



UNIVERSIDAD LAICA ELOY ALFARO DE MANABÍ
EXTENSIÓN EN EL CARMEN CARRERA DE INGENIERÍA DE SOFTWARE

Creada Ley No. 10 – Registro Oficial 313 de Noviembre 13 de 1985

TEMA

**SISTEMA INFORMÁTICO CON ARQUITECTURA DEEP LEARNING PARA
EVALUACIÓN DE SALUD VEGETAL DE THEOBROMA CACAO EN LA
GRANJA EXPERIMENTAL EL CARMEN**

PROYECTO INTEGRADOR:

MONITOREO FITOSANITARIO DIGITAL DEL CACAO

AUTOR

PINCAY GOMEZ LILIBETH CAROLINA

TUTOR

ING. REASCOS PINCHAO RAÚL SAED, MG

EL CARMEN, 2026



CERTIFICACIÓN DEL DIRECTOR DEL TRABAJO DE TITULACIÓN

	NOMBRE DEL DOCUMENTO:	CÓDIGO: PAT-04-F-004
	CERTIFICADO DE TUTOR	
	PROCEDIMIENTO: TITULACIÓN DE ESTUDIANTES DE GRADO BAJO LA UNIDAD DE INTEGRACIÓN CURRICULAR	REVISIÓN: 1 Página 1 de 1

CERTIFICACIÓN

En calidad de docente tutor de la Extensión El Carmen de la Universidad Laica "Eloy Alfaro" de Manabí, CERTIFICO:

Haber dirigido, revisado y aprobado preliminarmente el Trabajo de Integración Curricular bajo la autoría de la estudiante PINCAY GOMEZ LILIBETH CAROLINA, legalmente matriculada en la carrera de Ingeniería de Software, período académico 2025(2)-2026(1), cumpliendo el total de 288 horas, cuyo tema del proyecto es "SISTEMA INFORMÁTICO CON ARQUITECTURA DEEP LEARNING PARA EVALUACIÓN DE SALUD VEGETAL DE THEOBROMA CACAO EN LA GRANJA EXPERIMENTAL EL CARMEN".

La presente investigación ha sido desarrollada en apego al cumplimiento de los requisitos académicos exigidos por el Reglamento de Régimen Académico y en concordancia con los lineamientos internos de la opción de titulación en mención, reuniendo y cumpliendo con los méritos académicas, científicos y formales, y la originalidad del mismo, requisitos suficientes para ser sometida a la evaluación del tribunal de titulación que designe la autoridad competente.

Particular que certifico para los fines consiguientes, salvo disposición de Ley en contrario.

El Carmen, 31 de julio de 2026.

Lo certifico,



Ing. Saed Reascos Pinchao, Mg.
Docente Tutor
Área: Tecnologías de la Información & Software



Universidad Laica Eloy Alfaro de Manabí Extensión El Carmen

Carrera de Ingeniería de Software

TRIBUNAL DE SUSTENTACIÓN

Títulos de trabajo de titulación:

Sistema Informático con Arquitectura Deep Learning para Evaluación de Salud Vegetal de Theobroma cacao en la Granja Experimental El Carmen.

Modalidad:

Proyecto integrador

Autor:

Pincay Gómez Lilibeth Carolina

Tutor:

Ing. Reascos Pinchao Raúl Saed

Tribunal de sustentación:

Presidente: Ing. Minaya Macias Renelino Wladimir

Miembro: Ing. Arévalo Hermida Rómulo Danilo

Miembro: Ing. Mendoza Villamar Rocio Alexandra

Fecha de sustentación:

27 de agosto de 2026

DECLARACIÓN EXPRESA DE AUTORÍA

Yo, **PINCA Y GOMEZ LILIBETH CAROLINA** con Cédula Nacional de Identidad N.º 131163021-2; declaro que este trabajo de titulación con el tema: “**SISTEMA INFORMÁTICO CON ARQUITECTURA DEEP LEARNING PARA EVALUACIÓN DE SALUD VEGETAL DE THEOBROMA CACAO EN LA GRANJA EXPERIMENTAL EL CARMEN**”, es original de mi autoría, que se han citado las fuentes correspondientes y respetado las disposiciones legales que protegen los derechos de autor vigentes.

Asimismo, autorizo a la ULEAM realizar la digitalización y publicación de este trabajo de titulación en el repositorio digital de conformidad con lo dispuesto en el Art. 144 de la Ley Orgánica de Educación Superior.

Lo certifica,



PINCA Y GOMEZ LILIBETH CAROLINA
C.I.: 131163021-2

DEDICATORIA

Dedico este trabajo en primer lugar a Dios por darme la fortaleza, la sabiduría y la perseverancia necesaria para llegar hasta aquí. A mi madre, por haber estado siempre conmigo, por su amor incondicional, por su esfuerzo, por haber sido padre y madre para mis hermanos y para mí, enseñándonos con su ejemplo el valor de salir adelante aún en los momentos más difíciles. A mi abuela, que aunque ya no se encuentra físicamente conmigo, permanece presente en mis recuerdos y mi corazón, porque fue una persona muy especial en mi vida que siempre me consintió brindándome amor en cada etapa de mi vida. A mis hijas, quienes son mi mayor motivación y la razón por la que decidí formarme profesionalmente, con el deseo de ser para ellas un ejemplo de lucha, constancia y superación.

Lilibeth Pincay

AGRADECIMIENTOS

Expreso mi más sincero agradecimiento a Dios, por permitirme culminar esta etapa tan importante de mi vida y por darme la fortaleza necesaria para seguir adelante en cada momento de dificultad. A mis amigas Anita y Jeny, por su amistad, compañía y por estar presentes cuando más las necesité brindándome ánimo en este proceso. A mi tutor por su paciencia, orientación y apoyo durante el desarrollo de este trabajo, así como por compartir conmigo experiencias y anécdotas de vida que me motivaron a creer que si otros pudieron salir adelante, yo también podía lograrlo. A la Universidad Laica Eloy Alfaro de Manabí, Extensión El Carmen, por brindarme la oportunidad de formarme profesionalmente y permitirme alcanzar una meta tan significativa en mi vida. Finalmente, agradezco a todas las personas que de una u otra manera me apoyaron durante esta etapa, aportando con su tiempo, sus palabras, su compañía o sus enseñanzas para hacer posible este logro.

Lilibeth Pincay

ÍNDICE DE CONTENIDOS

PORTADA	
CERTIFICACIÓN DEL DIRECTOR DEL TRABAJO DE TITULACIÓN.....	I
TRIBUNAL DE SUSTENTACIÓN	¡Error! Marcador no definido.
DECLARACIÓN EXPRESA DE AUTORÍA	III
DEDICATORIA.....	IV
AGRADECIMIENTOS.....	V
ÍNDICE DE GRÁFICOS E ILUSTRACIONES.....	XI
RESUMEN	XIII
ABSTRACT	XIV
CAPÍTULO I.....	1
1 INTRODUCCIÓN.....	1
1.1 Introducción.....	1
1.2 Presentación del tema	2
1.3 Contextualización de problemática	2
1.4 Planteamiento de problema	3
1.4.1 Problematización	3
1.4.2 Génesis del problema.....	3
1.4.3 Estado actual del problema.....	4
1.5 Diagrama de causa – efecto del problema.....	5
1.6 Objetivos.....	5

1.6.1	Objetivo General.....	5
1.6.2	Objetivos Específicos	5
1.7	Justificación	6
1.8	Impactos esperados.....	6
CAPÍTULO II.....		8
2	MARCO TEÓRICO	8
2.1	Antecedentes históricos y del estado del arte	8
2.2	Investigaciones relacionadas	8
2.2.1	Clasificación de imágenes médicas con técnicas de Deep Learning	8
2.2.2	Desarrollo de una aplicación móvil para la detección de enfermedades en la mazorca de cacao utilizando Deep Learning.....	9
2.3	Definiciones conceptuales	9
2.3.1	Variable independiente - Sistema Informático	9
2.3.2	Variable dependiente Salud Vegetal.....	15
2.4	Conclusiones relacionadas al marco teórico.....	20
CAPÍTULO III		22
3	MARCO INVESTIGATIVO.....	22
3.1	Introducción al capítulo	22
3.1.1	Presentación.....	22
3.1.2	Justificación Metodológica.....	23
3.2	Tipo de investigación.....	23
3.2.1	Métodos de investigación	23

3.2.2	Método inductivo-deductivo.....	24
3.2.3	Método de análisis-síntesis.....	24
3.2.4	Método de observación.....	25
3.3	Fuentes de información de datos	25
3.3.1	Fuentes Primarias	25
3.4	Estrategias Operacionales para la recolección de datos	27
3.4.1	Población	27
3.4.2	Muestra	28
3.4.3	Análisis de herramientas de recolección de datos a utilizar	29
3.4.4	Plan de recolección de datos.....	30
3.5	Análisis y presentación de resultados	30
3.5.1	Presentación y descripción de los resultados obtenidos	30
CAPÍTULO IV		37
4	MARCO PROPOSITIVO.....	37
4.1	Introducción.....	37
4.2	Descripción de la propuesta.....	37
4.3	Determinación de recursos	38
4.3.1	Humanos.....	38
4.3.2	Tecnológicos.....	39
4.3.3	Económicos	40
4.4	Desarrollo del sistema de trabajo bajo la metodología.....	30
4.4.1	Descripción del Producto.....	30

4.4.2	Historias del Usuario	31
4.4.3	Product Backlog	32
4.4.4	Planificación de Sprints	32
CAPÍTULO V.....		50
5	EVALUACIÓN DE RESULTADOS.....	50
5.1	Introducción.....	50
5.2	Presentación y monitoreo de los resultados.....	50
5.2.1	Planificación de evaluación	50
5.2.2	Ejecución del monitoreo	51
4.1	Interpretación objetiva	52
CAPÍTULO VI		54
6	CONCLUSIONES Y RECOMENDACIONES	54
6.1	Conclusiones.....	54
6.2	Recomendaciones	56
BIBLIOGRAFÍA		59
ANÉXOS		65
GLOSARIO		69

ÍNDICE DE TABLAS

Tabla 1 - Recursos humanos del proyecto CacaoHealth AI.....	38
Tabla 2 - Recursos Tecnológicos del proyecto CacaoHealth AI.....	39
Tabla 3 - Recursos Económicos del proyecto CacaoHealth AI.....	40
Tabla 4 - Historias de usuario del sistema CacaoHealth AI.....	31
Tabla 5 - Product Backlog del sistema CacaoHealth AI	32
Tabla 6 - Planificación y configuración del entorno	32
Tabla 7 - Recolección de imágenes en campo.....	34
Tabla 8 - Preparación y ampliación del dataset.....	36
Tabla 9 - Entrenamiento del modelo de clasificación	38
Tabla 10 - Optimización del modelo con clase No_Cacao	40
Tabla 11 - Desarrollo de la interfaz gráfica.....	42
Tabla 12 - Pruebas de validación.....	43
Tabla 13 - Exportación e integración del modelo.....	46
Tabla 14 - Evaluación de efectividad del sistema frente a causas del problema	49
Tabla 15 - Comparación entre diagnóstico manual y diagnóstico automatizado	49
Tabla 16 - Planificación de la evaluación del sistema CacaoHealth AI.....	50
Tabla 17 - Resultados de tiempo de respuesta del sistema CacaoHealth AI.....	51
Tabla 18 - Resultados de calidad de clasificación por clases (modelo_v4)	51
Tabla 19 - Interpretación objetiva de los resultados del sistema CacaoHealth AI	
Elaboración propia basada en los resultados de las pruebas de validación ejecutadas ..	53

ÍNDICE DE GRÁFICOS E ILUSTRACIONES

Figura 1- Instalación de Python en VS Code	33
Figura 2 - Instalación TensorFlow.....	33
Figura 3 – VS Code con las dependencias instaladas.....	34
Figura 4 - Visita a la Granja para Captura de Fotografías del cacao.....	35
Figura 5 - Primera organización de imágenes en carpetas	35
Figura 6 - Imágenes organizadas en carpetas	35
Figura 7 - aplicación de data augmentation.....	36
Figura 8 – Incorporación de rotaciones, zoom, brillo y volteo de imágenes.....	37
Figura 9 - Creación de carpetas train y validation.....	37
Figura 10 - • Verificación del balance del dataset por clase.....	38
Figura 11 - Capas densas, Dropout y BatchNormalization	39
Figura 12 - Terminal mostrando el entrenamiento del modelo con 4 clases (dataset_v2)	39
Figura 13 - Evaluación de resultados y generación de gráficas de precisión (modelos anteriores)	40
Figura 14 - Incorporación de clase No_Cacao al dataset	41
Figura 15 - Aplicación de class weights para balancear clases	41
Figura 16 - Rentrenamiento del modelo dataset V4	41
Figura 17 - Evaluación de mejoras y comparación conversiones anteriores.....	42
Figura 18 - Funcionamiento de la interfaz en Tkinter	43
Figura 19 - Tiempo de respuesta	44
Figura 20 - Calidad de respuesta	45
Figura 21 - Terminal con resultados por clase	45
Figura 22 - • Conversión del modelo Keras a formato TensorFlow Lite	46

Figura 23 - archivo clase.txt con las 5 clases del modelo.....	47
Figura 24 - Redacción de la documentación de integración para el equipo móvil	47
Figura 25 - Entrega del modelo TFLite y documentación al responsable de la interfaz móvil.....	48
Figura 26 - Aplicación en funcionamiento (Android)	48
Figura 27 - Anexo 1 Tutor asignado.....	65
Figura 28 - Anexo 2 pruebas iniciales del prototipo en Tkinter	66
Figura 29 - Anexo 3 programando prototipo.....	66
Figura 30 - Anexo 4 Entrevista.....	67
Figura 31- Certificado de análisis.....	68

RESUMEN

El presente trabajo de titulación describe el desarrollo e implementación de CacaoHealth AI, un sistema informático basado en Deep Learning orientado a la evaluación del estado de salud vegetal del cacao (*Theobroma cacao*) en la Granja Experimental El Carmen de la Uleam, Extensión El Carmen. El sistema fue desarrollado en Python y se construyó utilizando la arquitectura MobileNetV2 mediante Transfer Learning, implementado con TensorFlow/Keras y es capaz de clasificar fotografías de plantas de cacao en cinco categorías: Estrés_Abiótico, Fito, Monilia, No_Cacao y Sana. La construcción del modelo se basó en un conjunto propio de 2.097 imágenes reales, tomadas en campo y luego ampliadas a través de técnicas de aumento de datos. El modelo entrenado, que realizó un ajuste fino en las capas superiores de MobileNetV2, logró una precisión general del 81.09% sobre un conjunto completo de 312 imágenes de validación. El sistema incorpora una interfaz gráfica en Tkinter con un tiempo de respuesta promedio de 361 ms y fue exportado a formato TensorFlow Lite para su integración en la aplicación móvil Android del proyecto grupal. La validación de campo se realizó mediante un enfoque mixto, aplicando una encuesta a una muestra de 29 estudiantes de Agropecuaria (seleccionados mediante muestreo discrecional de una población de 389 estudiantes) y una entrevista al Ing. Marco Vinicio de la Cruz como fuente experta, cuyos resultados confirmaron la pertinencia y necesidad del sistema como herramienta de apoyo al diagnóstico fitosanitario del cultivo cacaotero.

Palabras clave: Deep Learning, Transfer Learning, MobileNetV2, salud vegetal, cacao, clasificación de imágenes.

ABSTRACT

This thesis describes the development and implementation of CacaoHealth AI, a computer system based on Deep Learning aimed at the evaluation of the plant health status of cocoa (*Theobroma cacao*) at the El Carmen Experimental Farm of Uleam, El Carmen Extension. The system was developed in Python and was built using the MobileNetV2 architecture through Transfer Learning, implemented with TensorFlow/Keras and is capable of classifying photographs of cocoa plants into five categories: Estrés_Abiótico, Fito, Monilia, No_Cacao and Sana. The construction of the model was based on a set of 2,097 real images, taken in the field and then expanded through data augmentation techniques. The trained model, which fine-tuned the upper layers of MobileNetV2, achieved an overall accuracy of 81.09% over a full set of 312 validation images. The system incorporates a graphical interface in Tkinter with an average response time of 361 ms and was exported to TensorFlow Lite format for integration into the group's Android mobile application. The field validation was carried out through a mixed approach, applying a survey to a sample of 29 Agricultural students (selected through discretionary sampling of a population of 389 students) and an interview with Engineer Marco Vinicio de la Cruz as an expert source, whose results confirmed the relevance and need of the system as a tool to support the phytosanitary diagnosis of cocoa crops.

Keywords: Deep Learning, Transfer Learning, MobileNetV2, plant health, cocoa, image classification.

CAPÍTULO I

1 INTRODUCCIÓN

1.1 Introducción

El Carmen es un cantón perteneciente a la provincia de Manabí, Ecuador, el cual se considera uno de los principales cantones en la producción de cacao de calidad debido a la participación de los agricultores en el cultivo, la producción, la comercialización y la exportación de este producto. Sin embargo, a pesar de que cuentan con buenas condiciones para el cultivo en cuanto a las condiciones microclimáticas los agricultores dependen de una inspección visual para controlar el estado de sus cultivos, lo cual está asociado a métodos de diagnósticos subjetivos y el tiempo que se necesita para recorrer las grandes áreas cultivadas.

Esto provoca un retraso en el diagnóstico de enfermedades, deficiencias nutricionales y deshidratación de las plantas lo cual resulta en pérdidas económicas y errores en los procesos de toma de decisiones relacionados con la gestión de los cultivos. Como alternativa a este problema, se proponen tecnologías basadas en la inteligencia artificial como una solución para acelerar la detección de estos problemas y minimizar la posibilidad de algún diagnóstico retrasado o erróneo. Este es el contexto de esta investigación la cual se enmarca dentro del estado del módulo de la salud de la planta de un proyecto completo que consta de cinco módulos, el cual tiene como objetivo diseñar un sistema computacional basado en un aprendizaje profundo usando una red neuronal convolucional que pueda evaluar el estado de las plantas de *Theobroma cacao* usando imágenes digitales.

1.2 Presentación del tema

El proyecto se desarrolló en la Granja Experimental de la ULEAM, situada en la región del cantón El Carmen en la provincia de Manabí, Ecuador. El Cantón El Carmen se destaca como una de las zonas con mayor producción de cacao de alta calidad en Ecuador. Participando activamente en los procesos como la siembra, la producción, la comercialización y la exportación de este producto agrícola tan importante en la región.

1.3 Contextualización de problemática

El Carmen es una región conocida por su favorable microclima para el cultivo de cacao, presenta un clima tropical húmedo, con lluvias frecuentes y temperaturas que ayudan a que los árboles de cacao crezcan de forma saludable. Sin embargo, en la zona los agricultores de la zona no cuentan aún con suficiente información técnica o herramientas tecnológicas para detectar a tiempo las enfermedades que afectan el cultivo, lo que retrasa muchas veces el control oportuno y por ende afecta el manejo adecuado de las plantaciones.

El proyecto integrador se dividió en cinco módulos, cada uno fue asignado a un integrante del grupo; el módulo de salud vegetal correspondió al presente trabajo. Este módulo se refiere principalmente a determinar si la planta está sana o presenta estrés abiótico, es decir, si tiene deficiencia o exceso de riego, luz o nutrientes. Los módulos restantes del proyecto integrador se orientan a la detección de plagas y hongos.

Descomposición del proyecto integrador por temas:

- a. Entrenamiento IA plagas
- b. Entrenamiento IA para hongos-líquenes
- c. Entrenamiento de salud vegetal
- d. Interfaz para la detección de anomalías

e. **Recomendaciones de Chatbot**

El objetivo principal es desarrollar y aplicar el componente tres del módulo relacionado con la salud vegetal del cacao, enfocado en el diagnóstico del estado de salud de las plantas mediante el análisis de sus hojas, considerando aspectos relacionados con el agua, la luz, la temperatura y los factores del suelo, entre otros factores que afectan el crecimiento de la planta.

1.4 Planteamiento de problema

La alta demanda de cacao en El Carmen ha incrementado la presión sobre los agricultores, que se ven obligados a cultivar más y obtener producciones de mejor calidad. Aun así, el monitoreo de salud vegetal sigue realizándose en gran parte de forma manual, lo cual impide la detección de enfermedades o carencias nutricionales en las plantas. Esto repercute en daños financieros y dificulta tomar decisiones adecuadas sobre cómo gestionar los cultivos. De ahí surge que las nuevas tecnologías basadas en inteligencia artificial puedan ayudar a detectar las condiciones problemáticas de los cultivos de cacao con mayor rapidez, mejorar el tratamiento y reducir riesgos.

