



UNIVERSIDAD
DE LA REPÚBLICA
URUGUAY



FACULTAD DE
INGENIERÍA

Sistema de detección temprana del Picudo Rojo en palmeras Phoenix Canariensis

Informe de Proyecto de Grado presentado por

Matías Forcelledo, Bruno Gandós y Guzmán Pieroni

en cumplimiento parcial de los requerimientos para la graduación de la carrera
de Ingeniería en Computación de Facultad de Ingeniería de la Universidad de
la República

Supervisores

Ignacio Ramírez y Gonzalo Tejera

Montevideo, 1 de diciembre de 2025



**Sistema de detección
temprana del Picudo Rojo
en palmeras Phoenix**

Canariensis por Matías Forcelledo, Bruno Gandós y Guzmán
Pieroni tiene licencia [CC Atribución 4.0](#).

Agradecimientos

Queremos expresar nuestro más profundo agradecimiento a nuestras familias, parejas y amigos, quienes nos acompañaron incondicionalmente a lo largo de este proceso. Su apoyo constante, paciencia y aliento fueron fundamentales para mantenernos motivados y llegar hasta aquí. También recordamos con especial cariño a nuestros seres queridos que ya no están presentes, pero cuyo ejemplo y recuerdo nos inspiran a seguir adelante.

Agradecemos especialmente a nuestros tutores, Gonzalo Tejera e Ignacio Ramírez, por la guía brindada durante el desarrollo de este proyecto. Sus orientaciones, consejos y experiencia fueron clave para encaminar nuestro trabajo y mantener un avance constante.

Finalmente, extendemos nuestro reconocimiento a todas las personas e instituciones que contribuyeron de distintas formas a este proyecto, ya sea colaborando en la toma de datos, compartiendo sus conocimientos o facilitando recursos técnicos. Su aporte fue esencial para concretar esta investigación.

Resumen

El Picudo Rojo (*Rhynchophorus ferrugineus*) constituye actualmente una de las plagas más dañinas para las palmeras en Uruguay. Su detección temprana resulta crítica, ya que las larvas se desarrollan de manera interna durante varios meses sin manifestar síntomas visibles, reduciendo así las posibilidades de control y recuperación de los ejemplares afectados. En este contexto, el presente proyecto de grado propone el desarrollo de un sistema de detección temprana del picudo rojo basado en imágenes térmicas y RGB, apoyado en técnicas de procesamiento digital e inteligencia artificial. Para ello, se llevó a cabo una campaña de recolección de datos en diversas zonas de Montevideo, Uruguay, obteniendo un conjunto de imágenes de palmeras en distintos estados sanitarios (sanas, afectadas y muertas). Además, se incluyó procesamiento de datos contemplando técnicas de *inpainting*, recorte, conversión de escalas de temperatura a representaciones en escala de grises, y detección automática de la regiones de interés.

Se implementaron distintos enfoques de clasificación, evaluando modelos clásicos de aprendizaje automático (SVM y Random Forest) basados en atributos térmicos estadísticos, modelos híbridos que combinaron dichas características con descriptores extraídos mediante redes convolucionales preentrenadas, y modelos end-to-end que aprenden directamente a partir de las imágenes. Asimismo, se exploraron esquemas de fusión de predicciones entre modalidades térmicas y visuales.

Los resultados obtenidos muestran que la información térmica constituye el indicador más relevante para distinguir el estado sanitario de las palmeras. Los clasificadores basados en atributos térmicos alcanzaron desempeños superiores a los modelos puramente visuales, mientras que los enfoques híbridos, que combinan características tabulares y representaciones profundas extraídas mediante redes convolucionales, lograron el mejor equilibrio entre precisión y estabilidad. Entre ellos, el modelo híbrido destacó por su capacidad para reducir errores en la detección de palmeras afectadas. En conjunto, el sistema desarrollado constituye una prueba de concepto válida para la detección temprana del Picudo Rojo en palmeras, y representa un avance hacia el uso de metodologías de inteligencia artificial en el monitoreo fitosanitario urbano.

Palabras clave: Proyecto de Grado, Computación, Detección temprana, Picudo rojo, Termografía, Visión por computadora, Aprendizaje profundo, Palmeras

Índice general

1. Introducción	1
1.1. Contexto del problema	1
1.2. Objetivo general	2
1.3. Objetivo deseado	2
2. Antecedentes	3
2.1. Estado del Arte	4
2.1.1. An intelligent approach for detecting palm trees diseases using Image Processing and Machine Learning	4
2.1.2. Detection of Red Palm Weevil infected trees using thermal imaging	5
2.1.3. Water stress detection as an indicator of red palm weevil attack using WorldView data	7
2.1.4. Red Palm Weevil Detection in Date Palm Using Temporal UAV Imagery	7
2.2. Síntesis	8
3. Marco teórico	11
3.1. Fundamentos agronómicos	12
3.1.1. El picudo rojo de las palmeras	12
3.1.2. Palmeras Phoenix Canariensis	12
3.1.3. Importancia de la detección temprana	13
3.1.4. Técnicas de detección del picudo rojo	14
3.2. Técnicas de tratamiento de las palmeras	15
3.2.1. Tratamientos químicos	16
3.2.2. Control biológico	16
3.2.3. Métodos mecánicos y culturales	16
3.2.4. Métodos biotecnológicos	16
3.3. Fundamentos técnicos	16
3.3.1. Vehículos Aéreos no Tripulados (UAV)	17
3.3.2. Inercia térmica	17
3.4. Visión por computadora	18
3.4.1. Redes neuronales	18
3.4.2. Aprendizaje profundo	18

3.4.3. Aumento de datos	19
3.5. Procesamiento digital de imágenes	20
3.5.1. Inpainting	20
3.5.2. Homografía	21
3.5.3. Normalización y recorte	22
3.6. Técnicas de detección de objetos	22
3.6.1. You Only Look Once	23
3.7. Técnicas clasificadoras	24
3.7.1. ResNet18	25
3.7.2. ResNet50	25
3.7.3. DenseNet121	25
3.7.4. Random Forest	26
3.7.5. Support Vector Machine (SVM)	26
3.8. Métricas de evaluación de modelos	26
3.8.1. Matriz de confusión	26
3.8.2. Precisión	27
3.8.3. Exhaustividad	27
3.8.4. F1 Score	27
3.8.5. F1 Macro, Micro y Weighted	28
3.8.6. Mean Average Precision	28
3.8.7. Exactitud	29
3.8.8. Exactitud Balanceada	29
3.8.9. Coeficiente Kappa de Cohen	29
3.8.10. Coeficiente de Correlación de Matthews	30
3.8.11. Log Loss	30
3.8.12. Área bajo la curva ROC	30
3.8.13. Área bajo la curva Precisión-Recall	30
3.8.14. Exactitud Top-k	31
4. Desarrollo	33
4.1. Adquisición de datos	34
4.1.1. Descripción del entorno de captura	34
4.1.2. Sesiones y protocolos de captura	37
4.1.3. Especies y ejemplares estudiados	38
4.1.4. Instrumentación utilizada	40
4.1.5. Estructura del dataset	41
4.1.6. Preprocesamiento inicial	42
4.2. Preprocesamiento	45
4.2.1. Limpieza y conversión de imágenes térmicas	46
4.2.2. Transformación homográfica entre imágenes RGB y térmicas	48
4.2.3. Cambio de gradiente RGB a escala de grises	50
4.2.4. Aumento de datos	51
4.2.5. Marcador de regiones de palmeras	54
4.2.6. Variaciones del modelo: detección de tronco y hojas	57
4.2.7. Determinador de temperatura promedio	60
4.3. Clasificación	64

4.3.1. Consideraciones generales	64
4.3.2. Clasificador de Línea Base	65
4.3.3. Clasificador con Machine Learning	65
4.3.4. Clasificador híbrido con CNN y Machine Learning	67
4.3.5. Clasificador basado en CNN entrenada de extremo a extremo	68
4.3.6. Clasificador híbrido con optimización y reentrenamiento	69
4.3.7. Consolidación de los modelos generados	71
5. Experimentación	73
5.1. Diseño experimental	74
5.1.1. Hipótesis	74
5.1.2. Dataset y etapas empleadas	75
5.1.3. Métricas de evaluación	76
5.1.4. Control de variables y supuestos	76
5.1.5. Procedimiento general	77
5.2. Etapas de experimentación	78
5.2.1. Etapa inicial	78
5.2.2. Etapa intermedia	80
5.2.3. Etapa final	88
5.3. Discusión general de resultados	102
6. Conclusiones	105
7. Trabajo Futuro	107
A. Distribuciones de temperaturas con dataset completo	109
B. Resultados experimentales complementarios	111
Referencias	113

Capítulo 1

Introducción

1.1. Contexto del problema

La detección temprana del picudo rojo (*Rhynchophorus ferrugineus*) cobra una importancia crítica en Uruguay debido a su reciente aparición y rápido avance desde 2022. Este insecto, un coleóptero (de la familia de los escarabajos) de color rojizo de entre 2–4 cm de longitud, ataca palmeras ornamentales y nativas excavando galerías profundas que destruyen el *meristemo* (principal centro de crecimiento en las palmeras, ubicado en el centro de la corona) y la estructura interna del tronco. Las larvas son las principales responsables del daño, ya que pueden crecer hasta 3 cm y alimentarse durante más de 90 días, provocando la muerte de las palmeras infectadas.

Según el Ministerio de Ganadería, Agricultura y Pesca, la plaga está actualmente establecida en departamentos como Montevideo, Canelones, Maldonado, Florida, Flores, San José, Colonia, Lavalleja y Durazno, y afecta especies como *Phoenix canariensis*, *Butia odorata* y *Syagrus romanzoffiana* ([Ministerio de Ganadería, Agricultura y Pesca - Gobierno del Uruguay, 2024](#)). Desde su detección inicial en 2022 (Canelones, Montevideo, San José), ha colonizado múltiples departamentos y se estima que ha afectado a más de 10.000 ejemplares, aumentando la preocupación a escala nacional.

La detección oportuna se complica por la naturaleza sigilosa del picudo rojo: durante varios meses, el insecto se desarrolla dentro del tronco sin manifestar síntomas externos visibles, lo cual reduce significativamente el margen para tratamientos efectivos ([Ministerio de Ganadería, Agricultura y Pesca - Gobierno del Uruguay, 2024](#)). En este contexto, adquiere especial relevancia el uso de métodos que permitan detectar anomalías internas antes de la aparición de síntomas externos evidentes.

Dada la rápida propagación, la capacidad de vuelo del insecto (hasta 5km/día) y su alta tasa reproductiva (hasta tres generaciones por año en climas templados), la respuesta nacional requiere detección precoz, georreferenciación precisa y gestión activa en campo. El plan de contingencia nacional prevé estrategias

coordinadas entre intendencias y organismos como SINAE para monitoreo, remoción de palmeras afectadas y uso autorizado de insecticidas sistémicos o foliares ([Sistema Nacional de Emergencias \(SINAE\), 2024](#)).

La magnitud del problema ha generado una respuesta urgente por parte de la Intendencia de Montevideo, que ha manifestado la necesidad de contar con herramientas confiables para la detección temprana de ejemplares infestados. En este marco, distintos actores institucionales, como la Facultad de Ingeniería de la Universidad de la República y la Armada Nacional, han impulsado proyectos complementarios que exploran diversas técnicas de diagnóstico. Entre estas se encuentran el uso de drones equipados con cámaras multiespectrales, la identificación acústica de sonidos asociados a la actividad de las larvas en el interior del tronco, y el análisis visual mediante algoritmos de visión por computadora. Estas iniciativas reflejan la importancia de abordar el problema desde múltiples enfoques tecnológicos, buscando soluciones integrales que permitan contener el avance del picudo rojo y preservar el patrimonio forestal y paisajístico de la ciudad.

1.2. Objetivo general

Investigar y evaluar una metodología basada en imágenes térmicas y RGB, combinadas con técnicas de procesamiento digital y aprendizaje automático, para la detección de infestaciones por picudo rojo en palmeras. La investigación abarca desde el diseño del protocolo de captura y la construcción del conjunto de datos, hasta el desarrollo de una secuencia de preprocesamiento y análisis, con el fin de validar la factibilidad técnica y científica de este enfoque en el contexto uruguayo.

1.3. Objetivo deseado

Como objetivo deseado, se plantea que la metodología desarrollada pueda aplicarse de manera práctica en campo para la detección temprana de palmeras recientemente afectadas, es decir, en etapas iniciales de la infestación en las que aún es posible revertir el daño mediante tratamientos como la *endoterapia* ([Fertinyect, 2025](#)). De este modo, se busca contribuir con una herramienta tecnológica que permita a las autoridades y equipos de gestión urbana actuar de forma preventiva y más eficiente frente al avance de la plaga.

Capítulo 2

Antecedentes

En los últimos años, el avance de la inteligencia artificial y la visión por computadora ha permitido el desarrollo de soluciones innovadoras para la detección automática de plagas en la agricultura. Esta tendencia ha sido impulsada por la necesidad de identificar de forma temprana y precisa amenazas como en este caso, el picudo rojo, cuya detección visual suele ser tardía debido a la ausencia de síntomas externos en las etapas iniciales de infestación. En este contexto, diversos trabajos han abordado la aplicación de técnicas como el aprendizaje profundo, el uso de imágenes térmicas, la superresolución, y la integración de sensores acústicos para mejorar el monitoreo fitosanitario. A continuación, se presentan los estudios más relevantes que han sido analizados en profundidad por su aporte técnico y metodológico con respecto a los objetivos del presente proyecto.

2.1. Estado del Arte

Es importante señalar que, si bien se elaboró un documento específico de estado del arte (Pieroni, Gando, y Frocelledo, 2025) en una etapa temprana del proyecto, dicho trabajo fue desarrollado cuando el tema de investigación se centraba en el control general de plagas en el sector agrícola. Posteriormente, el enfoque del proyecto se redefinió hacia esta dirección. Por este motivo, esta sección incluye apartados de estado del arte, actualizados y más directamente vinculados con el nuevo alcance del proyecto. De esta forma se garantiza que el marco teórico y la revisión bibliográfica reflejen con mayor precisión los avances más relevantes para el caso de estudio abordado.

2.1.1. An intelligent approach for detecting palm trees diseases using Image Processing and Machine Learning

Este trabajo (Alaa y cols., 2020) constituyó una referencia inicial para la detección temprana de palmeras afectadas por el picudo rojo. A partir de sus resultados, se formuló la hipótesis de que las palmeras infestadas por el picudo rojo presentan una temperatura interna superior a la de las palmeras sanas. Además, el artículo sirvió como base metodológica para el presente estudio, proponiendo la captura de la temperatura interna de las palmeras mediante cámaras infrarrojas.

El estudio no se limitó únicamente al picudo rojo, sino que también abordó otras enfermedades comunes de las palmeras como *Leaf Spots* y *Blight Spots*, las cuales afectan principalmente las hojas. Para estas, los autores implementaron un modelo de *Red Neuronal Convolucional* (Convolutional Neural Networks, por su traducción del inglés) (IBM, 2025c), mientras que para la detección del picudo rojo se optó por un clasificador basado en *Máquinas de Vectores de Soporte* (Support Vector Machine, según su traducción del inglés) (IBM, 2025b), dada la escasez de datos disponibles. El sistema propuesto integraba imágenes normales y térmicas, lo que permitió abarcar tanto patologías visibles en hojas como plagas internas que afectan al estípite (nombre otorgado al tronco de las palmeras). Incluso se propuso una aplicación móvil para la adquisición de imágenes y diagnóstico en tiempo real, con el objetivo de facilitar la implementación en campo.

En cuanto a la adquisición de datos, el conjunto relacionado al picudo rojo se construyó a partir de 16 imágenes de palmeras sanas y 24 infectadas, capturadas en diferentes momentos del día (12 pm y 10 pm) en el Palm Research Center de Giza, Egipto, utilizando una cámara Testo 890-2. Debido al reducido tamaño de la muestra, se aplicaron técnicas de *aumento de datos* (Data Augmentation, según su traducción) (IBM Corporation, 2025) para expandir cada clase a 1200 imágenes. Asimismo, se construyó un dataset foliar a partir de bases públicas y capturas adicionales con cámaras móviles, generando más de 5000 imágenes

por clase tras aplicar aumentos.

El procesamiento de las imágenes térmicas incluyó la extracción de características estadísticas y texturales mediante la Matriz de Co-ocurrencia de Nivel de Gris (GLCM), lo que permitió alimentar al modelo SVM. Los resultados experimentales reportaron una precisión del 97.9 % en la clasificación de manchas foliares con CNN y un 92.8 % en la detección de palmeras infestadas por picudo rojo con SVM, siendo considerados los mejores resultados conocidos en el área hasta la fecha. Además, los autores validaron su sistema en diferentes condiciones de iluminación y horarios, demostrando robustez frente a variaciones ambientales. No obstante, dado el reducido tamaño inicial del conjunto de datos (40 imágenes) y la dependencia del aumento artificial de muestras, es posible que los modelos reportados presenten cierto grado de sobreajuste, lo cual plantea dudas sobre su capacidad de generalización en condiciones de campo diversas. Esta limitación refuerza la necesidad de generar un nuevo relevamiento en el contexto local, con un protocolo de captura propio y adaptado a las condiciones ambientales de Uruguay.

Entre las principales contribuciones del trabajo se destacan: la creación del primer dataset público de imágenes térmicas de palmeras infectadas con picudo rojo, la demostración de la viabilidad de la termografía como método no invasivo y la integración de modelos de aprendizaje automático en una aplicación práctica. Como líneas de trabajo futuro, los autores señalaron la necesidad de ampliar la base de datos, incorporar imágenes aéreas mediante drones o satélites, y explorar técnicas hiperespectrales e inteligencia de enjambre para mejorar la precisión y escalabilidad del sistema.

Sin embargo, debido a las diferencias de temperatura entre Egipto y nuestro país, fue descartada la idea de utilizar el dataset creado por los autores del trabajo, aunque sus resultados continúan siendo una referencia fundamental en la validación de nuestra metodología.

2.1.2. Detection of Red Palm Weevil infected trees using thermal imaging

Este estudio (Golomb, Alchanatis, Cohen, Levin, y Soroker, 2015) explora una solución innovadora: el uso de imágenes térmicas para identificar árboles infectados antes de que aparezcan síntomas visibles. La hipótesis central es que las larvas, al destruir el sistema vascular de la palmera, generan condiciones locales de estrés hídrico que se reflejan en incrementos de temperatura del follaje, medibles mediante cámaras infrarrojas. Este razonamiento se apoya en trabajos previos sobre la relación entre temperatura foliar y estrés hídrico en cultivos (Gates, 1964; Tanner, 1963).

Durante tres años (2012–2014), los investigadores llevaron a cabo seis experimentos controlados en condiciones de cuarentena en la estación Eden (Israel),

infestando intencionalmente ejemplares de *Phoenix canariensis* y *Phoenix dactylifera*. En cada campaña se compararon palmeras controladas con palmeras infestadas, realizando mediciones periódicas desde la segunda hasta la decimoquinta semana posterior a la infestación. Las imágenes se obtuvieron con una cámara térmica FLIR SC655 (640×480 px, sensibilidad $0,1^\circ\text{C}$) montada en plataforma, junto con imágenes RGB para identificación de la copa. Como referencia fisiológica, las palmeras fueron disecadas al final de cada experimento para evaluar el grado real de daño causado por las larvas.

El procesamiento de las imágenes térmicas se realizó con el software *Thermal-CAM Researcher* (Gauthier, Marin, Bodnar, y Stubbe, 2014) y posteriormente en *Matlab* (Ktep, 2024) y *ArcGIS* (Pucha-Cofrep y Fries, 2025). Para cuantificar el estrés hídrico se utilizó el *Crop Water Stress Index* (CWSI, según sus siglas en inglés) (Jones, 2014). Los resultados mostraron que en 64 % de las campañas las palmeras infestadas presentaron valores de CWSI superiores a las controladas, con mayor consistencia entre la séptima y octava semana tras la infestación, cuando hasta el 71 % de los árboles infestados mostraron diferencias estadísticamente significativas. Sin embargo, en etapas muy tempranas (semanas 1–4) las diferencias fueron menos claras, debido a que el daño interno aún no era suficiente para afectar significativamente el transporte de agua.

Además de los ensayos en invernadero, en 2013 se realizaron pruebas en un palmeral comercial de dátiles (*Hayani*) en Israel, donde se adquirieron imágenes aéreas a 600 m de altura con la misma cámara. Mediante un algoritmo de segmentación por *watershed* y un umbral adaptado por el método de Otsu, se separaron las copas de las palmeras del fondo del suelo y se calcularon los valores de CWSI por árbol. Utilizando como criterio de detección el límite superior del intervalo de confianza del grupo control, el sistema logró identificar correctamente el 75 % de las palmeras infestadas antes de que se manifestaran síntomas visibles. No obstante, también clasificó erróneamente como infestadas a un 25 % de los ejemplares sanos, lo que pone de manifiesto la sensibilidad del método a la selección del umbral de CWSI y la influencia de factores externos como el riego, la edad o el estado hídrico general de las plantas.

El estudio concluye que la termografía, tanto terrestre como aérea, es una herramienta prometedora para la detección temprana del picudo rojo en plantaciones extensivas, ya que permite generar productos de teledetección cuantitativos y no invasivos. No obstante, los autores remarcan limitaciones como la dependencia de la severidad del daño, la variabilidad en el estado hídrico por riego, edad o enfermedades no relacionadas, y la sensibilidad de los resultados al umbral de CWSI utilizado.

2.1.3. Water stress detection as an indicator of red palm weevil attack using WorldView data

Este trabajo (Bannari, Abdulaziz, y El Battay, 2017) propone por primera vez el uso de índices espectrales de estrés hídrico derivados de datos simulados del satélite *WorldView-3* (WV-3, según su abreviación) (European Space Agency, 2025), para discriminar diferentes niveles de infestación en palmeras datileras mediante sensores remotos.

La investigación se llevó a cabo en Bahréin, un entorno árido y altamente afectado por la reducción de palmeras a causa de factores climáticos, limitaciones de agua y la propagación del picudo rojo (Bannari y cols., 2017). En este contexto, más de 400 palmeras fueron muestreadas y clasificadas en cinco categorías: sanas, atacadas-tratadas, atacadas-no tratadas, severamente atacadas y muertas. Para cada árbol se tomaron muestras foliares y se adquirieron mediciones espectroradiométricas en laboratorio utilizando un *espectroradiómetro ASD* (Universidad del País Vasco (UPV/EHU) - Grupo de Investigación IBeA, 2025), con un rango espectral de 350 a 2500 nm. Estos espectros fueron posteriormente simulados de acuerdo a las características de las bandas del WV-3 mediante el código de transferencia radiativa canadiense *CAM5S*.

El estudio evaluó seis índices de estrés hídrico previamente establecidos en la literatura (NDWI, SRWI, SIWSI-1, SIWSI-2, NDII y NMDI), ampliamente vinculados a variaciones en el contenido de agua y pigmentación foliar (Gao, 1996; L. Wang, Qu, Hao, y Zhu, 2007; Fensholt y Sandholt, 2003; Hardisky, Klemas, y Smart, 1983). Los resultados mostraron que índices como NDII y SRWI presentaron una correlación significativa con los niveles de estrés ($R^2 \approx 0,95$), aunque su rango dinámico de detección de contenido hídrico estuvo limitado entre el 10 % y el 55 %.

Los autores desarrollaron cinco nuevos *índices de estrés hídrico específicos para palmeras* (PTWSI-1 a PTWSI-5), basados en combinaciones de bandas SWIR del WV-3. Estos nuevos índices alcanzaron coeficientes de determinación de hasta 0.90 con funciones polinomiales de segundo orden y mejoraron notablemente el rango dinámico, alcanzando valores de hasta el 100 %. Esta capacidad permitió diferenciar con mayor precisión entre clases de ataque, incluso en casos de diferencias sutiles entre palmeras jóvenes y adultas.

El estudio concluye que la teledetección mediante índices de estrés hídrico derivados de sensores de alta resolución, como el WV-3, constituye un enfoque prometedor y no invasivo para la detección temprana del RPW. En particular, el monitoreo del contenido hídrico actúa como un *proxy* del estado fisiológico del árbol, siendo capaz de detectar alteraciones en etapas tempranas de infestación cuando los síntomas externos aún no son visibles.

2.1.4. Red Palm Weevil Detection in Date Palm Using Temporal UAV Imagery

En este caso, el artículo (Delalieux y cols., 2023) no solo presenta una revisión, sino que reporta un experimento práctico para evaluar cómo la tele-

detección mediante Vehículos Aéreos No Tripulados puede emplearse para la detección temprana del picudo rojo en *palmeras datileras* (*Phoenix dactylifera*, por su nombre científico) (Raíz Verde, 2025). La investigación se llevó a cabo en Aspe (Alicante, España) entre 2020 y 2021, en un lote experimental de 18 palmeras (9 infestadas intencionalmente y 9 controles tratadas) que fueron monitoreadas durante dos temporadas de cultivo.

El artículo explora diversas tecnologías de detección remota, pero centra su aporte en un protocolo estandarizado de observaciones para UAV's, que incluyó imágenes RGB, *multiespectrales* (VNIR, por sus siglas en inglés) (Aerial Productions, 2025) y *térmicas* (Mesurex, 2025) adquiridas a lo largo del tiempo. A partir de estas imágenes, se aplicaron procesos de segmentación de copas, cálculo de índices espectrales y extracción de métricas térmicas como la temperatura máxima y la desviación estándar dentro de cada copa. El análisis permitió detectar patrones anómalos en palmeras infestadas, tales como la ausencia del incremento estacional de 'verdor' observado en palmeras sanas y la presencia de firmas térmicas diferenciadas incluso antes de que aparecieran síntomas visibles.

Los resultados mostraron que las diferencias térmicas fueron estadísticamente significativas ($\alpha = 0,05$) en etapas severas de infestación. Asimismo, la serie temporal de índices espectrales demostró ser una herramienta robusta para distinguir árboles infestados de sanos, ya que las palmeras enfermas no siguieron el mismo ciclo de desarrollo en cuanto a pigmentación y contenido hídrico. La combinación de estas técnicas con algoritmos de clasificación, como máquinas de soporte vectorial (SVM, por sus siglas en inglés) (IBM, 2025b) o redes neuronales (OCW, 2008), ha mejorado significativamente la precisión de detección, alcanzando valores superiores al 85 % en algunos ensayos controlados.

El trabajo destaca además la importancia de la reproducibilidad metodológica: se propone un flujo de procesamiento objetivo y estandarizado, replicable en futuras campañas, que incluye desde la adquisición de imágenes hasta la integración de mapas en plataformas de visualización como Mapeo®. Como conclusión, los autores enfatizan que el uso combinado de imágenes térmicas y espectrales, junto con el análisis de evolución a lo largo del tiempo y el empleo de algoritmos de aprendizaje automático, constituye el enfoque más prometedor para el monitoreo a gran escala del picudo rojo.

2.2. Síntesis

En conjunto, los trabajos revisados evidencian que la detección del picudo rojo requiere un abordaje complementario que combine distintas modalidades de sensado y técnicas de análisis. La termografía ha demostrado ser una herramienta eficaz para identificar anomalías de temperatura asociadas al daño interno antes de que aparezcan síntomas visibles, mientras que los índices es-

pectrales permiten capturar variaciones en el contenido hídrico y la actividad fotosintética que también reflejan la presencia de la plaga. A su vez, el uso de teledetección mediante drones permite realizar un seguimiento en el tiempo, enriqueciendo el diagnóstico al observar la evolución de cada ejemplar. Finalmente, las técnicas de aprendizaje automático facilitan la automatización del proceso de detección y mejoran la robustez frente a la variabilidad ambiental y de captura. Esta revisión respalda la pertinencia de la metodología adoptada en el presente proyecto y señala caminos claros de mejora futura, como la integración de datos multiespectrales, el seguimiento temporal y la incorporación de algoritmos más avanzados de visión por computadora.

Capítulo 3

Marco teórico

En esta sección se presentan los principales antecedentes, conceptos y enfoques que permiten articular la dimensión biológica y agronómica de la plaga estudiada con los fundamentos tecnológicos que hacen posible su detección y análisis.

3.1. Fundamentos agronómicos

3.1.1. El picudo rojo de las palmeras

El picudo rojo (*Rhynchophorus ferrugineus*, Olivier, 1790) es un coleóptero de la familia *Dryophthoridae* (véase la Figura 3.1); es considerado la plaga más destructiva de palmeras a nivel mundial. Originario del sudeste asiático y Melanesia, se ha expandido hacia Medio Oriente, el norte de África, Europa y América, provocando graves pérdidas económicas y ecológicas (Faleiro, 2006; Giblin-Davis, 2001). Las hembras depositan sus huevos en fisuras, heridas o axilas foliares del estípite (tronco de la palmera, mencionado anteriormente), de donde emergen larvas sin alas (*ápteras*) que excavan galerías en el tejido interno de la palmera (*parénquima*), rico en reservas y con funciones de almacenamiento. Estas galerías interrumpen la circulación de agua y nutrientes, debilitando progresivamente a la planta. El ciclo larval, que puede durar entre 1 y 3 meses, transcurre completamente en el interior del tronco, lo que hace que la infección sea difícil de detectar en sus primeras etapas. Se señala que la actividad metabólica de las larvas pueden generar un aumento de temperatura detectable en el cogollo de la palmera durante etapas avanzadas de infestación (El-Sabea, Faleiro, y Abo-El-Saad, 2011). Cuando los síntomas externos son evidentes (colapso de la corona, exudación, caída de hojas) el daño interno suele ser irreversible, llevando finalmente a la muerte de la palmera. Esta característica convierte a la detección temprana en un aspecto crítico para la supervivencia del ejemplar y la gestión sanitaria de palmerales (Abbas, 2016).



Figura 3.1: *Rhynchophorus ferrugineus*

3.1.2. Palmeras *Phoenix Canariensis*

La *palmera canaria* (*Phoenix canariensis*, según su nombre científico) es una especie endémica de las Islas Canarias (ver Figura 3.2), ampliamente cultivada en regiones mediterráneas, subtropicales y templadas como especie ornamental. Se distingue por su robustez, longevidad y su porte imponente, con un estípite que puede superar los 15 metros de altura y una corona foliar densa que le

confiere alto valor paisajístico y patrimonial (Rivas-Martínez, 2012). Además de su relevancia ornamental, esta especie cumple funciones ecológicas en entornos urbanos, como la provisión de sombra, la captura de carbono y la conservación de biodiversidad asociada a su hábitat (Martín y Cabello, 2009). Sin embargo, presenta alta susceptibilidad al ataque del picudo rojo, debido a la estructura compacta de su estípite y la abundancia de tejido parenquimático en la corona, que constituye un ambiente favorable para el desarrollo de larvas (Dembilio y Jacas, 2009). Este factor, junto con su amplia distribución en ciudades y espacios públicos, la convierte en una de las especies más afectadas por la plaga, motivo por el cual ha sido objeto de numerosos programas de investigación y control fitosanitario.



Figura 3.2: *Phoenix canariensis*

3.1.3. Importancia de la detección temprana

La detección temprana del picudo rojo es un aspecto crítico en la gestión fitosanitaria de palmeras, ya que los daños ocasionados por la plaga son inicialmente internos y, por tanto, invisibles a simple vista. Cuando los síntomas externos son evidentes (hojas caídas, exudados, colapso de la corona), el daño vascular suele ser irreversible y la probabilidad de recuperación de la palmera es mínima (Abbas, 2016; Hoddle, Hoddle, Faleiro, El-Sabea, y Jeske, 2013). En cambio, la identificación en etapas iniciales permite aplicar medidas de control efectivas como *endoterapia* (método que consiste en inyectar el tratamiento directamente en el tronco del árbol para que se distribuya por su interior y combatir plagas o enfermedades) (Fertinfect, 2025), tratamientos químicos localizados o eliminación selectiva de ejemplares infestados, reduciendo la dispersión del insecto y los costos económicos asociados (Faleiro, 2006). Además, en entornos urbanos, la detección temprana no solo tiene valor económico, sino también social y patrimonial, al preservar el paisaje cultural vinculado a las palmeras.

ornamentales ([Martín y Cabello, 2009](#)).

3.1.4. Técnicas de detección del picudo rojo

La detección del picudo rojo en palmeras ha sido objeto de numerosas investigaciones debido a la dificultad de identificar la plaga en sus fases iniciales, cuando los daños aún no se manifiestan externamente. Con el fin de superar esta limitación, se han desarrollado distintos métodos que abarcan desde técnicas tradicionales de inspección directa hasta tecnologías avanzadas de teledetección y análisis de imágenes (ver Figura 3.3). Estas herramientas permiten aproximarse al estado fisiológico de la planta de manera no invasiva, identificando indicadores tempranos de infestación que facilitan la toma de decisiones en el manejo fitosanitario. A continuación, se presentan las principales técnicas de detección, clasificadas en métodos de contacto directo y métodos de teledetección.



Figura 3.3: Técnicas de detección del picudo rojo

Inspección visual

La inspección visual es el método más extendido y de menor costo, basado en la observación de síntomas externos como marchitez o caída de hojas centrales, exudados marrones y orificios en el estípite ([Faleiro, 2006](#)). Sin embargo, este método es limitado, ya que tales síntomas suelen aparecer en estados avanzados, cuando el daño es irreversible ([El-Sabea y cols., 2011](#)). Por ello, aunque útil como primera aproximación, la inspección visual es insuficiente para la detección temprana.

Detección acústica

La detección acústica se basa en captar los sonidos generados por las larvas al alimentarse dentro del *estípite* (definido anteriormente), mediante micrófonos o sensores piezoelectrónicos ([Mouli Chandra y cols., 2022](#)) insertados en la base

de la palmera. Se han desarrollado algoritmos de procesamiento de señal capaces de diferenciar estos sonidos de otros ruidos ambientales (Mankin, 2011). Si bien es una técnica no invasiva, su eficacia disminuye en ambientes ruidosos o con palmeras de gran tamaño, ya que las señales se atenúan con la distancia (Potamitis, Ganchev, y Kontodimas, 2009).

Sensado multiespectral e hiperespectral

Los sensores *multiespectrales* e *hiperespectrales* permiten identificar cambios fisiológicos relacionados con el estrés hídrico y la actividad fotosintética de las palmeras. Estos cambios se reflejan en variaciones de la reflectancia en el *rango del infrarrojo cercano* (NIR, según sus siglas en inglés) y el infrarrojo de onda corta (SWIR, según sus siglas en inglés) (Bannari y cols., 2017). Índices como NDWI, NDII, SRWI y los más recientes PTWSI (Palm Tree Water Stress Index, por su nomenclatura en inglés) han demostrado correlación con los niveles de infestación, constituyendo un indicador indirecto del estado de salud del árbol incluso antes de la aparición de síntomas visibles (Gao, 1996; L. Wang y cols., 2007). Estas técnicas, aplicadas desde UAV (anteriormente mencionados) o satélites de alta resolución, tienen potencial para el monitoreo de grandes superficies.

