

APLICACIÓN DE TÉCNICAS DE INTELIGENCIA ARTIFICIAL PARA LA
DETECCIÓN DE ENFERMEDADES FOLIARES EN EL CULTIVO DE TOMATE
(SOLANUM LYCOPERSICUM) DURANTE LA ETAPA DE CRECIMIENTO
VEGETATIVO EN LA REGIÓN ANDINA

OSCAR SANTIAGO BOLIVAR RONCANCIO
SERGIO ALBERTO SÁNCHEZ CASTAÑEDA

FUNDACIÓN UNIVERSITARIA AGRARIA DE COLOMBIA - UNIAGRARIA
FACULTAD INGENIERÍA
PROGRAMA DE INGENIERÍA MECATRÓNICA
BOGOTÁ DC
2023

APLICACIÓN DE TÉCNICAS DE INTELIGENCIA ARTIFICIAL PARA LA
DETECCIÓN DE ENFERMEDADES FOLIARES EN EL CULTIVO DE TOMATE
(SOLANUM LYCOPERSICUM) DURANTE LA ETAPA DE CRECIMIENTO
VEGETATIVO EN LA REGIÓN ANDINA

OSCAR SANTIAGO BOLIVAR RONCANCIO
SERGIO ALBERTO SÁNCHEZ CASTAÑEDA

TRABAJO DE GRADO PARA OPTAR POR EL TÍTULO DE INGENIERO
MECATRÓNICO

DIRECTOR
Ing. VALENTINA REYES MARTINEZ DOCENTE MT

FUNDACIÓN UNIVERSITARIA AGRARIA DE COLOMBIA - UNIAGRARIA
FACULTAD INGENIERÍA
PROGRAMA DE INGENIERÍA MECATRÓNICA
BOGOTÁ DC
2023

Nota de Aceptación

Presidente del Jurado

Jurado

Jurado

Ciudad y Fecha (día, mes, año) (Fecha de entrega)

DEDICATORIA

A mis padres, fuente inagotable de amor, apoyo y sacrificio. Cada logro en esta travesía académica es un reflejo de su dedicación y enseñanzas.

A mis amigos y seres queridos, por su paciencia, comprensión y aliento constante. Su presencia ha iluminado los días más desafiantes.

A mis profesores y mentores, cuyas enseñanzas han guiado mi camino hacia el conocimiento y el crecimiento personal.

Este trabajo está dedicado a todos aquellos que han sido parte integral de mi camino, contribuyendo con su amor y apoyo a mi éxito académico.

AGRADECIMIENTOS

En primer lugar, les agradecemos a nuestros padres que siempre nos han brindado su apoyo incondicional para poder cumplir todos los objetivos personales y académicos. Ellos son los que con su cariño nos han impulsado siempre a perseguir nuestras metas y nunca abandonarlas frente a las adversidades. También son los que nos han brindado el soporte material y económico para poder concentrarnos en los estudios y nunca abandonarlos.

Le agradecemos muy profundamente a nuestra directora Valentina Reyes y nuestro exdirector Alejandro Quiroga por su dedicación y paciencia, sin sus palabras y correcciones precisas no habiésemos podido lograr llegar a esta instancia tan anhelada. Gracias por su guía y todos sus consejos, los llevaremos grabados para siempre en la memoria y en nuestro futuro profesional.

Son muchos los docentes que han sido parte de nuestro camino universitario, y a todos ellos les queremos agradecer por transmitirnos los conocimientos necesarios para hoy poder estar aquí. Sin ustedes los conceptos serían solo palabras, y las palabras ya sabemos quién se las lleva, el viento.

Agradecerles a todos nuestros compañeros los cuales muchos de ellos se han convertido en nuestros amigos, cómplices y hermanos. Gracias por las horas compartidas, los trabajos realizados en conjunto y las historias vividas.

Por último, agradecer a la universidad que nos ha exigido tanto, pero al mismo tiempo nos ha permitido obtener nuestro tan ansiado título. Agradecer a cada directivo por su trabajo y por su gestión, sin lo cual no estarían las bases ni las condiciones para aprender conocimientos.

CONTENIDO

	Pág.
1. FORMULACIÓN DEL PROBLEMA	12
2. JUSTIFICACIÓN	16
3. OBJETIVOS	20
3.1 OBJETIVO GENERAL	20
3.2 OBJETIVOS ESPECÍFICOS.....	20
4. MARCO REFERENCIAL	21
4.1 MARCO TEÓRICO	21
4.2 MARCO CONCEPTUAL	23
4.3 ESTADO DEL ARTE / MARCO HISTORICO	31
5. DISEÑO METODOLÓGICO	39
5.1 TIPO DE INVESTIGACIÓN	39
5.2 METODOLOGIA	39
5.3 FASES DEL PROYECTO	42
6. RECURSOS DISPONIBLES	53
7. CRONOGRAMA.....	54
8. BIBLIOGRAFÍA	55
9. ANEXOS	60

LISTA DE TABLAS

Tabla 1 Produccion mundial de tomate.....	13
Tabla 2 Cantidad de imágenes de tomates con enfermedad y sanos	43
Tabla 3 Numero de etiqueta y etiqueta a la que pertenece	44
Tabla 4 Resultados de las predicciones con nuevas imágenes.....	49
Tabla 5 Costos de elementos implementados	53

LISTA DE FIGURAS

Figura 1 Antioquia y norte de Santander son los departamentos lideres en la producción de tomate	19
Figura 2 ¿Qué es el algoritmo de k vecinos más cercanos?	26
Figura 3 Estructura de una red neuronal convolucional para la clasificación de una imagen	28
Figura 4 Arquitectura del modelo ResNet50	29
Figura 5 Diagrama de metodología a seguir	41
Figura 6 Fases del proyecto	42
Figura 7 Resultados del clasificador Naive Bayes (Raw Data y Feature Extracted).	45
Figura 8 Resultados del clasificador KNN (Raw Data y Feature Extracted).	46
Figura 9 Resultados del clasificador SVM (Raw Data y Feature Extracted).	47
Figura 10 Resultados de la red ResNet-50 (Feature Extracted & Predictions).	48

LISTA DE ANEXOS

Anexo 9.1.....	60
Anexo 9.2.....	62
Anexo 9.3.....	66
Anexo 9.4.....	67

GLOSARIO

Inteligencia Artificial- Campo de estudio de la informática que busca crear sistemas y máquinas capaces de emular la inteligencia y comportamientos racionales del ser humano para la resolución de problemas. Utiliza disciplinas como ciencia de datos, aprendizaje automático, procesamiento de lenguaje natural, etc (Russell y Norvig, 2004).

Solanum Lycopersicum- Nombre científico de la especie botánica del tomate cultivado comúnmente en agricultura. Pertenece a la familia Solanaceae, que incluye otras especies de gran importancia alimenticia como papa, berenjena y pimentón (Blanca et al., 2015).

Clasificador- Algoritmo de aprendizaje automático que asigna categorías o clases a datos nuevos o no vistos previamente, después de haber sido entrenado con datos etiquetados o clasificados previamente. Permite organizar y estructurar conjuntos de datos (MathWorks, 2023).

Enfermedades Foliare- Enfermedades que afectan el tejido foliar de las plantas, manifestándose a través de síntomas visibles en las hojas como manchas, tizones, pudriciones o deformaciones. Pueden ser causadas por hongos, bacterias, virus o condiciones ambientales adversas (Agris, 2005).

Métricas de Evaluación- Medidas cuantitativas que permiten valorar y comparar el rendimiento de modelos de aprendizaje automático después de su entrenamiento. Algunas métricas comunes son precisión, exactitud, sensibilidad, curva ROC, error cuadrático medio, etc (Tharwat, 2021).

Crecimiento Vegetativo- Etapa del crecimiento de las plantas en la que se llevan a cabo los procesos fisiológicos de expansión foliar, elongación del tallo, desarrollo de ramas y enraizamiento. Precede a la etapa reproductiva de floración y fructificación (Taiz y Zeiger, 2006).

RESUMEN

El cultivo de tomate se ve frecuentemente afectado por enfermedades foliares que impactan negativamente su productividad. La detección temprana de estas enfermedades permite tomar medidas de control oportunas y evitar pérdidas económicas. En este trabajo se desarrolló una herramienta para la detección automática de enfermedades foliares en tomate mediante técnicas de inteligencia artificial.

Se utilizó un conjunto de 13478 imágenes de hojas sanas y con enfermedades para entrenar y evaluar clasificadores como Naive Bayes, KNN, SVM y ResNet-50. Los mejores resultados se obtuvieron con la red neuronal convolucional profunda ResNet-50, logrando una precisión de validación de 0.9754.

Esto demuestra la efectividad de las CNN para extraer características discriminativas directamente de las imágenes sin requerir selección manual de características. La herramienta desarrollada representa una solución prometedora para el diagnóstico automatizado de enfermedades foliares en el cultivo de tomate, permitiendo la detección temprana para reducir impactos económicos negativos.

PALABRAS CLAVE: detección de enfermedades, cultivo de tomate, inteligencia artificial, aprendizaje profundo, redes neuronales convolucionales, ResNet-50.

1. FORMULACIÓN DEL PROBLEMA

Hoy en día, el aumento de la población mundial va ligado a la alta demanda de alimentos. Por consiguiente, la agricultura es una fuente de desarrollo para el crecimiento económico de muchos países. Por esta razón, es de vital importancia darle el reconocimiento necesario al papel que juega el sector agro en cuanto al suministro de alimentos, el desarrollo económico, la sostenibilidad ambiental y la seguridad alimentaria (Perfett et al, 2013).

Según datos de la Organización de Naciones Unidas para la Alimentación y la Agricultura (FAO) y el Banco Mundial, las pérdidas y desperdicio de alimentos representan cerca de un tercio de los alimentos producidos en el mundo (Gustavsson et al, 2012). La FAO dio a conocer que la agricultura mundial registra pérdidas cercanas a los 300 mil millones de dólares (mdd) por plagas y enfermedades en las plantas, y estimó que hasta el 40 % de los cultivos alimentarios se pierde cada año (en Alimentos, 2022). El *Solanum lycopersicum* (tomate) es uno de los alimentos más importantes de la familia Solanaceae en Colombia y el mundo, debido a su elevado consumo en fresco y como producto transformado por la agroindustria (Carreño et al, 2007; Salazar et al, 2011).

Por otro lado, según Agro Bayer Colombia (s.f.), el tomate es plantado en al menos 21 departamentos del país, alcanzando un área de siembra de aproximadamente 9 mil hectáreas y una producción de 512 mil toneladas por año, con un promedio de producción de 62,3 toneladas por hectárea. El 90% de la producción de tomate está concentrada en 10 departamentos de la región Andina siendo Boyacá el de mayor productividad y el primer productor de tomate bajo invernadero seguido de caldas, Risaralda y Cundinamarca. Además, este es uno de los cultivos más tecnificados y productivos del país alcanzando hasta 100 Ton/ha.

Según Laura Brenes (2015), a partir de la cumbre de la Comunidad de Estados Latinoamericanos y Caribeños (CELAC), se incluyó la reducción de Pérdidas y Desperdicios de Alimentos como línea de acción de su Plan de Acción para la Seguridad Alimentaria, Nutrición y Erradicación del Hambre 2025, por lo que la temática sigue cobrando relevancia como herramienta puntual para lograr estas metas y apoyar el consumo y producción sostenibles. El tomate es considerado uno

de los principales cultivos a nivel mundial debido a su elevado potencial alimenticio y altos contenidos de licopeno, vitaminas C y A y flavonoides. Estos compuestos son considerados antioxidantes y se encuentran asociados con la prevención de enfermedades de tipo carcinogénicas y cardiovasculares.

El consumo regular de vegetales, incluyendo el tomate, está asociado con numerosos beneficios para la salud, como la prevención de enfermedades crónico-degenerativas y cardiovasculares. Además, el contenido y estabilidad de estos compuestos antioxidantes depende del cultivar o de la variedad utilizada, condiciones ambientales de cultivo, estado de maduración del fruto y tratamientos postcosecha. Por lo tanto, el tomate es importante por su valor nutricional y por sus propiedades antioxidantes que pueden contribuir a la prevención de enfermedades (Guevara & Alvarado, 2014).

El tomate se ha consolidado como una de las hortalizas de mayor demanda a nivel mundial, su área de cultivo se ha incrementado en los últimos años significativamente. En 2008, a nivel mundial se tenían sembradas 5.227.883 hectáreas con una producción de 129.649.883 toneladas Si bien se cultiva tomate en más de cien países, tanto para consumo fresco como para industria, los diez principales productores concentran más del 70 % del total mundial. Colombia ocupa el puesto No. 34 en la producción mundial de Tomate.

No.	País	Producción (Ton)	Observaciones
1	China	48.576.853	*
2	India	16.826.000	
3	Estados Unidos de América	12.624.700	
4	Turquía	11.003.400	
5	Egipto	8.105.260	
6	Irán (República Islámica del)	6.824.300	Im
7	Italia	5.950.220	
8	Brasil	4.416.650	
9	España	3.821.490	
10	Uzbekistán	2.585.000	*
34	Colombia	595.299	

Tabla 1 = Cifras no oficiales | [] = Datos oficiales | F = Estimación FAO | Im = Datos de FAO basados en una metodología de imputación | M = Datos no disponibles

Fuente: FAOSTAT | © FAO Dirección de Estadística 2010

En Colombia, para el mismo año, el área cultivada fue de 14.855 hectáreas con una producción de 595.299 toneladas (FAO, 2010). La producción de tomate es común en casi todas las zonas, y el departamento del Norte de Santander lidera la producción participando con el 29% del total de la producción nacional, seguido de Antioquia con 11,4%, Boyacá con 11,3%, Santander con 10,4% y Cundinamarca con el 6,5%. En el año 2013 hubo exportaciones de tomate desde Colombia por valor de 298.565 dólares, siendo los principales países destino de las importaciones Antillas Holandesas, Aruba, Ecuador y España. Los principales departamentos exportadores fueron Cundinamarca con una participación del 52,8% del total exportado, Atlántico con 37,9%, Bogotá con 6,3%, Santander con 2,3% y Risaralda con 0,8% (Cubillos, 2015).

El crecimiento de este cultivo en el país se debe principalmente al mejoramiento de los circuitos comerciales aunado a la tecnificación de los cultivos, lo que ha facilitado el incremento en área sembrada de la hortaliza (Perilla et al., 2011). Este cultivo se implementa en diferentes regiones del país como plantación comercial o como apoyo de economías familiares en pequeños huertos. Aunque la producción de tomate en el país ha sido suficiente para satisfacer la demanda interna, el productor se ve afectado económicamente por pérdidas causadas por diferentes plagas y enfermedades (Amaya et al, 2009).

En Colombia, gracias a la investigación, se ha logrado en buena parte mejorar los niveles de productividad de diversos cultivos. Sin embargo, para la fase de crecimiento vegetativo es poca la tecnología disponible, razón por la cual las pérdidas presentes en los procesos siguientes a la recolección cada día son mayores. Según Hurtado (1984), dentro de las causas biológicas, los hongos, insectos y aves tuvieron una importancia relativamente similar, pues las pérdidas fueron de 1.49%, 1.31% y 1.60% respectivamente.

Su presencia en esta etapa es atribuible a los recolectores quienes cosechan toda clase de frutos sin importarles el estado sanitario en que se encuentren. Posteriormente estos frutos mezclados con los sanos se convertirán en fuentes de contaminación en las subsiguientes etapas, aumentando de esa manera las pérdidas de postcosecha. Esta situación demuestra la importancia de la sanidad de los cultivos de tomate durante su periodo de crecimiento vegetativo y principalmente en la etapa productiva de la planta.