1.4.1 Problematicación

¿Cuáles son las principales limitaciones que enfrentan los agricultores del cantón El Carmen al depender de los métodos tradicionales de monitoreo fitosanitario en los cultivos de cacao?

¿De qué manera la falta de herramientas digitales accesibles afecta la capacidad de los productores locales para mantener rendimientos estables y aplicar prácticas ambientalmente responsables?

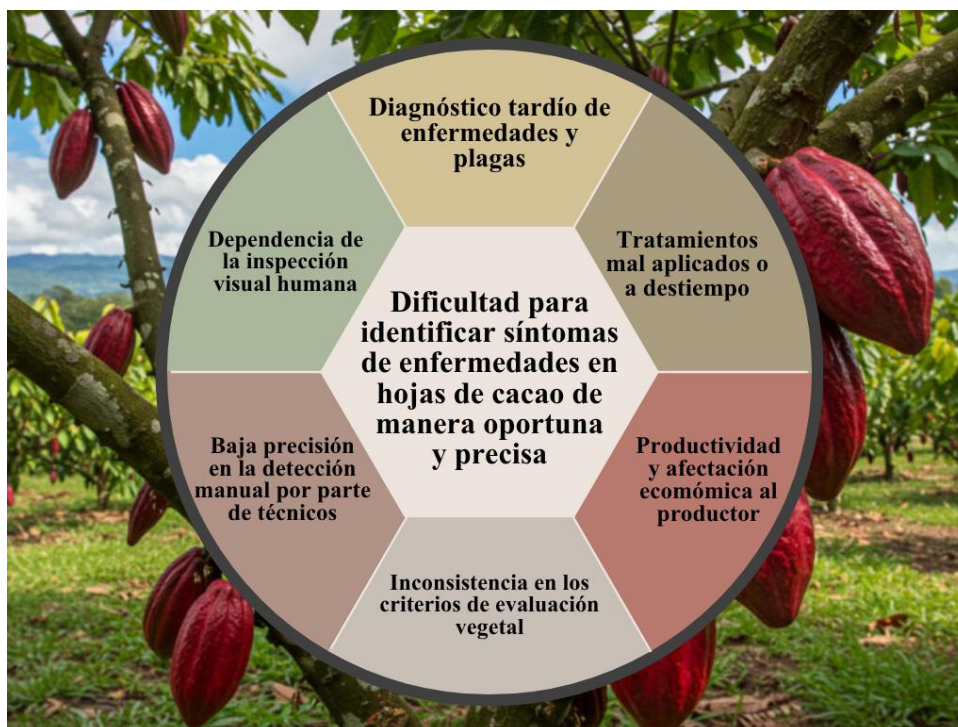
1.4.2 Génesis del problema

El origen del problema radica en la dependencia de métodos tradicionales de observación visual y la ausencia de soluciones tecnológicas económicas. El incremento en la producción exige una capacidad de respuesta ágil por parte de los productores, pero la falta de sistemas automatizados retrasa los diagnósticos y puede resultar en el uso excesivo o incorrecto de productos fitosanitarios. Se requiere una plataforma que permita el entrenamiento de modelos de diagnóstico automático y que ofrezca soporte técnico integral al agricultor, facilitando información sobre el estado real de las plantas.

1.4.3 Estado actual del problema

Hoy en día, los agricultores de cacao del cantón El Carmen dependen casi totalmente de inspecciones visuales para identificar problemas. La subjetividad y el límite de tiempo restringen efectividad en el monitoreo de grandes áreas. Esto dificulta la identificación de enfermedades como la moniliasis, escoba de bruja, hongos, deficiencias nutricionales y estrés hídrico. Las innovaciones mundiales no han llegado a la región de manera práctica y asequible, limitando la capacidad de respuesta y la sostenibilidad productiva local.

1.5 Diagrama de causa – efecto del problema



1.6 Objetivos

1.6.1 Objetivo General

Desarrollar un sistema informático basado en Deep Learning que permita entrenar modelos para la evaluación eficiente automatizada de la salud vegetal en cultivos de *Theobroma cacao* usando imágenes digitales, optimizando el manejo preventivo, las decisiones técnicas.

1.6.2 Objetivos Específicos

- Analizar fundamentos teóricos de redes neuronales convolucionales (CNN) especializadas en diagnóstico fitosanitario
- Analizar la aplicación de las CNN al diagnóstico de la salud vegetal en los cultivos de cacao.
- Recopilar un conjunto de imágenes representativas de hojas de cacao en diferentes estados sanitarios, empleando técnicas adaptadas a la realidad de campo.

- Validar el conjunto de imágenes recopiladas mediante criterios de calidad y representatividad, asegurando que sea adecuado para el entrenamiento del modelo.
- Implementar un prototipo móvil basado en el modelo entrenado, capaz de detectar y clasificar la salud vegetal de las hojas de cacao a partir de imágenes digitales.
- Evaluar en campo, qué tan precisa, sensible, útil resulta la aplicación del sistema en condiciones reales del cultivo.

1.7 Justificación

Este proyecto propone un estudio que busca incrementar el manejo fitosanitario del cacao en el cantón El Carmen. Con el empleo de la inteligencia artificial, se busca fortalecer una agricultura más eficiente e integral, ofreciendo precisión en los procesos. Asimismo, busca simplificar el acceso a diagnóstico de calidad para agricultores menores y medianos. Facilitar recomendaciones inmediatas y confiables intensifica el impacto social y económico positivo que puede lograrse con herramientas tecnológicas apropiadas, contribuyendo a una agricultura sostenible y competitiva.

1.8 Impactos esperados

1.8.1 Impacto tecnológico

El sistema impulsará la adopción de IA en agricultura de precisión, entrenando modelos de redes neuronales convolucionales (CNN) para identificar la salud vegetal de los cultivos con precisión mayor al 85%, tomando como referencia las tecnologías ya existentes para detección de malezas. Facilitará el acceso a prototipos móviles de bajo costo para productores de pequeña escala, integrando monitoreo en tiempo real y reduciendo la desigualdad tecnológica en zonas rurales como El Carmen.

1.8.2 Impacto social

El proyecto mejorará la capacidad de agricultores locales para tomar decisiones oportunas al facilitar el acceso a diagnósticos apoyados en la tecnología de inteligencia artificial. Esto contribuirá a reducir pérdidas económicas asociadas a enfermedades fitosanitarias, fortalecerá las capacitaciones técnicas de los productores y podrá generar nuevas oportunidades de formación y empleo vinculadas al sector agro tecnológico. Asimismo, el uso de una herramienta accesible promoverá mayor autonomía en las comunidades cacaoteras y reforzará su capacidad productiva.

1.8.3 Impacto ecológico

Desde el punto de vista ecológico el sistema permitirá optimizar el uso de insumos fitosanitarios mediante la detección temprana de enfermedades, reduciendo la aplicación innecesaria de productos químicos. Esto favorecerá la implementación de prácticas de manejo integrado de plagas, con efectos positivos en la conservación de la biodiversidad, la salud del suelo y la sostenibilidad del agroecosistema. Impulsará prácticas sostenibles, reduciendo la contaminación y fortaleciendo la resiliencia climática en cultivos de *Theobroma cacao*.

CAPÍTULO II

2 MARCO TEÓRICO

2.1 Antecedentes históricos y del estado del arte

La evolución de las redes neuronales artificiales, actualmente más conocidas como aprendizaje profundo (Deep Learning), se inicia en los años 40. El término Deep Learning cobró fuerza cerca del 2000, consolidándose en 2006 gracias a innovaciones en el entrenamiento de estructuras profundas. Un punto de inflexión en la evolución de la visión artificial fue la aparición de AlexNet en 2012, modelo que mejoró notablemente el desempeño en ImageNet y consolidó el uso de las CNN como una referencia en el procesamiento de imágenes (Berzal, 2018). A partir de ese momento el Deep Learning comenzó a ganar protagonismo por su capacidad para reconocer patrones visuales complejos con mayor precisión que los métodos tradicionales.

Estos avances han sido trasladados a la agricultura de precisión, especialmente en el monitoreo fitosanitario de cultivos como el cacao. Gracias al análisis de imágenes digitales de las hojas, es posible identificar a tiempo enfermedades como la moniliasis y la escoba de bruja, así como deficiencias nutricionales y condiciones de estrés abiótico o biótico que afectan el desarrollo del cultivo. Las tecnologías de Deep Learning aportan a este proceso agilizando el diagnóstico, disminuyendo los errores de evaluación y apoyando un manejo más eficiente y sostenible a los insumos agrícolas.

2.2 Investigaciones relacionadas

2.2.1 Clasificación de imágenes médicas con técnicas de Deep Learning

Para esta tarea se han desarrollado diferentes modelos de Inteligencia Artificial, basándose en Deep Learning y, dentro de este, en redes neuronales convolucionales.

Durante el desarrollo del proyecto se ha empleado una metodología de Data Science (Ciencia de Datos). Esta metodología es ampliamente utilizada en proyectos que conlleven realizar un modelo basado en Inteligencia Artificial. En términos generales, consiste en comprender el problema, extraer los datos que van a ser utilizados, procesarlos de forma que sean utilizables para el proyecto, y clasificarlos en función de la variable deseada (García, 2023).

2.2.2 Desarrollo de una aplicación móvil para la detección de enfermedades en la mazorca de cacao utilizando Deep Learning

Aunque existen diversas estrategias para prevenir las enfermedades del cacao, ninguna garantiza por completo la eliminación de la afección. Por ello, resulta necesario contar con métodos que permitan identificar de manera manual las mazorcas afectadas. En este contexto, el estudio de Ruilova Cumbicos (2022) propone desarrollar un modelo de Deep Learning capaz de reconocer Fitóftora y Monilia en mazorca de cacao. Debido a su capacidad para detectar varios objetos dentro de una misma imagen, se eligió Yolo como algoritmo de detección, y el modelo fue integrado posteriormente en una aplicación móvil que permite capturar imágenes desde la cámara o la galería para ejecutar el análisis.

2.3 Definiciones conceptuales

2.3.1 Variable independiente - Sistema Informático

La variable independiente de este estudio es el sistema computacional, el cual se refiere al conjunto de componentes de software y arquitectura del sistema de inteligencia artificial capaz de realizar la evaluación automatizada de la salud de las plantas de cacao. A continuación se consideran los fundamentos teóricos que sustentan el desarrollo del sistema, los cuales incluyen los principios de la gobernanza ética de los sistemas inteligentes, la arquitectura de redes convolucionales y recurrentes para la visión por

computadora, los modelos de atención, la integración con sistemas distribuidos y el aprendizaje por refuerzo profundo.

2.3.1.1 Ética y gobernanza de sistemas inteligentes

La gobernanza ética para agricultura en sistemas inteligentes debe priorizar la rendición de cuentas y la transparencia, en aplicaciones agrícolas especialmente cuando la IA toma decisiones sobre salud vegetal autónomamente. En este sentido Mittelstadt et al. (2016) argumenta que los marcos de gobernanza deben incluir auditorías algorítmicas para identificar sesgos en modelos de Deep Learning, garantizando equidad en el monitoreo de cultivos como lo es el cultivo de cacao.

Además, la gobernanza responsable en IA agrícola debe integrar principios que estén orientados como la privacidad por diseño y la explicabilidad, para proteger a las personas y mejorar la comprensión del sistema, lo cual resulta especialmente importante en arquitecturas distribuidas y en fincas remotas donde se capturan los datos y la toma de decisiones atraviesan distintos actores y fases del flujo técnico. En esa línea, Jobin et al., (2019) indica que la integración efectiva de la IA y valores humanos exige multidisciplinariedad, capaz de traducir principios éticos que sean aplicables y prácticos al monitoreo agrícola.

2.3.1.2 Redes convolucionales (CNN) en visión y procesamiento inteligente

Las CNN han revolucionado la visión artificial mediante extracción jerárquica de características locales mediante convoluciones, permitiendo detectar enfermedades en hojas de cacao desde imágenes aéreas con precisiones superiores al 95%, esta capacidad se potencia con arquitecturas como ResNet, que para evitar la degradación del gradiente en redes ultra-profundas utiliza conexiones residuales. Según Goodfellow et al., (2016) y He et al., (2016), las estructuras permiten procesar dataset complejos de agricultura de

manera eficaz garantizando un rendimiento alto en tareas críticas con la clasificación de estrés hídrico.

Para el monitoreo de plantaciones remotas en tiempo real, la CNN eficiente como MobileNet es clave, reduce el costo computacional mediante convoluciones separables y así facilita su despliegue en dispositivos móviles o drones. Como indica Howard et al., (2017), este tipo de arquitectura resulta útil cuando se busca equilibrio entre precisión y eficiencia.

2.3.1.3 Redes recurrentes (RNN, LSTM, GRU) y procesamiento secuencial

Las Redes Neuronales Recurrentes (RNN) son importantes para analizar la evolución de datos agrícolas, ya que permite identificar relaciones y dependencias en secuencias temporales o procesos que cambian con el tiempo, esenciales para series de monitoreo de la salud del cacao. Sin embargo, para superar estas limitaciones técnicas surgieron las LSTM. Este tipo de redes usan puertas de entrada y puertas de olvido para memorizar patrones prolongados, lo que resulta esencial para predecir brotes epidémicos o analizar sensores IoT sin perder información relevante en periodos extendidos del cultivo (Chollet, 2021).

Las GRU simplifican LSTM al fusionar puertas en update y reset, ofreciendo un rendimiento similar con menos parámetros para despliegues en tareas de predicción meteorológicas, esto para dispositivos de baja potencia que requieren mayor eficiencia. Por otro lado, la aplicación de CNN-LSTM (para integración de modelos hídricos) permite procesar datos que son simultáneamente espacio-temporales en imágenes satelitales para agricultura de precisión (Goodfellow et al., 2016).

2.3.1.4 Modelos transformadores y atención (Attention Mechanisms)

Los transformadores representan un cambio de paradigma que elimina la recurrencia en favor de mecanismos de auto-atención, permitiendo procesamiento paralelo y captura de dependencias globales en secuencias extensas de datos agrícolas. La atención de producto escalar y la atención de múltiples cabezas permiten que estos modelos identifiquen que información es más relevante dentro de un contexto específico y capten relaciones complejas en conjuntos de datos variados sobre salud vegetal. De acuerdo con Rothman (2021) y Tunstall et al. (2022), esta capacidad es fundamental para analizar de manera eficiente información no estructurada y apoyar el diagnóstico en agricultura de precisión.

Por otro lado, para manejar la alta resolución de capturas en plantaciones de cacao, los avances como Swin Transformers implementan una atención jerárquica de ventana desplazada que permite una inferencia notable más rápida en dispositivos de supervisión. La arquitectura es capaz de procesar imágenes en tiempo real en aplicaciones móviles agrícolas sin requerir equipos de alto rendimiento, lo cual facilita el diagnóstico fitosanitario en una finca (Rothman, 2021).

2.3.1.5 Integración de Deep Learning con sistema distribuido e inferencia en la nube

La convergencia entre el Deep Learning y los sistemas distribuidos es fundamental para escalar modelos masivos, ya que permite la carga de cómputo en clústeres híbridos, lo que posibilita inferencias en tiempo para el monitoreo agrícola remoto. De acuerdo con los principios de diseños de sistemas intensivos en datos, el procesamiento distribuido mejora la escalabilidad, la disponibilidad y la capacidad de respuesta en entornos con grandes volúmenes de información. Asimismo, la arquitectura en la nube permite atender múltiples solicitudes de drones y sensores, optimizando el

desempeño sobre conjuntos de datos a gran escala (Kleppmann, 2017; Tanenbaum & Van Steen, 2007).

Por otra parte, la orquestación de servicios en contenedores y la coordinación de recursos de cómputo contribuyen a aumentar el throughput de las predicciones y hacer más robustos el monitoreo autónomo. En ese sentido, las arquitecturas distribuidas favorecen el despliegue eficiente de modelos de visión por computadora en escenarios agrícolas donde la escalabilidad y estabilidad operativa son factores esenciales (Beyer et al., 2016; Murphy et al., 2016).

2.3.1.6 Deep Reinforcement Learning (DRL) en sistemas autónomos

El aprendizaje por refuerzo profundo entrena agentes autónomos mediante un proceso de ensayo y error, en el que el comportamiento se guía por recompensas. Esto lo hace adecuado para optimizar rutas de drones en la inspección de plantaciones de cacao. Sutton y Barto (2018) explican que el Q-Learning profundo permite que los agentes aprendan estrategias de navegación eficientes en entornos cambiantes, sin depender de supervisión constante.

Asimismo, el uso de algoritmos de control continuo como PPO y SAC mejora la estabilidad al gestionar recursos, lo cual es especialmente útil en sistemas de riego autónomo del cultivo de cacao. Schulman et al., (2024) señala que PPO facilita entrenar de forma segura en simulación antes de llevarlos a hardware real, manteniendo un buen equilibrio entre explorar el entorno aprovechar lo aprendido.

2.3.1.7 MobileNetV2 y arquitecturas eficientes para dispositivos móviles.

Mientras que los grandes modelos de CNN requieren una alta capacidad de cómputo MobileNetV2 fue creado especialmente para ser usado en zonas con recursos limitados ya que utiliza la convolución separable por profundidad y técnicas de bloques

residuales invertidos y esto disminuye en gran medida el número de parámetros sin perder mucho de la precisión(Sandler et al., 2018). Esta característica lo hace más aplicable en dispositivos móviles para ser usado por agricultores en el campo. En el modelo que hemos usado en este estudio para diagnosticar enfermedades de plantas en plantas de cacao.

Otra importante fortaleza de esta arquitectura es que puede controlar el equilibrio entre la precisión y el rendimiento con la ayuda del multiplicador de ancho de parámetro, el cual permite variar el tamaño de modelos según el hardware disponible sin tener que reconstruir toda la arquitectura(Sandler et al., 2018). Este tipo de proyectos donde el producto final debe funcionar en teléfonos inteligentes de gama media de pequeños y medianos productores en el cantón El Carmen se vuelve esencial.

2.3.1.8 Transfer Learning y aumento de datos en visión agrícola

El aprendizaje por transferencia permite reutilizar el conocimiento de modelos preentrenados en grandes conjuntos de imágenes y adaptarlo a tareas específicas con datasets más reducidos, como es el caso del diagnóstico de enfermedades en hojas de cacao(Panigrahi et al., 2021). De forma complementaria, las técnicas de aumento de datos como rotaciones, cambio de brillo y volteos horizontales permiten aplicar ampliar artificialmente la variedad de un conjunto de imágenes limitado, mejorando la capacidad de generalización del modelo, un aspecto especialmente relevante cuando ciertas clases, como Monilia, cuentan con menos muestras disponibles que otras.

Esta combinación de técnicas resulta particularmente útil en contextos agrícolas donde la recolección de imágenes en campo está limitada por factores como el acceso a las plantaciones, las condiciones climáticas y el tiempo disponible para el muestreo. Al partir de los pesos preentrenados del MobileNetV2 y reentrenar únicamente las capas finales con el conjunto propio de imágenes de hojas de cacao, es posible obtener un

modelo funcional incluso con un número reduciendo de muestras por clase, reduciendo así el riesgo de sobreajuste frente al entrenamiento de una red desde cero.

2.3.2 Variable dependiente Salud Vegetal

La variable dependiente del presente estudio corresponde a la salud vegetal, entendida como el estado fisiológico y sanitario de las plantas de cacao, determinada por la interacción entre factores bióticos y abióticos que inciden en su desarrollo. A continuación, se abordan los fundamentos teóricos relacionados con el microbioma vegetal y la salud del suelo, el control biológico de enfermedades, el impacto del cambio climático en la sanidad del cultivo y las aplicaciones de nanotecnología en la detección temprana de patologías, aspectos que sustentan la comprensión del fenómeno evaluado por el sistema desarrollado.

2.3.2.1 Microbioma vegetal y salud del suelo

El microbioma rizosférico es vital para la nutrición y defensa del cacao. Microorganismos como *Bacillus* spp. optimizan la solubilización del fósforo en suelos ácidos tropicales, mejorando la absorción de nutrientes (Mendes et al., 2013). Esta diversidad microbiana previene enfermedades mediante competencia nutricional y producción de antibióticos naturales; estudios metagenómicos en *Theobroma cacao* confirman que una alta biodiversidad en el rizobioma está directamente vinculada con una mayor resistencia frente a patógenos como *Phytophthora* (Trivedi et al., 2020).

