



TECNOLÓGICO
NACIONAL DE MÉXICO®



Instituto **Tecnológico**®
de Aguascalientes

INSTITUTO TECNOLÓGICO DE AGUASCALIENTES

TESIS PARA OBTENER EL GRADO DE:
MAESTRO EN CIENCIAS DE LA INGENIERÍA
**Algoritmo Basado en Deep Learning para
Detección de Plagas**

PRESENTA: ING. RAFAEL CAMARILLO BARBA
DIRECTOR: DR. FRANCISCO JAVIER LUNA ROSASCO
CO-DIRECTOR: MARCO ANTONIO HERNANDEZ
VARGAS



Educación
Secretaría de Educación Pública



TECNOLÓGICO
NACIONAL DE MÉXICO



Instituto Tecnológico de Aguascalientes
División de Estudios de Posgrado e Investigación

Aguascalientes, Ags., **29/octubre/2025**

Oficio No. DEPI 670/2025

Asunto: Autorización impresión tesis

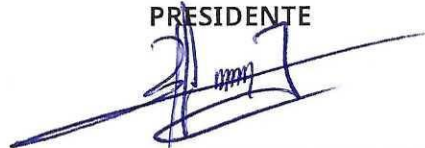
C. RAFAEL CAMARILLO BARBA
PRESENTE

Los abajo firmantes, Miembros del Jurado para Examen de Grado de Maestría, hacemos **constar** que, habiendo revisado el Trabajo de Tesis desarrollado por Usted, bajo el título: **"Algoritmo Basado en Deep Learning para Detección de Plagas"**, hemos dictaminado que éste es de aceptarse, por lo que se autoriza su impresión.

ATENTAMENTE

Excelencia en Educación Tecnológica®
Ingenio, Cultura y Saber, que conducen a la Excelencia®

PRESIDENTE


DR. FRANCISCO JAVIER LUNA ROSAS

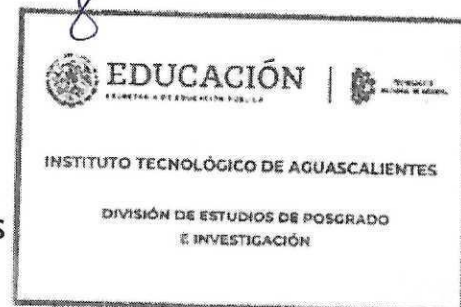
SECRETARIO


DR. JULIO CÉSAR MARTÍNEZ ROMO

VOCAL


DR. MARCO ANTONIO HERNÁNDEZ VARGAS

ccp. Archivo
JFLH/yevo



2025
Año de
La Mujer
Indígena

Av. Adolfo López Mateos #1801. Fracc. Bona Gens
C.P. 20256 Aguascalientes, Ags. Tel. 449 910 5002 ext. 127
e-mail: jefatura.depi@aguascalientes.tecnm.mx
tecnm.mx | www.aguascalientes.tecnm.mx



AUTORIZACIÓN DE USO DE DERECHOS DE AUTOR OTORGADO POR

Rafael Camarillo Barba, mayor de edad, con domicilio ubicado en DEL AGUILA 142, BARRIO EL ENCINO, AGUASCALIENTES, AGS, en mi calidad de titular y autor de la tesis denominada **Algoritmo Basado en Deep Learning para Detección de Plagas** quien para todos los fines del presente documento se denominará **EL AUTOR Y/O TITULAR**, suscribo el presente documento de autorización de uso de derechos patrimoniales de autor a favor del Tecnológico Nacional de México/Instituto Tecnológico de Aguascalientes el cual se registrará por las clausulas siguientes:

PRIMERA – AUTORIZACIÓN: **EL AUTOR Y/O TITULAR**, mediante el presente documento autoriza la utilización de los derechos patrimoniales de autor al Tecnológico Nacional de México/Instituto Tecnológico de Aguascalientes, de la tesis denominada **Algoritmo Basado en Deep Learning para Detección de Plagas**, a través del Repositorio Institucional del Tecnológico Nacional de México (en lo sucesivo TecNM) y en el Repositorio Nacional, que puede ser consultado en la liga electrónica: (www.repositorionacionalcti.mx/).

SEGUNDA – OBJETO: Por medio del presente escrito, **EL AUTOR Y/O TITULAR** autoriza al Tecnológico Nacional de México/Instituto Tecnológico de Aguascalientes, a través del Repositorio Institucional del TecNM y en el Repositorio Nacional para que de conformidad con la Ley Federal del Derecho de Autor y la Ley de la Propiedad Industrial, use los derechos del documento antes referido, con fines exclusivamente académicos.

TERCERA – TERRITORIO: Los derechos aquí autorizados se dan sin limitación geográfica o territorial alguna.

CUARTA – ALCANCE: La presente autorización se da tanto para formato o soporte material, y se extiende a la utilización en medio óptico, magnético, electrónico, en red, mensajes de datos o similar conocido o por conocer, del ejemplar o número respectivo de la publicación.

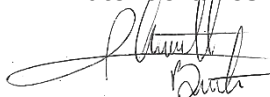
QUINTA – EXCLUSIVIDAD: La autorización de uso aquí establecida no implica exclusividad en favor del Tecnológico Nacional de México/Instituto Tecnológico de Aguascalientes. Por lo tanto, **EL AUTOR Y/O TITULAR** en su carácter de autor de la obra objeto del presente documento se reserva el derecho de publicar directamente, u otorgar a cualquier tercero, autorizaciones de uso similares o en los mismos términos aquí acordados.

SEXTA – DERECHOS MORALES (Créditos y mención): La Autorización de los derechos antes mencionados no implica la cesión de los derechos morales sobre los mismos por cuanto en conformidad con lo establecido en los artículos 18, 19, 20, 21, 22 y 23 de la Ley Federal de Derechos de Autor, dada la cuenta que estos derechos son inalienables, imprescriptibles, irrenunciables e inembargables. Por lo tanto, los mencionados derechos seguirán radicados en cabeza de **EL AUTOR Y/O TITULAR**, y siempre deberá mencionarse su nombre cuando se utilice la obra.

SÉPTIMA – AUTORÍA: **EL AUTOR Y/O TITULAR**, declara y ratifica que el material objeto de la presente fue realizada por él o por ella sin violar o usurpar derechos de Propiedad Intelectual de terceros.

Aguascalientes, Aguascalientes, a 29 días del mes de octubre de 2025.

Autor de la Tesis



Rafael Camarillo Barba

Nombre y firma

Director de Tesis



Dr. Francisco Javier Luna Rosas

Nombre y firma



Instituto Tecnológico de Aguascalientes
División de Estudios de Posgrado e Investigación

Aguascalientes, Ags., **29/octubre/2025**
Asunto: Constancia de Originalidad

**CONSEJO DE POSGRADO
MAESTRÍA EN CIENCIAS DE LA INGENIERÍA
PRESENTE**

El director de tesis de **Rafael Camarillo Barba**, estudiante de **Maestría en Ciencias de la Ingeniería** con número de control **G23153203**, hace constar que, después de haber realizado la revisión del informe de originalidad generado con la herramienta Turnitin, el trabajo **"Algoritmo Basado en Deep Learning para Detección de Plagas"** es original conforme a los criterios establecidos. El porcentaje de similitud es de **28%** y se adjunta resumen del reporte.

ATENTAMENTE

Excelencia en Educación Tecnológica®

Ingenio, Cultura y Saber, que conducen a la Excelencia

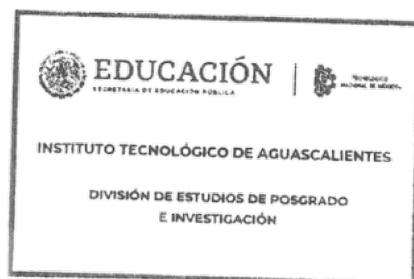
Vo.Bo.

DR. FRANCISCO JAVIER LUNA ROSAS
DIRECTOR DE TESIS

mardel

DR. MARIO ALBERTO RODRÍGUEZ DÍAZ
**COORDINADOR DE LA MAESTRÍA EN CIENCIAS
DE LA INGENIERÍA**

ccp. Archivo



2025
Año de
**La Mujer
Indígena**

Av. Adolfo López Mateos #1801. Fracc. Bona Gens
C.P. 20256 Aguascalientes, Ags. Tel. 449 910 5002 ext. 127
e-mail: jefatura.depi@aguascalientes.tecnm.mx
tecnm.mx | www.aguascalientes.tecnm.mx



Instituto Tecnológico
de Aguascalientes



Tesis_MCI_Rafael Camarillo Barba.pdf

-  Rafael Camarillo Barba
-  Rafael Camarillo Barba
-  Tecnológico Nacional de Mexico

Detalles del documento

Identificador de la entrega

trn:oid:::20755:519946146

Fecha de entrega

29 oct 2025, 9:01 a.m. GMT-6

Fecha de descarga

29 oct 2025, 9:08 a.m. GMT-6

Nombre del archivo

Tesis_MCI_Rafael Camarillo Barba.pdf

Tamaño del archivo

1.9 MB

62 páginas

16.043 palabras

98.948 caracteres

28% Similitud general

El total combinado de todas las coincidencias, incluidas las fuentes superpuestas, para ca...

