Texto

Descripción generada automáticamente

Universidad Internacional de La Rioja

Escuela Superior de Ingeniería y

Tecnología

Máster Universitario en Inteligencia artificial

Categorizar imágenes de plagas en cultivos de cacao utilizando métodos de aprendizaje profundo supervisado

|  |  |
| --- | --- |
| Trabajo fin de estudio presentado por: | Bryan Ricardo Armas Loyaga |
| Tipo de trabajo: | Comparativa de soluciones |
| Director/a: | Alvaro Alexander Martinez Navarro |
| Fecha: | 21-04-2025 |

Resumen

La Inteligencia Artificial tienen gran impacto en cada uno de las actividades de uso diario, sobre todo en el campo laboral, con la elaboración de este TFM permite seleccionar un método de aprendizaje profundo supervisado capaz de evaluar, detectar y categorizar las plagas en los cultivos de cacao. Es por ello que se ha llevado a cabo una comparativa de soluciones para lograr establecer cual, de los modelos evaluados, genera mejores resultados para el tema propuesto, utilizando el mismo dataset obtenido desde kaggle pero adaptado a las diferentes estructuras que cada uno de los modelos necesita como parámetros para su entrenamiento.

Para ello se ha realizado la transformación mediante Python para que los diferentes modelos puedan ejecutar su entrenamiento y validaciones sin inconvenientes por tema de datos.

El resultado obtenido luego de este trabajo de fin de master muestra un modelo que arrojó excelentes resultados para el dataset utilizado, dándolo como ganador del modelo comparativo de soluciones, manteniendo un equilibrio entre interpretabilidad y precisión.

**Palabras Clave:** Detección de plagas, Resnet-50, Aprendizaje Profundo, Aprendizaje supervisado, CocoaPest.

Abstract

Artificial Intelligence has a significant impact on our daily activities, especially in the workplace. This Master's Final Project (TFM) aims to select a supervised deep learning method capable of evaluating, detecting, and categorizing pests in cocoa crops. Therefore, a comparative analysis of different solutions was carried out to determine which of the evaluated models yields the best results for the proposed topic. The same dataset from Kaggle was used, but it was adapted to the specific structures each model requires for its training parameters.

The data transformation was performed using Python to ensure the various models could execute their training and validation processes without any data-related issues.

The final result of this master's project shows that one model produced excellent results for the dataset used, making it the winner of the comparative analysis while maintaining a balance between interpretability and precision.

**Keywords**: Pest Detection, Resnet-50, Deep Learning, Supervised Learning, CocoaPest.

Índice de contenidos

[1. Introducción 1](#_Toc208413417)

[1.1. Motivación 2](#_Toc208413418)

[1.2. Planteamiento del trabajo 3](#_Toc208413419)

[1.3. Estructura del trabajo 4](#_Toc208413420)

[2. Contexto y estado del arte 6](#_Toc208413421)

[2.1. Contexto del problema 6](#_Toc208413422)

[2.2. Estado del arte 7](#_Toc208413423)

[2.3. Conclusiones 12](#_Toc208413424)

[3. Objetivos concretos y metodología de trabajo 12](#_Toc208413425)

[3.1. Objetivo general 12](#_Toc208413426)

[3.2. Objetivos específicos 12](#_Toc208413427)

[3.3. Metodología del trabajo 13](#_Toc208413428)

[4. Planteamiento de la comparativa 14](#_Toc208413429)

[4.1. Identificación y Contextualización del Problema 14](#_Toc208413430)

[4.2. Soluciones Alternativas a Evaluar: Modelos de Deep Learning Supervisado 14](#_Toc208413431)

[4.2.1. YOLO (You Only Look Once) V8: 15](#_Toc208413432)

[4.2.2. Faster R-CNN (Region-based Convolutional Neural Network): 15](#_Toc208413433)

[4.2.3. Mask R-CNN: 15](#_Toc208413434)

[4.2.4. ResNet-50: 15](#_Toc208413435)

[4.3. Criterios de Éxito para la Comparativa 16](#_Toc208413436)

[4.3.1. F1-Curva de confianza 16](#_Toc208413437)

[4.3.2. Precisión-Curva de confianza 16](#_Toc208413438)

[4.3.3. Matriz de confusión 17](#_Toc208413439)

[4.3.4. Precisión promedio 17](#_Toc208413440)

[4.3.5. Velocidad 17](#_Toc208413441)

[5. Desarrollo de la comparativa 17](#_Toc208413442)

[6. Discusión y análisis de resultados 26](#_Toc208413443)

[7. Conclusiones y trabajos futuros 29](#_Toc208413444)

[7.1. Conclusiones 29](#_Toc208413445)

[7.2. Trabajos futuros 30](#_Toc208413446)

[8. Trabajos citados 31](#_Toc208413447)

Índice de figuras

[Ilustración 1 *Distribución de información del dataset de origen* 18](#_Toc208413448)

[Ilustración 2 *Matriz de confusión del modelo YOLOv8* 21](#_Toc208413449)

[Ilustración 3 *F1-Curva de confianza inicial modelo YOLOv8* 21](#_Toc208413450)

[Ilustración 4 *Precision-Curva de confianza modelo YOLOv8* 22](#_Toc208413451)

[Ilustración 5 *Matriz de confusión modelo Faster R-CNN* 22](#_Toc208413452)

[Ilustración 6 *F1-Curva de confianza del modelo Faster R-CNN* 23](#_Toc208413453)

[Ilustración 7 *Precision-Curva de confianza del modelo Faster R-CNN* 23](#_Toc208413454)

[Ilustración 8 *Matriz de confusión del modelo Mask R-CNN* 24](#_Toc208413455)

[Ilustración 9 *F1-Curva de confianza para el modelo Mask R-CNN* 24](#_Toc208413456)

[Ilustración 10 *Precisión-Curva de confianza para el modelo Mask R-CNN* 25](#_Toc208413457)

[Ilustración 11 *Matriz de confusión para el modelo Resnet-50* 25](#_Toc208413458)

[Ilustración 12 *Resultado del modelo Resnet-50 con el set de datos de test* 26](#_Toc208413459)

Índice de tablas

[Tabla 1 Proceso a realizar para la investigación 13](#_Toc208413460)

[Tabla 2 *Resumen comparativo de criterios de éxito* 17](#_Toc208413461)

[Tabla 3 Resultados del modelo YOLOv8 26](#_Toc208413462)

[Tabla 4 Resultado del modelo Faster R-CNN 27](#_Toc208413463)

[Tabla 5 Resultados del modelo Resnet-50 28](#_Toc208413464)

[Tabla 6 Tabla comparativa de resultados para los sets de datos en el modelo Resnet-50 29](#_Toc208413465)

# Introducción

El incremento acelerado en la generación de datos visuales incrustados en imágenes, impulsado por la propagación de dispositivos de captura de imágenes y la expansión de plataformas digitales, ha permitido posicionar la categorización de las imágenes como una tarea de importancia crítica en el panorama tecnológico actual. Esta capacidad de analizar y comprender automáticamente el contenido visual dentro las imágenes se han convertido en un componente esencial para la eficiencia operativa, la toma de decisiones estratégicas y la innovación en diversas industrias, que van desde la medicina hasta la agricultura, pasando por la seguridad, banca, finanzas y el comercio electrónico.

Con el aprendizaje supervisado profundo, al ser una rama avanzada del aprendizaje automático, se ha considerado como la metodología dominante para abordar los desafíos inherentes a la categorización de imágenes. Esto ha permitido aprovechar las bondades de los grandes conjuntos de datos etiquetados, los modelos de aprendizaje profundo supervisado, especialmente las Redes Neuronales Convolucionales (CNNs), han demostrado una capacidad sin precedentes para extraer características complejas y establecer patrones complejos dentro de los datos visuales. Esta habilidad permite, en muchos casos, superar la precisión del ojo humano en tareas específicas de clasificación, abriendo nuevas fronteras en la automatización y la inteligencia artificial.

En esta investigación se ahonda en el núcleo de la categorización de imágenes mediante el uso del aprendizaje profundo supervisado. A través de una exploración exhaustiva, se desglosa los conceptos fundamentales que sustentan este campo, se analiza las arquitecturas de CNNs más influyentes que han marcado su evolución, se examina los conjuntos de datos de referencia que sirven como pilares dentro de la investigación y la evaluación, permite la discusión de las métricas esenciales para cuantificar el rendimiento de los modelos. Además, se incluye el tema de los desafíos persistentes, las limitaciones inherentes a esta disciplina, destacando las tendencias emergentes y los avances recientes los cuales prometen dar forma al futuro en relación con la categorización de imágenes. Finalmente, se establecerá los pasos esenciales utilizados en el desarrollo de un sistema de categorización de imágenes los cuales se basan en aprendizaje supervisado profundo.

## Motivación

En el territorio ecuatoriano varias empresas dedicadas a la producción de cacao se han visto afectados por pérdidas económicas y del producto, dando como resultado que el causante de esta perdida son las plagas dentro de los cultivos.

Una de las problemáticas detectadas según (Instituto Nacional Autónomo de Investigaciones Agropecuarias, n.d.) se debe a diversos problemas ocasionados por insectos, ácaros, malezas, hongos, bacterias, virus, nematodos entre otros y al no disponer de información exacta sobre sus causas no pueden tomar medidas apropiadas, con perjuicio al rendimiento, incrementado los costos y ocasionando deterioro de la calidad de sus cosechas.

A su vez (Pérez-Martínez et al., 2017) en su artículo señala que esos síntomas de marchitamiento o clorosis que suelen manifestarse en los viveros de cacao son, en gran medida, resultado de la agresiva acción de hongos. Estos microorganismos indeseables se propagan a través de vías de transmisión horizontal, creando un verdadero desafío para los cultivadores.

Por su parte (Sanchez Alvarado et al., 2023) indica que Diversas plagas se hacen resistentes a plaguicidas, fungicidas y herbicidas; se emplean agroquímicos porque se cree que no es viable producir arroz sin un alto nivel de sustancias químicas para controlar las plagas. Los plaguicidas causan y suscitan problemas ecológicos y de salud para humanos y animales.

Dentro de la investigación previa realizada por mi parte para encontrar trabajos con características similares dentro de Ecuador, si bien se encuentra trabajos que utilizan detección con procesamiento de imágenes multiespectrales, manejo de plagas con redes de sensores inalámbricos o análisis de plagas en rosas con redes neuronales convolucionales, sin embargo, no existen proyectos que se enfocan en la clasificación de imágenes para detección de plagas con aprendizaje profundo supervisado.

La utilización de imágenes dentro de los cultivos de cacao en el día a día, la creciente difusión y la importancia estratégica de la información visual ha permitido centrar el estudio y optar por este tema de titulación. La proliferación de cámaras en dispositivos móviles, sistemas de vigilancia dentro de las plantaciones, genera un flujo constante de imágenes que contienen información valiosa, susceptible de ser analizada para obtener conocimientos profundos y facilitar la automatización de tareas.