Además, los microbiomas funcionales ayudan a mitigar el estrés abiótico por factores del cambio climático mediante la síntesis de fitohormonas y el priming de defensas sistémicas (Berendsen et al., 2024). Cuando se aplican biofertilizantes basados en consorcios rizobiales no solo restaura suelos degradados, sino que impulsa la productividad sostenible. Según Olanrewaju et al., (2025), reporta que en fincas de

Manabí el uso de estos inoculantes ha incrementado el rendimiento hasta un 25%, posicionando a la biotecnología microbiana como un pilar para la eficiencia agrícola actual.

2.3.2.2 Control biológico y agentes antagonistas

El control biológico en el cacao se basa en el empleo antagonista como *Trichoderma* spp., los cuales actúan mediante micoparasitismo, antibiosis y competencia para inhibir y lisar patógenos como la *Moniliophthora rorerri* (Leiva et al., 2020). Estos mecanismos han mostrado ser útiles en el desarrollo de estrategias de control biológico aplicadas al cultivo de cacao, especialmente por su capacidad para reducir la incidencia de la enfermedad y mejorar la sanidad en la de las mazorcas. De manera complementaria, bacterias del género *Pseudomonas* han sido ampliamente estudiadas por su capacidad de inducir resistencia sistemática y suprimir enfermedades mediante la producción de sideróforos, antibióticos y otros mecanismos de biocontrol (Bakker, Pieterse, & van Loon, 2007). En este sentido el uso de bacterias beneficia la rizosfera contribuye a reforzar la defensa de la planta frente a diversos patógenos.

2.3.2.3 Impacto del cambio climático en fitopatología

El cambio climático se ha convertido en un factor que complica cada vez más la producción de cacao. Porque modifica condiciones básicas para el cultivo, como la temperatura y la humedad, y eso puede favorecer la presencia y expansión de patógenos. En Ecuador se ha documentado que estas variaciones climáticas alteran la distribución potencial de hongos asociados al cacao y aumentan el riesgo para las zonas predictoras (Plasencia-Vázquez et al., 2022; Ferrer-Sánchez et al., 2022). Además, algunos estudios muestran que ciertos patógenos del cacao responden de forma directa a las condiciones

ambientales, especialmente cuando hay más humedad o temperaturas favorables para su desarrollo (Troya Guerrero & Pino Melendez, 2023).

A esto se suma que la sequía prolongada debilita a la planta porque le generan estrés hídrico y reducen su capacidad para defenderse frente a enfermedades. Cuando el cacao está bajo estas condiciones, se vuelve más vulnerable al ataque de patógenos oportunistas y a la pérdida de vigor general del cultivo. Por eso, el cambio climático no solo afecta al rendimiento del cacao, sino también su estabilidad sanitaria y su resistencia natural frente a enfermedades (Ferrer-Sánchez et al., 2022).

2.3.2.4 Nanotecnología aplicada a sanidad vegetal

La implementación de nanoformulaciones permite la encapsulación de fungicidas para una liberación controlada en el cultivo de cacao, lo que puede mejorar la eficiencia del tratamiento y reducir pérdidas por degradación. En estudios con nanopartículas de cobre se ha observado un mejor desempeño frente a *Phytophthora*, incluso con menores dosis formuladas que los productos comerciales convencionales (Khan & Rizvi, 2024; Gogos et al., 2012). De forma paralela, las nanopartículas de plata han mostrado una amplia acción antimicrobiana, ya que pueden generar especies reactivas de oxígeno que eliminan patógenos (Davila Costa & Romero, 2025). Asimismo, la síntesis verde de este tipo de materiales ha sido estudiada para mejorar su estabilidad y su uso en aplicaciones agrícolas.

Por otro lado, los nanocarriers facilitan el transporte de moléculas funcionales en plantas y se están investigando como vehículos para aplicaciones dirigidas en sanidad vegetal (Gogos et al., 2012). Además, el desarrollo de nanosensores permite la detección temprana de biomarcadores de estrés y toxinas; en este ámbito, los nanosensores basados

en nanopartículas de oro han demostrado utilidad para identificar aflatoxinas de forma rápida y precisa (Li et al., 2021).

2.3.2.5 Tecnologías de diagnóstico molecular y avanzado

El diagnóstico de patógenos en el cacao ha evolucionado hacia métodos rápidos y de alta precisión, como la amplificación isotérmica, que permite detectar *Phytophthora* en campo en menos de una hora sin necesidad de laboratorios complejos (Ivanov et al., 2021; Zou et al., 2020). Esta técnica ofrece una alternativa práctica para el diagnóstico oportuno de enfermedades en plantas (Yang et al., 2025; Németh & Kovács, 2025).

La metagenómica shotgun permite identificar comunidades microbianas complejas y detectar patógenos asociados a disbiosis en la rizosfera, aportando información valiosa sobre la salud del cultivo (Jansson et al., 2024). A esto se suma la secuenciación mediante nanoporos, que permite obtener información genómica en tiempo real y facilita la caracterización rápida de variantes patógenas, incluyendo aquellas vinculadas con resistencia a fungicidas y manejo fitosanitario (Ivanov et al., 2021).

2.3.2.6 Bioseguridad vegetal y organismos cuarentenarios

La biodiversidad en el cultivo de cacao se fundamenta en los protocolos de las NPPO (Organizaciones Nacionales de Protección Fitosanitaria), los cuales regulan la introducción de germoplasma para asegurar que esté libre de patógenos como *Ceratocystis* (End et al., 2024; IPPC, 2024). En este sentido, el establecimiento de barreras fitosanitarias en países como Ecuador es vital para prevenir la propagación de enfermedades transfronterizas (IPPC, 2024).

Por otra parte, el tratamiento de desinfección con vapor caliente ha demostrado ser una solución eficaz para eliminar endófitos cuarentenarios en las semillas de cacao sin comprometer su viabilidad. La integración de sistemas de trazabilidad basados en

blockchain asegura la transferencia en la cadena de custodia de materiales en cuarentena. De acuerdo con Kshetri (2021), la combinación de metadatos de pruebas PCR con registro inmutables permite un seguimiento más exhaustivo de patógenos, fortaleciendo la bioseguridad del sector cacaotero.

2.3.2.7 Epidemiología y modelización de enfermedades emergentes

La epidemiología moderna emplea modelos SIR (susceptibles-infectados-recuperados) de carácter espacial para predecir la dispersión de patógenos como la *Moniliophthora*, impulsada principalmente por factores ambientales como el viento y la lluvia (Keeling & Rohani, 2008; Cressie & Wikle, 2011). La calibración de estos modelos con datos locales permite fijar umbrales críticos para intervenir oportunamente a nivel fitosanitario (Gilligan, 2008).

Asimismo, la simulación de brotes en cacao se basa en el análisis de redes dinámicas, que permite estimar parámetros epidemiológicos relevantes para comprender la propagación del patógeno (Railsback & Grimm, 2019). Estos modelos se integran en enfoques basados en agentes, combinando el Manejo Integrado de Plagas de forma predictiva para anticipar enfermedades emergentes.

2.3.2.8 Manejo integrado de enfermedades y plagas (MIP/IDM)

El manejo integrado de Plagas y Enfermedades (MIP/IDM) constituye un enfoque holístico para la producción de cacao sostenible, al combinar tácticas culturales, biológicas y químicas fundamentales en un monitoreo constante. El IDM (Manejo Integrado de Enfermedades) es un modelo que se define como un sistema que integra diversas estrategias de manejo (naturales, químicas y físicas). La clave es el seguimiento, monitoreo continuo del cultivo y la supervisión para poder implementar acciones de control solo en un momento preciso activándose al alcanzar límites de daño económico.

Los componentes clave de este plan (IDM) incluye pronóstico de enfermedades y la aplicación de tecnologías innovadoras.

Según la literatura técnica manejo integrado, esta estrategia permite optimizar los umbrales económicos de acción, asegurando que las intervenciones solo se realicen cuando sea estrictamente necesario (Kogan, 1998; Pedigo & Rice, 2009). Por otro lado, la integración de tecnologías avanzadas, como el internet de las cosas (IoT) y los sistemas de pronóstico, permite dirigir aplicaciones fitosanitarias focalizadas con mayor precisión (Mahlein, 2016). Este modelo de MIP predictivo se fortalece con el uso de resistencia genética y la aplicación de bioinsumos, elementos clave para desarrollar plantaciones de cacao resilientes (Savary et al., 2019; Proyecto Chocolate4all, s. f.).

2.4 Conclusiones relacionadas al marco teórico

La revisión teórica realizada fundamenta de manera sólida el desarrollo de una plataforma de diagnóstico fitosanitario impulsada por IA para el cultivo de cacao en el cantón El Carmen provincia de Manabí en Ecuador. El sistema busca proveer a los agricultores y técnicos locales una herramienta digital que es capaz de identificar plagas y enfermedades mediante el análisis visual de hojas, frutos y suelo. Para lograr esto la arquitectura del sistema implementará modelos avanzados de Deep Learning, como son las redes neuronales y transformadores adaptados a la visión computacional, entrenados a las condiciones agroecológicas específicas del Ecuador. Esta tecnología se implementó en un entorno móvil diseñado para entornos rurales, garantizando una interfaz intuitiva en español y con capacidad de capturar datos en campo a través de dispositivos cotidianos.

Con el objetivo de mejorar las limitaciones de conectividad en las zonas agrícolas de El Carmen, la plataforma integrará una infraestructura híbrida, distribuyendo la carga de procesamiento entre el dispositivo móvil y servidores externos para optimizar los

tiempos de respuestas. Asimismo, la detección automatizada, el sistema cruzará registros de manejo y datos climáticos provistos por los usuarios para generar recomendaciones personalizadas de manejo integrado de plagas y enfermedades alineadas con las normativas agronómicas del Ecuador. En otras palabras, este marco conceptual valida la integración de la IA, el conocimiento fitopatológico regional y los sistemas distribuidos como una solución tecnológica factible para reducir la pérdida de cosechas, fomentar la sostenibilidad y agilizar la toma de decisiones del sector cacaotero en Ecuador.

CAPÍTULO III

3 MARCO INVESTIGATIVO

3.1 Introducción al capítulo

Este capítulo analiza la metodología de la investigación de esta tesis, revelando la metodología que constituye la base del desarrollo de este estudio. En este capítulo se discutirá la metodología utilizada en esta investigación, los métodos de análisis del problema, las fuentes de información revisadas y los métodos utilizados para la recolección y el procesamiento de los datos, para garantizar la validez y la confiabilidad de los resultados discutidos en este estudio.

3.1.1 Presentación

En el presente capítulo se desarrolla el marco investigativo de la tesis, el cual permite establecer la base metodológica utilizada para el análisis del problema planteado. La investigación se orienta a la detección de enfermedades en hojas de cacao mediante una propuesta tecnológica basada en Deep Learning, con el propósito de identificar las principales dificultades que presentan los estudiantes de agropecuaria al momento de realizar un diagnóstico visual. Este apartado resulta importante porque organiza el proceso de estudio y permite explicar de qué manera se obtuvo y analizó la información recolectada en campo.

La investigación considera como población de estudio a los estudiantes de la carrera de agropecuaria de la Uleam , Extensión El Carmen, y como fuente experta a un docente especializado en el área agrícola. La combinación de estos actores permitió obtener información tanto cuantitativa como cualitativa sobre el problema investigado. Por esta razón, el capítulo se estructura en varias secciones que incluyen la presentación,

la justificación metodológica, los tipos de investigación, las fuentes de información, las estrategias para la recolección de datos y el análisis de los resultados obtenidos.

3.1.2 Justificación Metodológica

La justificación metodológica de esta investigación se fundamenta en la necesidad de emplear un enfoque mixto para comprender de forma más completa la problemática estudiada. En este caso, no resulta suficiente medir únicamente la frecuencia de errores en la identificación de síntomas, sino que también es necesario analizar las razones técnicas y contextuales que explican dichas dificultades. Por ello, se combinan herramientas cuantitativas y cualitativas, ya que esta integración permite incrementar la confianza en los hallazgos y obtener una interpretación más sólida del fenómeno investigado.

Para la parte cuantitativa se aplicó una encuesta a los estudiantes de la carrera de agropecuaria de la Uleam, Extensión El Carmen, con el fin de recopilar información sobre la confusión entre enfermedades, la dependencia de la inspección visual y los errores frecuentes al momento del diagnóstico. En la parte cualitativa, se realizó una entrevista al Ing. Marco Vinicio de la Cruz, quien compartió su experiencia sobre la precisión del diagnóstico manual, las condiciones ambientales que influyen en la observación y la conveniencia de incorporar herramientas tecnológicas en el análisis fitosanitario.

3.2 Tipo de investigación

3.2.1 Métodos de investigación

Para el desarrollo de la presente investigación se emplearon los métodos inductivo-deductivo, análisis-síntesis y observación, los cuales permitieron estudiar el problema desde diferentes perspectivas y obtener una comprensión más completa del fenómeno analizado. La aplicación de estos métodos fue necesaria porque el estudio

requiere no solo recopilar información, sino también interpretarla, relacionarla y contrastarla con la realidad observada en el campo de estudio.

3.2.2 Método inductivo-deductivo

El método inductivo-deductivo es una estrategia de razonamiento científico que combina la inducción, que parte de casos particulares para llegar a conclusiones generales, con la deducción, que aplica principios generales para explicar situaciones específicas (Hernández Sampieri et al., 2018).

El método inductivo-deductivo permitió partir de casos particulares observados en los estudiantes de Agropecuaria y el criterio técnico del Ingeniero entrevistado, para luego establecer conclusiones generales sobre las dificultades en la identificación de enfermedades en hojas de cacao. A su vez, este método facilitó contrastar la información obtenida con la teoría revisada, lo cual ayudó a determinar la importancia de implementar una propuesta tecnológica basada en Deep Learning para mejorar el diagnóstico visual.

3.2.3 Método de análisis-síntesis

El método de análisis-síntesis consiste en descomponer un fenómeno en sus partes constitutivas para estudiarlas de forma individual y luego integrarlas en una comprensión global del objeto de estudio (Bernal, 2010, p. 60).

Por su parte, el método análisis-síntesis permitió descomponer la información recolectada en sus elementos principales, tales como la confusión entre enfermedades, la dependencia de la inspección visual y la influencia de las condiciones ambientales. Posteriormente, esta información fue integrada para obtener una visión más clara y organizada del problema investigado. Este método fue útil para estudiar cada parte del fenómeno y luego unir las en una comprensión general del estudio.

3.2.4 Método de observación

El método de observación es una técnica de investigación que consiste en percibir de manera directa y sistemática los fenómenos, hechos o situaciones del objeto de estudio, sin manipularlos, con el fin de obtener información real sobre el contexto analizado (Arias, 2012, p. 69).

El método de observación fue utilizado como apoyo para reconocer de manera directa las condiciones del entorno y los aspectos visibles relacionados con la detección de síntomas en las hojas de cacao. A través de este método fue posible complementar la información obtenida mediante la encuesta y la entrevista, ya que permitió identificar ciertos detalles del proceso de diagnóstico visual en campo. De esta manera, la observación fortaleció el análisis del problema investigado.

3.3 Fuentes de información de datos

3.3.1 Fuentes Primarias

Las fuentes primarias o directas constituyen el objetivo de la revisión de la literatura, debido a que proporcionan datos de primera mano. Entre ellas se encuentran libros, artículos de publicaciones periódicas, reportes de asociaciones, trabajos presentados en conferencias, testimonios de expertos y otros materiales originales que sirven como base para la investigación (Hernández Sampieri et al., 2014). En el presente trabajo, las fuentes primarias utilizadas fueron la encuesta aplicada a los estudiantes, y la entrevista realizada a un docente especializado en el área agrícola. Ambos instrumentos permitieron recolectar datos de primera mano sobre las dificultades en la identificación visual de enfermedades del cacao.

3.3.1.1 Encuesta

La encuesta es una técnica utilizada para recolectar datos, mediante la cual se obtiene información de un grupo de personas a través de preguntas previamente organizadas, con el propósito de conocer sus opiniones, conocimientos o comportamientos sobre un tema específico (Arias, 2012, p.67).

La encuesta fue uno de los instrumentos principales de la recolección de datos en esta investigación. Esta se aplicó a los estudiantes de la carrera de Agropecuaria de la Uleam, Extensión El Carmen, para obtener su opinión sobre cómo identificar enfermedades en hojas de cacao y determinar las dificultades con las que tienen que lidiar al identificarlas. Se recogió información cuantitativa sobre elementos como la confusión entre enfermedades, la dependencia de la observación visual y la probabilidad de errores en el diagnóstico.

3.3.1.2 Entrevista

La entrevista es una técnica de recolección de información basada en una conversación planificada entre el investigador y una persona que posee conocimientos o experiencias sobre el tema estudiado, con el propósito de profundizar en aspectos relevantes de la investigación (Hernández-Sampieri & Mendoza, 2018).

La entrevista fue el segundo instrumento utilizado en la investigación y estuvo dirigida al ingeniero Marco Vinicio de la Cruz. Este instrumento permitió obtener información cualitativa y más profunda sobre la problemática del diagnóstico visual de enfermedades en cacao. Mediante la entrevista se recogieron criterios técnicos, opiniones y experiencias profesionales relacionadas con la identificación de síntomas, la influencia de las condiciones ambientales y la necesidad de herramientas de apoyo para mejorar la precisión del análisis.

3.4 Estrategias Operacionales para la recolección de datos

Las estrategias de recolección de datos deben planificarse con precisión para garantizar que los datos obtenidos sean efectivos, fiables y concuerden con los objetivos de la investigación. Para ello, es importante definir el instrumento de recolección, organizar el trabajo de campo, capacitar a quienes participarán en la aplicación de los instrumentos y realizar una prueba piloto que permita corregir posibles errores antes de la ejecución definitiva (Hernández Sampieri et al., 2018). Además, la selección de las técnicas debe responder a la naturaleza del estudio, al tamaño de la población y al tipo de información que se desea obtener, ya que estos aspectos influyen directamente en la calidad de los resultados (Arias, 2012).

3.4.1 Población

La población de estudio estuvo integrada por los estudiantes de agropecuaria de la Uleam Extensión El Carmen, en un rango de 389 estudiantes, quienes tienen relación directa con la temática investigada. Esta población fue seleccionada porque sus integrantes poseen conocimientos básicos sobre el cultivo de cacao y pueden aportar información relevante acerca de las dificultades que presentan al momento de identificar enfermedades en las hojas. Su participación fue fundamental para conocer la situación real del problema desde el ámbito educativo y práctico.

Además, dentro del proceso investigativo se consideró la participación del Ing. Marco Vinicio de la Cruz como fuente experta, debido a su conocimiento técnico y experiencia en el área agrícola. Aunque no forma parte de la población encuestada, su criterio fue importante dentro del componente cualitativo de la investigación, ya que permitió completar la información obtenida de los estudiantes y aportar su visión profesional sobre la temática analizada.

La población es el conjunto total de elementos que comparten características comunes y que forman parte del objeto de estudio de una investigación (Tamayo y Tamayo, 2004, p. 176).

3.4.2 Muestra

Muestreo discrecional, también conocido como muestreo por juicio o muestreo intencional, es una forma de muestreo no probabilístico en el cual el investigador selecciona las muestras dependiendo de su juicio tomado en consideración, factores que debido a su experiencia y cercanía al fenómeno que se estudia pueden ofrecer información útil para los objetivos de la investigación (Hernández Sampieri et al., 2014). Fue sobre esta base que se utilizó el criterio para dar prioridad a los participantes que ya habían interactuado con el problema de la investigación en términos de la práctica de la producción del cacao.

La muestra de investigación estuvo conformada por 29 respuestas obtenidas de los estudiantes de Agropecuaria de la Universidad Laica Eloy Alfaro de Manabí Extensión El Carmen que participaron en la encuesta. Con esta muestra, se pudo adquirir información valiosa y relevante sobre las dificultades que los estudiantes experimentan al detectar enfermedades en hojas de cacao, sin tener que recurrir a todos los estudiantes.

En cuanto a la entrevista, la muestra cualitativa incluyó al Ing. Marco Vinicio de la Cruz, elegido por su experiencia y conocimiento del sector agrario. Su colaboración resultó fundamental para recopilar datos con un fuerte componente técnico y especializado contribuyendo a analizar la investigación. De esta manera la muestra utilizada en la investigación fue adecuada para cumplir con los objetivos planteados y para garantizar la obtención de datos relevantes sobre la problemática investigada.