Un estudio realizado en cultivos de tomate de árbol en Silvania y Fusagasugá demostró la importancia de la clasificación adecuada de los frutos en el proceso de postcosecha. Se observó que el 87.5% de los productores en Silvania y el 94% de los productores en Fusagasugá sufrían pérdidas económicas debido a la falta de clasificación en la postcosecha. Además, se identificaron pérdidas físicas, biológicas y mecánicas, siendo el transporte el factor principal para las pérdidas físicas y mecánicas. Mientras que las enfermedades eran el factor principal en las pérdidas biológicas que contribuía a las pérdidas irrecuperables (William, 2021).

La importancia de utilizar inteligencia artificial en la detección de enfermedades en frutas radica en la posibilidad de detectar enfermedades en una etapa temprana, lo que puede reducir las pérdidas económicas en la industria alimentaria. Además, la detección temprana de enfermedades en frutas puede ayudar a prevenir la propagación de enfermedades a otras plantas y reducir la necesidad de pesticidas y otros productos químicos (Aysun, 2020). Por esta razón, el uso de inteligencia artificial para la identificación de enfermedades en cultivos de tomate durante la etapa de crecimiento vegetativo se ha convertido en una alternativa prometedora.

¿Cómo puede la inteligencia artificial ser utilizada para identificar enfermedades en la etapa de crecimiento vegetativo en el cultivo de *solanum lycopersicum* (tomate) en la región andina?

2. JUSTIFICACIÓN

Hoy en día, la necesidad de utilizar técnicas de aprendizaje profundo para clasificar enfermedades de plantas de manera eficiente y precisa se debe a que los métodos tradicionales de diagnóstico de enfermedades de plantas son costosos, requieren mucho tiempo y a menudo no son precisos. El uso de la inteligencia artificial ha demostrado ser altamente efectivo en la clasificación de imágenes en otros campos, como la visión por computadora y el reconocimiento de voz. Por lo tanto, la investigación propuesta sobre la identificación de enfermedades en plantas de tomate durante la etapa de crecimiento vegetativo a través del análisis de hojas tiene un gran impacto en la agricultura y en la sociedad.

Esta investigación beneficiará a los agricultores, debido a que les proporcionará una herramienta para la detección temprana de enfermedades en las plantas y les permitirá tomar medidas preventivas o curativas antes de que se produzcan daños irreparables en los cultivos. Además, esta herramienta también beneficiará a los consumidores, ya que ayudará a garantizar la calidad y la seguridad de los productos alimenticios. Si se detectan enfermedades en las plantas durante la etapa de crecimiento vegetativo, se pueden tomar medidas para controlarlas, lo que reduce la necesidad de pesticidas y otros productos químicos para el control de enfermedades en etapas posteriores del ciclo de vida de las plantas (Hio, 2003).

La transformación que se producirá si los resultados de la investigación se transfieren al contexto real es significativa. La detección temprana de enfermedades en plantas mediante el análisis de hojas es un enfoque prometedor para mejorar la calidad y la cantidad de la producción de cultivos (Leal, 2021). Esta investigación puede transformar la forma en que se manejan los cultivos, ya que los agricultores tendrán una herramienta precisa y fácil de usar para la identificación temprana de enfermedades, lo que les permitirá tomar decisiones informadas y oportunas sobre el manejo de los cultivos.

En términos teóricos, la investigación propuesta se basa en la aplicación de algoritmos de inteligencia artificial para el análisis de imágenes de hojas de plantas de *Solanum lycopersicum*. La metodología combina técnicas de visión artificial y aprendizaje automático para crear un modelo predictivo que pueda detectar enfermedades en las hojas de las plantas (Romero, 2019).

Este proyecto puede tener el potencial de cambiar la forma en que se realizan las pruebas de diagnóstico en las plantas, lo que a su vez puede reducir el tiempo y el costo de los análisis de laboratorio tradicionales. Desde una perspectiva práctica, la herramienta desarrollada en esta investigación puede ayudar a mejorar la productividad de los agricultores y reducir los costos de producción de los cultivos. También puede mejorar la calidad y la seguridad de los productos alimenticios, lo que es fundamental en un mundo en el que la seguridad alimentaria es una preocupación creciente.

En cuanto a la relevancia del estudio para las políticas institucionales, la investigación propuesta se alinea con los objetivos de desarrollo sostenible de las Naciones Unidas, especialmente con el objetivo 2: Hambre Cero. El uso de técnicas de inteligencia artificial en la agricultura puede ayudar a mejorar la eficiencia y la sostenibilidad de la producción de alimentos, lo que es fundamental para lograr el objetivo de Hambre Cero (Moran, 2020).

En términos de relevancia para la carrera o profesión, la investigación propuesta es importante para los ingenieros agroindustriales y los profesionales en inteligencia artificial. La combinación de estas dos disciplinas puede mejorar significativamente la capacidad de los agricultores para detectar enfermedades en las plantas, lo que es fundamental para la productividad y la sostenibilidad de la agricultura. Para la Universidad, la investigación propuesta puede contribuir al avance del conocimiento en el campo de la agricultura sostenible y la aplicación de tecnologías de vanguardia para la identificación temprana de enfermedades en cultivos.

Por ejemplo, la implementación de la herramienta podría permitir la optimización de los recursos destinados a la vigilancia de enfermedades en cultivos, lo que podría tener un impacto positivo en la eficiencia y sostenibilidad de las prácticas agrícolas en la región. La inteligencia artificial se basa en la creación de modelos matemáticos a partir de grandes cantidades de datos, que permiten el reconocimiento de patrones y la identificación de posibles patógenos de manera rápida y precisa (Intel, s.f.).

La inteligencia artificial puede ser utilizada para identificar las especies y si presentan enfermedades a través de la creación de modelos matemáticos a partir

de grandes cantidades de datos. Los datos pueden ser recolectados de diversas fuentes, como fotografías de las hojas, imágenes de microscopio y pruebas de laboratorio. A través de técnicas de aprendizaje automático, la IA puede aprender a reconocer patrones y características de las especies, permitiendo una identificación más rápida y precisa (Zhou, 2022).

La IA también puede ser utilizada para predecir si presentan alguna enfermedad o no. En general, el uso de la IA puede mejorar significativamente la eficacia y eficiencia de la identificación de especies en las hojas de cultivos, lo que puede ayudar a prevenir la propagación de enfermedades y mejorar la calidad y rentabilidad de los cultivos, así como mejorar su productividad. Por último, es importante mencionar que la investigación propuesta también puede tener implicaciones más amplias para la sociedad en general, en términos de la seguridad alimentaria y la producción agrícola sostenible.

La identificación temprana de enfermedades en cultivos puede ayudar a los agricultores a tomar medidas preventivas para reducir la propagación de enfermedades y mejorar la calidad de sus cultivos, lo que puede tener un impacto positivo en la seguridad alimentaria y en la economía de la región y del país en general. La investigación propuesta tiene una alta relevancia y beneficios potenciales para la región Andina de Colombia. Esta zona tiene una larga tradición en el cultivo de tomate, siendo uno de los principales productores a nivel nacional como se muestra en la siguiente figura:

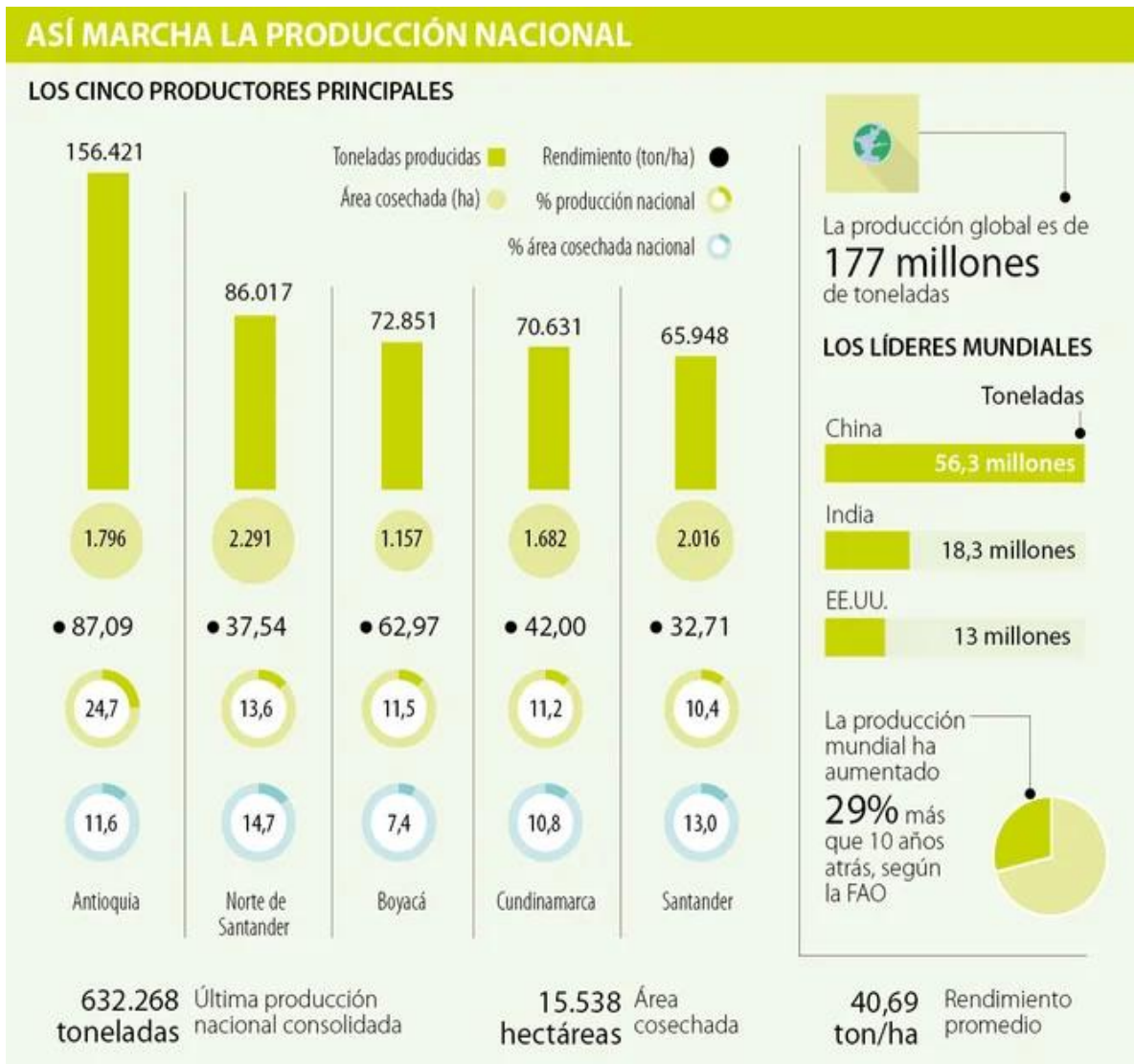


Figura 1: Antioquia y norte de Santander son los departamentos líderes en la producción de tomate. *Obtenido de: Agronegocios (2018)*

La implementación de una herramienta de detección temprana de enfermedades basada en inteligencia artificial podría tener un impacto transformador en la productividad y sostenibilidad de este cultivo en la región. Los productores andinos se beneficiarían enormemente al poder identificar y controlar plagas y enfermedades de manera rápida y efectiva. Esto reduciría pérdidas económicas por daños en la producción, mejoraría la calidad y sanidad de los cultivos, e incrementaría la competitividad de los agricultores de la zona.

3. OBJETIVOS

3.1 OBJETIVO GENERAL

Desarrollar una herramienta basada en algoritmos de inteligencia artificial para identificar enfermedades en la etapa de crecimiento vegetativo de las hojas en el cultivo de *Solanum lycopersicum* (tomate) para mejorar la productividad, la sostenibilidad y la eficiencia general de la agricultura, al tiempo que se reduce el impacto ambiental.

3.2 OBJETIVOS ESPECÍFICOS

- Identificar los requerimientos necesarios para el desarrollo de la herramienta de inteligencia artificial, considerando aspectos como la obtención y preparación de datos relevantes para el entrenamiento del modelo, así como el preprocesamiento de imágenes mediante técnicas de segmentación, filtrado y extracción de características relevantes.
- Diseñar la arquitectura, el flujo de trabajo de la herramienta, la selección de los algoritmos de inteligencia artificial más adecuados para la tarea de detección, entre otros aspectos.
- Implementar la codificación de los algoritmos de inteligencia artificial y los recursos de hardware y software necesarios para el desarrollo de la herramienta.
- Evaluar la precisión de la herramienta mediante la ejecución de pruebas con nuevos datos para determinar si se cumple el objetivo del proyecto.

4. MARCO REFERENCIAL

4.1 MARCO TEÓRICO

Es importante destacar que la inteligencia artificial, en términos generales, se refiere a la metodología que utilizan las máquinas o computadoras para simular de manera precisa o similar las decisiones tomadas por un operador humano. Esto se logra mediante el análisis y procesamiento de datos proporcionados al sistema por el creador de la inteligencia artificial (Intel, s.f.).

La inteligencia artificial es una ciencia compleja y extensa en el campo de la ingeniería. Se presentan diferentes tipos de clasificaciones en la inteligencia artificial, que ofrecen resultados variados según los datos a los que se someten. Estos resultados pueden ser mejores o peores dependiendo de la complejidad de los algoritmos utilizados y del tipo de datos que se procesan.

Para continuar, estas son las tres clasificaciones principales en las cuales está dividida la inteligencia artificial:

- IA basada en capacidades humanas:
 1. IA débil: También conocida como Artificial Narrow Intelligence (ANI), se refiere a sistemas de IA diseñados para realizar tareas específicas y limitadas. Estos sistemas no poseen una comprensión o conciencia general del mundo y se centran en tareas específicas (Ferreira, 2021).
 2. IA fuerte: También conocida como IA general (AGI), se refiere a sistemas de IA que poseen un nivel de inteligencia similar o superior al humano en múltiples áreas y pueden comprender y realizar una amplia gama de tareas (Ferreira, 2021).
- Aprendizaje automático (Machine Learning, ML): Es una rama de la IA que se enfoca en el desarrollo de algoritmos y modelos que permiten a las máquinas aprender a partir de datos y realizar tareas sin ser explícitamente programadas. Los modelos de ML pueden hacer predicciones o tomar decisiones basadas en patrones identificados en los datos (Zhou, 2022).
- Aprendizaje profundo (Deep Learning, DL): Es una subcategoría del aprendizaje automático que utiliza redes neuronales artificiales profundas para modelar y comprender datos complejos. El DL se inspira en la estructura

y función del cerebro humano, y se ha utilizado con éxito en tareas como el reconocimiento de imágenes y voz, el procesamiento del lenguaje natural y la conducción autónoma (Goodfellow, Bengio, & Courville, 2016).

Además de los diferentes tipos de inteligencia artificial, también existen varios enfoques de desarrollo utilizados en la creación de sistemas de inteligencia artificial. Estos enfoques, popularmente conocidos como "clasificadores", se encargan de clasificar los datos suministrados de diversas formas para obtener resultados específicos.

