenes de hojas de cafeto adecuadamente estructurado y balanceado, a partir de capturas realizadas en campo y posteriormente etiquetadas y validadas. Esto permitió contar con una base confiable para el entrenamiento y la evaluación de los modelos. Además, la aplicación de técnicas de aumento de datos, de forma controlada, ayudó a incrementar el volumen de información disponible sin afectar el equilibrio entre las clases.
3. Se determinó que los modelos híbridos basados en transferencia de aprendizaje con ResNet18 como extractor de características y clasificadores supervisados (SVM, Random Forest y XGBoost) son viables y efectivos, siendo ResNet18 + SVM el modelo con mejor equilibrio entre precisión, estabilidad y capacidad discriminativa frente a las demás combinaciones evaluadas.
4. La evaluación mediante métricas de rendimiento y validación cruzada mostró resultados altos y estables, y el análisis estadístico inferencial (ANOVA y prueba post hoc) confirmó diferencias significativas entre los modelos, sustentando la selección de ResNet18 + SVM como alternativa óptima para la clasificación de las tres condiciones fitosanitarias.

RECOMENDACIONES

1. Se recomienda implementar el modelo ResNet18 + SVM como base de una herramienta de apoyo al diagnóstico (web o móvil), incorporando un flujo de captura estandarizado y retroalimentación clara, para facilitar decisiones oportunas de manejo fitosanitario en contextos cafetaleros como Saposoá.
2. Se recomienda ampliar el dataset incorporando mayor variabilidad real de campo (iluminación, fondos, severidad, oclusiones, diferentes parcelas y campañas) y registrar metadatos básicos (fecha, nivel de afectación), a fin de fortalecer la generalización del modelo.
3. Se recomienda evaluar extractores alternativos ligeros (p. ej., MobileNet o EfficientNet) manteniendo el esquema híbrido, y comparar adicionalmente el desempeño considerando consumo de memoria, tiempo de inferencia y facilidad de despliegue en dispositivos con recursos limitados.
4. Se recomienda realizar validación externa con imágenes nuevas de otra campaña o zona distinta, y priorizar el análisis por clase (especialmente el recall de roya) para reducir falsos negativos; además, se sugiere analizar casos difíciles (síntomas incipientes) para orientar mejoras de muestreo y preprocesamiento.

REFERENCIAS BIBLIOGRÁFICAS

- 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. (2023a). 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>
- Abuhayi, B. M., & Mossa, A. A. (2023b). 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>
- Acosta Trinidad, L. T., Ortiz Claro, W. R., Acosta Huaman, G. C., Alhuay Malpartida, D. S., Torres Lino, E. S., Vivas Garcia, R. P., & Cano Arevalo, K. L. (2025). Relación entre incidencia y severidad de roya amarilla, ojo de gallo y cercospora en cafetos de chilachi (Huancabamba, Oxapampa), Perú. *Anales Científicos*, 86(2), 47–56. <https://doi.org/10.21704/ac.v86i2.2322>
- 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>
- Alvarado-Huamán, L., Borjas-Ventura, R. R., Castro-Cépero, V., García-Nieves, L., Jimenez-Dávalos, J., Julca-Otiniano, A., & Gómez-Pando, L. (2020). Dynamics of severity of coffee leaf rust (*Hemileia vastatrix*) on Coffee, in Chanchamayo (Junin-Peru). *Agronomía Mesoamericana*, 517–529. <https://doi.org/10.15517/am.v31i3.39726>
- 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. (2015a). 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>
- Avelino, J., Cristancho, M., Georgiou, S., Imbach, P., Aguilar, L., Bornemann, G., Läderach, P., Anzueto, F., Hruska, A. J., & Morales, C. (2015b). 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>
- 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>