La muestra es una parte de la población que se selecciona para obtener información cuando no resulta posible estudiar a todos sus integrantes (Sabino, 1992, p. 94).

3.4.3 Análisis de herramientas de recolección de datos a utilizar

3.4.3.1 Encuesta

La encuesta fue una de las principales herramientas utilizadas para la recolección de datos, debido a que permitió obtener información de los estudiantes de manera organizada y precisa. Este instrumento estuvo conformado por preguntas dirigidas a identificar el nivel de conocimiento que poseen sobre las enfermedades que afectan las hojas de cacao, así como las dificultades que enfrentan al momento de reconocer sus síntomas. Su aplicación fue relevante porque facilitó la obtención de datos concretos que sirvieron para el análisis estadístico de la investigación.

3.4.3.2 Entrevista

La entrevista se aplicó al Ing. Marco Vinicio de la Cruz como fuente experta, con el fin de recopilar criterios técnicos relacionados con la problemática investigada. Ese instrumento permitió profundizar en aspectos específicos sobre el diagnóstico de enfermedades en hojas de cacao, así como conocer su opinión profesional acerca de la importancia de implementar herramientas tecnológicas para mejorar este proceso. La información obtenida mediante la entrevista fue valiosa para complementar y enriquecer el análisis realizado con la encuesta.

3.4.3.3 Estructura de los instrumentos de recolección de datos aplicados

El instrumento utilizado para los estudiantes agropecuaria consistió en un cuestionario de 10 preguntas cerradas en escala de frecuencia (nunca, a veces siempre)

para detectar errores típicos del diagnóstico visual, por ejemplo, la confusión de los síntomas, la falta de coherencia entre los análisis y la influencia de los factores ambientales. Además, se realizó una entrevista con 10 preguntas abiertas con el Ing. Marco Vinicio de la Cruz sobre los mismos ejes para conocer en profundidad las causas de estos errores desde el punto de vista profesional.

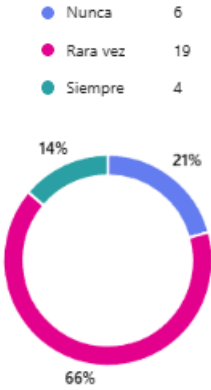
3.4.4 Plan de recolección de datos

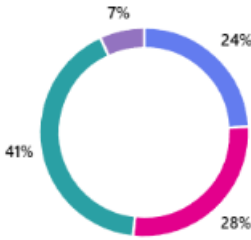
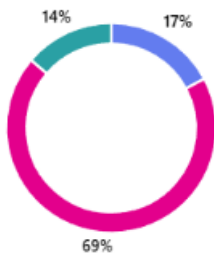
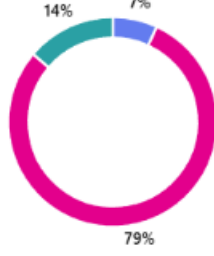
ACTIVIDAD	RESPONSABLE	FECHA
Encuesta a estudiantes	Pincay Gómez Lilibeth Carolina	03/06/2026
Entrevista a Ing. Marco Vinicio de la Cruz	Pincay Gómez Lilibeth Carolina	03/06/2026

3.5 Análisis y presentación de resultados

3.5.1 Presentación y descripción de los resultados obtenidos

3.5.1.1 Análisis de encuesta dirigida a los estudiantes de Agropecuaria

Pregunta	Resultados	Análisis
¿Confunde los síntomas de moniliasis con los de escoba de bruja en las hojas?	 ● Nunca 6 ● Rara vez 19 ● Siempre 4	La mayoría expresa que rara vez confunde los síntomas de moniliasis con escoba de bruja, aunque existe un grupo menor que sí presenta confusión. Esto evidencia que todavía hay dificultades en la identificación visual de enfermedades similares.

Pregunta	Resultados	Análisis
¿El mismo técnico da diagnósticos distintos en dos revisiones del mismo árbol?	- Nunca 3 - A veces 25 - Siempre 1 	Predomina la respuesta “a veces”, esto demuestra que en algunos casos el diagnóstico no es totalmente consistente. Esto puede deberse a la subjetividad de la observación visual y a la similitud entre síntomas.
¿Ha diagnosticado una enfermedad inexistente por similitud de manchas?	- Nunca 7 - Rara vez 8 - A veces 12 - Siempre 2 	La mayor frecuencia se concentra en “a veces”, lo que refleja que se presentan errores al confundir manchas parecidas en enfermedades reales. Esto puede ocasionar diagnósticos incorrectos y tratamientos innecesarios.
¿La revisión del cultivo se hace solo mirando a simple vista?	- Nunca 5 - A veces 20 - Siempre 4 	La mayoría consideran que la revisión se realiza sólo a simple vista en algunas ocasiones. Esto evidencia una dependencia de la observación visual, lo cual limita la precisión del diagnóstico.
¿Considera confiable el diagnóstico visual cuando se repite la revisión del mismo árbol en distintos momentos?	- Nunca 2 - A veces 23 - Siempre 4 	Los resultados muestran que en la mayoría de los casos el mismo técnico puede emitir diagnósticos distintos en revisiones diferentes. Esto confirma que el proceso de identificación no siempre es uniforme.

Pregunta	Resultados	Análisis
¿Se omiten plantas enfermas al inspeccionar visualmente una parcela grande?	● Nunca 8 ● A veces 16 ● Siempre 5	Predomina la respuesta “a veces”, lo que indica que en parcelas grandes puede haber plantas enfermas que no son detectadas durante la inspección visual. Esto representa una limitación importante en el monitoreo manual.
¿Ha tratado plantas sanas por error al confundirlas con enfermas?	● Nunca 8 ● A veces 19 ● Siempre 2	La mayoría afirma que esto ocurre “a veces”, lo que demuestra que existen errores al diferenciar plantas sanas de enfermas. Este problema puede ocasionar tratamientos innecesarios y afectar el manejo adecuado del cultivo.
¿La falta de herramientas como lupa, guías entre otras reduce la precisión por debajo del 70%?	● Nunca 2 ● No sé 13 ● A veces 9 ● Siempre 5	La mayoría de los estudiantes respondió “no sé”, lo que indica que no tienen claro el impacto real de no contar con herramientas de apoyo. Aun así, al sumar las respuestas positivas se observa que sí existe la percepción de que la precisión del diagnóstico disminuye sin estas herramientas.
¿La detección visual del cultivo falla más cuando hay alta humedad o poca luz?	● Nunca 2 ● A veces 22 ● Siempre 5	La mayoría indica que esto ocurre “a veces”, lo que demuestra que las condiciones ambientales influyen en la detección visual. La humedad y la poca luz dificultan observar correctamente los síntomas.

Pregunta	Resultados	Análisis
¿Ha logrado siempre la misma identificación en días soleados que en días nublados?	● Nunca 3 ● Rara vez 12 ● A veces 12 ● Siempre 2	Los resultados muestran que la identificación no siempre es igual en distintas condiciones climáticas. Esto confirma que la iluminación y el clima afectan la consistencia del diagnóstico visual.

3.5.1.2 Análisis de entrevista dirigida al Ing. Marco Vinicio de la Cruz

Pregunta	Resultados	Análisis
¿Qué enfermedades del cacao son las que los productores confunden más con mayor frecuencia al observar lo que son las hojas del cacao?	El entrevistado indicó que la enfermedad que más confunden los estudiantes y productores es la escoba de bruja, especialmente con la flor de pajarito. Explicó que es causada por un hongo que afecta la parte apical de la planta y ocasiona daños en el desarrollo vegetal y de la producción.	La respuesta muestra que la principal confusión en el cacao se presenta con la escoba de bruja, debido a la similitud de sus síntomas con otras alteraciones de la planta. Esto evidencia la dificultad de realizar un diagnóstico visual preciso.
¿Considera que la identificación visual de enfermedades foliares sin apoyo de herramientas como lupas, guías, etc. lleva normalmente errores de diagnósticos y en qué porcentaje estima esos errores?	El entrevistado señaló que la identificación visual sin apoyo de herramientas sí puede generar errores de diagnóstico. Indicó que es importante contar con las lupas, tomar fotografías y revisar el análisis junto con los productores para ir descartando si el problema corresponde a plagas, enfermedades o deficiencias nutricionales.	Esto demuestra que la observación visual por sí sola no es suficiente para lograr un diagnóstico confiable, y que el uso de herramientas de apoyo ayuda a reducir errores y a identificar mejor la causa real del problema.
¿En campos con alta humedad o baja luminosidad, la fiabilidad de la inspección visual humana disminuye	El entrevistado indicó que, en zonas con clima subtropical, alta humedad y menor radiación solar, la inspección visual se vuelve	Esta respuesta confirma que la fiabilidad de la inspección visual disminuye en ambientes con mucha humedad y poca luz. En esas

Pregunta	Resultados	Análisis
significativamente? ¿qué tipo de fallos son más comunes en esas condiciones?	más compleja y menos confiable. Señaló que en esas condiciones las plagas y enfermedades se diseminan con mayor rapidez.	condiciones son más frecuentes los errores al identificar enfermedades.
¿Ha observado que un mismo técnico o productor emita diagnósticos distintos sobre el mismo árbol en 2 revisiones separadas? ¿A qué lo atribuye?	El entrevistado manifestó que a veces un mismo técnico puede emitir diagnósticos distintos sobre el mismo árbol en revisiones separadas. Atribuyó esto al uso de parámetros diferentes al momento de evaluar y a la influencia de recomendaciones asociadas a casas comerciales.	La respuesta evidencia que el diagnóstico puede variar según el criterio del evaluador y la forma en que interpreta los síntomas. Esto refuerza la necesidad de criterios más estandarizados.
¿Cuándo se inspecciona una parcela grande de cacao visualmente, ¿es común omitir plantas enfermas? ¿Qué factores aumentan esa omisión?	El entrevistado señaló que en una parcela grande es posible omitir plantas enfermas durante la inspección visual. Indicó que esto ocurre especialmente cuando no se realiza un estudio minucioso del área afectada y cuando no se toma en cuenta la prevención alrededor de la planta infectada.	La respuesta muestra que en parcelas extensas la observación visual puede pasar por altos focos de infección, lo que favorece la diseminación de la enfermedad.
¿Qué tan frecuente es que un productor trate plantas sanas porque las confundió con enfermas? ¿qué consecuencias productivas ha visto por esto?	El entrevistado explicó que este error puede ocurrir cuando los síntomas visibles no corresponden a una enfermedad sino a factores abióticos o deficiencias nutricionales. Señaló que, si no se realiza un buen diagnóstico, se pueden recomendar productos innecesarios que afectan el bolsillo del productor.	La respuesta evidencia que la confusión entre plantas sanas y enfermas puede provocar tratamientos incorrectos y gastos adicionales. Esto demuestra la importancia de diferenciar con precisión la causa real del problema.
¿Considera que la falta de herramientas sencillas (lupa de bolsillo, cartilla de síntomas, aplicación móvil) reduce la precisión del diagnóstico por ejemplo del	El entrevistado señaló que la falta de herramientas sencillas si reduce la precisión del diagnóstico, y afirmó que al menos una lupa sería indispensable para observar mejor los	La respuesta confirma que los productores no capacitados dependen mucho de la observación directa y que, sin herramientas de apoyo,

Pregunta	Resultados	Análisis
70% en productores no capacitados?	síntomas e incluso de detectar insectos muy pequeños que causan daños severos.	pueden pasar por alto detalles importantes.
¿Qué diferencias observa en la precisión del diagnóstico entre un día soleado y uno nublado? ¿La variabilidad climática afecta más de lo que se cree?	En días soleados puede ser más difícil encontrar ciertas plagas porque muchas se esconden en la hoja rasca o en hábitats donde la radiación solar les afecta. En cambio, en días nublados considera que hay más ventajas para el análisis, porque se observan con mayor presencia plagas y enfermedades.	La respuesta muestra que la variabilidad climática sí influye en la precisión del diagnóstico visual. Los días nublados permiten observar mejor las plantas y, en consecuencia, facilitan un diagnóstico más preciso.
¿Cree que la dependencia exclusiva de inspección visual humana, sin apoyo tecnológico o de laboratorio, es una causa principal de propagación de enfermedades en cacao?	El entrevistado indicó que depender solo de la inspección visual puede retrasar la detección de enfermedades y favorecer su propagación. También señaló que lo más adecuado es contar con la revisión de un técnico en campo, ya que llevar únicamente muestras a una agropecuaria puede llevar a diagnósticos poco precisos al no observarse la planta completa ni el contexto real del cultivo.	Esto indica que depender únicamente de la inspección visual o de muestras aisladas puede llevar a errores de diagnóstico y a gastos innecesarios, lo cual resalta la importancia del acompañamiento técnico en campo para identificar correctamente las enfermedades.
¿Si tuviera que recomendar una solución práctica y de bajo costo para mejorar la precisión en la detección de síntomas foliares en campo, cuál sería?	El entrevistado recomendó mantener un monitoreo constante en el campo, recolectar y enterrar los frutos infectados, realizar podas sanitarias de ramas dañadas y desinfectar adecuadamente las herramientas de trabajo.	Esta respuesta sugiere que una solución práctica y de bajo costo es aplicar un protocolo básico de monitoreo y manejo sanitario en la parcela. Estas acciones ayudan a reducir la incidencia de enfermedades y evitar su propagación.

3.5.1.3 Informe final del análisis de datos

Según la pregunta 2 de la encuesta, la mayoría menciona que a veces un mismo técnico da diagnósticos diferentes sobre el mismo árbol, y según la pregunta 4 de la

entrevista, esto se debe a que cada evaluador maneja parámetros distintos al momento de revisar la planta. Esto señala que la falta de uniformidad en el diagnóstico visual es una de las causas que explican los errores de la identificación de enfermedades.

Según la pregunta 8 de la encuesta, la falta de herramientas de apoyo reduce la precisión del diagnóstico, y según la pregunta 7 de la entrevista, una lupa y el análisis en campo ayudan a diferenciar mejor entre plagas, enfermedades y deficiencias nutricionales. Esto señala que la carencia de herramientas básicas es otra causa que afecta la exactitud del diagnóstico de cacao.

CAPÍTULO IV

4 MARCO PROPOSITIVO

4.1 Introducción

En este capítulo se expone la propuesta de solución tecnológica desarrollada dentro del proyecto integrador, así como los elementos principales que forman parte de su implementación orientada a la detección del estado de salud vegetal de cacao (*Theobroma cacao*) mediante técnicas de inteligencia artificial. Se detalla la arquitectura del sistema, los recursos empleados durante el desarrollo y la metodología ágil Scrum aplicada mediante Sprints (Schwaber & Sutherland, 2020), que permitió organizar el trabajo de manera iterativa e incremental. La propuesta surge como respuesta a la necesidad de los estudiantes de la Granja Experimental de El Carmen de contar con una herramienta tecnológica accesible que permita identificar de forma temprana alteraciones en el estado de salud de sus cultivos, contribuyendo así a la toma de decisiones oportunas y a la reducción de pérdidas en la producción cacaotera.

4.2 Descripción de la propuesta

El sistema informático se desarrolló basado en Deep Learning, utilizando la arquitectura MobileNetV2 con Transfer Learning (Sandler et al., 2018; Pan & Yang, 2010), con el fin de clasificar el estado de salud vegetal de cacao a partir de imágenes fotográficas. El desarrollo se realizó en Python (Van Rossum & Drake, 2009), utilizando el framework TensorFlow/Keras (Abadi et al., 2016). El sistema cuenta además con una interfaz gráfica elaborada en Tkinter, que permite al usuario cargar una imagen de la planta y obtener un diagnóstico de forma instantánea. El modelo fue entrenado con un

dataset propio de 2.097 imágenes, distribuidas en cinco categorías: Estrés Abiótico, Fito, Monilia, No_Cacao y Sana.

Para mejorar el rendimiento del modelo, el entrenamiento fue dividido en dos etapas. Durante la primera fase, se mantuvieron fijadas las capas de base de MobileNetV2 y solo se optimizaron las nuevas capas adicionadas al final, como GlobalAveragePooling2D, Dense y Dropout. Esta acción permitió adaptar rápidamente a los nuevos pesos sin alterar la red ya entrenada en ImageNet. Posteriormente, en el segundo paso, se utilizó el método de fine-tuning al descongelar las últimas 30 capas del modelo e iniciar un nuevo entrenamiento con una mayor tasa de aprendizaje. Con esto se buscó que la red adaptara mejor a las características particulares de las hojas y mazorcas de cacao, logrando una mejor generalización con el dataset utilizado.

El sistema genera como salida el diagnóstico de la clase detectada junto con el porcentaje de confianza y una recomendación de acción para el estudiante. Adicionalmente, se exportó el modelo a un formato TensorFlow Lite para su integración en la aplicación móvil que fue desarrollada por un integrante del grupo de trabajo.

4.3 Determinación de recursos

4.3.1 Humanos

Recursos Humanos	Función
Investigadora	Encargada de planificar el trabajo, recolectar los datos, entrenar el modelo, desarrollar la interfaz y realizar la validación del sistema.
Tutor de tesis	Acompaña el proceso de investigación y brinda apoyo en la revisión metodológica y en la redacción del documento.
Grupo de trabajo	Apoya en la recolección de imágenes en la Granja Experimental El Carmen y en la integración del módulo dentro de la aplicación móvil.

Tabla 1 - Recursos humanos del proyecto CacaoHealth AI

El desarrollo del sistema estuvo a cargo de la autora, estudiante de la carrera de Ingeniería en Software de la Uleam, Extensión El Carmen, bajo la supervisión del Ingeniero Raúl Saaed Reascos Pinchao. Adicionalmente, se contó con el apoyo del equipo de trabajo conformado por los demás integrantes del grupo de tesis, quienes colaboraron en la recolección de imágenes en la Granja Experimental El Carmen y en la integración de mi módulo dentro de la aplicación móvil general del proyecto.

4.3.2 Tecnológicos

Hardware	Software
Computadora portátil (Laptop)	Python 3.10
Smartphone para captura de imágenes	TensorFlow / Keras
Conexión Internet (Wifi doméstico y universitario)	Visual Studio Code
	MobileNetV2 (ImageNet)
	TensorFlow Lite

Tabla 2 - Recursos Tecnológicos del proyecto CacaoHealth AI

Para el desarrollo del sistema se utilizaron los siguientes recursos tecnológicos:

- Computadora portátil con procesador Ryzen 7, 16 GB de RAM y sistema operativo Windows 11
- Python 3.10 como lenguaje de programación principal
- TensorFlow 2.x / Keras para el entrenamiento del modelo de Deep Learning
- MobileNetV2 como arquitectura base con pesos preentrenados de ImageNet
- Visual Studio Code como entorno de desarrollo integrado
- Conexión a internet mediante red Wifi doméstica y red universitaria
- Smartphone para la captura de imágenes del dataset en campo

4.3.3 Económicos

Recursos	Características	Costos	Descripción
Computadora Portátil	Lenovo Legion 5	\$ 850.00	Equipo personal utilizado para el desarrollo, entrenamiento y pruebas del sistema.
Celular	Tecno Spark 30C	\$ 130.00	Dispositivo personal utilizado para pruebas del sistema.
Internet	Net Home	\$ 80.00	Uso de red Wifi doméstica y universitaria para investigación, consultas técnicas y desarrollo del sistema.
Software Gratuito	Python, TensorFlow, Visual Studio Code, MobileNetV2, entre otros	\$ 0.00	Todas las herramientas de desarrollo utilizadas son de código abierto y sin costo económico.
Transporte	Moto - Yamaha Crypton	\$ 20.00	Medio de transporte utilizado para trasladarse a la Granja Experimental El Carmen para recolectar imágenes útiles para el dataset.
Horas de Programación	288 horas por \$10 la hora	\$ 2.880.00	Horas invertidas en la planificación, desarrollo, pruebas y ajustes del sistema.
Total	Recursos invertidos	\$ 3.960.00	Monto total de los recursos económico invertidos en el desarrollo del sistema CacaoHealth AI

Tabla 3 - Recursos Económicos del proyecto CacaoHealth AI

El desarrollo del proyecto no implicó inversión económica elevada, ya que se trabajó principalmente con recursos personales y herramientas de software gratuitas. Los gastos considerados fueron mínimos y se reflejan en la tabla de costos, donde se evidencia que los recursos utilizados corresponden al equipo de trabajo, el servicio de internet y el transporte requerido para la recolección de imágenes. En general, el proyecto pudo desarrollarse con un presupuesto accesible, lo que demuestra que es posible implementar soluciones de inteligencia artificial sin necesidad de una inversión alta.