Filtrado desde el informe

- Bibliografía
- Texto citado
- Texto mencionado
- Coincidencias menores (menos de 8 palabras)

Fuentes principales

- 20%  Fuentes de Internet
- 2%  Publicaciones
- 17%  Trabajos entregados (trabajos del estudiante)

Marcas de integridad

N.º de alertas de integridad para revisión

No se han detectado manipulaciones de texto sospechosas.

Los algoritmos de nuestro sistema analizan un documento en profundidad para buscar inconsistencias que permitirían distinguirlo de una entrega normal. Si advertimos algo extraño, lo marcamos como una alerta para que pueda revisarlo.

Una marca de alerta no es necesariamente un indicador de problemas. Sin embargo, recomendamos que preste atención y la revise.





Agradecimientos

Deseo expresar mi más profundo agradecimiento a mi familia, cuyo apoyo incondicional ha sido fundamental para la culminación de este trabajo. Gracias por su comprensión, paciencia y motivación constante durante todo este proceso.

A mis padres, por enseñarme el valor del esfuerzo y la perseverancia, por creer siempre en mí y brindarme la fortaleza necesaria para seguir adelante en los momentos difíciles.

De manera muy especial, quiero agradecer a mi hermano, por su compañía, apoyo y aliento durante el tiempo en que atravesé una etapa de enfermedad. Su presencia y palabras fueron una fuente invaluable de ánimo y esperanza que me permitió continuar y alcanzar esta meta.

Este logro no habría sido posible sin el respaldo y el amor de cada uno de ellos, a quienes dedico con gratitud este trabajo.

Resumen

El presente documento describe el desafío de la detección temprana y precisa de enfermedades en cultivos agrícolas mediante el desarrollo de un algoritmo basado en deep learning. Se centra específicamente en el uso de redes neuronales convolucionales (CNN) para identificar patrones complejos en imágenes de plantas y diagnosticar la presencia de enfermedades de manera automatizada y eficiente.

La necesidad de realizar esta investigación surge debido a la importancia crítica de la detección temprana y precisa de enfermedades en cultivos agrícolas. Las enfermedades de las plantas representan una amenaza significativa para la seguridad alimentaria y la sostenibilidad agrícola, ya que pueden causar pérdidas masivas de cultivos y disminuir la calidad y cantidad de alimentos disponibles. La detección manual de enfermedades en cultivos es un proceso laborioso y propenso a errores, lo que resalta la necesidad de desarrollar métodos automatizados y eficientes para esta tarea. Por lo tanto, la investigación sobre el desarrollo de un algoritmo basado en deep learning para la detección de enfermedades en cultivos agrícolas se justifica por su potencial para abordar este desafío de manera efectiva.

El proceso de desarrollo del algoritmo implica la recopilación de un amplio conjunto de datos de imágenes de plantas afectadas por diferentes enfermedades, así como de imágenes de plantas sanas como referencia. Estos datos se utilizan para entrenar y validar la CNN, que es capaz de aprender y reconocer características distintivas asociadas con cada enfermedad.

Abstract

This paper describes the challenge of early and accurate detection of diseases in agricultural crops through the development of an algorithm based on deep learning. It focuses specifically on the use of convolutional neural networks (CNNs) to identify complex patterns in plant images and diagnose the presence of diseases in an automated and efficient manner.

The need to conduct this research arises due to the critical importance of early and accurate disease detection in agricultural crops. Plant diseases pose a significant threat to food security and agricultural sustainability, as they can cause massive crop losses and decrease the quality and quantity of food available. Manual disease detection in crops is a laborious and error-prone process, highlighting the need to develop automated and efficient methods for this task. Therefore, research into the development of a deep learning-based algorithm for disease detection in agricultural crops is justified by its potential to address this challenge effectively.

The algorithm development process involves collecting a broad dataset of images of plants affected by different diseases, as well as images of healthy plants for reference. This data is used to train and validate the CNN, which is capable of learning and recognizing distinctive features associated with each disease.

Indices de contenido

Agradecimientos.....	2
Resumen	3
Abstract	4
1 Protocolo de investigación.....	11
1.1 Introducción.....	11
1.2 Antecedentes	12
1.3 Marco teórico	14
1.4 Problemática	16
1.5 Objetivos	16
1.5.1 Objetivo general	16
1.5.2 Objetivos particulares	16
1.5.3 Objetivos específicos.....	17
1.6 Metodología	17
1.7 Justificación de los resultados esperados.....	19
2 inteligencia Artificial.....	19
2.1 Introducción	19
2.2 Inteligencia Artificial General (AGI)	20
2.2.1 Características de la inteligencia artificial general.....	20
2.2.2 Aplicaciones de la Inteligencia Artificial General	21
2.3 Inteligencia Artificial Específica (ANI).....	22
2.3.1 Características de la inteligencia artificial específica	23
2.3.2 Aplicaciones de la Inteligencia Artificial Específica	23
3 Machine learning.....	24
3.1 Introducción	25
3.2 Características principales del Machine learning	25
3.3 Principales tipos de Machine Learning.....	25
3.4 Tipos de datos	27
4 Redes neuronales	28
4.1 Introducción	28
4.2 Concepto y estructura de una red neuronal	29
4.3 Tipos de redes neuronales	29
4.4 Funcionamiento de una red neuronal	30

5 Redes Neuronales convolucionales	31
5.1 Introducción	31
5.1.1 Arquitectura de una CNN	32
5.1.2 Operación de Convolución	32
5.1.3 Tipos de datos	33
5.1.4 Algoritmos de Agrupación (Pooling)	33
5.1.5 Entrenamiento y Optimización de la CNN	34
5.1.6 Algoritmos de Convolución Eficientes	34
5.1.7 Aplicación de CNN para la Detección de Plagas en Hojas de Cultivo	35
6 Transfer learning	35
6.1 Características Clave del Transfer Learning	36
6.1.1 Reutilización de modelos	36
6.1.3 Tipos de modelos	37
7 Redes Generativas Antagónicas (GANs)	38
7.1 Características de las GANs	39
7.2 Modelos de GANs	40
8 Metodología y Resultados	41
8.1 Clasificación con Base de Datos	41
8.1.1 Etapa 1: Adquisición y preprocesamiento de imágenes	41
8.1.2 Etapa 2: Entrenamiento de modelos	43
8.1.3 Transformaciones aplicadas a imágenes	44
8.1.4 Transferencia de aprendizaje	46
8.1.5 Ajuste fino (Fine tuning)	47
8.2 Configuración de hiperparámetros	48
8.3 Validación de modelos	48
8.4 Análisis comparativo de modelos	49
8.5 Predicción de nuevas imágenes	52
8.6 Optimización del modelo	52
8.7 Conclusiones parciales	58
8.8 Conclusiones generales	59
8.9 Referencias	60
Bibliografía	60

Índice de figuras

Fig. 1.6.1 Proceso Metodológico.....	17
fig. 2.2.1 Inteligencia Artificial General.....	20
fig. 2.3.1 Tipos Inteligencia Artificial Especifica	22
fig. 3.3.1 Modelo de funcionalidad del aprendizaje supervisado.....	26
fig. 3.3.2 Modelo de funcionalidad del aprendizaje no supervisado	26
fig. 3.3.3 Modelo de funcionalidad del aprendizaje por refuerzo	26
fig. 4.2.1 estructura de una red neuronal recuperado de (https://openwebinars.net/blog/que-son-las-redes-neuronales-y-sus-aplicaciones/)	29
fig. 4.4.1 Esquema de funcionamiento de una red neuronal recuperado de (https://www.diegocalvo.es/definicion-de-red-neuronal/).....	30
Fig. 5.1 Funcionamiento de una CNN Cortesía de “Jacobs, R. A., Jordan, M. I., Nowlan, S. J., and Hinton, G. E. (1991). Adaptive mixtures of local experts. Neural Computation”	31
Fig. 5.1.4.1 Los componentes de una capa típica de una red neuronal convolucional cortesía de “Jacobs, R. A., Jordan, M. I., Nowlan, S. J., and Hinton, G. E. (1991). Adaptive mixtures of local experts. Neural Computation”.....	34
Fig. 7.1 Diagrama de bloques del funcionamiento de una GAN	38
Fig. 7.1.1 imagen de ruido inicio vs imagen de procesamiento final de una GAN	39
Fig. 8.1.1.1Imágenes originales del Dataset con los 3 tipos de hoja con las que se trabajan los modelos	41
Fig. 8.1.1.2 imagen inicial de ruido que arroja la red GAN.....	42
Fig- 8.1.1.3 imágenes desarrolladas por la red GAN para la base de datos	43
Fig. 8.1.3.4 Imágenes originales del Dataset vs imágenes con Data Augmentation	45
Fig. 8.1.4.5Modelo con Transfer learning.....	47
Fig. 8.1.5.6 Modelo con Fine tuning.....	48
Fig. 8.4.1 Exactitud global por modelo	50
Fig. 8.4.2 Error global por modelo	51
Fig. 8. 5.1 Predicción de nuevas imágenes con modelo MOBILENET V3.....	52
Fig. 8. 6.1 Características del ordenador.....	53
Figura 8.6.2 Comportamiento de la precisión durante el entrenamiento de los modelos propuestos.	54
Figura 8.6.3 Comparativa de desempeño entre los modelos CNN Normal FT y MobileNet FT.	55
Figura 8.6.4. Comparativa de matrices de confusión entre los modelos CNN Normal con Transfer Learning y MobileNet con Fine Tuning.	56



Figura 8.6.5 Comparativa de predicciones entre los modelos CNN Normal con Transfer Learning FT y MobileNet con Fine Tuning.57



Índice de tablas

Tabla 5.1.3.1: Ejemplos de diferentes formatos de datos que pueden utilizarse con redes convolucionales	33
Tabla 8.2: Arquitectura CNN simplificada	44
Tabla 8.4 Comparación de modelos	49



1 Protocolo de investigación

1.1 Introducción

La agricultura y el cuidado de la vegetación son fundamentales para el bienestar de la sociedad, especialmente ante dos desafíos cruciales que enfrenta la humanidad en la actualidad: el cambio climático y la necesidad de alimentar a una población en constante crecimiento. Esto genera una preocupación por mejorar la productividad de los cultivos.