Termografía infrarroja

La termografía infrarroja mide las diferencias de temperatura en la superficie de las copas de las palmeras, asociadas al estrés hídrico inducido por el daño vascular que generan las larvas del picudo (Golomb y cols., 2015). Estudios en Egipto e Israel han demostrado que las palmeras infestadas pueden presentar temperaturas foliares superiores a las sanas entre la 7^a y 8^a semana de infestación, cuando aún no existen síntomas externos visibles (Alaa y cols., 2020). Tanto en plataformas terrestres como aéreas, esta técnica ha mostrado alta efectividad para identificar palmeras afectadas de manera no invasiva, aunque factores como el riego desigual, la edad del árbol, o condiciones ambientales pueden generar falsos positivos (Delalieux y cols., 2023).

3.2. Técnicas de tratamiento de las palmeras

El manejo del picudo rojo en palmeras requiere la aplicación de diferentes técnicas de control que buscan frenar la propagación de la plaga y, cuando es posible, salvar a los ejemplares infestados en fases tempranas. Estas prácticas incluyen tratamientos químicos, biológicos, mecánicos y biotecnológicos, cuya elección depende del grado de infestación, la especie afectada y el contexto ambiental en el que se implementan (Faleiro, 2006; Abbas, 2016).

3.2.1. Tratamientos químicos

Los tratamientos químicos han sido los más empleados históricamente y consisten en la aplicación de insecticidas sistémicos o de contacto mediante pulverización, riego al suelo o *endoterapia* (inyecciones en el estípite) (Fertinfect, 2025). Aunque son efectivos en fases iniciales, presentan desventajas como el alto costo, la necesidad de repeticiones frecuentes, y el impacto ambiental y sanitario asociado al uso intensivo de plaguicidas (Dembilio y Jaques, 2012).

3.2.2. Control biológico

El control biológico busca reducir las poblaciones de picudo rojo mediante el uso de organismos antagonistas. Entre ellos destacan los hongos entomopatógenos como *Beauveria bassiana* (Moura Mascarin y Jaronski, 2016) y *Metarhizium anisopliae* (Parjane, Kabre, Mahale, Shejale, y Nirgude, 2020), así como *nemátodos entomopatógenos* del género *Steinernema* (Schwarz y cols., 2025). Estos agentes han mostrado eficacia en condiciones de laboratorio y campo, representando una alternativa sostenible frente al uso de químicos (Gindin, Levski, Glazer, y Soroker, 2006).

3.2.3. Métodos mecánicos y culturales

Incluyen la eliminación y destrucción de palmeras severamente infestadas para evitar la dispersión de la plaga, la poda y eliminación de tejidos afectados en casos leves, y el uso de trampas con feromonas agregativas para capturar adultos. Si bien son técnicas efectivas para reducir la propagación, requieren logística y coordinación a gran escala (Abuagla y Al-Deeb, 2012). Esto último se debe a que el uso de feromonas puede atraer una gran cantidad de individuos adultos desde largas distancias, lo que, sin un adecuado control y monitoreo, puede incrementar el riesgo de infestación en áreas circundantes en lugar de reducirlo.

3.2.4. Métodos biotecnológicos

La investigación reciente ha explorado el uso de *feromonas sintéticas* (Rosas Licona, Lara, Carballo, y Bueno, 2024) y *kairomonas* (Komala, Manda, y Seram, 2021) para atraer y monitorear poblaciones adultas, así como el desarrollo de variedades de palmeras resistentes. Asimismo, se estudian estrategias de control integrado que combinan biotecnología con teledetección y tratamientos localizados, con el objetivo de incrementar la eficacia y reducir los impactos colaterales (Vidyasagar, Aldosari, Sultan, y Al-Dobai, 2016).

3.3. Fundamentos técnicos

Esta sección presenta los conceptos técnicos necesarios para comprender el marco de desarrollo del proyecto. Se abordan nociones clave que, si bien no

constituyen el núcleo central de la investigación, resultan indispensables para contextualizar la metodología empleada y entender las herramientas utilizadas.

3.3.1. Vehículos Aéreos no Tripulados (UAV)

Los *Vehículos Aéreos no Tripulados* (UAV's, según sus siglas en inglés) son aeronaves que operan sin piloto humano a bordo, controladas de manera remota o mediante sistemas autónomos. Suelen formar parte de un sistema mayor denominado *Sistema de Aeronaves No Tripuladas* (UAS, por sus siglas en inglés), que incluye la plataforma aérea, la estación de control en tierra y los enlaces de comunicación necesarios para su operación ([Wikipedia, 2025](#); “[Unmanned Aerial Vehicle \(UAV\)](#)”, 2024).

Estos sistemas han ganado gran relevancia tanto en el ámbito militar como civil, siendo empleados en tareas de vigilancia, monitoreo ambiental, agricultura de precisión, inspección de infraestructuras y respuesta a emergencias, entre otras aplicaciones ([Glosario de tecnología: qué significa vehículo aéreo no tripulado \(UAV\)](#), 2024; [Ramírez y Torres, 2019](#)). Su versatilidad y bajo costo relativo los convierten en una tecnología clave en la evolución de sistemas autónomos aéreos, aunque aún enfrentan desafíos en autonomía energética, seguridad de comunicaciones y regulaciones (“[Unmanned Aerial Vehicles: Opportunities and Challenges](#)”, 2023; [H. Wang y Zhao, 2018](#)).

3.3.2. Inercia térmica

La *inercia térmica* es la resistencia de un objeto a cambiar de temperatura. Una inercia térmica baja implica que un objeto cambia de temperatura rápidamente en respuesta a los cambios en su entorno. Se calienta y enfriá con facilidad. Una inercia térmica alta implica que un objeto cambia de temperatura lentamente; retiene el calor o el frío por más tiempo ([Kočí, Šesták, y Černý, 2022](#)). En el contexto de las palmeras, este concepto es fundamental, ya que el estado biológico de la planta (sana o muerta) altera drásticamente cómo interactúa con su entorno térmico.

Durante el día, las palmeras sanas regulan su temperatura mediante la transpiración (un proceso equivalente a la sudoración en humanos), liberando vapor de agua que genera un efecto de enfriamiento evaporativo. Esto hace que su temperatura se mantenga más baja que la del aire y que la de objetos inertes. Las palmeras muertas, al carecer de este mecanismo, se comportan como cuerpos inertes que absorben radiación solar, alcanzando temperaturas más elevadas.

Durante la noche, las palmeras sanas descienden gradualmente su temperatura hacia el equilibrio con el ambiente, pero retienen calor por su mayor contenido de agua y su actividad metabólica residual, lo que les confiere una inercia térmica más alta. Las palmeras muertas, en cambio, irradian con rapidez el calor acumulado durante el día y su temperatura cae con mayor velocidad. En condiciones de invierno, como en Uruguay, estas diferencias se atenúan durante el día debido a la menor intensidad de la transpiración, aunque se mantienen

evidentes durante la noche, cuando la pérdida de calor de las palmeras muertas es más marcada.

3.4. Visión por computadora

La visión por computadora es un campo multidisciplinario que estudia cómo las máquinas pueden procesar, interpretar y comprender imágenes y videos del mundo real de manera similar —o complementaria— a la percepción humana (Szeliski, 2010). Su objetivo es extraer información significativa de datos visuales y utilizarla en tareas como la clasificación, la detección de objetos, el reconocimiento de patrones o el seguimiento de movimiento.

En el contexto de la agricultura y la sanidad vegetal, la visión por computadora se ha consolidado como una herramienta clave para la detección de plagas, el monitoreo del crecimiento y la evaluación del estado fisiológico de cultivos (Kamilaris y Prenafeta-Boldú, 2018). El avance en hardware de procesamiento gráfico y el desarrollo de algoritmos basados en aprendizaje automático han impulsado la adopción de esta disciplina en escenarios prácticos.

3.4.1. Redes neuronales

Las *Redes Neuronales Artificiales* (ANN, por sus siglas en inglés) (Basheer y Hajmeer, 2000) son modelos computacionales inspirados en la estructura y funcionamiento del sistema nervioso biológico. Están formadas por unidades básicas llamadas neuronas artificiales, organizadas en capas conectadas mediante pesos sinápticos que se ajustan en un proceso de entrenamiento (Haykin, 1998). Cada neurona aplica una función de activación sobre la suma ponderada de sus entradas, lo que permite modelar relaciones no lineales entre los datos. Las arquitecturas más simples, como el *perceptrón multicapa* (MLP, por sus siglas en inglés) (Popescu, Balas, Perescu-Popescu, y Mastorakis, 2009), han demostrado capacidad para resolver problemas de clasificación, regresión y predicción en numerosos campos, incluyendo el procesamiento de imágenes.

En aplicaciones de visión por computadora, las redes neuronales sentaron las bases para modelos más complejos como las *Redes Neuronales Convolucionales* (CNN's, por sus siglas en inglés) (IBM, 2025c), especialmente diseñadas para capturar patrones espaciales en imágenes (LeCun, Bottou, Bengio, y Haffner, 1998).

3.4.2. Aprendizaje profundo

El *aprendizaje profundo* (Deep Learning, según su traducción del inglés) es un subcampo del aprendizaje automático que se caracteriza por el uso de redes neuronales profundas, es decir, con múltiples capas ocultas. Estas arquitecturas permiten aprender representaciones jerárquicas de los datos, donde niveles más bajos capturan características simples (bordes, texturas) y niveles superiores

aprenden patrones más abstractos y complejos (Goodfellow, Bengio, y Courville, 2016).

El éxito del aprendizaje profundo en visión por computadora ha sido potenciado por tres factores principales: (i) la disponibilidad de grandes volúmenes de datos, (ii) los avances en capacidad de cómputo, en particular las *Unidades de Procesamiento Gráfico* (GPU, según sus siglas en inglés) (Misic, Durdevic, y Tomasevic, 2012), y (iii) la mejora en técnicas de optimización y regularización. Modelos como *AlexNet* (Krizhevsky, Sutskever, y Hinton, 2012a), *VGG* (Simonyan y Zisserman, 2015), *ResNet* (He, Zhang, Ren, y Sun, 2016a) y *YOLO* (Sohan, Ram, y Ch, 2024) han marcado hitos en la última década, demostrando un rendimiento sin precedentes en tareas de detección y clasificación de imágenes (Krizhevsky, Sutskever, y Hinton, 2012b; He, Zhang, Ren, y Sun, 2016b).

3.4.3. Aumento de datos

En el contexto del aprendizaje profundo, el *aumento de datos* (Data Augmentation, según su traducción del inglés) es una técnica esencial para mejorar la capacidad de generalización de los modelos, especialmente en escenarios donde la cantidad de datos es limitada. Consiste en la generación artificial de nuevas muestras de entrenamiento mediante la aplicación de transformaciones sobre los datos originales, tales como rotaciones, traslaciones, cambios de escala, recortes, adición de ruido o modificaciones de color en imágenes. De esta manera, se incrementa la diversidad del conjunto de entrenamiento sin necesidad de recolectar datos adicionales, lo que reduce el riesgo de sobreajuste y mejora el rendimiento del modelo en datos no vistos (Shorten y Khoshgoftaar, 2019). Además, su aplicación resulta especialmente relevante en tareas de visión por computadora, donde incluso pequeñas variaciones pueden aportar robustez al modelo frente a cambios en condiciones de iluminación, perspectiva o ruido en la captura (Perez y Wang, 2017).

Las distintas estrategias de aumento de datos pueden clasificarse en dos grandes grupos. Por un lado, las técnicas tradicionales, las basadas en transformaciones geométricas y fotométricas simples tales como rotación, escalado, traslación o cambios de brillo y contraste; éstas son ampliamente utilizadas debido a su bajo costo computacional y facilidad de implementación. Por otro lado, han surgido enfoques más avanzados que aprovechan la generación automática de datos sintéticos mediante *redes generativas adversarias* (Generative Adversarial Networks, según su traducción en inglés) (Amazon Web Services, Inc., 2025) o métodos de búsqueda automática de políticas de aumento, como *AutoAugment* y *RandAugment*, los cuales aprenden combinaciones óptimas de transformaciones para maximizar el desempeño del modelo. Estas técnicas modernas ofrecen mayor flexibilidad y capacidad para producir ejemplos altamente realistas, aunque a costa de una mayor complejidad y demanda computacional (Cubuk, Zoph, Mane, Vasudevan, y Le, 2019; Cubuk, Zoph, Shlens, y Le, 2020; Antoniou, Storkey, y Edwards, 2017).

3.5. Procesamiento digital de imágenes

El procesamiento digital de imágenes se centra en el análisis y la manipulación de imágenes mediante algoritmos, con el objetivo de mejorar su calidad, extraer información relevante o adaptarlas para tareas posteriores de análisis automático (Gonzalez y Woods, 2008). En el ámbito de la visión por computadora, estas técnicas constituyen un paso esencial para preparar los datos de entrada, garantizando que los modelos de detección o clasificación trabajen sobre información estandarizada y sin artefactos que puedan afectar el rendimiento.

3.5.1. Inpainting

El *inpainting* es una técnica de procesamiento de imágenes que busca llenar o reconstruir áreas faltantes, dañadas u ocultas de una imagen de manera coherente con la información circundante (Bertalmío, Sapiro, Caselles, y Ballester, 2000). Su origen está inspirado en el trabajo de restauradores de arte, que reparan secciones deterioradas de pinturas utilizando el contexto visual disponible. En aplicaciones de visión por computadora, el inpainting se emplea para eliminar objetos no deseados, recuperar zonas ocultas o corregir datos faltantes en capturas visuales. En el contexto agrícola, puede utilizarse para restaurar regiones afectadas por ruido en imágenes multiespectrales o para completar áreas con pérdida de información debido a reflejos, sombras, fallas en el sensor o marcas de agua de la cámara (Criminisi, Pérez, y Toyama, 2004). La Figura 3.4 ilustra un caso práctico de aplicación de inpainting. En este ejemplo, una imagen inicial es transformada para llenar el espacio deseado de forma consistente, lo que da como resultado la segunda vista.

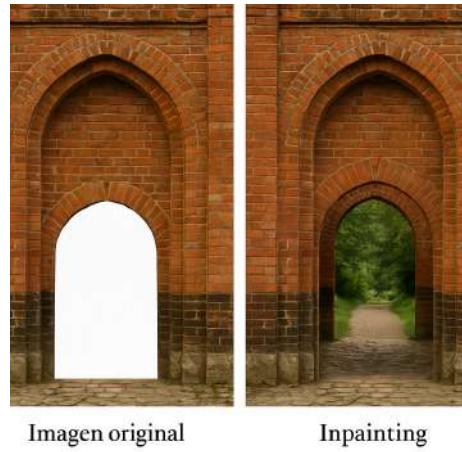


Figura 3.4: Ejemplo de Inpainting

3.5.2. Homografía

Una *homografía* es una transformación proyectiva que establece la correspondencia entre dos planos, permitiendo mapear un punto de una imagen a su ubicación equivalente en otra mediante una matriz de transformación 3×3 (Hartley y Zisserman, 2004). Este concepto constituye la base de numerosos algoritmos de visión por computadora, especialmente en tareas de registro de imágenes, mosaicos y corrección de perspectiva. En aplicaciones prácticas, las homografías permiten alinear imágenes capturadas desde distintos ángulos o en diferentes momentos, facilitando la comparación espacial entre ellas. Esta propiedad ha sido ampliamente utilizada en teledetección y *fotogrametría* (permite obtener información precisa de un objeto o un terreno a partir de fotografías) para la construcción de *ortoimágenes* (corregidas geométricamente) y mapas georreferenciados, donde la corrección de la perspectiva y las distorsiones geométricas resulta esencial (Brown, 1992; Lowe, 2004). La Figura 3.5 ilustra un caso práctico de aplicación de homografía. En este ejemplo, una imagen inicial es transformada para corregir la perspectiva y alinearse con una segunda vista del mismo objeto.

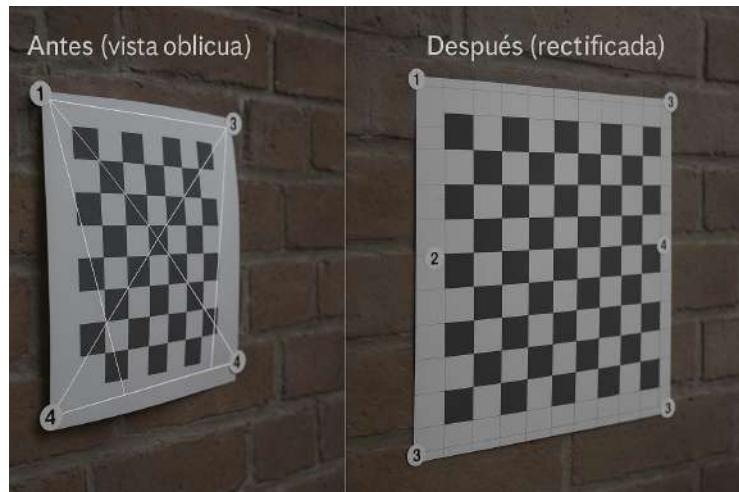


Figura 3.5: Ejemplo de Homografía. En la imagen de la izquierda (Antes) se observan los cuatro puntos numerados (1-4) que marcan las esquinas del tablero, conectados por segmentos que ilustran la relación geométrica original en una vista oblicua. Estos puntos definen las correspondencias utilizadas para calcular la matriz de homografía, la cual transforma el plano inclinado en un plano frontal. En la imagen de la derecha (Después), el patrón aparece rectificado: las líneas se vuelven ortogonales y el tablero recupera su forma cuadrada.

3.5.3. Normalización y recorte

La *normalización* es un procedimiento mediante el cual los valores de intensidad o de color de una imagen se ajustan a un rango común, generalmente con el objetivo de reducir variaciones debidas a cambios de iluminación, exposición o condiciones de captura (Gonzalez y Woods, 2008). Este paso resulta clave para garantizar que las comparaciones entre imágenes o la alimentación a modelos de aprendizaje automático no se vean sesgadas por diferencias de escala en los valores de los píxeles. El recorte, por su parte, consiste en la delimitación de una *región de interés* (Region of Interest, según su traducción en inglés) (Hossain y cols., 2023) dentro de la imagen, descartando la información irrelevante para el análisis. En aplicaciones agrícolas, el recorte puede usarse para enfocar exclusivamente en la copa de la palmera, eliminando el fondo o elementos que no aportan al diagnóstico del estado del árbol. En conjunto, la normalización y el recorte permiten reducir el ruido en los datos y optimizar el rendimiento de los algoritmos de detección y clasificación. La Figura 3.6 ilustra un caso práctico de cada aplicación. En este ejemplo, una imagen inicial es normalizada respecto a los colores en la segunda vista, y posteriormente recortada, teniendo como resultado la tercera vista.

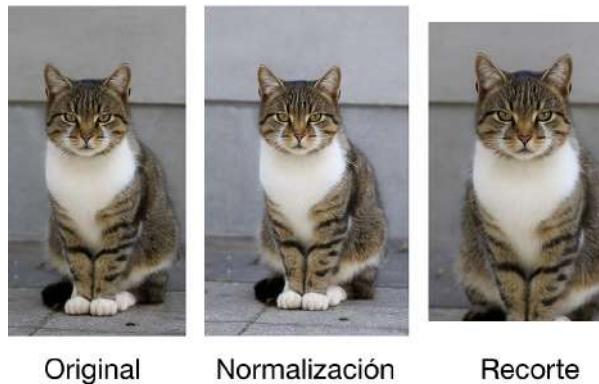


Figura 3.6: Ejemplo de normalización y recorte

3.6. Técnicas de detección de objetos

La detección de objetos es una rama fundamental de la visión por computadora que combina dos tareas: la localización de regiones de interés en una imagen y la clasificación de los objetos contenidos en ellas. A diferencia de la clasificación de imágenes, donde se asigna una etiqueta global, la detección de objetos proporciona información espacial mediante *cuadros delimitadores* (Bounding Boxes, tal como marca su traducción del inglés) (ScienceDirect, 2025). En los

últimos años, los avances en aprendizaje profundo han permitido desarrollar modelos de detección con un rendimiento cercano al humano, capaces de identificar y localizar múltiples objetos en tiempo real. Entre ellos, destacan arquitecturas como R-CNN (*Modelo Faster R-CNN — IBM Visual Insights, 2025*), y YOLO (*Sohan y cols., 2024*), cada una con diferentes enfoques en cuanto a precisión y velocidad.

3.6.1. You Only Look Once

El algoritmo *You Only Look Once* (YOLO, según sus siglas en inglés) introdujo un cambio de paradigma en la detección de objetos al formular el problema como una tarea de regresión única, que predice simultáneamente las clases y las coordenadas de los cuadros delimitadores en una sola pasada por la red neuronal (*Redmon, Divvala, Girshick, y Farhadi, 2016*). A diferencia de métodos anteriores basados en regiones propuestas, como R-CNN, YOLO divide la imagen en una cuadrícula y, para cada celda, predice un número fijo de cuadros y sus probabilidades asociadas. Este enfoque permite alcanzar velocidades de procesamiento en tiempo real, lo que lo convierte en una herramienta ideal para aplicaciones donde la rapidez es un requisito fundamental.

Con el tiempo, se han desarrollado múltiples versiones de YOLO que han mejorado tanto la precisión como la eficiencia. YOLOv2 (*IBM, 2025a*) introdujo mejoras en la resolución de entrada y técnicas como *batch normalization*; YOLOv3 (*Redmon y Farhadi, 2018*) incorporó redes residuales y predicciones multiescala; YOLOv4 (*Bochkovskiy, Wang, y Liao, 2020*) y YOLOv5 (*Jocher y cols., 2022*) optimizaron aún más la arquitectura para alcanzar un balance entre exactitud y velocidad; mientras que versiones recientes como YOLOv7 (*Ultralytics, 2025*) y YOLOv8 (*Jocher, Chaurasia, y Qiu, 2023*) integran módulos de aprendizaje profundo más avanzados y se adaptan mejor a escenarios complejos. Versiones más recientes, como YOLOv9, YOLOv10 y YOLOv11 (*Murat y Kiran, 2025*), continúan la evolución del modelo con innovaciones significativas. YOLOv9, una versión orientada a mitigar la pérdida de información en redes profundas y mejorar la detección de objetos en condiciones adversas. YOLOv10 realiza todo el proceso de detección de objetos dentro de una sola arquitectura unificada en tiempo real con menor carga computacional. Finalmente, YOLOv11 incorpora bloques optimizados junto a la detección de objetos orientados, mejorando la precisión en escenas con objetos rotados o de baja resolución.

En el contexto agrícola, YOLO ha sido empleado con éxito para la detección de plagas, el conteo de frutos y la clasificación de cultivos, mostrando un gran potencial en *Sistemas de Pronóstico Fitosanitarios Automatizados* (*Ministerio de Ganadería, Agricultura y Pesca (MGAP), 2025*) (herramientas de control fitosanitario cuya utilidad se fundamenta en la disminución del riesgo de tomar decisiones poco acertadas técnicamente y no favorables económicamente). Su capacidad para procesar imágenes en tiempo real lo convierte en un candidato especialmente adecuado para la detección temprana de infestaciones en palmeras mediante UAV o cámaras terrestres.

La Figura 3.7 representa esquemáticamente el flujo de procesamiento de YOLO. La imagen de entrada se divide en una cuadricula de celdas, y para cada celda se generan tres tipos de salida: (i) las probabilidades de objeto, que indican el grado de confianza de que exista un objeto en dicha celda; (ii) las probabilidades de clase, que determinan la categoría del objeto detectado; y (iii) los cuadros delimitadores, definidos mediante coordenadas y evaluados con la métrica de solapamiento. Las predicciones redundantes son posteriormente filtradas mediante la técnica de Supresión de No-Máximos (*Non-Maximum Suppression*, NMS, por sus siglas en inglés), lo que permite conservar únicamente los cuadros delimitadores más representativos. El resultado final son los cuadros delimitadores de salida, que identifican en la imagen los objetos detectados junto con su clase y nivel de confianza.

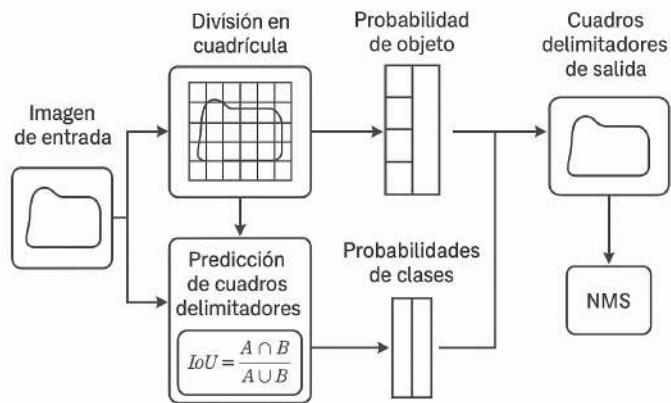


Figura 3.7: Proceso de clasificación de YOLO

3.7. Técnicas clasificadoras

Las redes neuronales convolucionales han demostrado ser especialmente aptas para tareas de clasificación de imágenes debido a su capacidad para aprender jerarquías de características visuales directamente de los datos de entrada. En particular, dos de las arquitecturas con mayor relevancia en el campo son las Redes Residuales (Residual Networks, según sus traducción en inglés) (He y cols., 2016a) y las redes densamente conectadas (Huang, Liu, y Weinberger, 2016) (Densely Connected Networks, según su traducción). Se ha demostrado que el uso de modelos preentrenados permite aprovechar el conocimiento adquirido a partir de millones de imágenes generales, lo que facilita que la red ya disponga de representaciones útiles de texturas, formas y patrones visuales. De esta forma, en lugar de entrenar desde cero, el modelo parte de una base sólida y puede adaptarse más rápidamente al dominio específico.

Además de estos modelos de aprendizaje profundo, existen modelos de aprendizaje automático tradicionales como Random Forest (IBM Cloud Education, 2020) y Support Vector Machine (SVM, según sus siglas). Estos clasificadores se basan en enfoques estadísticos y matemáticos que permiten la identificación de patrones a partir de características previamente definidas, sin requerir grandes volúmenes de datos ni arquitecturas complejas. En el caso de Random Forest, se utiliza un conjunto de árboles de decisión entrenados de manera aleatoria para mejorar la generalización y reducir el sobreajuste, mientras que SVM busca encontrar el hiperplano óptimo que maximiza la separación entre clases en el espacio de características. Estos métodos continúan siendo alternativas eficientes y robustas en contextos donde los conjuntos de datos son limitados o el costo computacional es una restricción relevante.

3.7.1. ResNet18

La arquitectura ResNet18 (Residual Network 18) (He, Zhang, Ren, y Sun, 2016c) es una red neuronal convolucional que introduce el *concepto de conexiones residuales* o “saltos” (skip connections, según su traducción del inglés) (Xu, Wang, Wu, Leng, y Xu, 2025). Estas conexiones permiten que las capas de la red se salten una o más capas, facilitando el entrenamiento de redes muy profundas al mitigar el problema del desvanecimiento del gradiente. Este problema ocurre cuando, durante el proceso de entrenamiento de una red neuronal profunda mediante *retropropagación del error* (backpropagation, según su traducción del inglés) (Sekhar y Meghana, 2020), los gradientes de las capas más profundas tienden a hacerse muy pequeños (se “desvanecen”) al propagarse hacia atrás. ResNet18 tiene un total de 18 capas de aprendizaje y es una opción ligera y eficiente para la clasificación de imágenes.

3.7.2. ResNet50

ResNet50 es una versión más profunda de la arquitectura ResNet, con 50 capas de aprendizaje. A diferencia de ResNet18, utiliza *bloques de cuello de botella* (bottleneck blocks, según su traducción en inglés) (Rehman, 2024) para reducir la complejidad computacional. Esto permite una mayor profundidad con menos parámetros, logrando un equilibrio entre rendimiento y eficiencia. ResNet50 es ampliamente utilizada en la clasificación de imágenes debido a su alto rendimiento (He y cols., 2016c).

3.7.3. DenseNet121

DenseNet121 es una arquitectura que aborda el problema del desvanecimiento del gradiente y la reutilización de características de una manera diferente a ResNet. En DenseNet, cada capa se conecta directamente a todas las capas subsiguientes de forma densa. Esto promueve la propagación de características y gradientes, permitiendo que cada capa tenga acceso a las características de todas las capas anteriores (Huang, Liu, Van Der Maaten, y Weinberger, 2017). Con

121 capas, DenseNet121 es conocida por su eficiencia en el uso de parámetros y su precisión.

3.7.4. Random Forest

El Random Forest (Breiman, 2001) es un algoritmo de aprendizaje supervisado que construye múltiples árboles de decisión durante el entrenamiento y los fusiona para obtener una clasificación más precisa y estable. Este método es robusto frente al sobreajuste y puede manejar un gran número de características. En este proyecto, se utilizó para clasificar las imágenes preprocesadas, aprovechando su capacidad para identificar patrones complejos en los datos.

3.7.5. Support Vector Machine (SVM)

El Support Vector Machine (SVM, según sus siglas en inglés) es un algoritmo de clasificación que busca el hiperplano óptimo que separa las clases en un espacio multidimensional (Cortes y Vapnik, 1995). El objetivo es maximizar el margen entre el hiperplano y los puntos de datos más cercanos de cada clase, conocidos como vectores de soporte. Para este proyecto, se utilizó un kernel (función matemática que permite transformar los datos de entrada a un espacio de mayor dimensión) (Lai, Li, Huang, y Lin, 2024) para manejar la naturaleza de los datos de imagen y evaluar su efectividad en la clasificación de palmeras afectadas.

3.8. Métricas de evaluación de modelos

La evaluación del desempeño de modelos de clasificación y detección requiere el uso de métricas objetivas y estandarizadas. Estas permiten cuantificar la calidad de las predicciones bajo distintos criterios, considerando tanto la exactitud global como el balance entre errores y aciertos en cada clase. A continuación se describen las métricas más utilizadas en aprendizaje automático y visión por computadora, junto con sus variantes más relevantes.

3.8.1. Matriz de confusión

La *matriz de confusión* (Confusion Matrix, según su traducción en inglés) ofrece una representación estructurada de los resultados de clasificación, mostrando cómo se distribuyen las predicciones correctas e incorrectas entre las distintas clases (Erbani, Portier, Egyed-Zsigmond, y Nurbakova, 2024). Cada fila corresponde a las instancias reales y cada columna a las predicciones del modelo, permitiendo identificar errores específicos como falsos positivos y falsos negativos.

La Tabla 3.1 presenta un ejemplo de matriz de confusión para el caso de tres clases: palmeras sanas, afectadas y muertas. Las filas corresponden a las clases reales y las columnas a las clases predichas por el modelo. Se observa que de

las 50 instancias reales de la clase sana, 45 fueron clasificadas correctamente, mientras que 3 fueron asignadas erróneamente a la clase afectada y 2 a la clase muerta. En el caso de la clase afectada, 38 instancias fueron correctamente identificadas, mientras que 4 fueron clasificadas como sanas y 8 como muertas. Finalmente, para la clase muerta, 44 instancias fueron predichas correctamente, 5 se confundieron con la clase afectada y 1 con la clase sana. Este ejemplo evidencia cómo la matriz de confusión permite analizar no solo el desempeño global del modelo, sino también la naturaleza específica de los errores de clasificación entre categorías.

Clase real	Predicha: Sana	Predicha: Afectada	Predicha: Muerta
Sana	45	3	2
Afectada	4	38	8
Muerta	1	5	44

Tabla 3.1: Ejemplo de matriz de confusión para tres clases de palmeras

3.8.2. Precisión

La *precisión* (Precision, según la traducción en inglés) mide la proporción de predicciones positivas que son correctas ([Sokolova y Lapalme, 2009](#)). Se define como:

$$Precision = \frac{TP}{TP + FP}$$

donde TP son los verdaderos positivos y FP los falsos positivos.

3.8.3. Exhastividad

La *exhastividad* (Recall, según la traducción en inglés) mide la capacidad del modelo para identificar correctamente todos los elementos positivos presentes en los datos ([Sokolova y Lapalme, 2009](#)). Se expresa como:

$$Recall = \frac{TP}{TP + FN}$$

donde FN son los falsos negativos.

3.8.4. F1 Score

El *F1 Score* combina precisión y recall en una única métrica armónica, ofreciendo un equilibrio entre ambas ([Powers, 2011](#)). Se define como:

$$F1 = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall}$$

3.8.5. F1 Macro, Micro y Weighted

El F1 Score combina precisión y recall en una única métrica armónica. Sin embargo, en problemas multiclase existen diferentes formas de calcularlo, que responden a cómo se ponderan las clases:

- **F1 Macro:** calcula el F1 de cada clase de manera independiente y luego obtiene el promedio simple. Todas las clases tienen el mismo peso, sin importar cuántas instancias posean. Es especialmente útil cuando se busca evaluar el desempeño en clases minoritarias.

$$F1_{macro} = \frac{1}{N} \sum_{i=1}^N F1_i$$

- **F1 Micro:** agrega los verdaderos positivos, falsos positivos y falsos negativos de todas las clases antes de calcular el F1. De esta manera, las clases con más instancias dominan el resultado.

$$F1_{micro} = \frac{2 \cdot \sum TP}{2 \cdot \sum TP + \sum FP + \sum FN}$$

- **F1 Weighted:** es similar al F1 macro que se definió anteriormente, pero cada clase se pondera según la cantidad de instancias que posee. De esta forma se corrige el sesgo hacia clases minoritarias, manteniendo la influencia proporcional de las clases en el conjunto de datos.

$$F1_{weighted} = \sum_{i=1}^N w_i \cdot F1_i, \quad w_i = \frac{n_i}{\sum_{j=1}^N n_j}$$

donde n_i es el número de instancias de la clase i .

Estas variantes permiten adaptar la evaluación al objetivo del problema: mientras que F1 Macro refleja el equilibrio entre clases, F1 Micro mide el rendimiento global favoreciendo a las clases mayoritarias, y F1 Weighted constituye un punto intermedio al ponderar según la frecuencia de cada clase ([Sokolova y Lapalme, 2009](#); [Powers, 2011](#); [Opitz y Burst, 2019](#)).