Los clasificadores desempeñan un papel fundamental en la inteligencia artificial y el aprendizaje automático. Son herramientas vitales para asignar etiquetas a los datos, automatizar la toma de decisiones, analizar patrones y mejorar la precisión. Por otro lado, permiten organizar y comprender los datos, así como identificar tendencias y relaciones entre variables. Los clasificadores tienen diversas aplicaciones prácticas, como reconocimiento de patrones, diagnósticos médicos, detección de fraudes y recomendaciones de productos. Su importancia radica en su capacidad para procesar y extraer conocimiento de los datos, facilitando la toma de decisiones informadas y el desarrollo de sistemas inteligentes (Gavilán, 2017).

Los Sistemas Expertos se basan en la captura del conocimiento y la experiencia de expertos humanos en un dominio específico para crear sistemas capaces de realizar tareas similares a las que realizaría un experto humano. Estos sistemas utilizan reglas lógicas y algoritmos de inferencia para procesar la información y ofrecer soluciones o recomendaciones.

En el caso de la economía agraria, los Sistemas Expertos pueden ser de gran utilidad para abordar diversas problemáticas relacionadas con la producción agrícola, el manejo de plagas y enfermedades, la optimización de recursos, la toma de decisiones con base a datos y variables ambientales, entre otros aspectos. Al aprovechar el conocimiento experto y aplicar técnicas de IA, estos sistemas pueden ofrecer soluciones eficientes y precisas, contribuyendo así al desarrollo y mejora de la actividad agraria (Vecino, N/N).

La inteligencia artificial (IA) ha revolucionado la industria agrícola al ofrecer soluciones innovadoras y eficientes para mejorar la productividad, la sostenibilidad y la rentabilidad de los cultivos. La aplicación de la IA en el sector agrícola ha abarcado diversas áreas, desde la detección y diagnóstico de enfermedades en los cultivos hasta la optimización del uso de recursos, el monitoreo de plagas y malezas, el pronóstico de cosechas y la agricultura de precisión. Estas aplicaciones han permitido a los agricultores tomar decisiones informadas y basadas en datos, optimizando sus prácticas agrícolas y maximizando los rendimientos.

En este contexto, una de las prácticas clave en la agricultura es la detección temprana y precisa de enfermedades que afectan a los cultivos, como es el caso del cultivo de tomate, conocido comúnmente como tomate. Enfermedades como el mildiú vellosa, la mancha foliar y la marchitez bacteriana pueden tener un impacto significativo en la calidad y el rendimiento de los tomates durante su etapa de crecimiento vegetativo.

El tomate es una fruta/hortaliza que se cultiva en todo el mundo por su valor nutricional y su uso en la cocina. pertenece a la familia de las Solanáceas. El tomate es una planta perenne anual de porte arbustivo que se desarrolla de forma rastrera semirrecta o erecta. Según el hábito de crecimiento, las variedades se dividen en determinadas e indeterminadas. El fruto del tomate es una fuente importante de vitaminas A, B1, B2, B6, C y E, y de minerales como fósforo, potasio, magnesio, manganeso, zinc, cobre, sodio, hierro y calcio. Además, tiene un importante valor nutricional ya que incluye proteínas, hidratos de carbono, fibra, ácido fólico, ácido tartárico, ácido succínico y ácido salicílico. El tomate se utiliza en una amplia variedad de platos, desde ensaladas hasta salsas y guisos, y es un ingrediente esencial en la cocina mediterránea.

4.2 MARCO CONCEPTUAL

Dentro de la inteligencia artificial, existen diferentes tipos de clasificadores que son ampliamente utilizados para la clasificación y predicción de datos. Estos clasificadores incluyen algoritmos como el Naive Bayes, K-Nearest Neighbors (KNN), Support Vector Machines (SVM), Redes Neuronales Convolucionales (CNN) y ResNet-50. Cada uno de estos clasificadores tiene sus propias características y

se basa en diferentes principios teóricos y metodologías para realizar la clasificación de los datos.

En el campo del aprendizaje automático y la inteligencia artificial, existen diversos tipos de clasificadores que desempeñan un papel fundamental en la organización y comprensión de datos. Estos clasificadores son herramientas esenciales para asignar etiquetas a los datos basándose en sus características y patrones. Los principales clasificadores se pueden visualizar en el anexo 9.2. Cada uno de estos clasificadores tiene sus propias ventajas y aplicaciones específicas, y su elección depende del problema y los datos particulares. Explorar estos clasificadores proporciona una visión más profunda de las técnicas disponibles para el análisis y la clasificación de datos en el campo de la inteligencia artificial y el aprendizaje automático.

Para la identificación de las enfermedades del tomate, es necesario establecer una selección específica de las enfermedades a diagnosticar. En este proyecto, se han identificado y se estudiarán ocho tipos de estados de salud de la planta del tomate. Estos estados abarcan una amplia variedad de enfermedades comunes que afectan al tomate, permitiendo un análisis exhaustivo y preciso de su detección. Mediante la identificación y estudio de estas enfermedades, se busca desarrollar una herramienta efectiva y confiable para el diagnóstico temprano y la implementación de soluciones adecuadas en el manejo de cultivos de tomate. Adicionalmente, se puede visualizar una tabla detallada que enumera las nueve enfermedades del tomate evaluadas en los anexos 9.3. Esta tabla proporcionará información clave sobre cada enfermedad, incluyendo su nombre científico, síntomas característicos, así como una imagen de muestra.

4.2.1 *CLASIFICADORES*

Las técnicas de clasificación propuestas son Naive Bayes, K-nearest neighbors, Support Vector Machine, Convolutional Neural Network y ResNet-50. Las redes neuronales son muy útiles para, a partir de un dato de entrada, hacer una clasificación o regresión a partir de un modelo entrenado. Esta misma idea se puede aplicar a una imagen, para obtener, por ejemplo, qué representa la imagen (clasificación) o detectar un objeto que forme parte de ella (segmentación, detección, identificación, etc) (Desarrollosidn, 2022).

Naive Bayes (NB): Este algoritmo es ampliamente utilizado en problemas de clasificación, especialmente cuando se trabaja con conjuntos de datos grandes. Es eficiente y rápido en términos de tiempo de entrenamiento y predicción. Además, asume independencia condicional entre las características, lo cual es útil cuando se tienen múltiples características que pueden influir en la detección de enfermedades foliares en plantas de tomate.

Utiliza el teorema de Bayes para calcular la probabilidad condicional de que un objeto pertenezca a una clase dada en función de sus características. Matemáticamente, se usa la fórmula de Bayes:

$$P(\text{Clase}|\text{Características}) = \frac{P(\text{Características}|\text{Clase}) * P(\text{Clase})}{P(\text{Características})} \quad (1)$$

Donde $P(\text{Clase}|\text{Características})$ es la probabilidad de que el objeto pertenezca a una clase dada dado un conjunto de características, $P(\text{Características}|\text{Clase})$ es la probabilidad de observar las características dadas la clase, $P(\text{Clase})$ es la probabilidad a priori de la clase y $P(\text{Características})$ es la probabilidad de observar las características.

K-nearest neighbors (KNN): Es un método de aprendizaje automático basado en la similitud de las instancias. En este caso, se clasificarán las plantas de tomate según su proximidad a otras plantas ya etiquetadas con enfermedades foliares conocidas. KNN es adecuado cuando la estructura de los datos es importante y la clasificación se basa en la similitud de características.

Para clasificar un punto, se calcula la distancia entre ese punto y todos los puntos de entrenamiento. Luego, se seleccionan los k puntos más cercanos. La clase más común entre estos k vecinos se asigna al punto de prueba. La métrica de distancia puede variar, pero la distancia euclidiana es comúnmente utilizada.

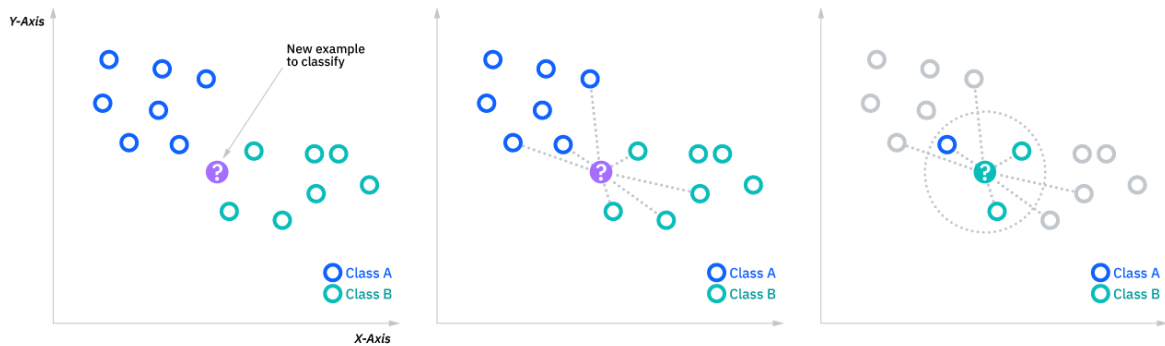


Figura 2: ¿Qué es el algoritmo de k vecinos más cercanos? *Obtenido de:* IBM (s.f.)

La fórmula de la distancia euclidiana entre dos puntos en un espacio n-dimensional se expresa de la siguiente manera. Si tienes dos puntos, A y B, en un espacio n-dimensional, donde las coordenadas de A son (a_1, a_2, \dots, a_n) y las coordenadas de B son (b_1, b_2, \dots, b_n) , entonces la distancia euclidiana (d) entre A y B se calcula de la siguiente manera:

$$d(A, B) = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + \dots + (a_n - b_n)^2} \quad (2)$$

Esta fórmula representa la distancia euclidiana, es decir, la distancia más corta entre dos puntos en un espacio n-dimensional. La distancia euclidiana se utiliza comúnmente en algoritmos de aprendizaje automático, como K-Nearest Neighbors (K-NN) y algoritmos de clustering, así como en diversas aplicaciones de análisis de datos y geometría.

Support Vector Machine (SVM): Este algoritmo es adecuado para problemas de clasificación en espacios de alta dimensión. SVM busca encontrar un hiperplano óptimo que separe las muestras de diferentes clases. En este caso, se puede utilizar para separar las plantas de tomate afectadas por enfermedades foliares de las plantas sanas. SVM tiene una buena capacidad de generalización y es efectivo incluso cuando los datos son no linealmente separables. Matemáticamente, SVM busca la distancia entre el hiperplano de separación y los puntos de datos más cercanos (vectores de soporte). Utiliza funciones kernel para transformar los datos

en un espacio de características de mayor dimensión si es necesario para lograr la separación óptima. SVM busca encontrar el hiperplano de separación óptimo. La función objetivo se define como:

$$\text{Minimizar} = \frac{1}{2} * ||w||^2 \quad (3)$$

Donde "w" es un vector de pesos que define el hiperplano de separación. Las restricciones de SVM están relacionadas con el margen (distancia entre el hiperplano y los puntos más cercanos de cada clase) y la clasificación correcta de los datos de entrenamiento. Estas restricciones se expresan matemáticamente como:

$$y_i * ((w * x_i) + b) \geq 1 \quad (4)$$

Donde " y_i " es la etiqueta de clase del punto de entrenamiento " x_i " y " b " es un término de sesgo. El objetivo es encontrar " w " y " b " que minimicen la función objetivo mientras se satisfacen las restricciones anteriores.

Convolutional Neural Network (CNN): Esta técnica es especialmente útil cuando se trabajan datos de imágenes, como las hojas de las plantas de tomate. CNN puede aprender y extraer características importantes de las imágenes de manera automática. Además, puede capturar patrones y características complejas en las imágenes para una clasificación precisa. Las capas de convolución aplican filtros a regiones locales de la imagen de entrada y producen mapas de características. Luego, se utilizan capas de agrupación (pooling) para reducir la dimensionalidad. Finalmente, capas totalmente conectadas se utilizan para la clasificación. Las operaciones de convolución y agrupación están definidas matemáticamente y se aplican de manera secuencial a través de la red.

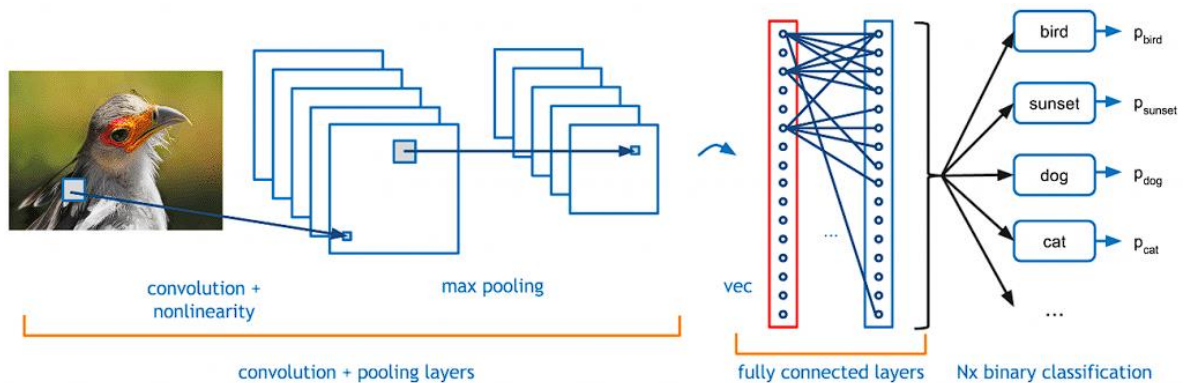


Figura 3. Estructura de una red neuronal convolucional para la clasificación de una imagen. *Obtenido de:* Desarrollosidn (2022).

Las ecuaciones en una CNN son inherentemente más complicadas y dependen de la arquitectura específica de la red y las capas utilizadas. Sin embargo, se muestra una descripción general simplificada de cómo funcionan las capas de convolución en una CNN. En la capa de convolución, se aplica una operación de convolución para obtener mapas de características. Dada una imagen de entrada "I" y un filtro (kernel) "K", la operación de convolución se expresa como:

$$(I * K)(x, y) = \sum I(x + i, y + j) * K(i, j) \quad (5)$$

Donde "x" y "y" son las coordenadas espaciales, "i" y "j" son las coordenadas del filtro y "Σ" denota la suma sobre todas las coordenadas. Después de la convolución, se aplica una función de activación, como ReLU (Rectified Linear Unit), que introduce no linealidad a la red:

$$ReLU(z) = \max(0, z) \quad (6)$$

Donde "z" es la salida de la convolución.

ResNet-50: es una arquitectura de red neuronal profunda que se utiliza comúnmente en tareas de clasificación de imágenes. Se caracteriza por la introducción de conexiones residuales que ayudan a mitigar el problema de desvanecimiento del gradiente en redes profundas. Matemáticamente, estas conexiones residuales se expresan como funciones de identidad que agregan la salida de una capa a la salida de una capa anterior.

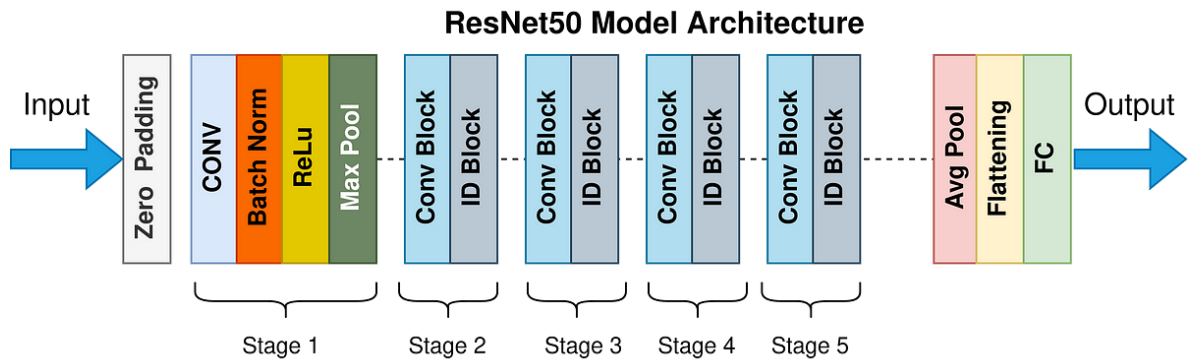


Figura 4. Arquitectura del modelo ResNet50. Obtenido de: Mukherjee (2022).