Hay varias formas de incrementar la productividad agrícola, pero también existen numerosas amenazas. Una de ellas son las enfermedades que pueden afectar a las plantas, poniendo en riesgo su salud, tanto las utilizadas en la agricultura como las que forman parte de los ecosistemas naturales, como bosques y selvas, vitales para la producción de oxígeno.

Las plagas de plantas representan una preocupación constante para la agricultura en México, un país con una gran diversidad de cultivos y ecosistemas. Estas plagas pueden causar daños significativos a los cultivos, reduciendo la producción y generando pérdidas económicas para los agricultores. Además, pueden afectar la seguridad alimentaria y el medio ambiente al alterar los ecosistemas naturales y los sistemas agrícolas.

Lograr identificar las enfermedades que puedan estar atacándolas de manera efectiva y temprana, esto implica la capacidad de identificar rápidamente los signos y síntomas de las enfermedades en sus etapas iniciales, antes de que se propaguen y causen daños significativos, esto es clave para llevar a cabo las medidas correctivas que logren preservar la salud de cada una de ellas con la finalidad de contribuir con satisfacción a la agricultura y las necesidades de conservación de los cultivos. Para lograr esto, es crucial enfrentar diversas amenazas, entre ellas las enfermedades que afectan a las plantas, tanto las cultivadas como las que forman parte de ecosistemas naturales. La rápida identificación de estas enfermedades es esencial para tomar medidas correctivas que salvaguarden la salud de las plantas, lo cual implica comprender la causa subyacente de la enfermedad y tomar medidas específicas para controlarla, lo que ayuda a minimizar su impacto en la producción agrícola contribuyendo así al éxito de la agricultura y a la conservación de la vegetación. (Omar Mercado, 2021).

Planteamiento del Problema

El problema fundamental radica en la necesidad de desarrollar estrategias efectivas y sostenibles para el manejo de plagas de plantas, es decir llevar una adecuada implementación de prácticas y estrategias que permitan abordar de manera efectiva y continua el problema de las plagas de plantas, teniendo en cuenta la conservación del cultivo.

Formulación del Problema

¿Cómo se puede desarrollar e implementar estrategias efectivas y sostenibles para el manejo de plagas de plantas que mitiguen los impactos negativos en la agricultura?

1.2 Antecedentes

Desarrollo de un prototipo de nodo WSN con capacidad visual para la detección temprana de plagas.

Se desarrolló un prototipo funcional de un sistema de detección temprana de plagas en cultivos frutícolas mediante tecnologías de redes de sensores inalámbricos con capacidad visual. El mismo se enmarca en el convenio FPTA entre el Grupo de Microelectrónica del IIE-FING-UdelaR, el INIA y JUMECAL para el desarrollo de soluciones de WSN para el uso en el sector agropecuario.

La necesidad surge por parte de los productores para mejorar el control de poblaciones de polillas *Cydia Pomonella* y *Cydia Molesta* en los cultivos. Tradicionalmente estos controles se realizaron mediante el conteo manual y local de los especímenes atrapados en trampas que en su piso contienen un adhesivo y feromonas de la especie. Éstas se encuentran ubicadas de manera distribuida en el terreno, colgadas de las ramas de los árboles. En este contexto, el diferencial que ofrece esta solución es obtener en tiempo real imágenes distribuidas de las trampas, para así generar una base de datos mediante una red inalámbrica integrada por nodos de costo accesible con autonomía energética de varios meses. La información obtenida permite tomar decisiones en torno al uso de pesticidas y control mediante confusión sexual, repercutiendo en una mejor calidad de las cosechas.

En el proyecto se exploró en técnicas de procesamiento de imágenes en hardware reducido y se alcanzó a implementar, por fuera del alcance estipulado inicialmente, un algoritmo de decodificación parcial de imágenes JPEG y de detección de polillas mediante un clasificador lineal Fisher. El procesamiento permite reducir la cantidad de datos a ser transmitidos, al enviar únicamente el conteo de polillas o tomar decisiones locales sobre cuándo enviar las imágenes, repercutiendo en una mayor autonomía energética y en una simplificación de la red.

El principal desafío en torno al estudio, diseño e implementación de la red consistió en la necesidad de encontrar protocolos orientados a conexión y con cables en WSN que tuvieran el menor consumo energético posible. La necesidad de una calidad de servicio orientado a conexión y con cable surge por tratarse de imágenes JPEG donde no son admisibles pérdidas, duplicados o desorden de los paquetes.

Se realizaron ensayos de caracterización del consumo de corriente y se estimó la carga consumida en cada ciclo de ejecución. Se logró obtener niveles de bajo consumo de corriente, lo que permite una autonomía energética superior a la planteada en los objetivos del proyecto.

DIAGNÓSTICO DE ENFERMEDADES EN PLANTAS

Una red neuronal es una familia de funciones inspirada por redes neuronales naturales con los principales objetivos de poder clasificar o de realizar ciertas predicciones. Una red neuronal se compone por la conexión de un diverso número de neuronas donde cada neurona recibe un número de entradas y tiene asociados el mismo número de pesos más un bias. La topología de una red del tipo MLP es una capa de entrada que se conecta a todas las entradas de datos y un número de capas intermedias que reciben como entrada

de información la salida generada por las neuronas anteriores, finalmente se tiene una capa de salida que tiene como objetivo generar una clasificación o predicción. Cada neurona realiza un ajuste de sus pesos y esa suma ponderada de los pesos multiplicados por las entradas sirve de entrada para una función de activación que en la mayoría de las ocasiones es del tipo sigmoideal (logística), tanh o relu. Se tiene que realizar este paso debido a que, si no se realiza sin importar el número de neuronas que se tengan, la salida se comportara como una gran neurona, este se debe a que cualquier combinación de funciones lineales produce una función lineal.

Se observa un estimado de la porción de presencia de puntos blancos, necrosis, clorosis y tejido sano de la planta u hoja, ubicados en la parte superior de la imagen original y en la parte superior de la imagen que corresponde a la segmentación alcanzada, se observa el tiempo de ejecución en segundos.

Con el objetivo de comparar el desempeño del diagnóstico automático de enfermedades en plantas con una inspección humana y debido a la nula disponibilidad de una base de datos que tuviera la cuantificación de la porción de planta cubierta con algunas de las enfermedades mencionadas con anterioridad, se realizó la encuesta para el público en general. Debido a que la encuesta se enfocó en un público que no posee experiencia en el diagnóstico de enfermedades en plantas, el instrumento fue diseñado para ser respondido haciendo uso de la opción múltiple con 4 opciones en total para cada planta, cada una de las opciones mostrando estimados de la porción de presencia de algunas de las enfermedades, para todos los casos una de las opciones corresponde al diagnóstico automático, por lo tanto el interés se encuentra en evaluar el porcentaje respuesta que tuvo cada una las opciones que corresponde al diagnóstico automático, se considera valida la evaluación por el público en general, puesto que es únicamente identificación de colores.

el trabajo presente realiza el diagnostico cuando las fotos son tomadas inclusive en exteriores, presentando mucho ruido lo que dificulta el uso de técnicas de segmentación tradicionales y se implementa el concepto de prominencia bayesiana, dado que el éxito de muchos de las técnicas bayesianas tienen que ver con la forma en la cual se incorpora la información a priori, se propone pasar de usar una codificación de la imagen en escala de grises o en RGB para pasar a una codificación haciendo uso de la red neuronal que esta entrenada para reconocer los colores que se consideran que a priori pueden formar parte de una planta u hoja. La propuesta de usar la red para pasar a una codificación que capture la información disponible con prioridad, es decir que codifica de una forma adecuada a aquellos pixeles que presuntamente pueden ser una planta mejoro de forma significativa el desempeño y la robustez del algoritmo propuesto, ya que en cierta manera busca asemejarse a la forma en la cual un ser humano encontrara una planta u hoja en una imagen, primero buscando los colores que considera son típicos en una planta y después decidiendo cuales de los objetos que tiene esas características son en realidad prominentes, es decir en el contexto del trabajo actual, cuales de ellos son planta.

Por lo tanto se considera que la capacidad de generalización de la propuesta generada en el trabajo ya sea mediante el uso de segmentación por GMM cuando los fondos no tienen ruido o el uso de Prominencia Bayesiana con fondos que poseen mucho ruido es una de las fortalezas principales de la propuesta y que lo posicionan como un ejemplo interesante de lo que se puede lograr cuando se logran incorporar un número reducido de supuestos que de alguna manera u otra permitan que el algoritmo propuesto reaccione de forma adecuada ante una diversidad en el tipo de plantas y en el tipo de fondos observados.