3.8.6. Mean Average Precision

La métrica *precisión promedio media* (Mean Average Precision, mAP, según su traducción en inglés) es el estándar en la evaluación de algoritmos de detección de objetos ([Everingham, Van Gool, Williams, Winn, y Zisserman, 2010](#)). Se basa en el cálculo del *Average Precision* (AP), que corresponde al área bajo la curva de precisión frente a la exhaustividad para una clase determinada. El mAP se obtiene promediando los valores de AP de todas las clases:

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i$$

donde N es el número de clases. La definición de un acierto depende del solapamiento entre predicción y anotación real, medido mediante el *Intersection over Union* (IoU):

$$IoU = \frac{|B_{\text{pred}} \cap B_{\text{gt}}|}{|B_{\text{pred}} \cup B_{\text{gt}}|}$$

donde B_{pred} es la caja predicha y B_{gt} la caja de referencia. En la práctica se usan variantes como $mAP@0,5$, que considera un acierto cuando $IoU \geq 0,5$, y $mAP@[0,5 : 0,95]$, que promedia sobre umbrales de IoU en el rango $[0,5, 0,95]$ con incrementos de 0,05.

3.8.7. Exactitud

La *exactitud* (Accuracy, según su traducción en inglés) mide la proporción de predicciones correctas sobre el total de instancias evaluadas (Ferri, Hernández-Orallo, y Modroiu, 2009). Se calcula como:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

3.8.8. Exactitud Balanceada

La *exactitud balanceada* (Balanced Accuracy, según su traducción) corresponde al promedio del recall de cada clase (Brodersen, Ong, Stephan, y Buhmann, 2010). Se calcula como:

$$Balanced\ Accuracy = \frac{1}{C} \sum_{i=1}^C \frac{TP_i}{TP_i + FN_i}$$

donde C es el número de clases.

3.8.9. Coeficiente Kappa de Cohen

El *Coeficiente Kappa de Cohen* (Cohen's Kappa, según la nomenclatura en inglés) mide el grado de acuerdo entre predicciones y etiquetas verdaderas corrigiendo el azar (Cohen, 1960):

$$\kappa = \frac{p_o - p_e}{1 - p_e}$$

donde p_o es el *acuerdo observado* y p_e el *acuerdo esperado*. El acuerdo observado (p_o) corresponde a la proporción de instancias en que el clasificador y la referencia coinciden efectivamente, es decir, la suma de los aciertos en la diagonal principal de la matriz de confusión dividida entre el total de observaciones. En cambio, el acuerdo esperado (p_e) representa la proporción de coincidencias que se obtendrían únicamente por azar, tomando como base las distribuciones marginales de las clases. De esta manera, el coeficiente Kappa ajusta el nivel de

acuerdo real en función de lo que sería esperable por casualidad, ofreciendo una medida más rigurosa que la exactitud simple.

3.8.10. Coeficiente de Correlación de Matthews

El *Coeficiente de Correlación de Matthews* (Matthews Correlation Coefficient, MCC, según su traducción en inglés) evalúa la calidad global de la clasificación binaria o multiclas (Matthews, 1975). Se define como:

$$MCC = \frac{TP \cdot TN - FP \cdot FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}$$

Sus valores varían entre -1 (predicción opuesta) y 1 (predicción perfecta).

3.8.11. Log Loss

La *pérdida logarítmica* (Log Loss, por su traducción en inglés) evalúa el error de las predicciones probabilísticas (Bishop, 2006). Penaliza más los errores de alta confianza y se define como:

$$Log\ Loss = -\frac{1}{N} \sum_{i=1}^N \sum_{c=1}^C y_{i,c} \cdot \log(p_{i,c})$$

donde $y_{i,c}$ es la etiqueta real y $p_{i,c}$ la probabilidad predicha.

3.8.12. Área bajo la curva ROC

El *Área bajo la curva ROC* (ROC-AUC, por sus siglas en inglés) mide la capacidad del modelo para discriminar entre clases (Fawcett, 2006). Existen variantes como:

- *ROC-AUC Macro*: promedio no ponderado por clase.
- *ROC-AUC Weighted*: promedio ponderado por número de instancias.

Valores más cercanos a 1 indican mejor discriminación.

3.8.13. Área bajo la curva Precisión-Recall

El *Área bajo la curva Precisión-Recall* (PR-AUC, según sus siglas en inglés) es especialmente útil en escenarios de desbalance (Saito y Rehmsmeier, 2015). Evalúa la relación entre precisión y recall. Existen variantes Macro y Weighted, según el promedio se haga uniforme o ponderado.

3.8.14. Exactitud Top-k

La *exactitud Top-k* (Top-k Accuracy, según su traducción en inglés) mide si la clase correcta se encuentra dentro de las k predicciones más probables del modelo (Bengio, Weston, y Grangier, 2010).

Matemáticamente, la exactitud Top- k se define como:

$$\text{Top-}k \text{ Accuracy} = \frac{1}{N} \sum_{i=1}^N \mathbb{1}\{y_i \in \hat{Y}_i^{(k)}\} \quad (3.1)$$

donde:

- N es el número total de muestras.
- y_i es la clase verdadera de la muestra i .
- $\hat{Y}_i^{(k)}$ es el conjunto de las k clases con mayor probabilidad predicha por el modelo para la muestra i .
- $\mathbb{1}\{\cdot\}$ es la función indicadora, que vale 1 si la condición es verdadera y 0 en caso contrario.

En el caso particular de la *Top-2 Accuracy*, se toma $k = 2$, lo que mide la fracción de muestras en las que la clase correcta se encuentra entre las dos predicciones más probables.

Capítulo 4

Desarrollo

El desarrollo del proyecto se basó en la construcción de una solución práctica para la detección temprana del picudo rojo en palmeras mediante imágenes térmicas (imágenes obtenidas a partir de la radiación infrarroja emitida por los objetos, que reflejan diferencias de temperatura no visibles al ojo humano) (Fluke Corporation, 2025). Con este propósito, se estableció una metodología que abarca desde la recolección de datos en campo, pasando por el procesamiento de las imágenes obtenidas, hasta su clasificación. Se realizaron capturas con una cámara FLIR T1020 (FLIR Systems, 2025) que genera imágenes térmicas y RGB (Red, Green, Blue según sus siglas) (ScienceDirect Topics, 2025) simultáneamente, y se organizaron los datos con sus respectivas anotaciones. Sobre esta base, se implementaron algoritmos de preprocesamiento, detección y análisis, con el objetivo de evaluar la temperatura como posible indicador de presencia del picudo rojo. Las distintas etapas y los resultados obtenidos durante la experimentación se describen en las secciones siguientes.

4.1. Adquisición de datos

La etapa de adquisición de datos constituye la base del presente proyecto, ya que determina la calidad, representatividad y confiabilidad de la información que será utilizada en el posterior procesamiento y análisis. A continuación, se detallan los distintos aspectos involucrados en esta fase.

4.1.1. Descripción del entorno de captura

La adquisición de datos se llevó a cabo en un entorno urbano, específicamente en la ciudad de Montevideo, capital de Uruguay. Dado que esta ciudad concentra una gran parte de la población de palmeras *Phoenix canariensis* que están siendo afectadas, se seleccionó como área de estudio para garantizar representatividad y maximizar las posibilidades de registrar ejemplares infestados por el picudo rojo.

Las capturas se realizaron en diversas zonas de la capital, que abarcan tanto avenidas principales como espacios públicos y áreas costeras. En la Figura 4.1 se detallan las ubicaciones relevadas. A esto se le asocia la Tabla 4.1 que referencia sus coordenadas geográficas de inicio y fin, permitiendo una referencia espacial precisa de los puntos de muestreo.

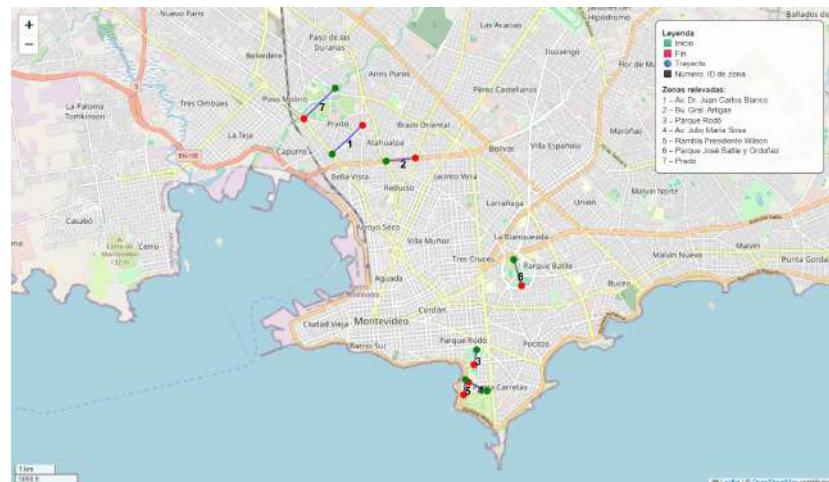


Figura 4.1: Mapa de las zonas relevadas enumeradas según la Tabla 4.1.

Ubicación	Coordenadas de Inicio		Coordenadas de Fin	
	Latitud	Longitud	Latitud	Longitud
Av. Dr. Juan Carlos Blanco	-34.869831	-56.204355	-34.863713	-56.196222
Bv. Gral. Artigas	-34.871371	-56.190139	-34.870788	-56.182322
Parque Rodó	-34.912110	-56.166315	-34.915409	-56.166857
Av. Julio María Sosa	-34.920924	-56.163425	-34.919173	-56.168242
Rambla Presidente Wilson	-34.918540	-56.169090	-34.921856	-56.169605
Parque José Batlle y Ordóñez	-34.892714	-56.156425	-34.898317	-56.154407
Prado	-34.855581	-56.203532	-34.862226	-56.211769

Tabla 4.1: Zonas relevadas y sus coordenadas geográficas

En cuanto a las condiciones ambientales, el relevamiento se desarrolló principalmente durante el período otoño-invierno (desde mayo hasta comienzos de septiembre), lo cual contrasta con los estudios previos en regiones áridas y cálidas del Mediterráneo y Medio Oriente. El clima local presentó características propias de la región: jornadas con cielo nublado, elevada humedad ambiental propia de Montevideo y, en ocasiones, precipitaciones que obligaron a realizar capturas bajo condiciones de lluvia. Esta variabilidad climática constituye un elemento diferenciador frente a los antecedentes internacionales, donde las adquisiciones fueron realizadas en condiciones de calor seco.

Las imágenes fueron tomadas en distintos rangos horarios para evaluar el impacto de la iluminación natural y artificial en la calidad de los datos recolectados. Se definieron dos franjas principales:

- **Entornos diurnos:** entre las 9:00 y las 13:30 horas.
- **Entornos nocturnos:** entre las 18:00 y las 22:00 horas.

La elección de Montevideo como entorno de captura respondió a múltiples factores complementarios. En primer lugar, la ciudad concentra una de las mayores poblaciones de palmeras *Phoenix canariensis* del país, lo que permite contar con un número suficiente de ejemplares para la adquisición de datos y, a su vez, incrementa la probabilidad de encontrar casos efectivos de infestación por el picudo rojo. En segundo lugar, el acceso directo a ejemplares urbanos infestados constituye una buena oportunidad de análisis, ya que en contextos rurales o productivos el monitoreo suele estar en un estado borde, siendo los casos de estudio de palmeras muertas o sanas.

Asimismo, la localización en la capital ofreció ventajas logísticas significativas, como la proximidad geográfica para los traslados del equipo, la disponibilidad por un tiempo acotado de la cámara, y la posibilidad de realizar capturas en horarios variados sin requerir grandes desplazamientos. Estos aspectos facilitaron un relevamiento sistemático y sostenido en el tiempo, condición fundamental para construir un conjunto de datos representativo.

Finalmente, la elección del entorno urbano tiene un valor práctico y social relevante. Montevideo es actualmente el foco principal de la infestación en Uruguay, por lo que cualquier avance en la detección temprana tiene un impacto directo en la preservación del arbolado urbano y en la protección del patrimonio

paisajístico de la ciudad. De esta manera, la adquisición de datos en estas zonas no solo responde a criterios técnicos y logísticos, sino también a la urgencia de aportar soluciones aplicables a la realidad local, en consonancia con los esfuerzos de gestión y control impulsados por la Intendencia de Montevideo.

La Figura 4.2 muestra diversos ejemplos de imágenes capturadas en condiciones de iluminación diurna y nocturna.

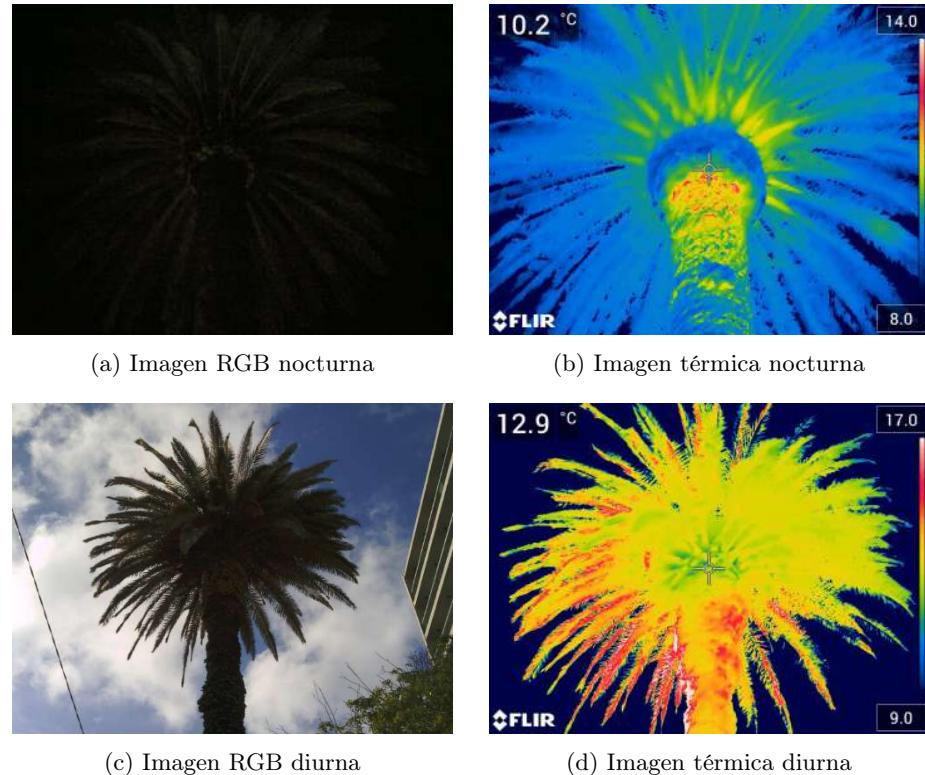


Figura 4.2: Comparación de imágenes RGB y térmicas capturadas de día y de noche.

4.1.2. Sesiones y protocolos de captura

Las capturas se llevaron a cabo en el transcurso de nueve sesiones, realizadas en diferentes franjas horarias, tanto en entornos diurnos como nocturnos, con el propósito de analizar las variaciones en la calidad de los datos en función de las condiciones de iluminación y temperatura. En la Tabla 4.2 se presenta un esquema diseñado para registrar las sesiones de adquisición, el cual incluye la fecha, la cantidad de imágenes obtenidas y la franja horaria de captura.

Fecha	Cantidad de fotos	Comienzo
2025-05-29	158	17:20
2025-06-01	100	17:40
2025-06-05	124	18:10
2025-06-06	132	16:35
2025-06-08	122	10:21
2025-08-07	134	19:17
2025-08-11	148	19:00
2025-08-28	158	18:15
2025-08-29	208	19:20

Tabla 4.2: Registro de sesiones de captura de imágenes

La etapa de captura de imágenes se realizó siguiendo un protocolo estandarizado que buscó asegurar la coherencia, calidad y representatividad de los datos obtenidos. Todas las adquisiciones fueron efectuadas desde el suelo a nivel peatonal, posicionando la cámara a una distancia aproximada de entre 5 y 8 metros de la palmera. En cada ejemplar se tomó una única fotografía compuesta (RGB y térmica), procurando seleccionar el ángulo más consistente y con menor ruido visual, de forma de maximizar la visibilidad del cogollo, las hojas principales y parte del tronco.

Durante esta etapa de recolección se implementó, además, un registro paralelo y riguroso que permitió contextualizar cada adquisición y asegurar la trazabilidad de los datos. Para cada sesión se documentaron los siguientes parámetros:

- Día de la toma.
- Hora de inicio de la sesión.
- Temperatura ambiente al momento de cada captura.
- Sensación térmica estimada.
- Ubicación geográfica precisa de las capturas.
- Rango fijo de temperaturas configurado en la cámara térmica (escala).
- Identificador único asignado a cada imagen capturada.
- Clasificación visual preliminar del estado de la palmera (sana, afectada o muerta).

Este procedimiento sistemático permitió no solo estandarizar la metodología de adquisición, sino también asegurar que cada imagen estuviera acompañada de metadatos confiables que posteriormente facilitaron las etapas de preprocesamiento, análisis y validación de resultados.

El objetivo de este protocolo fue garantizar que la zona del cogollo, de particular interés por ser la región más afectada por el picudo rojo, quedara representada con la mayor claridad posible en cada toma. Para ello, se evitó capturar imágenes en las que las palmeras estuvieran completamente obstruidas por construcciones, mobiliario urbano u otras palmeras. Asimismo, se controló que fuentes de luz intensas (tanto naturales como artificiales) no incidieran directamente sobre la lente, a fin de reducir distorsiones en la información térmica.

4.1.3. Especies y ejemplares estudiados

La especie analizada en este estudio corresponde a la palmera canaria, *Phoenix canariensis*, ampliamente utilizada como elemento ornamental en espacios urbanos de Montevideo. Tal como se detalla en el marco teórico, esta especie se caracteriza por un tronco robusto y columnar que puede alcanzar más de 15 metros de altura, con una copa formada por largas hojas pinnadas de hasta 5 metros de longitud. Su valor paisajístico y su presencia masiva en avenidas, parques y la rambla costera explican la importancia de preservarla frente a la amenaza que supone la infestación por el picudo rojo.

La selección de ejemplares se realizó de forma planificada en zonas con alta concentración de palmeras, priorizando la diversidad de situaciones observadas en campo. Para esta tarea resultó fundamental la colaboración con la Intendencia de Montevideo, que proporcionó un mapa georreferenciado de las palmeras

de la ciudad, actualizado periódicamente con información sobre el estado sanitario de cada ejemplar¹. Este recurso permitió identificar con mayor precisión la ubicación de ejemplares sanos, afectados y muertos, facilitando la planificación de las salidas de campo y garantizando que la muestra recogida fuera lo más representativa de la situación real en el arbolado urbano de la ciudad. A pesar de poseer esta información, es relevante considerar el plazo de actualización que posee, ya que se estima una actualización por zona de alrededor de dos meses, lo que conllevó a tener que utilizar en muchas situaciones la percepción visual de la palmera para predecir su estado a la hora de tomar la foto. Por otro lado, el equipo técnico de la Intendencia ha informado que las campañas de monitoreo incluyen la vacunación sistemática de todas las palmeras recuperables, siguiendo un esquema que comienza con una primera dosis, una segunda a los 30 días y aplicaciones bimensuales posteriores, esto es llevado a cabo en distintas etapas del año según la zona.

Cabe señalar que, para los fines de este proyecto, se optó por simplificar la clasificación de los ejemplares en tres categorías principales: sanas, afectadas y muertas. Esta decisión respondió tanto a la necesidad de reducir la complejidad del problema de clasificación, como al hecho de que las diferencias entre algunos estados intermedios resultan difíciles de discernir sin un entrenamiento especializado en diagnóstico fitosanitario. En contraste, la Intendencia de Montevideo emplea un esquema de categorización más detallado que incluye estados como: palmera retirada, tratamiento, muerta por picudo, síntomas drásticos, síntomas claros, síntomas leves, para retirar, sin copa, seca por otros agentes, cirugía y otros/sin datos. Si bien esta clasificación ofrece un nivel de granularidad mayor, para el presente trabajo se consideró más apropiado comenzar con una segmentación más general, que permitirá validar la metodología y sentar bases sólidas para futuras investigaciones con categorías más específicas.

La progresión del daño ocasionado por el picudo rojo en las palmeras sigue un patrón relativamente predecible, aunque de rápida evolución. Se ha identificado que el ciclo desde la aparición de síntomas hasta la muerte de la palmera puede completarse en un plazo de cuatro a seis meses, con una aceleración marcada durante el verano. En algunos casos, se ha observado que una palmera puede morir en tan solo 20 días al finalizar la estación cálida. Resulta clave intervenir en la etapa de síntomas leves, ya que es el único estado en el cual se ha logrado revertir la afección mediante tratamiento. Las palmeras que alcanzan síntomas claros pierden la capacidad de conducir los productos necesarios para el tratamiento, por lo que la endoterapia deja de ser eficaz.

En este contexto, la detección temprana adquiere carácter fundamental. Las experiencias en zonas donde se ha implementado un plan preventivo en etapas tempranas ha sido mediante endoterapia, demostrando que las pérdidas pueden reducirse hasta tres veces en comparación con áreas sin intervención, donde la tasa de pérdida ha alcanzado el 25 %. Sin embargo, en estados más avanzados, la cirugía vegetal es el único recurso disponible, aunque su alto costo y

¹https://www.google.com/maps/d/u/0/viewer?mid=1yCQ986yfEy6SmUXEREmedCEptZEA_h0&l1=-34.82702551894993%2C-56.1596163240187&z=12

complejidad logística limitan su aplicación. Se estima que una cirugía cuesta aproximadamente la mitad que el retiro de una palmera muerta, siendo esta última la opción más costosa.

En cuanto a los síntomas visibles, los especialistas han identificado como señales de alarma la caída de hojas, resecamiento, pérdida de simetría en la copa y la muerte progresiva de hojas desde la base hacia la parte superior. Si bien existen hongos que pueden generar síntomas similares, el diagnóstico visual resulta menos preciso. La palmera canaria (*Phoenix canariensis*) es la especie más susceptible, y es por ello que constituye el foco principal de este estudio. Aunque el Picudo Rojo no afecta otras especies arbóreas, se ha advertido su capacidad de adaptación, en particular en casos confirmados de especies como la Butiá (Butiá Odorata, según su nombre científico) ([Equitec Tecnologías y Control, 2025](#)).

Desde el punto de vista térmico, se ha confirmado por profesionales de la Intendencia de Montevideo que existe una diferencia en la retención de temperatura entre palmeras vivas y muertas. Las palmeras vivas, debido a su contenido de agua, presentan una inercia térmica mayor, es decir, no cambian de temperatura tan rápido como sus contrapartes muertas. En particular, se ha observado que los ejemplares muertos tienden a enfriarse más rápidamente, lo cual respalda el uso de imágenes térmicas como herramienta de diagnóstico en esta línea de trabajo.

Por tanto, el uso de imágenes térmicas como soporte para la detección automatizada temprana representa una herramienta valiosa, con potencial para identificar ejemplares afectados en etapas reversibles. Este tipo de solución resulta particularmente útil para especies emblemáticas que se busca preservar. Dado que las campañas de endoterapia no alcanzan una cobertura completa, una herramienta que contribuya a priorizar palmeras en riesgo tendría un impacto significativo en la gestión urbana de esta problemática.

4.1.4. Instrumentación utilizada

Para la adquisición de imágenes térmicas y RGB de las palmeras, se utilizó la cámara **FLIR T1020**, un dispositivo de alta gama diseñado para inspección termográfica profesional. Esta cámara cuenta con un detector de resolución nativa de 1024×768 píxeles. Cuando las imágenes se procesan con el modo *UltraMax* de FLIR —un esquema de fusión de múltiples fotogramas con desplazamientos subpíxel— se obtiene un archivo con mayor resolución y nitidez percibida. Este incremento es por software (post-procesado) y no modifica el tamaño real del detector, que permanece en 1024×768 ([FLIR Systems, 2020a, 2025](#)). Además, permite capturar de manera simultánea imágenes en espectro visible con una resolución de 1280×960 píxeles, lo cual resulta fundamental para disponer de un par RGB-térmico en cada toma.

La FLIR T1020 opera en un rangopectral de $[7,5 - 14] \mu\text{m}$, con una sensibilidad térmica (NETD) inferior a 20 mK a $+30^\circ\text{C}$, lo que le permite detectar diferencias mínimas de temperatura entre zonas de la misma palmera. La precisión térmica declarada es de $\pm 1^\circ\text{C}$ o $\pm 1\%$ dentro del rango de 5°C a 150°C ,

lo que garantiza la confiabilidad de las mediciones (FLIR Systems, 2020b).

Adicionalmente, se empleó un **termómetro FLUKE 971** para registrar la temperatura ambiente en cada sesión de captura. Este dispositivo portátil permite medir tanto temperatura como humedad relativa, brindando un contexto ambiental necesario para la interpretación de las imágenes térmicas (Fluke Corporation, 2016).

Las imágenes RGB y térmicas fueron almacenadas de forma conjunta en cada captura, utilizando una tarjeta micro-SD. Todas las adquisiciones se realizaron de manera terrestre, manipulando manualmente la cámara, ajustando el ángulo y la distancia según cada ejemplar. No se emplearon UAV, priorizando la precisión y control directo de las capturas en entornos urbanos.

Con el fin de ilustrar los equipos empleados durante la etapa de adquisición de datos, en la Figura 4.3 se presentan las imágenes oficiales de los dos instrumentos principales utilizados en este proyecto.



Figura 4.3: Instrumentos utilizados en la adquisición de datos: (izquierda) cámara FLIR T1020 (fuente: Teledyne FLIR), (derecha) termómetro FLUKE 971 (fuente: Fluke).

4.1.5. Estructura del dataset

El conjunto de datos resultante está compuesto por imágenes térmicas y RGB capturadas de cada ejemplar, acompañadas de un registro sistemático de metadatos. El conjunto de datos está representado por un archivo CSV, donde cada entrada del mismo contiene tanto la ubicación de la imagen a procesar como los datos visuales de la palmera (variables asociadas que permiten contextualizar y analizar los datos de forma estandarizada, descritas debajo).

La Tabla 4.3 resume las variables almacenadas en el conjunto de datos, las cuales fueron definidas con el objetivo de facilitar posteriores procesos de análisis, anotación y clasificación automática.

Variable	Descripción
<code>image_path</code>	Ruta del archivo de imagen térmica.
<code>min_t</code>	Temperatura mínima registrada en la escala (°C).
<code>max_t</code>	Temperatura máxima registrada en la escala (°C).
<code>cls</code>	Clasificación de la palmera: <i>sana</i> , <i>afectada</i> o <i>muerta</i> .
<code>ambient</code>	Temperatura ambiente al momento de la captura (°C).

Tabla 4.3: Descripción de las variables del dataset de imágenes térmicas

Para ilustrar cómo se registran los datos, la Tabla 4.4 presenta ejemplos concretos de tres capturas.

<code>image_path</code>	<code>min_t</code>	<code>max_t</code>	<code>cls</code>	<code>ambient</code>
IR_1999-12-31_1323.jpg	8.0	14.0	sana	11.0
IR_1999-12-31_1543.jpg	9.0	16.0	afectada	10.0
IR_1999-12-31_1673.jpg	8.5.0	15.0	muerta	9.0

Tabla 4.4: Ejemplo de registros individuales dentro del dataset

4.1.6. Preprocesamiento inicial

El preprocesamiento inicial del dataset tuvo como objetivo asegurar la calidad y consistencia de los datos, antes de su utilización en algoritmos de visión por computadora. Esta etapa se estructuró en los siguientes pasos:

1. Identificación de problemas en las capturas. Durante la recolección se detectaron dificultades frecuentes que afectaron la calidad de ciertas imágenes, como iluminación desigual en entornos urbanos, presencia de ruido (aves u objetos en movimiento), obstrucciones parciales de las palmeras, desenfoques y variaciones climáticas (lluvia o alta humedad). Estas imágenes fueron etiquetadas para su eventual descarte o tratamiento específico.

2. Separación por condiciones de captura. Dado que las capturas incluían tanto sesiones diurnas como nocturnas, se decidió dividir el dataset en dos subconjuntos diferenciados. Esta decisión responde a la alta varianza visual que introducen las distintas condiciones de iluminación, lo que puede dificultar el entrenamiento de modelos de aprendizaje automático.

3. Evaluación de ventajas y desventajas. Las capturas nocturnas mostraron una ventaja clara en la información térmica, al eliminar el efecto del calentamiento solar desigual. Sin embargo, las imágenes RGB nocturnas presentaron baja calidad visual. Por el contrario, las capturas diurnas ofrecen imágenes RGB nítidas, pero las imágenes térmicas pueden verse afectadas por zonas de calentamiento irregular. La Tabla 4.5 muestra esta dualidad, planteando la

necesidad de elegir cuidadosamente qué subconjuntos poseen información más valiosa según el tipo de análisis.

Aspecto	Capturas diurnas	Capturas nocturnas
Calidad de imágenes RGB	Alta nitidez y buena iluminación natural, lo que facilita el análisis visual.	Baja calidad por falta de luz ambiental, colores poco representativos y ruido elevado.
Calidad de imágenes térmicas	Afectadas por radiación solar, con riesgo de zonas de calentamiento espurias en el follaje y el tronco.	Mayor uniformidad térmica al no existir incidencia solar directa, lo que mejora la precisión en la detección de anomalías.
Condiciones ambientales	Influencia de factores externos como nubosidad, ángulos de incidencia del sol o reflejos.	Menor interferencia ambiental, aunque dependiente de condiciones como humedad o temperatura nocturna.
Utilidad para modelos de Machine Learning	Adecuadas para algoritmos que dependen de información visual (RGB).	Adecuadas para algoritmos basados en análisis térmico.

Tabla 4.5: Comparación entre capturas diurnas y nocturnas en el dataset

4. Resultados del filtrado. Tras el proceso de selección y limpieza, el dataset quedó conformado por 489 imágenes térmicas utilizables, compuestas por 78 capturas diurnas y 411 nocturnas. Tal y como se mencionó en el punto 2 anterior, debido a la gran diferencia que presentan las capturas nocturnas y diurnas, se decidió dividir el dataset original en un conjunto de datos diurno y otro nocturno. El dataset diurno presenta un total de 51 sanas, 14 afectadas y 13 muertas mientras que el conjunto de datos nocturno consta de 108 sanas, 137 afectadas y 166 muertas.

5. Generación dinámica de subconjuntos. Con el fin de facilitar la experimentación, se desarrolló un script denominado `generar_datasets.py`, que permite crear de forma dinámica nuevos conjuntos de datos a partir de los conjuntos de datos generales. La herramienta ofrece flexibilidad para seleccionar combinaciones de clases específicas sin necesidad replicar información y automatizando un proceso manual el cual es tedioso y propenso a errores. La Figura 4.4 presenta el esquema general seguido para la organización y generación de los diferentes subconjuntos.

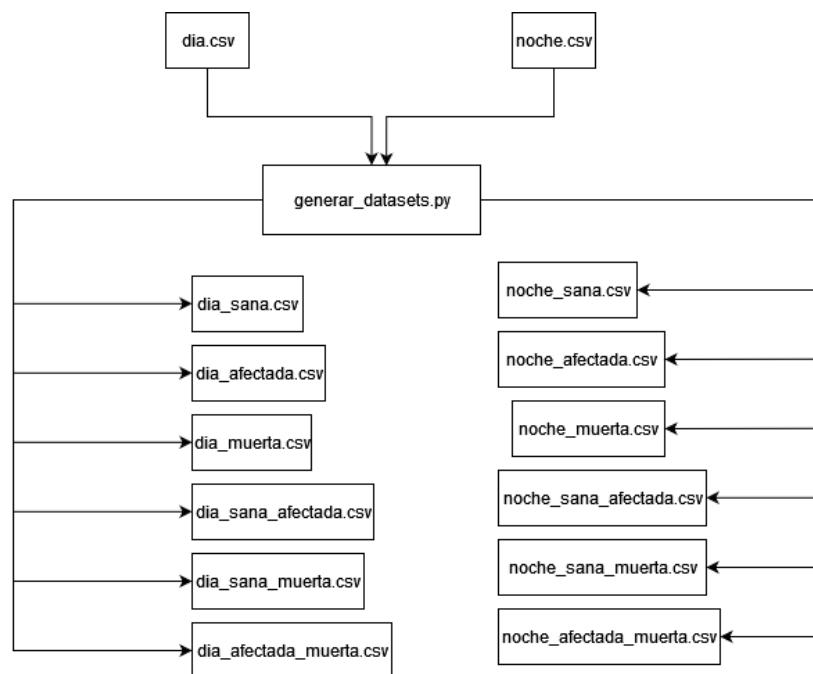


Figura 4.4: Diagrama de la estructura de los datasets

4.2. Preprocesamiento

El preprocesamiento constituye una etapa fundamental en el flujo de trabajo, ya que asegura la calidad, homogeneidad y usabilidad de las imágenes antes de aplicar algoritmos de detección y clasificación. En esta sección se detallan los pasos realizados, desde la estandarización hasta la preparación de subconjuntos de datos para su experimentación.

La Figura 4.5 presenta una visión general del flujo de procesamiento y clasificación propuesto en el trabajo a partir de imágenes RGB y térmicas. El sistema se compone de una etapa inicial de preprocesamiento, donde se recortan las imágenes térmicas para eliminar elementos añadidos por la cámara, se convierten las imágenes RGB a escala de grises, y se alinea la imagen RGB con su correspondiente térmica mediante una transformación homográfica. A partir de la imagen térmica, se utiliza un modelo YOLO (Sohan y cols., 2024) entrenado para detectar las regiones de la palmera (Cogollo, hojas, y tronco), donde se espera observar las primeras señales de infestación. Luego, sobre la imagen térmica en escala de grises, se calcula la temperatura media de dicha región y se la entrega al módulo clasificador, el cual determina el estado de la palmera (sana, afectada o muerta).

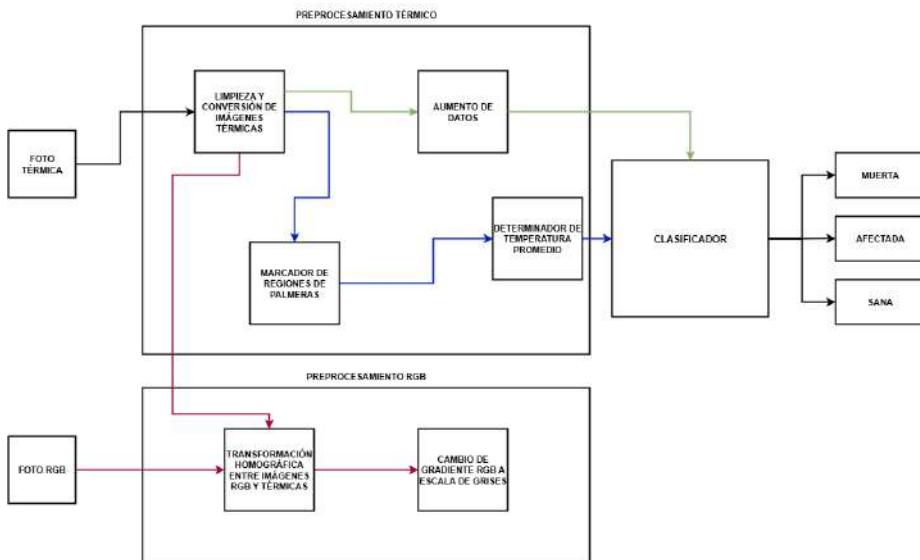


Figura 4.5: Diagrama de la estructura del algoritmo completo y sus diferentes flujos de trabajo

Por otra parte de lo mencionado, se ilustran los tres flujos de preprocesamiento implementados en el proyecto.