- Barbedo, J. G. A. (2016a). A review on the main challenges in automatic plant disease identification based on visible range images. *Biosystems Engineering*, 144, 52–60. <https://doi.org/10.1016/j.biosystemseng.2016.01.017>
- Barbedo, J. G. A. (2016b). A review on the main challenges in automatic plant disease identification based on visible range images. *Biosystems Engineering*, 144, 52–60. <https://doi.org/10.1016/j.biosystemseng.2016.01.017>
- Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer New York, NY.
- 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>
- Chelangat, S., Anirwoth, R., Mayanja, K. N., & Sserwadda, A. (2025). *A Machine Learning Dataset for Classification of Common Coffee Leaf Diseases in Uganda*. <https://doi.org/10.17632/k36wnd6knb.1>
- Dietterich, T. G. (1998). Approximate Statistical Tests for Comparing Supervised Classification Learning Algorithms. *Neural Computation*, 10(7), 1895–1923. <https://doi.org/10.1162/089976698300017197>
- 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>
- Ferentinos, K. P. (2018a). Deep learning models for plant disease detection and diagnosis. *Computers and Electronics in Agriculture*, 145, 311–318. <https://doi.org/10.1016/j.compag.2018.01.009>
- Ferentinos, K. P. (2018b). Deep learning models for plant disease detection and diagnosis. *Computers and Electronics in Agriculture*, 145, 311–318. <https://doi.org/10.1016/j.compag.2018.01.009>
- 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>
- Gichuru, E., Alwora, G., Gimase, J., & Kathurima, C. (2021). Coffee Leaf Rust (*Hemileia vastatrix*) in Kenya—A Review. *Agronomy*, 11(12). <https://doi.org/10.3390/agronomy11122590>
- 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. (2024a). New Races of *Hemileia vastatrix* Detected in Peruvian Coffee Fields. *Agronomy*, 14(8), 1811. <https://doi.org/10.3390/agronomy14081811>
- 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. (2024b). New Races of *Hemileia vastatrix* Detected in Peruvian Coffee Fields. *Agronomy*, 14(8). <https://doi.org/10.3390/agronomy14081811>
- Karunathilake, E. M. B. M., Le, A. T., Heo, S., Chung, Y. S., & Mansoor, S. (2023a). The Path to Smart Farming: Innovations and Opportunities in Precision Agriculture. *Agriculture*, 13(8). <https://doi.org/10.3390/agriculture13081593>
- Karunathilake, E. M. B. M., Le, A. T., Heo, S., Chung, Y. S., & Mansoor, S. (2023b). The Path to Smart Farming: Innovations and Opportunities in Precision Agriculture. *Agriculture*, 13(8). <https://doi.org/10.3390/agriculture13081593>
- Koutouleas, A., Collinge, D. B., & Boa, E. (2024). The coffee leaf rust pandemic: An ever-present danger to coffee production. *Plant Pathology*, 73(3), 522–534. <https://doi.org/10.1111/ppa.13846>
- Kumar, M., Gupta, P., Madhav, P., & Sachin. (2020). Disease Detection in Coffee Plants Using Convolutional Neural Network. *2020 5th International Conference on Communication and Electronics Systems (ICCES)*, 755–760. <https://doi.org/10.1109/ICCES48766.2020.9138000>
- LeCun, Y., Bengio, Y., & Hinton, G. (2015a). Deep learning. *Nature*, 521(7553), 436–444. <https://doi.org/10.1038/nature14539>
- LeCun, Y., Bengio, Y., & Hinton, G. (2015b). Deep learning. *Nature*, 521(7553), 436–444.