Esto permite que los gradientes fluyan más fácilmente durante el entrenamiento y permite la construcción de redes muy profundas sin degradación del rendimiento. Las conexiones residuales se expresan como funciones de identidad y se suman a las salidas de las capas anteriores. La fórmula simplificada para una conexión residual es:

$$F(x) = x + Convolution(x) \quad (7)$$

Donde " x " es la entrada original, " $Convolution(x)$ " es la salida de una capa de convolución y " $F(x)$ " es la salida final de la capa. La idea clave es que esta estructura permite que los gradientes fluyan más fácilmente durante el entrenamiento de la red profunda.

Es importante destacar que estas son simplificaciones de las ecuaciones y fórmulas reales utilizadas en estas técnicas. Las implementaciones reales de SVM, CNN y ResNet-50 involucran muchas más capas y detalles matemáticos. En resumen, cada una de estas técnicas tiene sus propias ventajas y capacidades, y su elección

depende de la naturaleza del problema y los datos disponibles. En este estudio, utilizar una combinación de estas técnicas permitirá evaluar el rendimiento de los clasificadores en la detección de enfermedades foliares en cultivos de tomate, brindando una evaluación exhaustiva y comparativa de su desempeño.

4.2.2 METRICAS

Las métricas que se proponen para evaluar los resultados de los clasificadores son Accuracy, Unweighted Average Recall, Recall, Matriz de confusión y F1 Score. Estas métricas son buenas para evaluar los clasificadores propuestos en el estudio debido a las siguientes razones:

Accuracy (Precisión): La precisión general del clasificador es una métrica fundamental para evaluar su desempeño global. En este caso, indica la proporción de instancias clasificadas correctamente en relación con el total de instancias evaluadas. Es importante tener una alta precisión para asegurar que las plantas de tomate afectadas por enfermedades foliares sean identificadas correctamente y se tomen las medidas necesarias.

Unweighted Average Recall (Tasa de clasificación correcta ponderada): Esta métrica calcula el promedio de las tasas de clasificación correcta para cada etiqueta o clase. Proporciona una visión general del rendimiento global del clasificador y toma en cuenta el desempeño en todas las clases. Es útil para evaluar si el clasificador tiene un buen desempeño en la detección de todas las enfermedades foliares en los cultivos de tomate.

Recall (Recuperación): La capacidad del clasificador para detectar correctamente instancias de una clase específica es esencial en la detección de enfermedades foliares. Recall o tasa de recuperación es la proporción de instancias positivas correctamente clasificadas en relación con el total de instancias positivas. Es importante tener un alto recall para asegurar que se detecten y clasifiquen correctamente las plantas de tomate afectadas por cada enfermedad específica.

Matriz de confusión: Esta matriz muestra el número de instancias clasificadas correcta e incorrectamente para cada etiqueta o clase. Permite identificar errores de clasificación y evaluar el desempeño del clasificador en cada clase. Es útil para comprender en qué clases se cometen más errores y cómo se distribuyen los errores de clasificación en general.

4.3 ESTADO DEL ARTE / MARCO HISTORICO

La disciplina de Inteligencia Artificial (IA) surgió en 1956 durante la Conferencia de Darmouth con el propósito de investigar actividades humanas que carecían de métodos bien definidos para describir cómo se llevaban a cabo. Esta necesidad surgía debido a la falta de algoritmos capaces de describir actividades cognitivas aparentemente simples, como el reconocimiento visual de objetos, la comprensión del lenguaje hablado y escrito, el diagnóstico de enfermedades, entre otros.

A partir de finales de la década de 1960, los esfuerzos en el campo de la IA se orientaron hacia el estudio de sistemas inteligentes para la robótica, como la visión artificial y la navegación sin piloto. Esta evolución llevó a la necesidad de incorporar una gran cantidad de conocimientos específicos a los sistemas, especialmente en relación con los problemas que se buscaban resolver. Es en este contexto donde surge el estudio de los Sistemas Expertos (SE), una rama de la IA que puede resultar de particular interés para la economía agraria.

El uso de inteligencia artificial, específicamente redes neuronales convolucionales, ha mostrado resultados prometedores en la detección y diagnóstico de enfermedades de plantas en el sector agrícola. Sin embargo, el costo del entrenamiento computacional y la necesidad de grandes cantidades de datos siguen siendo un desafío. Otras técnicas, como los árboles de decisión con lógica difusa, también pueden ser efectivas en la detección de enfermedades. La implementación de estas tecnologías puede llevar a un diagnóstico oportuno y a una reducción de las pérdidas económicas para los agricultores. En este marco de la agricultura de precisión podemos clasificar las aplicaciones de IA que han tenido un gran impacto en cuatro grandes bloques: las enfermedades y plagas, la calidad del grano, la fenología y el fenotipado (Patrício & Rieder, 2018).

Esta se ve afectada por la aparición de enfermedades o infecciones provocadas por insectos y plagas. Lo que repercute de manera directa produciendo aspectos negativos en el desarrollo del cultivo y en la calidad del grano, provocando una variación en el precio de mercado. Por tanto, esta identificación, gestión y detección de enfermedades de forma efectiva garantiza una agricultura productiva y sostenible (Fouquet, 2021).

La implementación de la inteligencia artificial (IA) en el ámbito agrícola, enfocada principalmente en el análisis del suelo y la gestión de cultivos, resalta una serie de ventajas significativas. Entre ellas se incluyen el notable aumento de la productividad y eficiencia, así como la capacidad de prever los momentos óptimos para la siembra y cosecha. También se discuten los desafíos que enfrenta el sector agrícola, como el tratamiento inadecuado del suelo y la infestación de enfermedades, y cómo la IA puede ayudar a abordar estos problemas. El documento concluye con una revisión de las fortalezas y limitaciones de la IA en la agricultura y el uso de sistemas expertos para aumentar la productividad (Bonilla, 2021).

Según un estudio de Ortega, Biswal y Sanchez (2019), se expone varios trabajos de investigación sobre la posibilidad de detectar enfermedades en los cultivos agrícolas, utilizando técnicas de inteligencia artificial. Los autores de dichos trabajos hacen propuestas de técnicas de obtención de características de las hojas o frutos de las plantas, así como también el uso de algoritmos clasificadores o de agrupación, todo esto con el fin de determinar si una hoja, presenta signos de alguna enfermedad. Al haber diversos tipos de enfermedades y diversas variedades de plantas, los autores hacen propuestas para utilizar el algoritmo que ellos consideran el que obtendrá mejores resultados.

Los autores proponen varias técnicas para obtener características de hojas o frutas de plantas y utilizar algoritmos de clasificación o agrupamiento para determinar si una hoja muestra signos de enfermedad. El artículo revisa varios trabajos significativos relacionados con la detección de enfermedades en varios tipos de plantas utilizando diferentes algoritmos de aprendizaje automático. Los autores concluyen que el diagnóstico oportuno de enfermedades puede beneficiar a los agricultores al reducir el riesgo de pérdidas económicas.

El documento discute varios estudios sobre el uso de inteligencia artificial para la detección de enfermedades en cultivos. Se utilizaron diferentes métodos como lógica difusa, SVM, KNN y CNN para la clasificación, siendo SVM y CNN los que muestran los mejores resultados. Los estudios muestran que la IA puede ser una herramienta útil para la detección de enfermedades en cultivos, pero se necesita más investigación para mejorar la precisión y aplicabilidad de estos métodos.

Un estudio realizado por Aysun et al (2020), utilizaron técnicas de procesamiento de imágenes y aprendizaje automático para detectar enfermedades en frutas. Se creó un conjunto de datos de 22 manzanas y 18 membrillos, y las imágenes se clasificaron utilizando redes neuronales convolucionales (CNN). Se logró una precisión del 83,3% con la arquitectura propuesta de CNN, mientras que el modelo pre-entrenado AlexNet logró una precisión del 81,3%.

El estudio destaca la importancia de la detección temprana de enfermedades en frutas para reducir las pérdidas económicas en la industria alimentaria. Por otro lado, el estudio discute el uso de inteligencia artificial y aprendizaje profundo en la identificación y clasificación de enfermedades en frutas. Además, se menciona que el conjunto de datos se ampliará en el futuro para incluir otras enfermedades en frutas y se creará un sistema de identificación basado en video.

En otro estudio realizado para la detección de enfermedades con las hojas del tomate usando inteligencia artificial, los autores concluyeron que el uso de redes neuronales convolucionales y aprendizaje profundo es una técnica efectiva para detectar enfermedades en hojas de tomate utilizando cámaras de teléfonos inteligentes. El modelo entrenado con un conjunto de datos de 9000 imágenes de hojas de tomate infectadas y saludables logró una precisión del 99,84% en un conjunto de prueba. Además, los autores encontraron que el modelo que utilizó imágenes a color tuvo un mejor rendimiento que el modelo que utilizó imágenes en escala de grises.

Los autores también visualizaron las activaciones intermedias de los modelos para obtener información sobre cómo identifican las manchas de enfermedades. En general, el estudio sugiere que el uso de CNN y aprendizaje profundo puede ser

una herramienta útil para la detección temprana de enfermedades en plantas como el tomate (Ashqar, 2018).

Para continuar, un proyecto realizado por Gaspar y Escobar (2018) proponen un sistema de monitoreo basado en propiedades eléctricas y biológicas para controlar la recolección de feijoas. Utiliza el método KNN para establecer relaciones entre las variables eléctricas y las características fisicoquímicas para mejorar la calidad y cantidad de la fruta recolectada en la Sabana de Bogotá. Se desarrollo un modelamiento matemático y un sistema de diagnóstico basado en estas variables.

El método KNN, siendo este el más óptimo y usando los 5 vecinos más cercanos, se encontraron porcentajes de acierto en un 89% para frutos verdes y 100% para frutos maduros, para un total de 94% de acierto total. En el caso de la RNA con la información más concreta empleando la aplicación Pattern Recognition con 14 neuronas en su única capa oculta se encontraron porcentajes de acierto de 95.1% para frutos verdes y 88,6% de frutos maduros, para un total de 92.1 % de acierto total.

Revisando un contexto local, en la Fundación Universitaria Agraria de Colombia se realizó un proyecto de grado con los lineamientos de la inteligencia artificial para desarrollar un sistema de monitoreo, el proyecto de grado se centra en el desarrollo de un sistema para el monitoreo y control cuantitativo del sector ganadero en Colombia. Utilizando tecnologías de código abierto e integración de inteligencia artificial, se busca caracterizar e identificar los hatos bovinos en zonas de pastoreo, mejorando las técnicas tradicionales y generando informes automáticos con vistas aéreas del estado actual del hato en términos del volumen de cabezas presentes.

El sistema busca combatir el abigeato, un delito común en Colombia, proporcionando reportes eficientes y precisos para que los agricultores puedan tomar medidas oportunas. Las pruebas experimentales se realizaron en entornos controlados y no controlados para evaluar la viabilidad y robustez del sistema. Los resultados de la eficacia en el monitoreo y control cuantitativo del sector ganadero en Colombia permiten caracterizar e identificar los hatos bovinos en zonas de pastoreo, mejorando las técnicas tradicionales y generando informes automáticos del estado del hato. Además, se ha destacado la importancia del procesamiento de

imágenes, la evaluación de diferentes metodologías y la implementación de un sistema automático.

Las pruebas experimentales han demostrado la viabilidad y robustez del sistema en entornos controlados y no controlados. Como resultado, se han obtenido importantes aprendizajes y se han identificado recomendaciones para futuros trabajos, en aras de mejorar la implementación del sistema en diferentes zonas de pastoreo (Alexander, 2022).

Según un artículo publicado por Vitida, Jignesh y Chetan (2013). Propone un enfoque integral para el manejo de plagas en la agricultura, conocido como Manejo Integrado de Plagas (MIP). El MIP busca minimizar las pérdidas en la producción agrícola debido a plagas y enfermedades utilizando una combinación de prácticas que optimizan el uso de pesticidas. Una de las estrategias clave en el MIP es la detección temprana de plagas y su control. El enfoque propuesto en este artículo se centra en utilizar datos meteorológicos para prever el desarrollo de enfermedades de las plantas.

Estos datos meteorológicos incluyen temperatura, humedad y duración de la humedad en las hojas (LWD, por sus siglas en inglés). Estos factores climáticos desempeñan un papel crucial en el crecimiento de microorganismos responsables de enfermedades en las plantas. El artículo sugiere que, al prever las enfermedades de las plantas en función de los datos climáticos, los agricultores pueden tomar acciones preventivas a tiempo para controlar las enfermedades. Esto también puede ayudar a reducir la cantidad de pesticidas utilizados, lo que a su vez puede mitigar la contaminación del suelo.

El sistema utiliza lógica difusa, un enfoque matemático que permite manejar la incertidumbre y la imprecisión en los datos, para desarrollar un sistema de pronóstico de enfermedades de las plantas. El objetivo es construir un Sistema de Soporte a la Decisión Agrícola (ADSS, por sus siglas en inglés) basado en conocimientos, que ayude a los agricultores a tomar decisiones informadas sobre la gestión de plagas y enfermedades en sus cultivos.

Otro artículo llevado a cabo por Boyd y Sun (1994), describe el desarrollo de un sistema experto (ES) prototipo diseñado para gestionar la producción de semillas de papas. El prototipo se enfoca en la emulación de la experiencia humana en el diagnóstico de enfermedades de la papa basado en diversos síntomas. El sistema experto fue desarrollado utilizando un intérprete de comandos llamado PC Expert Professional de Software Artistry, Inc., y es operable en computadoras personales compatibles. Contiene 127 reglas para diagnosticar once enfermedades patógenas y seis enfermedades no patógenas de las papas, utilizando ocho bases de conocimientos (KB).

El proceso de adquisición de conocimientos implicó la participación de dos ingenieros de conocimiento y cinco expertos en el dominio del diagnóstico de enfermedades de la papa. El desarrollo pasó por seis iteraciones y se realizó a través de actividades como entrevistas de adquisición de conocimientos, construcción de un demo del sistema experto para mostrar a los productores de papas, actualización del demo a un prototipo y revisiones periódicas de las bases de conocimientos por parte de los expertos.

Por otro lado, el artículo de Mohammed, Kamel y Abdelouahab (2017), trata sobre el uso del aprendizaje profundo (deep learning) para la clasificación de enfermedades del tomate y la visualización de síntomas. Los autores utilizaron 14828 imágenes de entrenamiento con nueve enfermedades diferentes y presentaron la Red Neuronal Convolutiva (CNN) como algoritmo de aprendizaje. Los resultados mostraron una precisión del 99,18%, superando a los modelos superficiales.

Los autores también propusieron un método para la detección y visualización de síntomas para ayudar a los agricultores a proteger sus cultivos. El estudio destaca las ventajas del aprendizaje profundo en la automatización del proceso de clasificación y la extracción de características directamente de los datos crudos. Por otro lado, el estudio compara el rendimiento de los modelos profundos con características hechas a mano con los modelos superficiales en la clasificación de enfermedades de las plantas. Se utilizaron dos modelos profundos famosos, AlexNet y GoogleNet, y se compararon los resultados con los definidos por expertos. El estudio también evalúa el rendimiento de los modelos pre-entrenados y utiliza

métodos de visualización para extraer conocimiento biológico de los modelos entrenados de manera completamente automática.