El primer flujo (línea azul) constituye el procedimiento original y principal. Comienza con la limpieza y conversión de las imágenes térmicas, para luego aplicar un modelo YOLO encargado de identificar y marcar las regiones de interés correspondientes de la palmera. A partir de esta región se calcula la temperatura promedio, y con los atributos visuales de la imagen, se utilizan en conjunto como entrada para un modelo de clasificación. Este enfoque integra la información estadística de las imágenes térmicas con características visuales extraídas de la región detectada, representando la estrategia base del sistema desarrollado.

El segundo flujo (línea roja) corresponde a una línea experimental que finalmente fue descartada por no aportar resultados de utilidad. En este caso, además de la limpieza y conversión de las imágenes térmicas, se incorporaba una transformación homográfica entre imágenes RGB y térmicas, con el objetivo de lograr una alineación espacial entre ambas modalidades. Posteriormente, las imágenes RGB eran convertidas a escala de grises para su procesamiento. Sin embargo, las limitaciones prácticas asociadas al uso de imágenes RGB en condiciones urbanas y de baja iluminación redujeron su aplicabilidad, por lo que el enfoque quedó deprecado.

El tercer flujo (línea verde) plantea una variante basada exclusivamente en redes neuronales convolucionales (CNN). Tras la limpieza y conversión de las imágenes térmicas, se aplica un esquema de aumento de datos que busca incrementar la variabilidad del conjunto de entrenamiento y mejorar la capacidad de generalización del modelo. Estas imágenes procesadas se utilizan directamente como entrada al clasificador CNN, sin cálculo intermedio de estadísticas térmicas. Este flujo representa un acercamiento alternativo más dependiente del aprendizaje profundo y complementa la estrategia inicial.

4.2.1. Limpieza y conversión de imágenes térmicas

Las imágenes térmicas capturadas con la cámara FLIR T1020 contienen diversos elementos superpuestos que, si bien son útiles para la interpretación visual humana, resultan problemáticos en el análisis computacional. Entre ellos se encuentran: (i) el puntero central con su respectiva lectura de temperatura en la esquina superior izquierda, (ii) la barra de colores vertical ubicada a la derecha junto con las temperaturas mínima (iii) y máxima (iv) de la escala, y (v) la marca de agua de la cámara en la esquina inferior izquierda. Cada uno de estos elementos puede observarse en la Figura 4.6.

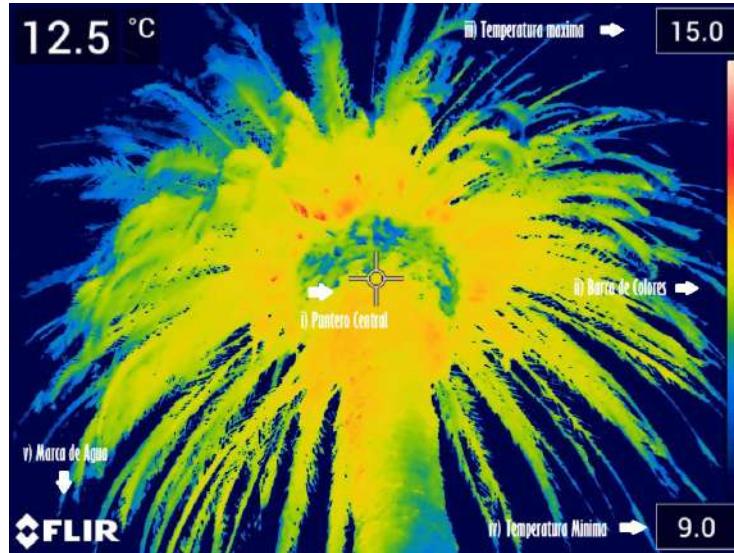


Figura 4.6: Imagen capturada con la cámara FLIR

Asimismo, la imagen se muestra utilizando una paleta de falso color perceptual, en la cual los tonos representan una temperatura que depende de la escala establecida. Esto implica que dos imágenes con diferentes escalas, tendrán píxeles que reflejan temperaturas distintas. Por otra parte, todo elemento incorporado en la imagen por parte de la cámara introducen píxeles ajenos a la distribución térmica real de la palmera, por lo que su eliminación resulta indispensable.

Para abordar esta problemática se diseñó un algoritmo en Python que integra tanto la limpieza de artefactos como la conversión de la imagen térmica a un mapa de temperaturas en escala de grises. La detección del puntero se realiza mediante una máscara fija adaptada a las orientaciones posibles de la cámara (vertical u horizontal), sobre la cual se aplica una técnica de *inpainting biharmonic* de la biblioteca `scikit-image`, que reconstruye la zona a partir de la información circundante. La elección de dicha biblioteca se debe a que es de código abierto, proporciona un conjunto de herramientas amplio, y es eficiente para el procesamiento de imágenes en Python. A su vez, se aplica un recorte sistemático con parámetros predefinidos, eliminando márgenes, la barra de colores y demás superposiciones gráficas, de modo que solo se conserva la región central correspondiente a la escena térmica. Como resultado, la resolución de las imágenes pasa de 1024×768 píxeles a 992×600 , reduciendo el tamaño de los datos y manteniendo la región de interés.

El segundo paso consiste en la conversión del mapa térmico en color a un mapa numérico normalizado en escala de grises global para todas las imágenes. Para ello, el algoritmo extrae la barra de color incluida en la imagen y asocia cada valor cromático a un valor real de temperatura en función de los límites

mínimo y máximo registrados en la captura. Esta asociación se implementa mediante un árbol de búsqueda `KDTree` de la biblioteca `scikit-learn`, que permite identificar de manera eficiente el color más cercano para cada píxel de la imagen y mapearlo a su temperatura correspondiente. Posteriormente, los valores de temperatura son normalizados en un rango de 0 a 255 y almacenados como imágenes de un solo canal, listas para su procesamiento posterior. De esta manera, cada nivel de gris representa un incremento aproximado de $0,1\text{ }^{\circ}\text{C}$ en la escala térmica, de modo que el valor numérico de cada píxel puede interpretarse directamente como una medida relativa de temperatura.

El proceso se apoyó en un conjunto de bibliotecas de código abierto ampliamente utilizadas en visión por computadora y procesamiento de imágenes: `NumPy` para la manipulación matricial, `Pillow` para la carga y guardado de imágenes, `Matplotlib` para visualización, `scikit-learn` para la construcción del `KDTree`, y `scikit-image` para las operaciones de *inpainting*. El resultado final de este proceso es un conjunto de imágenes limpias y estandarizadas en escala de grises que representan de manera fiel el mapa térmico de cada palmera, sin elementos ajenos y con dimensiones homogéneas, lo que asegura su consistencia para las etapas de alineación, detección y clasificación. La Figura 4.7 muestra ejemplos visuales del proceso desde el estado base de la imagen hasta la conversión y recorte de la misma.

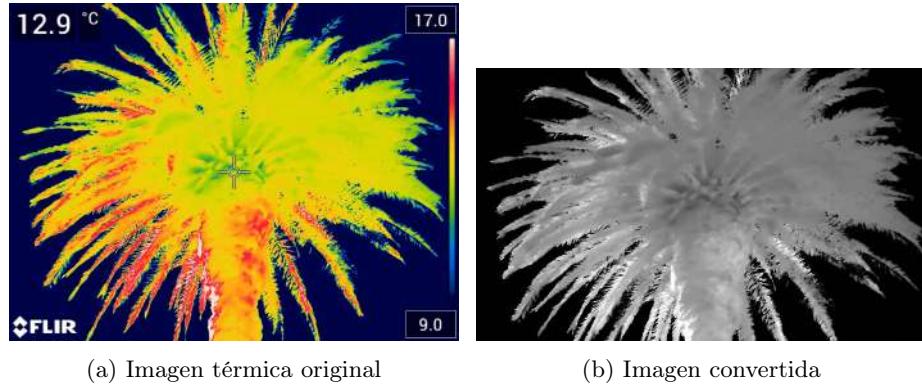


Figura 4.7: Conversión y limpieza de imagen térmica.

4.2.2. Transformación homográfica entre imágenes RGB y térmicas

Este proceso requiere como insumo las imágenes térmicas previamente procesadas en la etapa de limpieza y conversión, ya que la presencia de artefactos gráficos como el puntero, la barra de colores o la marca de agua impediría una alineación precisa. Una vez obtenidas las imágenes térmicas normalizadas en escala de grises, se procedió a realizar la alineación geométrica con las imágenes RGB correspondientes.

Aunque la cámara FLIR T1020 captura simultáneamente imágenes RGB y térmicas de un mismo ejemplar, ambas modalidades provienen de sensores diferentes y, en consecuencia, no se encuentran alineadas píxel a píxel. Esta desalineación impide comparar directamente regiones de interés entre ambas imágenes, por lo que resulta necesario aplicar una transformación geométrica que permita superponerlas de manera coherente. Para este fin se utilizó una homografía, que modela la relación entre dos planos proyectivos mediante una matriz de transformación 3×3 .

El procedimiento comenzó con la selección manual de un conjunto de puntos de correspondencia entre la imagen RGB y su contraparte térmica, identificando esquinas y bordes característicos de la palmera. Estos puntos se emplearon para calcular una matriz homográfica inicial H utilizando las funciones de la biblioteca `OpenCV`. Sin embargo, debido a la naturaleza aproximada de la selección manual, se diseñó un proceso de optimización para refinar la transformación.

La optimización consistió en minimizar una función de costo basada en la diferencia absoluta de las derivadas verticales entre la imagen térmica y la RGB transformada. Este criterio permitió alinear las estructuras internas de ambas modalidades, priorizando la coincidencia de las formas de la palmera en lugar de los valores absolutos de intensidad. La optimización fue implementada con el método `Powell`, disponible en la biblioteca `SciPy`, lo que dio lugar a una matriz homográfica optimizada H_{opt} .

Una vez obtenida la matriz H_{opt} , se automatizó su aplicación en lote mediante un proceso que transforma todas las imágenes RGB de una carpeta de entrada, generando versiones alineadas con sus correspondientes térmicas. De esta manera, cada par de imágenes quedó espacialmente registrado, lo que permite utilizar directamente las detecciones realizadas en la modalidad RGB para extraer valores térmicos en las mismas regiones de interés. Esta capacidad es fundamental para el análisis posterior del cogollo de la palmera, ya que garantiza que las estadísticas térmicas se calculen exactamente sobre la región previamente identificada por el modelo de detección.

La Figura 4.8 muestra ejemplos visuales del proceso de alineación, incluyendo la imagen RGB original, la imagen térmica, la imagen RGB transformada mediante la homografía y una visualización de las diferencias entre ambas modalidades.

Esta última imagen combina las diferencias de intensidad o detalle entre la imagen RGB transformada y térmica y las codifica con colores para hacer visible el error. El color rojo representa áreas donde la imagen RGB tiene mayor relevancia, el color verde indica donde es más relevante la imagen térmica en escala de grises y el color negro indica áreas donde la alineación es muy buena.

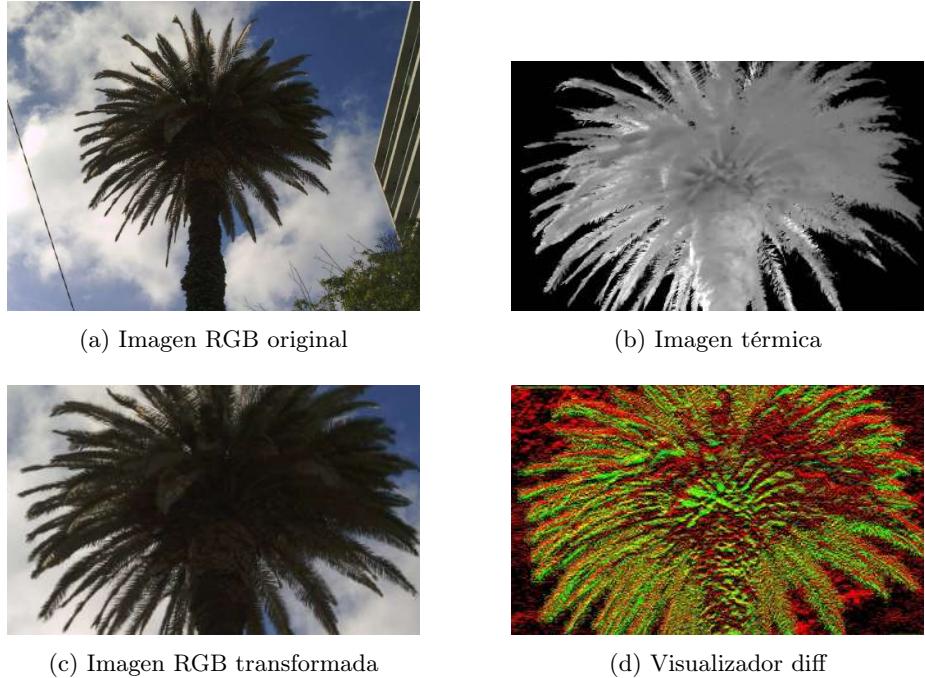


Figura 4.8: Comparación de imágenes RGB y térmicas en distintas etapas del proceso.

4.2.3. Cambio de gradiente RGB a escala de grises

Este procedimiento depende de la etapa de homografía, ya que la conversión a escala de grises se aplica sobre las imágenes RGB previamente alineadas por la etapa anterior. Si bien el orden de aplicación no altera el resultado final, conceptualmente se consideró más adecuado realizar primero la alineación geométrica y posteriormente la normalización de intensidades, ya que la homografía requiere la correspondencia espacial entre puntos característicos definidos sobre la imagen RGB original, los cuales podrían perderse o volverse ambiguos tras la conversión a escala de grises. De este modo, se garantiza que la estimación de la matriz de transformación se base en información completa de color y textura antes de reducir la dimensionalidad de la imagen.

El proceso consistió en transformar todas las imágenes RGB en escala de grises, con el objetivo de reducir la dimensionalidad (de tres canales a uno solo), eliminar la variabilidad cromática debida a las condiciones de iluminación y generar un conjunto más homogéneo para su eventual utilización en algoritmos de visión por computadora. Este tipo de normalización estandariza los datos y facilita la comparación entre capturas realizadas en distintos contextos.

No obstante, el aporte práctico de esta etapa resultó limitado. Debido a que el relevamiento priorizó las capturas nocturnas para garantizar la calidad térmica,

las imágenes RGB obtenidas en esas condiciones presentan baja calidad visual y escaso contenido informativo. En consecuencia, las imágenes RGB normalizadas no tuvieron un rol central en los análisis posteriores, siendo la modalidad térmica la principal fuente de información para la detección temprana del picudo rojo.

La Figura 4.9 muestra un ejemplo representativo del procedimiento. A la izquierda se observa una imagen RGB original y a la derecha su versión convertida a escala de grises. Como puede apreciarse, el proceso conserva las formas y estructuras principales de la palmera, pero elimina la información cromática, generando una representación más homogénea y estandarizada para etapas de procesamiento automático.

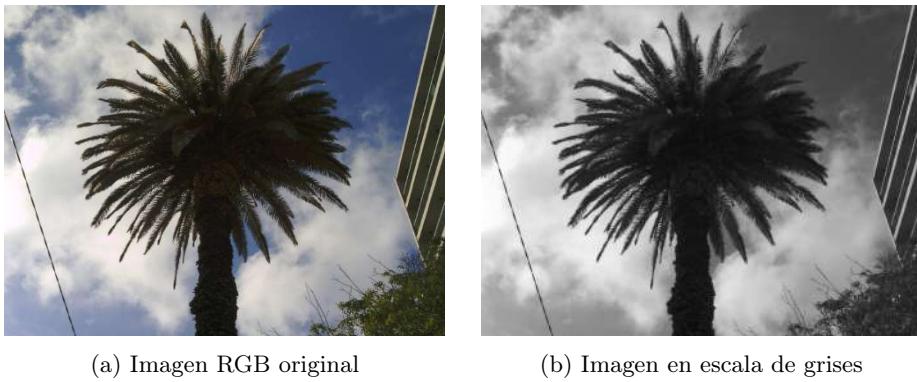


Figura 4.9: Transformación de imagen RGB a escala de grises.

4.2.4. Aumento de datos

Una de las principales limitaciones de este proyecto es la reducida cantidad de imágenes térmicas disponibles para entrenamiento y validación de los modelos clasificadores. Como se ha señalado en la literatura, los métodos de aprendizaje profundo y de clasificación basados en patrones de píxeles suelen requerir conjuntos de datos amplios y diversos para evitar sobreajuste y mejorar la capacidad de generalización (Shorten y Khoshgoftaar, 2019).

Con el fin de mitigar este problema, se implementó una etapa de aumento de datos, aplicada exclusivamente a las imágenes térmicas y posterior a su limpieza y conversión. En caso de no respetar este orden, ocurrirán errores inesperados como por ejemplo, el algoritmo que se encarga de eliminar elementos de la cámara tratará de recortar una imagen rotada.

Para esta tarea, se desarrolló un procedimiento llamado `augmentation.py` el cual recibe como parámetro un archivo CSV de origen y la cantidad de aumentos por cada clase. Tras realizar una serie de transformaciones a cada imagen del conjunto de datos de entrada, el algoritmo devuelve un nuevo conjunto de datos con los aumentos para dicha clase.

Las transformaciones implementadas incluyen operaciones básicas de aumento de datos ampliamente utilizadas en visión por computadora:

- **Rotaciones:** giros aleatorios en un rango entre -5° y 5° .
- **Escalado:** redimensionamiento aleatorio de la imagen en un rango entre el 100 % y el 140 % de su tamaño original.

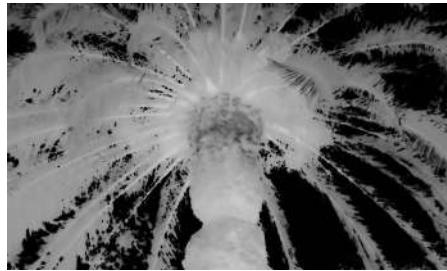
Se decidió intencionalmente no aplicar transformaciones que generen rotaciones pronunciadas ni alteren la intensidad de los píxeles en las imágenes. Esto se debe a que, en el contexto del presente proyecto, dichos valores representan directamente la temperatura normalizada de la palmera. En consecuencia, cualquier modificación artificial en la intensidad podría distorsionar la información térmica real y generar un sesgo en los modelos de clasificación posteriores, comprometiendo así la validez de los resultados obtenidos. Por este motivo, el aumento de datos se limitó a transformaciones geométricas (como rotaciones leves y cambios de escala) que preservan la correspondencia entre valores de intensidad y temperatura.

A su vez, no se aplicó la transformación de traslación, dado que las redes neuronales convolucionales presentan una cierta invariancia a traslaciones (Kauderer-Abrams, 2017). Por otra parte, las reflexiones fueron consideradas en etapas tempranas de desarrollo pero fueron posteriormente descartadas.

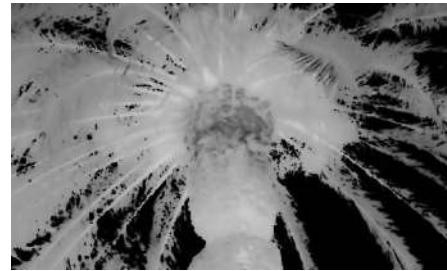
Estas técnicas incrementan artificialmente el tamaño del dataset, generando múltiples variantes de cada imagen térmica original. Si bien no aportan información nueva en sentido estricto, enriquecen el conjunto de entrenamiento al introducir variabilidad controlada, lo cual contribuye a que los clasificadores aprendan representaciones más robustas.

La Figura 4.10 ejemplifica el proceso de aumento de datos aplicado sobre las imágenes térmicas. En la subfigura 4.10a se muestra una imagen limpia tras las etapas de preprocesamiento, utilizada como punto de partida. A partir de ella, se generan cinco nuevas variantes sintéticas mediante las rotaciones y escalado, como puede observarse en las subfiguras 4.10b hasta 4.10f. Estas imágenes aumentadas son el resultado de aplicar ambas transformaciones de forma aleatoria.

Tras el proceso de aumento de datos, se adquirieron un total de 1970 imágenes térmicas en escala de grises, de las cuales 680 son de la clase afectada, 530 de la clase sana y 760 de la clase muerta.



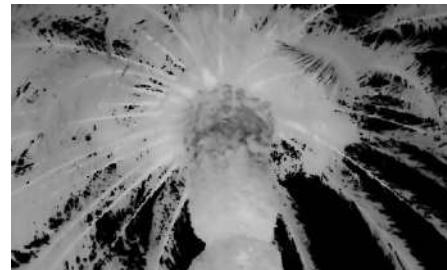
(a) Imagen en escala de grises original



(b) Imagen aumentada 1



(c) Imagen aumentada 2



(d) Imagen aumentada 3



(e) Imagen aumentada 4



(f) Imagen aumentada 5

Figura 4.10: Proceso de aumento de datos

4.2.5. Marcador de regiones de palmeras

Para la detección de regiones de la palmera —zonas críticas donde se observan los principales indicios de infección— se entrenó un modelo de la familia YOLO.

La elección de la versión YOLOv8 se fundamenta en su amplia adopción en la literatura científica reciente dedicada a la detección de objetos, donde ha demostrado un gran rendimiento en términos de precisión y velocidad. Este modelo mantiene la filosofía de detección de una sola pasada (*single-shot*), característica de la familia YOLO, permitiendo identificar simultáneamente la ubicación y la clase de los objetos en una única inferencia. Diversos estudios han validado su eficiencia computacional y capacidad de generalización en múltiples dominios, lo que respalda su elección para el presente proyecto enfocado en la detección de regiones específicas en palmeras a partir de imágenes.

El modelo utilizado fue YOLOv8 versión 'small' (YOLOv8s), que ofrece un equilibrio adecuado entre precisión y eficiencia, considerando la disponibilidad limitada de imágenes etiquetadas. Si bien en la actualidad existen versiones más recientes de la familia YOLO, al momento de iniciar el proyecto, YOLOv8 era la versión más consolidada, con documentación extensa, bibliotecas estandarizadas y una comunidad activa que respalda su uso en distintos ámbitos, incluyendo aplicaciones con objetivos similares. Versiones posteriores presentan mejoras, pero muchas aún se encuentran en etapas tempranas de adopción, con menor validación en casos de uso reales y menor estabilidad de bibliotecas asociadas. En este sentido, optar por YOLOv8 resultó en la mejor opción para el uso objetivo que fue dado, garantizando un marco de trabajo robusto, reproducible y suficientemente probado.

El entrenamiento fue realizado a partir de anotaciones en formato CSV, generadas manualmente para cada imagen del conjunto de datos. Si bien existen herramientas de etiquetado, se optó por la generación manual de las anotaciones debido a la naturaleza particular de las imágenes térmicas. En este tipo de imágenes, las zonas de interés como el cogollo o las hojas requieren una delimitación basada en el conocimiento contextual del observador (conocimiento del dominio), lo cual no se creyó que pudiera ser replicado de manera fiable por herramientas automáticas. De este modo, el etiquetado manual garantizó una mayor exactitud en la definición de las regiones relevantes para el entrenamiento del modelo.

Las anotaciones mencionadas incluyen las coordenadas x_{\min} , y_{\min} , x_{\max} y y_{\max} que definen la región de interés (*bounding box*). Posteriormente, dichas coordenadas fueron transformadas al formato requerido por YOLO, basado en coordenadas normalizadas de centro (x , y) junto con ancho y alto del cuadro delimitador, manteniendo una única clase asignada a todos los ejemplos, denominada “cogollo”, “hojas” o “tronco”.

El archivo `train.py` implementa el proceso de entrenamiento utilizando la biblioteca `ultralytics` ([Jocher y cols., 2023](#)). Se calibró de forma empírica mediante entrenamientos preliminares la cantidad de épocas según la estadísticas generadas por el modelo. Para otro tipo de parámetros, el modelo está diseña-

do para optimizarlos de forma automática. El modelo se entrena durante 100 épocas con imágenes redimensionadas a 992×992 píxeles, donde originalmente las imágenes poseen una resolución de 992×600 . Cuando se entrena un modelo YOLO con una dimensión de entrada fija pero las imágenes originales tienen un tamaño diferente, el algoritmo redimensiona cada imagen para que se ajuste a las dimensiones esperadas por la red neuronal. Este ajuste no deforma la imagen directamente, sino que mantiene su relación de aspecto y rellena con píxeles adicionales negros para completar la dimensión cuadrada. En este caso, la imagen de 992×600 se centra en un lienzo de 992×992 , agregando bordes arriba y abajo (padding vertical) para llenar el espacio restante. Este procedimiento garantiza que el modelo reciba entradas consistentes, sin distorsionar los objetos, lo cual es clave para que la detección funcione correctamente. Por otra parte, el archivo de configuración `base.yaml` define la estructura del conjunto de datos y la ubicación de los archivos de entrenamiento y validación. El tamaño de batch fue establecido en 16 para favorecer una utilización eficiente de la GPU disponible. La Tabla 4.6 muestra las especificaciones del equipo con el que se llevó a cabo el entrenamiento.

Componente	Especificación
Procesador (CPU)	AMD Ryzen 9 9900X (12 núcleos, 4.40 GHz)
Memoria RAM	32 GB
Tarjeta gráfica (GPU)	NVIDIA RTX 4080 Super

Tabla 4.6: Especificaciones del equipo utilizado para el entrenamiento.

Una vez finalizado el entrenamiento, el modelo se utiliza en el script `palm_marker.py` para procesar nuevas imágenes. A partir de un archivo CSV que contiene los nombres de las imágenes y los valores térmicos asociados, cada imagen es cargada y se ejecuta la inferencia mediante el modelo entrenado. En caso de detección, las coordenadas de la región de interés (en formato centro y dimensiones) se agregan a la correspondiente fila del CSV. En caso contrario, se registra una fila con valores nulos. Este procedimiento permite automatizar el etiquetado y facilitar tareas posteriores de análisis térmico en la zona marcada.

La Figura 4.11 muestra visualmente las predicciones del modelo sobre el conjunto de validación. Las imágenes del lado izquierdo representan las etiquetas reales con las regiones de interés esperadas para el cogollo. Mientras que las imágenes del lado derecho muestran las predicciones del modelo, junto con su respectiva confianza. Se observa que el modelo logra detectar correctamente los cogollos en la mayoría de los casos, aunque existen algunas predicciones con baja confianza (alrededor de 0.4) que pueden indicar ambigüedad o dificultad visual en ciertas imágenes. Esta visualización permite verificar la efectividad del modelo para focalizarse en la región de interés que será utilizada para el análisis térmico. Si bien la etiqueta manual del “cogollo” abarca una región ligeramente más amplia que el centro exacto, esta decisión fue intencional. En las imágenes térmicas, los límites del cogollo no son totalmente nítidos, y por este motivo, se

optó por incluir un margen adicional alrededor del área central, abarcando las zonas contiguas. Esta elección busca asegurar que el modelo aprenda patrones térmicos representativos del entorno inmediato del cogollo.

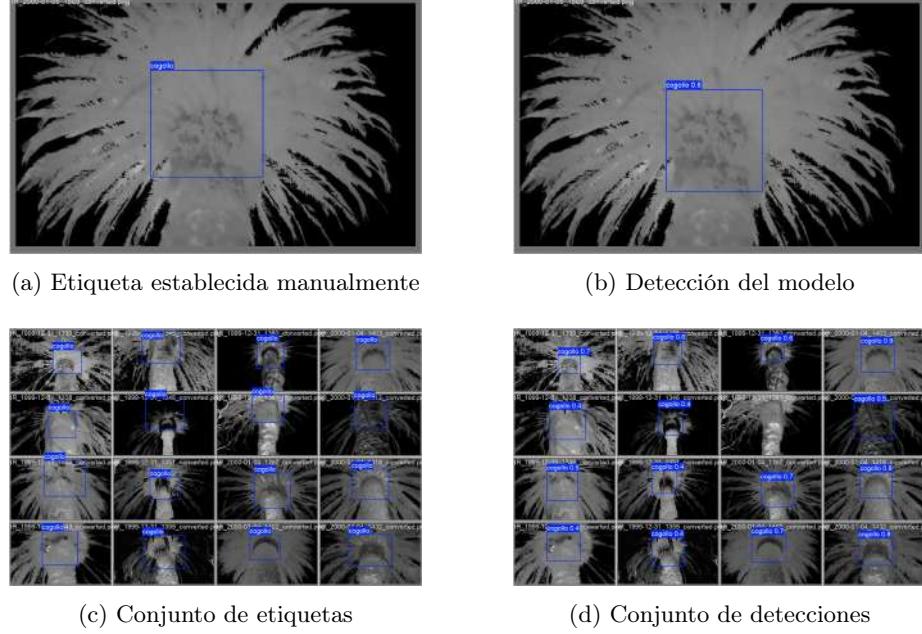


Figura 4.11: Resultados de la detección del cogollo.

La Figura 4.12 contiene gráficas de la evolución de las métricas y pérdidas durante los 100 épocas de entrenamiento sobre un conjunto de 165 imágenes térmicas etiquetadas manualmente, utilizando una división del 80 % para entrenamiento (132 imágenes) y 20 % para validación (33 imágenes).

Las métricas `train/box_loss`, `train/cls_loss` y `train/dfl_loss` presentan una disminución consistente a lo largo del entrenamiento, lo que indica que el modelo mejora progresivamente su capacidad para ajustar las cajas delimitadoras y clasificar correctamente los objetos. Por otro lado, las métricas correspondientes al conjunto de validación (`val/box_loss`, `val/cls_loss`, `val/dfl_loss`) muestran un comportamiento más ruidoso, como es habitual en problemas de visión por computadora con conjuntos de datos reducidos, aunque mantienen una tendencia general decreciente, lo que sugiere que el modelo está generalizando adecuadamente.

Las métricas *precision* (sección 3.8.2) y *recall* (sección 3.8.3) aumentan a medida que avanza el entrenamiento, estabilizándose en torno a 0.8. Esto indica que el modelo adquiere una buena capacidad para detectar los cogollos y predecirlos correctamente, minimizando tanto los falsos positivos como los falsos negativos.

En cuanto a las métricas estándar de precisión media, *mAP@0,5* y *mAP@[0,5 :*

0,95], detalladas en la sección 3.8.6, se observa también una evolución positiva. La primera alcanza valores cercanos a 0.75, mientras que la segunda se estabiliza alrededor de 0.4. Estos resultados son razonables considerando que el conjunto de datos utilizado consta de un número relativamente pequeño de imágenes etiquetadas manualmente.

En conjunto, los resultados sugieren que el modelo ha aprendido de forma efectiva a identificar los cogollos de las palmeras con una precisión aceptable. Sin embargo, aún existe margen de mejora, ya sea mediante la incorporación de un mayor número de imágenes, el refinamiento de las anotaciones existentes o la aplicación de técnicas de aumento de datos para robustecer el entrenamiento.

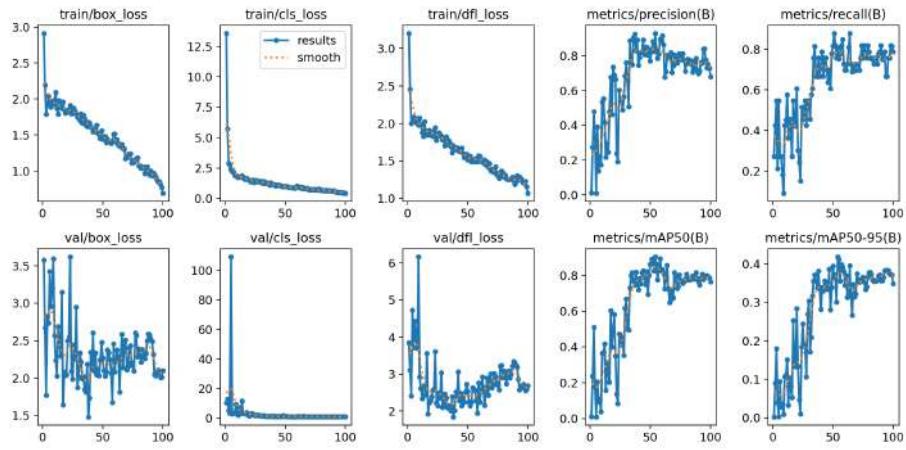
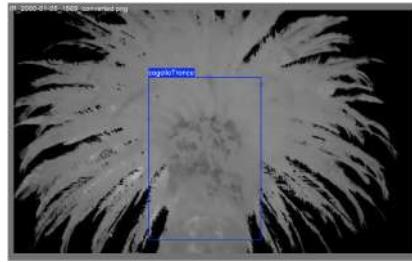


Figura 4.12: Evolución de métricas

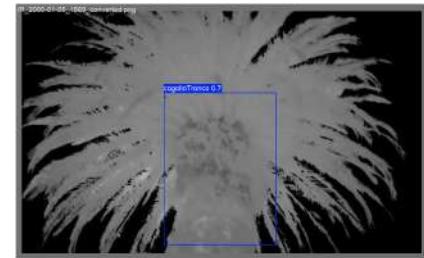
4.2.6. Variaciones del modelo: detección de tronco y hojas

Además del modelo entrenado para la detección exclusiva del cogollo, se exploraron dos variaciones adicionales con el objetivo de ampliar el rango de información visual capturada en cada imagen. Al igual que en el caso anterior, las anotaciones se realizaron manualmente en formato CSV y posteriormente se transformaron al esquema de coordenadas normalizadas requerido por YOLO. En estas variaciones, se definieron clases alternativas de regiones más grandes que abarcan distintas zonas de la palmera:

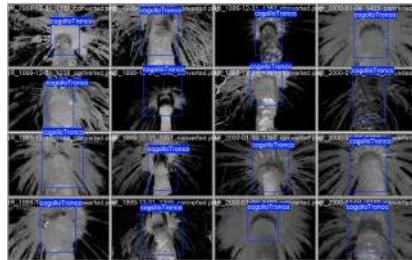
- **Primera variación (tronco y cogollo):** se entrenó un modelo para detectar regiones que incluyen el tronco y parte de la base del cogollo. Esta configuración busca capturar patrones asociados a la estructura central de la palmera, permitiendo un análisis contextual que podría complementar la identificación de indicios de infección en el cogollo.