- <https://doi.org/10.1038/nature14539>
- M., M., & Zahraddeen Yakoo, I. (2024). *Artificial Intelligence-Based Computer Vision on Medical Images* (pp. 125–147). <https://doi.org/10.4018/979-8-3693-2426-4.ch007>
- Mansouri, N., Guessmi, H., & Alkhalil, 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. (2022a). 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>
- Martinez, F., Montiel, H., & Martinez, F. (2022b). 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>
- Martínez S., F. H., Montiel A., H., & G., E. J. (2022a). Identification of diseases and lesions in the coffee leaf using a convolutional model. *ARNP Journal of Engineering and Applied Sciences*, 17(8).
- Martínez S., F. H., Montiel A., H., & G., E. J. (2022b). Identification of diseases and lesions in the coffee leaf using a convolutional model. *ARNP Journal of Engineering and Applied Sciences*, 17(8).
- Mehta, B., Bharany, S., Elkamchouchi, D. H., Rehman, A. U., Singh, R., & Adem, S. H. (2025). EffResViT-SE FusionNet: A Hybrid Deep Learning Framework for Accurate Classification of Coffee Leaf Diseases. *Food Science & Nutrition*, 13(12). <https://doi.org/10.1002/fsn3.71311>
- Mendes, J., Pinho, T. M., Neves dos Santos, F., Sousa, J. J., Peres, E., Boaventura-Cunha, J., Cunha, M., & Morais, R. (2020). Smartphone Applications Targeting Precision Agriculture Practices—A Systematic Review. *Agronomy*, 10(6). <https://doi.org/10.3390/agronomy10060855>
- Mohanty, S. P., Hughes, D. P., & Salathé, M. (2016a). Using Deep Learning for Image-Based Plant Disease Detection. *Frontiers in Plant Science*, 7. <https://doi.org/10.3389/fpls.2016.01419>
- Mohanty, S. P., Hughes, D. P., & Salathé, M. (2016b). Using Deep Learning for Image-Based Plant Disease Detection. *Frontiers in Plant Science*, 7. <https://doi.org/10.3389/fpls.2016.01419>
- Montalbo, F. J. P. (2022a). Automated diagnosis of diverse coffee leaf images through a stage-wise aggregated triple deep convolutional neural network. *Machine Vision and Applications*, 33(19). <https://doi.org/10.1007/s00138-022-01277-y>
- Montalbo, F. J. P. (2022b). Automated diagnosis of diverse coffee leaf images through a stage-wise aggregated triple deep convolutional neural network. *Machine Vision and Applications*, 33(19). <https://doi.org/10.1007/s00138-022-01277-y>
- Nawaz, M., Nazir, T., Javed, A., Tawfik Amin, S., Jeribi, F., & Tahir, A. (2024). CoffeeNet: A deep learning approach for coffee plant leaves diseases recognition. *Expert Systems with Applications*, 237, 121481. <https://doi.org/10.1016/j.eswa.2023.121481>
- 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>
- Pan, S. J., & Yang, Q. (2010a). A Survey on Transfer Learning. *IEEE Transactions on Knowledge and Data Engineering*, 22(10), 1345–1359. <https://doi.org/10.1109/TKDE.2009.191>
- Pan, S. J., & Yang, Q. (2010b). A Survey on Transfer Learning. *IEEE Transactions on Knowledge and Data Engineering*, 22(10), 1345–1359. <https://doi.org/10.1109/TKDE.2009.191>
- Parraga-Alava, J., Cusme, K., Loor, A., & Santander, E. (2019). *RoCoLe: A robusta coffee leaf images dataset*. <https://doi.org/10.17632/c5yvn32dzg.2>
- Paulos, E. B., & Woldeyohannis, M. M. (2022). Detection and Classification of Coffee Leaf Disease using Deep Learning. *2022 International Conference on Information and*