En el ámbito de la agricultura, la categorización de imágenes permite la detección temprana de anomalías en cultivos, la clasificación en relación de la calidad del cacao, de los procesos de cosecha, lo que se traduce en una mayor eficiencia y una reducción de costos.

El aprendizaje profundo supervisado ha demostrado ser una herramienta excepcionalmente poderosa para abordar estos desafíos en diversos dominios. Su capacidad para aprender representaciones jerárquicas de características directamente de los datos de la imagen, sin la necesidad de una ingeniería de características manual exhaustiva, lo convierte en un enfoque atractivo y efectivo (LeCun et al., 2015).

Las CNNs, esas fascinantes redes neuronales que han revolucionado el campo, han alcanzado logros notables en la categorización de imágenes a gran escala. ¿Quién no se siente impresionado al ver cómo brillan en competiciones de referencia como ImageNet? Los resultados son innegables y son testimonio del increíble trabajo realizado por investigadores como ImageNet (Khan et al., 2022).

Sin embargo, a pesar de los éxitos alcanzados, la categorización de imágenes con aprendizaje profundo supervisado aún presenta desafíos importantes. La necesidad de grandes cantidades de datos etiquetados para entrenar modelos robustos, el problema del sobreajuste, la interpretabilidad limitada de las redes neuronales profundas y la demanda de recursos computacionales significativos son solo algunos de los obstáculos que aún requieren investigación y desarrollo continuo.

En este trabajo de investigación la motivación es el deseo de ahondar en la comprensión de los fundamentos teóricos y las aplicaciones prácticas del aprendizaje profundo supervisado para la categorización de imágenes en detección de plagas en los cultivos de cacao. Al explorar las arquitecturas de CNNs más relevantes, los conjuntos de datos de referencia, las métricas de evaluación y los desafíos actuales, se busca proporcionar una visión integral del estado del arte en este campo y sentar las bases para futuras investigaciones y avances.

## Planteamiento del trabajo

El presente trabajo se centra en la exploración de la categorización de imágenes para la detección de plagas en cultivos de cacao utilizando métodos de aprendizaje profundo supervisado. La premisa de que el aprendizaje profundo, y en particular las CNNs, representan el estado del arte en la resolución de problemas complejos de clasificación visual. Lo cual permitirá evaluar y determinar el mejor método de aprendizaje profundo supervisado.

**Conceptos Fundamentales**: Definir y clarificar los conceptos esenciales relacionados con la categorización de imágenes y el aprendizaje profundo supervisado, estableciendo un marco teórico sólido para la investigación.

**Arquitecturas de CNNs**: Investigar y describir las arquitecturas de CNNs más influyentes y ampliamente utilizadas en la categorización de imágenes, analizando sus innovaciones, ventajas y limitaciones.

**Conjuntos de Datos de Referencia**: Identificar y caracterizar los conjuntos de datos de imágenes de referencia más relevantes y utilizados para el entrenamiento y la evaluación de modelos de categorización, destacando sus características y su importancia en la investigación.

**Métricas de Evaluación**: Explicar y analizar las métricas de evaluación estándar utilizadas para cuantificar el rendimiento de los modelos de categorización de imágenes, debatiendo su significado y aplicabilidad en diferentes contextos.

**Desafíos y Limitaciones**: Explorar y discutir los desafíos y limitaciones propias a la categorización de imágenes con aprendizaje profundo supervisado, incluyendo el sobreajuste, la necesidad de datos etiquetados y la interpretabilidad de los modelos.

**Tendencias y Avances Recientes**: Investigar, destacar las tendencias y los avances más recientes en la investigación de la categorización de imágenes, como el uso de mecanismos de atención y redes transformadoras.

**Pasos de Desarrollo**: Describir los pasos involucrados en el desarrollo de un sistema de categorización de imágenes utilizando aprendizaje profundo supervisado, desde la recopilación de datos hasta la evaluación del modelo.

En resumen, este trabajo proporciona una comprensión profunda y actualizada del campo de la categorización de imágenes utilizando aprendizaje profundo supervisado, sirviendo como un recurso valioso para aquellos interesados en investigar y aplicar estas técnicas en diversos dominios.

## Estructura del trabajo

• **Capítulo 1**: Introducción: Este capítulo proporciona una visión general del tema de la categorización de imágenes de plagas en los cultivos de cacao utilizando métodos de aprendizaje profundo supervisado, incluyendo la motivación, el planteamiento del trabajo y la estructura del TFM.

• **Capítulo 2**: Contexto y estado del arte: En este capítulo se describe el contexto del problema de la categorización de imágenes de plagas en los cultivos de cacao, se presenta un análisis del estado del arte en el campo del aprendizaje profundo supervisado aplicado a esta tarea. Se revisarán los antecedentes, los estudios actuales y las comparativas de herramientas existentes.

• **Capítulo 3**: Objetivos concretos y metodología de trabajo: Este capítulo detalla el objetivo general y los objetivos específicos de la investigación, formulados bajo los criterios SMART. Además, se describe la metodología de trabajo que se seguirá para alcanzar estos objetivos.

• **Capítulo 4**: Planteamiento de la comparativa: Este capítulo detalla el trabajo previo realizado para identificar el problema y las posibles soluciones a evaluar.

• **Capítulo 5**: Desarrollo de la comparativa: Este capítulo detalla la comparativa realizada, incluyendo resultados y mediciones obtenidos.

• **Capítulo 6**: Discusión y análisis de resultados: Este capítulo se centrará en la discusión, análisis de las ventajas y desventajas de las soluciones evaluadas.

• **Capítulo 7**: Conclusiones y trabajos futuros.

# Contexto y estado del arte

## Contexto del problema

La categorización o clasificación de imágenes, es una tarea fundamental dentro del campo de la visión por computador que se centra en la comprensión automática del contenido visual de una imagen. Implica la asignación de una categoría o etiqueta a una imagen completa basándose en su contenido visual. Esta tarea es fundamental y juega un papel clave en un sinfín de aplicaciones que abarcan múltiples campos.

Desde la agricultura donde ayuda a optimizar cultivos hasta la medicina, donde puede marcar la diferencia en la atención al paciente. Sin olvidar la seguridad que se beneficia enormemente de esta tecnología, al igual que la teledetección que proporciona datos vitales sobre nuestro entorno.

En el ámbito de la banca y las finanzas su influencia es palpable, mejorando la gestión del riesgo y la toma de decisiones. Y por supuesto no podemos pasar por alto su impacto en el comercio electrónico, donde está revolucionando la manera en que compramos y vendemos.

En la agricultura, las técnicas de aprendizaje supervisado se aplican para la detección y clasificación de enfermedades y defectos en imágenes de frutas, tal como se revisa en el trabajo de (Ynfante Martínez et al., 2022), quienes destacan la relevancia de estas técnicas para asegurar y garantizar la calidad de productos como la fruta bomba y la guayaba dentro del mercado internacional.

Asimismo, (Sanchez Pardo & More Villegas, 2023) han desarrollado métodos basados en procesamiento de imágenes digitales y aprendizaje automático para la clasificación por madurez de la fresa, buscando optimizar la gestión de cultivos y la comercialización.

En el ámbito de la medicina, el aprendizaje profundo supervisado ha sido utilizado para la clasificación de tractos nerviosos mediante técnicas supervisadas, como se visualiza en la tesis de (Escalante Belmonte, 2024). En esta investigación se observa la implementación de técnicas basadas en aprendizaje profundo para distinguir tractos específicos en imágenes de resonancia magnética por difusión, con el propósito de su análisis posterior y potencial uso en aplicaciones clínicas.

Además, (Hinojosa Merlo, 2022) explora cómo se están aplicando las técnicas de Deep Learning en la clasificación de radiografías de COVID-19. Este enfoque ha demostrado ser una herramienta valiosa, mostrando la impresionante capacidad de la Inteligencia Artificial para no solo ayudar, sino realmente enriquecer la detección y el diagnóstico rápido de esta enfermedad que tanto nos ha impactado.

La intersección entre tecnología y salud está, sin duda, creando nuevas oportunidades para mejorar la atención médica. Los documentos de (Ruíz & Domínguez, n.d.) también exploran el Deep Learning aplicado en imágenes fotoacústicas para la identificación del cáncer de seno, destacando el potencial de estas técnicas en el diagnóstico médico.

En el emocionante ámbito de la seguridad, se utilizan técnicas de aprendizaje profundo supervisado que permiten identificar a personas en imágenes y detectar rostros de manera asombrosa. Este tema se explora a fondo en la investigación de (Salvadó i Gómez, 2024). Este proyecto se enfoca en entrenar un modelo de IA que detecte e identifique personas utilizando el algoritmo YOLO y técnicas de Deep Learning.

En la teledetección, el aprendizaje profundo supervisado se aplica para la clasificación automática de cobertura de suelo utilizando imágenes satelitales, como se presenta en el artículo de (Suárez Londoño et al., 2017).

En este trabajo, se explora de manera fascinante el uso de Redes Neuronales Convolucionales, esas maravillas de la inteligencia artificial, para llevar a cabo la clasificación y el mapeo automático de las coberturas del suelo en imágenes satelitales. Es un viaje al corazón de la tecnología que nos ayuda a desentrañar los secretos del paisaje desde arriba en forma satelital.

Así mismo, (Jiménez Cleves et al., 2024) realiza una evaluación comparativa del efecto del tamaño de entrenamiento de algoritmos de machine learning en la clasificación de imágenes Landsat-9, destacando la relevancia de la ingeniería geomática en la aplicación efectiva de modelos de clasificación de imágenes satelitales.

El trabajo de (Baquerizo & Ventocilla, 2022) también evalúan algoritmos de machine learning en la clasificación de imágenes satelitales multiespectrales, en el caso de la Amazonía peruana, comparando el rendimiento de diferentes técnicas de aprendizaje supervisado.

En el comercio electrónico, la categorización de imágenes mediante aprendizaje profundo supervisado se utiliza para la clasificación de productos, la búsqueda visual y la recomendación de artículos, mejorando la experiencia del usuario y optimizando las ventas. Estos ejemplos ilustran la vasta aplicabilidad de la categorización de imágenes utilizando métodos de aprendizaje profundo supervisado en diversos campos, lo que destaca la importancia de esta área de investigación.

## Estado del arte

El aprendizaje profundo supervisado ha marcado un antes y un después en el ámbito de la categorización de imágenes, y dentro de este fascinante escenario, las Redes Neuronales Convolucionales (CNNs) han emergido como auténticas protagonistas (Cireşan et al., 2011).

Con el paso del tiempo, han surgido variadas arquitecturas de CNNs, cada una aportando su propia chispa innovadora, lo que ha llevado a alcanzar desempeños realmente impresionantes en las tareas de categorizar imágenes. Es asombroso ver cómo la tecnología ha avanzado tanto.