APLICACIÓN DE MÉTODOS BIO-ACÚSTICOS PARA LA DETECCIÓN TEMPRANA DEL PICUDO ROJO (*Rhynchophorus ferrugineus* Olivier)

El Picudo Rojo (*Rhynchophorus ferrugineus* Olivier) es actualmente la plaga más dañina de las palmeras. La detección temprana de la infestación resulta casi imposible, de tal forma que, cuando los efectos que producen sobre las palmeras son visibles, normalmente es demasiado tarde para recuperarlas. Las normativas estatales y autonómicas para el control del Picudo Rojo obligan al arranque y destrucción de las palmeras en las que se haya detectado la presencia de plaga.

Las técnicas bio-acústicas han demostrado su potencialidad para la detección de la actividad de distintas plagas que se desarrollan en el interior de las plantas. El Centro de AgroIngeniería del IVIA ha desarrollado un dispositivo electrónico basado en sensores acústicos, autónomo y adaptado a las características particulares de la actividad de las larvas del Picudo Rojo en el interior de las palmeras. Los ensayos han demostrado que éstas producen un sonido característico alrededor de 2.250 Hz, que no producen otras plagas, y cuya magnitud depende tanto de la población existente dentro de la palmera como de la edad de las larvas.

El dispositivo desarrollado representa una opción esperanzadora para la detección temprana del Picudo Rojo, que permitiría una acción curativa sobre el árbol. El coste económico del dispositivo, respecto a los aparatos actualmente comercializados, es bajo. Como no requiere personal especializado para interpretar la respuesta, podrá servir como base para desarrollar un sistema de monitorización autónomo capaz de cubrir un área extensa de palmeras.

1.3 Marco teórico

La identificación de plagas en hojas de cultivo constituye una tarea esencial y crítica en el ámbito agrícola, debido a que la presencia y proliferación de plagas pueden causar daños severos en las plantas y tener efectos devastadores en la productividad de los cultivos. Estas plagas pueden afectar la planta en múltiples aspectos: desde la decoloración de las hojas hasta la aparición de perforaciones, manchas y deformaciones que impactan directamente en su capacidad de realizar la fotosíntesis. Esto reduce la salud general de la planta, afecta su crecimiento y, en última instancia, disminuye el rendimiento de la cosecha. Los daños provocados por las plagas también incrementan los costos de producción, ya que los agricultores deben invertir en medidas de control como pesticidas, lo cual no solo tiene un costo económico sino también un impacto ambiental.

Tradicionalmente, la detección de plagas en las hojas de las plantas ha dependido de la inspección visual, una metodología realizada por agricultores, técnicos o especialistas en campo. Este proceso, aunque efectivo en ciertos casos, es a menudo laborioso y consume una gran cantidad de tiempo y recursos humanos. Además, la precisión del diagnóstico manual puede variar significativamente, ya que se encuentra sujeta a la experiencia y capacidad del inspector, así como a factores ambientales como la iluminación y las condiciones meteorológicas. La inspección manual también es propensa a errores humanos y tiene limitaciones en términos de escalabilidad, especialmente cuando se trata de grandes extensiones de cultivo. Esto representa una problemática significativa para la agricultura moderna, donde la eficiencia y la rapidez en el diagnóstico son factores esenciales.

La necesidad de una detección temprana, precisa y eficiente de plagas ha impulsado el desarrollo de métodos automatizados que aprovechan tecnologías avanzadas como las redes neuronales convolucionales (CNN, por sus siglas en inglés). Las CNN son un tipo especializado de red neuronal artificial diseñada para procesar datos visuales, como imágenes. A través de estas redes, es posible analizar imágenes de hojas y detectar patrones específicos que indican la presencia de plagas, permitiendo así una identificación rápida y precisa sin intervención humana directa. Las CNN se han utilizado con éxito en diversas aplicaciones de visión artificial debido a su capacidad para capturar y reconocer patrones complejos en imágenes, como texturas, formas y colores, que pueden ser indicativos de infestación por plagas.

Este marco teórico se centra en la implementación de CNN para abordar la problemática de detección de plagas en hojas de cultivo, proporcionando un fundamento conceptual y técnico para el desarrollo de un sistema de detección automatizada. La implementación de una CNN para esta tarea implica un enfoque sistemático de captura de datos, entrenamiento del modelo y optimización para asegurar que el sistema pueda distinguir entre hojas sanas y hojas con presencia de plagas. De esta manera, se busca no solo mejorar la precisión del diagnóstico, sino también facilitar un sistema que sea escalable y capaz de operar en tiempo real, permitiendo a los agricultores realizar monitoreo constante y tomar decisiones informadas de manera oportuna.

1. Problemas en la Detección de Plagas en Hojas de Cultivo

Las plagas representan un obstáculo importante para la agricultura, ya que comprometen tanto la salud de los cultivos como su productividad. En el caso de las hojas de las plantas, las plagas producen una serie de daños visibles, como manchas, perforaciones, deformaciones y otros síntomas característicos que deterioran la calidad del cultivo. Estos daños no solo afectan la apariencia de la planta, sino que también tienen un impacto directo en su capacidad para realizar la fotosíntesis de manera efectiva, lo que limita el crecimiento de la planta y, en última instancia, reduce el rendimiento de la cosecha.

Identificar las plagas en etapas tempranas es crucial para mitigar estos efectos negativos, ya que una detección oportuna permite aplicar medidas de control antes de que las plagas se propaguen y causen daños irreparables. Sin embargo, la detección manual de plagas es un proceso que consume tiempo, requiere personal capacitado y puede ser ineficiente en términos de precisión, especialmente cuando se trata de monitorear grandes extensiones de cultivo. Esto se debe a que la inspección visual depende de factores humanos, como la habilidad y experiencia del inspector, así como de factores ambientales que pueden dificultar la identificación correcta de los síntomas.

Además, los métodos tradicionales de detección de plagas no siempre son accesibles para pequeños y medianos productores, quienes podrían carecer de los recursos necesarios para contratar personal especializado o realizar inspecciones frecuentes. En este contexto, la implementación de un sistema automatizado de detección de plagas mediante CNN se presenta como una solución innovadora y rentable, que tiene el potencial de mejorar la precisión y eficiencia del proceso de detección, brindando así una herramienta de gran valor para la agricultura moderna.

1.4 Problemática

La agricultura enfrenta constantes desafíos relacionados con el control de plagas y enfermedades, los cuales impactan directamente en la productividad y calidad de los cultivos. La identificación de plagas en las hojas es crucial, ya que la infestación puede causar daños significativos, como decoloración, perforaciones y caída prematura de las hojas, limitando la capacidad de la planta para realizar la fotosíntesis. Esto no solo reduce el rendimiento del cultivo, sino que también incrementa los costos al requerir intervenciones adicionales, como la aplicación de pesticidas o técnicas de manejo más intensivas.

A pesar de la importancia del monitoreo temprano, los métodos tradicionales de inspección manual son lentos, costosos y propensos a errores humanos. Además, estos métodos suelen requerir personal capacitado y recursos significativos, lo cual no siempre es viable para productores de menor escala. Por ello, existe la necesidad de una solución automatizada y precisa que permita identificar plagas y enfermedades en etapas tempranas, optimizando los recursos y facilitando una respuesta oportuna.

Este proyecto aborda esta problemática mediante el desarrollo de un sistema de detección automatizada de plagas en hojas de cultivo utilizando redes neuronales convolucionales (CNN). Estas redes son capaces de analizar imágenes de hojas y detectar patrones específicos asociados con la presencia de plagas, permitiendo una evaluación rápida y precisa. La implementación de esta tecnología puede reducir la dependencia en el monitoreo manual, mejorar la eficiencia en la gestión de plagas y contribuir a la sostenibilidad y productividad de la agricultura moderna.

1.5 Objetivos

1.5.1 Objetivo general

Desarrollar un algoritmo efectivo y eficiente para la detección temprana de enfermedades en plantas, con el fin de mejorar la salud y productividad de los cultivos.

1.5.2 Objetivos particulares

- Investigar y recopilar información sobre las principales enfermedades que afectan a las plantas en la región objetivo.
- Analizar las técnicas y herramientas existentes para la detección de enfermedades en plantas, incluyendo métodos tradicionales y tecnológicos.
- Seleccionar y adaptar las técnicas de detección más adecuadas para desarrollar el algoritmo, considerando factores como precisión, rapidez.
- Diseñar e implementar un prototipo del algoritmo de detección temprana de enfermedades en plantas, utilizando técnicas de procesamiento de imágenes y aprendizaje automático.
- Evaluar la efectividad y precisión del algoritmo mediante pruebas en condiciones controladas y en campo, comparando sus resultados con métodos tradicionales de detección.

- Optimizar el algoritmo en función de los resultados de las pruebas, incorporando retroalimentación y mejoras continuas.