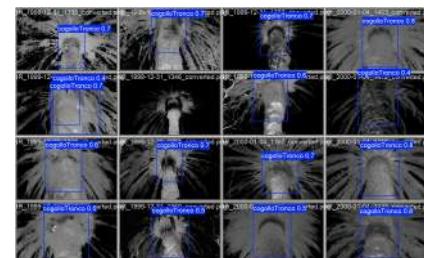
(a) Etiqueta establecida manualmente para el tronco y cogollo



(b) Detección del modelo



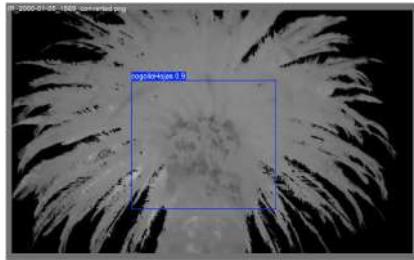
(c) Conjunto de etiquetas para el tronco y cogollo



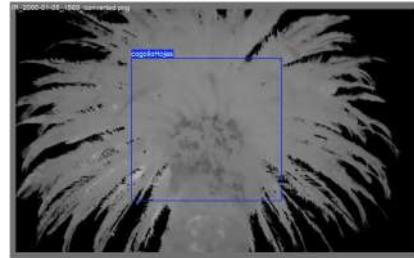
(d) Conjunto de detecciones

Figura 4.13: Resultados de la detección del tronco y cogollo.

- **Segunda variación (hojas):** en esta alternativa, las regiones abarcan las hojas que rodean al cogollo. La hipótesis es que, en ciertas condiciones, las hojas presentan cambios que pueden anticipar o reforzar la detección de problemas en la zona crítica.



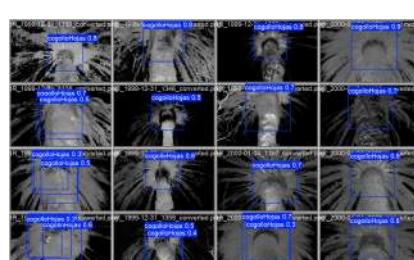
(a) Etiqueta establecida manualmente para las hojas y el cogollo



(b) Detección del modelo



(c) Conjunto de etiquetas para las hojas y el cogollo



(d) Conjunto de detecciones

Figura 4.14: Resultados de la detección de hojas y cogollo.

En ambos casos se utilizó el mismo procedimiento de entrenamiento con el script `train.py`, ajustando únicamente la definición de clases y las regiones de interés anotadas.

Estas variaciones no buscan reemplazar la detección del cogollo, sino complementarla. La detección del tronco y las hojas podría servir como referencia adicional en análisis multiclase, por ejemplo, combinando la información térmica de distintas zonas de la palmera o desarrollando modelos jerárquicos donde la detección del cogollo esté condicionada por la presencia simultánea de tronco y hojas correctamente identificados.

El análisis comparativo de las tres variantes permite observar diferencias claras en el alcance de las detecciones:

- En la **detección del cogollo** (Figura 4.11) el modelo se focaliza en la región central, lo que facilita un análisis térmico más preciso en la zona crítica de la palmera. Sin embargo, esta aproximación depende de que el cogollo esté claramente visible, lo cual no siempre ocurre en imágenes con ruido o condiciones de iluminación desfavorables.
- En la **detección del cogollo y las hojas** (Figura 4.14), las regiones de interés abarcan un área más amplia que incluye el follaje inmediato alrededor del cogollo. Esto permite capturar indicios adicionales presentes en las hojas, como variaciones térmicas o de su forma natural, aunque introduce mayor variabilidad dentro de la clase, lo que puede afectar la precisión del modelo.
- En la **detección del cogollo y el tronco** (Figura 4.13), la caja delimitadora se extiende hacia abajo para incluir parte del tronco. Esta configuración puede resultar útil para estudios que busquen correlacionar el estado del cogollo con el de la estructura central de la palmera. No obstante, al abarcar una región más heterogénea, el modelo puede perder especificidad en la detección exclusiva del cogollo.

En conjunto, las variaciones muestran un compromiso entre *focalización* y *contexto*: mientras la detección del cogollo maximiza la precisión en la zona crítica, las variantes con hojas y tronco ofrecen información contextual que podría resultar valiosa.

4.2.7. Determinador de temperatura promedio

Este paso del preprocesamiento consistió en estimar la temperatura promedio dentro de las regiones de interés. Este punto resulta clave, dado que la hipótesis principal del proyecto establece que las palmeras afectadas por el picudo rojo presentan, en promedio, temperaturas superficiales mayores que las sanas.

Es importante remarcar que la métrica calculada no corresponde necesariamente al cogollo en sí mismo, sino al área rectangular definida por las zonas de interés.

Para llevar a cabo esta tarea se desarrollaron tres herramientas principales: `promedio_píxeles.py`, `boxplot.py` y `statistics.py`, cada una con un rol específico en el flujo de procesamiento.

1. Cálculo de temperaturas promedio. La Figura 4.7 describe la estructura del archivo CSV que utiliza el procedimiento `promedio_píxeles.py`.

Variable	Descripción
<code>image_path</code>	Ruta a la imagen térmica en escala de grises.
<code>min_t</code>	Temperatura mínima de la escala térmica de la cámara (°C).
<code>max_t</code>	Temperatura máxima de la escala térmica de la cámara (°C).
<code>cls</code>	Clase o clasificación de la palmera.
<code>ambient</code>	Temperatura ambiente a la hora de la captura (°C).
<code>x_min</code>	Coordenada x mínima de la región de interés.
<code>x_max</code>	Coordenada x máxima de la región de interés.
<code>y_min</code>	Coordenada y mínima de la región de interés.
<code>y_max</code>	Coordenada y máxima de la región de interés.

Tabla 4.7: Descripción de las variables del dataset de imágenes térmicas

A partir de esta información, el algoritmo recorta la región correspondiente a la zona de interés y filtra los píxeles con intensidades inferiores al umbral predefinido de 15. Este filtrado permite excluir regiones no representativas, como cielo o bordes fríos, que podrían sesgar la estimación. Es importante destacar que, durante la conversión inicial de la imagen a escala de grises, se estableció un mapeo directo entre los valores de intensidad de cada píxel y el rango de temperaturas presente en la captura, de modo que cada nivel de gris corresponde a un valor específico en °C.

El resultado se guarda en un nuevo archivo CSV que incluye la ruta de la imagen, el valor promedio de píxeles en escala de grises, la temperatura promedio estimada dentro de la región, la clasificación de la palmera (muerta, sana, afectada) y la diferencia entre la temperatura ambiente y la de la región de interés. Esta última métrica resulta especialmente valiosa para comparar capturas realizadas en diferentes condiciones ambientales, ya que reduce el sesgo introducido por la variabilidad climática.

2. Visualización y análisis exploratorio. El archivo CSV generado en la etapa anterior se utiliza como entrada del procedimiento `boxplot.py`, encargado de producir diagramas de caja (boxplots) diferenciados por la clasificación de la palmera. Se representan dos métricas, la temperatura promedio absoluta dentro de la región delimitada y la diferencia con respecto a la temperatura ambiente. Este enfoque gráfico facilita la identificación de patrones y la evaluación preliminar de la hipótesis.

De manera complementaria, el procedimiento `statistics.py` genera histogramas de temperatura para cada clase de palmera, redondeando los valores de temperatura a múltiplos de $0,5^{\circ}\text{C}$. Estos gráficos permiten observar la distribución de valores térmicos en cada categoría (sanas, afectadas y muertas), aportando una visión más detallada sobre posibles solapamientos o separaciones entre clases.

Las Figuras 4.15 y 4.16 presentan un ejemplo ilustrativo de ambas visualizaciones. En conjunto, las herramientas mencionadas permiten explorar las diferencias térmicas entre categorías y evaluar de manera preliminar la validez de la hipótesis planteada.

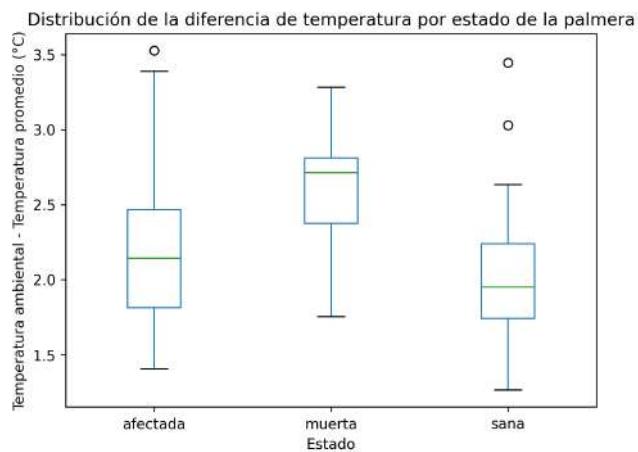


Figura 4.15: Ejemplo de diagrama de cajas con la distribución de temperatura promedio por clase de palmera.

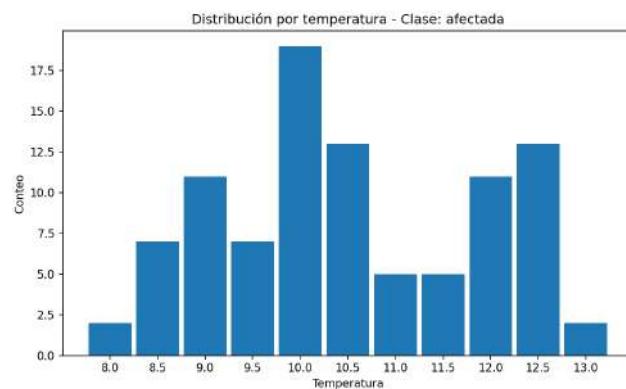


Figura 4.16: Ejemplo de histograma de la temperatura promedio de palmeras afectadas.

3. Rol dentro del proceso. La determinación de la temperatura promedio dentro de las áreas de interés (*bounding boxes*) constituye un puente entre las etapas de preprocesamiento y los modelos clasificadores. Por un lado, permite realizar un análisis estadístico exploratorio que valida la existencia de diferencias térmicas entre categorías; por otro, proporciona insumos numéricos que pueden integrarse en clasificadores tradicionales (e.g., SVM o Random Forest) o como variables adicionales en modelos de aprendizaje profundo.

Esta etapa permitió sistematizar la extracción de una métrica térmica clave y contrastarla con la temperatura ambiente. Si bien los resultados preliminares sugieren diferencias detectables entre clases, la variabilidad observada indica la necesidad de emplear modelos más robustos que integren información espacial adicional. Además, al no limitarse únicamente al cogollo, esta metodología abre la posibilidad de extender el análisis a distintas regiones.

4.3. Clasificación

El objetivo de esta etapa es clasificar cada palmera como *sana*, *afectada* o *muerta* en base a un archivo CSV de entrada. Para lograrlo, se diseñaron y entrenaron distintos clasificadores que aprovechan tanto la información numérica (a través de temperaturas) como las características visuales presentes en las imágenes procesadas.

Los enfoques explorados se agrupan en cinco líneas principales:

- **Modelos de línea base**, sin entrenamiento previo, que asignan clases siguiendo un criterio arbitrario.
- **Modelos de Machine Learning tradicionales**, entrenados sobre valores derivados de las imágenes térmicas (temperatura promedio, mínima, máxima, varianza y diferencia de temperatura con el ambiente).
- **Modelos basados en extracción de características**, en los cuales una red neuronal convolucional preentrenada se emplea únicamente como extracto de descriptores visuales. Dichos vectores de características se combinan con información tabular y son posteriormente clasificados mediante algoritmos tradicionales de aprendizaje automático.
- **Redes Neuronales Convolucionales** (CNN, según sus siglas en inglés), aplicadas de comienzo a fin, entrenadas exclusivamente con imágenes del conjunto de datos, con el objetivo de identificar directamente patrones visuales asociados al estado sanitario.
- **Modelos híbridos**, que integran simultáneamente información tabular y visual, combinando la salida probabilística de modelos tradicionales y de CNN para reforzar la discriminación entre clases.

En las siguientes subsecciones se presentan los distintos clasificadores desarrollados siguiendo esta progresión, describiendo en cada caso la metodología empleada, los resultados alcanzados y las mejoras

4.3.1. Consideraciones generales

Antes de detallar los enfoques particulares explorados, resulta pertinente presentar las decisiones generales que guiaron el desarrollo de los clasificadores.

En primer lugar, el conjunto de datos se dividió de manera estratificada en tres particiones: *entrenamiento*, *validación* y *prueba*. La partición se realizó de modo que el conjunto de entrenamiento contenga un 70 % de las instancias, el conjunto de validación un 10 % y el conjunto de prueba un 20 % de las instancias, garantizando siempre la preservación de la proporción de clases en cada subconjunto. El conjunto de entrenamiento se utilizó para el ajuste de los parámetros internos de los modelos, mientras que el de validación permitió seleccionar hiperparámetros y evitar el sobreajuste. Finalmente, el conjunto de prueba se reservó

estRICTAMENTE PARA LA EVALUACIÓN FINAL, ASEGURANDO UNA ESTIMACIÓN NO SESGADA DEL DESEMPEÑO.

EN CUANTO AL PREPROCESAMIENTO, TODAS LAS ESTRATEGIAS SIGUIERON EL PROTOCOLO COMÚN DE LIMPIEZA DE DATOS, CONVERSIÓN DE IMÁGENES TÉRMICAS A ESCALA DE GRISSES Y NORMALIZACIÓN DE VALORES. EN LOS MODELOS BASADOS EN ATRIBUTOS TABULARES, LAS CARACTERÍSTICAS CONSIDERADAS INCLUYERON TEMPERATURA PROMEDIO, MÍNIMA, Y MÁXIMA. POR OTRA PARTE, LAS REDES NEURONALES CONVOLUCIONALES FUERON ADAPTADAS PARA TRABAJAR CON IMÁGENES DE UN SOLO CANAL, PRESERVANDO LA COMPATIBILIDAD CON ARQUITECTURAS PREENTRENADAS.

EL DESEMPEÑO DE CADA MODELO SE EVALUÓ MEDIANTE MÉTRICAS ESTÁNDAR DE CLASIFICACIÓN (*precision*, *recall*, *f1-score*, *accuracy*, *macro avg* y *weighted avg*), OBTENIDAS A PARTIR DEL CONJUNTO DE PRUEBA.

ASIMISMO, SE PRIORIZÓ LA REPRODUCIBILIDAD DE LOS EXPERIMENTOS MEDIANTE LA FIJACIÓN DE SEMILLAS ALEATORIAS Y LA APLICACIÓN DE CRITERIOS DE BALANCEO DE CLASES EN LOS PROCESOS DE ENTRENAMIENTO. ESTA ÚLTIMA DECISIÓN BUSCÓ MITIGAR EL IMPACTO DEL DESBALANCE EXISTENTE ENTRE LAS CATEGORÍAS DE PALMERAS (SANAS, AFECTADAS Y MUERTAS), QUE CONSTITUYE UNO DE LOS PRINCIPALES DESAFÍOS DEL PROBLEMA.

4.3.2. Clasificador de Línea Base

CON EL FIN DE ESTABLECER UN PUNTO DE REFERENCIA MÍNIMO, SE INCLUYÓ UN *clásificador aleatorio* (*C0_baseline.py*) QUE NO UTILIZA INFORMACIÓN DE LAS IMÁGENES NI ATRIBUTOS DERIVADOS. ESTE MODELO ÚNICAMENTE REALIZA PREDICCIONES DE MANERA ALEATORIA Y UNIFORME ENTRE LAS CLASES DEFINIDAS.

ESTE MODELO BASE CUMPLE EL PROPÓSITO FUNDAMENTAL DE PROVEER UN UMBRAL INFERIOR DE DESEMPEÑO QUE CUALQUIER MODELO POSTERIOR DEBE SUPERAR.

SE ESPERA QUE, EN CONJUNTOS BALANCEADOS, LA EXACTITUD TIENDA A $1/C$, DONDE C ES EL NÚMERO DE CLASES.

4.3.3. Clasificador con Machine Learning

EL SEGUNDO ENFOQUE DE CLASIFICACIÓN DESARROLLADO (*C1_rf_svc.py*) SE CENTRA EXCLUSIVAMENTE EN INFORMACIÓN TABULAR DERIVADA DE LAS IMÁGENES TÉRMICAS PROCESADAS. UNA VEZ OBTENIDA LA IMAGEN PREVIAMENTE PROCESADA, SE CALCULARON ESTADÍSTICAS SOBRE LOS PÍXELES VÁLIDOS DE LA IMAGEN, ES DECIR, AQUELLOS PÍXELES DE LA IMAGEN CUYO VALOR SEA SUPERIOR A UN UMBRAL DEFINIDO (15 UNIDADES). DE ESTE MODO, CADA INSTANCIA QUEDÓ REPRESENTADA POR LA TEMPERATURA PROMEDIO, EL PROMEDIO DE LOS VALORES DE LOS PÍXELES VÁLIDOS, LA VARIANZA DE LOS VALORES VÁLIDOS Y EL VALOR MÍNIMO Y MÁXIMO DE TEMPERATURA DETECTADOS EN LA IMAGEN EN SU TOTALIDAD.

ESTAS CARACTERÍSTICAS, QUE COMBINAN INFORMACIÓN DE TEMPERATURA GLOBAL CON MEDIDAS DE DISPERSIÓN Y EXTREMOS, CONFORMARON EL VECTOR DE ENTRADA PARA LOS MODELOS DE CLASIFICACIÓN SUPERVISADA:

- **Support Vector Machine (SVM)**
- **Random Forest**

En el caso de la SVM, se implementó utilizando la clase SVC de la biblioteca `sklearn.svm`. Se optó por un kernel radial (RBF) debido a su capacidad para capturar fronteras de decisión no lineales, adecuadas para las relaciones complejas entre las variables térmicas. El parámetro de penalización C se fijó en 1, valor intermedio que equilibra la maximización del margen con la tolerancia a errores de clasificación, evitando tanto el sobreajuste como el subajuste. El parámetro γ se dejó en modo automático, permitiendo que el modelo ajuste la influencia de cada muestra en función de la dimensionalidad de los datos. Además, se emplearon *class weights* balanceados para contrarrestar la desigualdad en la distribución de clases, asegurando que las categorías minoritarias contribuyan proporcionalmente al entrenamiento.

Para Random Forest, se utilizó la clase RandomForestClassifier de `sklearn.ensemble`. Se estableció un número de 200 árboles de decisión, dado que un mayor número de estimadores tiende a estabilizar el modelo y reducir la varianza sin un incremento prohibitivo del costo computacional. La profundidad de los árboles se dejó sin restricciones explícitas para permitir que el algoritmo se adapte a la complejidad inherente de los datos. Asimismo, se incorporaron pesos balanceados entre clases, lo cual resulta esencial para mitigar el sesgo hacia categorías mayoritarias y mejorar la sensibilidad hacia las minoritarias.

La Tabla 4.8 resume los parámetros seleccionados para cada modelo junto con su justificación.

Modelo	Parámetro	Justificación
SVM	Kernel RBF	Captura fronteras no lineales adecuadas para relaciones complejas entre variables térmicas.
SVM	$C = 1$	Compromiso entre margen amplio y tolerancia a errores, evitando sobreajuste o subajuste.
SVM	γ automático	Ajuste adaptativo de la influencia de cada muestra según la dimensionalidad de los datos.
SVM	Pesos balanceados	Corrige el desbalance de clases, dando importancia proporcional a categorías minoritarias.
Random Forest	200 árboles	Aumenta la estabilidad del modelo y reduce varianza sin excesivo costo computacional.
Random Forest	Profundidad ilimitada	Permite capturar relaciones complejas sin restringir el crecimiento de los árboles.
Random Forest	Pesos balanceados	Mitiga el sesgo hacia clases mayoritarias y mejora la detección de clases menos representadas.

Tabla 4.8: Parámetros y justificación de los clasificadores en `C1_rf_svc`

Este enfoque constituye el punto de partida de la experimentación, basándose en un conjunto reducido de variables derivadas de las imágenes térmicas y deja de lado información visual más compleja. Su principal ventaja radica en la simplicidad y en la interpretabilidad de los atributos empleados.

4.3.4. Clasificador híbrido con CNN y Machine Learning

El tercer modelo desarrollado (*C2_features_cnn.py*) integra información visual y tabular para la clasificación del estado sanitario de las palmeras. En este caso, las imágenes térmicas se procesan mediante arquitecturas de redes neuronales convolucionales preentrenadas sobre el conjunto *ImageNet* (Deng y cols., 2009). Se consideraron tres modelos representativos de la familia *ResNet* (He y cols., 2016a) y *DenseNet* (Huang y cols., 2016):

- **ResNet18**
- **ResNet50**
- **DenseNet121**

Cada red se ajustó para aceptar imágenes de un solo canal (escala de grises) y su capa de salida (*fc* o *classifier*) fue reemplazada por una identidad, de manera que la red actúe únicamente como extractor de características. De este modo, para cada imagen se obtiene un vector de representación profunda (*deep features*, según su traducción en inglés) que captura patrones visuales relevantes.

A dichas representaciones se les concatenaron características tabulares como lo son la temperatura promedio, media, varianza, valor mínimo y máximo de los píxeles válidos, y la diferencia de temperatura con el ambiente. Todas estas métricas fueron calculadas directamente de los píxeles válidos de la imagen en cuestión, tal y como se explicó en la subsección anterior.

Las variables fueron normalizadas con *StandardScaler* y las etiquetas codificadas mediante *LabelEncoder*. El vector combinado (deep features + características tabulares) fue utilizado como entrada para clasificadores Support Vector Machine (SVM) y Random Forest.

La elección de ResNet18, ResNet50 y DenseNet121 responde a su amplia validación en la literatura como arquitecturas capaces de aprender representaciones visuales robustas. ResNet18 ofrece un modelo ligero y eficiente, mientras que ResNet50 incorpora mayor profundidad y capacidad expresiva; DenseNet121, por su parte, permite un mayor reaprovechamiento de características mediante conexiones densas. En todos los casos, se emplearon pesos preentrenados en *ImageNet*, lo cual favorece la transferencia de conocimiento y mejora la generalización al trabajar con conjuntos de datos reducidos.

Por otra parte, en el caso de SVM y Random Forest, se mantuvieron las mismas decisiones que en el segundo clasificador.

La Tabla 4.9 resume los parámetros seleccionados en este enfoque híbrido junto con su justificación.

Modelo	Parámetro	Justificación
ResNet18/50	Conv1 modificado	Adaptación a imágenes en escala de grises (1 canal).
ResNet18/50	Capa <i>fc</i> → identidad	Uso de la red como extractor de características en lugar de clasificador final.
DenseNet121	Conv0 modificado	Ajuste para entrada de un canal.
DenseNet121	Capa <i>classifier</i> → identidad	Extracción de representaciones profundas en lugar de predicción directa.

Tabla 4.9: Parámetros y justificación de los clasificadores en *C2_features_cnn*

Este clasificador híbrido representa una evolución respecto al modelo anterior; mientras que el primer enfoque se basaba exclusivamente en medidas estadísticas derivadas de las temperaturas, aquí se incorporan descriptores visuales aprendidos por CNN preentrenadas. De esta forma, se combina la información global de la palmera (texturas, formas, patrones visuales) con la información térmica local de las zonas de interés. El resultado es un modelo más robusto, con potencial para capturar tanto diferencias sutiles en la temperatura como variaciones visuales asociadas al estado de la planta.

4.3.5. Clasificador basado en CNN entrenada de extremo a extremo

El cuarto enfoque (*C3_cnn.py*) consistió en entrenar directamente redes neuronales convolucionales sobre las imágenes térmicas procesadas. De este modo, la red aprende de forma automática las representaciones necesarias para distinguir entre las clases *sana*, *afectada* y *muerta*.

Para este experimento se implementó un modelo que utiliza la arquitectura ResNet18 preentrenada y modificada para aceptar imágenes en escala de grises y con la capa de clasificación reemplazada por una capa lineal adaptada al número de clases.

Cada modelo se entrenó utilizando una carga de datos configurada con lotes de tamaño 8, aplicando transformaciones básicas de preprocesamiento (redimensionado a 224×224 píxeles y conversión a tensor). El entrenamiento se realizó con el optimizador Adam, una tasa de aprendizaje inicial de 0,001, regularización mediante *weight decay* de 10^{-4} y función de pérdida *CrossEntropyLoss*. Se ejecutaron 50 épocas completas sobre el conjunto de entrenamiento, registrando la pérdida promedio por época como indicador de convergencia.

Al finalizar el entrenamiento, los modelos se evaluaron sobre un conjunto de prueba, utilizando las probabilidades de salida de la capa *softmax* para calcular las predicciones finales. El desempeño se reportó mediante métricas estándar de clasificación (*precision*, *recall*, *f1-score*), lo que permitió comparar el rendimiento relativo entre arquitecturas.

La elección de ResNet18 responde a su eficacia probada en múltiples dominios de visión por computadora y a la disponibilidad de pesos preentrenados, lo que favorece la transferencia de conocimiento en contextos con conjuntos de datos limitados. ResNet18, al ser más ligera, permite entrenamientos rápidos con menor riesgo de sobreajuste.

El batch size de 8 se estableció como un compromiso entre la estabilidad de las actualizaciones de gradiente y las limitaciones de memoria de GPU disponibles. Se fijaron 50 épocas de entrenamiento para asegurar la convergencia del modelo sin caer en un sobreentrenamiento, monitorizando la pérdida como criterio de estabilidad. La tasa de aprendizaje inicial de 0,001 es un valor comúnmente recomendado en combinación con el optimizador Adam, mientras que el *weight decay* de 10^{-4} se empleó como regularización para reducir el riesgo de sobreajuste y mejorar la generalización. La función de pérdida *CrossEntropyLoss* se utilizó por ser el estándar en tareas de clasificación multiclase.

La Tabla 4.10 resume los parámetros seleccionados y su justificación en este enfoque de entrenamiento de extremo a extremo.

Modelo	Parámetro	Justificación
ResNet18	Conv1 modificado	Adaptación a imágenes en escala de grises (1 canal).
ResNet18	Capa final lineal	Ajuste del número de salidas a las tres clases definidas.
Entrenamiento	Batch size = 8	Compromiso entre estabilidad y limitaciones de memoria GPU.
Entrenamiento	Epochs = 50	Garantiza convergencia sin caer en sobreajuste prematuro.
Entrenamiento	Optimizador Adam	Método robusto y eficiente para optimizar redes profundas.
Entrenamiento	LR = 0.001	Valor recomendado para Adam que equilibra rapidez y estabilidad.
Entrenamiento	Weight decay = 10^{-4}	Regularización que reduce sobreajuste y mejora generalización.
Entrenamiento	CrossEntropyLoss	Función estándar para clasificación multiclas.

Tabla 4.10: Parámetros y justificación de los clasificadores en *C3_cnn.py*

Mientras que en la versión anterior las CNN actuaban como extractores de características fijos complementados con clasificadores clásicos, en este caso las redes se entrenan de extremo a extremo, especializándose directamente en el dominio de imágenes térmicas de palmeras.

4.3.6. Clasificador híbrido con optimización y reentrenamiento

Esta versión del clasificador (*C4_cnn_rf.py*) combina dos enfoques complementarios. Por un lado, se entrena una red convolucional ResNet18 adaptada para trabajar con imágenes en escala de grises, responsable de extraer características visuales profundas. Por otro, se entrena un modelo Random Forest sobre descriptores tabulares derivados de las imágenes térmicas (temperatura promedio, varianza, mínimo y máximo). Ambos clasificadores producen vectores de probabilidad sobre las clases; el Random Forest (RF) se calibra mediante *isotonic regression* para mejorar la fiabilidad de sus probabilidades, mientras que la CNN (ResNet18) entrega distribuciones *softmax*. La fusión inicial se realiza como un promedio ponderado:

$$P_{\text{fusión}} = \lambda P_{\text{CNN}} + (1 - \lambda) P_{\text{RF}},$$

donde $\lambda \in [0, 1]$ controla el peso relativo de cada clasificador. Adicionalmente, se aplica un refuerzo específico sobre la clase *afectada* para priorizar su detección. El uso de este refuerzo es debido a que el modelo Random Forest tiende a clasificar mejor la clase afectada, ya que ésta presenta cambios térmicos detectables incluso cuando el patrón visual no es evidente. Denotando por $c = \text{afectada}$ y por $\beta \in [0, 1]$ el factor de refuerzo, la probabilidad combinada en esa clase se redefine como:

$$P_{\text{final}}(c) = \underbrace{\lambda(1 - \beta)}_{\text{peso CNN reducido}} P_{\text{CNN}}(c) + \underbrace{(1 - \lambda(1 - \beta))}_{\text{peso RF aumentado}} P_{\text{RF}}(c).$$

Para el resto de las clases se mantiene $P_{\text{fusión}}$. Finalmente, las probabilidades se renormalizan por fila para asegurar que $\sum_k P_{\text{final}}(k) = 1$.

Fase 1: Optimización de hiperparámetros

En esta etapa, implementada en el procedimiento `optimize_hparams.py`, se utiliza la biblioteca Optuna ([Akiba, Sano, Yanase, Ohta, y Koyama, 2019](#)) para explorar diferentes configuraciones de hiperparámetros, puntualmente los siguientes:

- **Random Forest:** número de árboles, profundidad máxima, tamaño mínimo de hojas y divisiones internas.
- **CNN (ResNet18):** tasa de aprendizaje, *weight decay*, tamaño de batch y número de épocas de entrenamiento de la cabeza de clasificación.

El uso de Optuna permite una búsqueda sistemática y eficiente en espacios de hiperparámetros complejos, evitando la selección manual y reduciendo la probabilidad de configuraciones subóptimas. En el caso del Random Forest, la exploración del número de árboles y la profundidad busca equilibrar estabilidad y capacidad de generalización. En la red convolucional, la optimización de la tasa de aprendizaje y el *weight decay* apunta a mejorar la convergencia y la regularización, mientras que el tamaño de lotes (*batch size*) ajusta la estabilidad del gradiente a las restricciones de hardware. La fusión ponderada garantiza que tanto la información visual como la tabular aporten de forma proporcional, y el refuerzo explícito de la clase *afectada* responde a la prioridad práctica de detectar tempranamente los ejemplares en riesgo.

La Tabla 4.11 resume los principales hiperparámetros optimizados para cada componente del modelo (Random Forest, ResNet18 y el esquema de fusión), junto con la justificación de su elección.

Componente	Parámetro	Justificación
Random Forest	<i>n_estimators</i>	Aumentar número de árboles mejora estabilidad sin sobreajuste excesivo.
Random Forest	<i>max_depth</i>	Controla la complejidad de cada árbol, evitando sobreajuste.
Random Forest	<i>min_samples_split</i>	Previene divisiones demasiado pequeñas.
Random Forest	<i>min_samples_leaf</i>	Previene divisiones demasiado pequeñas, mejorando generalización.
CNN (ResNet18)	Learning rate	Ajusta la velocidad de convergencia; valores óptimos mejoran eficiencia del entrenamiento.
CNN (ResNet18)	Weight decay	Actúa como regularización para evitar sobreajuste.
CNN (ResNet18)	Batch size	Compromiso entre estabilidad del gradiente y limitaciones de GPU.
CNN (ResNet18)	Épocas (cabeza)	Entrenamiento inicial de la cabeza para adaptar la red al dominio.
CNN (ResNet18)	Épocas (final = 100)	Entrenamiento completo tras optimización, asegurando convergencia estable.
CNN (ResNet18)	OneCycleLR	Estrategia de tasa de aprendizaje que acelera la convergencia y mejora generalización.
Fusión	Peso λ	Controla el aporte relativo de la CNN y el RF en la decisión final.

Tabla 4.11: Parámetros optimizados y justificación en *C4_cnn_rf*

Cada configuración se evaluó sobre el conjunto de validación empleando una métrica compuesta que prioriza en un 70 % el *macro-F1* y en un 30 % el *F1* de la clase *afectada*, priorizando así la detección de casos de interés agronómico. Tras pruebas de 200 y 400 ensayos, se seleccionaron los mejores hiperparámetros, que

se almacenaron en un archivo CSV (`mejores_hiperparametros.csv`) para su uso posterior.

Fase 2: Reentrenamiento y evaluación final

La segunda fase, implementada en el procedimiento principal, toma los hiperparámetros óptimos desde el archivo generado en la fase anterior y entrena nuevamente ambos clasificadores con mayor profundidad:

1. El **Random Forest** se reentrena sobre la unión de los conjuntos de entrenamiento y validación, y se calibra mediante *isotonic regression* ([Niculescu-Mizil y Caruana, 2005](#)) para ajustar las probabilidades de salida.
2. La **ResNet18** se entrena durante 100 épocas completas utilizando el planificador de tasa de aprendizaje **OneCycleLR** ([Smith y Topin, 2019](#)), con regularización mediante *weight decay* y un aumento de datos ligero (rotación aleatoria y recorte central).
3. Los vectores de probabilidades de ambos clasificadores se combinan con el peso α optimizado previamente.
4. Se aplica un ajuste adicional sobre las probabilidades asociadas a la clase *afectada*, incrementando su valor en la combinación final mediante un factor de refuerzo ($\beta = 0.25$), aumentando la sensibilidad del modelo ante casos tempranos de infestación, sin alterar significativamente el equilibrio general entre clases.

Este flujo de trabajo híbrido combina la capacidad de las CNN para capturar patrones visuales complejos con la estabilidad de un modelo calibrado, integrando además un mecanismo de refuerzo para la clase más crítica. La separación en dos fases —optimización automática y reentrenamiento final— garantiza tanto la eficiencia en la búsqueda de hiperparámetros como la solidez en la evaluación sobre el conjunto de prueba.

4.3.7. Consolidación de los modelos generados

Para poner en marcha el uso de los modelos de clasificación comentados anteriormente, se creó el procedimiento `clasificador.py`. Este se encarga de utilizar los modelos previamente entrenados (la versión que se consideró más calificada), con Random Forest calibrado, escalador de características y la red neuronal convolucional ResNet18, para clasificar nuevas imágenes a partir de un archivo CSV, proveniente del flujo trabajado. El algoritmo genera las predicciones del Random Forest y la CNN, combina ambas probabilidades según el hiperparámetro `cnn_weight`, y produce un CSV con las predicciones finales y las probabilidades asociadas a cada clase (muerta, sana, afectada). Además, guarda un reporte de métricas asociados a la ejecución.

Capítulo 5

Experimentación

En este capítulo se presentan los experimentos realizados con el objetivo de validar la hipótesis de trabajo que las palmeras afectadas por el Picudo Rojo presentan, en promedio, temperaturas superficiales mayores que las sanas. Para ello, se utilizaron los conjuntos de datos previamente obtenidos en la Sección 4.1, se aplicaron los algoritmos de preprocesamiento descritos en la Sección 4.2, y se diseñó una serie de pruebas controladas que permitieron evaluar la eficacia de los diferentes clasificadores definidos la Sección 4.3.