Teoría sobre el aprendizaje profundo

De acuerdo con (Díaz-Ramírez, 2021) en la actualidad, la Inteligencia Artificial, Aprendizaje Automático y el Aprendizaje Profundo están en una etapa de constante crecimiento, serán seguramente los ejes principales para el desarrollo de la ciencia de la computación y la humanidad donde estarán fundidas, no sólo hardware y software, sino también varias tecnologías, tales como nanotecnología, computación cuántica, automatización, entre otras.

Por otra parte (Díaz Rodríguez et al., 2022) en su investigación manifiesta que gracias al aprendizaje automático las máquinas son capaces de aprender la relación oculta existente entre los datos de entrada.

También (Ramos & Santiago, 2022) comenta que el aprendizaje profundo es una de las poderosas técnicas de aprendizaje automático impulsadas por la IA. Las técnicas de aprendizaje profundo pueden procesar una gran cantidad de información presente en los conjuntos de datos de manera eficiente.

(Türkoğlu & Hanbay, 2019) afirma que las CNN utilizadas como herramientas básicas de aprendizaje profundo han obtenido un éxito significativo en varios estudios de detección de enfermedades de las plantas.

De igual manera (Wang et al., 2025) declara que los modelos de aprendizaje profundo han demostrado un rendimiento superior, ofreciendo mayor precisión, consistencia y escalabilidad en la detección de plagas y enfermedades. Estas ventajas resaltan la necesidad de explorar y sintetizar la investigación sobre aplicaciones de aprendizaje profundo en este campo, abordando las deficiencias críticas de los métodos tradicionales y sentando las bases para futuros avances. La Red Neuronal Convolucional (CNN), que aprovecha mecanismos como el aprendizaje jerárquico de características, la percepción local y la distribución de pesos, extrae eficazmente características complejas y discriminativas de imágenes sin procesar, lo que permite la identificación y segmentación de plagas y enfermedades de las plantas.

Sin embargo (Díaz-Gaxiola et al., 2019) apuntan a un desafío que se alza como el más formidable en el uso de las redes neuronales convolucionales: la enorme cantidad de datos que se necesita para llevar a cabo su entrenamiento. Es como intentar llenar un inmenso océano con un pequeño río; se requiere de una cantidad ingente de información para que estos modelos logren aprender y adaptarse efectivamente.

(Krizhevsky et al., 2011) marcó un hito en el campo de la visión por computador al ser una de las primeras CNNs profundas en lograr un éxito sobresaliente en el ImageNet Large Scale Visual Recognition Challenge (ILSVRC) 2012. Su arquitectura, compuesta por ocho capas (cinco convolucionales y tres totalmente conectadas), introdujo varias innovaciones clave. Utilizó la función de activación ReLU, que demostró acelerar el entrenamiento y mitigar el problema del desvanecimiento del gradiente, superando a las funciones sigmoid y tanh utilizadas anteriormente; también empleó capas de max-pooling con solapamiento para reducir la dimensionalidad espacial y controlar el sobreajuste, así como la técnica de dropout en las capas totalmente conectadas para mejorar la generalización. Además, fue pionera en el uso de unidades de procesamiento gráfico (GPUs) para acelerar significativamente el entrenamiento en conjuntos de datos masivos como ImageNet. A pesar de estos avances, AlexNet podía ser susceptible al sobreajuste en conjuntos de datos más pequeños.

Desde la perspectiva agrícola, es interesante notar que, como señalaron (Nassr & Abu Naser, 2018) en diversas regiones del globo, los expertos en agricultura no se ocupan realmente de las enfermedades que aquejan a la piña. Sorprendentemente, estos especialistas y los centros dedicados al tratamiento de dichas enfermedades son raros, casi escasos, en la mayor parte del mundo. Esta situación plantea interrogantes sobre el futuro de este cultivo tan apreciado.

(Ploetz, 2016) afirma que las enfermedades constituyen limitaciones significativas para la producción comercial y reducen el rendimiento en un 20 %, según estimaciones, o una producción estimada de 1,3 millones de toneladas de granos en 2012. Al ser trasladado fuera del neotrópico, el cacao se liberó de sus dos principales enemigos americanos: Moniliophthora roreri (causante de la mazorca helada) y M. perniciosa (escoba de bruja).

Podemos mencionar la Teoría de la Generalización (John Lu, 2010) ya que en el contexto del aprendizaje automático y la inteligencia artificial es un concepto fundamental que se refiere a la capacidad que tiene un modelo de inteligencia artificial o de aprendizaje automático se considera exitoso si, después de su entrenamiento con un conjunto de datos específico, es capaz de aplicar lo aprendido para procesar y rendir de manera óptima con datos completamente nuevos y nunca antes analizados.

Por su parte (Goodfellow et al., 2016) en su libro menciona que el desafío principal en el aprendizaje automático es lograr que los algoritmos funcionen eficazmente con datos nuevos, no vistos previamente, más allá de los datos con los que fueron entrenados. A esta capacidad se le denomina "generalización”. La generalización está estrechamente ligada a la capacidad del modelo, así como a los fenómenos de sobreajuste (overfitting) y subajuste (underfitting).

A su vez (Vasileiou et al., 2025) explora las capacidades de generalización de las redes neuronales de paso de mensajes donde se refiere a la habilidad de las MPNNs para hacer predicciones significativas más allá de los datos de entrenamiento. Esto implica su capacidad para adaptarse eficazmente a nuevos grafos no vistos previamente que provienen de la misma distribución que el conjunto de entrenamiento.

La representación de características es el proceso de transformar los datos de entrada en un formato adecuado que sea de utilidad para un algoritmo de aprendizaje automático, en ese sentido se puede encontrar los embeddings y la comprensión de características.

En su estudio (Mikolov, 2013) determina que una limitación de los vectores de incrustación estática de word2vec es que las palabras pueden significar cosas diferentes en distintos contextos. Sin embargo, con las incrustaciones estáticas, cada palabra se representa mediante un único punto en el espacio vectorial, aunque pueda tener diversos significados.

(Li et al., 2022) en su artículo destaca la capacidad para aprender representaciones jerárquicas y de menor dimensión (embeddings) a partir de datos de imágenes, lo cual es crucial para tareas de clasificación, como la identificación de plagas.

La compresión no es solo una optimización sino una parte integral de cómo las redes profundas aprenden representaciones de alto nivel y semánticas, ignorando detalles irrelevantes de bajo nivel para enfocarse en los patrones clave que definen una plaga.

La evaluación de un modelo de aprendizaje profundo supervisado va más allá de solo ver su precisión. La robustez y la elección de métricas de evaluación adecuadas son cruciales para entender el verdadero rendimiento en un entorno real, específicamente en la detección de plagas.

En su artículo (Hancox-Li, 2020) alude que este desacuerdo subyace a debates sobre otras desideratas, como la robustez de las explicaciones ante pequeñas perturbaciones en los datos de entrada. Sostengo que la robustez es deseable en la medida en que nos preocupa encontrar patrones reales en el mundo.

A su vez (Oyen et al., 2022) en su estudio enfatiza que se ha demostrado, tanto empírica como teóricamente, que los clasificadores de aprendizaje automático son robustos al ruido de etiqueta bajo ciertas condiciones; en particular, se asume que el ruido de etiqueta es independiente de las características dada la etiqueta de clase.

En el tema agrícola (Ferentinos, 2018) resalta las variaciones en las condiciones de luz, el ángulo de la cámara, la presencia de otras plantas o suciedad, y las diferencias en el desarrollo de la enfermedad pueden afectar la precisión de los modelos. Esto subraya directamente la necesidad de modelos robustos en entornos agrícolas.

La cognición visual humana según (Pinker, 1984) se refiere a los procesos complejos por los cuales nuestro cerebro interpreta la información de la luz que llega a nuestros ojos, transformándola en una comprensión significativa del mundo que nos rodea. Este proceso no es pasivo, sino que involucra atención, memoria, reconocimiento de patrones y razonamiento.

A su vez (LeCun et al., 2015) describe que las CNNs fueron concebidas con una arquitectura que imita el procesamiento jerárquico de la corteza visual. Las capas convolucionales extraen características de bajo nivel en las primeras etapas y estas a su vez se combinan en capas posteriores para formar representaciones de alto nivel, lo cual es fundamental para el reconocimiento de objetos complejos.

De la misma manera (LeCun et al., 2002) en su investigación explican cómo estas capas reducen la resolución espacial de los mapas de características, lo que les permite a las redes ser menos sensibles a pequeñas traslaciones y distorsiones en la entrada, contribuyendo así a la invariancia frente a la traslación.

Por el contrario (Smeulders et al., 2000) en su artículo formalizó el concepto de la brecha semántica. Lo define explícitamente como la diferencia entre la información de bajo nivel que se puede extraer de los datos visuales y la interpretación de alto nivel de esos datos por parte de un usuario.

A si mismo (Datta et al., 2005) en su trabajo reitera que la brecha semántica es el principal desafío en este campo, la describe como la dificultad de traducir características visuales primitivas estas pueden ser color, textura y forma, en conceptos semánticos que los humanos utilizan para describir y buscar imágenes.

Otro tema interesante es el Fairness en Machine Learning, se refiere a la aspiración de diseñar, desarrollar y desplegar sistemas de IA de manera que no discriminen, no reproduzcan ni amplifiquen sesgos existentes en los datos o en la sociedad, y traten a todos los individuos o grupos de manera justa. Según (Mehrabi et al., 2021) fairness busca garantizar que los sistemas de IA no discriminen contra individuos o grupos específicos y que traten a todos de manera justa.

El tema ético es crucial, por ello (Floridi, 2023) en su artículo aborda la ética de la IA de manera más amplia, incluye la equidad como un pilar fundamental. Resalta la responsabilidad social de los modelos de IA, la necesidad de evitar la amplificación de sesgos sistémicos cuando se despliegan en el día a día de las personas.

## Conclusiones

El estado del arte para la categorización de imágenes utilizando aprendizaje profundo supervisado ha tenido un avance significativamente en los últimos años. Las arquitecturas de CNNs han demostrado ser altamente efectivas, y la investigación continúa explorando nuevas formas de mejorar la precisión, la eficiencia y la interpretabilidad de estos modelos. La disponibilidad de grandes conjuntos de datos etiquetados y el desarrollo de potentes recursos computacionales han sido fundamentales para estos avances.

Sin embargo, todavía existen desafíos relacionados con el sobreajuste, la necesidad de grandes cantidades de datos y la interpretabilidad de los modelos, que continúan impulsando la investigación en este campo.

El conocimiento adquirido en este capítulo proporciona el contexto necesario para comprender los objetivos y la metodología de trabajo que se abordarán en los siguientes capítulos de esta investigación.

# Objetivos concretos y metodología de trabajo

## Objetivo general

Identificada la escases de categorización de imágenes de plagas dentro del país y al considerar que esto ocasiona pérdidas significativas a los productores de cacao, el presente trabajo realiza una comparativa para determinar la efectividad de diferentes arquitecturas de aprendizaje automático profundo supervisado para la categorización automática de imágenes de plagas en cultivos de cacao, con el fin de mejorar la detección temprana y el manejo de estas amenazas agrícolas.