1.5.3 Objetivos específicos

- Identificar los parámetros clave para la detección temprana de enfermedades en imágenes de plantas.
- Desarrollar algoritmo de segmentación de imágenes para aislar regiones sospechosas de estar afectadas por plagas o enfermedades.
- Implementar un modelo de clasificación de enfermedades basados en características visuales de las plantas, como color, textura y forma.
- Investigar y aplicar técnicas de aprendizaje automático, como redes neuronales convolucionales, para mejorar la precisión y robustez del algoritmo.
- Realizar pruebas piloto en diferentes tipos de cultivos y condiciones ambientales para validar la eficacia del algoritmo en situaciones reales.

1.6 Metodología

La metodología propuesta abarca desde la revisión de la literatura científica hasta la implementación práctica del algoritmo en plataformas o dispositivos relevantes para su uso en el campo. A lo largo de este proceso, se seguirán los siguientes pasos para su implementación exitosa. Fig 1

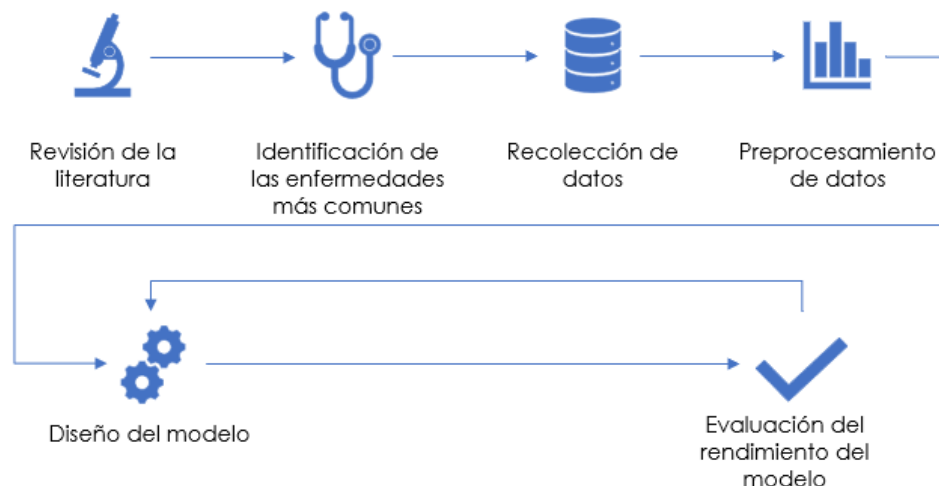


Fig. 1.6.1 Proceso Metodológico.

Hacer relación número de figura con número de capítulo- add diseño de modelo por entrenamiento, quitar ajustes e implementación – mod metodología -

- La revisión exhaustiva de la literatura científica es crucial para comprender el estado actual de la detección de enfermedades en plantas. Se analizarán investigaciones previas sobre métodos de detección, técnicas de procesamiento de imágenes y algoritmos de aprendizaje automático aplicados en la agricultura. Esto proporcionará información sobre las mejores prácticas, los desafíos comunes y las posibles áreas de mejora para el desarrollo del proyecto. Además, se identificarán las tendencias emergentes y las tecnologías innovadoras que podrían ser aplicables en el contexto específico del proyecto.
- La identificación de las enfermedades más comunes en la región objetivo es esencial. Se recopilará información detallada sobre las especies de enfermedades que afectan a los cultivos locales, así como sus ciclos de vida y factores de propagación. Esto permitirá priorizar las amenazas más relevantes y diseñar el algoritmo de detección para abordar específicamente esos problemas.
- La recolección de datos es un paso crítico para el desarrollo del algoritmo de detección temprana. Se buscará obtener un conjunto de datos diverso y representativo que contenga imágenes de plantas sanas y afectadas por una variedad de enfermedades. Se utilizarán diversas fuentes para obtener estos datos, como bases de datos públicas y estudios de campo. Es importante asegurar que los datos capturados reflejen fielmente las condiciones reales del entorno agrícola.
- El preprocesamiento de datos es necesario para garantizar la calidad y consistencia de las imágenes antes de su análisis. Se realizarán varias tareas, como ajustes de tamaño, corrección de iluminación, eliminación de ruido y normalización de color, según sea necesario. Esto ayudará a mejorar la calidad visual de las imágenes y a reducir la variabilidad entre las muestras, lo que facilitará la extracción de características por parte del algoritmo de detección.
- El diseño del algoritmo implica diseñar, desarrollar, entrenar e implementar técnicas de procesamiento de imágenes y aprendizaje automático para detectar enfermedades en las plantas. Esto puede incluir la segmentación de las regiones afectadas, la extracción de características relevantes y la clasificación de estas. Se explorarán diferentes enfoques y algoritmos, adaptándolos según las necesidades específicas del proyecto y las características de los datos disponibles. El objetivo es desarrollar un algoritmo eficiente que pueda detectar las enfermedades con precisión y rapidez.
- Basándose en los resultados de la evaluación, se realizarán ajustes y mejoras en el algoritmo para optimizar su rendimiento. Esto puede incluir la modificación de parámetros, la incorporación de nuevas características o la exploración de diferentes algoritmos y técnicas. Además, se optimizará la eficiencia computacional y el tiempo de procesamiento del algoritmo para su implementación práctica.

1.7 Justificación de los resultados esperados

La implementación de un algoritmo basado en deep learning para la detección de plagas en cultivos promete ser un catalizador significativo para mejorar la eficiencia y la sostenibilidad en la agricultura. Al aprovechar el potencial del Deep Learning, este algoritmo ofrece una capacidad única para analizar grandes conjuntos de datos con una precisión y rapidez sin precedentes. Esta característica es esencial para una detección temprana y precisa de las plagas, lo que permite una respuesta rápida y efectiva por parte de los usuarios.

La detección temprana de plagas es fundamental para reducir las pérdidas de cultivos y minimizar la necesidad de intervenciones drásticas, al identificar las plagas en sus etapas iniciales, se pueden implementar estrategias de control más específicas y menos invasivas, lo que no solo beneficia a los agricultores al reducir los costos de producción, sino que también tiene un impacto positivo en el medio ambiente y la salud humana al disminuir la exposición a productos químicos nocivos. De la misma manera la detección temprana de plagas ayuda a prevenir la propagación de enfermedades y la destrucción de cultivos, lo que garantiza la disponibilidad y calidad de los alimentos.

Este avance tecnológico impulsará la investigación científica en áreas relacionadas, como la biología de las plagas, es decir características, comportamientos y ciclos de vida de los organismos que afectan negativamente a los cultivos, y la ecología de los ecosistemas agrícolas el cual nos habla del cómo se comportan los cultivos con otros organismos en el entorno en el que se encuentren. La convergencia entre la inteligencia artificial y la agricultura no solo proporciona herramientas innovadoras para abordar los desafíos actuales, sino que también abre nuevas oportunidades para comprender mejor la complejidad de los sistemas agrícolas y desarrollar soluciones más efectivas y sostenibles a largo plazo.

2 inteligencia Artificial

La inteligencia artificial se puede definir como “la capacidad de las computadoras u ordenadores para hacer actividades que requieren la inteligencia humana” (Rouhiainen, 2018). Se podría decir que la inteligencia artificial es la capacidad de las máquinas para usar algoritmos, aprender de los datos y utilizar lo aprendido en la toma de decisiones tal y como un humano lo realizaría. Los dispositivos a diferencia de las personas no necesitan descansar para poder continuar con el trabajo y pueden analizar grandes cantidades de información a la vez lo cual los lleva a un error significativamente menor. (Rouhiainen, 2018)

2.1 Introducción

La inteligencia artificial hoy en día se ha vuelto un tema de suma importancia, debido al uso que ha tomado dentro de la sociedad, logrando llevar a cabo gran parte de las tareas dentro del área domestica como en las grandes áreas de empresas a nivel mundial. Las tecnologías de la inteligencia artificial están beneficiando de manera significativa a los humanos logrando una mayor eficiencia en casi todos los ámbitos de la vida.

Existen diversidades de situaciones en las que la inteligencia artificial puede ser aplicada, tales como:

- Reconocimiento de imágenes estáticas, clasificación y etiquetado

- Mejoras del desempeño de la estrategia algorítmica comercial
- Procesamiento eficiente y escalable de datos de pacientes
- Mantenimiento predictivo
- Detección y clasificación de objetos
- Distribución del contenido en las redes sociales
- Protección contra amenazas de seguridad cibernética.

2.2 Inteligencia Artificial General (AGI)

La Inteligencia Artificial General (AGI) se define como un sistema capaz de realizar cualquier tarea cognitiva que un ser humano pueda ejecutar, con capacidad de razonamiento, adaptación y aprendizaje autónomo. La AGI tendría la habilidad de transferir conocimientos entre distintos dominios y operar de manera independiente. Aunque su desarrollo aún no se ha logrado, su potencial impacto ha generado un debate sobre los riesgos éticos y de seguridad que podrían surgir si se implementa sin restricciones adecuadas.



fig. 2.2.1 Inteligencia Artificial General

2.2.1 Características de la inteligencia artificial general

- **Razonamiento y resolución de problemas**
Capacidad de analizar situaciones nuevas y resolver problemas sin programación previa.
- **Aprendizaje autónomo**
Puede aprender de la experiencia y mejorar su desempeño sin intervención humana.
- **Transferencia de conocimiento**
Usa aprendizajes previos en diferentes contextos, como lo hace un ser humano.
- **Creatividad e intuición**
No solo analiza datos, sino que puede generar ideas innovadoras y tomar decisiones en situaciones ambiguas.
- **Conciencia y autoevaluación**
Algunos teóricos sugieren que la AGI podría desarrollar cierto nivel de autoconciencia o metacognición.