El documento muestra que el aprendizaje profundo es una herramienta efectiva para la clasificación de enfermedades de las plantas y la visualización de síntomas, lo que puede ayudar a los agricultores a proteger sus cultivos de manera más eficiente. Se puede encontrar una tabla en el anexo 9.1 donde se resume la información relevante de los artículos mencionados anteriormente y a continuación se mencionan las investigaciones en orden cronológico:

1983 - 2017: Durante este período, se lleva a cabo una serie de investigaciones en el campo de la agricultura y la inteligencia artificial. Las técnicas de IA utilizadas incluyen sistemas expertos, redes neuronales artificiales y sistemas difusos. Se desarrolla una base sólida para aplicar la IA en la agricultura.

2017: Mohammed, Kamel y Abdelouahab (2017) presentan un estudio que utiliza aprendizaje profundo (deep learning) para la clasificación de enfermedades en hojas de tomate. Introducen una Red Neuronal Convolucional (CNN) y obtienen una precisión del 99,18% en la clasificación de enfermedades de las plantas. También proponen un método para la detección y visualización de síntomas.

2018: el artículo "Artificial Intelligence in Agriculture: A Literature Survey", que revisa 100 contribuciones importantes relacionadas con el uso de IA en la agricultura. El documento aborda las aplicaciones de la IA en diversos aspectos de la agricultura y destaca la evolución de las técnicas de IA a lo largo de los años (Sarkar, et al 2018). Por otro lado, Gaspar y Escobar (2018) proponen un sistema de monitoreo basado en propiedades eléctricas y biológicas para la recolección de feijos. Utilizan algoritmos como K-Nearest Neighbors (KNN) y Redes Neuronales Artificiales (RNA) para lograr altos porcentajes de acierto en la clasificación de frutos verdes y maduros.

2019: Ortega, Biswal y Sanchez (2019) presentan un estudio que revisa varios trabajos relacionados con la detección de enfermedades en diferentes tipos de plantas utilizando algoritmos de aprendizaje automático. Destacan la importancia del diagnóstico temprano de enfermedades en la agricultura para reducir las pérdidas económicas.

2020: Aysun et al (2020) utilizan técnicas de procesamiento de imágenes y aprendizaje automático para detectar enfermedades en frutas como manzanas y

membrillos. Logran una precisión del 83,3% en la clasificación de enfermedades utilizando redes neuronales convolucionales (CNN).

2021: En un estudio sobre la detección de enfermedades en hojas de tomate, se concluye que las redes neuronales convolucionales y el aprendizaje profundo son efectivos para esta tarea. Se logra una precisión del 99,84% en un conjunto de prueba.

2021: En otro estudio sobre la agricultura de precisión, se destaca la importancia de la identificación y detección de enfermedades en cultivos de tomate de árbol para garantizar una agricultura productiva y sostenible.

2022: En la Fundación Universitaria Agraria de Colombia, se desarrolla un sistema de monitoreo y control cuantitativo del sector ganadero en Colombia utilizando inteligencia artificial y tecnologías de código abierto. El sistema busca combatir el abigeato y proporcionar informes eficientes y precisos a los agricultores.

4.4 MARCO LEGAL

Para el presente proyecto, que utiliza una IA para detectar enfermedades en los cultivos en crecimiento vegetativo en las hojas de *solanum lycopersicum*, hay varios aspectos legales que se deben tener en cuenta para el desarrollo del proyecto en Colombia. A continuación, se presentan algunos aspectos relevantes:

4.5.1 Protección de datos personales: En Colombia, la normativa vigente es la Ley Estatutaria 1581 de 2012 y su decreto reglamentario 1377 de 2013. Hay que asegurarse de obtener el consentimiento explícito y previo de los usuarios para recopilar, almacenar y procesar sus datos personales, explicando claramente cómo se utilizarán y protegerán.

4.5.2 Responsabilidad legal: Aunque el proyecto se desarrolle con fines académicos o de investigación, es importante considerar la responsabilidad legal en caso de que la IA brinde resultados inexactos o se produzcan consecuencias adversas para los usuarios. Se debe incluir cláusulas de exención de responsabilidad y limitaciones claras en el uso de la IA, dejando claro que no se brinda asesoramiento médico ni garantía de precisión en los resultados.

5. DISEÑO METODOLÓGICO

5.1 TIPO DE INVESTIGACIÓN

El presente estudio se enmarca en una investigación experimental de tipo cuantitativo. El objetivo principal es evaluar el rendimiento de clasificadores en la detección de enfermedades foliares en cultivos de tomate en su etapa de crecimiento vegetativo. Para ello, se realizarán experimentos utilizando datos recopilados de muestras de plantas de tomate afectadas por diversas enfermedades, así como de plantas sanas.

Se aplicarán métodos estadísticos para analizar los resultados y se obtendrán datos cuantitativos que permitirán comparar el desempeño de los clasificadores. Adicionalmente, el estudio sigue los lineamientos institucionales y de investigación académica debido a que busca utilizar la inteligencia artificial para identificar la especie del cultivo y si padece alguna enfermedad. Un modelo de inteligencia artificial bien entrenado para identificar las posibles enfermedades foliares puede ayudar a la conservación del cultivo de tomate, tomar medidas preventivas y de control antes de que se propaguen entre el cultivo. La solución propuesta permitirá identificar con mayor rapidez y precisión, mejorando así la productividad en el cultivo.

5.2 METODOLOGIA

Para obtener una evaluación precisa del desarrollo de la inteligencia artificial, se requiere implementar técnicas de recolección de datos que permitan analizar cada clasificador y determinar cuál ofrece los mejores resultados. Este análisis comparativo resulta fundamental para seleccionar el clasificador más efectivo.

La base de datos utilizada, PlantVillage, es una base de datos ampliamente utilizada en la comunidad de investigación agrícola y de aprendizaje automático. Contiene una gran cantidad de imágenes de hojas de plantas afectadas por diversas enfermedades, incluidas las enfermedades foliares en los cultivos de tomate. En total se utilizaron 13478 imágenes y se capturaron en diferentes condiciones y representan una variedad de síntomas asociados con enfermedades foliares en las hojas de tomate.

A las enfermedades *Leaf Mold* y *Mosaic Virus* (ver anexo 9.3) se les aplico un data augmentation para que la cantidad de imágenes en cada enfermedad sea mínima de 1000 imágenes para reducir la diferencia en total que hay por cada enfermedad, el número de imágenes utilizado por cada enfermedad se puede ver en el anexo 9.5. Con base en esto, las imágenes por cada enfermedad varían entre 1000 y 2127 imágenes para evitar saturar de información que pueda afectar el entrenamiento de la red.

Para continuar, las imágenes se pueden utilizar para entrenar y validar los clasificadores propuestos, como Naive Bayes, K-nearest neighbors (KNN), Support Vector Machine (SVM) y Convolutional Neural Network (CNN). Los clasificadores se entrenarían utilizando imágenes de hojas de tomate etiquetadas con enfermedades foliares específicas y también imágenes de hojas de tomate sanas como clase negativa. Los clasificadores Naive Bayes, KNN y SVM son módulos de código abierto que se encuentran disponibles desde la librería scikit-learn la cual se puede usar por medio de Python. Por otro lado, la CNN ResNet-50 es un modelo preentrenado de código abierto que se encuentra disponible en la librería de Tensorflow y se usa por medio de Python.

Para continuar, Las imágenes de hojas se utilizarán como datos de entrada para entrenar y evaluar un modelo de CNN con el fin de extraer características relevantes de las imágenes de las hojas de tomate, como texturas, formas y patrones, que pueden ser indicadores de enfermedades foliares. La extracción de características será la información de entrada al entrenamiento de los clasificadores. Una vez que los clasificadores estén entrenados, se utilizarán para clasificar nuevas imágenes de hojas de tomate y detectar la presencia de enfermedades foliares. El rendimiento de los clasificadores se evaluará utilizando las métricas mencionadas anteriormente, como Accuracy, Unweighted Average Recall, Recall y la Matriz de confusión.

Estas métricas proporcionarían información sobre la precisión, la capacidad de detección y el rendimiento global de los clasificadores en la detección de enfermedades foliares en los cultivos de tomate. Continuando con lo anterior, una vez obtenidos los resultados de las métricas de evaluación se procede a analizar los resultados obtenidos y se toma la decisión final de elegir el mejor clasificador para evaluarlo nuevamente con imágenes nuevas que no pertenecen a la base de datos con las que se entrenó el clasificador y obtener los resultados con esas imágenes para concluir el proyecto.

Por último, los clasificadores y las métricas de evaluación fueron programados usando el lenguaje de programación Python con una estructura de Notebook de Jupyter a través de Google Colab el cual permitió adquirir hardware como la GPU NVIDIA A100 Tensor Core, 81 GB de RAM y 250 GB de almacenamiento con el fin de entrenar la CNN (ResNet-50) y los otros clasificadores sin limitarse con el hardware que proporciona Colab por defecto.

A continuación, se muestra un diagrama que resume la metodología a seguir:



Figura 5. Diagrama de metodología a seguir. *Obtenido de:* Autores.

5.3 FASES DEL PROYECTO



Figura 6. Fases del proyecto. *Obtenido de:* Autores.

6. RESULTADOS Y ANALISIS

Para la obtención de los resultados, se realizó una extensa recopilación y curación de datos correspondientes a imágenes de hojas de tomate. Se consultaron diversas fuentes públicas que contenían conjuntos de datos etiquetados y verificados por expertos agrónomos. En total se reunieron 13478 imágenes de alta resolución provenientes principalmente de la base de datos PlantVillage con 8 tipos de enfermedades y 1 de tomates sanos como se muestra a continuación:

ENFERMEDADES	CANTIDAD
Tomate, Mancha bacteriana	2127
Tomate Tizón temprano	1000
Tomate sano	1591
Tomate Tizón tardío	1909
Moho de la hoja del tomate	1000
Manchas foliares del tomate Septoriosis	1771
Tomate Araña roja Dos	1676
Mancha diana del tomate	1404
Virus del mosaico del tomate	1000
Total	13478

Tabla 2. Cantidad de imágenes de tomates con enfermedad y sanos. Fuente: Autores.

Para la extracción de características (Feature Extraction), las imágenes fueron filtradas y procesadas para garantizar su calidad, variedad y representatividad de los diferentes estados de salud del cultivo, esto se llevó a cabo por medio del preprocesamiento que consistía en normalizar los valores de las imagen en un rango de 0 a 1 y redimensionar las imágenes a 224x224, con el objetivo de usar la CNN ResNet-50 para el Feature Extraction por medio de patrones y características relevantes de la imagen para su posterior vectorización de 1x2048, el 1 representa cada imagen o fila de la base de datos y el 2048 representa la cantidad de características en el vector por imagen. A continuación, se muestra una tabla que contiene el número de etiqueta y la etiqueta a la que pertenece ese valor con el fin de identificar en la matriz de confusión de cada clasificador la etiqueta a la que pertenece esos valores:

NUMERO DE ETIQUETA	ETIQUETA
0	Tomato Bacterial Spot
1	Tomato Early Blight
2	Tomato Late Blight
3	Tomato Leaf Mold
4	Tomato Septoria Leaf Spot
5	Tomato Spider Mites Two-Spotted Spider Mite
6	Tomato Target Spot
7	Tomato Healthy
8	Tomato Mosaic Virus

Tabla 3. Numero de etiqueta y etiqueta a la que pertenece. Fuente: Autores.

Antes de pasar a los resultados obtenidos por cada clasificador, se presenta la escala de evaluación para la precisión de validación (VAL_ACC) que se usó para determinar el rendimiento del clasificador y poder seleccionar el mejor clasificador para realizar nuevas pruebas. Se considera rendimiento bajo si los valores están por debajo de 0.7, rendimiento aceptable entre 0.7 y 0.8, rendimiento bueno entre 0.8 y 0.9, rendimiento excelente entre 0.9 y 1.

Naive Bayes:

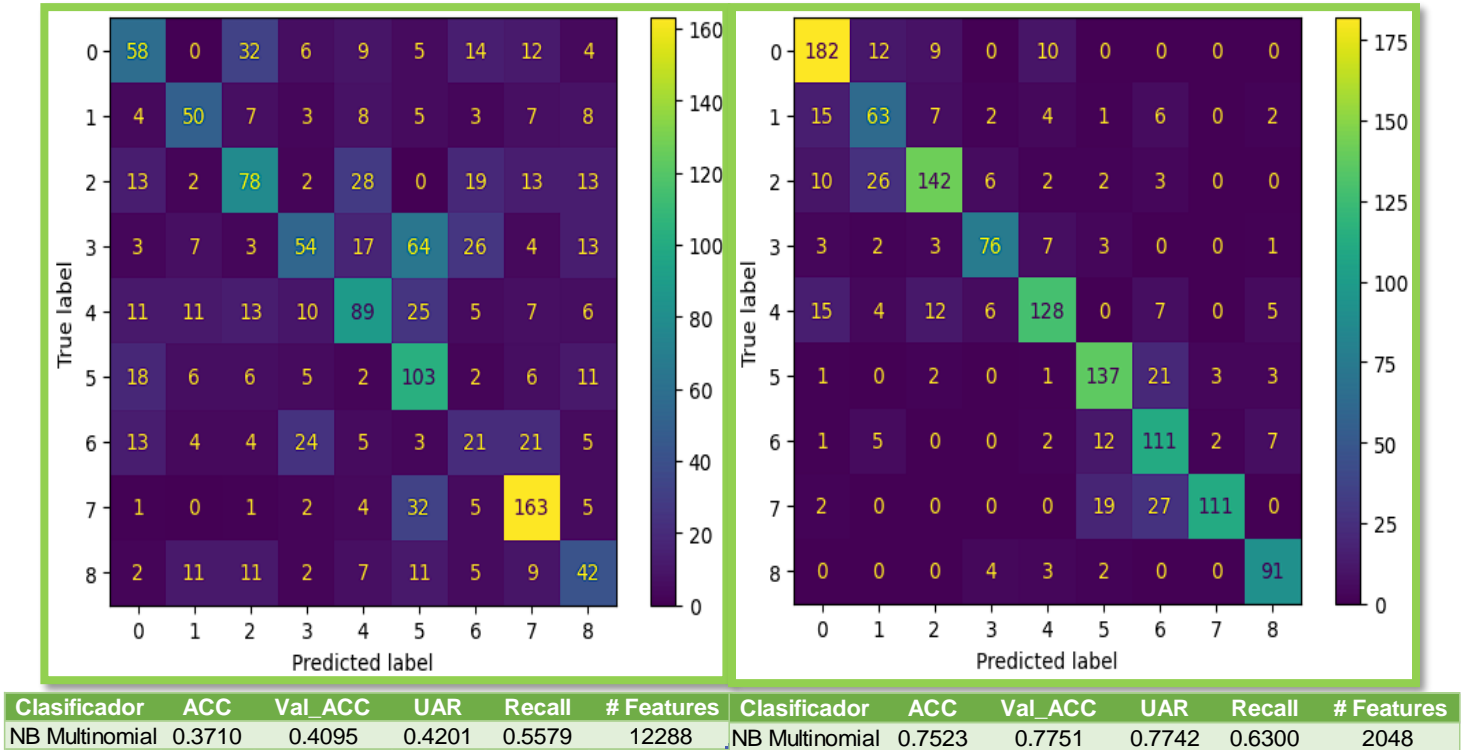


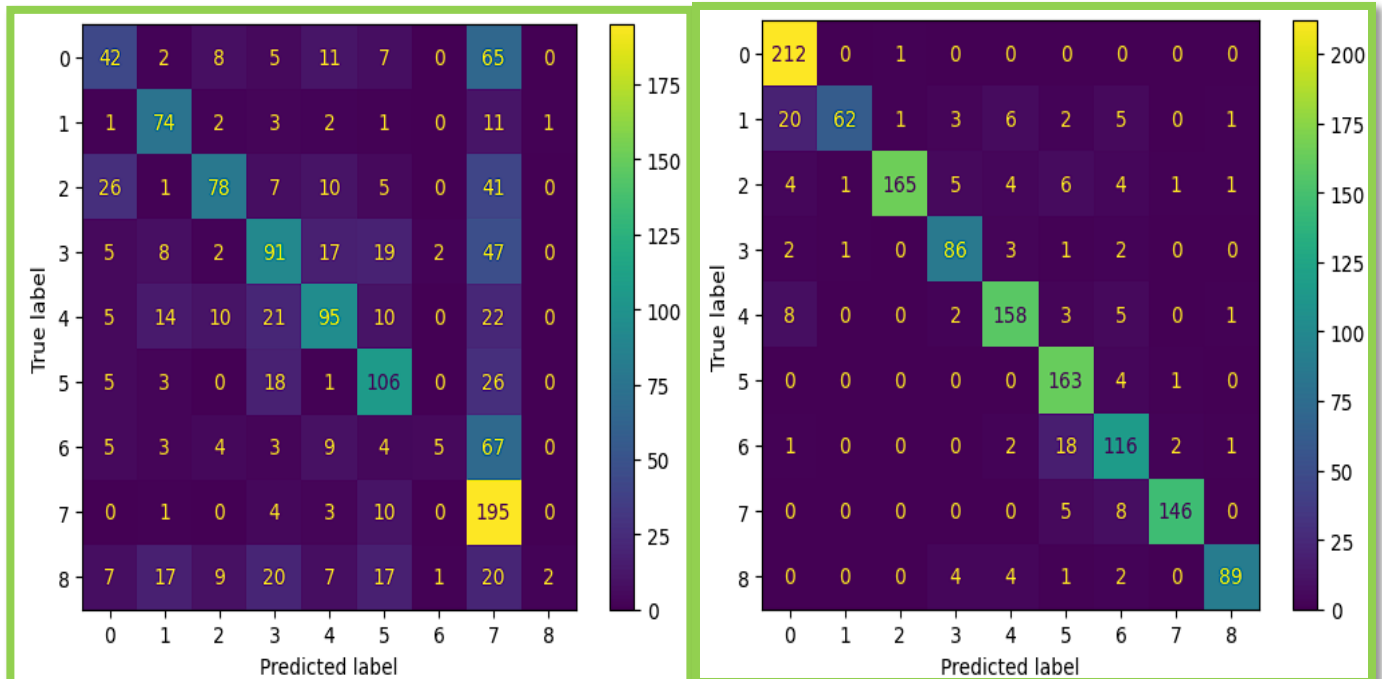
Figura 7. Resultados del clasificador Naive Bayes (Raw Data y Feature Extracted).
Obtenido de: Autores.

La imagen de la izquierda muestra los resultados del clasificador con los datos en crudo, es decir, que no se le hizo ningún tipo de extracción de características. La imagen de la derecha muestra los resultados para los datos que se les realizó una extracción de características. Como se puede observar, los resultados de la extracción de características en general tienen unos valores más altos que los datos en crudo (entre más grande o se aproxime más a 1 los valores, mejor).

Para continuar, los datos en crudo tienen que la matriz de confusión no es uniforme, tiene muchos valores como falsos positivos, falsos negativos y verdaderos negativos, razón por la cual la precisión de validación (VAL_ACC) es baja porque hay muchos datos que están prediciendo erróneamente. El UAR indica que a pesar de que su valor es un poco más alto que la precisión de validación, esta sigue siendo un valor bajo, esto quiere decir que, aunque se tenga el mismo número de imágenes para cada categoría, los resultados no mejoraran mucho.

En los datos con la extracción de características, se puede ver una matriz más uniforme a pesar de que sigue habiendo datos que no se predijeron correctamente. Su precisión de validación es aceptable y su UAR indica que no importa si se tiene o no el mismo número de imágenes por categoría se estaría obteniendo una precisión similar.

K-nearest neighbors (KNN):



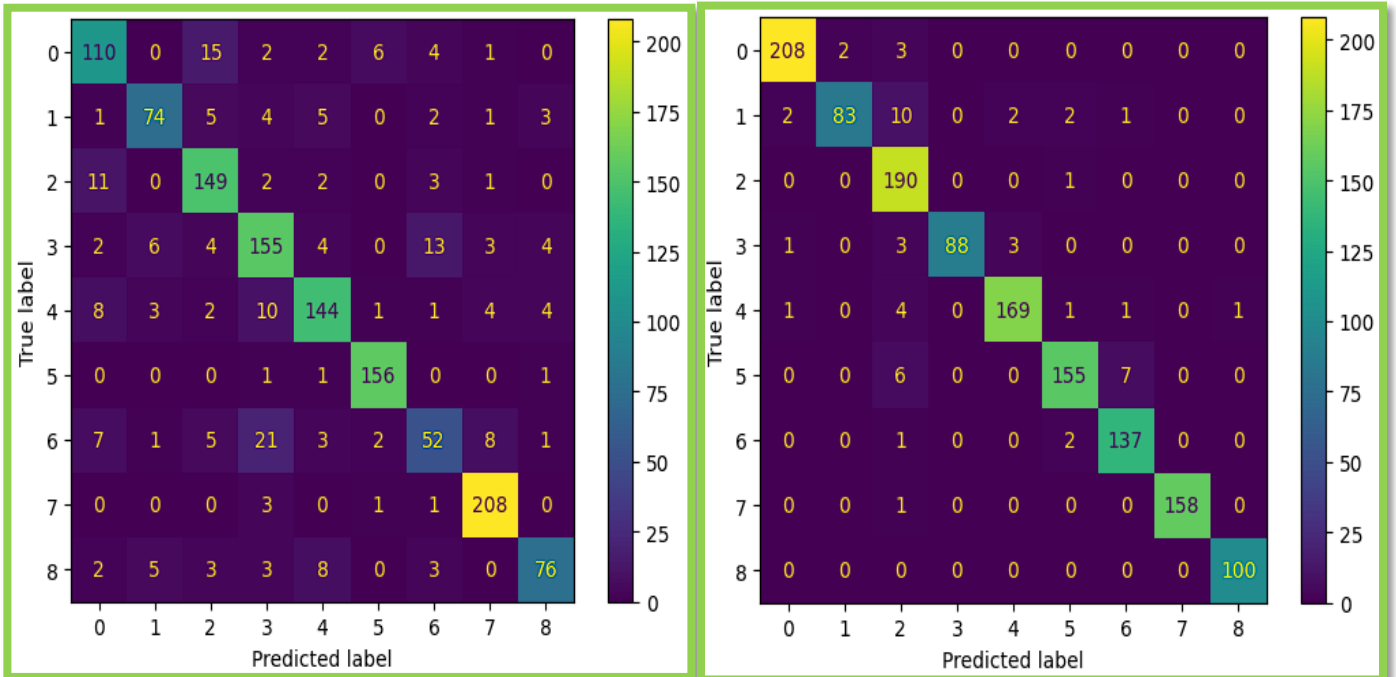
Clasificador	ACC	Val_ACC	UAR	Recall	# Features	Vecino	Clasificador	ACC	Val_ACC	UAR	Recall	# Features	Vecino
KNN	0.5714	0.5123	0.4676	0.7789	12288	6	KNN	0.9197	0.8913	0.8760	0.6200	2048	6

Figura 8. Resultados del clasificador KNN (Raw Data y Feature Extracted).
Obtenido de: Autores.

Lo primero que se puede notar es que los valores en la imagen de la derecha son mejores que la imagen de la izquierda porque se pueden observar más datos dispersos de la diagonal principal lo que indica que tiene más errores en la predicción y se ve reflejado en la precisión de predicción (VAL_ACC) en comparación a la imagen de la derecha que tiene una diagonal más definida y con menos datos dispersos en toda la matriz viendo esta “Uniformidad” reflejada en la precisión de predicción por lo cual su rendimiento es bueno por lo que este clasificador con estos datos se vuelve una opción prometedora para elegir el clasificador de mejor rendimiento.

En la imagen de la izquierda, la precisión de entrenamiento es mayor que la precisión de validación, lo que da a entender que al clasificador aun le cuesta predecir con los datos en crudo, el UAR no es mayor que la precisión de validación por lo que tener el mismo número de imágenes para cada categoría hará que su precisión disminuya un poco, también es bueno revisar el Recall que para este clasificador tiene un valor aceptable con los datos en crudo y un valor bueno con las características extraídas.

Support Vector Machine (SVM):



Clasificador	ACC	Val_ACC	UAR	Recall	# Features	Clasificador	ACC	Val_ACC	UAR	Recall	# Features
SVM	0.9592	0.8369	0.8127	0.7789	12288	SVM	0.9959	0.9590	0.9530	0.8300	2048

Figura 9. Resultados del clasificador SVM (Raw Data y Feature Extracted).
Obtenido de: Autores.

Lo primero que se observa es que los valores de la imagen derecha son mejores que los de la imagen izquierda y la razón es similar a la justificación que se dio en el clasificador KNN. Además, las matrices en ambos casos se ven uniformes con pocos valores incorrectamente predichos, por lo que su precisión de validación está entre rendimiento bueno (para la imagen izquierda) y rendimiento excelente (para la imagen de la derecha).

En la imagen de la izquierda, la precisión de entrenamiento es mayor a la de validación, esto indica que al clasificador aún le cuesta predecir con los datos crudos. El UAR no supera la precisión de validación, por lo que tener el mismo número de imágenes por categoría podría disminuir un poco su rendimiento. El Recall tiene un valor aceptable para datos crudos y bueno para datos procesados.

En la imagen de la derecha, se observa que los puntajes obtenidos abarcan desde niveles buenos hasta excelentes, lo que sugiere que el uso de Support Vector Machine (SVM) con extracción de características se presenta como una opción altamente prometedora al considerar la elección del clasificador óptimo. La consistente obtención de resultados positivos y de alta calidad en este contexto respalda la viabilidad y eficacia de la implementación de SVM con extracción de características como una estrategia clave para la optimización del rendimiento del clasificador en la tarea en cuestión. Por último, este clasificador se convierte en el mejor candidato en comparación al de NB y KNN según los resultados obtenidos, solo queda ver el rendimiento de ResNet-50 y decidir el mejor clasificador.

ResNet-50:

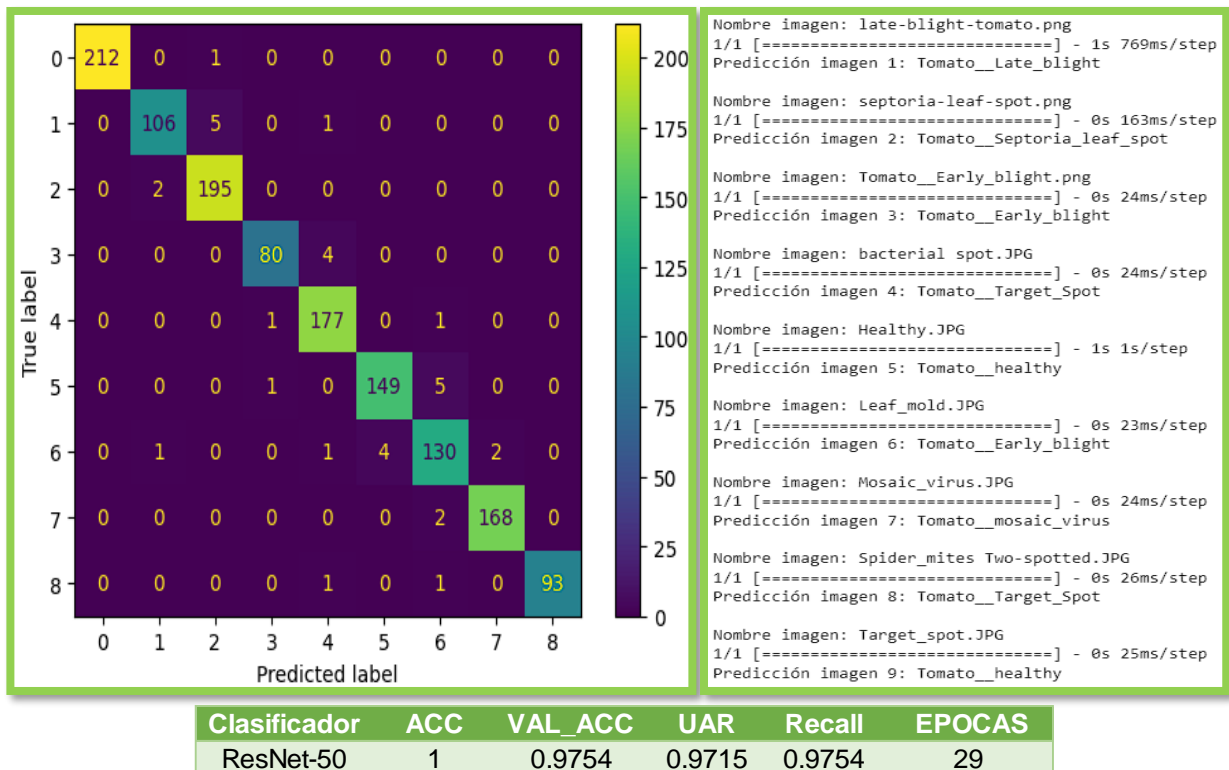


Figura 10. Resultados de la red ResNet-50 (Feature Extracted & Predictions).
Obtenido de: Autores.

Los resultados de la red ResNet-50 muestran una matriz de confusión muy cercana a la uniformidad completa en la diagonal principal, lo que indica que la mayoría de las imágenes usadas para validar los datos de entrenamiento fueron clasificadas correctamente. Esto se ve reflejado en los valores numéricos obtenidos en la precisión de validación que obtuvo un 0.9754, el UAR con un 0.9715 y un Recall de 0.9754 obteniendo así los resultados más altos de todos los clasificadores. Para conseguir esto, la red se tuvo que entrenar en 29 épocas (ciclos) para aumentar su precisión de entrenamiento e incluso desde la época número 1 ya mostraba resultados prometedores por encima de 0.8494.

Con base en lo mencionado anteriormente, el paso siguiente a seguir es seleccionar el clasificador que haya dado mejor rendimiento que, como ya se mencionó anteriormente, ese puesto le pertenece a la red ResNet-50 por obtener unos valores excelentes en todas sus métricas de evaluación respecto a los demás clasificadores. Una vez elegido este clasificador se procede a evaluarlo con nuevas imágenes, en total se evaluó con 1 imagen para cada categoría para determinar si es capaz de predecir nueva información que no está en la base de datos. Los resultados se muestran a continuación:

REPRESENTACIÓN	PREDICCIÓN ACERTADA POR ¿ESTA ENFERMA O SALUDABLE?	PREDICCIÓN ¿ESTA ENFERMA O SALUDABLE? Y ¿CUAL ENFERMEDAD?
CANTIDAD	8/9	6/9
PORCENTAJE	88,88%	66,66%

Tabla 4. Resultados de las predicciones con nuevas imágenes. Fuente: Autores.