## Objetivos específicos

Establecer mediante un estudio documental las técnicas de Machine Learning a comparar.

Desarrollar la comparativa de métodos de aprendizaje profundo supervisado para la detección de plagas en cultivos de cacao.

Establecer ventajas y desventajas de los métodos de aprendizaje profundo supervisado comparados para la detección de plagas en cultivos de cacao.

## Metodología del trabajo

De cara a alcanzar los objetivos específicos se explicará en este capítulo las fuentes de información utilizadas, la descripción de las técnicas e instrumentos utilizados para la recopilación de los datos, los enfoques para el análisis de la información recabada y los resultados previstos para completar cada propósito de investigación.

Tabla 1 Proceso a realizar para la investigación

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Objetivo Específico** | **Fuentes** | **Técnica de Recolección de Información** | **Instrumento** | **Técnica de Análisis de Información** | **Resultado** |
| **1. Establecer mediante un estudio documental las técnicas de Machine Learning a comparar.** | Artículos | Revisión documental | Ficha de revisión documental | Análisis documental | Listado top 4 de modelos de Deep Learning supervisado |
| **2. Desarrollar la comparativa de métodos de aprendizaje profundo supervisado para la detección de plagas en cultivos de cacao.** | Datasets de imágenes de plagas de cacao | Experimentación controlada | Bitácora | Análisis | Informe de resultados |
| **3. Establecer ventajas y desventajas de los métodos de aprendizaje profundo supervisado comparados para la detección de plagas en cultivos de cacao.** | Resultados de la comparativa, Métricas de evaluación estándar en visión por computadora | Evaluación y Análisis | Bitácora | Análisis estadístico | Informe de ventajas y desventajas |

# Planteamiento de la comparativa

Este capítulo tiene como propósito fundamental contextualizar la investigación, delineando el problema específico a abordar y las soluciones tecnológicas alternativas que serán evaluadas. Asimismo, se establecerán los criterios explícitos de éxito para la comparativa de modelos de Deep Learning, las métricas cuantitativas que se emplearán para evaluar su rendimiento y las consideraciones generales que guiarán el estudio experimental.

## Identificación y Contextualización del Problema

La agricultura es un pilar económico fundamental en diversas regiones, incluyendo aquellas donde el cultivo de cacao es predominante. Sin embargo, estos cultivos son constantemente amenazados por diversas plagas, que pueden causar pérdidas significativas en la producción, degradación de la calidad del producto, en casos extremos, la destrucción total de las cosechas. Tradicionalmente, la detección de estas plagas se ha basado en la inspección visual por parte de agrónomos o técnicos especializados.

Este método es intrínsecamente subjetivo, consume mucho tiempo, es costoso y, a menudo, la detección se produce cuando el daño ya es considerable, limitando la efectividad de las intervenciones.

La identificación tardía de plagas no solo incrementa el uso de pesticidas, con sus consecuentes impactos ambientales y en la salud humana, sino que también dificulta la implementación de estrategias de manejo integrado de plagas (MIP) que buscan soluciones más sostenibles. Por lo tanto, existe una necesidad apremiante de desarrollar sistemas de detección de plagas que sean más rápidos, precisos, automatizados y capaces de operar a gran escala en el campo.

## Soluciones Alternativas a Evaluar: Modelos de Deep Learning Supervisado

En la última década, los avances en el campo del Deep Learning, particularmente en las Redes Neuronales Convolucionales, han revolucionado el área de la visión por computadora. Estos modelos han demostrado una capacidad excepcional para aprender características complejas directamente de los datos de imagen, superando a menudo los métodos de procesamiento de imágenes tradicionales y el Machine Learning clásico en tareas de clasificación y detección de objetos.

Para abordar el problema de la detección de plagas en cultivos de cacao, esta investigación se centrará en la evaluación comparativa de las siguientes soluciones alternativas, todas ellas basadas en arquitecturas de Deep Learning supervisado, reconocidas por su estado del arte en tareas de visión por computadora.

Se ha utilizado para este TFM un conjunto de datos llamado Cocoa Diseases, del repositorio de Kaggle correspondiente al año 2020 en su versión 2, la particularidad de este dataset es que su estructura estaba basada para ser utilizado en el modelo Yolo especialmente en la versión 4.

### YOLO (You Only Look Once) V8:

Durante el entrenamiento arrojó un rendimiento considerable por su alta velocidad y buena precisión en la detección de objetos en tiempo real. Ya lo dijo (Redmon et al., 2015) YOLO es extremadamente rápido. Dado que enmarcamos la detección como un problema de regresión, no necesitamos una secuencia compleja.

### Faster R-CNN (Region-based Convolutional Neural Network):

Implementarlo es un tanto más complejo en comparación a YOLOv8 ya que se requirió más pasos de configuración, pero mejora significativamente la precisión y robustez en la detección de objetos, aunque generalmente más lento que YOLO. Como lo indicó (Ahmed & Das, 2025) los resultados muestran que Faster R-CNN logra la máxima precisión en la detección de peatones y ciclistas, mientras que YOLOv5 tiene un buen equilibrio entre velocidad y precisión.

### Mask R-CNN:

Es una extensión de Faster R-CNN que añade la capacidad de segmentación de instancias a nivel de píxel. (He et al., 2016) menciona que el método genera eficientemente objetos en una imagen mientras simultáneamente produce una máscara de segmentación de alta calidad para cada instancia. Lograr su ejecución fue un poco más complicado que los modelos anteriores debido a que las imágenes que son de alta resolución consumen mucha más memoria, para ello se hizo un ajuste en las dimensiones de la imagen a 512x512 pixeles.

### ResNet-50:

Se manifiesta que estas arquitecturas son fundamentales como extractores de características y como base para clasificadores de imágenes. Ya lo afirma (He et al., 2015) donde explica cómo las redes residuales profundas ResNets facilitan el entrenamiento de redes mucho más profundas, lo que conduce a una mayor precisión en la clasificación de imágenes. El concepto de saltos residuales permite que la información fluya directamente a través de capas, mitigando el problema de la desaparición o explosión de gradientes.

A diferencia de los otros modelos Resnet-50 también contiene o espera otro tipo de directorio para poder realizar el entrenamiento.

Estas arquitecturas representan un espectro de soluciones desde las más rápidas (YOLO) hasta las más precisas y detalladas como Resnet-50 que ya la revisaremos a detalle en el apartado de conclusiones, permitiendo una comparativa en el contexto específico de la detección de plagas en los cultivos de interés.

## Criterios de Éxito para la Comparativa

El éxito de esta comparativa se determinará por la capacidad que tenga el modelo para localizar las plagas u objetos con precisión y tener la capacidad de clasificarlos de manera correcta, que permitirán determinar cuál o cuáles modelos de Deep Learning son los más adecuados para la detección de plagas en cultivos de cacao. Los criterios principales son:

### F1-Curva de confianza

Es la métrica que combina la precisión y el recall. El puntaje en F1 cercano a 1.0 da a entender que el modelo tiene un buen equilibrio entre no equivocarse al detectar objetos y no perderse ninguno.

El umbral de confianza es el valor mínimo que una detección debe tener para ser considerada válida.

### Precisión-Curva de confianza

Esta curva es trascendental para comprender cómo varía la precisión del modelo a medida que se realiza ajustes en el umbral de confianza.

En el eje X se encuentra el umbral de confianza que no es más que es el puntaje mínimo que debe tener la detección para que el modelo se considere cómo válido.

Y en el eje Y se encuentra la precisión que mide la proporción de detecciones que son correctas. Un valor de precisión de 1.0 significa que todas las detecciones que nuestro modelo hizo a un umbral dado son correctas.

### Matriz de confusión

Este criterio es fundamentalmente clave para conocer si el modelo no solamente detecta los objetos, sino también ver si los clasifica de manera correcta.

### Precisión promedio

Se puede mencionar que estas métrica AP y mAP son las más importantes para la detección de objetos:

AP: esta métrica nos ayuda con la medición de una clase específica, es decir un AP alto para cada clase sea Fito, Monilia o Sana nos da un indicador de criterio exitoso.

mAP: esta métrica a su vez es el promedio de las AP de todas las clases empleadas, esto se interpreta en que si el valor del mAP se encuentra encima del 0.5 el modelo tiene un buen rendimiento.

### Velocidad

Este criterio es crucial para YOLOv8 que se destaca por su velocidad, si su rendimiento de FPS es suficientemente alto cumplirá con los requisitos del sistema.

Tabla 2 *Resumen comparativo de criterios de éxito*

|  |  |  |
| --- | --- | --- |
| **Modelo** | **Tipo** | **Criterios de Éxito** |
| YOLOv8, Faster R-CNN | Detección de Objetos | - mAP alto.  - Curvas de F1-Confianza y Precisión-Confianza altas.  - Velocidad de inferencia (FPS) alta (especialmente para YOLOv8). |
| Mask R-CNN | Detección y Segmentación de Objetos | - Mismos criterios que la detección (mAP, F1-Confianza).  - Precisión de segmentación (mAP de máscaras) alta. |
| ResNet/Inception | Clasificación de Imágenes | - Matriz de Confusión con la mayoría de los valores en la diagonal.  - Precisión, Recall y F1-Score altos (por clase y promediados). |

# Desarrollo de la comparativa

En este capítulo se muestra la información referente a los modelos evaluados, así como la información pertinente del dataset utilizado en el desarrollo de esta comparativa de soluciones.

Este dataset fue tomado de Kaggle y contiene 627 archivos distribuidas de la siguiente estructura de carpetas:

Enfermedades Cacao

* Fito

- archivos .jpg

- archivos .txt

* Monilia

- archivos .jpg

- archivos .txt

* Sana

- archivos .jpg

- archivos .txt

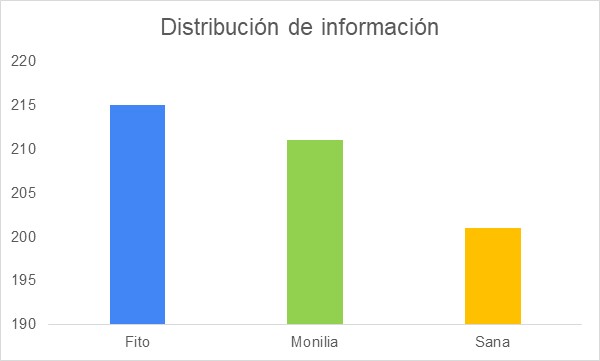
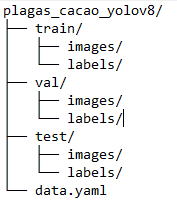


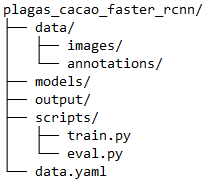
Ilustración 1 *Distribución de información del dataset de origen*

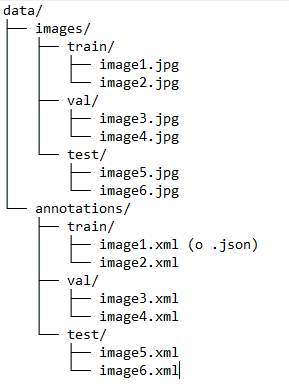
Uno de los desafíos iniciales al realizar este trabajo fue el formato esperado por cada uno de los modelos, esto llevo a realizar el tratamiento y homologación de los datos utilizando el mismo notebook para colocar los datos en el formato esperado por el modelo.