- Communication Technology for Development for Africa (ICT4DA)*, 1–6.
<https://doi.org/10.1109/ICT4DA56482.2022.9971300>
- 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>
- Ristaino, J. B., Anderson, P. K., Bebber, D. P., Brauman, K. A., Cunniffe, N. J., Fedoroff, N. V., Finegold, C., Garrett, K. A., Gilligan, C. A., Jones, C. M., Martin, M. D., MacDonald, G. K., Neenan, P., Records, A., Schmale, D. G., Tateosian, L., & Wei, Q. (2021). The persistent threat of emerging plant disease pandemics to global food security. *Proceedings of the National Academy of Sciences*, 118(23). <https://doi.org/10.1073/pnas.2022239118>
- Rocha-Calderón, C. E., Barón-Velandia, J., & Vanegas-Ayala, S. C. (2023a). Supervised Model for the Detection of Coffee Leaf Diseases by Image Analysis. *International Journal of Intelligent Systems and Applications in Engineering*, 11(3), 405–411.
- Rocha-Calderón, C. E., Barón-Velandia, J., & Vanegas-Ayala, S. C. (2023b). Supervised Model for the Detection of Coffee Leaf Diseases by Image Analysis. *International Journal of Intelligent Systems and Applications in Engineering*, 11(3), 405–411.
- Salazar-Navarro, A., Ruiz-Valdiviezo, V., Joya-Dávila, J., & Gonzalez-Mendoza, D. (2024a). Coffee Leaf Rust (Hemileia vastatrix) Disease in Coffee Plants and Perspectives by the Disease Control. *Phyton*, 93(5), 923–949. <https://doi.org/10.32604/phyton.2024.049612>
- Salazar-Navarro, A., Ruiz-Valdiviezo, V., Joya-Dávila, J., & Gonzalez-Mendoza, D. (2024b). Coffee Leaf Rust (Hemileia vastatrix) Disease in Coffee Plants and Perspectives by the Disease Control. *Phyton*, 93(5), 923–949. <https://doi.org/10.32604/phyton.2024.049612>
- Santa Maria Pinedo, J. C. (2026). *Coffee leaf dataset by phytosanitary class*.
<https://doi.org/10.17632/mfpxg4y65r.1>
- SENASICA. (2014). *Ficha Técnica No.49 Ojo de gallo Mycena citricolor (Berkeley & Curtis)*.
- Sera, G. H., de Carvalho, C. H. S., de Rezende Abrahão, J. C., Pozza, E. A., Matiello, J. B., de Almeida, S. R., Bartelega, L., & dos Santos Botelho, D. M. (2022). Coffee Leaf Rust in Brazil: Historical Events, Current Situation, and Control Measures. *Agronomy*, 12(2), 496. <https://doi.org/10.3390/agronomy12020496>
- Shetty, A. (2024). Coffee Leaf Disease Detection Using CNN. *International Journal for Research in Applied Science and Engineering Technology*, 12(8), 439–444.
<https://doi.org/10.22214/ijraset.2024.63940>
- Shmueli, G. (2010). To Explain or to Predict? *Statistical Science*, 25(3).
<https://doi.org/10.1214/10-STS330>
- 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>
- Thakur, D., Gera, T., Aggarwal, A., Verma, M., Kaur, M., Singh, D., & Amoon, M. (2024). SUNet: Coffee Leaf Disease Detection Using Hybrid Deep Learning Model. *IEEE Access*, 12, 149173–149191. <https://doi.org/10.1109/ACCESS.2024.3476211>
- Too, E. C., Yujian, L., Njuki, S., & Yingchun, L. (2019a). A comparative study of fine-tuning deep learning models for plant disease identification. *Computers and Electronics in Agriculture*, 161, 272–279. <https://doi.org/10.1016/j.compag.2018.03.032>
- Too, E. C., Yujian, L., Njuki, S., & Yingchun, L. (2019b). A comparative study of fine-tuning deep learning models for plant disease identification. *Computers and Electronics in Agriculture*, 161, 272–279. <https://doi.org/10.1016/j.compag.2018.03.032>
- Zheng, R., & Kamruzzaman, M. (2024). Applications of hyperspectral imaging in the coffee industry: Current research and future outlook. *Applied Spectroscopy Reviews*, 59(8), 1061–1085. <https://doi.org/10.1080/05704928.2023.2295351>

ANEXOS.

1. Codificación:

<https://colab.research.google.com/drive/1xGNyHKKHQD96KiFlaj-xuLJ9mTo8TJy6P?usp=sharing>

2. Artículos publicados / aceptados: <https://doi.org/10.51252/rcsi.v6i1.1349>