5.1. Diseño experimental

El objetivo de esta sección es describir con precisión la hipótesis a validar, los conjuntos de datos empleados, las métricas consideradas y el procedimiento general seguido para producir los resultados. Todas las pruebas se realizaron sobre imágenes térmicas previamente procesadas (recorte de artefactos, inpainting del puntero, normalización a escala de grises) y con regiones de interés definidas por el detector basado en YOLOv8.

5.1.1. Hipótesis

1. **Diferenciación térmica:** La hipótesis central de trabajo sostiene que las diferencias fisiológicas entre palmeras *sanas*, *afectadas* y *muertas* se manifiestan como patrones térmicos diferenciables en imágenes infrarrojas. En particular, se plantea que, en promedio, las palmeras *afectadas* presentan temperaturas superficiales mayores que las *sanas* en regiones críticas de la copa, con especial énfasis en el cogollo.
2. **Inercia térmica:** La inercia térmica de las palmeras juega un papel muy importante dependiendo de su estado sanitario. Las palmeras sanas presentan una mayor inercia térmica que sus contrapartes muertas, producto de mecanismos internos como un flujo continuo de savia. Esto les permite calentarse más lentamente durante el día, y enfriarse más lentamente durante la noche. Las palmeras muertas por el contrario, al haber cesado su actividad biológica y flujo de savia debido a la intervención del picudo y sus larvas, han perdido su capacidad de regulación volviéndose un reflejo de su entorno, adoptando las temperaturas ambientales de forma más pronunciada. Las palmeras afectadas sin embargo, presentan una inercia térmica más impredecible, debido a que depende de su grado de infestación. Esto produce cambios de temperatura inesperados dependiendo del grado de infestación y del ciclo del día y noche.
3. **Variabilidad diurna y nocturna:** Se anticipa que las imágenes tomadas durante el día presenten mayor variabilidad en los resultados debido a los reflejos y a las fluctuaciones térmicas ocasionadas por la radiación solar. En contraste, se espera que las imágenes nocturnas ofrezcan una mayor estabilidad y consistencia en las mediciones, lo cual debería traducirse en un mejor desempeño de los modelos en dicho escenario. No obstante, debe considerarse que un conjunto de datos que mezcle indiscriminadamente capturas diurnas y nocturnas podría introducir un alto nivel de inconsistencia, ya que el comportamiento térmico de las palmeras difiere naturalmente entre estas etapas del ciclo diario.
4. **Limitación visual:** Se prevé que los modelos basados exclusivamente en información visual (texturas y formas) enfrenten mayores dificultades para discriminar entre palmeras sanas y afectadas, dado que la plaga no siempre produce síntomas visibles en las fases iniciales. En cambio, la identificación

de palmeras muertas se considera más sencilla por sus rasgos externos notorios.

5. **Cogollo como zona crítica:** Respecto a las temperaturas promedios de diferentes zonas de interés, el cogollo resultará más determinante para la clasificación que el uso de regiones más amplias que incluyan tronco u hojas, ya que es en esta zona donde la plaga genera las primeras alteraciones fisiológicas.
6. **Reducción de clases:** La reducción artificial del problema a dos clases (por ejemplo *sana* y *afectada* o *muerta* y *afectada*) podría ofrecer métricas de desempeño aparentemente más altas al simplificar la tarea. Sin embargo, este planteo no sería realista para un uso práctico, pues conduciría a una clasificación errónea de la categoría *muerta*, cuya identificación resulta fundamental para una gestión adecuada de la plaga.

5.1.2. Dataset y etapas empleadas

Se trabajó únicamente con las imágenes térmicas capturadas en Montevideo (entorno urbano), abarcando franjas diurnas y nocturnas. El conjunto global recolectado es de 489 fotos térmicas utilizables, donde 411 forman parte del conjunto nocturno y 78 son del conjunto de datos diurno (véase 4.1).

El proceso experimental se estructuró en tres etapas sucesivas, definidas en función del volumen de datos disponible en el momento y del grado de madurez de los modelos aplicados. Cada etapa representa una ampliación tanto en la cantidad de imágenes como en el alcance analítico de las pruebas realizadas.

- **Etapa inicial:** conformada por un total de 60 imágenes térmicas diurnas. Su propósito principal fue evaluar la factibilidad de utilizar capturas realizadas bajo iluminación solar directa, observando el impacto de los reflejos y del calentamiento ambiental sobre la medición térmica. En esta fase no se implementaron modelos clasificadores; el análisis se centró exclusivamente en las clases muerta y sana sobre las temperaturas promedio de la zona del cogollo y la diferencia de la misma con la temperatura ambiente registrada en campo.
- **Etapa intermedia:** En esta instancia se ampliaron los experimentos a 78 imágenes diurnas y a un conjunto nocturno que inicialmente contaba con 87 imágenes, el cual fue incrementándose progresivamente hasta alcanzar 228 capturas en el último experimento de la etapa. El objetivo fue determinar qué momento de captura y región de interés ofrecía los resultados más consistentes y precisos.
- **Etapa final:** basada en la integración de todo el conocimiento previo, esta etapa consolidó el conjunto de datos completo, manteniendo las 78 imágenes diurnas e incrementando el total de imágenes nocturnas hasta 411. En esta fase se introdujo la evaluación sistemática de modelos clasificadores

C0 (véase 4.3.2), C1 (véase 4.3.3), C2 (véase 4.3.4), C3 (véase 4.3.5) C4 (véase 4.3.6) sobre subconjuntos de imágenes nocturnas. Además, se incorporó la optimización automática de hiperparámetros para cada modelo, junto con aumento de datos sobre el modelo C3. Los experimentos de esta etapa buscaron validar la robustez de diversos modelos optimizados frente a un conjunto final de datos, más amplio y heterogéneo.

Este diseño escalonado permitió avanzar desde pruebas exploratorias sobre la viabilidad del uso de imágenes térmicas diurnas hasta la construcción de un sistema híbrido optimizado y validado sobre un volumen considerable de datos, asegurando la trazabilidad metodológica de los resultados obtenidos.

5.1.3. Métricas de evaluación

Se utilizaron dos familias de métricas:

1. **Métricas térmicas:** (i) *Temperatura promedio* dentro de las regiones de interés; (ii) *Diferencia térmica* respecto de la temperatura ambiente ($\Delta T = T_{\text{amb}} - T_{\text{BB}}$). Estas métricas permiten contrastar cuantitativamente las distribuciones por clase mediante diagramas de caja (boxplots) e histogramas.
2. **Métricas de desempeño de modelos:** Para clasificadores, *Precision* (véase 3.8.2), *Recall* (véase 3.8.3), *F1-score* (véase 3.8.4), *mAP* (véase 3.8.6), *Accuracy* (véase 3.8.7), *F1 Macro*, *Micro* y *Weighted* (véase 3.8.5), *Balanced Accuracy* (véase 3.8.8), *Cohen's Kappa* (véase 3.8.9), *Log Loss* (véase 3.8.11), *MCC* (véase 3.8.10) y *ROC-AUC* (véase 3.8.12). En esta sección se reportan principalmente los resultados térmicos; las métricas del detector se discuten donde corresponde.

5.1.4. Control de variables y supuestos

- **Cálculo de temperatura promedio:** se evalúa la temperatura promedio por sí sola y además considerando su diferencia con la temperatura ambiente.
- **Conjuntos de datos:** se separan análisis diurnos y nocturnos por sus efectos térmicos diferenciales. Las dimensiones del conjunto de datos varían según la etapa del proyecto, alcanzando nuevos resultados.
- **Región medida:** la temperatura se calcula sobre el *bounding box* detectado; por hipótesis, este suele incluir el cogollo y área circundante, pero no se restringe exclusivamente a él. Se consideran, por un lado, regiones centradas en el tronco, y por otro, el marcado del cogollo con extensión hacia las hojas.
- **Clasificadores:** se analiza el uso de diversos clasificadores para determinar cual posee mejor rendimiento.

- **Hiperparámetros:** se contempla la variación de hiperparámetros tanto en clasificadores tradicionales como en redes neuronales, ya sea mediante configuraciones predefinidas (por ejemplo, número de árboles en Random Forest, tasa de aprendizaje en CNN) como a través de optimización automática (Optuna). Este control permite evaluar hasta qué punto los resultados dependen de la configuración de entrenamiento y garantiza la reproducibilidad de los experimentos.

5.1.5. Procedimiento general

El flujo experimental siguió estos pasos:

1. **Preprocesamiento térmico** (requisito previo de todo el análisis): eliminación de artefactos (puntero, barra de colores, marcas de agua) mediante *inpainting* y recorte fijo; conversión de la escala de color a escala de grises sobre las imágenes térmicas.
2. **Detección de regiones de interés:** a partir de las imágenes térmicas, se obtienen las regiones de interés con YOLOv8. En todos los análisis térmicos, la temperatura se calcula sobre una región, pero se prueba en diversas regiones de la palmera.
3. **Cálculo de temperatura promedio:** para cada imagen, se computa la media térmica, excluyendo píxeles fríos asociados al cielo mediante un umbral de intensidad fijo. Además, se calcula ΔT restando la temperatura ambiente registrada en campo. Los resultados se vuelcan a un CSV de salida junto con la clase de la palmera.
4. **Análisis estadístico y visualizaciones:** con el CSV resultante, se generan *boxplots* por clase (temperatura promedio y ΔT) e *histogramas* de frecuencia. Estas visualizaciones permiten evaluar la hipótesis y comparar condiciones.
5. **Aumento de datos:** paso opcional donde se obtiene el CSV del paso anterior y se produce el aumento de datos para cada clase. El CSV resultante se utiliza como entrada de la etapa final.
6. **Entrenamiento y evaluación de clasificadores:** finalmente, las variables derivadas del análisis térmico (temperatura promedio y ΔT) se utilizan como insumos para distintos modelos de clasificación (umbral térmico simple, SVM, Random Forest y redes neuronales). Este paso constituye el núcleo de la validación experimental, ya que permite cuantificar qué tan bien los modelos predicen el estado sanitario de las palmeras. Para su evaluación se emplean métricas estandarizadas (véase [5.1.3](#)).

5.2. Etapas de experimentación

Con el fin de analizar de manera progresiva la evolución del sistema y validar la hipótesis en distintos contextos de captura, las pruebas se organizaron en tres etapas sucesivas: inicial, intermedia y final. Cada una se compone de una serie de experimentos numerados, los cuales reflejan la secuencia temporal de desarrollo, el incremento gradual del volumen de datos y la complejidad de los modelos evaluados.

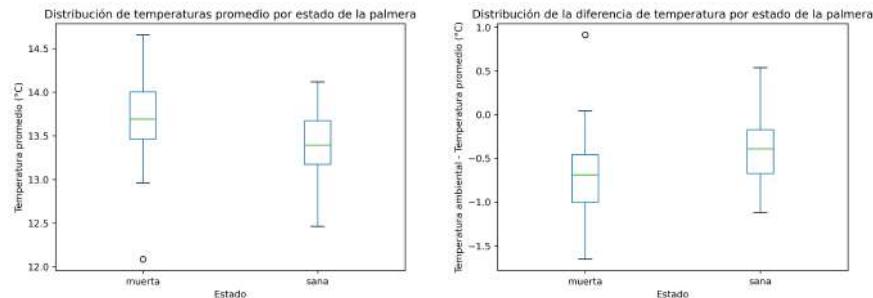
5.2.1. Etapa inicial

Expectativas

Se esperaba observar una alta variabilidad en las mediciones diurnas, principalmente por efectos de radiación solar y reflejos, lo cual serviría para establecer la necesidad de priorizar capturas nocturnas en etapas posteriores, más estables y consistentes en términos térmicos (Alaa y cols., 2020; Delalieux y cols., 2023).

Experimentos

1. **Experimento 1:** Cálculo de la temperatura promedio en regiones de interés (comparando las distribuciones térmicas entre palmeras sanas y muertas). Análisis de la diferencia térmica entre la temperatura promedio de la región de interés y la temperatura ambiente medida en campo. La Figura 5.1 muestra los resultados de la experimentación para esta etapa inicial.



(a) Distribución de temperatura promedio para las clases muerta y sana, donde se aprecia una tendencia a valores más altos en palmeras muertas respecto a las sanas. Se identificó un valor atípico en el grupo de palmeras muertas.

(b) Distribución de la diferencia de temperatura para palmeras sanas y muertas, donde se observa que las palmeras muertas indican un cogollo más cálido que el ambiente; en contraste, las sanas muestran una mediana menor, con un calentamiento menos marcado.

Figura 5.1: Distribuciones de temperatura en palmeras - Fase Inicial - Experimento 1

Resultados del cálculo de temperatura promedio: En la Figura 5.1a se puede apreciar el diagrama de cajas por temperatura de palmera. Se observó una diferencia térmica ligera entre algunas palmeras muertas y otras aparentemente sanas, lo cual respalda la hipótesis de diferenciación térmica (1). No obstante, los resultados obtenidos no permiten aún establecer umbrales confiables, debido a la cantidad limitada de ejemplos y a la variabilidad de las condiciones externas. Prosiguiendo con el análisis, las palmeras muertas presentan temperaturas promedio más elevadas que las palmeras sanas, incluso considerando que el conjunto de imágenes analizado es reducido y presenta incertidumbre en la clasificación. La mediana de las temperaturas correspondientes a palmeras muertas se sitúa por encima de la observada en palmeras sanas, lo que respalda dicha tendencia. Sin embargo, la diferencia entre medianas no puede considerarse estadísticamente significativa con base en la información disponible, ya que la dispersión térmica (rango intercuartílico) es relativamente baja y las muestras son escasas. También se identificó un valor atípico (*outlier*) en el grupo de palmeras muertas (alrededor de 12°C), asociado a un ejemplar sin follaje visible.

Resultados del cálculo de la diferencia térmica: En la Figura 5.1b se aprecia el diagrama de cajas por diferencia de temperatura entre el cogollo de la palmera y el ambiente. Las palmeras muertas muestran una mediana de -0.7 °C, sugiriendo que su cogollo es más caliente que el ambiente. En contraste, las palmeras sanas tienen una mediana de -0.4 °C, indicando que su cogollo también es más caliente pero en menor medida. Esta diferencia respalda la hipótesis de inercia térmica (2). Ambas clases muestran dispersión, siendo mayor para las muertas, con un rango intercuartílico más amplio y extremos o bigotes que alcanzan valores extremos.

Conclusiones de la Etapa

- La cantidad de datos de esta etapa no es suficiente para inferir conclusiones, generando así la necesidad de recabar más datos.
- Si bien esto sirve como un análisis de la factibilidad de los resultados futuros, la cantidad de clases utilizadas respalda lo esperado por la hipótesis de reducción de clases (6).

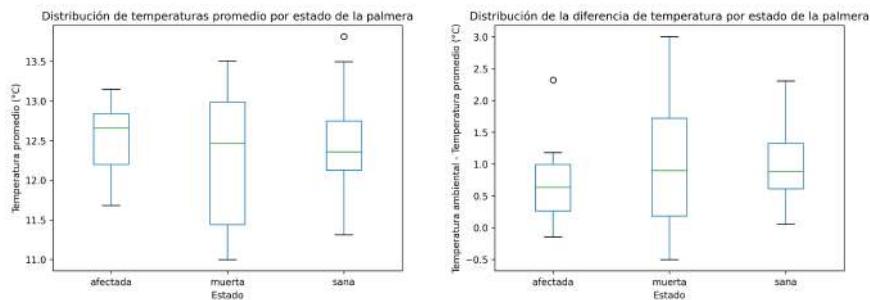
5.2.2. Etapa intermedia

Expectativas

Se esperaba que los modelos entrenados con imágenes nocturnas mostraran mayor estabilidad y desempeño que aquellos basados en imágenes diurnas, debido a la menor influencia de variaciones ambientales. Además, era de esperar que los experimentos en otras regiones de la palmera (hojas, tronco) aportaran información valiosa.

Experimentos

1. **Experimento 1:** En este experimento se analizaron conjuntamente los conjuntos de datos diurno y nocturno, considerando las tres clases disponibles: *sana*, *afectada* y *muerta*. Las distribuciones de temperatura promedio y de diferencia térmica se muestran en las Figuras (véase la Figura 5.2 y véase la Figura 5.3).

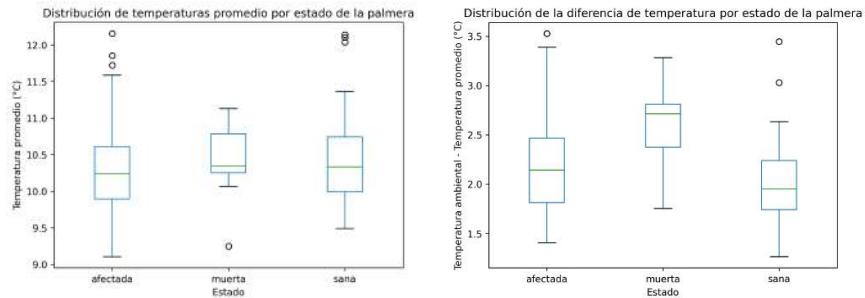


(a) Distribución de temperatura promedio por diurna para las tres clases. Se aprecia la mayor dispersión en la clase muerta y un valor atípico en la clase sana, indicando una palmera con cogollo más caliente que el resto.

(b) Distribución de la diferencia de temperatura diurna para las tres clases. La clase afectada presenta un valor atípico de una palmera más fría que el resto. La clase muerta presenta la mayor dispersión entre las tres clases.

Figura 5.2: Distribuciones de temperatura en palmeras durante el día - Fase intermedia - Experimento 1.

Resultados del cálculo de temperatura promedio: El análisis de las temperaturas promedio capturadas en condiciones diurnas permite observar tendencias térmicas que se correlacionan con el estado de salud de las palmeras. En la distribución de temperaturas promedio (véase la Figura 5.2a), se aprecia que las palmeras muertas presentan una dispersión térmica considerablemente mayor que las demás clases, mientras que las palmeras sanas exhiben un rango más acotado y concentrado. Las afectadas muestran un comportamiento intermedio, con temperaturas algo más



(a) Distribución de temperaturas por promedio en horario nocturno. Se aprecian valores atípicos en las tres clases y una menor dispersión en las temperaturas comparado con la distribución diurna.

(b) Distribución de la diferencia de temperatura nocturna para las tres clases. Se observan valores atípicos en las clases sana y afectada y una menor dispersión en todas las clases comparado a la distribución por diferencia diurna.

Figura 5.3: Distribuciones de temperatura en palmeras durante la noche - Fase intermedia - Experimento 1.

elevadas respecto a las sanas, lo que sugiere un incremento térmico asociado al daño fisiológico. Estas observaciones respaldan parcialmente la hipótesis vinculada a la inercia térmica (2), ya que no se tuvo una diferencia estadísticamente significativa en esas temperaturas.

En las capturas nocturnas, las distribuciones son más consistentes (véase la Figura 5.3a). Las palmeras sanas mantienen un rango de temperatura estable (aproximadamente entre 10.0 °C y 10.7 °C), con una mediana cercana a 10.3 °C. Las afectadas presentan una mediana ligeramente menor (10.2 °C) y un rango intercuartílico (IQR) entre 9.8 °C y 10.6 °C, junto a varios valores atípicos por encima de 11.5 °C. Las palmeras muertas muestran una mediana de 10.4 °C, un IQR más amplio (9.9 °C a 10.7 °C) y algunos outliers que superan 12.0 °C. La similitud general entre categorías indica que, aunque existen diferencias, estas no son significativamente marcadas, posiblemente debido a condiciones térmicas ambientales homogéneas.

Resultados del cálculo de la diferencia térmica: El análisis de la diferencia entre la temperatura ambiental y la temperatura promedio de las palmeras (véase la Figura 5.2b) muestra que las palmeras sanas presentan una mediana de aproximadamente 0.85 °C, con un rango intercuartílico entre 0.6 °C y 1.3 °C. Esto implica que, en promedio, el cogollo de las palmeras sanas se mantiene entre 0.6 °C y 1.3 °C más frío que el ambiente, con una variabilidad que oscila entre 0.05 °C y 2.3 °C.

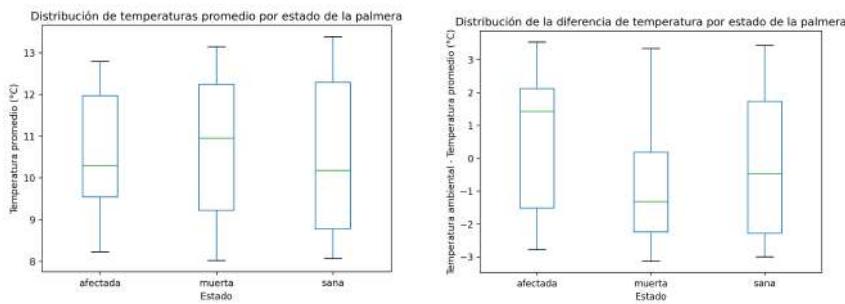
Por su parte, las palmeras afectadas y muertas tienden a mostrar diferencias de temperatura más pronunciadas respecto al ambiente. El cogollo

de las afectadas se mantiene, en promedio, más cálido que el de las sanas, mientras que las muertas alcanzan los valores de temperatura más elevados. Este comportamiento respalda a la hipótesis basada en la diferenciación térmica (1), que asocia el incremento térmico a la pérdida de capacidad de transpirar y al deterioro del flujo de savia, reduciendo la eficiencia de disipación del calor y generando un sobrecalentamiento en la copa, especialmente en el cogollo.

Si bien las diferencias entre sanas y afectadas no son drásticas, la mayor variabilidad observada en las muertas concuerda parcialmente con la hipótesis 1 mencionada anteriormente. En las sesiones nocturnas (véase la Figura 5.3b), el comportamiento se invierte, las palmeras afectadas exhiben una mayor diferencia térmica con el ambiente, mientras que las sanas se mantienen más estables. Las muertas presentan una mediana de aproximadamente 2.7 °C —la más alta de todas—, con un rango de 2.3 °C a 2.8 °C y sin presencia de valores atípicos. Este resultado respalda a la hipótesis vinculada a la inercia térmica (2), indicando que la capacidad de regulación térmica disminuye progresivamente con el deterioro fisiológico.

En general, las mediciones nocturnas se destacan por su mayor estabilidad y mejor separación entre categorías, lo que reduce inconsistencias y refuerza la validez de la hipótesis de variabilidad diurna y nocturna (3).

2. **Experimento 2:** En este experimento se incorporaron nuevas imágenes y se trabajó exclusivamente con el conjunto de datos nocturno, con el objetivo de evaluar su desempeño en contraste con el conjunto diurno utilizado previamente. La detección de la región del cogollo se realizó de forma automática mediante el modelo *YOLOv8* previamente entrenado, reemplazando el proceso manual empleado en la fase anterior. Los resultados obtenidos se presentan en las Figuras (véase la Figura 5.4a) y (véase la Figura 5.4b).



(a) Distribución de temperaturas promedio para las tres clases. Se observa una mayor dispersión en las temperaturas de todas las clases debido a la incorporación de nuevos datos.

(b) Distribución de la diferencia de temperatura por estado de la palmera. Se observa que las palmeras muertas son las que poseen la menor dispersión entre las tres clases y su temperatura en la zona del cogollo es la más caliente.

Figura 5.4: Comparación de distribuciones de temperatura promedio y diferencia de temperatura por estado de la palmera — Fase intermedia — Experimento 2.

Resultados del cálculo de temperatura promedio: Las palmeras afectadas y muertas presentaron temperaturas promedio ligeramente superiores a las de las palmeras sanas (véase la Figura 5.4a). Las medianas observadas se ubicaron aproximadamente en 10.5 °C para las afectadas, 11 °C para las muertas y 10 °C para las sanas, con una dispersión relativamente similar entre grupos. Si bien las diferencias no son marcadas, se identifica una tendencia creciente de temperatura en función del nivel de deterioro, coherente con la hipótesis basada en la diferenciación térmica (1).

La superposición parcial de los rangos intercuartílicos y la similitud en la dispersión sugieren, sin embargo, que las diferencias observadas no son estadísticamente significativas.

Resultados del cálculo de la diferencia térmica: El análisis de la diferencia entre la temperatura ambiental y la temperatura promedio

del cogollo (véase la Figura 5.4b) muestra un comportamiento particular. Las palmeras afectadas presentan valores positivos cercanos a 1.5 °C, lo que indica que su temperatura promedio es menor que la ambiental (el cogollo se mantiene más frío que el aire). En cambio, las palmeras sanas y muertas exhiben valores negativos, de aproximadamente -0.5 °C y -1.5 °C respectivamente, lo que refleja cogollos más calientes que el ambiente.

Este resultado contradice parcialmente lo planteado en la hipótesis de diferenciación térmica (1), según la cual las palmeras afectadas deberían presentar temperaturas más altas que las sanas. No obstante, tanto las palmeras sanas como las muertas sí se comportan conforme a dicha hipótesis. El patrón observado en las afectadas —con valores positivos más amplios y una dispersión notoria— podría asociarse a la variabilidad fisiológica interna dentro de la categoría, dependiente del grado de infestación y de la fase de avance de la plaga. Este fenómeno coincide con lo planteado en la hipótesis de inercia térmica (2), que sugiere que la respuesta térmica de las palmeras afectadas puede ser inconsistente debido a la progresiva pérdida de regulación térmica.

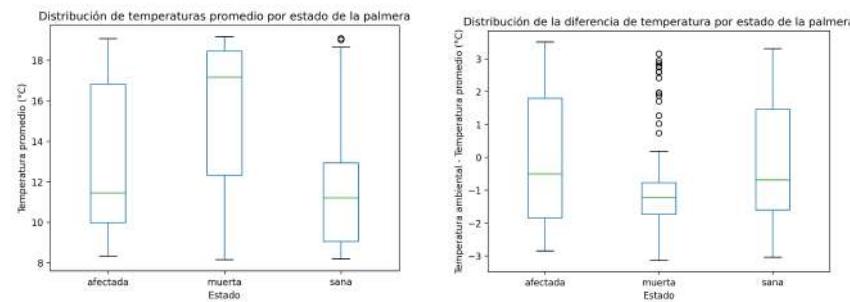
Este experimento permitió validar la eficacia del procedimiento automático de detección del cogollo y, además, confirmó que las mediciones nocturnas tienden a ofrecer mayor estabilidad térmica y mejor diferenciación entre clases.

3. **Experimento 3:** Este experimento utilizó el mismo conjunto de datos del experimento anterior, pero el análisis se centró en determinar si la variación en las regiones de interés aportaba información térmica adicional. Para ello, se entrenaron tres modelos independientes, cada uno enfocado en una región específica: cogollo, tronco y hojas. Los resultados permiten contrastar el poder discriminante de cada zona en relación con las hipótesis previamente planteadas.

Región analizada: Cogollo

El análisis de esta región se corresponde con el ya realizado en el Experimento 2 de esta misma etapa.

Regiones analizadas: Cogollo y tronco, con mayor énfasis en el segundo.



(a) Distribución de temperaturas promedio por estado de la palmera — cogollo y tronco. Se observa un valor atípico en la clase sana y gran dispersión en las clases afectada y muerta.

(b) Distribución de diferencia de temperaturas — cogollo y tronco. Es de destacar la gran cantidad de valores atípicos en la clase muerta, así como su baja dispersión.

Figura 5.5: Comparación de distribuciones de temperatura promedio y diferencia de temperatura por estado de la palmera en cogollo y tronco — Fase intermedia — Experimento 3.

Resultados del cálculo de temperatura promedio: En la región del cogollo y tronco (véase la Figura 5.5a), las palmeras muertas presentan las temperaturas promedio más altas, con una mediana claramente superior a las de las sanas y afectadas. Estas dos últimas muestran medianas similares y una superposición considerable de sus rangos intercuartílicos, observándose además un valor atípico alto en la clase sana.

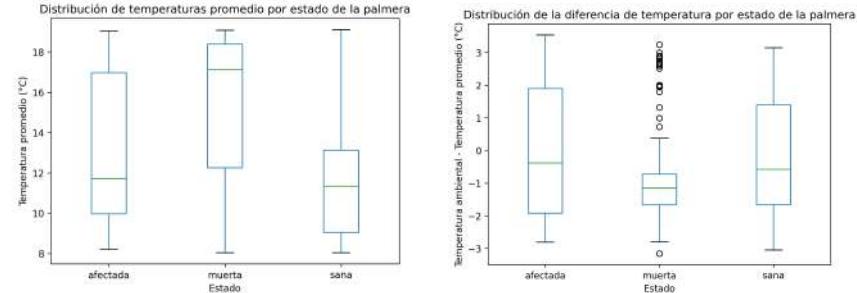
Este patrón respalda la hipótesis de diferenciación térmica (1), que anticipaba que las afectadas serían más cálidas que las sanas. Si bien sus medianas son similares, la dispersión presentada en la clase afectada marca claramente una diferencia significativa entre ambas. En este caso, la

categoría que se separa hacia valores más altos es la de palmeras muertas. Asimismo, el aumento de superposición entre categorías al ampliar la región de análisis refuerza la hipótesis basada en la relevancia del cogollo como zona crítica (5), ya que sugiere que el cogollo aislado ofrece mayor capacidad de discriminación que el cogollo y tronco o las áreas extendidas.

Resultados del cálculo de la diferencia térmica: En este análisis (véase la Figura 5.5b), se observa que las palmeras sanas poseen una mediana negativa, lo que indica que su temperatura promedio es mayor que la ambiental. Las muertas presentan una mediana próxima a cero (temperatura similar al entorno), mientras que las afectadas exhiben una mediana ligeramente negativa, acompañada de una mayor dispersión.

Este comportamiento respalda a la hipótesis de inercia térmica (2), ya que las palmeras sanas mantienen temperaturas más estables y ligeramente superiores al ambiente, reflejando su mayor inercia térmica. Las muertas, al haber perdido la capacidad de regulación, tienden a ajustarse a la temperatura ambiental, mientras que las afectadas muestran variabilidad térmica acorde a un estado intermedio de deterioro fisiológico.

Regiones analizadas: Cogollo y hojas (región extendida hacia hojas).



(a) Distribución de temperaturas promedio por estado de la palmera — cogollo y hojas. Resultados extremadamente similares a los del experimento anterior, salvo por un valor atípico en la clase sana.

(b) Distribución de diferencia de temperaturas — cogollo y hojas. Resultados extremadamente similares a los del experimento anterior a excepción por un valor atípico en la clase muerta.

Figura 5.6: Comparación de distribuciones de temperatura promedio y diferencia de temperatura por estado de la palmera en el cogollo y las hojas — Fase intermedia — Experimento 3.

Resultados del cálculo de temperatura promedio: En este análisis las palmeras muertas vuelven a presentar las temperaturas promedio más altas, con una mediana superior a las sanas y afectadas. En general, la Figura 5.6a coincide prácticamente en su totalidad con la Figura 5.5a salvo por un pequeño valor atípico en la clase sana, lo cual no garantiza suficiente evidencia estadística para preferirla sobre la zona del cogollo.

Resultados del cálculo de la diferencia térmica: Al igual que para el promedio, en este caso la Figura 5.6b es prácticamente igual a la 5.5b ya analizada, salvo por un valor atípico en la clase muerta.

Conclusiones de la Etapa

- Las mediciones nocturnas ofrecen un escenario más estable y adecuado para distinguir entre las tres clases.
- La disponibilidad de un mayor volumen de imágenes contribuye de manera significativa en el análisis de resultados.
- Al ampliar la región hacia el tronco o las hojas, el contraste entre las clases se redujo, la superposición entre grupos aumentó. Por tanto, los modelos que focalizan el análisis en el cogollo conservan mayor poder discriminante.

En conjunto, los resultados de la etapa respaldan las hipótesis de diferenciación térmica (1), inercia térmica (2), variabilidad diurna y nocturna (3) y cogollo como zona crítica (5).

5.2.3. Etapa final

Expectativas

Posterior a la extensión del conjunto de datos y una distribución de temperatura con variaciones más notorias (véase el Anexo A), se esperaba que esta etapa final permitiera evidenciar de forma sólida sobre qué tipo de enfoque (térmico, visual o híbrido) resulta más adecuado para la detección temprana de palmeras afectadas. En particular, se anticipaba que los modelos C1 (clasificador con Machine Learning) y C2 (clasificador híbrido con CNN y Machine Learning) presentaran mejoras significativas respecto al clasificador aleatorio (C0), gracias a la inclusión de descriptores térmicos y representaciones profundas extraídas mediante redes convolucionales. Asimismo, se esperaba que el modelo C3 (clasificador basado en CNN entrenada de extremo a extremo) mostrara un aprendizaje más especializado ante la clase muerta pero peor hacia la clase afectada. Además, la incorporación de aumento de datos podría contribuir a una mayor estabilidad y capacidad de generalización. Finalmente, el modelo híbrido C4 (clasificador híbrido con optimización y reentrenamiento) debía alcanzar el mejor equilibrio entre precisión y robustez, beneficiándose de la combinación ponderada de probabilidades entre redes convolucionales y clasificadores tabulares.

La inclusión de la optimización automática de hiperparámetros tomaría el papel de mejorar la estabilidad y consistencia de los resultados, reduciendo la dependencia de configuraciones manuales.

Dentro de las métricas obtenidas en cada modelo, consideramos que **Weighted Avg F1-score** y **Recall** son las más relevantes. Por un lado, el Weighted Avg F1-score nos proporciona una visión global del desempeño del modelo, teniendo en cuenta el desbalance de las clases como también una combinación entre Precision y Recall. El Recall o Exhaustividad cobra mucha importancia en nuestro proyecto, ya que un modelo con bajo Recall implica una gran cantidad de falsos negativos. Para el caso de la clase afectada, un alto número de falsos negativos se traduce en que el modelo clasifica erróneamente palmeras que sí estaban enfermas. Esto constituye el error de mayor costo, ya que implica dejar focos de plaga sin tratar, permitiendo su propagación.

Experimentos

1. **Experimento 1:** el primer experimento consistió en la ejecución del clasificador C0, un modelo aleatorio sin entrenamiento previo, diseñado para establecer una línea base de desempeño. Este modelo selecciona aleatoriamente una clase entre las tres posibles (*sana*, *afectada*, *muerta*), sin utilizar información térmica ni visual. Su objetivo fue determinar el nivel de precisión esperable por azar y proporcionar un punto de referencia para los clasificadores subsiguientes.

Clase	Precision	Recall	F1-score	Support
Afectada	0.40	0.29	0.33	28
Muerta	0.30	0.32	0.31	31
Sana	0.26	0.33	0.29	21
Accuracy			0.31	80
Macro Avg	0.32	0.31	0.31	80
Weighted Avg	0.33	0.31	0.31	80

Tabla 5.1: Reporte de clasificación del modelo C0 (véase Etapa Final - Experimento 1).

Resultados: Los resultados obtenidos reflejan un comportamiento puramente aleatorio, con una precisión promedio de 31 % y valores de *F1-score* y *recall* similares (véase la Tabla 5.1). Ninguna clase presentó un desempeño superior al resto, lo que confirma que el modelo no logra identificar patrones diferenciables en los datos. En términos prácticos, este resultado representa el nivel mínimo de desempeño que cualquier modelo entrenado debe superar para considerarse informativo.