**YOLOv8**:

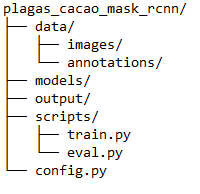


**Faster R-CNN:**

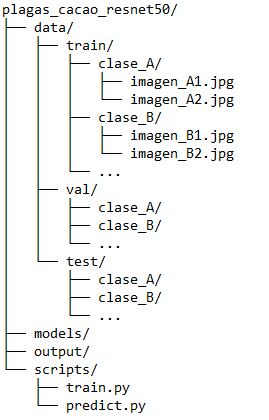




**Mask R-CNN:**



**Resnet-50:**



Como se ha hecho mención en los apartados, por cada uno de los modelos empleados en la comparativa de soluciones, se ha realizado el ajuste de dichos parámetros asociados para ser entrenados con el conjunto de datos de entrenamiento.

En las Ilustración 1, 5, 8 y 11 se puede visualizar las matrices de confusión para cada uno de los modelos empleados YOLOv8, Faster R-CNN, Mask R-CNN y finalmente Resnet-50

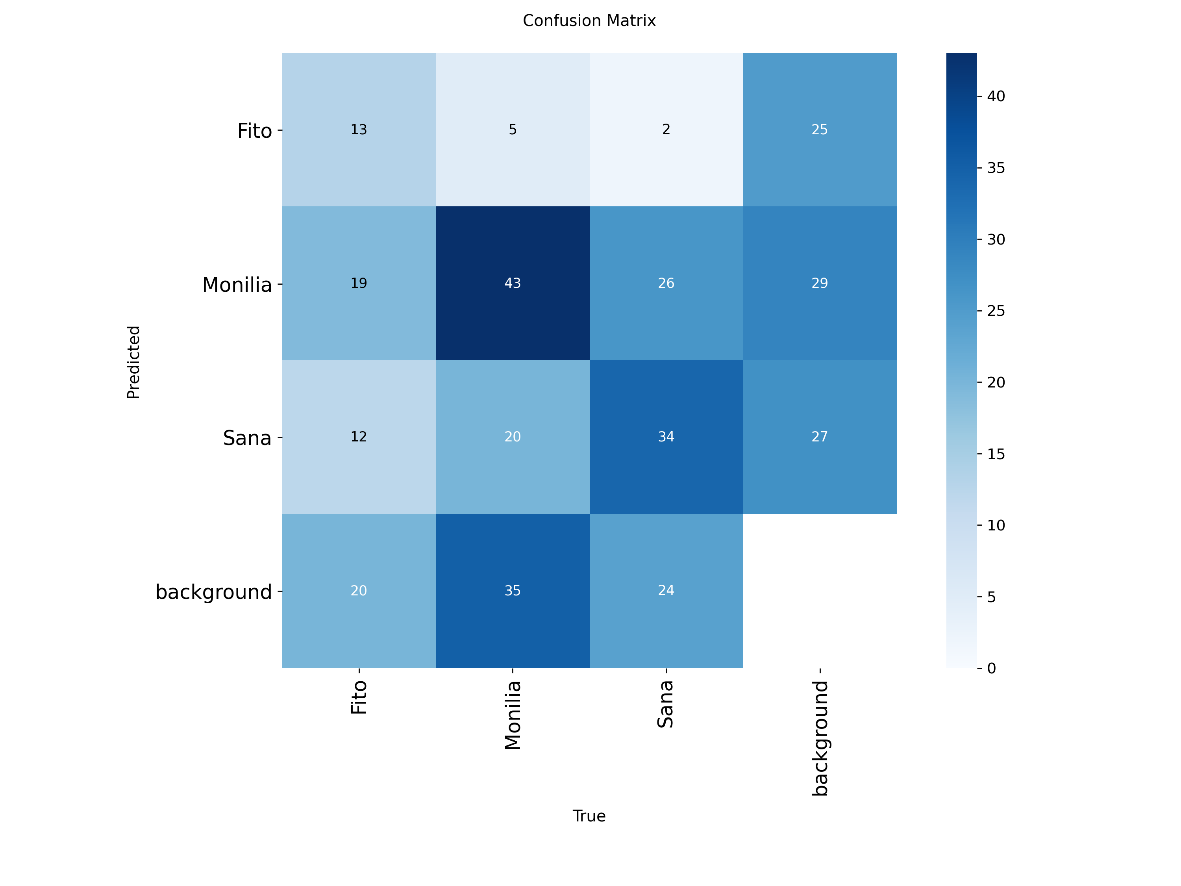


Ilustración *Matriz de confusión del modelo YOLOv8*

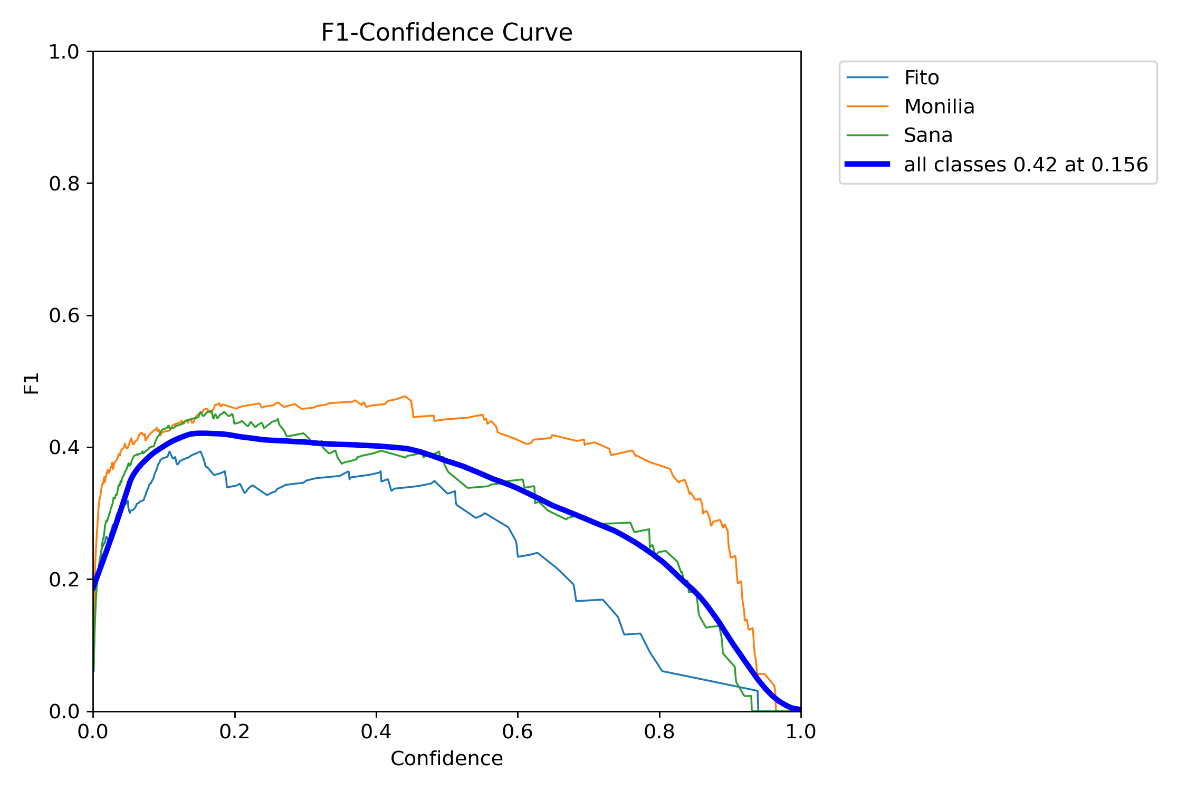


Ilustración *F1-Curva de confianza inicial modelo YOLOv8*

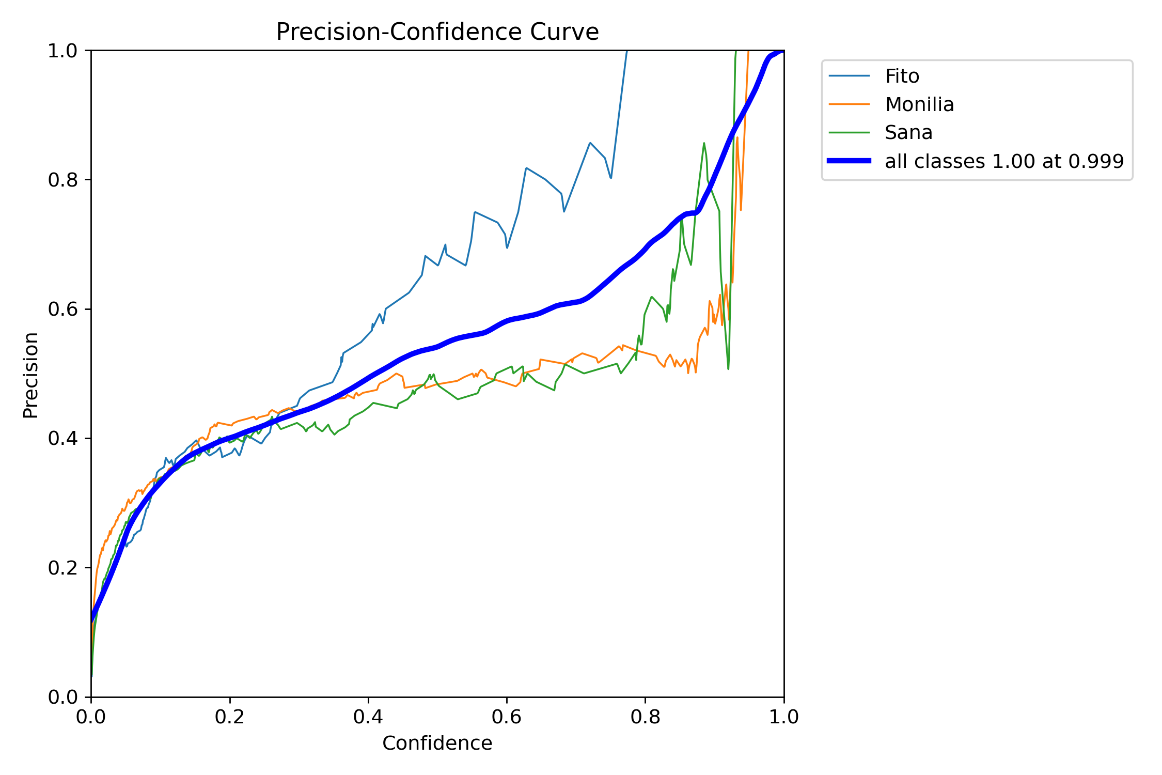


Ilustración *Precision-Curva de confianza modelo YOLOv8*

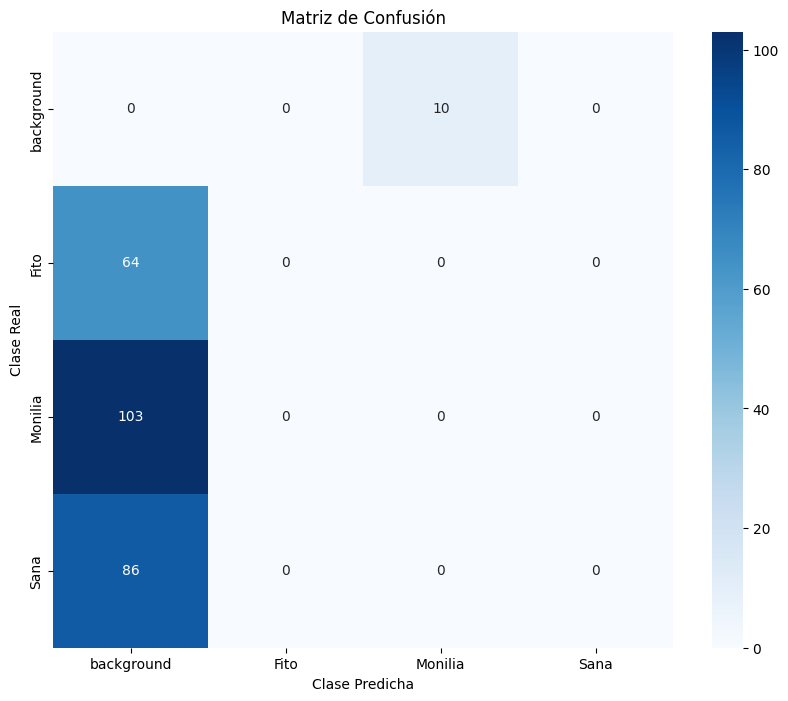


Ilustración *Matriz de confusión modelo Faster R-CNN*

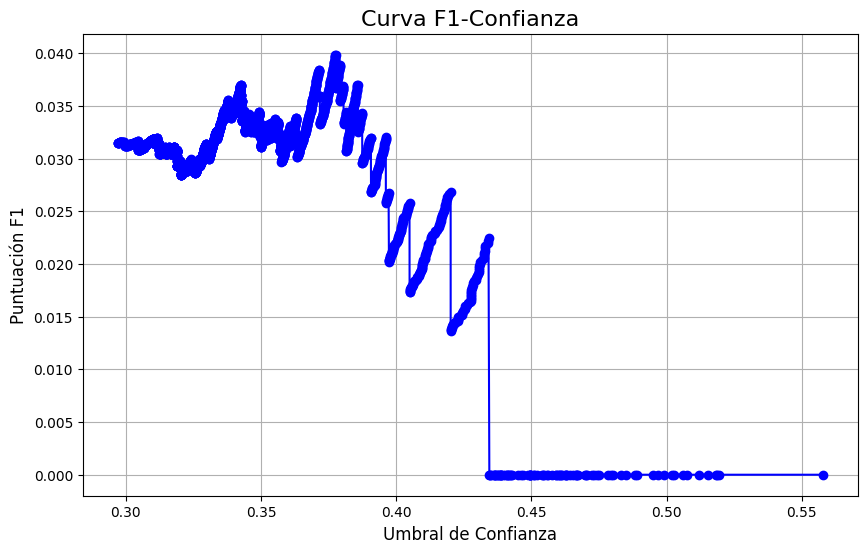


Ilustración *F1-Curva de confianza del modelo Faster R-CNN*

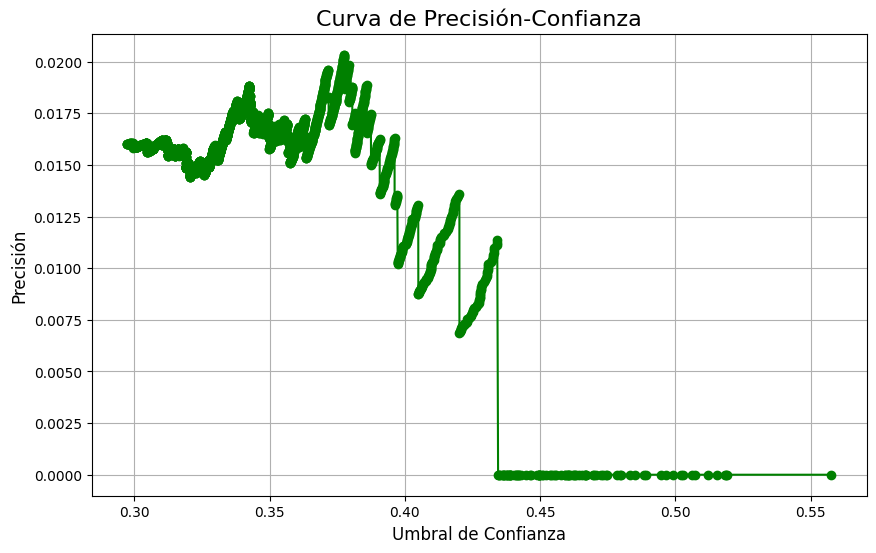


Ilustración *Precision-Curva de confianza del modelo Faster R-CNN*

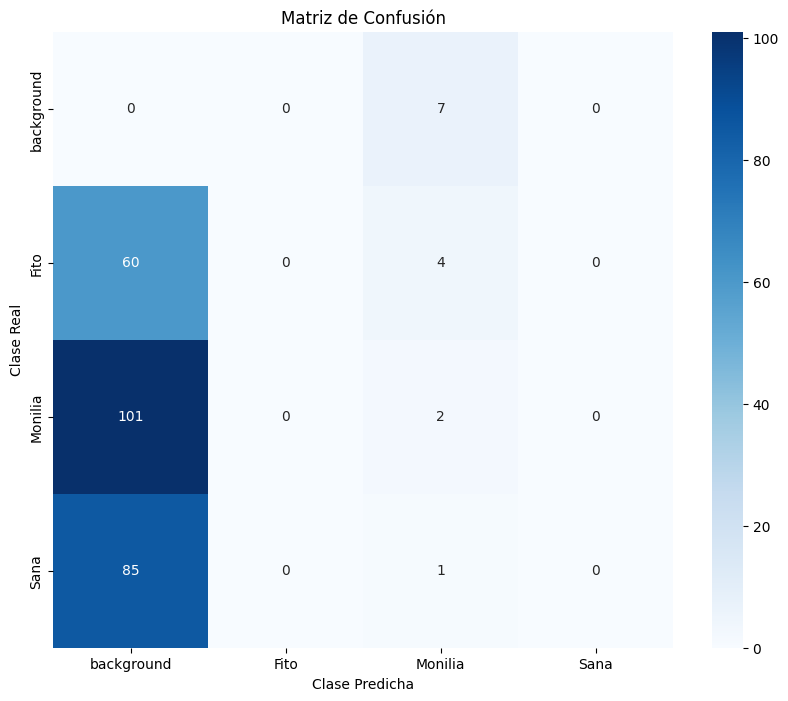


Ilustración 8 *Matriz de confusión del modelo Mask R-CNN*

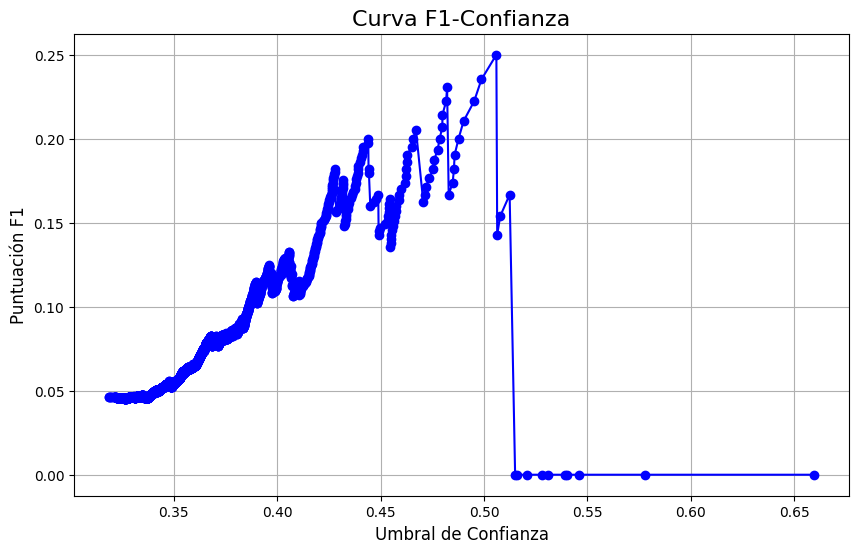


Ilustración 9 *F1-Curva de confianza para el modelo Mask R-CNN*

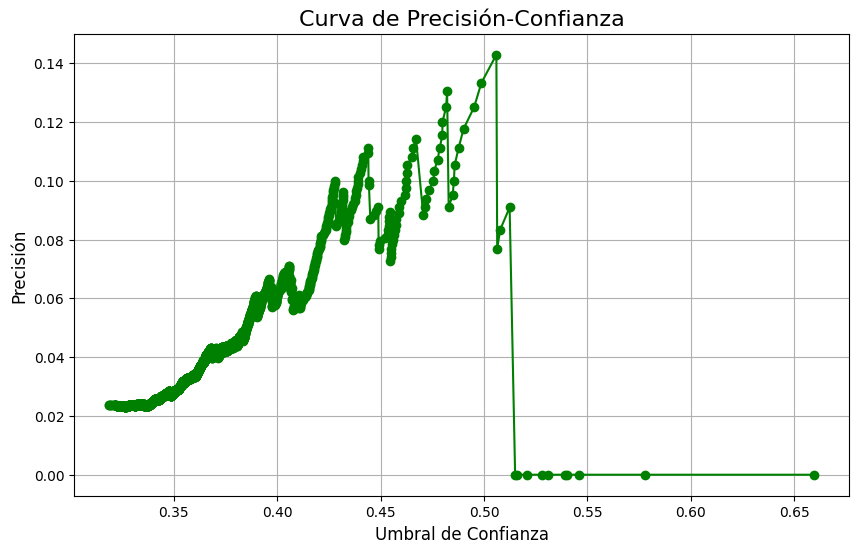


Ilustración 10 *Precisión-Curva de confianza para el modelo Mask R-CNN*

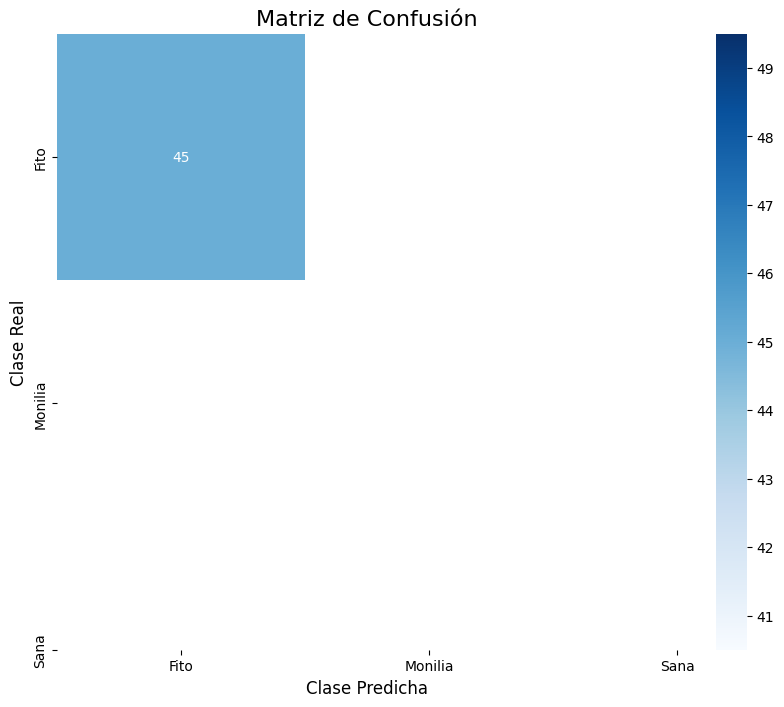


Ilustración 11 *Matriz de confusión para el modelo Resnet-50*

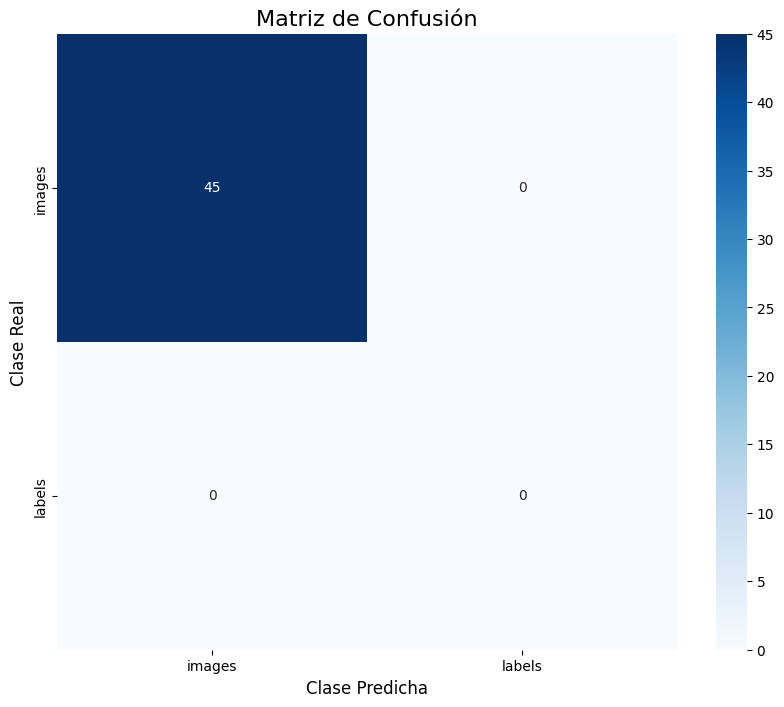


Ilustración 12 *Resultado del modelo Resnet-50 con el set de datos de test*

# Discusión y análisis de resultados

En base a los resultados que se obtuvieron en la evaluación respectiva para cada modelo citado en los puntos anteriores, se puede llegar a determinar datos valiosos.

Para el modelo YOLOv8, el resultado que se obtuvo para cada clase y a nivel general fue:

Tabla 3 Resultados del modelo YOLOv8

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Clase** | **Imágenes** | **Instancias** | **Precisión(P)** | **Recall(R)** | **mAP50** | **mAP50-95** |
| all | 45 | 253 | 0.387 | 0.49 | 0.373 | 0.215 |
| Fito | 17 | 64 | 0.388 | 0.377 | 0.347 | 0.209 |
| Monilia | 16 | 103 | 0.399 | 0.534 | 0.406 | 0.239 |
| Sana | 15 | 86 | 0.373 | 0.558 | 0.367 | 0.196 |

Este modelo tuvo 72 capas y un aproximado de 3 millones de parámetros que la red fue aprendiendo durante el entrenamiento. La capacidad de operaciones que utilizó para hacer su predicción fue de 8.1 GFLOPs.

En líneas generales el modelo tuvo una Precisión(P) de un 38.7% en detecciones correctas, un Recall(R) del 49% que nos indica que el modelo solo fue capaz de encontrar ese porcentaje de objetos existente en cada imagen.