4.4 Desarrollo del sistema de trabajo bajo la metodología

Para desarrollar el sistema se trabajó la metodología ágil Scrum, organizando las actividades de nueve sprints de dos semanas cada uno. Esto permitió llevar un orden en el proceso, revisar lo avanzado en cada etapa y hacer ajustes necesarios antes de continuar con la siguiente fase.

Cabe aclarar que durante el desarrollo, el dataset pasó por versiones intermedias antes de llegar a su versión final: la versión 2 (dataset_v2) contó con cuatro clases y un volumen menor de imágenes; la versión 3 (dataset_v3) incorporó más fotografías y añadió la clase No_Cacao, alcanzando 5 clases; y la versión 4 (dataset_v4) utilizada para el entrenamiento del modelo final, amplió aún más el número de imágenes por clase, logrando una precisión del 81.09%.

4.4.1 Descripción del Producto

El módulo de inteligencia artificial que se desarrolló para la evaluación del estado de salud vegetal de cacao como parte del sistema integrador de monitoreo fitosanitario digital del grupo. El producto permite al usuario cargar una fotografía de una hoja o planta de cacao y recibir un diagnóstico automático, indicando si la planta presenta Estrés abiótico, síntomas de Fito, Monilia o si se encuentra Sana. El sistema incorpora un mecanismo de validación que rechaza imágenes que no corresponden a plantas de cacao, mostrando un mensaje de imagen no válida cuando la confianza del modelo es inferior al 40%. El modelo fue entrenado con fotografías reales tomadas en la Granja Experimental El Carmen y exportado a formato TensorFlow Lite para su funcionamiento en dispositivos móviles Android sin necesidad de conexión a internet.

4.4.2 Historias del Usuario

ID	Historia de usuario	Criterio de aceptación.
HU-01	Como usuaria, quiero cargar una fotografía de mi planta de cacao para obtener un diagnóstico automático del estado de salud.	El sistema muestra el diagnóstico en menos de 200 ms con el porcentaje de confianza.
HU-02	Como usuaria, quiero recibir el nombre de la clase detectada y el porcentaje de confianza del diagnóstico.	El sistema muestra la clase predicha y el porcentaje de confianza en la interfaz.
HU-03	Como usuaria, quiero que el sistema rechace imágenes que no corresponden a plantas de cacao.	El sistema muestra 'Imagen no válida' cuando la confianza es menor al 40% o la imagen es No Cacao.
HU-04	Como usuaria, quiero recibir un aviso cuando el sistema no está seguro del diagnóstico.	El sistema muestra 'verificar manualmente' cuando la confianza está entre 40% y 70%.
HU-05	Como usuaria, quiero recibir una recomendación específica junto al diagnóstico para saber qué hacer con mi cultivo.	El sistema muestra una recomendación diferente y específica para cada clase detectada.
HU-06	Como usuaria, quiero visualizar la imagen que cargue en la interfaz para confirmar que es la correcta.	La imagen se muestra en la interfaz manteniendo su proporción original.
HU-07	Cómo desarrolladora, necesito entrenar el modelo con un dataset propio de imágenes de cacao etiquetadas por clase.	El modelo es entrenado con 2097 imágenes organizadas en 5 clases, con un criterio de aceptación general $\geq 85\%$, conforme a lo establecido e la planificación de evaluación del sistema. (tabla 16, capítulo V)
HU-08	Cómo desarrolladora, necesito exportar el modelo entrenado a formato TFLite para su uso de la aplicación móvil.	El archivo modelo_salud_vegetal.tflite es generado correctamente y funciona en Android.
HU-09	Cómo desarrolladora, necesito ejecutar pruebas de validación para medir la precisión del modelo.	Se ejecutan 30 pruebas de calidad y 30 de tiempo de respuesta, cuyos resultados se guardan en archivos CSV.
HU-10	Cómo desarrolladora, necesito integrar mi módulo con la aplicación móvil desarrollada por el grupo de trabajo.	El archivo TFLite y el archivo clases.txt son entregados a responsables de la interfaz móvil junto con la documentación de integración.

Tabla 4 - Historias de usuario del sistema CacaoHealth AI

4.4.3 Product Backlog

Nº	HU	Tarea	Prioridad	Esfuerzo (días)
1	HU-07	Recolectar imágenes de la granja experimental El Carmen.	Alta	10
2	HU-07	Organizar y etiquetar el dataset por clases.	Alta	5
3	HU-07	Aplicar data augmentation para implementar el dataset.	Alta	5
4	HU-07	Entrenar modelo_v2 con MobileNetV2 y Transfer Learning	Alta	8
5	HU-07	Ampliar el dataset_v3, incorporar clase No_Cacao y reentrenar modelo v4.	Alta	8
6	HU-01, HU-02, HU-06	Desarrollar la interfaz gráfica en Tkinter.	Alta	6
7	HU-03, HU-04	Implementar validación de confianza e imágenes no válidas.	Alta	5
8	HU-5	Agregar mensaje de diagnóstico y recomendaciones por clase.	Alta	3
9	HU-09	Ejecutar pruebas de tiempo de respuesta y calidad.	Alta	4
10	HU-08, HU-10	Exportar modelo a TFLite y entregar a equipo de integración.	Alta	3

Tabla 5 - Product Backlog del sistema CacaoHealth AI

4.4.4 Planificación de Sprints

4.4.4.1 Sprint 1

Campo	Detalle
Sprint	1
Duración	2 semanas (10 días hábiles)
Semana	1-2
Objetivo	Planificar el proyecto, definir arquitectura tecnológica y configurar el entorno de desarrollo.
HU	HU-07
Entregable	Entorno de desarrollo configurado y arquitectura definida.

Tabla 6 - Planificación y configuración del entorno

Tareas del Sprint

- Definición del alcance del módulo.
- Instalación de Python, TensorFlow, Keras y Visual Studio Code.

- Selección de MobileNetV2 con arquitectura base.
- Definición de las 4 clases iniciales del modelo.

```

Downloading wrapt-2.1.2-cp313-cp313-win_amd64.whl (60 kB)
Downloading namex-0.1.0-py3-none-any.whl (5.9 kB)
Downloading optree-0.19.0-cp313-cp313-win_amd64.whl (337 kB)
Downloading rich-15.0.0-py3-none-any.whl (310 kB)
Downloading pygments-2.20.0-py3-none-any.whl (1.2 MB)
 1.2/1.2 MB 9.3 MB/s eta 0:00:00
Downloading markdown_it_py-4.0.0-py3-none-any.whl (87 kB)
Downloading mdurl-0.1.2-py3-none-any.whl (10.0 kB)
Downloading setuptools-82.0.1-py3-none-any.whl (1.0 MB)
 1.0/1.0 MB 20.3 MB/s eta 0:00:00

```

Figura 1- Instalación de Python en VS Code

```

PS D:\Usuario Lilibeth\Desktop\tesis_cacao> py -3.11 -m pip install tensorflow numpy matplotlib opencv
-python scikit-learn pillow
Using cached markdown_it_py-4.0.0-py3-none-any.whl (87 kB)
Using cached pygments-2.20.0-py3-none-any.whl (1.2 MB)
Using cached mdurl-0.1.2-py3-none-any.whl (10.0 kB)
Installing collected packages: namex, libclang, flatbuffers, wrapt, urllib3, typing_extensions, thread
poolctl, termcolor, six, pyparsing, pygments, protobuf, pillow, packaging, opt_einsum, numpy, mdurl, k
iwisolver, joblib, idna, gast, fonttools, cycler, charset_normalizer, certifi, absl-py, wheel, scipy,
requests, python-dateutil, optree, opencv-python, ml_dtypes, markdown-it-py, h5py, grpcio, google_pasta,
contourpy, scikit-learn, rich, matplotlib, astunparse, keras, tensorflow
Successfully installed absl-py-2.4.0 astunparse-1.6.3 certifi-2026.4.22 charset_normalizer-3.4.7 conto
urpy-1.3.3 cycler-0.12.1 flatbuffers-25.12.19 fonttools-4.62.1 gast-0.7.0 google_pasta-0.2.0 grpcio-1.
80.0 h5py-3.14.0 idna-3.13 joblib-1.5.3 keras-3.14.0 kiwisolver-1.5.0 libclang-18.1.1 markdown-it-py-4
.0.0 matplotlib-3.10.9 mdurl-0.1.2 ml_dtypes-0.5.4 namex-0.1.0 numpy-2.4.4 opencv-python-4.13.0.92 opt
_einsum-3.4.0 optree-0.19.0 packaging-26.2 pillow-12.2.0 protobuf-7.34.1 pygments-2.20.0 pyparsing-3.3
.2 python-dateutil-2.9.0.post0 requests-2.33.1 rich-15.0.0 scikit-learn-1.8.0 scipy-1.17.1 six-1.17.0
tensorflow-2.21.0 termcolor-3.3.0 threadpoolctl-3.6.0 typing_extensions-4.15.0 urllib3-2.6.3 wheel-0.4
7.0 wrapt-2.1.2

[notice] A new release of pip is available: 24.0 -> 26.1
[notice] To update, run: python.exe -m pip install --upgrade pip
PS D:\Usuario Lilibeth\Desktop\tesis_cacao>

```

Figura 2 - Instalación TensorFlow

The image shows a VS Code editor with two files open: `modelo_cacao_v4.py` and `pruebas_validacion.py`. The `modelo_cacao_v4.py` file contains the following Python code:

```

1  import os
2  import numpy as np
3  import matplotlib.pyplot as plt
4  from sklearn.utils.class_weight import compute_class_weight
5  from tensorflow.keras.applications import MobileNetV2
6  from tensorflow.keras.models import Model
7  from tensorflow.keras.layers import Dense, GlobalAveragePooling2D, Dropout, BatchNormalization
8  from tensorflow.keras.preprocessing.image import ImageDataGenerator
9  from tensorflow.keras.optimizers import Adam
10 from tensorflow.keras.callbacks import ModelCheckpoint, EarlyStopping, ReduceLROnPlateau
11 import warnings
12 warnings.filterwarnings('ignore')
13

```

The terminal window shows the output of the command `pip show tensorflow keras`:

```

PS D:\Usuario Lilibeth\Desktop\tesis_cacao> pip show tensorflow keras
---
Name: keras
Version: 3.14.0
Summary: Multi-backend Keras
Home-page:
Author:
Author-email: Keras team <keras-users@googlegroups.com>
License: Apache License 2.0
Location: C:\Users\user\AppData\Local\Programs\Python\Python313\Lib\site-packages
Requires: absl-py, h5py, ml-dtypes, namex, numpy, optree, packaging, rich
Required-by: tensorflow
PS D:\Usuario Lilibeth\Desktop\tesis_cacao>

```

Figura 3 – VS Code con las dependencias instaladas

4.4.4.2 Sprint 2

Campo	Detalle
Sprint	2
Duración	2 semanas (10 días hábiles)
Semana	3 - 4
Objetivo	Recolectar imágenes en la granja experimental El Carmen para construir el dataset.
HU	HU-07
Entregable	Dataset inicial con imágenes reales de cacao organizadas por clase.

Tabla 7 - Recolección de imágenes en campo

Tareas del Sprint

- Visita a la granja experimental de Carmen
- Captura de fotografías de hojas y plantas con diferentes estados de salud
- Organización de imágenes en carpetas por clase
- Revisión y filtrado de imágenes de baja calidad



Figura 4 - Visita a la Granja para Captura de Fotografías del cacao

```

● PS D:\Usuario Lilibeth\Desktop\tesis_cacao> python organizar_dataset.py
Fito: 85 train | 22 validación ✓
Monilia: 84 train | 21 validación ✓
Sana: 80 train | 20 validación ✓

¡Dataset organizado correctamente!
○ PS D:\Usuario Lilibeth\Desktop\tesis_cacao> 

```

Figura 5 - Primera organización de imágenes en carpetas

tesis_cacao > dataset_v2 > train					Buscar en train
Ordenar					Ver
Nombre	Fecha de modificación	Tipo	Tamaño		
Estres_Abiotico	1/5/2026 18:01	Carpeta de archivos			
Fito	1/5/2026 17:26	Carpeta de archivos			
Monilia	1/5/2026 17:26	Carpeta de archivos			
Sana	1/5/2026 17:26	Carpeta de archivos			

Figura 6 - Imágenes organizadas en carpetas

4.4.4.3 Sprint 3

Camp	Detalle
Sprint	3
Duración	2 semanas (10 días hábiles)
Semana	5 - 6
Objetivo	Preparar y ampliar el dataset aplicando las técnicas de data augmentation.
HU	HU-07
Entregable	Dataset v2 ampliado mediante técnicas de data augmentation, con las 4 clases iniciales listas para entrenamiento.

Tabla 8 - Preparación y ampliación del dataset

Tareas del Sprint

- Aplicación de técnicas de data augmentation.
- Incorporación de rotaciones, zoom, brillo y volteo de imágenes.
- Creación de carpetas train y validation con proporción 80/20.
- Verificación del balance del dataset por clase.

```
PS D:\Usuario Lilibeth\Desktop\tesis_cacao> py -3.11 aumentar_estres.py

Total generadas: 96 imagenes de Estres_Abiotico
Validacion: 20 imagenes
Listo!
PS D:\Usuario Lilibeth\Desktop\tesis_cacao>
```

Figura 7 - aplicación de data augmentation

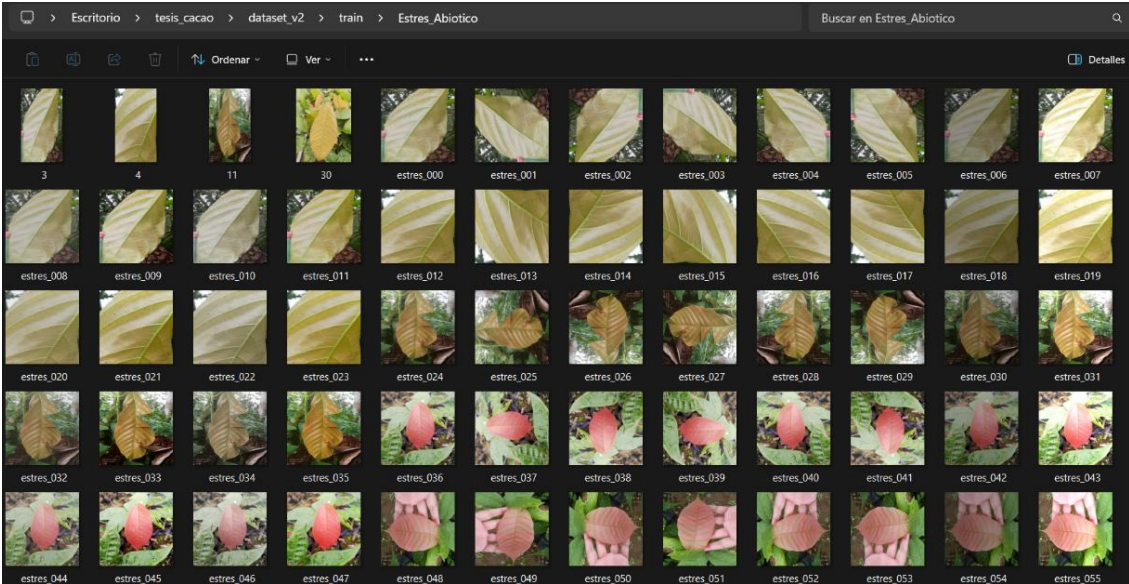


Figura 8 – Incorporación de rotaciones, zoom, brillo y volteo de imágenes

Nombre	Fecha de modificación	Tipo	Tamaño
test	1/5/2026 17:53	Carpeta de archivos	
train	1/5/2026 17:53	Carpeta de archivos	
valid	1/5/2026 17:53	Carpeta de archivos	
README.dataset	1/5/2026 17:53	Documento de tex...	1 KB
README.roboflow	1/5/2026 17:53	Documento de tex...	1 KB

Figura 9 - Creación de carpetas train y validation

```

Monilia
Fotos originales : 84
Multiplicador    : x5
Fotos a generar  : ~420
Guardando en     : dataset_v3/train\Monilia
✅ Total en carpeta destino: 504 fotos

Sana
Fotos originales : 89
Multiplicador    : x3
Fotos a generar  : ~267
Guardando en     : dataset_v3/train\Sana
✅ Total en carpeta destino: 356 fotos

Fito
Fotos originales : 99
Multiplicador    : x3
Fotos a generar  : ~297
Guardando en     : dataset_v3/train\Fito
✅ Total en carpeta destino: 396 fotos

Estres_Abiotico
Fotos originales : 100
Multiplicador    : x2
Fotos a generar  : ~200
Guardando en     : dataset_v3/train\Estres_Abiotico
✅ Total en carpeta destino: 300 fotos

```

Figura 10 - • Verificación del balance del dataset por clase

4.4.4.4 Sprint 4

Campo	Detalle
Sprint	4
Duración	2 semanas (10 días hábiles)
Semana	7 - 8
Objetivo	Entrenar el modelo de clasificación con MobileNetV2 y Transfer Learning.
HU	HU-07
Entregable	Modelo v2 entrenado con 4 clases y precisión inicial del 74%.

Tabla 9 - Entrenamiento del modelo de clasificación

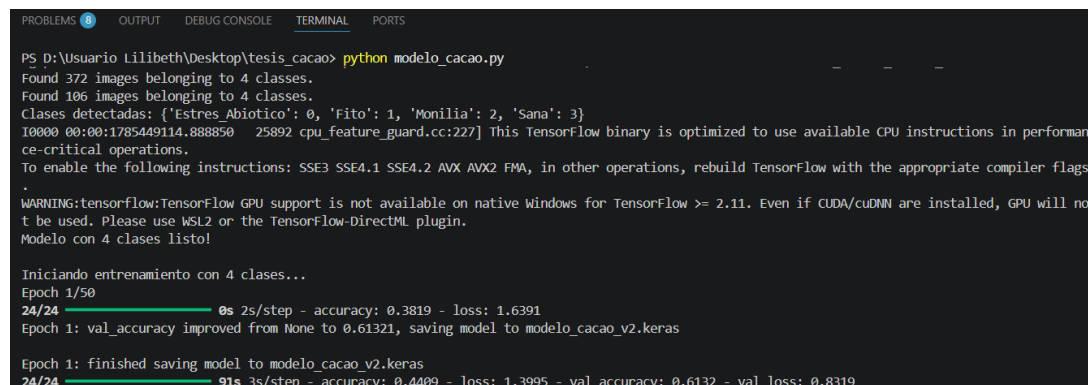
Tareas del Sprint

- Implementación del modelo con MobileNetV2 preentrenado.