El modelo aleatorio constituye una referencia metodológica necesaria. Su bajo rendimiento confirma la dificultad del problema de clasificación multicategoría y la importancia de incorporar características térmicas o visuales para alcanzar resultados significativamente mejores. En el contexto del proyecto, un desempeño de este nivel sería inaceptable en una aplicación real, ya que implicaría la clasificación errática de palmeras afectadas, con el riesgo de omitir focos activos de la plaga.

2. **Experimento 2:** el segundo experimento corresponde al modelo C1, implementado mediante algoritmos de *Machine Learning* tradicionales entrenados sobre información tabular derivada de las imágenes térmicas. Este experimento permite analizar la efectividad de las variables estadísticas como la temperatura promedio del cogollo, la media y varianza de los píxeles válidos, y los valores extremos (mínimo y máximo) dentro del área de interés. Dichas variables fueron normalizadas y empleadas como insumo para clasificadores supervisados, entre los cuales se evaluaron un *Support Vector Machine* (SVM) y un *Random Forest*.

Clase	Precision	Recall	F1-score	Support
Afectada	0.63	0.68	0.66	28
Muerta	0.77	0.77	0.77	31
Sana	0.58	0.52	0.55	21
Accuracy			0.68	80
Macro Avg		0.66	0.66	80
Weighted Avg		0.67	0.68	80

Tabla 5.2: Reporte de clasificación del modelo C1 (véase Etapa Final - Experimento 2).

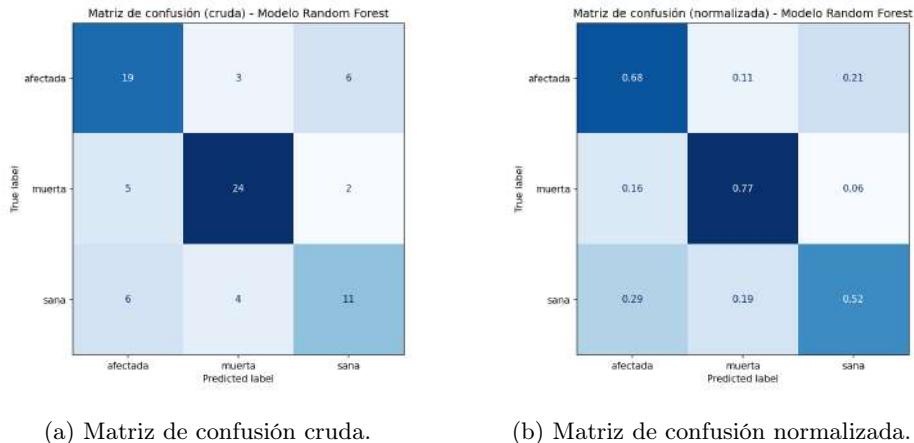


Figura 5.7: Matrices de confusión del modelo C1 — Random Forest.

Resultados: Durante la etapa de validación, el modelo Random Forest mostró un desempeño claramente superior al SVM, tanto en estabilidad como en capacidad de generalización. En consecuencia, este modelo fue seleccionado para la evaluación final sobre el conjunto de prueba.

Mientras que el modelo SVM alcanzó en validación una exactitud del 49 %, con valores de F1-score relativamente bajos y equilibrados entre clases, sugiriendo una capacidad de separación limitada. Los resultados obtenidos por Random Forest (véase la Tabla 5.2) evidencian un comportamiento equilibrado entre las tres clases, alcanzando una exactitud global de **0.68** y un *F1-score* ponderado de **0.67**. La clase *muerta* obtuvo la mejor discriminación, con valores de *precision* y *recall* de 0.77, lo que indica una adecuada capacidad del modelo para reconocer ejemplares en estado terminal. Por otro lado, las clases *afectada* y *sana* presentan ligeras confusiones entre sí, reflejadas en la matriz de confusión (Figuras 5.7a y 5.7b). En particular, el modelo tiende a clasificar erróneamente algunas palmeras afectadas como sanas, lo que implica un riesgo práctico, ya que estos casos representan focos potenciales de propagación no detectados.

Este experimento demuestra que las características térmicas extraídas de las imágenes infrarrojas contienen información discriminativa suficiente para distinguir entre estados sanitarios de las palmeras. Si bien no alcanza una separación perfecta, logra capturar tendencias térmicas coherentes con la hipótesis de diferenciación térmica (1), confirmando la validez del uso de modelos basados en aprendizaje supervisado clásico cuando se dispone de un conjunto de variables térmicas adecuadamente procesadas.

3. **Experimento 3:** el tercer experimento evalúa el modelo C2, cuyo propósito fue evaluar si la integración de información térmica y visual mejoraba la capacidad de separación entre clases respecto a los modelos puramente tabulares. En este experimento se combinaron características extraídas mediante redes convolucionales preentrenadas (ResNet18, ResNet50 y DenseNet121) con variables térmicas provenientes de los datos de entrada. Cada arquitectura se utilizó como extractor de características fijas, alimentando posteriormente a Support Vector Machine y Random Forest.

Clase	Precision	Recall	F1-score	Support
Afectada	0.63	0.43	0.51	28
Muerta	0.70	0.84	0.76	31
Sana	0.58	0.67	0.62	21
Accuracy			0.65	80
Macro Avg	0.64	0.64	0.63	80
Weighted Avg	0.65	0.65	0.64	80

Tabla 5.3: Reporte de clasificación del modelo C2 (véase Etapa Final - Experimento 3).

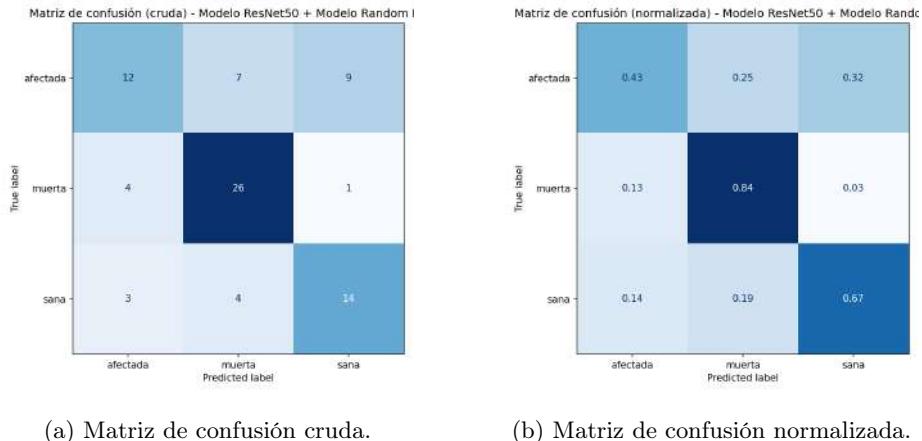


Figura 5.8: Matrices de confusión del modelo C2 — CNN + Random Forest.

Resultados: Durante la etapa de validación, los resultados mostraron un rendimiento variable según la red utilizada. Entre las combinaciones evaluadas, el mejor desempeño se obtuvo con la pareja ResNet50 y Random Forest alcanzando una exactitud global de 0.65 y un F1-score ponderado de 0.64 (véase la Tabla 5.3). La clase *muerta* fue la más consistente,

con un recall de 0.84, lo que refleja la facilidad para identificar patrones visuales y térmicos asociados a estados avanzados de deterioro. En contraste, la clase *afectada* mostró un menor equilibrio entre precisión (0.63) y recall (0.43), lo que indica que el modelo sigue presentando dificultades para detectar las primeras etapas de infestación.

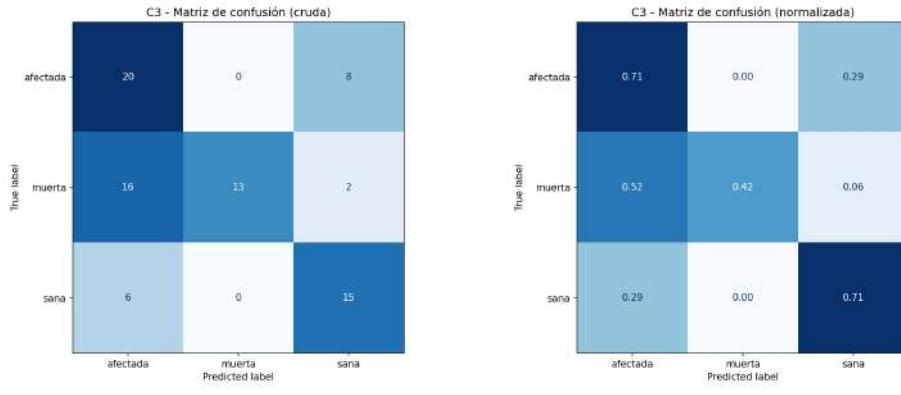
Las Figuras 5.8a y 5.8b presentan las matrices de confusión cruda y normalizada, respectivamente. En ellas se observa que parte de los ejemplares afectados son clasificados como sanos o muertos, lo que sugiere que, si bien el modelo capta información térmica útil, las diferencias visuales entre estados intermedios aún son sutiles para el clasificador.

El experimento C2 evidencia que la fusión de información visual y térmica puede ser una estrategia prometedora, pero su desempeño depende en gran medida de la calidad y cantidad de datos disponibles. El modelo logra mantener una buena estabilidad general y refleja el potencial de integrar características derivadas de CNNs con descriptores físicos interpretables. Aun así, la baja sensibilidad en la clase *afectada* sugiere que las características extraídas de las imágenes térmicas en escala de grises no capturan plenamente las señales tempranas de infestación, posiblemente debido a la homogeneidad visual de los ejemplares y a las limitaciones del conjunto de entrenamiento.

4. **Experimento 4:** el cuarto experimento corresponde al modelo C3, una red neuronal convolucional entrenada de extremo a extremo directamente sobre las imágenes térmicas en escala de grises. En este caso, la red convolucional aprende directamente las representaciones necesarias para la clasificación sin recurrir a descriptores tabulares.

Clase	Precision	Recall	F1-score	Support
Afectada	0.48	0.71	0.57	28
Muerta	1.00	0.42	0.59	31
Sana	0.60	0.71	0.65	21
Accuracy			0.60	80
Macro Avg	0.69	0.62	0.60	80
Weighted Avg	0.71	0.60	0.60	80

Tabla 5.4: Reporte de clasificación del modelo C3 (véase Etapa Final - Experimento 4).



(a) Matriz de confusión cruda.

(b) Matriz de confusión normalizada.

Figura 5.9: Matrices de confusión del modelo C3 — CNN entrenada de extremo a extremo.

Resultados El modelo alcanzó una exactitud global de 0.60 y un F1-score ponderado de 0.60 (véase la Tabla 5.4). Si bien los resultados se encuentran por debajo de los obtenidos por modelos con información tabular, el desempeño demuestra una capacidad razonable de aprendizaje visual a partir de imágenes térmicas.

La clase *muerta* fue la más sencilla de identificar, lo cual coincide con la hipótesis de limitación visual (4), dado que las palmeras en este estado

presentan rasgos visuales distintivos y contrastes térmicos más pronunciados. En cambio, las clases *afectada* y *sana* mostraron mayor superposición en las predicciones, reflejando las dificultades para detectar síntomas tempranos únicamente a partir de patrones visuales.

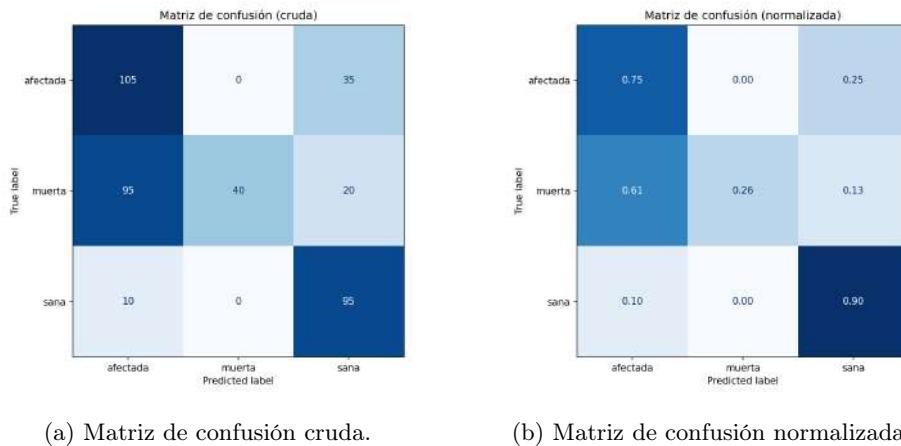
Las matrices de confusión cruda y normalizada (Figuras 5.9a y 5.9b) confirman esta tendencia: el modelo confunde con frecuencia ejemplares *afectados* con *sanos*, aunque logra mantener una cobertura aceptable de la clase *muerta*. La distribución de errores sugiere que el modelo aprende representaciones coherentes, pero no suficientemente específicas para capturar los cambios térmicos sutiles asociados a la infestación inicial.

El modelo C3 demuestra la viabilidad de emplear redes convolucionales entrenadas directamente sobre imágenes térmicas, sin necesidad de atributos complementarios. Sin embargo, su desempeño revela los límites de un enfoque puramente visual en este tipo de problema: las diferencias entre palmeras sanas y afectadas no siempre se traducen en contrastes visuales evidentes dentro del rango infrarrojo. Aun así, el comportamiento observado indica que la red no incurre en sobreajuste significativo, manteniendo una generalización estable en el conjunto de prueba. Este modelo sirvió como base para el experimento siguiente, donde se introdujo aumento de datos para evaluar su impacto en la capacidad de generalización.

5. **Experimento 5:** el quinto experimento corresponde a una extensión del modelo C3, en la que se aplicaron técnicas de aumento de datos con el objetivo de incrementar la capacidad de generalización del clasificador y reducir la sensibilidad a variaciones menores en las imágenes térmicas. El conjunto de datos fue quintuplicado mediante transformaciones geométricas simples, que incluyeron rotaciones aleatorias y pequeños desplazamientos de recorte, preservando la coherencia térmica de las regiones de interés.

Clase	Precision	Recall	F1-score	Support
Afectada	0.50	0.75	0.60	140
Muerta	1.00	0.26	0.41	155
Sana	0.63	0.90	0.75	105
Accuracy			0.60	400
Macro Avg		0.71	0.64	0.59
Weighted Avg		0.73	0.60	0.56

Tabla 5.5: Reporte de clasificación del modelo C3 con aumento de datos (véase Etapa Final - Experimento 5).



(a) Matriz de confusión cruda. (b) Matriz de confusión normalizada.

Figura 5.10: Matrices de confusión del modelo C3 con aumento de datos.

Resultados: El modelo alcanzó una exactitud global de 0.60 y un F1-score ponderado de 0.56 (véase la Tabla 5.5). A pesar de que los valores numéricos no mejoraron de forma significativa respecto al experimento previo, se observó una mayor estabilidad en las métricas y una reducción en la dispersión de los errores por clase. La clase *sana* obtuvo los mejores resultados, con un *recall* de 0.90 y un *precision* de 0.63, lo que indica que

el modelo logró reconocer con mayor seguridad los ejemplares sin signos de infestación. Por su parte, la clase *afectada* presentó una sensibilidad más alta (0.75) pero una precisión limitada (0.50), lo que evidencia una tendencia a clasificar erróneamente otras clases como afectadas. La clase *muerta*, en cambio, mostró un descenso en la sensibilidad, probablemente asociado a la sobrerepresentación del resto de las clases tras el aumento.

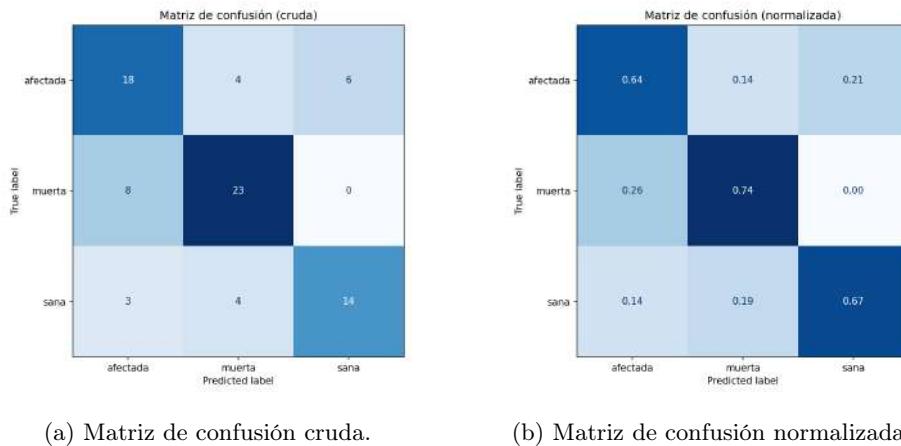
Las matrices de confusión cruda y normalizada (Figuras 5.10a y 5.10b) reflejan una distribución de errores más homogénea en comparación con el modelo original, lo que sugiere una mejora en la consistencia global del aprendizaje. No obstante, la confusión entre *afectada* y *muerta* persiste, lo que refuerza la idea de que las diferencias térmicas entre estos estados son más sutiles de lo que la red puede captar sin información tabular complementaria.

El aumento de datos permitió reforzar la robustez del modelo frente a pequeñas variaciones de captura y mejorar la estabilidad de sus resultados, aunque sin traducirse en un incremento sustancial del rendimiento global. Este comportamiento sugiere que el desempeño de la red no estaba limitado por la cantidad de ejemplos, sino por la complejidad intrínseca del problema y la similitud térmica entre clases. Aun así, este experimento confirmó que la estrategia de aumento no introduce sobreajuste ni degradación del rendimiento, sentando una base sólida para su integración en modelos híbridos más complejos.

6. **Experimento 6:** el sexto experimento corresponde al modelo C4, la versión más avanzada en la que se combinaron clasificadores térmicos y visuales. El enfoque se basó en una fusión ponderada de probabilidades entre un modelo Random Forest (entrenado sobre variables tabulares derivadas de las imágenes térmicas) y una ResNet18 adaptada a escala de grises (entrenada de extremo a extremo). Además, se incorporó un refuerzo explícito de la clase *afectada*, dada su importancia desde el punto de vista fitosanitario.

Clase	Precision	Recall	F1-score	Support
Afectada	0.62	0.64	0.63	28
Muerta	0.74	0.74	0.74	31
Sana	0.70	0.67	0.68	21
Accuracy			0.69	80
Macro Avg		0.69	0.68	80
Weighted Avg		0.69	0.69	80

Tabla 5.6: Reporte de clasificación del modelo C4 (véase Etapa Final - Experimento 6).



(a) Matriz de confusión cruda. (b) Matriz de confusión normalizada.

Figura 5.11: Matrices de confusión del modelo C4 — Híbrido optimizado.

Resultados: El modelo alcanzó una exactitud global de 0.69 y un F1-score ponderado de 0.69 (véase la Tabla 5.6). Las métricas complementarias, como *Balanced Accuracy* (0.68) y *Cohen's Kappa* (0.52), reflejan un nivel de acuerdo sustancial entre las predicciones del modelo y las etiquetas reales (véase la Tabla B.6 del Anexo). Las matrices de confusión cruda y normalizada (Figuras 5.11a y 5.11b) muestran una distribución

equilibrada de aciertos y errores, con mejoras en la clase *sana* respecto de experimentos previos. La clase *afectada*, aunque sigue siendo la más difícil de distinguir, presenta un rendimiento más estable y una reducción visible de falsos negativos.

El modelo C4 constituye la culminación del proceso experimental, integrando las ventajas de los modelos tabulares y visuales dentro de un esquema optimizado de fusión de decisiones. Los resultados evidencian un aumento de la estabilidad y coherencia del sistema, confirmando que la combinación de fuentes de información —temperatura promedio, características estadísticas y patrones visuales— ofrece una descripción más completa del estado fisiológico de la palmera. Asimismo, la inclusión de un refuerzo específico sobre la clase *afectada* permitió priorizar los casos de mayor interés agronómico, reduciendo el impacto de los errores más costosos.

Si bien el modelo no alcanza altos valores de desempeño, sus resultados se consideran robustos y consistentes con las condiciones experimentales reales, marcando un avance significativo respecto a las versiones anteriores. Este experimento consolida la validez del enfoque híbrido y sienta las bases para futuras mejoras basadas en adquisición de datos multiespectrales y expansión del conjunto de entrenamiento.

Conclusiones de la Etapa

- La información térmica se confirma como el descriptor más relevante para la detección de palmeras afectadas, especialmente en estados avanzados, donde las diferencias de temperatura respecto a ejemplares sanos resultan más pronunciadas.
- Los modelos como el C3 basados exclusivamente en información visual demostraron una capacidad limitada para identificar etapas tempranas de afectación, aunque resultan eficaces en la detección de palmeras muertas debido a su marcada diferenciación morfológica (véase hipótesis de limitación visual 4).
- La incorporación de aumento de datos en C3 mejoró la estabilidad del aprendizaje sin incrementar significativamente el rendimiento global, confirmando que el problema no reside en la cantidad de ejemplos, sino en la complejidad térmica y visual del fenómeno.
- Los clasificadores que integran información térmica tabular (C1 y C4) mostraron el mejor equilibrio entre precisión y robustez, siendo los más adecuados para un uso práctico orientado a la detección temprana (véase las Tablas 5.8 y 5.7).

Clase	C1			C4		
	Precision	Recall	F1-Score	Precision	Recall	F1-Score
Afectada	0.63	0.68	0.66	0.62	0.64	0.63
Muerta	0.77	0.77	0.77	0.74	0.74	0.74
Sana	0.58	0.52	0.55	0.70	0.67	0.68

Tabla 5.7: Comparación de métricas por clase entre clasificadores C1 y C4

Métrica	C1	C4
Accuracy	0.68	0.69
Balanced Accuracy	0.66	0.68
F1-Weighted Avg	0.67	0.69
Enfoque	Térmico	Híbrido

Tabla 5.8: Comparación de métricas entre clasificadores C1 y C4

- La optimización automática de hiperparámetros permitió, en base a la experimentación, alcanzar un comportamiento más consistente entre clases y una reducción significativa de los falsos negativos en la categoría *afectada*, lo cual adquiere especial importancia.

- La clase *afectada* continúa siendo la más desafiante para todos los modelos, reflejando la sutileza de las señales térmicas y visuales asociadas a las fases iniciales de la plaga (véase la Tabla 5.9).

Modelo	Precision	Recall	F1-score
C0	0.40	0.29	0.33
C1	0.63	0.68	0.66
C2	0.63	0.43	0.51
C3	0.48	0.71	0.57
C3-A	0.50	0.75	0.60
C4	0.62	0.64	0.63

Tabla 5.9: Métricas para la clase *Afectada* en los modelos C0–C4

- En términos generales, los resultados evidencian la pertinencia del enfoque híbrido y el valor de combinar distintas fuentes de información para representar de forma más completa el estado fisiológico de las palmeras.

5.3. Discusión general de resultados

El análisis integral de los resultados experimentales permite establecer una visión comparativa del desempeño de los distintos enfoques propuestos, así como evaluar la validez de las hipótesis planteadas en la Sección 5.1.1.

Desde una perspectiva metodológica, la evolución entre etapas refleja un aprendizaje progresivo en la formulación y refinamiento de los modelos. El proyecto transitó desde aproximaciones simples y altamente interpretables hacia configuraciones complejas y optimizadas, sin perder de vista la trazabilidad experimental. Este enfoque incremental permitió identificar las fuentes principales de error —reflejos, variabilidad ambiental y similitud térmica entre clases— y ajustar tanto la estrategia de captura como los procedimientos de entrenamiento.

Los resultados de la etapa intermedia confirmaron que las capturas nocturnas son preferibles para un análisis de temperatura interna debido a la regularización de las temperaturas, generando resultados más consistentes que las diurnas las cuales presentan intervención externa climática, como lo puede ser la reflexión solar. Además, la evidencia obtenida respalda que la zona del cogollo es el área de interés más relevante para la tarea de clasificación, ya que es donde se manifiestan de forma más crítica las alteraciones fisiológicas. La validación de muchas de nuestras hipótesis de trabajo, como lo son la diferenciación térmica por inercia y la relevancia crítica del cogollo confirma la solidez del marco teórico que guió este proyecto.

Los modelos de referencia (C0 y C1) evidenciaron la relevancia de la información térmica como fuente primaria para la discriminación entre clases. Mientras el clasificador aleatorio (C0) alcanzó resultados cercanos al azar, el modelo C1, basado en descriptores tabulares derivados de las imágenes térmicas, logró una mejora sustancial en todas las métricas. Este resultado confirma que las diferencias térmicas —especialmente en el cogollo— constituyen indicadores válidos del estado fisiológico de las palmeras y respaldan la hipótesis de diferenciación térmica (1).

En cuanto a los modelos híbridos (C2 y C4), su desempeño muestra que la combinación de información térmica y visual aporta una representación más rica y estable del fenómeno. En particular, el modelo C4 logró el mejor equilibrio global entre precisión, *recall* y consistencia, lo que sugiere que la fusión de decisiones entre clasificadores tabulares y convolucionales es una estrategia eficaz para enfrentar la heterogeneidad de las condiciones de captura. La optimización automática de hiperparámetros permitió además reducir la sensibilidad del sistema a configuraciones iniciales, reforzando la reproducibilidad del método.

Por su parte, los modelos puramente visuales (C3 y C3 con aumento de datos) evidenciaron las limitaciones de las CNN al operar sobre imágenes térmicas en ausencia de información tabular complementaria. Aunque fueron capaces de identificar correctamente las palmeras muertas —debido a su contraste visual evidente—, su capacidad para distinguir entre palmeras sanas y afectadas fue reducida. Este comportamiento valida la hipótesis relacionada a la limitación visual (4). No obstante, el aumento de datos demostró mejorar la estabilidad del

aprendizaje y reducir la varianza entre ejecuciones, indicando que la red es capaz de aprovechar mejor la diversidad de ejemplos cuando dispone de suficientes variaciones geométricas y contextuales.

Desde una perspectiva cuantitativa, las métricas incluidas en la Tabla 5.10 permiten caracterizar de forma integral el desempeño de cada modelo. El *Accuracy* y la *Balanced Accuracy* reflejan la proporción general de aciertos, evidenciando que los modelos híbridos (C2 y C4) mantienen un rendimiento estable. Las métricas basadas en la armonización entre precisión y exactitud (*F1 Macro* y *F1 Weighted*) confirman la capacidad de estos modelos para equilibrar la correcta detección de cada clase sin favorecer de forma excesiva a las más representadas. Por su parte, los coeficientes de concordancia (*Cohen's Kappa* y *MCC*) revelan un nivel de acuerdo moderado a sustancial entre predicciones y etiquetas reales, siendo C4 el modelo más consistente en este sentido. El *Log Loss* y las curvas *ROC-AUC* y *PR-AUC* aportan información complementaria sobre la calibración probabilística y la discriminación entre clases, indicadores clave en donde la certeza del modelo es tan importante como la predicción en sí. Finalmente, la métrica *Top-2 Accuracy* destaca la capacidad para incluir la clase correcta entre sus dos predicciones más probables, un aspecto valioso en escenarios donde las diferencias visuales o térmicas son sutiles.

En conjunto, y teniendo en prioridad las métricas como Recall y Weighted Avg F1-Score, los mejores modelos clasificadores fueron C1 y C4. Dentro de estos dos modelos, identificamos que el que obtuvo mejores resultados es el clasificador C4 debido a su estructura que representa la combinación de los enfoques de los modelos anteriores, obteniendo los valores más elevados en gran parte de las métricas utilizadas. No obstante, la mejoría es mínima y puede no ser justificable su complejidad si la comparamos con C1, el cual es significativamente más simple y rápido.

Métrica	C0	C1	C2	C3	C3+Aug	C4
Accuracy	0.31	0.68	0.65	0.60	0.60	0.69
Balanced Accuracy	0.31	0.66	0.64	0.62	0.64	0.68
F1 Macro Avg	0.31	0.66	0.63	0.60	0.59	0.69
F1 Weighted Avg	0.31	0.67	0.64	0.60	0.56	0.69
Cohen's Kappa	-	0.50	0.47	0.40	0.41	0.52
MCC	-	0.51	0.48	0.43	0.45	0.52
Log Loss	-	0.73	0.78	0.80	1.47	0.70
ROC-AUC (Macro)	-	0.85	0.84	0.84	0.83	0.86
PR-AUC (Macro)	-	0.75	0.71	0.72	0.71	0.75
Top-2 Accuracy	-	0.88	0.93	0.96	0.90	0.95

Tabla 5.10: Comparativa general de métricas de desempeño entre los modelos C0–C4.

Capítulo 6

Conclusiones

El presente proyecto constituye un aporte relevante en el camino hacia la detección temprana del picudo rojo mediante técnicas de visión por computadora y aprendizaje automático. A lo largo del trabajo se abordaron distintas estrategias de modelado, desde clasificadores tradicionales como *Support Vector Machines* y *Random Forest*, hasta arquitecturas híbridas que integran redes convolucionales con modelos basados en características estadísticas. Esta diversidad metodológica permitió contrastar enfoques y evidenciar tanto sus fortalezas como sus limitaciones.

El proceso desarrollado representó en sí mismo un aprendizaje integral. El trabajo con imágenes constituyó un terreno completamente nuevo, que exigió la construcción de un flujo de preprocesamiento propio y la adaptación de metodologías existentes a un entorno local con condiciones particulares. Del mismo modo, la experiencia de trabajo de campo resultó clave para comprender las dificultades prácticas que impone la recolección de datos en contextos urbanos, así como la necesidad de desarrollar protocolos rigurosos de captura y anotación que aseguren la calidad del dataset. En conjunto, estas instancias permitieron convertir la investigación en un proceso formativo que trascendió lo meramente técnico.

En términos de impacto, el proyecto aporta evidencia clara sobre la viabilidad del enfoque y abre un amplio margen de mejora. Su utilidad potencial trasciende el ámbito académico, dado que puede servir de base para iniciativas institucionales en curso, tales como las impulsadas por la Fuerza Aérea, la Intendencia de Montevideo y el Ministerio de Ambiente, que trabajan en la resolución del problema principal abarcado en este proyecto mediante la utilización de otras herramientas como drones o imágenes satelitales. En el contexto uruguayo, donde el picudo rojo ha ocasionado la pérdida de un número considerable de palmeras y amenaza el patrimonio paisajístico de ciudades como Montevideo, disponer de un sistema de alerta temprana constituye una herramienta de gran relevancia para la gestión fitosanitaria.

En este sentido, la mayor contribución del proyecto radica en haber sentado las bases para un esquema metodológico adaptado a las condiciones del país,

integrando aprendizaje automático con imágenes térmicas en escenarios reales de campo. Más allá de las métricas alcanzadas, el valor del trabajo se encuentra en haber demostrado la factibilidad técnica de la propuesta, visibilizando los principales desafíos del preprocesamiento y la adquisición de datos, y delineando un camino concreto hacia soluciones más robustas y escalables en el futuro.

Capítulo 7

Trabajo Futuro

A partir de la experiencia desarrollada, se identifican varias líneas de mejora que pueden fortalecer significativamente la propuesta. En primer lugar, resulta indispensable ampliar el conjunto de datos disponible, incorporando un número mayor de ejemplares y garantizando la certeza en su clasificación. La construcción de un dataset más robusto, acompañado de un protocolo profesional de captura de imágenes, permitirá reducir el sesgo generado por condiciones incontroladas y mejorar la generalización de los modelos.

En segundo lugar, la incorporación de sensores montados en drones aparece como un camino natural para optimizar la adquisición de imágenes. Este enfoque no solo aumentaría la cobertura espacial, sino que también posibilitaría añadir información de geolocalización de cada palmera, lo cual resulta de gran utilidad para el diseño de sistemas de monitoreo y alerta temprana a gran escala, aplicables por municipios y ministerios.

Asimismo, la integración de modalidades adicionales, como imágenes multiespectrales, abre la puerta a modelos híbridos que combinen clasificadores tradicionales como *Random Forest* con arquitecturas de aprendizaje profundo, explotando de manera complementaria la información proveniente de diferentes tipos de sensores (térmicos y multiespectrales). Esta línea representa una oportunidad clara para avanzar hacia sistemas más precisos y robustos.

Finalmente, un aspecto crítico a futuro es la colaboración interdisciplinaria. La participación de entomólogos y agrónomos permitirá asegurar que la anotación de las palmeras en categorías como “sanas”, “afectadas” o “muertas” se realice con la mayor rigurosidad científica posible, lo que redundará en un aumento de la confiabilidad de los modelos entrenados.

En conjunto, estas acciones plantean un camino que mejora en el punto más sensible del proyecto, siendo la adquisición de datos. Mientras que por otro lado, la incorporación de nuevas modalidades sensoriales y la vinculación con actores institucionales y científicos, podrían consolidar un sistema escalable, estable y aplicable al contexto nacional.

Anexo A

Distribuciones de temperaturas con dataset completo

Estas son las distribuciones de temperatura para el conjunto de datos completo, utilizado en la Etapa final (véase [5.2.3](#)).

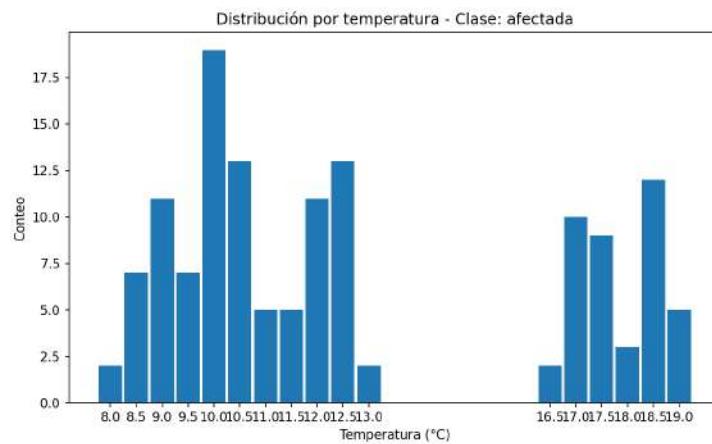


Figura A.1: Distribución de temperaturas para la clase *afectada*. Se observa una concentración de valores en dos rangos principales, uno entre 8.5°C y 12.5°C y otro entre 17°C y 19°C, indicando posibles diferencias térmicas entre niveles de afectación.