Por otro lado, en las métricas mAP50 se obtuvo un resultado del 0.373 cuando lo esperado es 0.5, es decir que se encuentra por debajo de los valores ideales esperados y la métrica mAP50-95 se obtuvo un 0.215 que de igual manera es bajo y nos sugiere que el modelo no puede delimitar el contorno de las imágenes con mucha precisión.

Analizando el conjunto de resultados para YOLOv8 se puede decir que si bien el modelo en líneas generales es muy rápido pero la precisión y el recall son moderados.

Para el modelo Faster R-CNN, los resultados obtenidos se pueden observar en la Tabla 4

Tabla 4 Resultado del modelo Faster R-CNN

|  |  |
| --- | --- |
| **Métrica** | **Valor** |
| Average Precision(AP) | @[ IoU=0.50:0.95 | area= all | maxDets=100 ] = 0.004 |
| Average Precision(AP) | @[ IoU=0.50 | area= all | maxDets=100 ] = 0.020 |
| Average Precision(AP) | @[ IoU=0.75 | area= all | maxDets=100 ] = 0.001 |
| Average Precision(AP) | @[ IoU=0.50:0.95 | area= small | maxDets=100 ] = -1.000 |
| Average Precision(AP) | @[ IoU=0.50:0.95 | area=medium | maxDets=100 ] = 0.002 |
| Average Precision(AP) | @[ IoU=0.50:0.95 | area= large | maxDets=100 ] = 0.005 |
| Average Recall(AR) | @[ IoU=0.50:0.95 | area= all | maxDets= 1 ] = 0.005 |
| Average Recall(AR) | @[ IoU=0.50:0.95 | area= all | maxDets= 10 ] = 0.037 |
| Average Recall(AR) | @[ IoU=0.50:0.95 | area= all | maxDets=100 ] = 0.083 |
| Average Recall(AR) | @[ IoU=0.50:0.95 | area= small | maxDets=100 ] = -1.000 |
| Average Recall(AR) | @[ IoU=0.50:0.95 | area=medium | maxDets=100 ] = 0.006 |
| Average Recall(AR) | @[ IoU=0.50:0.95 | area= large | maxDets=100 ] = 0.088 |

En la métrica de precisión AP @[ IoU=0.50:0.95 | area= all | maxDets=100 ] = 0.004 general, se promedió la precisión obtenida a través de los diferentes umbrales de intersección desde el 0.50 hasta 0.95 dando como resultado un valor de 0.004 que es extremadamente bajo, lo que nos da como conclusión que el modelo tiene un rendimiento muy pobre.

Ahora estas métricas también se dividen por tamaño de los objetos, en una sub división de small, médium y large. Los resultados que arrojó el modelo en el tamaño small fue de -1000 que nos indica que no se encontraron objetos pequeños en el conjunto de datos empleado o a su vez que el modelo no detectó ninguno. En el tamaño médium se obtuvo un resultado de 0.002 lo que indica que el rendimiento para los objetos de tamaño mediano es demasiado bajo. Finalmente, en el tamaño large se obtuvo un resultado de 0.005 que nos da un indicador de que el modelo tiene un rendimiento ligeramente mejor con los objetos de tamaño grande pese a que el valor obtenido sigue siendo considerablemente bajo.

Con respecto a la métrica de exhaustividad (R) el valor del [maxDets=100]=0.083 conociendo que esto mide el R promediado en todos los IoU con un máximo de 100 detecciones por cada imagen, entendiendo que el resultado significa que el modelo solo está encontrando aproximadamente el 8.3% de los objetos reales, nuevamente evidenciamos que es muy bajo y el modelo no está detectando la mayoría de los objetos que debería encontrar.