Los resultados se muestran de forma simple según la utilidad que se quiera dar a la red. Si lo que se quiere es que la red sea capaz solamente de detectar si las hojas del cultivo se encuentran enfermas o saludables, su precisión es del 88,88%. Por otro lado, si lo que se quiere es saber si está enferma o saludable la hoja del cultivo y adicionalmente saber que enfermedad contiene, su precisión es del 66,66%.

Estos resultados fueron más bajos que la precisión de validación que se obtuvieron en las pruebas con la base de datos. Sin embargo, hay que tener en cuenta que las imágenes no provienen de la base de datos de Plant Village, sino que fueron extraídas directamente de Internet. Sin embargo, se ha realizado una verificación para asegurar su correspondencia con la enfermedad en evaluación y sean una

enfermedad del cultivo de tomate en la región andina, las imágenes usadas no fueron tomadas cuidadosamente como lo están en la base de datos, aspectos como el fondo, la iluminación, el encuadre, entre otros aspectos hacen que las fotos varíen a la base de datos y por consiguiente el resultado de predicción.

Por otro lado, solo se probó una imagen por categoría debido a que encontrar estas imágenes en línea no es fácil porque se deben verificar aspectos como la enfermedad y que si corresponda al tomate. Lo ideal hubiese sido poder tomar fotos directamente para la predicción, pero por disposición de tiempo y recursos no se pudo realizar de esta forma.

Con base en lo mencionado anteriormente, la red aún puede mejorar y bastante, si se decide mejorar la base de datos o aplicar más filtros o preprocesamiento a las imágenes de entrada para obtener datos más limpios. Además, la herramienta demuestra ser mucho más rápida para predecir enfermedades de lo que lo haría una persona manualmente, debido a que hacer esas 9 predicciones le tomo tan solo 2.2 segundos en total y 6 de las 9 imágenes identifique la enfermedad correcta por lo que esto puede dar una mejor idea o base para saber a qué enfermedad se enfrentan y ganar tiempo para tratarla. En la sección de recomendaciones de este proyecto se mencionarán cambios que se pueden hacer al proyecto para una futura continuación o expansión y también observaciones que podrían ayudar a mejorar el proyecto.

7. CONCLUSIONES

Se ha logrado una identificación exitosa de los requisitos fundamentales para el desarrollo de la herramienta, destacando la importancia de obtener un conjunto de datos representativo, aplicar técnicas efectivas de preprocesamiento de imágenes y seleccionar algoritmos de clasificación adecuados. El diseño meticuloso de la arquitectura y el flujo de trabajo ha facilitado una implementación eficiente de la herramienta, confirmando la elección acertada de algoritmos, especialmente ResNet-50, según los resultados obtenidos. La exitosa implementación de la codificación en Python, aprovechando recursos de computación en la nube, ha posibilitado el entrenamiento efectivo de modelos complejos de deep learning.

La aplicación de técnicas de inteligencia artificial para la detección temprana de enfermedades foliares en el cultivo de tomate se postula como una alternativa prometedora para mitigar las pérdidas económicas asociadas a estas patologías. Los clasificadores basados en redes neuronales convolucionales profundas, como ResNet-50, se han destacado al exhibir el mejor rendimiento para la tarea propuesta, alcanzando una impresionante precisión de validación del 0.9754. ResNet-50 ha demostrado su capacidad para extraer características discriminativas directamente de las imágenes, eliminando la necesidad de selección manual de características gracias a su arquitectura profunda.

En cuanto al clasificador SVM, se vislumbra como una opción interesante para futuras investigaciones debido a su menor demanda de recursos computacionales. Aunque su precisión de validación fue ligeramente inferior a la de ResNet-50, alcanzando el 0.9590, destaca como una alternativa viable con margen de mejora y un tiempo de entrenamiento notablemente menor. Este conjunto de resultados respalda la relevancia y la eficacia de las técnicas empleadas, consolidando el potencial de la inteligencia artificial para contribuir significativamente a la detección y gestión de enfermedades foliares en la agricultura.

El análisis de la matriz de confusión revela patrones significativos en el rendimiento de los clasificadores. Inicialmente, sin procesamiento de imágenes, la etiqueta 7 ("Tomato Healthy") destaca con la mayor precisión en todos los clasificadores. Sin embargo, tras la aplicación de técnicas de preprocesamiento, la etiqueta 0 ("Tomato Bacterial Spot") emerge como la mejor clasificada, indicando una mejora sustancial en la capacidad del modelo para identificar esta enfermedad específica.

8. RECOMENDACIONES

Se recomienda ampliar la base de datos de imágenes, incorporando una mayor diversidad de fondos y condiciones de iluminación, con el objetivo de mejorar la capacidad de generalización de los modelos. Además, se sugiere probar técnicas adicionales de aumentación de datos, tales como rotaciones, recortes y ajustes de brillo/contraste, para introducir mayor variabilidad en los datos de entrenamiento.

En el proceso de mejora, se aconseja evaluar diferentes arquitecturas de redes neuronales convolucionales, como VGG, InceptionNet, EfficientNet, entre otras, con el fin de comparar su rendimiento con el modelo actual, ResNet-50. Asimismo, se propone la implementación de técnicas de Transfer Learning, aprovechando modelos CNN preentrenados en bases de datos extensas y adaptándolos al problema específico de detección de enfermedades foliares en tomates.

Para facilitar la accesibilidad y utilidad de la herramienta, se plantea el desarrollo de una interfaz gráfica amigable con el propósito de simplificar el proceso de diagnóstico de plantas.

Además, se sugiere explorar el uso de técnicas de aprendizaje por refuerzo para mejorar de manera iterativa el rendimiento del modelo a través de la experiencia. Para optimizar la implementación en campo, se aconseja evaluar la viabilidad de ejecutar el modelo en dispositivos móviles, lo que facilitaría su uso por parte de los agricultores.

Con el objetivo de evaluar la robustez del modelo en condiciones reales, se propone realizar pruebas con imágenes capturadas directamente en el campo. Finalmente, se destaca la importancia de la transparencia y la colaboración en la investigación al publicar el código, datos y el modelo entrenado de la herramienta desarrollada, con el fin de fomentar la investigación y el avance en esta área específica.

9. RECURSOS DISPONIBLES

Para la elaboración del proyecto se requieren del uso de computadores para la programación y documentación del proyecto. A continuación, se muestran los elementos que se consideran necesarios para el desarrollo del proyecto.

Cantidad	Rubro	Detalle	Unidad	Valor	Valor Total
2	Personal	Estudiante Investigador	Moneda - COP	\$0	\$0
1		Director	Moneda - COP	\$0	\$0
0		Asesores Externos	Moneda - COP	\$0	\$0
1	Propiedades, planta y equipo	Computador de mesa	Moneda - COP	\$ 6.000.000	\$ 6.000.000
1		Computador de mesa	Moneda - COP	\$ 5.400.000	\$ 5.400.000
1		Computador portátil	Moneda - COP	\$ 3.800.000	\$ 3.800.000
2		Dispositivo móvil	Moneda - COP	\$0	\$0
2	Software	Python	Moneda - COP	\$0	\$0
2		TensorFlow	Moneda - COP	\$0	\$0
2		Sk-learn	Moneda - COP	\$0	\$0
2		Visual Studio Code	Moneda - COP	\$0	\$0
2		Google Collab	Moneda - COP	\$43.165	\$43.165
	Transportes	Gasolina	Moneda - COP	\$60.000	\$60.000
0	Libros, revistas, folletos, guías, entre otros.	No aplica	Moneda - COP	\$0	\$0
0	Impresiones, fotocopias y Empaste	No aplica	Moneda - COP	\$0	\$0
0	Imprevistos	No aplica	Moneda - COP	\$0	\$0

Tabla 5. Costos de elementos implementados. Fuente: Autores.

10. CRONOGRAMA

Ver anexo 9.4 para visualizar el cronograma.

11. BIBLIOGRAFÍA

- Perfett, J. J., Balcázar, Á., Hernández, A., & Leibovich, J. (2013). *Políticas para el desarrollo de la agricultura en Colombia*. Bogotá DC: La Imprenta Editores S.A.
- Gustavsson, J., Cederberg, C., Sonesson, U., van Otterdijk, R., & Meybeck, A. (2012). *Pérdidas y Desperdicio de Alimentos en el Mundo: Alcances Causas y Prevención*. Roma, Italia: FAO.
- Laura Brenes-Peralta, I., Maria Jimenez-Morales, I. F., & Marianella Gamboa-Murillo, I. (2015). *Diagnóstico de Pérdidas y Desperdicio Alimenticio en dos canales de comercialización de la Agrocadena de Tomate Costarricense para su posterior Disminución*.
- Carreño, N., Vargas, A., Bernal, A. J. & Restrepo, S. (2007). Problemas fitopatológicos en especies de la familia Solanaceae causados por los géneros *Phytophthora*, *Alternaria* y *Ralstonia* en Colombia. Una revisión. *Agronomía colombiana*, 25(2), 320-329.
- Amaya, P., Peña, L., Mosquera, A., Villada, H. & Villada, D. (2010). Efecto del uso de recubrimientos sobre la calidad del tomate (*Lycopersicon esculentum* mill). *Dyna*, 77(162), 67-73.
- Cultivo de Tomate en Colombia | Agro Bayer Colombia. (n.d.). Bayer. <https://www.agro.bayer.co/es-co/cultivos/tomate.html>
- Agronegocios. (2018, May 18). Antioquia y Norte de Santander son los departamentos líderes en la producción de tomate. *AGRONEGOCIOS*. <https://www.agronegocios.co/agricultura/cuales-son-las-regiones-que-mas-producen-tomate-2728689>
- Aysun Yilmaz Kizilboga | Atilla Ergüzen | Erdal Erdal. (2020). Determination of Various Diseases in Two Most Consumed Fruits using Artificial Neural Networks and Deep Learning Techniques. *International Journal of Trend in Scientific Research and Development*, 5(1), 1007–1010. <https://www.ijtsrd.com/engineering/computer-engineering/38128/determination-of-various-diseases-in-two-most-consumed-fruits-using-artificial-neural-networks-and-deep-learning-techniques/aysun-yilmaz-kizilboga%0APdf> URL: <https://www.ijtsrd.com/papers/ijtsrd38128.pdf>
- Ashqar, B. A. M., & Abu-Naser, S. S. (2018). Image-Based Tomato Leaves Diseases Detection Using Deep Learning. *International Journal of Academic Engineering Research*, 2(12), 10–16. www.ijeais.org/ijaer

- William. (2021, March 12). *Pérdidas presentadas por manejo de postcosecha en tomate de árbol cyphomandra betacea en Fusagasugá y Silvania*. Monografias.com.
<https://www.monografias.com/trabajos13/analecon/analecon>
- Hío, J. C., Jaramillo Noreña, J. E., & Rodríguez, V. P. (2003). *Manejo integrado de enfermedades*. 369–425. <http://hdl.handle.net/20.500.12324/1348>
- Red neuronal convolucional ResNet-50 - MATLAB resnet50 - MathWorks España*. (n.d.). <https://es.mathworks.com/help/deeplearning/ref/resnet50.html>
- Delgado, L. D. M. (2019). EVALUACIÓN DE INDUCCIÓN DE RESISTENCIA EN TOMATE (*Solanum lycopersicum* L.) CON SILICIO Y ANTAGONISMO DE *Trichoderma viride* CONTRA LA MARCHITEZ VASCULAR CAUSADA POR *Fusarium oxysporum* f.sp *lycopersici*.
- FAO, (2010). Faostat. Área cosechada, producción y rendimiento de tomate. En: <http://faostat.fao.org/site/567/DesktopDefault.aspx?PageID=567#ancor>, consulta: noviembre de 2015.
- Perilla, A., Rodríguez, L. F. & Bermúdez, L. T. (2011). Estudio técnico económico del sistema de producción de tomate bajo invernadero en Guateque, Sutatenza y Tenza (Boyacá). *Revista colombiana de ciencias hortícolas*, 5 (2), 220-232.
- Andrade, L. (2019). *Cultivo de tomate (Solanum lycopersicum L.) y maíz (Zea mays) como alternativa de sostenimiento para familias campesinas en Algeciras Huila*. 1–65.
https://ciencia.lasalle.edu.co/ingenieria_agronomicahttps://ciencia.lasalle.edu.co/ingenieria_agronomica/133
- Leal-Lara, Daniel-David, Barón-Velandia, Julio, & Rocha-Calderón, Camilo-Enrique. (2021). Interpretability in the Field of Plant Disease Detection: A Review. *Revista Facultad de Ingeniería*, 30(58), e100. Epub December 22, 2021.<https://doi.org/10.19053/01211129.v30.n58.2021.13495>
- Romero, A. C., & Cortes, H. H. (2019). Machine learning in plant disease detection. *Tecnología, Investigación y Academia*, 7, 55–61.
<https://revistas.udistrital.edu.co/index.php/tia/issue/archive>
- Moran, M. (2020, July 24). *Hambre y seguridad alimentaria - Desarrollo Sostenible*. Desarrollo Sostenible.
<https://www.un.org/sustainabledevelopment/es/hunger/>

- Sarkar, U., Bannerjee, G., Das, S., & Ghosh, I. (2018). Artificial Intelligence in Agriculture: A Literature Survey. *International Journal of Scientific Research in Computer Science Applications and Management Studies IJSRCSAMS*, 7(3). www.ijsrcsams.com
- Brahimi, M., Boukhalfa, K., & Moussaoui, A. (2017). Deep Learning for Tomato Diseases: Classification and Symptoms Visualization. *Applied Artificial Intelligence*, 31(4), 299–315. <https://doi.org/10.1080/08839514.2017.1315516>
- Shoaib, M., Hussain, T., Shah, B., Ullah, I., Shah, S. M., Ali, F., & Park, S. H. (2022). Deep learning-based segmentation and classification of leaf images for detection of tomato plant disease. *Frontiers in Plant Science*, 13. <https://doi.org/10.3389/fpls.2022.1031748>
- Guevara, L., & Alvarado, D. (2014). Importancia, contribución y estabilidad de antioxidantes en frutos y productos de tomate. *Avances De Investigacion Agropecuaria*, 18(1), 51–66.
- Roldán Ortega, B., Roshan Biswal, R., & Sánchez Delacruz, E. (2019). Detección de enfermedades en el sector agrícola utilizando Inteligencia Artificial. *Research in Computing Science*, 148(7), 419–427. <https://doi.org/10.13053/rcs-148-7-31>
- Bonilla Segovia, J. S., Dávila Rojas, F. A., & Villa Quishpe, M. W. (2021). Estudio del uso de técnicas de inteligencia artificial aplicadas para análisis de suelos para el sector agrícola. *Recimundo*, 5(1), 4–19. [https://doi.org/10.26820/recimundo/5.\(1\).enero.2021.4-19](https://doi.org/10.26820/recimundo/5.(1).enero.2021.4-19)
- Patrício, D. I., & Rieder, R. (2018). Computer vision and artificial intelligence in precision agriculture for grain crops: A systematic review. *Computers and Electronics in Agriculture*, 153(August), 69–81. <https://doi.org/10.1016/j.compag.2018.08.001>
- Fouquet, F. (2021). Inteligencia Artificial aplicada a la agricultura de precisión. Control de hongos en la planta de tomate.
- Abe, S. (2001). *Support Vector Machines for Pattern Classification*. Springer.
- Alexander, T. V. (2022). *DESARROLLO DE UN SISTEMA UAS PARA EL MONITOREO CUANTITATIVO*. Bogota: Uniagraria.
- Blancard, D. (2015). *Tomato Diseases: Identification, Biology and Control*. . Wiley-Blackwell.
- Chilealimentos. (15 de Abril de 2021). *Chilealimentos*. Obtenido de TOMATES – ITALIA: VIRUS CAUSA PÉRDIDAS DE PRODUCCIÓN DE HASTA EL 40%: <https://chilealimentos.com/tomates-italia-virus-causa-perdidas-de-produccion-de-hasta-el-40/>