- Configuración de capas densas, Dropout y BatchNormalization.
- Entrenamiento con Callbacks EarlyStopping y ReduceLROnPlateau.
- Evaluación de resultados y generación de gráficas de precisión.

```
x = base_model.output
x = GlobalAveragePooling2D()(x)
x = BatchNormalization()(x)
x = Dropout(0.3)(x)
x = Dense(256, activation='relu', kernel_regularizer=l2(0.001))(x)
x = BatchNormalization()(x)
x = Dropout(0.3)(x)
x = Dense(128, activation='relu', kernel_regularizer=l2(0.001))(x)
x = Dropout(0.2)(x)
output = Dense(NUM_CLASSES, activation='softmax')(x)
```

Figura 11 - Capas densas, Dropout y BatchNormalization



```
PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS
PS D:\Usuario Lilibeth\Desktop\tesis_cacao> python modelo_cacao.py
Found 372 images belonging to 4 classes.
Found 106 images belonging to 4 classes.
Clases detectadas: {'Estres_Abiotico': 0, 'Fito': 1, 'Monilia': 2, 'Sana': 3}
I0000 00:00:1785449114.888850 25892 cpu_feature_guard.cc:227] This TensorFlow binary is optimized to use available CPU instructions in performan
ce-critical operations.
To enable the following instructions: SSE3 SSE4.1 SSE4.2 AVX AVX2 FMA, in other operations, rebuild TensorFlow with the appropriate compiler flags.
WARNING:tensorflow:TensorFlow GPU support is not available on native Windows for TensorFlow >= 2.11. Even if CUDA/cuDNN are installed, GPU will no
t be used. Please use WSL2 or the TensorFlow-DirectML plugin.
Modelo con 4 clases listo!

Iniciando entrenamiento con 4 clases...
Epoch 1/50
24/24 ----- 0s 2s/step - accuracy: 0.3819 - loss: 1.6391
Epoch 1: val_accuracy improved from None to 0.61321, saving model to modelo_cacao_v2.keras
Epoch 1: finished saving model to modelo_cacao_v2.keras
24/24 ----- 91s 3s/step - accuracy: 0.4409 - loss: 1.3995 - val accuracy: 0.6132 - val loss: 0.8319
```

Figura 12 - Terminal mostrando el entrenamiento del modelo con 4 clases (dataset_v2)

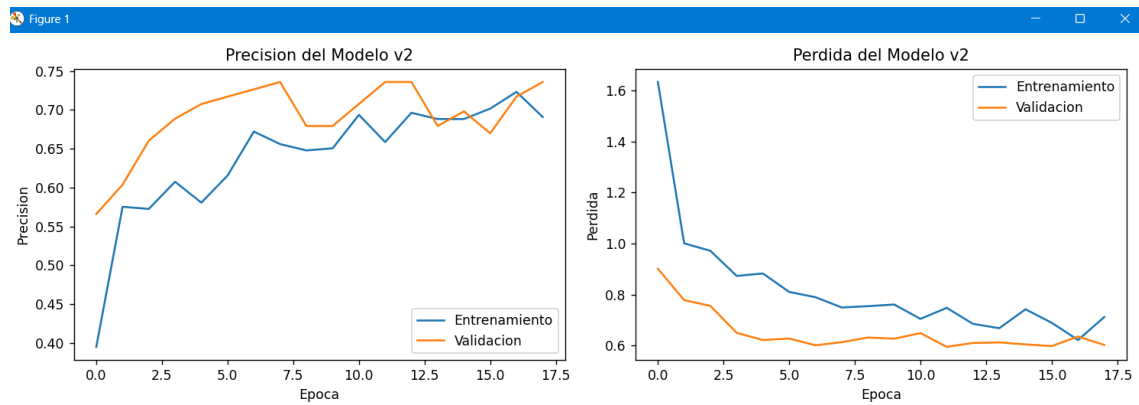


Figura 13 - Evaluación de resultados y generación de gráficas de precisión (modelos anteriores)

4.4.4.5 Sprint 5

Campo	Detalle
Sprint	5
Duración	2 semanas (10 días hábiles)
Semana	9 - 10
Objetivo	Optimizar el modelo incorporado, la clase no cacao y class weights para mejorar la precisión.
HU	HU-03, HU-07
Entregable	Dataset v3/v4 ampliado con 2.097 imágenes en 5 clases, modelo v4 entrenado con una precisión del 81.09% tras ajuste fino.

Tabla 10 - Optimización del modelo con clase No_Cacao

Tareas del Sprint

- Incorporación de clase No_Cacao al dataset.
- Aplicación de class weights para balancear clases.
- Reentrenamiento del modelo dataset_V4.
- Evaluación de mejoras y comparación con versiones anteriores.

```

No_Cacao
Fotos originales : 28
Multiplicador      : x8
Fotos a generar    : ~224
✅ Total en destino: 252 fotos

```

Figura 14 - Incorporación de clase No_Cacao al dataset

```

# -----
# CLASS WEIGHTS
# -----
etiquetas = train_gen.classes
pesos = compute_class_weight(
    class_weight='balanced',
    classes=np.unique(etiquetas),
    y=etiquetas
)
class_weights = dict(enumerate(pesos))
print(f"\nPesos por clase: {class_weights}")

```

Figura 15 - Aplicación de class weights para balancear clases

```

PS D:\Usuario_Lilibeth\Desktop\tesis_cacao> python modelo_cacao_v4.py
113/113 34s 300ms/step - accuracy: 0.7406 - loss: 0.5473 - val_accuracy: 0.7080 - val_loss: 0.6666 - learning_rate: 2.5000e-04
Epoch 19/80
113/113 0s 254ms/step - accuracy: 0.7155 - loss: 0.5853
Epoch 19: val_accuracy did not improve from 0.73451
113/113 34s 303ms/step - accuracy: 0.7306 - loss: 0.5560 - val_accuracy: 0.6903 - val_loss: 0.6610 - learning_rate: 2.5000e-04
Epoch 20/80
113/113 0s 251ms/step - accuracy: 0.7405 - loss: 0.5484
Epoch 20: val_accuracy did not improve from 0.73451
113/113 34s 301ms/step - accuracy: 0.7306 - loss: 0.5521 - val_accuracy: 0.6903 - val_loss: 0.6566 - learning_rate: 2.5000e-04
Epoch 21/80
113/113 0s 252ms/step - accuracy: 0.7616 - loss: 0.5090
Epoch 21: val_accuracy did not improve from 0.73451
113/113 35s 305ms/step - accuracy: 0.7539 - loss: 0.5113 - val_accuracy: 0.6637 - val_loss: 0.6675 - learning_rate: 2.5000e-04
Epoch 22/80
113/113 0s 255ms/step - accuracy: 0.7393 - loss: 0.5253
Epoch 22: val_accuracy did not improve from 0.73451

Epoch 22: ReduceLROnPlateau reducing learning rate to 0.0001250000059371814.
113/113 34s 305ms/step - accuracy: 0.7528 - loss: 0.5100 - val_accuracy: 0.6814 - val_loss: 0.6697 - learning_rate: 2.5000e-04
Epoch 23/80
113/113 0s 252ms/step - accuracy: 0.7485 - loss: 0.5105
Epoch 23: val_accuracy did not improve from 0.73451
113/113 34s 301ms/step - accuracy: 0.7428 - loss: 0.5146 - val_accuracy: 0.6991 - val_loss: 0.6718 - learning_rate: 1.2500e-04
Epoch 24/80
113/113 0s 253ms/step - accuracy: 0.7619 - loss: 0.4836
Epoch 24: val_accuracy did not improve from 0.73451
113/113 34s 302ms/step - accuracy: 0.7489 - loss: 0.5049 - val_accuracy: 0.6903 - val_loss: 0.6742 - learning_rate: 1.2500e-04
Epoch 25/80
113/113 0s 252ms/step - accuracy: 0.7391 - loss: 0.5356
Epoch 25: val_accuracy did not improve from 0.73451
113/113 34s 301ms/step - accuracy: 0.7434 - loss: 0.5307 - val_accuracy: 0.6991 - val_loss: 0.6694 - learning_rate: 1.2500e-04
Epoch 25: early stopping
Restoring model weights from the end of the best epoch: 10.

```

Figura 16 - Rentrenamiento del modelo dataset V4

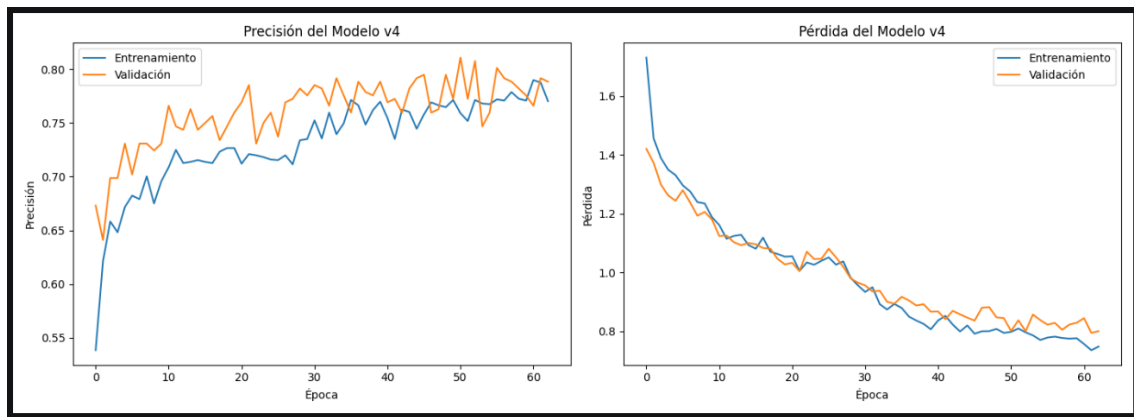


Figura 17 - Evaluación de mejoras y comparación conversiones anteriores

4.4.4.6 Sprint 6

Campo	Detalle
Sprint	6
Duración	2 semanas (10 días hábiles)
Semana	11 – 12
Objetivo	Desarrollar la interfaz gráfica del sistema con visualización de imagen y diagnóstico.
HU	HU-01, HU-02, HU-05, HU-06
Entregable	Interfaz gráfica funcional desarrollada en Tkinter.

Tabla 11 - Desarrollo de la interfaz gráfica

Tareas del Sprint

- Diseño y desarrollo de la interfaz gráfica en Tkinter.
- Implementación del botón de carga de imagen.
- Visualización de imagen con proporción original.
- Mostrar diagnóstico, confianza y recomendación por clase.

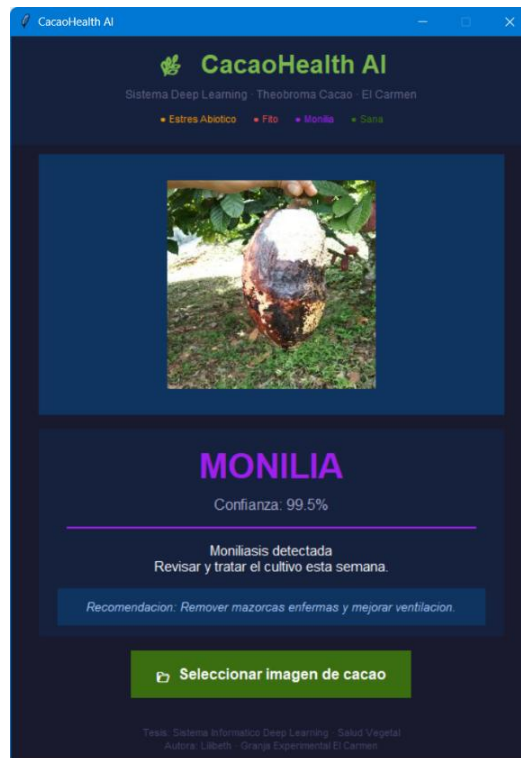


Figura 18 - Funcionamiento de la interfaz en Tkinter

4.4.4.7 Sprint 7

Campo	Detalle
Sprint	7
Duración	2 semanas (10 días hábiles)
Semana	13 - 14
Objetivo	Ejecutar prueba de validación del modelo y la interfaz para medir precisión y tiempo de respuesta.
HU	HU-09
Entregable	Resultados de pruebas guardados en archivos csv con 81.09% de precisión general sobre conjunto completo de validación.

Tabla 12 - Pruebas de validación

Tareas del Sprint

- Ejecución de 30 pruebas de tiempo de respuesta.
- Ejecución de 30 pruebas de calidad de respuesta.
- Análisis de resultados por clase.
- Guardado de resultados en archivos prueba_tiempo_v4.csv y prueba_calidad_v4.csv.

```
=====
PRUEBA 1: TIEMPO DE RESPUESTA – MODELO V4
=====
Prueba 01: 2859.21 ms
Prueba 02: 256.61 ms
Prueba 03: 322.97 ms
Prueba 04: 262.72 ms
Prueba 05: 277.9 ms
Prueba 06: 251.69 ms
Prueba 07: 230.85 ms
Prueba 08: 512.08 ms
Prueba 09: 230.84 ms
Prueba 10: 181.73 ms
Prueba 11: 175.27 ms
Prueba 12: 255.5 ms
Prueba 13: 239.32 ms
Prueba 14: 230.51 ms
Prueba 15: 239.64 ms
Prueba 16: 221.04 ms
Prueba 17: 240.23 ms
Prueba 18: 237.85 ms
Prueba 19: 239.44 ms
Prueba 20: 477.42 ms
Prueba 21: 233.09 ms
Prueba 22: 254.93 ms
Prueba 23: 468.98 ms
Prueba 24: 198.37 ms
Prueba 25: 269.55 ms
Prueba 26: 271.49 ms
Prueba 27: 474.32 ms
Prueba 28: 257.16 ms
Prueba 29: 230.97 ms
Prueba 30: 238.29 ms

Promedio : 361.33 ms
Minimo   : 175.27 ms
Maximo   : 2859.21 ms
Guardado en prueba_tiempo_v4.csv OK
```

Figura 19 - Tiempo de respuesta

PRUEBA 2: CALIDAD DE RESPUESTA – MODELO V4				
Prueba 01:	Real=Fito	Pred=Fito	87.45%	Correcto
Prueba 02:	Real=Monilia	Pred=Monilia	67.28%	Correcto
Prueba 03:	Real=Fito	Pred=Fito	84.01%	Correcto
Prueba 04:	Real=Estres_Abiotico	Pred=Estres_Abiotico	99.99%	Correcto
Prueba 05:	Real=No_Cacao	Pred=No_Cacao	100.0%	Correcto
Prueba 06:	Real=Estres_Abiotico	Pred=Estres_Abiotico	91.27%	Correcto
Prueba 07:	Real=Fito	Pred=Fito	94.6%	Correcto
Prueba 08:	Real=Sana	Pred=Sana	36.17%	Correcto
Prueba 09:	Real=Estres_Abiotico	Pred=Estres_Abiotico	99.98%	Correcto
Prueba 10:	Real=Fito	Pred=Fito	74.66%	Correcto
Prueba 11:	Real=Fito	Pred=Fito	61.66%	Correcto
Prueba 12:	Real=Sana	Pred=Sana	76.66%	Correcto
Prueba 13:	Real=Sana	Pred=Monilia	55.09%	Error
Prueba 14:	Real=Sana	Pred=Sana	50.87%	Correcto
Prueba 15:	Real=Monilia	Pred=Monilia	96.36%	Correcto
Prueba 16:	Real=Monilia	Pred=Monilia	95.78%	Correcto
Prueba 17:	Real=Monilia	Pred=Monilia	48.42%	Correcto
Prueba 18:	Real=Fito	Pred=Fito	99.98%	Correcto
Prueba 19:	Real=Monilia	Pred=Monilia	60.85%	Correcto
Prueba 20:	Real=Fito	Pred=Monilia	86.33%	Error
Prueba 21:	Real=Monilia	Pred=Monilia	91.12%	Correcto
Prueba 22:	Real=No_Cacao	Pred=No_Cacao	100.0%	Correcto
Prueba 23:	Real=Fito	Pred=Fito	81.04%	Correcto
Prueba 24:	Real=No_Cacao	Pred=No_Cacao	100.0%	Correcto
Prueba 25:	Real=No_Cacao	Pred=No_Cacao	100.0%	Correcto
Prueba 26:	Real=Monilia	Pred=Monilia	97.18%	Correcto
Prueba 27:	Real=Fito	Pred=Fito	42.01%	Correcto
Prueba 28:	Real=Sana	Pred=Sana	98.86%	Correcto
Prueba 29:	Real=Monilia	Pred=Monilia	65.14%	Correcto
Prueba 30:	Real=Sana	Pred=Sana	77.29%	Correcto
Aciertos : 28/30				
Errores : 2/30				
Porcentaje: 93.33%				
--- Por clase ---				
Estres_Abiotico : 3/3 (100.0%)				
Fito : 8/9 (88.9%)				
Monilia : 8/8 (100.0%)				
No_Cacao : 4/4 (100.0%)				
Sana : 5/6 (83.3%)				
Guardado en prueba_calidad_v4.csv OK				

Figura 20 - Calidad de respuesta

```

PS D:\Usuario lilibeth\Desktop\tesis cacao> python verificar_precision.py
WARNING: All log messages before absl::InitializeLog() is called are written to STDERR
I0000 00:00:1785711912.299100 30328 port.cc:153] oneDNN custom operations are on. You may see slightly different numerical results due to floating-point round-off
f errors from different computation orders. To turn them off, set the environment variable 'TF_ENABLE_ONEDNN_OPTS=0'.
WARNING: All log messages before absl::InitializeLog() is called are written to STDERR
I0000 00:00:1785711915.913078 30328 port.cc:153] oneDNN custom operations are on. You may see slightly different numerical results due to floating-point round-off
f errors from different computation orders. To turn them off, set the environment variable 'TF_ENABLE_ONEDNN_OPTS=0'.
I0000 00:00:1785711918.201281 30328 cpu_feature_guard.cc:227] This TensorFlow binary is optimized to use available CPU instructions in performance-critical opera
tions.
To enable the following instructions: SSE3 SSE4.1 SSE4.2 AVX AVX2 FMA, in other operations, rebuild TensorFlow with the appropriate compiler flags.
WARNING:tensorflow:TensorFlow GPU support is not available on native Windows for TensorFlow >= 2.11. Even if CUDA/cuDNN are installed, GPU will not be used. Please
use WSL2 or the TensorFlow-DirectML plugin.
Found 312 images belonging to 5 classes.
20/20 17s 778ms/step - accuracy: 0.8109 - loss: 0.8019
Total de imágenes evaluadas: 312
Precisión real sobre TODO el set de validación: 81.09%
PS D:\Usuario lilibeth\Desktop\tesis cacao>

```

Figura 21 - Terminal con resultados por clase

4.4.4.8 Sprint 8

Campo	Detalle
Sprint	8
Duración	2 semanas (10 días hábiles)
Semana	15 – 16
Objetivo	Exportar el modelo TFLite e integrarlo con la aplicación móvil del grupo de trabajo.
HU	HU-08, HU-10
Entregable	Archivo, modelo_salud_vegetal.tflite entregado al equipo de integración con documentación.

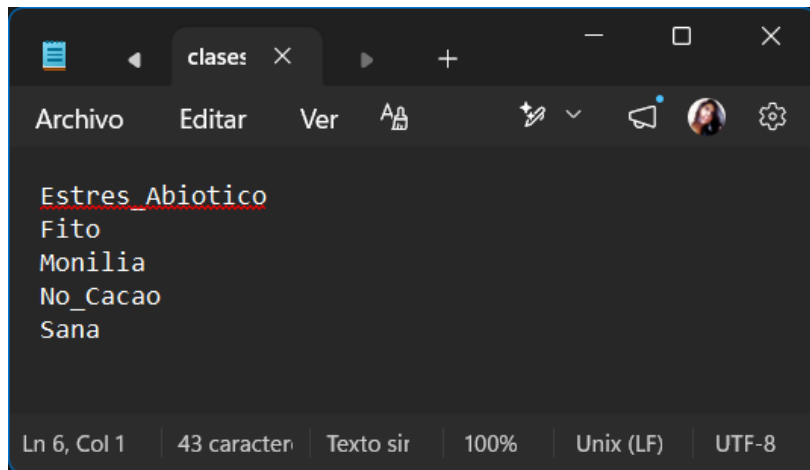
Tabla 13 - Exportación e integración del modelo

Tareas del Sprint

- Conversión del modelo Keras a formato TensorFlow Lite.
- Generación del archivo clase.txt con las 5 clases del modelo.
- Redacción de la documentación de integración para el equipo móvil.
- Entrega del modelo TFLite y documentación al responsable de la interfaz móvil.

```
PS D:\Usuario Lilibeth\Desktop\tesis_cacao> python convertir_tflite.py
1681325736784: TensorSpec(shape=(), dtype=tf.resource, name=None)
1681325739472: TensorSpec(shape=(), dtype=tf.resource, name=None)
1681325740048: TensorSpec(shape=(), dtype=tf.resource, name=None)
W0000 00:00:1785725266.713397 32284 tf_tfl_flatbuffer_helpers.cc:364] Ignored output_format.
W0000 00:00:1785725266.713653 32284 tf_tfl_flatbuffer_helpers.cc:367] Ignored drop_control_dependency.
I0000 00:00:1785725266.716328 32284 reader.cc:83] Reading SavedModel from: C:\Users\user\AppData\Local\Temp\tmp1814mhhhd
I0000 00:00:1785725266.730630 32284 reader.cc:52] Reading meta graph with tags { serve }
I0000 00:00:1785725266.730820 32284 reader.cc:147] Reading SavedModel debug info (if present) from: C:\Users\user\AppData\Local\Temp\tmp1814mhhhd
I0000 00:00:1785725266.852929 32284 mlir_graph_optimization_pass.cc:437] MLIR V1 optimization pass is not enabled
I0000 00:00:1785725266.876107 32284 loader.cc:236] Restoring SavedModel bundle.
I0000 00:00:1785725267.887674 32284 loader.cc:220] Running initialization op on SavedModel bundle at path: C:\Users\user\AppData\Local\Temp\tmp1814mhhhd
I0000 00:00:1785725268.098399 32284 loader.cc:471] SavedModel load for tags { serve }; Status: success: OK. Took 1387761 microseconds.
I0000 00:00:1785725268.343535 32284 dump_mlir_util.cc:269] disabling MLIR crash reproducer, set env var 'MLIR_CRASH_REPRODUCER_DIRECTORY' to enable.
✅ Conversión exitosa: modelo_cacao_v4.tflite
```

Figura 22 - • Conversión del modelo Keras a formato TensorFlow Lite

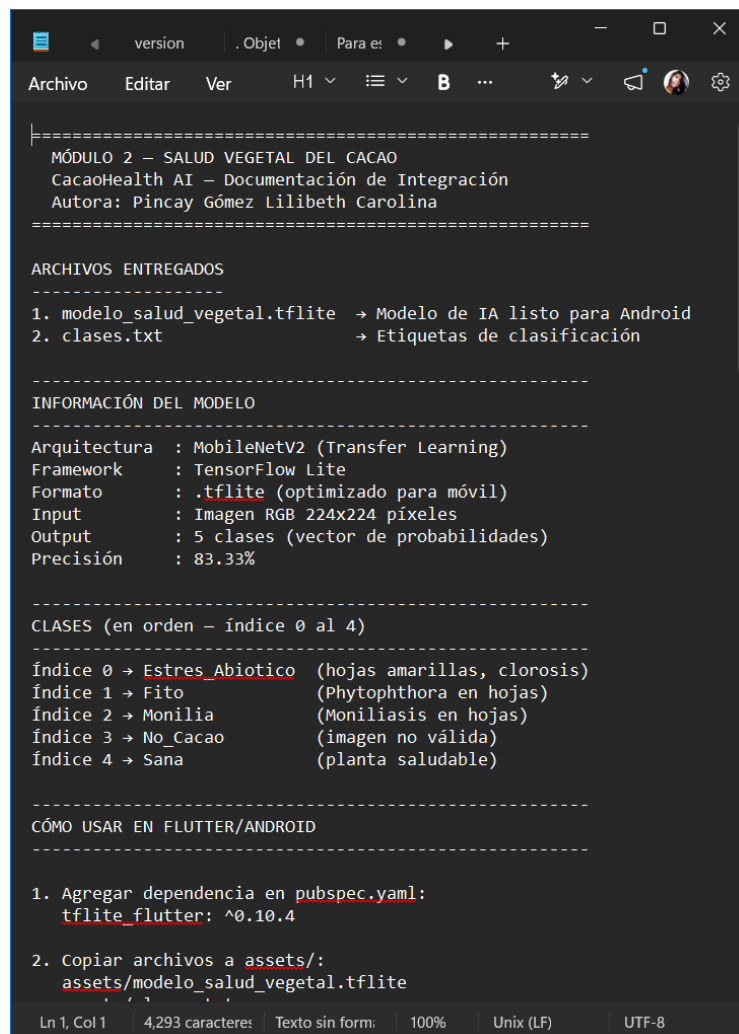


A screenshot of a code editor window titled 'clases'. The editor displays the following text:

```
Estres_Abiotico
Fito
Monilia
No_Cacao
Sana
```

The status bar at the bottom indicates 'Ln 6, Col 1', '43 caracteres', 'Texto sin', '100%', 'Unix (LF)', and 'UTF-8'.