Para el modelo Faster R-CNN los resultados globales son demasiado bajos, lo que nos da a entender que el modelo no ha logrado aprender a detectar objetos de manera efectiva.

El último modelo Resnet-50 obtuvo los siguientes resultados que los podemos visualizar en la Tabla 5.

Tabla 5 Resultados del modelo Resnet-50

|  |  |  |
| --- | --- | --- |
| **Época** | **Pérdida** | **Precisión de validación** |
| 1 | 0,3426 | 100% |
| 2 | 0,0548 | 100% |
| 3 | 0,0198 | 100% |
| 4 | 0,0085 | 100% |
| 5 | 0,0085 | 100% |
| 6 | 0,0063 | 100% |
| 7 | 0,0038 | 100% |
| 8 | 0,0039 | 100% |
| 9 | 0,0038 | 100% |
| 10 | 0,0035 | 100% |

El resultado nos muestra un rendimiento excepcional, pero es importante interpretar dichos resultados con mucha cautela ya que una validación del 100% y una pérdida tan baja pueden ser el resultado de que el modelo empleado en nuestro set de datos ha aprendido las características de los datos del set de validación de manera perfecta.

Sin embargo, aunque los resultados suenen excepcionales este puede ser un indicador de que puede existir un overfitting.

Para descartar el tema de overfitting se realiza la validación del modelo con el set de datos de test, obteniendo los siguientes resultados en comparación a la validación inicial:

Tabla 6 Tabla comparativa de resultados para los sets de datos en el modelo Resnet-50

|  |  |  |  |
| --- | --- | --- | --- |
| Set de datos | Precisión(P) | Recall(R) | F1-Score |
| val | 1 | 1 | 1 |
| Test | 1 | 1 | 1 |

# Conclusiones y trabajos futuros

## Conclusiones

Durante el desarrollo de este trabajo, se ha podido visualizar y dar el contexto necesario para los modelos para la detección de plagas en plantaciones de cacao, porque se evidencia los beneficios de llevar la implementación de la IA en el sector de la agricultura, si bien es cierto en cada país la situación puede variar, pero basándonos en el país donde fue concentrado el estudio “Ecuador” existen provincias del sector costero que se dedican a la producción del cacao, y en su gran mayoría son personas autóctonas del lugar que realizan este trabajo de manera empírica ya que la tierra donde realizan los sembríos y cultivan son herencia de sus familiares, pero cabe destacar que pese a su falta de estudios en la mayor parte de la población, es la experiencia lo que les ha llevado a trascender de generación en generación la producción agrícola.

Si se tuviera enfoque y apoyo por parte de los gobiernos autónomos descentralizados como son los municipios provinciales, la inteligencia artificial se convertiría en un aliado estratégico para este grupo de productores ya que tendría la facilidad de detectar de manera temprana cualquier plaga que esté afectando a sus cultivos de cacao, esto ahorraría tiempo y sobre todo dinero, lo que les permitiría invertir en más plantas o a su vez para seguir mejorando los procesos para su negocio.