- Cubillos, I. E. (2015). Tomate. *CCB (camara de comercio de Bogota)*, 56. enAlimentos. (18 de Mayo de 2022). *enAlimentos*. Obtenido de Agricultura mundial registra pérdidas por cerca de 300 mil mdd: FAO: <https://enalimentos.lat/noticias/5363-agricultura-mundial-registra-perdidas-por-cerca-de-300-mil-mdd-fao.html>
- Ferreira, K. (20 de 07 de 2021). *rockcontent*. Obtenido de Tipos de Inteligencia Artificial: <https://rockcontent.com/es/blog/tipos-de-inteligencia-artificial/>
- Fredy Giovanni Gaspar Lancheros, D. A. (2018). *Diseño de un sistema de monitoreo para la medición de madurez de feijoas en los cultivos de la sabana de bogotá*. Tesis, Fundación Universitaria Agraria de Colombia - UNIAGRARIA, Bogotá D.C, Bogotá D.C. Obtenido de <https://uniagraria.janium.net/janium-bin/detalle.pl?Id=20230910204006>
- Gavilán, I. G. (5 de 06 de 2017). *Blue chip*. Obtenido de Clasificadores: el encuentro entre Data Science, Machine Learning y Redes Neuronales: <http://bluechip.ignaciogavilan.com/2017/06/clasificadores-el-encuentro-entre-data.html#.ZGGhDXbMJPY>
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. EEUU: The MIT press.
- Hurtado, J. E. (1984). *Evaluación de perdidas de Post-cosecha en tomate (Lycopersicom esculentum Mill)*. Bogotá.
- Ian Goodfellow, A. C. (2015). *Deep Learning (Adaptive Computation and Machine Learning series)*. Estados Unidos: The Mit Press.
- Intel. (s.f.). *Intel*. Obtenido de Qué es el Modelado de Inteligencia Artificial: <https://www.intel.la/content/www/xl/es/analytics/data-modeling.html>
- Jiang, L., Zhang, H., & Cai, Z. (2009). A Novel Bayes Model: Hidden Naive Bayes. *IEEE Transactions on Knowledge and Data Engineering*, 1361-1371. doi:{10.1109/TKDE.2008.234}}
- Jones , J. B., & Zitter, T. A. (2014). *Compendium of Tomato Diseases and Pests (2nd ed.)*. APS Press.
- Kramer, O. (2013). *Dimensionality Reduction with Unsupervised Nearest Neighbors*. Berlin: Springer.
- Ritchie, P. S. (2007). *Phytopathogenic Prokaryotes: Plant-Associated Bacteria*. Springer.
- Ross, T. J. (2016). *Fuzzy Logic with Engineering Applications*. WILEY.
- Salas, R. (NN). *Redes Neuronales Artificiales*. Chile: Departamento de Computación. Universidad de Valparaíso.
- Sinclair, J. B., & Backman, P. A. (2011). *Diseases of Vegetable Crops*. APS Press.
- Trigiano, R. N., Windham, M. T., & Windham, A. S. (2017). *Plant Pathology: Concepts and Laboratory Exercises*. CRC Press.

- Vecino, J. B. (06 de N/N). Inteligencia artificial en la agricultura: Perspectivas de los sistemas expertos. pág. 77.
- Zhou, Z.-H. (2022). *Machine Learning*. China: Springer.
- Rouhiainen, L. (2018). *Inteligencia artificial: 101 cosas que debes saber hoy sobre nuestro futuro*. Alianza Editorial.
- Iglesias, R. B. (2013). *Inteligencia artificial avanzada*. Editorial UOC.
- Russell, S. J., & Norvig, P. (2004). *Inteligencia Artificial: Un enfoque moderno* (2da ed.). Pearson Educación.
- Blanca, J., Cañizares, J., Cordero, L., Fernández, S., Morales, R., Rodríguez, F., ... & Valpuesta, V. (2015). *Variedades tradicionales de tomate en España*. Ministerio de Agricultura, Alimentación y Medio Ambiente.
- MathWorks (2023). Classifier. <https://www.mathworks.com/help/deeplearning/ug/classifier.html>
- Agrios, G. N. (2005). *Plant pathology* (5th ed.). Elsevier Academic Press.
- Tharwat, A. (2021). Classification assessment methods. *Applied Computing and Informatics*, 17(1), 168-192. <https://doi.org/10.1016/j.aci.2018.08.003>
- Taiz, L., & Zeiger, E. (2006). *Fisiología vegetal* (3a ed.). Universitat Jaume I.
- ¿Qué es el algoritmo de k vecinos más cercanos? | IBM. (n.d.). <https://www.ibm.com/es-es/topics/knn>
- V. Tilva, J. Patel and C. Bhatt, "Weather based plant diseases forecasting using fuzzy logic," *2013 Nirma University International Conference on Engineering (NUiCONE)*, Ahmedabad, India, 2013, pp. 1-5, doi: 10.1109/NUiCONE.2013.6780173.
- D.W. Boyd, M.K. Sun, Prototyping an expert system for diagnosis of potato diseases, *Computers and Electronics in Agriculture*, Volume 10, Issue 3, 1994, Pages 259-267, ISSN 0168-1699, [https://doi.org/10.1016/0168-1699\(94\)90045-0](https://doi.org/10.1016/0168-1699(94)90045-0).
- Desarrollosidn. (2022, August 1). Deep Learning: clasificando imágenes con redes neuronales | LIS Data Solutions. LIS Data Solutions. <https://www.lisdatasolutions.com/es/blog/deep-learning-clasificando-imagenes-con-redes-neuronales/>
- Mukherjee, S. (2022, August 18). The annotated RESNet-50 - towards data science. Medium. <https://towardsdatascience.com/the-annotated-resnet-50-a6c536034758>

12. ANEXOS

Anexo 9.1:



Artículo	Tipo de red	Característica	Precisión
Uso de SVM para detección de enfermedades en hojas de tomate (Roldán et al., 2019)	SVM, KNN, lógica difusa.	Imágenes de hojas de tomate.	Mejor resultado con SVM: 97.62% de precisión.
Detección de enfermedades en manzanas y membrillos con CNN (Yilmaz Kizilboga et al., 2020)	CNN propia y AlexNet.	22 manzanas, 18 membrillos.	83.3% precisión con CNN propuesta. 81.3% con AlexNet.
Clasificación de enfermedades en hojas de tomate con CNN y teléfonos inteligentes (Ashqar & Abu-Naser, 2018).	CNN	9000 imágenes de hojas de tomate.	99.84% de precisión en validación.
Monitoreo de madurez en feijoa con variable eléctricas e IA (Fredy & Escobar, 2018).	KNN, RNA.	Mediciones eléctricas.	Hasta 95% de precisión.
Clasificación de enfermedades en tomate con deep learning (Brahimi et al., 2017)	CNN (AlexNet, GoogleNet)	9 clases de enfermedades foliares. 14,828 imágenes de entrenamiento.	Precisión de 99.18% con CNN.



Anexo 9.2:



Tipo de clasificador	Documento	Autor	Características
NB	A Novel Bayes Model: Hidden Naive Bayes	(Jiang, Zhang, & Cai, 2009)	Naive Bayes es un algoritmo de clasificación probabilístico basado en el teorema de Bayes. Es conocido por su simplicidad y eficiencia en el procesamiento de grandes volúmenes de datos. Este clasificador asume que todas las características son independientes entre sí, lo cual puede no ser realista en muchos casos, pero aun así ofrece buenos resultados en problemas de clasificación.
KNN	Dimensionality Reduction with Unsupervised Nearest Neighbors	(Kramer, 2013)	K-nearest neighbors (KNN) es un método de aprendizaje automático que se utiliza para clasificar objetos o predecir valores basándose en la proximidad de los vecinos más cercanos en el espacio de datos. En KNN, se seleccionan los K puntos de datos más cercanos a un nuevo punto de prueba y se utiliza su información para tomar decisiones



Tipo de clasificador	Documento	Autor	Características
SVM	Support Vector Machines for Pattern Classification	(Abe, 2001)	El clasificador SVM es un algoritmo de aprendizaje supervisado que busca encontrar un hiperplano óptimo para separar muestras de diferentes clases en un espacio de alta dimensión. Utiliza una técnica de maximización de margen para lograr una buena capacidad de generalización. Es efectivo en conjuntos de datos linealmente separables y puede extenderse a casos no lineales mediante el uso de trucos de kernel. SVM ha sido exitoso en reconocimiento de imágenes, clasificación de textos y diagnóstico médico.
CNN	Deep Learning	(Ian Goodfellow, 2015)	Las redes neuronales convolucionales (CNN) son un tipo de arquitectura de redes neuronales profundas diseñadas especialmente para el procesamiento de imágenes y reconocimiento visual. Utilizan capas de convolución para extraer características importantes de las imágenes y capas de pooling para reducir la dimensionalidad. Las CNN son capaces de aprender patrones complejos y jerarquías de características, lo que las hace altamente efectivas en tareas de clasificación y detección de objetos en imágenes.
ResNet-50	Redes Neuronales Artificiales	<i>(Red Neuronal Convolutiva ResNet-50 - MATLAB Resnet50 - MathWorks España, n.d.)</i>	Es una red neuronal convolutiva con 50 capas de profundidad. Puede cargar una versión preentrenada de la red neuronal entrenada con más de un millón de imágenes desde la base de datos de ImageNet. Como resultado, la red neuronal ha aprendido representaciones ricas en características para una amplia gama de imágenes. El tamaño de la entrada de imagen de la red neuronal es de 224 por 224.

Anexo 9.3:

Enfermedad	Documento	Autor	Característica	Muestra
<p>Mancha bacteriana del tomate (<i>Xanthomonas</i> spp.)</p>	<p>Phytopathogenic Prokaryotes: Plant-Associated Bacteria</p>	<p>(Ritchie, 2007)</p>	<p>Esta enfermedad se caracteriza por la aparición de manchas necróticas en las hojas, tallos y frutos del tomate. Es causada por bacterias del género <i>Xanthomonas</i> y puede resultar en la reducción del rendimiento del cultivo.</p>	
<p>Tizón temprano del tomate (<i>Alternaria solani</i>)</p>	<p>Diseases of Vegetable Crops</p>	<p>(Sinclair & Backman, 2011)</p>	<p>El tizón temprano es una enfermedad fúngica que afecta las hojas, tallos y frutos del tomate. Se manifiesta con la presencia de manchas marrones o negras con un halo amarillo alrededor. Puede causar defoliación y pérdida de frutos.</p>	

Enfermedad	Documento	Autor	Característica	Muestra
Tizón tardío del tomate (<i>Phytophthora infestans</i>)	Diseases of Vegetable Crops	(Sinclair & Backman, 2011)	El tizón tardío es una enfermedad causada por un oomiceto que afecta principalmente las hojas y los frutos del tomate. Provoca manchas irregulares y necrosis en las hojas, así como un rápido deterioro de los frutos. Es una enfermedad altamente destructiva y puede causar grandes pérdidas en los cultivos.	
Moho de la hoja del tomate (<i>Cladosporium</i> spp.)	Compendium of Tomato Diseases and Pests (2nd ed.)	(Jones & Zitter, 2014)	El moho de la hoja del tomate es causado por hongos del género <i>Cladosporium</i> . Se caracteriza por la aparición de manchas de color marrón o negro en las hojas del tomate. Puede afectar negativamente el crecimiento y desarrollo de la planta.	

Enfermedad	Documento	Autor	Característica	Muestra
<p>Manchas foliares del tomate (Septoria spp.)</p>	<p>Phytopathogenic Prokaryotes: Plant-Associated Bacteria</p>	<p>(Ritchie, 2007)</p>	<p>Las manchas foliares en las hojas de tomate son causadas por hongos del género Septoria y se caracterizan por pequeñas manchas circulares marrones en las hojas. Con el tiempo, estas manchas pueden agrandarse y provocar la caída de las hojas de la planta.</p>	
<p>Araña roja del tomate (Tetranychus artice)</p>	<p>Plant Pathology: Concepts and Laboratory Exercises</p>	<p>(Trigiano, Windham, & Windham, 2017)</p>	<p>La araña roja es una plaga común que afecta a los cultivos de tomate. Estos ácaros se alimentan de las hojas de la planta, provocando la aparición de manchas amarillas y un aspecto marchito. Además, producen una fina telaraña en la superficie de las hojas.</p>	

Enfermedad	Documento	Autor	Característica	Muestra
<p>Mancha diana del tomate (Corynespora cassicola)</p>	<p>Phytopathogenic Prokaryotes: Plant-Associated Bacteria</p>	<p>(Ritchie, 2007)</p>	<p>La mancha diana es una enfermedad causada por el hongo Corynespora cassicola. Se caracteriza por la aparición de manchas circulares con un centro marrón rodeado por un halo amarillo en las hojas y frutos del tomate. Puede afectar el rendimiento y la calidad de los cultivos.</p>	
<p>Virus del mosaico del tomate (Tomato mosaic virus)</p>	<p>Tomato Diseases: Identification, Biology and Control</p>	<p>(Blancard, 2015)</p>	<p>Este virus causa manchas claras y oscuras en las hojas del tomate, lo que resulta en un patrón de mosaico característico. También puede provocar deformaciones en los frutos y una disminución significativa en el rendimiento del cultivo.</p>	

Anexo 9.4:

PROYECTO DE GRADO						
ID	Nombre de tarea / Título	Asignado a	Fecha de inicio planificada	Fecha de finalización planificada	Nº Semana	Estado
1	Fase I (Documentación)		2/05/2023	30/06/2023	1-7	
1.1	Planteamiento del problema	Sergio sanchez y Santiago Bolivar	2/05/2023	5/05/2023	1	Terminado
1.2	Alcances del proyecto	Sergio sanchez y Santiago Bolivar	8/05/2023	12/05/2023	2	Terminado
1.3	Recopilación de información para iniciar proyecto	Sergio sanchez y Santiago Bolivar	15/05/2023	30/06/2023	3-7	Terminado
2	Fase II (Programación de datos y algoritmos)		3/07/2023	1/10/2023	8-21	
2.1	Preprocesamiento de base de datos	Sergio sanchez y Santiago Bolivar	3/07/2023	8/09/2023	8-18	Terminado
2.2	Pruebas de algoritmos con base de datos	Sergio sanchez y Santiago Bolivar	11/09/2023	22/09/2023	19-20	Terminado
2.3	Análisis de los resultados de algoritmos	Sergio sanchez y Santiago Bolivar	25/09/2023	1/10/2023	21	Terminado
3	Fase III (Prueba I.A)		2/10/2023	20/10/2023	22-24	
3.1	Elección del algoritmo más óptimo	Sergio sanchez y Santiago Bolivar	2/10/2023	13/10/2023	22-23	Terminado
3.2	Pruebas del algoritmo con nuevas imágenes	Sergio sanchez y Santiago Bolivar	16/10/2023	20/10/2023	24	Terminado
4	Fase IV (Finalización documentación)		23/10/2023	29/10/2023	25	
4.1	Resultados, conclusiones y recomendaciones	Sergio sanchez y Santiago Bolivar	23/10/2023	29/10/2023	25	Terminado