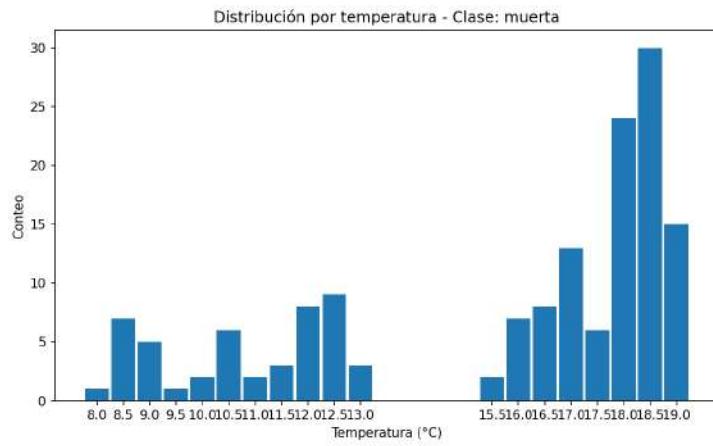


Figura A.2: Distribución de temperaturas para la clase *muerta*. La mayoría de las temperaturas se agrupan en el rango de 17°C a 19°C, con pocos valores dispersos por debajo de 13°C, lo que sugiere que las plantas muertas presentan una mayor temperatura promedio.

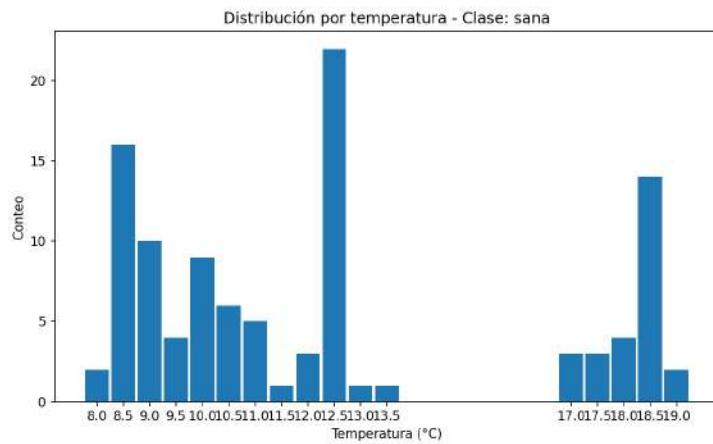


Figura A.3: Distribución de temperaturas para la clase *sana*. Se aprecian dos concentraciones principales: una en el rango de 8.5°C a 13°C y otra menor alrededor de 17°C a 19°C, lo que podría reflejar la variabilidad térmica natural en plantas sanas.

Anexo B

Resultados experimentales complementarios

En este anexo se presentan las métricas globales y adicionales correspondientes a los modelos evaluados (C0–C4). Cada bloque incluye los valores de *Accuracy Bal*, *Accuracy*, *F1 Macro*, *Kappa*, *MCC*, *Top-2 Accuracy* por clase.

Modelo C0 — Clasificador Aleatorio

Accuracy	Bal. Accuracy	F1 Macro	Kappa	MCC	Top-2 Acc.
0.31	0.33	0.31	—	—	—

Tabla B.1: Métricas globales del modelo C0.

Modelo C1 — Random Forest

Accuracy	Bal. Accuracy	F1 Macro	Kappa	MCC	Top-2 Acc.
0.68	0.66	0.66	0.50	0.51	0.88

Tabla B.2: Métricas globales del modelo C1.

Modelo C2 — CNN + Random Forest

Accuracy	Bal. Accuracy	F1 Macro	Kappa	MCC	Top-2 Acc.
0.65	0.64	0.63	0.47	0.48	0.93

Tabla B.3: Métricas globales del modelo C2.

Modelo C3 — CNN

Accuracy	Bal. Accuracy	F1 Macro	Kappa	MCC	Top-2 Acc.
0.60	0.62	0.60	0.40	0.43	0.96

Tabla B.4: Métricas globales del modelo C3.

Modelo C3 con aumento de datos

Accuracy	Bal. Accuracy	F1 Macro	Kappa	MCC	Top-2 Acc.
0.60	0.64	0.59	0.41	0.45	0.90

Tabla B.5: Métricas globales del modelo C3 con aumento de datos.

Modelo C4 — Híbrido optimizado

Accuracy	Bal. Accuracy	F1 Macro	Kappa	MCC	Top-2 Acc.
0.69	0.68	0.69	0.52	0.52	0.95

Tabla B.6: Métricas globales del modelo C4.

Referencias

- Abbas, M. S. T. (2016). Biology, ecology, and control of the red palm weevil, *rhynchophorus ferrugineus* (olivier) (coleoptera: Curculionidae: Rhynchophorinae). *Egyptian Journal of Biological Pest Control*, 26(1), 1–23. (Accedido: 2025-08-12)
- Abuagla, A. M., y Al-Deeb, M. A. (2012). Effect of bait quantity and trap color on the trapping efficacy of the pheromone trap for the red palm weevil, *rhynchophorus ferrugineus*. *Journal of Insect Science*, 12(120), 1–13. (Accedido: 2025-08-10) doi: 10.1673/031.012.12002
- Aerial Productions. (2025). *Cámara multiespectral: qué es y para qué sirve*. <https://aerialproductions.es/camara-multiespectral/>. (Accedido: 2025-07-20)
- Akiba, T., Sano, S., Yanase, T., Ohta, T., y Koyama, M. (2019). Optuna: A next-generation hyperparameter optimization framework. *CoRR*, *abs/1907.10902*. Descargado de <http://arxiv.org/abs/1907.10902> (Accedido: 2025-09-15)
- Alaa, H., Waleed, K., Samir, M., Alzahaby, M., Sobeah, H., y Salam, M. A. (2020, Jul). An intelligent approach for detecting palm trees diseases using image processing and machine learning. *International Journal of Advanced Computer Science and Applications*. Descargado de <https://thesai.org/Publications/ViewPaper?Volume=11&Issue=7&Code=IJACSA&SerialNo=57> (Accedido: 2025-06-24) doi: 10.14569/IJACSA.2020.0110757
- Amazon Web Services, Inc. (2025). ¿qué es una gan? Descargado 2025-05-23, de <https://aws.amazon.com/es/what-is/gan/> (Retrieved from AWS official website)
- Antoniou, A., Storkey, A., y Edwards, H. (2017). Data augmentation generative adversarial networks. *arXiv preprint arXiv:1711.04340*. (Accedido: 2025-07-17)
- Bannari, A., Abdulaziz, M., y El Battay, A. (2017, July). Water stress detection as an indicator of red palm weevil attack using worldview-3 data. En *2017 ieee international geoscience and remote sensing symposium (igarss)*. IEEE. Descargado de <https://ieeexplore.ieee.org/document/8127877> (Accedido: 2025-06-19) doi: 10.1109/IGARSS.2017.8127877

- Basheer, I. A., y Hajmeer, M. (2000). Artificial neural networks: fundamentals, computing, design, and application. *Journal of Microbiological Methods*, 43(1), 3–31. Descargado de <https://www.sciencedirect.com/science/article/pii/S0167701200002013> (Accedido: 2025-09-17) doi: 10.1016/S0167-7012(00)00201-3
- Bengio, S., Weston, J., y Grangier, D. (2010). Label embedding trees for large multi-class tasks. En *Advances in neural information processing systems* (Vol. 23, pp. 163–171). (Accedido: 2025-09-24)
- Bertalmío, M., Sapiro, G., Caselles, V., y Ballester, C. (2000). Image inpainting. En *Proceedings of the 27th annual conference on computer graphics and interactive techniques (siggraph)* (pp. 417–424). (Accedido: 2025-08-06) doi: 10.1145/344779.344972
- Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer. (Accedido: 2025-09-15)
- Bochkovskiy, A., Wang, C.-Y., y Liao, H.-Y. M. (2020). Yolov4: Optimal speed and accuracy of object detection. *arXiv preprint arXiv:2004.10934*. (Accedido: 2025-08-20)
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32. Descargado de <https://www.stat.berkeley.edu/~breiman/randomforest2001.pdf> (Accedido: 2025-07-26)
- Brodersen, K. H., Ong, C. S., Stephan, K. E., y Buhmann, J. M. (2010). The balanced accuracy and its posterior distribution. En *2010 20th international conference on pattern recognition* (pp. 3121–3124). (Accedido: 2025-09-06)
- Brown, L. G. (1992). A survey of image registration techniques. *ACM computing surveys (CSUR)*, 24(4), 325–376. (Accedido: 2025-09-31)
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1), 37–46. (Accedido: 2025-09-10)
- Cortes, C., y Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3), 273–297. Descargado de <https://link.springer.com/article/10.1007/BF00994018> (Accedido: 2025-07-30)
- Criminisi, A., Pérez, P., y Toyama, K. (2004). Region filling and object removal by exemplar-based image inpainting. *IEEE Transactions on Image Processing*, 13(9), 1200–1212. (Accedido: 2025-08-09) doi: 10.1109/TIP.2004.833105
- Cubuk, E. D., Zoph, B., Mane, D., Vasudevan, V., y Le, Q. V. (2019). Autoaugment: Learning augmentation policies from data. En *Proceedings of the ieee conference on computer vision and pattern recognition (cvpr)* (pp. 113–123). (Accedido: 2025-07-12)
- Cubuk, E. D., Zoph, B., Shlens, J., y Le, Q. V. (2020). Randaugment: Practical automated data augmentation with a reduced search space. En *Proceedings of the ieee/cvpr conference on computer vision and pattern recognition workshops* (pp. 702–703). (Accedido: 2025-07-15)
- Delalieux, S., Hardy, T., Ferry, M., Gomez, S., Kooistra, L., Culman, M., y Tits, L. (2023, Feb). Red palm weevil detection in date palm using temporal

- uav imagery. *Remote Sensing*. Descargado de <https://www.mdpi.com/2072-4292/15/5/1380> (Accedido: 2025-06-22) doi: 10.3390/rs15051380
- Dembilio, O., y Jacas, J. A. (2009). Biological control of the red palm weevil: a review. *Biocontrol Science and Technology*, 19(8), 767–788. (Accedido: 2025-08-21) doi: 10.1080/09583150903100865
- Dembilio, O., y Jaques, J. A. (2012). Biology and management of red palm weevil in the mediterranean area. *Journal of Pest Science*, 85(4), 417–429. (Accedido: 2025-08-01) doi: 10.1007/s10340-012-0449-3
- Deng, J., Dong, W., Socher, R., Li, L., Li, K., y Fei-Fei, L. (2009). Imagenet: A large-scale hierarchical image database. En *2009 ieee conference on computer vision and pattern recognition* (pp. 248–255). (Accedido: 2025-09-06) doi: 10.1109/CVPR.2009.5206848
- El-Sabea, A. M. R., Faleiro, J. R., y Abo-El-Saad, M. M. (2011). The red palm weevil rhynchophorus ferrugineus: biology, distribution and control: a review. *Outlooks on Pest Management*, 22(2), 56–72. (Accedido: 2025-08-08) doi: 10.1564/22apr07
- Equitec Tecnologías y Control. (2025). *Picudo rojo en butia odorata. palmeras nativas de uruguay [video]*. Descargado 2025-09-29, de https://www.youtube.com/watch?v=BrqMgq5PC_I (Accedido: 2025-09-15)
- Erbani, J., Portier, P.-E., Egyed-Zsigmond, E., y Nurbakova, D. (2024). Confusion matrices: A unified theory. *IEEE Access*, 12, 181372-181419. (Accedido: 2025-09-29) doi: 10.1109/ACCESS.2024.3507199
- European Space Agency. (2025). *Worldview-3 mission overview*. <https://earth.esa.int/eogateway/missions/worldview-3>. (Accedido: 2025-08-10)
- Everingham, M., Van Gool, L., Williams, C. K. I., Winn, J., y Zisserman, A. (2010). The pascal visual object classes (voc) challenge. *International Journal of Computer Vision*, 88(2), 303–338. (Accedido: 2025-08-03) doi: 10.1007/s11263-009-0275-4
- Faleiro, J. R. (2006). A review of the issues and management of the red palm weevil rhynchophorus ferrugineus (coleoptera: Rhynchophoridae) in coconut and date palm during the last one hundred years. *International Journal of Tropical Insect Science*, 26(3), 135–154. (Accedido: 2025-08-03) doi: 10.1079/IJT2006113
- Fawcett, T. (2006). An introduction to roc analysis. *Pattern Recognition Letters*, 27(8), 861–874. (Accedido: 2025-09-19)
- Fensholt, R., y Sandholt, I. (2003). Derivation of a shortwave infrared water stress index from modis near- and shortwave infrared data in a semiarid environment. *Remote Sensing of Environment*, 87(1), 111–121. (Accedido: 2025-08-24) doi: 10.1016/j.rse.2003.07.002
- Ferri, C., Hernández-Orallo, J., y Modroiu, R. B. (2009). An experimental comparison of performance measures for classification. *Pattern Recognition Letters*, 30(1), 27–38. (Accedido: 2025-10-04)
- Fertinyect. (2025). *Endoterapia*. <https://fertinyect.com/endoterapia/>. (Accedido: 2025-07-28)

- FLIR Systems. (2020a). Flir t1020 hd thermal imaging camera datasheet [Manual de software informático]. Descargado de <https://assets.testequity.com/te1/Documents/pdf/flir/flir-t1020-12-datasheet.pdf> (Accedido: 2025-08-16)
- FLIR Systems. (2020b). Flir t1020 user manual [Manual de software informático]. Descargado de <https://www.flir.com> (Accedido: 2025-08-19)
- FLIR Systems. (2025). *Flir t1020 high-performance hd thermal camera*. <https://www.flir.es/products/t1020/>. (Accedido: 2025-08-01)
- Fluke Corporation. (2016). Fluke 971 temperature humidity meter user manual [Manual de software informático]. Descargado de <https://www.fluke.com> (Accedido: 2025-08-13)
- Fluke Corporation. (2025). *How infrared cameras work*. <https://www.fluke.com/en/learn/blog/thermal-imaging/how-infrared-cameras-work>. (Accedido: 2025-08-06)
- Gao, B.-C. (1996). Ndwi—a normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sensing of Environment*, 58(3), 257–266. (Accedido: 2025-08-18) doi: 10.1016/S0034-4257(96)00067-3
- Gates, D. M. (1964). Leaf temperature and transpiration. *Agronomy Journal*, 56(3), 273–277. (Accedido: 2025-08-11) doi: 10.2134/agronj1964.00021962005600030019x
- Gauthier, E., Marin, T., Bodnar, J. l., y Stubbe, L. (2014, 07). Assessment of the pertinence of infrared thermography as a diagnostic tool in sinusitis cases study.. (Accedido: 2025-09-24) doi: 10.21611/qirt.2014.025
- Giblin-Davis, R. M. (2001). Borers of palms. *Invertebrate Biodiversity as Bioindicators of Sustainable Landscapes*, 267–305. (Accedido: 2025-08-06)
- Gindin, G., Levski, S., Glazer, I., y Soroker, V. (2006). Evaluation of the entomopathogenic fungi metarhizium anisopliae and beauveria bassiana against the red palm weevil, rhynchophorus ferrugineus. *Phytoparasitica*, 34(4), 370–379. (Accedido: 2025-08-04) doi: 10.1007/BF02981024
- Glosario de tecnología: qué significa vehículo aéreo no tripulado (uav)*. (2024). Descargado 2025-08, de <https://www.infobae.com/tecnologia/2024/08/24/glosario-de-tecnologia-que-significa-vehiculo-aereo-no-tripulado-uav> (Accedido: 2025-09-13)
- Golomb, O., Alchanatis, V., Cohen, Y., Levin, N., y Soroker, V. (2015, 07). Detection of red palm weevil infected trees using thermal imaging. En *Precision agriculture '15*. Wageningen Academic Publishers. (Accedido: 2025-06-28) doi: 10.3920/978-90-8686-814-8_80
- Gonzalez, R. C., y Woods, R. E. (2008). *Digital image processing* (3.^a ed.). Prentice Hall. (Accedido: 2025-08-02)
- Goodfellow, I., Bengio, Y., y Courville, A. (2016). *Deep learning*. MIT Press. Descargado de <http://www.deeplearningbook.org> (Accedido: 2025-08-22)
- Hardisky, M. A., Klemas, V., y Smart, R. M. (1983). The influence of soil salinity, growth form, and leaf moisture on the spectral radiance of Spartina

- alterniflora canopies. *Photogrammetric Engineering and Remote Sensing*, 49(1), 77–83. (Accedido: 2025-08-27)
- Hartley, R., y Zisserman, A. (2004). *Multiple view geometry in computer vision*. Cambridge university press. (Accedido: 2025-09-28)
- Haykin, S. (1998). *Neural networks: A comprehensive foundation* (2.^a ed.). Prentice Hall. (Accedido: 2025-08-15)
- He, K., Zhang, X., Ren, S., y Sun, J. (2016a). Deep residual learning for image recognition. En *Proceedings of the ieee conference on computer vision and pattern recognition* (pp. 770–778). (Accedido: 2025-09-01)
- He, K., Zhang, X., Ren, S., y Sun, J. (2016b). Deep residual learning for image recognition. En *Proceedings of the ieee conference on computer vision and pattern recognition (cvpr)* (pp. 770–778). (Accedido: 2025-08-28) doi: 10.1109/CVPR.2016.90
- He, K., Zhang, X., Ren, S., y Sun, J. (2016c). Deep residual learning for image recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770-778. Descargado de https://www.cv-foundation.org/openaccess/content_cvpr_2016/papers/He_Deep_Residual_Learning_CVPR_2016_paper.pdf (Accedido: 2025-07-21)
- Hoddle, M. S., Hoddle, C. D., Faleiro, J. R., El-Sabea, A. M. R., y Jeske, D. R. (2013). How far can the red palm weevil fly? computerized flight mill studies with field-captured weevils. *Journal of Economic Entomology*, 106(4), 1535–1540. (Accedido: 2025-08-24) doi: 10.1603/EC13017
- Hossain, M. S., Shahriar, G. M., Syeed, M. M. M., Uddin, M. F., Hasan, M., Shivam, S., y Advani, S. (2023, julio). Region of interest (roi) selection using vision transformer for automatic analysis using whole slide images. *Scientific Reports*, 13(1), 11314. (Accedido: 2025-09-15) doi: 10.1038/s41598-023-38109-6
- Huang, G., Liu, Z., Van Der Maaten, L., y Weinberger, K. Q. (2017). Densely connected convolutional networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 4700-4708. Descargado de <https://arxiv.org/abs/1608.06993> (Accedido: 2025-07-24)
- Huang, G., Liu, Z., y Weinberger, K. Q. (2016). Densely connected convolutional networks. *CoRR, abs/1608.06993*. Descargado de <http://arxiv.org/abs/1608.06993> (Accedido: 2025-09-04)
- IBM. (2025a). *Yolo v2 - ibm visual insights documentation*. <https://www.ibm.com/docs/es/visual-insights?topic=model-yolo-v2>. (Accedido: 2025-09-18)
- IBM. (2025b). *¿qué es una máquina de vectores de soporte (svm)?* — ibm. <https://www.ibm.com/mx-es/think/topics/support-vector-machine>. (Accedido: 2025-05-23)
- IBM. (2025c). *¿qué es una red neuronal convolucional?* — ibm. <https://www.ibm.com/es-es/think/topics/convolutional-neural-networks>. (Accedido: 2025-05-23)
- IBM Cloud Education. (2020). *What are random forests?* <https://www.ibm.com/topics/random-forest>. (Accedido: 2025-09-26)

- IBM Corporation. (2025). *What is data augmentation?* Descargado 2025-05-20, de <https://www.ibm.com/think/topics/data-augmentation> (Retrieved from IBM Think Blog)
- Jocher, G., Chaurasia, A., y Qiu, J. (2023). *Yolov8*. <https://github.com/ultralytics/ultralytics>. (Accedido: 2025-08-10)
- Jocher, G., Chaurasia, A., Stoken, A., Borovec, J., NanoCode012, Kwon, Y., ... others (2022). *ultralytics/yolov5: v7.0 - yolov5 sota realtime instance segmentation*. Zenodo. (Accedido: 2025-08-24) doi: 10.5281/zenodo.7347926
- Jones, H. G. (2014). Crop water stress index (cwsi). En N. V. Alfen (Ed.), *Encyclopedia of agriculture and food systems* (pp. 414–420). Elsevier. (Accedido: 2025-08-02) doi: 10.1016/B978-0-12-409548-9.10171-X
- Kamilaris, A., y Prenafeta-Boldú, F. X. (2018). Deep learning in agriculture: A survey. *Computers and Electronics in Agriculture*, 147, 70–90. (Accedido: 2025-08-31) doi: 10.1016/j.compag.2018.02.016
- Kauderer-Abrams, E. (2017). Quantifying translation-invariance in convolutional neural networks. *arXiv preprint arXiv:1801.01450*. (Accedido: 2025-09-06)
- Komala, G., Manda, R. R., y Seram, D. (2021, 04). Role of semiochemicals in integrated pest management. , 6, 247-253. (Accedido: 2025-09-11)
- Kočí, V., Šesták, J., y Černý, R. (2022). Thermal inertia and evaluation of reaction kinetics: A critical review. *Measurement*, 198, 111354. Descargado de <https://www.sciencedirect.com/science/article/pii/S0263224122005905> (Accedido: 2025-09-27) doi: 10.1016/j.measurement.2022.111354
- Krizhevsky, A., Sutskever, I., y Hinton, G. E. (2012a). Imagenet classification with deep convolutional neural networks. En *Advances in neural information processing systems (neurips)* (pp. 1097–1105). Curran Associates, Inc. Descargado de https://cvml.ista.ac.at/courses/DLWT_W17/material/AlexNet.pdf (Accedido: 2025-09-24)
- Krizhevsky, A., Sutskever, I., y Hinton, G. E. (2012b). Imagenet classification with deep convolutional neural networks. En *Advances in neural information processing systems (neurips)* (Vol. 25, pp. 1097–1105). (Accedido: 2025-08-24)
- Ktepi, B. (2024). *Matlab (matrix laboratory)*. Salem Press Encyclopedia of Science. Descargado de <https://research-ebsco-com.proxy.timbo.org.uy/linkprocessor/plink?id=10fa1cd9-df52-330d-9381-141c5946cbb2> (Accedido: 2025-09-26)
- Lai, J., Li, Z., Huang, D., y Lin, Q. (2024). *The optimality of kernel classifiers in sobolev space*. Descargado de <https://arxiv.org/abs/2402.01148> (Accedido: 2025-09-03)
- LeCun, Y., Bottou, L., Bengio, Y., y Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11), 2278–2324. (Accedido: 2025-08-19) doi: 10.1109/5.726791
- Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. En *International journal of computer vision* (Vol. 60, pp. 91–110). Springer. (Accedido: 2025-09-02)

- Mankin, R. W. (2011). Recent developments in the use of acoustic sensors and signal processing to detect and monitor insect pests. *Computers and Electronics in Agriculture*, 76(2), 125–132. (Accedido: 2025-08-26) doi: 10.1016/j.compag.2010.12.007
- Martín, M. M., y Cabello, T. (2009). Susceptibility of phoenix canariensis to red palm weevil attack. *Phytoparasitica*, 37(3), 243–251. (Accedido: 2025-08-17) doi: 10.1007/s12600-009-0032-y
- Matthews, B. W. (1975). Comparison of the predicted and observed secondary structure of t4 phage lysozyme. *Biochimica et Biophysica Acta (BBA)-Protein Structure*, 405(2), 442–451. (Accedido: 2025-09-13)
- Mesurex. (2025). *Cómo funciona una cámara termográfica*. <https://mesurex.com/como-funciona-una-camara-termografica/>. (Accedido: 2025-07-18)
- Ministerio de Ganadería, Agricultura y Pesca - Gobierno del Uruguay. (2024). *Información actualizada sobre el picudo rojo en palmeras*. Descargado de <https://www.gub.uy/ministerio-ganaderia-agricultura-pesca/comunicacion/noticias/informacion-actualizada-sobre-picudo-rojo-palmeras> (Accedido: 2025-07-29)
- Ministerio de Ganadería, Agricultura y Pesca (MGAP). (2025). *Sistema de pronóstico fitosanitario*. <https://www.gub.uy/ministerio-ganaderia-agricultura-pesca/politicas-y-gestion/sistema-pronostico-fitosanitario>. (Accedido: 2025-09-19)
- Misic, M., Durdevic, D., y Tomasevic, M. (2012, 01). Evolution and trends in gpu computing. En (p. 289-294). (Accedido: 2025-09-18)
- Modelo faster r-cnn — ibm visual insights*. (2025). <https://www.ibm.com/docs/es/visual-insights?topic=model-faster-cnn>. (Accedido: 2025-05-23)
- Mouli Chandra, B., H, A., O, P., B, A., T, V., y H, M. (2022, 05). Foot step power generation using piezoelectric sensors. *South Asian Journal of Engineering and Technology*, 12, 50-53. (Accedido: 2025-09-28) doi: 10.26524/sajet.2022.12.31
- Moura Mascarin, G., y Jaronski, S. (2016, 09). The production and uses of beauveria bassiana as a microbial insecticide. *World journal of microbiology & biotechnology*, 32, 177. (Accedido: 2025-09-31) doi: 10.1007/s11274-016-2131-3
- Murat, A. A., y Kiran, M. S. (2025). A comprehensive review on yolo versions for object detection. *Ciencia y Tecnología de la Ingeniería, una Revista Internacional*, 70, 102161. Descargado de <https://doi.org/10.1016/j.jestch.2025.102161> (Accedido: 2025-09-29) doi: 10.1016/j.jestch.2025.102161
- Niculescu-Mizil, A., y Caruana, R. (2005). Predicting good probabilities with supervised learning. En *Proceedings of the 22nd international conference on machine learning* (pp. 625–632). (Accedido: 2025-09-20)
- OCW. (2008). *Redes neuronales artificiales y sus aplicaciones*. OCW - Universidad del País Vasco. Descargado de <https://ocw.ehu.eus/>

- pluginfile.php/40137/mod_resource/content/1/redes_neuro/contenidos/pdf/libro-del-curso.pdf (Accedido: 2025-03-10)
- Opitz, J., y Burst, S. (2019). Macro f1 and macro f1. *arXiv preprint arXiv:1911.03347*. (Accedido: 2025-10-01)
- Parjane, N., Kabre, G., Mahale, A., Shejale, B., y Nirgude, S. (2020, 08). Compatibility of pesticides with metarhizium anisopliae. *JOURNAL OF ENTOMOLOGY AND ZOOLOGY STUDIES*, 8, 633-636. (Accedido: 2025-09-02)
- Perez, L., y Wang, J. (2017). The effectiveness of data augmentation in image classification using deep learning. *arXiv preprint arXiv:1712.04621*. (Accedido: 2025-07-08)
- Pieroni, G., Gandoms, B., y Frocelledo, M. (2025). *Relevamiento del estado del arte: Detección y georreferenciación de plagas en el sector agrícola*. https://gitlab.fing.edu.uy/bruno.gandoms/proyecto_de_grado/-/blob/main/Docs/P202501_25_-_Estado_del_Arte.pdf?ref_type=heads. (Accedido: 2025-10-08)
- Popescu, M.-C., Balas, V., Perescu-Popescu, L., y Mastorakis, N. (2009, 07). Multilayer perceptron and neural networks. *WSEAS Transactions on Circuits and Systems*, 8. (Accedido: 2025-09-21)
- Potamitis, I., Ganchev, T., y Kontodimas, D. (2009). On automatic bioacoustic detection of pests: the cases of rhynchophorus ferrugineus and sitophilus oryzae. *Journal of Economic Entomology*, 102(4), 1681–1690. (Accedido: 2025-08-30) doi: 10.1603/029.102.0435
- Powers, D. M. W. (2011). Evaluation: From precision, recall and f-measure to roc, informedness, markedness and correlation. *Journal of Machine Learning Technologies*, 2(1), 37–63. (Accedido: 2025-08-29)
- Pucha-Cofrep, F., y Fries, A. (2025). *Arcgis pro manual*. (Accedido: 2025-09-30)
- Ramírez, J., y Torres, M. (2019). Vehículos aéreos no tripulados: descripciones generales y aplicaciones. *ResearchGate*. Descargado 2025-08, de <https://www.researchgate.net/publication/332353911> (Accedido: 2025-09-15)
- Raíz Verde. (2025). ¿cuáles son los diferentes tipos de palmeras? <https://www.raizverde.com/cuales-son-los-diferentes-tipos-de-palmeras/>. (Accedido: 2025-08-10)
- Redmon, J., Divvala, S., Girshick, R., y Farhadi, A. (2016). You only look once: Unified, real-time object detection. En *Proceedings of the ieee conference on computer vision and pattern recognition (cvpr)* (pp. 779–788). (Accedido: 2025-08-15) doi: 10.1109/CVPR.2016.91
- Redmon, J., y Farhadi, A. (2018). Yolov3: An incremental improvement. En *arxiv preprint arxiv:1804.02767*. (Accedido: 2025-08-18)
- Rehman, Y. (2024). *How the bottleneck layers in the deep networks work and how do those layers reduce computational complexity*. <https://medium.com/@reh.yawar2/how-the-bottleneck-layers-in-the-deep-networks-work-and-how-do-those-layers-reduce-computational-7bc99c0d1e96>. (Accedido: 2025-09-16)
- Rivas-Martínez, S. (2012). *Phoenix canariensis: historia, biología y conservación*.

- ción. Gobierno de Canarias, Consejería de Medio Ambiente. (Accedido: 2025-08-15)
- Rosas Licona, P., Lara, L., Carballo, A., y Bueno, V. (2024, 08). Feromonas: la neuroquímica invisible de la comunicación animal. *Ciencia Latina Revista Científica Multidisciplinar*, 8, 5059-5078. (Accedido: 2025-09-09) doi: 10.37811/cl_rcm.v8i4.12729
- Saito, T., y Rehmsmeier, M. (2015). The precision-recall plot is more informative than the roc plot when evaluating binary classifiers on imbalanced datasets. *PloS one*, 10(3), e0118432. (Accedido: 2025-09-22)
- Schwarz, E., Baniya, A., Heppert, J., Schwartz, H., Tan, C.-H., Antoshechkin, I., ... Dillman, A. (2025, 08). Genomes of the entomopathogenic nematode steinerinema hermaphroditum and its associated bacteria. *Genetics*. (Accedido: 2025-09-06) doi: 10.1093/genetics/iyaf170
- ScienceDirect. (2025). *Bounding box — an overview*. <https://www.sciencedirect.com/topics/engineering/bounding-box>. (Accedido: 2025-08-29)
- ScienceDirect Topics. (2025). *Rgb image*. <https://www.sciencedirect.com/topics/computer-science/rgb-image>. (Accedido: 2025-07-28)
- Sekhar, C., y Meghana, P. (2020, 08). A study on backpropagation in artificial neural networks. *Asia-Pacific Journal of Neural Networks and Its Applications*, 4, 21-28. (Accedido: 2025-09-03) doi: 10.21742/AJNNIA.2020.4.1.03
- Shorten, C., y Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. *Journal of Big Data*, 6(1), 1-48. (Accedido: 2025-07-06)
- Simonyan, K., y Zisserman, A. (2015). *Very deep convolutional networks for large-scale image recognition*. Descargado de <https://arxiv.org/abs/1409.1556> (Accedido: 2025-09-20)
- Sistema Nacional de Emergencias (SINAЕ). (2024). *Rhynchophorus ferrugineus: el picudo rojo en palmeras*. Descargado de <https://www.gub.uy/sistema-nacional-emergencias/comunicacion/noticias/rhynchophorus-ferrugineus-picudo-rojo-palmeras> (Accedido: 2025-07-29)
- Smith, L. N., y Topin, N. (2019). Super-convergence: Very fast training of neural networks using large learning rates. En *Artificial intelligence and machine learning for multi-domain operations applications* (Vol. 11006, pp. 369–386). (Accedido: 2025-09-24)
- Sohan, M., Ram, T., y Ch, V. (2024, 01). A review on yolov8 and its advancements. En *Proceedings of the international conference on artificial intelligence and machine learning* (p. 529-545). Springer Nature Singapore. (Accedido: 2025-05-15) doi: 10.1007/978-981-99-7962-2_39
- Sokolova, M., y Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. *Information Processing & Management*, 45(4), 427–437. (Accedido: 2025-08-27) doi: 10.1016/j.ipm.2009.03.002
- Szeliski, R. (2010). *Computer vision: Algorithms and applications*. Springer. (Accedido: 2025-08-13) doi: 10.1007/978-1-84882-935-0

- Tanner, C. B. (1963). Plant temperatures. *Agronomy Journal*, 55(2), 210–211. (Accedido: 2025-08-15) doi: 10.2134/agronj1963.00021962005500020029x
- Ultralytics. (2025). *Yolov7 - ultralytics documentation*. <https://docs.ultralytics.com/es/models/yolov7/>. (Accedido: 2025-09-18)
- Universidad del País Vasco (UPV/EHU) - Grupo de Investigación IBeA. (2025). *Asd high-resolution fieldspec 4 spectroradiometer*. <https://www.ehu.eus/es/web/ibea/equipamiento/espectroscopia-microscopia-imagen/asd-spectroradiometer>. (Accedido: 2025-09-16)
- Unmanned aerial vehicles: Opportunities and challenges. (2023). *Frontiers in Robotics and AI*. Descargado 2025-08, de <https://pmc.ncbi.nlm.nih.gov/articles/PMC9841964> (Accedido: 2025-09-19)
- Unmanned aerial vehicle (uav). (2024). *ScienceDirect Topics*. Descargado 2025-08, de <https://www.sciencedirect.com/topics/engineering/unmanned-aerial-vehicle> (Accedido: 2025-09-10)
- Vidyasagar, P. S. P. V., Aldosari, S. A., Sultan, A., y Al-Dobai, S. (2016). Impact of mass pheromone trapping on red palm weevil: a key pest of date palm in the middle east. *Journal of Insect Science*, 16(1), 1–9. (Accedido: 2025-08-06) doi: 10.1093/jisesa/iew093
- Wang, H., y Zhao, H. (2018). Survey on unmanned aerial vehicle networks: A cyber physical system perspective. *arXiv preprint arXiv:1805.00881*. Descargado 2025-08, de <https://arxiv.org/abs/1805.00881> (Accedido: 2025-09-22)
- Wang, L., Qu, J. J., Hao, X., y Zhu, Q. (2007). Nmdi: A normalized multi-band drought index for monitoring soil and vegetation moisture with satellite remote sensing. *Geophysical Research Letters*, 34(20), L20405. (Accedido: 2025-08-20) doi: 10.1029/2007GL031021
- Wikipedia. (2025). *Unmanned aerial vehicle*. Descargado 2025-08, de https://es.wikipedia.org/wiki/Veh%C3%ADculo_a%C3%93reo_no_tripulado (Accedido: 2025-09-06)
- Xu, G., Wang, X., Wu, X., Leng, X., y Xu, Y. (2025, febrero). Development of residual learning in deep neural networks for computer vision: A survey. *Engineering Applications of Artificial Intelligence*, 142, 109890. Descargado de <http://dx.doi.org/10.1016/j.engappai.2024.109890> (Accedido: 2025-09-18) doi: 10.1016/j.engappai.2024.109890