## Trabajos futuros

Tomando como fuente de conocimiento toda la información recopilada y documentada, se ha podido percatar que en el país de donde soy “Ecuador” aún existen vacíos con respecto a las plagas en los cultivos agrícolas de cacao, si bien es cierto se tiene información general en comparación a otros países, no se encontró suficiente información a otra de las plagas que causa un impacto significativo al cultivo, llamado la escoba de bruja.

Esta plaga es muy conocida por los pequeños agricultores que se dedican en su día a día a trabajar con el cacao, para comprender un poco mas acerca de mi TFM y poder plasmarlo lo mas transparente posible, tuve la oportunidad de realizar trabajo de campo en unos de estos cantones de la zona costera, específicamente en la Provincia de Los Ríos cantón Quinsaloma.

Aquí pude conversar con varios agricultores locales quienes me indicaron los problemas que presentan en los cultivos y las consecuencias que generan las plagas, y me indicaban que no solo a nivel económico, sino también a nivel productivo ya que si una planta que está “en su punto” acorde al jergo coloquial y es infectada por cualquier plaga, tendrían que esperar que el resto de las mazorcas que aún están en crecimiento, lleguen a madurar lo que representan semanas si generar ingresos a su hogar.

Es por ello que este trabajo abre una pregunta ¿Se puede levantar banco de imágenes de la plaga escoba de bruja en el Ecuador?. Con esta información recopilada puede facilitar el entrenamiento de un modelo de machine learning y explorar aún más otros de los modelos que no fueron usados en este trabajo.

# Trabajos citados

Doran, G. T. (1981). There's a S.M.A.R.T. way to write management's goals and objectives. *Management Review (AMA FORUM), 70*, 35-36.

Pérez-Martínez, S. N. (2017). Descripción de plagas en viveros de cacao en el cantón Milagro a partir de diferentes fuentes de información. *Revista Ciencia UNEMI, 10(24)*, 19-38.

Baquerizo, N. C., & Ventocilla, E. J. V. (2022). Evaluación de algoritmos de machine learning en la clasificación de imágenes satelitales multiespectrales, caso: Amazonia Peruana. *Ciencia Latina Revista Científica Multidisciplinar*, *6*(1), 4946–4963.

Cireşan, D. C., Meier, U., Masci, J., Gambardella, L. M., & Schmidhuber, J. (2011). *High-Performance Neural Networks for Visual Object Classification*.

Datta, R., Li, J., & Wang, J. Z. (2005). Content-based image retrieval. *Proceedings of the 7th ACM SIGMM International Workshop on Multimedia Information Retrieval*, 253–262. https://doi.org/10.1145/1101826.1101866

Díaz Rodríguez, P., Han, A., & Merino Mayor, D. (2022). *Análisis de Tráfico en Dispositivos Móviles mediante Técnicas de Aprendizaje Profundo Supervisado.* Https://Docta.Ucm.Es/Entities/Publication/2893acdd-0571-45b2-8a99-19f908da3a5b.

Díaz-Gaxiola, E., Morales-Casas, Z. E., Castro-López, O., Beltrán-Gutiérrez, G., López, I. F. V., & Rendón, A. Y. (2019). Estudio comparativo de arquitecturas de CNNs en hojas de Pimiento Morrón infectadas con virus PHYVV o PEPGMV. *Res. Comput. Sci.*, *148*(7), 289–303.

Díaz-Ramírez, J. (2021). Machine Learning and Deep Learning. In *Ingeniare* (Vol. 29, pp. 182–183). Universidad de Tarapaca. https://doi.org/10.4067/S0718-33052021000200180

Escalante Belmonte, B. (2024). *Clasificación de tractos nerviosos mediante técnicas supervisadas de deep learning*. Https://Rua.Ua.Es/Dspace/Handle/10045/145798.

Ferentinos, K. P. (2018). Deep learning models for plant disease detection and diagnosis. *Computers and Electronics in Agriculture*, *145*, 311–318.

Floridi, L. (2023). *The ethics of artificial intelligence: Principles, challenges, and opportunities*.

Goodfellow, I., Bengio, Y., Courville, A., & Bengio, Y. (2016). *Deep learning* (Vol. 1, Issue 2). MIT press Cambridge.

Hancox-Li, L. (2020). Robustness in machine learning explanations: does it matter? *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 640–647.

He, K., Zhang, X., Ren, S., & Sun, J. (2015). *Deep Residual Learning for Image Recognition*. Https://Openaccess.Thecvf.Com/Content\_cvpr\_2016/Html/He\_Deep\_Residual\_Learning\_CVPR\_2016\_paper.Html.

Hinojosa Merlo, R. (2022). *Aplicación de técnicas de Deep Learning para clasificación de radiografías COVID-19*. Https://Oa.Upm.Es/71366/.

Instituto Nacional Autónomo de Investigaciones Agropecuarias, G. (Ecuador). E. E. L. Sur. D. N. de P. V. (n.d.). *Servicio de diagnóstico de problemas fitosanitarios para el sector agrícola del país*. Https://Repositorio.Iniap.Gob.Ec/Handle/41000/2032.

Jiménez Cleves, G., Londoño Pinilla, D. F., & Garzón Barrero, J. (2024). *Evaluación comparativa del efecto del tamaño de entrenamiento de dos algoritmos de machine learning en la clasificación de una imagen LANDSAT-9*. 1–10. https://doi.org/10.26507/paper.3750

John Lu, Z. Q. (2010). The Elements of Statistical Learning: Data Mining, Inference, and Prediction. *Journal of the Royal Statistical Society Series A: Statistics in Society*, *173*, 693–694. https://doi.org/10.1111/j.1467-985x.2010.00646\_6.x

Khan, S., Naseer, M., Hayat, M., Zamir, S. W., Khan, F., & Shah, M. (2022). Transformers in Vision: A Survey. *ACM Computing Surveys*, *54*. https://doi.org/10.1145/3505244

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2011). *ImageNet Classification with Deep Convolutional Neural Networks*. Https://Proceedings.Neurips.Cc/Paper/2012/Hash/C399862d3b9d6b76c8436e924a68c45b-Abstract.Html.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444. https://doi.org/10.1038/nature14539

LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (2002). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, *86*(11), 2278–2324.

Li, C., Zhen, T., & Li, Z. (2022). Image classification of pests with residual neural network based on transfer learning. *Applied Sciences*, *12*(9), 4356.

Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2021). A survey on bias and fairness in machine learning. *ACM Computing Surveys (CSUR)*, *54*(6), 1–35.

Mikolov, T. (2013). Efficient estimation of word representations in vector space. *ArXiv Preprint ArXiv:1301.3781*, *3781*.

Nassr, M. S., & Abu Naser, S. S. (2018). Knowledge based system for diagnosing pineapple diseases. *International Journal of Academic Pedagogical Research (IJAPR)*, *2*(7), 12–19.

Oyen, D., Kucer, M., Hengartner, N., & Singh, H. S. (2022). Robustness to label noise depends on the shape of the noise distribution. *Advances in Neural Information Processing Systems*, *35*, 35645–35656.

Pérez-Martínez, S., Noceda, C., Zambrano, O., Parra, D., Córdoba, L., & Sosa, D. (2017). Descripción de plagas en viveros de cacao en el cantón Milagro a partir de diferentes fuentes de información. *Revista Ciencia UNEMI*, *10*(24), 19–38.

Pinker, S. (1984). Visual cognition: An introduction. *Cognition*, *18*, 1–63. https://doi.org/10.1016/0010-0277(84)90021-0

Ploetz, R. (2016). The impact of diseases on cacao production: A global overview. In *Cacao Diseases: A History of Old Enemies and New Encounters* (pp. 33–59). Springer International Publishing. https://doi.org/10.1007/978-3-319-24789-2\_2

Ramos, A. R., & Santiago, O. L. (2022). *Sistema de Monitoreo basado en Aprendizaje Profundo en Sistemas Industriales*. Https://Dialnet.Unirioja.Es/Servlet/Articulo?Codigo=9513933.

Ruíz, E., & Domínguez, J. E. (n.d.). *Deep Learning aplicado en imágenes fotoacústicas para la Identificación del cáncer de seno*. Http://Scielo.Sld.Cu/Scielo.Php?Pid=S1684-18592022000100015&script=sci\_abstract&tlng=en.

Salvadó i Gómez, J. (2024). *Detección facial e identificación de personas en imágenes basado en técnicas Deep Learning*. Https://Upcommons.Upc.Edu/Handle/2117/408853.

Sanchez Alvarado, E., SUAREZ, F., Valarezo, N., Herrera, S., GAVILÁNEZ, F., & España Valencia, P. V. (2023). Monitoreo de insectos plaga mediante SIG aplicados al cultivo de Oryza sativa L. en Naranjal, Ecuador. *Manglar*, *20*, 59–67. https://doi.org/10.57188/manglar.2023.007

Sanchez Pardo, S., & More Villegas, F. J. (2023). *Desarrollo de métodos para la clasificación por madurez de la fresa utilizando procesamiento de imágenes digitales y machine learning*. Https://Repositorio.Uss.Edu.Pe/Handle/20.500.12802/13476?Show=full.

Smeulders, A. W. M., Worring, M., Santini, S., Gupta, A., & Jain, R. (2000). Content-based image retrieval at the end of the early years. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *22*(12), 1349–1380.

Suárez Londoño, A. S., Jiménez López, A. F., Castro Franco, M., & Cruz Roa, A. A. (2017). Clasificación y mapeo automático de coberturas del suelo en imágenes satelitales utilizando Redes Neuronales Convolucionales. *Orinoquia*, *21*, 64–75. https://doi.org/10.22579/20112629.432

Türkoğlu, M., & Hanbay, D. (2019). Plant disease and pest detection using deep learning-based features. *Turkish Journal of Electrical Engineering and Computer Sciences*, *27*, 1636–1651. https://doi.org/10.3906/elk-1809-181

Vasileiou, A., Jegelka, S., Levie, R., & Morris, C. (2025). *Survey on Generalization Theory for Graph Neural Networks*.

Wang, S., Xu, D., Liang, H., Bai, Y., Li, X., Zhou, J., Su, C., & Wei, W. (2025). Advances in Deep Learning Applications for Plant Disease and Pest Detection: A Review. In *Remote Sensing* (Vol. 17). Multidisciplinary Digital Publishing Institute (MDPI). https://doi.org/10.3390/rs17040698

Ynfante Martínez, M., Machado Molina, M., Machado García, N., & Velasteguí López, E. (2022). Técnicas de aprendizaje supervisado para la detección y clasificación de enfermedades y defectos en imágenes de frutas: revisión. *Magazine de Las Ciencias: Revista de Investigación e Innovación*, *7*, 1–16. https://doi.org/10.33262/rmc.v7i1.2330