2.2.2 Aplicaciones de la Inteligencia Artificial General

Aunque la AGI todavía no se ha alcanzado en su totalidad, su potencial para transformar distintas áreas de la vida cotidiana y la industria es significativo. Se presentan algunas aplicaciones posibles de la AGI,

- **Asistentes Inteligentes de Propósito General**

Una de las aplicaciones más esperadas de la AGI es la creación de asistentes inteligentes de propósito general, capaces de interactuar con los humanos de manera más fluida y natural que los asistentes actuales como Siri o Google Assistant. Estos asistentes no solo serían capaces de gestionar tareas simples (como la programación de citas o el control de dispositivos), sino también de realizar tareas complejas que requieren razonamiento, aprendizaje autónomo y toma de decisiones en tiempo real. Podrían ayudar en la gestión empresarial, asesoría legal, o incluso en la toma de decisiones estratégicas.

- **Innovación en Medicina y Biotecnología**

Con la capacidad de aprender y adaptarse a nuevas condiciones, la AGI podría revolucionar la medicina personalizada y la biotecnología. La AGI podría analizar grandes volúmenes de datos médicos, identificar patrones complejos entre enfermedades, tratamientos y genética, y sugerir enfoques personalizados para la atención de la salud. Además, tendría la capacidad de diseñar nuevas terapias y medicamentos de forma autónoma, optimizando la investigación biomédica y acelerando la innovación en la cura de enfermedades complejas.

- **Desarrollo de Tecnología y Ciencia Avanzada**

La AGI podría jugar un papel crucial en la investigación científica y el desarrollo tecnológico al ser capaz de realizar descubrimientos y avances por su cuenta. Desde la creación de nuevos materiales hasta la resolución de problemas complejos en áreas como la física cuántica, la ingeniería, la matemática y la astronomía, la AGI podría acelerar significativamente los procesos de investigación. Su capacidad para combinar conocimientos de diferentes disciplinas y generar hipótesis novedosas permitiría innovaciones en campos donde los humanos aún no han alcanzado el potencial total.

- **Automóviles Autónomos y Transporte**

En el campo del transporte, la AGI podría llevar los vehículos autónomos a un nivel superior, no solo gestionando la navegación y la conducción, sino también tomando decisiones complejas basadas en variables imprevistas, como el comportamiento humano o cambios repentinos en el entorno. Además, la AGI podría optimizar redes de transporte y logística a escala global, gestionando flotas de vehículos de manera más eficiente y reduciendo el impacto ambiental al optimizar las rutas y la distribución de recursos.

- **Educación Personalizada y Adaptativa**

La AGI también podría transformar el ámbito educativo. Al ser capaz de evaluar las necesidades cognitivas y emocionales de cada estudiante, la AGI podría proporcionar educación personalizada, adaptándose al ritmo, estilo de aprendizaje y áreas de interés de cada individuo. A través de un sistema de tutoría inteligente, la AGI sería capaz de enseñar no solo contenidos académicos, sino también habilidades emocionales, sociales y de resolución de problemas, mejorando la efectividad del aprendizaje.

- **Administración y Toma de Decisiones en Organizaciones**

La AGI podría ser utilizada en el ámbito de gestión empresarial y toma de decisiones. Con su capacidad para procesar y analizar grandes cantidades de datos, la AGI podría identificar oportunidades de mejora en la eficiencia de una organización, optimizar la distribución de recursos, predecir tendencias del mercado y tomar decisiones estratégicas en tiempo real. Esto permitiría a las empresas adaptarse rápidamente a cambios en el entorno económico y competitivo.

- **Interacción Humano-Máquina Avanzada**

Una de las aplicaciones más transformadoras de la AGI sería su capacidad para facilitar interacciones humano-máquina más sofisticadas y naturales. A través de interfaces de usuario avanzadas, los humanos podrían comunicarse con las máquinas no solo mediante comandos predefinidos, sino también mediante conversación natural, reconocimiento emocional y adaptabilidad a contextos cambiantes. Esto incluiría el uso de la AGI en robots sociales, terapeutas virtuales, y aplicaciones para la atención personalizada en sectores como la salud y la atención al cliente.

- **Resolución de Problemas Globales**

La AGI podría ser utilizada para abordar problemas globales complejos, como el cambio climático, la pobreza o la crisis energética. Con su capacidad para integrar y analizar datos de diversas fuentes, la AGI podría proponer soluciones innovadoras y optimizar los esfuerzos globales para resolver estos desafíos. Además, su habilidad para simular escenarios y prever impactos de diferentes decisiones podría ayudar a los gobiernos y organizaciones internacionales a tomar medidas más efectivas.

2.3 Inteligencia Artificial Específica (ANI)

La Inteligencia Artificial Específica (ANI) es un tipo de inteligencia artificial diseñada para ejecutar tareas concretas con alto rendimiento, pero sin la capacidad de transferir conocimientos entre distintos dominios de manera autónoma. A diferencia de la Inteligencia Artificial General (AGI), la ANI está programada para resolver problemas específicos dentro de un ámbito determinado. Actualmente, la ANI es la única forma de inteligencia artificial funcional, utilizada en diversas aplicaciones tecnológicas que optimizan procesos en sectores como la industria, la salud, la educación y el entretenimiento.

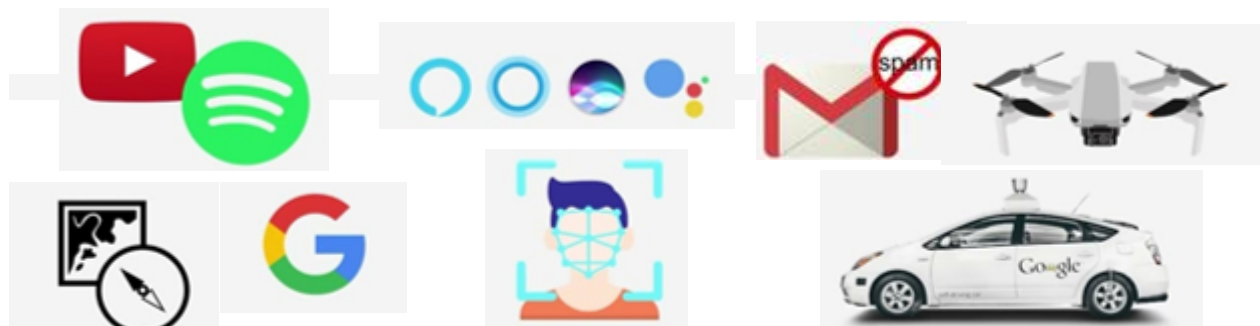


fig. 2.3.1 Tipos Inteligencia Artificial Específica

ESTRECHA GENERAL Y AMPLIA

ESTRECHA PARA TAREAS PRECIAS

AMPLIA PUNTO MEDIO ENTRE ESTRECHA Y GENERAL (ANÁLISIS DE DATOS EN EMPRESAS)

2.3.1 Características de la inteligencia artificial específica

Esta inteligencia artificial se caracteriza por su capacidad para realizar tareas concretas dentro de un dominio limitado, es decir no puede transferir conocimientos entre distintos contextos de manera autónoma ni razonar fuera de su ámbito de especialización.

- **Especialización en Tareas Concretas**
Diseñada para resolver problemas dentro de un área específica.
- **Alto Rendimiento en su Dominio**
Puede superar el desempeño humano en ciertas tareas especializadas.
- **Falta de Transferencia de Conocimiento**
No puede aplicar lo aprendido en un dominio a otro sin ser reentrenada.
- **Dependencia de Datos y Entrenamiento**
Requiere grandes volúmenes de datos y procesos de entrenamiento.
- **Ausencia de Razonamiento General**
No posee intuición, creatividad ni capacidad de análisis fuera de su programación.
- **Falta de Autoconciencia y Emociones**
No tiene conciencia de sí misma ni comprende emociones humanas.
- **Mejora Continua pero Supervisada**
Puede optimizar su desempeño, pero necesita ajustes y supervisión humana.
- **Aplicaciones en Diversos Sectores**
Se utiliza en salud, industria, entretenimiento, seguridad, entre otros.

2.3.2 Aplicaciones de la Inteligencia Artificial Específica

El desarrollo de la ANI ha permitido avances significativos en múltiples áreas. Algunos de los usos más destacados incluyen:

- **Procesamiento de Lenguaje Natural (PLN)**

El Procesamiento de Lenguaje Natural es una de las aplicaciones más extendidas de la ANI, permitiendo la interacción entre humanos y máquinas mediante el uso del lenguaje escrito y hablado. Tecnologías como asistentes virtuales (Siri, Alexa, Google Assistant) utilizan modelos de IA para interpretar comandos de voz, responder preguntas y ejecutar acciones basadas en texto. Además, estos sistemas han evolucionado para ofrecer respuestas contextuales más precisas, mejorando la experiencia del usuario.

- **Sistemas de Recomendación**

Los sistemas de recomendación utilizan algoritmos de IA para analizar patrones de comportamiento y sugerir contenido personalizado. Empresas como Netflix, Spotify y Amazon emplean esta tecnología para predecir las preferencias de los usuarios y ofrecer recomendaciones basadas en su historial de interacción. Este enfoque ha mejorado la retención de clientes y optimizado la experiencia de consumo de contenido.