Figura 23 - archivo clase.txt con las 5 clases del modelo



A screenshot of a code editor window showing integration documentation. The text is as follows:

```
=====
MÓDULO 2 – SALUD VEGETAL DEL CACAO
CacaoHealth AI – Documentación de Integración
Autora: Pincay Gómez Lilibeth Carolina
=====

ARCHIVOS ENTREGADOS
-----
1. modelo_salud_vegetal.tflite → Modelo de IA listo para Android
2. clases.txt → Etiquetas de clasificación

-----

INFORMACIÓN DEL MODELO
-----
Arquitectura : MobileNetV2 (Transfer Learning)
Framework : TensorFlow Lite
Formato : .tflite (optimizado para móvil)
Input : Imagen RGB 224x224 píxeles
Output : 5 clases (vector de probabilidades)
Precisión : 83.33%

-----

CLASES (en orden – índice 0 al 4)
-----
Índice 0 → Estres_Abiotico (hojas amarillas, clorosis)
Índice 1 → Fito (Phytophthora en hojas)
Índice 2 → Monilia (Moniliasis en hojas)
Índice 3 → No_Cacao (imagen no válida)
Índice 4 → Sana (planta saludable)

-----

CÓMO USAR EN FLUTTER/ANDROID
-----

1. Agregar dependencia en pubspec.yaml:
   tflite_flutter: ^0.10.4

2. Copiar archivos a assets/:
   assets/modelo_salud_vegetal.tflite
```

The status bar at the bottom indicates 'Ln 1, Col 1', '4,293 caracteres', 'Texto sin form:', '100%', 'Unix (LF)', and 'UTF-8'.

Figura 24 - Redacción de la documentación de integración para el equipo móvil

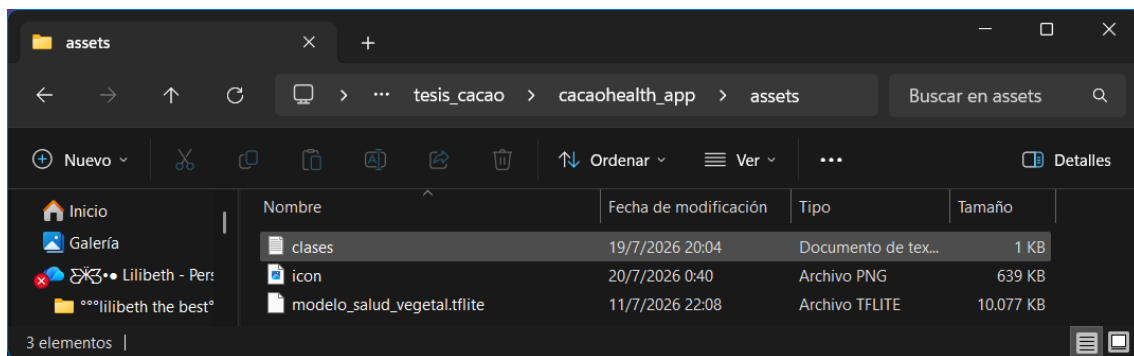


Figura 25 - Entrega del modelo TFLite y documentación al responsable de la interfaz móvil

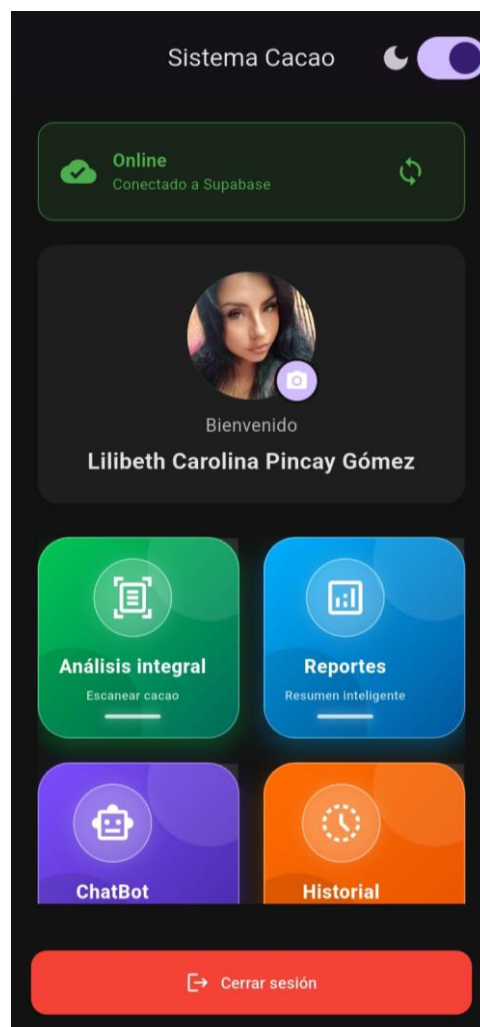


Figura 26 - Aplicación en funcionamiento (Android)

4.4.4.9 Sprint 9

Campo	Detalle
Sprint	9
Duración	1 semana (5 días hábiles)
Objetivo	Vincular las pruebas de evaluación del sistema con las causas del problema identificadas en el diagnóstico inicial y determinar el porcentaje de mejora obtenido respecto a la situación previa.
HU	HU-09
Entregable	Análisis comparativo entre el problema identificado y el porcentaje de mejora alcanzado con el sistema desarrollado.

Tabla 14 - Evaluación de efectividad del sistema frente a causas del problema

Tareas del Sprint

- Identificación de las causas raíz de problemas.
- Vinculación de las causas con pruebas técnicas del sistema.
- Simulación comparativa entre diagnóstico manual y diagnóstico automatizado.

Aspecto	Situación previa (diagnóstico manual)	Con CacaoHealth AI
Consistencia del diagnóstico	Variable con inconsistencias reportadas por 86.2% de los encuestados.	Constante. 81.09% de precisión en todas las pruebas.
Tiempo de diagnóstico	Depende de la experiencia del agricultor.	361.33 ms promedio por imagen.
Objetividad	Sujeta a percepción visual individual.	Basada en patrones aprendidos de forma consistente.

Tabla 15 - Comparación entre diagnóstico manual y diagnóstico automatizado

CAPÍTULO V

5 EVALUACIÓN DE RESULTADOS

5.1 Introducción

En el presente capítulo se exponen los resultados obtenidos tras el desarrollo, el entrenamiento y la validación del sistema CacaoHealth AI, orientado a la evaluación del estado de salud vegetal del cacao (*Theobroma cacao*) mediante técnicas de Deep Learning. El proceso de evaluación se llevó a cabo mediante pruebas técnicas de tiempo de respuesta y precisión de clasificación ejecutadas sobre el modelo entrenado, utilizando imágenes del conjunto de validación no empleadas durante el entrenamiento. Los resultados obtenidos permiten verificar el cumplimiento de los objetivos específicos planteados en el Capítulo I y demuestran la utilidad práctica del sistema para el diagnóstico fitosanitario del cacao en la región.

5.2 Presentación y monitoreo de los resultados

5.2.1 Planificación de evaluación

Componente	Detalle
Sistema evaluado	CacaoHealth AI – Módulo de salud vegetal del cacao
Modelo utilizado	modelo_cacao_v4.keras (MobileNetV2 con Transfer Learning)
Dataset de validación	dataset_v4/validation – 312 imágenes reales no usadas en entrenamiento
Muestra de prueba	30 imágenes seleccionadas aleatoriamente
Tipos de prueba	Prueba 1: Tiempo de respuesta / Prueba 2: Calidad de clasificación
Clases evaluadas	Estrés Abiótico, Fito, Monilia, No Cacao y Sana
Criterio de éxito	Precisión general $\geq 85\%$ y tiempo de respuesta ≤ 500 ms
Herramienta	Python 3.10 - TensorFlow/Keras, pruebas validación.py

Tabla 16 - Planificación de la evaluación del sistema CacaoHealth AI

5.2.2 Ejecución del monitoreo

Se ejecutaron las pruebas de validación sobre el modelo v4 utilizando 30 imágenes del conjunto de validación. A continuación se presentan los resultados obtenidos en cada tipo de prueba.

Prueba 1 - Tiempo de respuesta

Métrica	Valor obtenido	Interpretación
Tiempo promedio	361.33 ms	Tiempo adecuado para uso en campo dentro del criterio establecido (≤ 500 ms).
Tiempo mínimo	175.27 ms	Respuesta muy rápida en condiciones óptimas
Tiempo máximo	2859.21 ms	Primer diagnóstico, carga inicial del modelo en memoria
Total de pruebas	30 imágenes	Muestra representativa del dataset de validación

Tabla 17 - Resultados de tiempo de respuesta del sistema CacaoHealth AI

Nota: Resultados guardados en prueba_tiempo_v4.csv

Prueba 2 - Calidad de clasificación

Clases	Aciertos	Total	Precisión	Interpretación
Estrés_Abiótico	3	3	100%	Excelente - reconoce con alta confianza
Monilia	8	8	100%	Excelente - sin errores en esta muestra
No_cacao	4	4	100%	Excelente - rechaza imágenes inválidas correctamente.
Sana	5	6	83.3%	Bueno - leve confusión con Monilia.
Fito	8	9	88.9%	Bueno - leve confusión con Monilia.
Total (muestra)	28	30	93.33%	Resultado sobre muestra aleatoria de 30 imágenes.

Tabla 18 - Resultados de calidad de clasificación por clases (modelo_v4)

Resultados guardados en el archivo prueba_calidad_v4.csv

Nota: la precisión general reportada oficialmente en este trabajo corresponde a la evaluación sobre la totalidad de 312 imágenes del conjunto de validación (ver interpretación objetiva, más adelante en este capítulo), siendo esta última la métrica estadística más representativa.

5.3 Interpretación objetiva

Para obtener la precisión general del sistema, el modelo_cacao_v4.keras fue evaluado sobre la totalidad de 312 imágenes del conjunto de validación, un subconjunto de imágenes reales que no fue utilizado durante el entrenamiento. Para cada imagen el sistema generó una predicción de clase que fue comparada con la etiqueta real asignada previamente durante la organización del dataset. La precisión general se calculó como el cociente entre el número total de predicciones correctas y el número total de imágenes evaluada, expresado en porcentaje, mediante el script pruebas_validacion.py. Este procedimiento, al utilizar la totalidad del conjunto de validación en lugar de una muestra parcial, permite obtener una medición más representativa y estadísticamente confiable del desempeño real del modelo, en comparación con la evaluación puntual de 30 imágenes presentada en la tabla 18.

Los resultados obtenidos en las pruebas de validación del sistema CacaoHealth AI evidencian que el modelo entrenado con MobileNetV2 y Transfer Learning, alcanzó una precisión del 81.09% sobre el conjunto de 312 imágenes de validación, tras un proceso de ajuste fino aplicado sobre las capas superiores de la arquitectura base. Aunque este resultado no alcanza el criterio de éxito del 85% establecido en el capítulo I, se presenta una mejora significativa respecto al modelo inicial sin ajuste fino, que obtuvo un 70.80% de precisión. A continuación se presenta el análisis interpretativo de los resultados obtenidos.

Dimensión	Resultado	Análisis
Precisión general	81.09% (sobre 312 imágenes)	Resultado obtenido tras dos fases de entrenamiento. No alcanza el criterio del 85% establecido, aunque representa una mejora sustancial respecto al modelo inicial.
Tiempo de respuesta	361.33 ms promedio	Tiempo adecuado para uso en campo con variabilidad explicada por la carga inicial del modelo en algunas pruebas.
Clase más precisa	Estrés_Abiótico, No_Cacao y Monilia (100% en muestra)	El modelo reconoce con certeza estos casos y rechaza correctamente imágenes no relacionadas al cacao.
Clase con mayor dificultad	Sana (83.3% en muestra)	Presentó confusión ocasional con Monilia. Se requiere más imágenes de esta clase para mejorar el modelo.
Validación de imágenes	No_Cacao detectado al 100%	El sistema rechaza correctamente imágenes que no corresponden a plantas de cacao.
Objetivos específicos	Parcialmente Cumplidos	Se analizaron fundamentos de CNN, se construyó dataset propio, se implementó el prototipo, se evaluó su precisión; la validación en campo con agricultores quedó pendiente para una fase posterior.

*Tabla 19 - Interpretación objetiva de los resultados del sistema CacaoHealth AI
Elaboración propia basada en los resultados de las pruebas de validación ejecutadas*

En términos generales, el sistema CacaoHealth AI demostró ser una herramienta funcional para la evaluación del estado de salud vegetal del cacao. La precisión del 81.09% apoya el empleo de Transfer Learning junto con MobileNetV2 como una estrategia eficaz para la enseñanza de modelos de clasificación de imágenes agrícolas usando conjuntos de datos pequeños (Pan & Yang, 2010). En este estudio se encontraron las siguientes limitaciones fundamentales la clase Sana resultó ser muy poco precisa, con una confusión predominante con la clase Monilia. Para su mejora en futuras versiones, se recomienda agregar más imágenes reales de esta clase, etiquetadas por agricultores experimentados.

CAPÍTULO VI

6 CONCLUSIONES Y RECOMENDACIONES

6.1 Conclusiones

A partir del desarrollo, implementación y evaluación del sistema CacaoHealth AI, orientado a la evaluación del estado de salud vegetal del cacao (*Theobroma cacao*) mediante técnicas Deep Learning, se presentan las siguientes conclusiones:

Se detectó que los estudiantes de la carrera de Agropecuaria, Uleam Extensión El Carmen presentan dificultades en la identificación visual de enfermedades en plantas de cacao. Evidenciadas en la encuesta aplicada a 29 estudiantes de agropecuaria, donde el 86.2% reconoció inconsistencias en el diagnóstico visual, lo que justificó el desarrollo de un sistema informático basado en inteligencia artificial como solución tecnológica accesible y de bajo costo.

Los fundamentos teóricos revisados sobre Deep Learning, redes neuronales convolucionales y Transfer Learning permitieron seleccionar la arquitectura MobileNetV2 como base de modelo de clasificación, demostrando que esta arquitectura es adecuada para el desarrollo de sistemas de visión artificial en dispositivos móviles con recursos computacionales limitados.

La aplicación del método inductivo-deductivo, junto con la encuesta aplicada a estudiantes de Agropecuaria y la entrevista al Ing. Marco Vinicio de la Cruz, permitió recopilar información relevante sobre las necesidades reales del campo y validar la pertinencia del sistema desarrollado como herramienta de apoyo al diagnóstico fitosanitario del cacao.

Se desarrolló exitosamente el sistema CacaoHealth AI utilizando la arquitectura MobileNetV2 con Transfer Learning, implementado en Python con TensorFlow/Keras, aplicando la metodología ágil Scrum en nueve sprints que permitieron organizar el desarrollo de manera iterativa e incremental, logrando un prototipo funcional con interfaz gráfica en Tkinter y versión móvil en Flutter para Android.

Las pruebas de validación ejecutadas sobre el modelo v4 mostraron que el sistema alcanzó una precisión del 81.09 % sobre el conjunto completo de 312 imágenes de validación, mediante un proceso de ajuste fino (fine-tuning) sobre las capas superiores de MobileNetV2. Las clases Estrés_Abiótico y No_Cacao presentaron los niveles más altos de precisión, mientras que Fito mostró mayor dificultad de clasificación, lo que evidencia la necesidad de ampliar el dataset con más imágenes de esta clase en trabajos futuros.

En cuanto al primer objetivo específico se ha realizado una evaluación teórica de las redes neuronales convolucionales utilizadas para el monitoreo del estado de salud de las plantas lo cual ha permitido comprender las bases técnicas necesarias para elegir la arquitectura MobileNetV2 con aprendizaje transferido como base del sistema desarrollado.

En cuanto al segundo objetivo específico el análisis de la aplicación de las CNNs al diagnóstico del cacao ha permitido descubrir los enfoques al problema que se utilizan en casos análogos, como el enfoque de reconocimiento de la enfermedad en los vainos del cacao basado en el aprendizaje profundo, de esta manera guiando las elecciones técnicas de los desarrolladores en la construcción del modelo de clasificación de hojas de este proyecto.

En cuanto al tercer objetivo específico se ha recolectado 2097 imágenes de hojas de cacao en distintos estados de salud en la Granja Experimental de El Carmen divididas en cinco clases, capturadas utilizando los métodos ajustados a la práctica de campo.

En cuanto al cuarto objetivo específico las imágenes recolectadas han sido validadas en cuanto a su calidad y representatividad, completadas con los procedimientos de aumento de datos los cuales han permitido incrementar su variabilidad y por ende asegurar la idoneidad para entrenamiento del modelo.

En cuanto al quinto objetivo específico se ha creado e implementado exitosamente un prototipo funcional capaz de clasificar el estado de salud de la planta de cacao según sus imágenes, tanto en la interfaz desarrollada con Tkinter, como en la aplicación móvil para Android desarrollada con el equipo de trabajo.

En cuanto al sexto objetivo específico referente a la evaluación del sistema en el campo junto a los agricultores locales se ha realizado la prueba técnica de la precisión y el tiempo de respuesta del sistema en condiciones controladas; sin embargo la validación práctica del sistema junto a los agricultores en el campo no se ha logrado dentro del marco en este proyecto, por lo cual esta idea se plantea como una recomendación para las fases de implementación futura.