- **Reconocimiento de Imágenes y Visión Artificial**

El reconocimiento de imágenes ha revolucionado múltiples industrias mediante la implementación de redes neuronales capaces de procesar y analizar datos visuales. Algunas aplicaciones clave incluyen:

Seguridad: Sistemas de reconocimiento facial utilizados en vigilancia y control de acceso.

Diagnóstico médico: IA aplicada en la detección de enfermedades a partir de imágenes de rayos X, resonancias magnéticas y tomografías computarizadas.

Automóviles autónomos: Vehículos equipados con sistemas de visión artificial para detectar obstáculos, señales de tráfico y peatones.

- **Optimización de Procesos Industriales**

La ANI ha permitido mejorar la eficiencia en la industria mediante la implementación de robots y algoritmos de automatización. Algunas aplicaciones incluyen:

Producción y manufactura: Uso de robots industriales para ensamblaje y control de calidad en líneas de producción.

Logística y distribución: Algoritmos de IA que optimizan rutas de entrega y gestión de inventarios en almacenes.

Mantenimiento predictivo: Sensores inteligentes que detectan anomalías en maquinaria y previenen fallos antes de que ocurran.

- **Modelos de Lenguaje y Generación de Contenido**

Los modelos de IA han alcanzado un alto nivel de sofisticación en la generación de texto e imágenes. Herramientas como ChatGPT y DALL·E utilizan modelos de lenguaje entrenados en grandes volúmenes de datos para crear contenido escrito y visual con una calidad comparable a la producción humana. Estas tecnologías han encontrado aplicaciones en redacción de artículos, generación de imágenes creativas, desarrollo de asistentes conversacionales y producción de contenido multimedia.

3 Machine learning

“El aprendizaje automático (Machine Learning) es uno de los enfoques principales de la inteligencia artificial. En pocas palabras, se trata de un aspecto de la informática en el que los ordenadores o las máquinas tienen la capacidad de aprender sin estar programados para ello.” (Rouhiainen, 2018)

Es una subdisciplina de la inteligencia artificial que se enfoca en el desarrollo de algoritmos y modelos matemáticos capaces de aprender de datos y realizar predicciones o tomar decisiones basadas en esos datos, sin necesidad de instrucciones explícitas para cada tarea. En lugar de depender de reglas predefinidas, el aprendizaje automático permite que las máquinas identifiquen patrones y relaciones complejas en grandes volúmenes de datos, mejorando su rendimiento a medida que se exponen a más datos, lo que las hace más eficientes y adaptables a diferentes contextos.

3.1 Introducción

El aprendizaje automático, como una rama esencial de la inteligencia artificial, ha revolucionado diversas disciplinas al permitir que las máquinas aprendan directamente de los datos sin necesidad de programación manual explícita. En lugar de seguir un conjunto fijo de reglas, los sistemas basados en aprendizaje automático son capaces de analizar grandes volúmenes de datos, identificar patrones complejos y realizar predicciones basadas en dichos patrones. Este enfoque ha sido posibilitado por avances tecnológicos en el procesamiento de datos, como el almacenamiento en la nube y la capacidad de procesamiento, que permiten entrenar modelos con vastos conjuntos de información. A medida que las máquinas acceden a más datos, mejoran su rendimiento de forma autónoma, lo que abre nuevas oportunidades en múltiples campos.

El impacto de este avance se extiende a diversas áreas, como la automatización, la medicina, el comercio, la educación y las ciencias sociales. En la medicina, por ejemplo, los modelos de aprendizaje automático permiten diagnósticos más rápidos y precisos a partir de imágenes médicas y datos clínicos, mientras que, en el comercio, se emplean para personalizar las recomendaciones de productos. En el ámbito de la educación, el aprendizaje automático contribuye a crear plataformas adaptativas que ajustan el contenido a las necesidades de cada estudiante. Estos avances demuestran cómo el aprendizaje automático está transformando sectores clave, facilitando procesos más eficientes y mejorando la toma de decisiones.

3.2 Características principales del Machine learning

El aprendizaje automático se distingue por varios aspectos clave que lo hacen adecuado para resolver una amplia variedad de problemas, estos se pueden enlistar de la siguiente manera.

- Autonomía en la toma de decisiones: A través del aprendizaje automático, las máquinas pueden analizar datos de manera independiente y tomar decisiones sin intervención humana explícita.
- Adaptabilidad: Los modelos de aprendizaje automático tienen la capacidad de adaptarse a nuevos datos y mejorar su rendimiento con el tiempo, lo que los hace más eficientes a medida que se exponen a más ejemplos.
- Procesamiento de grandes volúmenes de datos: Los algoritmos de aprendizaje automático son capaces de manejar y extraer información útil de grandes cantidades de datos, algo que sería difícil para los humanos o los sistemas tradicionales.
- Identificación de patrones complejos: Los modelos son capaces de identificar relaciones no lineales y patrones complejos dentro de los datos, lo que permite hacer predicciones con mayor precisión en comparación con métodos tradicionales.

3.3 Principales tipos de Machine Learning

Aprendizaje Supervisado: Este tipo de aprendizaje se basa en un conjunto de datos etiquetados, donde las entradas están asociadas con las salidas correctas. El modelo aprende la relación entre las entradas y salidas para predecir resultados para nuevos datos. Este enfoque es ampliamente utilizado en tareas como

- clasificación: identificación de correos electrónicos como spam o no spam.
- regresión: predicción de valores numéricos como precios de viviendas



fig. 3.3.1 Modelo de funcionalidad del aprendizaje supervisado

Aprendizaje No Supervisado: En el aprendizaje no supervisado, los datos no están etiquetados, y el modelo debe identificar patrones y estructuras ocultas dentro de los datos. Técnicas comunes en este tipo de aprendizaje incluyen la agrupación, donde los datos se agrupan en categorías similares, y la reducción de dimensionalidad, que busca simplificar datos complejos manteniendo sus características esenciales.



fig. 3.3.2 Modelo de funcionalidad del aprendizaje no supervisado

Aprendizaje por Refuerzo: Este enfoque se basa en la interacción entre un agente y su entorno, donde el agente realiza acciones y recibe recompensas o penalizaciones en función de su comportamiento. El objetivo es que el agente aprenda a tomar decisiones que maximicen las recompensas acumuladas. Este tipo de aprendizaje es utilizado en aplicaciones como la robótica, los videojuegos y la optimización de procesos.



fig. 3.3.3 Modelo de funcionalidad del aprendizaje por refuerzo

3.4 Tipos de datos

Datos estructurados

Son aquellos que se organizan de manera clara y definida en un formato predecible, generalmente en tablas con filas y columnas, lo que facilita su almacenamiento, procesamiento y análisis. Este tipo de datos es altamente organizado y es comúnmente gestionado por bases de datos relacionales.

Características:

- **Formato:** Los datos estructurados suelen estar organizados en bases de datos y se representan en tablas con filas y columnas.
- **Accesibilidad:** Son fácilmente accesibles y analizables debido a su organización predecible.

Ejemplos:

- Datos financieros, como los registros de ventas, cuentas bancarias y transacciones.
- Información de inventarios, como el nombre del producto, la cantidad disponible y el precio.
- Registros de clientes, como el nombre, la dirección y el número de teléfono.

Ventajas:

- Fácil de almacenar y consultar mediante lenguajes de bases de datos estructuradas como SQL.
- Adecuado para análisis con herramientas y algoritmos tradicionales de datos tabulares.

Datos no estructurados

Los datos no estructurados, en cambio, no siguen un formato predefinido y son más difíciles de organizar y analizar de manera directa. Estos datos no se almacenan en una estructura de base de datos tradicional y no se pueden ingresar fácilmente en tablas o bases de datos relacionales sin procesamiento adicional.

Características:

- **Formato:** Los datos no estructurados pueden estar en formatos como texto libre, imágenes, videos, correos electrónicos, grabaciones de audio, etc.
- **Accesibilidad:** Son más difíciles de procesar y analizar sin utilizar técnicas especializadas de análisis de datos, como el procesamiento de lenguaje natural (PLN) o la visión por computadora.

Algunos de los más comunes suelen ser:

- Texto: Artículos de noticias, publicaciones en redes sociales, correos electrónicos, comentarios en línea.
- Imágenes y Videos: Fotografías, grabaciones de video, imágenes satelitales.
- Audio: Grabaciones de voz, podcasts, grabaciones de llamadas telefónicas.
- Documentos: PDFs, presentaciones, informes en formato Word o Excel.

Ventajas:

- Contienen información rica y detallada que no está limitada por una estructura rígida.
- Ofrecen una variedad más amplia de aplicaciones, como análisis de sentimientos, clasificación de imágenes y reconocimiento de voz.

Desafíos:

- Requieren técnicas y herramientas avanzadas (como PLN, procesamiento de imágenes, análisis de voz) para extraer valor de ellos.
- Son más difíciles de almacenar, organizar y buscar debido a la falta de estructura.