6.2 Recomendaciones

Con base en los resultados obtenidos y las limitaciones identificadas durante el desarrollo del sistema CacaoHealth AI, se formulan las siguientes recomendaciones:

Se recomienda al equipo de desarrollo del proyecto integrador ampliar el dataset de entrenamiento con un mínimo de 50 fotografías adicionales de hojas con síntomas de Fito, tomadas en diferentes condiciones de iluminación, ángulos y etapas de enfermedad, con el fin de mejorar la precisión del modelo en esta clase y reducir la confusión con la

clase Monilia y Sana, ya que Fito presentó apenas un 60% de precisión en las pruebas de validación frente al 90 – 100% alcanzado por las demás clases.

Se sugiere al estudiante responsable del módulo de recolección de datos incorporar fotografías de hojas enfermas como parte prioritaria , dado que la mayoría de las imágenes recolectadas corresponden a mazorcas, lo cual limita la capacidad del modelo para generalizar ante diferentes tipos de muestras visuales del cultivo.

Se recomienda al compañero desarrollador de la aplicación móvil extender la versión desarrollada en Flutter para Android incorporando funcionalidades adicionales como geolocalización GPS para registrar la ubicación exacta de cada diagnóstico realizado, permitiendo generar mapas de salud vegetal del cultivo y apoyar la agricultura de precisión en la Granja Experimental El Carmen.

Se sugiere al equipo de desarrollo mejorar el historial de diagnóstico que ya está implementado en la aplicación móvil, incorporando gráficas estadísticas que muestren la evolución del estado de salud del cultivo por fecha y clase detectada para facilitar el análisis histórico del comportamiento fitosanitario de la plantación.

Se recomienda a futuros investigadores o estudiantes que continúen este proyecto explorar técnicas avanzadas de fine-tuning en las capas profundas de MobileNetV2 utilizando un mayor volumen de datos, con el objetivo de superar el 85% de precisión general y mejorar especialmente el desempeño de la clase Fito, que presentó mayor dificultad de clasificación.

Se propone al tutor del proyecto y a la coordinación de la carrera gestionar la validación del sistema con agricultores reales en la Granja Experimental El Carmen, mediante pruebas de campo con imágenes capturadas en condiciones naturales. También

sería importante evaluar su usabilidad y aceptación por parte de los usuarios finales, para incluir sus observaciones en nuevas versiones de la aplicación.

BIBLIOGRAFÍA

- Bakker, P. A. H. M., Pieterse, C. M. J., & van Loon, L. C. (2007). Induced Systemic Resistance by Fluorescent *Pseudomonas* spp. *Phytopathology*®, 97(2), 239–243. <https://doi.org/10.1094/PHYTO-97-2-0239>
- Berendsen, R. L., Pieterse, C. M. J., & Bakker, P. A. H. M. (2012). The rhizosphere microbiome and plant health. *Trends in Plant Science*, 17(8), 478–486. <https://doi.org/10.1016/j.tplants.2012.04.001>
- Bernal Torres, C. A. (2010). *Metodología de la investigación: Administración, economía, humanidades y ciencias sociales* (Pearson Educación / Prentice Hall, Ed.; 3.^a edición).
- Berzal, F. (2018). *Fernando Berzal REDES NEURONALES & DEEP LEARNING*. 285–285.
- Chollet, F. (2022). Deep learning with Python (2^a ed) - François Chollet - Manning, outubro 2021, 504 pp. *Interações: Sociedade e as Novas Modernidades*, (42), 113–115. <https://doi.org/10.31211/interacoes.n42.2022.r1>
- Dávila Costa, J. S., & Romero, C. M. (2025). Nano-Biofungicides and Bio-Nanofungicides: State of the Art of Innovative Tools for Controlling Resistant Phytopathogens. *Biophysica*, 5(2), 1–16. <https://doi.org/10.3390/biophysica5020015>
- El estado del clima en América Latina y el Caribe 2023. (2024). In *El estado del clima en América Latina y el Caribe 2023*. United Nations. <https://doi.org/10.18356/9789263313515>

- Arias, F. G. (2012). El proyecto de investigación: *Introducción a la metodología científica* (6.^a ed.). Editorial Episteme.
- García, A. et al. (2023). Clasificación de imágenes médicas con técnicas de Deep Learning [Trabajo Fin de Grado, Universidad de Extremadura].
<https://dehesa.unex.es/server/api/core/bitstreams/c72d9a87-c16e-43c8-a476-a362cd4ff6c0/content>
- Gogos, A., Knauer, K., & Bucheli, T. D. (2012). Nanomaterials in plant protection and fertilization: Current state, foreseen applications, and research priorities. *Journal of Agricultural and Food Chemistry*, 60(39), 9781–9792.
<https://doi.org/10.1021/jf302154y>
- Hernández Sampieri, R., Fernández Collado, C., & Baptista Lucio, M. del P. (2018). Metodología de la investigación (1.^a ed.)
- Hernández Sampieri, R., Fernández Collado, C., & Baptista Lucio, P. (2014). *Metodología de la investigación*. McGraw-Hill/Interamericana.
- Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., & Adam, H. (2017). *MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications*.
<http://arxiv.org/abs/1704.04861>
- Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, 1(9), 389–399.
<https://doi.org/10.1038/s42256-019-0088-2>
- Kleppmann, M. (2017). Designing Data-Intensive Applications: The Big Ideas behind Reliable, Scalable, and Maintainable Systems. *O'Reilly Media, Inc.*,

569. <https://www.oreilly.com/library/view/designing-data-intensive-applications/9781491903063/>

Kogan, M. (1998). Integrated pest management: Historical perspectives and contemporary developments. In *Annual Review of Entomology* (Vol. 43, Number 1, pp. 243–270). <https://doi.org/10.1146/annurev.ento.43.1.243>

Leiva, S., Oliva, M., Hernández, E., Chuquibala, B., Rubio, K., García, F., & Torres de la Cruz, M. (2020). Assessment of the Potential of *Trichoderma* spp. Strains Native to Bagua (Amazonas, Peru) in the Biocontrol of Frosty Pod Rot (*Moniliophthora roreri*). *Agronomy*, 10(9), 1376. <https://doi.org/10.3390/agronomy10091376>

Li, R., Li, L., Huang, T., Liu, X., Chen, Q., Jin, G., & Cao, H. (2021). Gold nanoparticle-based colorimetric aptasensor for rapid detection of multiple mycotoxins in rice. *Analytical Methods*, 13(47), 5749–5755. <https://doi.org/10.1039/d1ay01809d>

Mahlein, A. K. (2016). Plant disease detection by imaging sensors – Parallels and specific demands for precision agriculture and plant phenotyping. In *Plant Disease* (Vol. 100, Number 2, pp. 241–254). <https://doi.org/10.1094/PDIS-03-15-0340-FE>

Mendes, R., Garbeva, P., & Raaijmakers, J. M. (2013). The rhizosphere microbiome: Significance of plant beneficial, plant pathogenic, and human pathogenic microorganisms. *FEMS Microbiology Reviews*, 37(5), 634–663. <https://doi.org/10.1111/1574-6976.12028>

- Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data and Society*, 3(2).
<https://doi.org/10.1177/2053951716679679>
- Montague, P. R. (1999). Reinforcement Learning: An Introduction, by Sutton, R.S. and Barto, A.G. *Trends in Cognitive Sciences*, 3(9), 360.
[https://doi.org/10.1016/s1364-6613\(99\)01331-5](https://doi.org/10.1016/s1364-6613(99)01331-5)
- Panigrahi, S., Nanda, A., & Swarnkar, T. (2021). A Survey on Transfer Learning. *Smart Innovation, Systems and Technologies*, 194, 781–789.
https://doi.org/10.1007/978-981-15-5971-6_83
- Papanastasiou, G., Dikaïos, N., Huang, J., Wang, C., & Yang, G. (2024). Is Attention all You Need in Medical Image Analysis? A Review. *IEEE Journal of Biomedical and Health Informatics*, 28(3), 1398–1411.
<https://doi.org/10.1109/JBHI.2023.3348436>
- Plasencia-Vázquez, A. H., Vilchez-Ponce, C. R., Ferrer-Sánchez, Y., & Veloz-Portillo, C. E. (2022). Effect of climate change on the potential distribution of the fungus *Moniliophthora roreri* and the cultivation of cacao (*Theobroma cacao*) in continental Ecuador. *Terra Latinoamericana*, 40.
<https://doi.org/10.28940/terra.v40i0.1151>
- Proyecto Chocolate4all Manual Manejo integrado de enfermedades en el cultivo de cacao.* (n.d.).
- Railsback, S. F., & Grimm, V. (2019). Agent-based and individual-based modeling: A practical introduction (2nd ed.). *Princeton University Press*.

- Rothman, D. (2021). Transformers for Natural Language Processing Build innovative deep neural network architectures for NLP & Python, PyTorch, TensorFlow, BERT, RoBERTa & more. *Packt Publishing*, 384. <https://www.packtpub.com/product/transformers-for-natural-language-processing/9781800565791>
- Ruilova Cumbicos, R. V. (2022). *Desarrollo de una aplicación móvil para la detección de enfermedades en la mazorca de cacao utilizando deep learning* [Trabajo de titulación, Universidad Técnica de Machala]. Repositorio Digital UTMACH. <http://repositorio.utmachala.edu.ec/handle/48000/18767>
- Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L. C. (2018). MobileNetV2: Inverted Residuals and Linear Bottlenecks. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 4510–4520. <https://doi.org/10.1109/CVPR.2018.00474>
- Savary, S., Willocquet, L., Pethybridge, S. J., Esker, P., McRoberts, N., & Nelson, A. (2019). The global burden of pathogens and pests on major food crops. *Nature Ecology and Evolution*, 3(3), 430–439. <https://doi.org/10.1038/s41559-018-0793-y>
- Sutton, R. S., & Barto, A. G. (2018). *Reinforcement Learning: An Introduction* (2.^a ed.). MIT Press.
- Tamayo y Tamayo, M. (2004). *El proceso de la investigación científica* (Lumisa, Ed.).
- Tanenbaum, A. S., & Van Steen, M. (2007). Distributed Systems: Principles and Paradigms, 2/E. In *Communication*.

- Trivedi, P., Leach, J. E., Tringe, S. G., Sa, T., & Singh, B. K. (2020). Plant–microbiome interactions: from community assembly to plant health. *Nature Reviews Microbiology*, 18(11), 607–621. <https://doi.org/10.1038/s41579-020-0412-1>
- Tunstall, L., von Werra, L., & Wolf, T. (2022). *Natural Language Processing with Transformers: Building Language Applications with Hugging Face*. O'Reilly Media.
- Wang, W., Li, Y., Zou, T., Wang, X., You, J., & Luo, Y. (2020). A novel image classification approach via dense-mobilenet models. *Mobile Information Systems*, 2020. <https://doi.org/10.1155/2020/7602384>

ANÉXOS

2/8/26, 22:06

Correo: PINCAY GOMEZ LILIBETH CAROLINA - Outlook

 Outlook

DPGA | Titulación | Periodo 2025-2 - Notificación de tutor asignado - SOFTWARE 2024 - NS - (EL CARMEN)

Desde NOTIFICACIONES TITULACION <notificaciones.titulacion@uleam.edu.ec>
Fecha Mié 13/08/2025 10:30
Para REASCOS PINCHAO RAUL SAED <raul.reascos@uleam.edu.ec>
CC PINCAY GOMEZ LILIBETH CAROLINA <e1311630212@live.uleam.edu.ec>; REASCOS PINCHAO RAUL SAED <raul.reascos@uleam.edu.ec>



Uleam
UNIVERSIDAD LAICA
ELOY ALFARO DE MANABÍ

Universidad Laica Eloy Alfaro de Manabí

Periodo 2025-2 - Notificación de tutor asignado - SOFTWARE 2024 - NS - (EL CARMEN)

Estimad@
Docente y Estudiante
Uleam

En cumplimiento de lo establecido en la Ley, el Reglamento de Régimen Académico y las disposiciones estatutarias de la Uleam, por medio de la presente se oficializa la dirección y tutoría en el desarrollo del Trabajo de Integración curricular / Trabajo de Titulación del siguiente estudiante:

Tema: SISTEMA INFORMÁTICO CON ARQUITECTURA DEEP LEARNING PARA EVALUACIÓN DE SALUD VEGETAL DE THEOBROMA CACAO EN LA GRANJA EXPERIMENTAL EL CARMEN

Estado de aprobación: Aprobado

Tipo de titulación: Trabajo de Integración Curricular

Tipo de proyecto: Trabajo de Integración Curricular / Trabajo de titulación se articula con proyectos y programas de Investigación.

Apellidos y nombres del tutor asignado: REASCOS PINCHAO RAUL SAED

Apellidos y nombres del estudiante: PINCAY GOMEZ LILIBETH CAROLINA

Carrera: SOFTWARE 2024 - NS - (EL CARMEN)

<https://outlook.cloud.microsoft/mail/d/AAQkAGVIMjQzYTBlWEyMzMINDQ1MC05ZTE2LTmWYmM2ZmE0ZmQwOAAQAOWRx4gUlpOgeQopt0pp...>

1/2

Figura 27 - Anexo 1 Tutor asignado

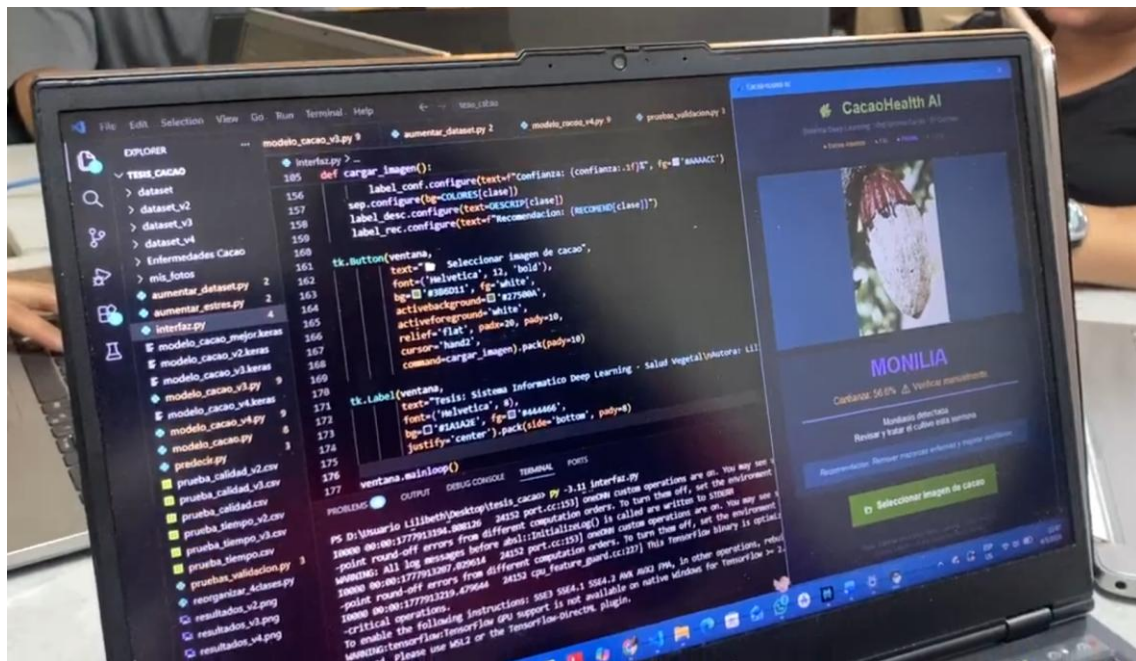


Figura 28 - Anexo 2 pruebas iniciales del prototipo en Tkinter



Figura 29 - Anexo 3 programando prototipo



Figura 30 - Anexo 4 Entrevista



Certificado de análisis
Compilatio Magister+ | ULEAM-ECU

PINCAY GOMEZ LILIBETH CAROLINA Compilatio 2026 08 04
ID : d1b00e33f3933761715a075195166681c902ebf4



8%
Textos sospechosos

Nombre del fichero : PINCAY GOMEZ LILIBETH CAROLINA Compilatio 2026 08 04.txt
Tamaño del archivo original : 4,92 MB
Número de palabras : 12.976

Depositante : RAUL REASCOS PINCHAO
Fecha de depósito : 4 de agosto de 2026
Tipo de carga : interface
fecha de fin de análisis : 4 de agosto de 2026

Resumen (sección 1/2)

Localización de los textos sospechosos en el documento :



Incluido en el porcentaje de textos sospechosos :

Similitudes <1%

Pasajes con similitudes a fuentes encontradas en diferentes colecciones.



Detección de IA 4%

Textos estilísticamente próximos a un texto generado por una IA.

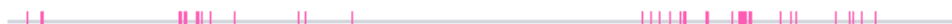
Este índice es un indicador y no una prueba. Comprueba con el autor si domina los conocimientos mencionados en el documento.



Idiomas no reconocidos 4%

Pasajes en los que parte del vocabulario utilizado no forma parte del diccionario de la lengua.

Puede tratarse de un intento del autor de modificar el texto para evitar ser detectado.



No incluido en el porcentaje de textos sospechosos :

Textos entre comillas 0%

Pasajes entre comillas, a menudo indicativos de una cita.

Figura 31- Certificado de análisis

GLOSARIO

CNN (Redes Neuronales Convolucionales): Tipo de red neuronal, diseñada especialmente para procesar imágenes, funciona identificando patrones visuales como bordes, textura y forma, que luego combina para reconocer objetos más complejos, como enfermedades en hojas de cacao.

class weights (Pesos de clase): Técnica utilizada cuando un dataset tiene un número desigual de imágenes por categoría, asigna mayor importancia a las clases con menor ejemplo durante el entrenamiento, evitando que el modelo favorezca únicamente a las clases con más datos disponibles.

confianza: Porcentaje que indica qué tan seguro está el modelo de que su predicción es correcta. Un valor de confianza alto significa que el sistema detectó patrones muy similares a los aprendidos durante el entrenamiento para esta clase específica.

conjunto : Conjunto de datos utilizado para entrenar y evaluar un modelo de inteligencia artificial. Un dataset se organiza realmente en categorías o clases, dependiendo de lo que el modelo debe aprender a reconocer.

¡Error! No se encuentra el origen de la referencia. **(Aumento de datos):** Técnica que genera variaciones artificiales de las imágenes originales en un dataset (rotándolas, cambiando su brillo, volteándolas, etc.), con el fin de aumentar la cantidad y diversidad de ejemplos disponibles para entrenar el modelo sin necesidad de recolectar más fotografías.

Deep Learning (Aprendizaje Profundo): Rama de la inteligencia artificial basada en redes neuronales con múltiples capas, capaz de aprender patrones complejos directamente a partir de grandes volúmenes de datos, como imágenes o texto.

Dropout: Técnica de regularización utilizada en redes neuronales que desactiva aleatoriamente un porcentaje de neuronas durante el entrenamiento, con el fin de evitar que el modelo dependa demasiado de patrones específicos y mejorar su capacidad de generalización.

fine-tuning (Ajuste fino): Técnica utilizada en Transfer Learning que consiste en reentrenar parcialmente un modelo ya preentrenado, ajustando sus parámetros internos para especializarlos en una nueva tarea específica, en lugar de entrenarlo completamente desde cero.

Flutter: Framework de desarrollo creado por Google que permite construir aplicaciones móviles multiplataforma a partir de un solo código fuente.

MobileNetV2 : Arquitectura de red neuronal convolucional diseñada para funcionar de forma eficiente en dispositivos con recursos limitados, como teléfonos móviles, manteniendo un buen equilibrio entre precisión y velocidad de procesamiento.

precisión (Accuracy): Métrica que indica el porcentaje de predicciones correctas realizadas por un modelo, en relación con el total de predicciones evaluadas. Es uno de los indicadores más comunes para medir el desempeño de un sistema de clasificación.

Scrum : Metodología ágil de gestión de proyectos que organiza el trabajo en ciclos cortos e iterativos llamados sprints, permitiendo revisar avances, hacer ajustes y entregar resultados de forma incremental.

TensorFlow Lite (TFLite): Versión optimizada del framework TensorFlow, diseñada para ejecutar modelos de inteligencia artificial en dispositivos móviles y de bajos recursos, reduciendo el tamaño del modelo sin afectar significativamente su desempeño.

Tkinter: Biblioteca estándar de Python, utilizada para crear interfaces gráficas de usuario (GUI) de escritorio empleada en proyectos para desarrollar la interfaz que permite cargar imágenes y visualizar el diagnóstico del sistema.

Transfer Learning (Aprendizaje por Transferencia): Técnica que aprovecha el conocimiento previamente aprendido por un modelo entrenado en un problema general (como el reconocimiento de miles de objetos distintos), para adaptarlo a resolver un problema más específico con menos datos y menos tiempo de entrenamiento.