4 Redes neuronales

Las redes neuronales artificiales (RNA) son un modelo computacional inspirado en el funcionamiento del cerebro humano “Una red neuronal multicapa es un aproximador universal. En principio, una red de este tipo es capaz de aprender cualquier cosa, siempre que la red sea lo suficientemente grande como para representar las peculiaridades de la función que pretendemos aprender. Si disponemos de un número suficiente de ejemplos de entrenamiento, una red multicapa de la capacidad suficiente será capaz de construir un modelo de los datos con los que la entrenamos.” (Berzal, 2018)

4.1 Introducción

El aprendizaje automático ha revolucionado la forma en que las máquinas pueden procesar, analizar y tomar decisiones a partir de datos. Dentro de este campo, las redes neuronales artificiales (RNA) han emergido como una de las herramientas más poderosas para resolver problemas complejos en diversas áreas del conocimiento. Inspiradas en el funcionamiento del cerebro humano, estas redes están compuestas por múltiples capas de neuronas interconectadas que aprenden de manera iterativa mediante algoritmos de optimización.

El desarrollo de redes neuronales ha permitido avances significativos en disciplinas como la visión por computadora, el procesamiento del lenguaje natural y la automatización industrial. Gracias a su capacidad para identificar patrones complejos en grandes volúmenes de datos, han sido aplicadas en sectores como la medicina, las finanzas, la seguridad y el entretenimiento. En este contexto, la presente investigación analiza la estructura, funcionamiento y aplicaciones de las redes neuronales artificiales, destacando su impacto en la transformación digital y la inteligencia artificial.

4.2 Concepto y estructura de una red neuronal

Una red neuronal artificial está compuesta por nodos o "neuronas" organizadas en capas interconectadas. Estas capas incluyen:

- Capa de entrada: Recibe los datos iniciales y los transmite a la siguiente capa.
- Capas ocultas: Procesan la información mediante pesos y funciones de activación para extraer patrones y relaciones.
- Capa de salida: Genera el resultado final, que puede ser una clasificación, una predicción numérica o una transformación de los datos.

Cada conexión entre neuronas tiene un peso, que se ajusta a través del proceso de entrenamiento para mejorar la precisión del modelo.

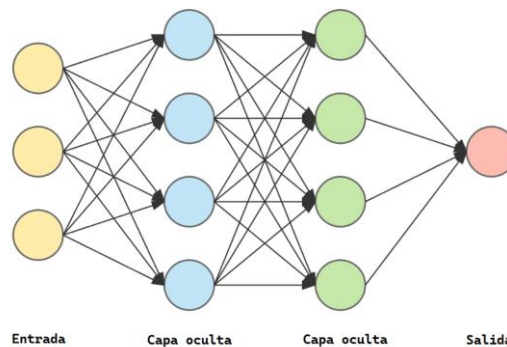


fig. 4.2.1 estructura de una red neuronal recuperado de (<https://openwebinars.net/blog/que-son-las-redes-neuronales-y-sus-aplicaciones/>)

4.3 Tipos de redes neuronales

Existen diferentes tipos de redes neuronales diseñadas para resolver problemas específicos:

- Redes Neuronales Artificiales (ANN, Artificial Neural Networks): Son las más básicas y están compuestas por capas densamente conectadas. Se utilizan en clasificación y regresión.
- Redes Neuronales Convolucionales (CNN, Convolutional Neural Networks): Son utilizadas principalmente en visión por computadora, procesamiento de imágenes y videos. Incorporan capas de convolución para detectar patrones espaciales.
- Redes Neuronales Recurrentes (RNN, Recurrent Neural Networks): Diseñadas para procesar datos secuenciales, como texto y series temporales. Permiten que la información fluya en bucles dentro de la red.
- Redes Neuronales Generativas (GAN, Generative Adversarial Networks): Se utilizan para generar nuevos datos a partir de un conjunto de entrenamiento, aplicándose en imágenes sintéticas y arte digital.

4.4 Funcionamiento de una red neuronal

El entrenamiento de una red neuronal se basa en un proceso iterativo de ajuste de pesos llamado retropropagación (backpropagation), que consta de los siguientes pasos:

- Propagación hacia adelante (forward propagation):
 - Los datos ingresan a la red desde la capa de entrada.
 - Cada neurona realiza una operación matemática aplicando pesos y una función de activación.
 - La información avanza a través de las capas ocultas hasta la capa de salida.
- Cálculo del error:
 - Se compara la salida de la red con la salida esperada mediante una función de pérdida.
- Retropropagación del error (backpropagation):
 - Se ajustan los pesos de la red en sentido inverso para minimizar el error.
 - Se utiliza el descenso del gradiente para optimizar los parámetros de la red.
- Repetición del proceso hasta que el error sea lo suficientemente bajo o la red alcance un número máximo de iteraciones.

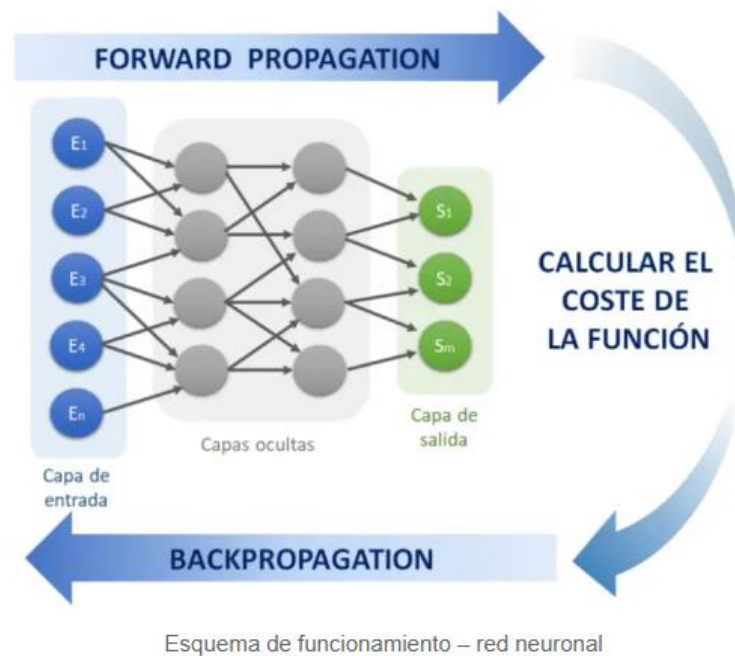


fig. 4.4.1 Esquema de funcionamiento de una red neuronal recuperado de (<https://www.diegocalvo.es/definicion-de-red-neuronal/>)

5 Redes Neuronales convolucionales

De acuerdo con Nowlan, S. J., quien cita (LeCun, 1989), Las redes convolucionales (LeCun, 1989), también conocidas como redes neuronales convolucionales o CNN, son un tipo especializado de red neuronal para procesar datos que tiene una topología conocida similar a una cuadrícula. Algunos ejemplos son los datos de series temporales, que pueden considerarse como una cuadrícula 1D que toma muestras a intervalos de tiempo regulares, y los datos de imagen, que pueden considerarse como una cuadrícula 2D de píxeles.

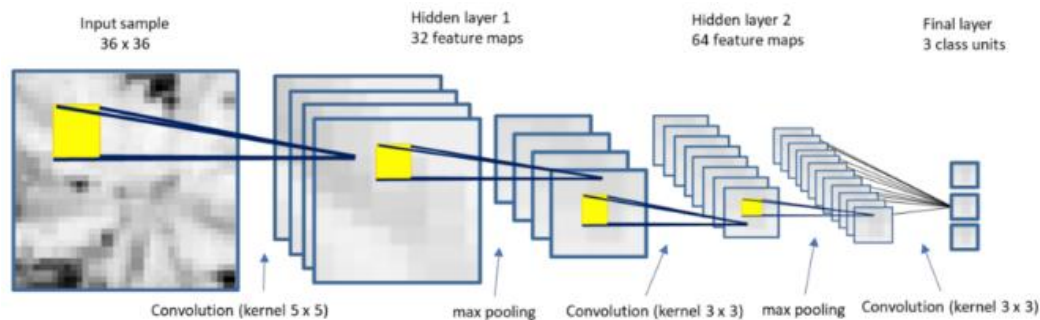


Fig. 5.1 Funcionamiento de una CNN Cortesía de "Jacobs, R. A., Jordan, M. I., Nowlan, S. J., and Hinton, G. E. (1991). Adaptive mixtures of local experts. Neural Computation"

Las redes convolucionales han tenido un éxito tremendo en aplicaciones prácticas. El nombre "red neuronal convolucional" indica que la red emplea una operación matemática llamada convolución. La convolución es un tipo especializado de operación lineal. Las redes convolucionales son simplemente redes neuronales que utilizan la convolución en lugar de la multiplicación general de matrices en al menos una de sus capas. (Jacobs, R. A., 1991)

5.1 Introducción

Las redes convolucionales han jugado un papel importante en la historia del aprendizaje profundo. Son un ejemplo clave de una aplicación exitosa de los conocimientos obtenidos mediante el estudio del cerebro a las aplicaciones de aprendizaje automático. También fueron algunos de los primeros modelos profundos en funcionar bien, mucho antes de que los modelos profundos arbitrarios se consideraran viables. (LeCun, 1989)

Las redes neuronales convolucionales, o CNN (Convolutional Neural Networks), son un tipo de red neuronal que ha demostrado ser particularmente eficaz en la identificación de patrones complejos en imágenes. Las CNN son ampliamente utilizadas en tareas de visión por computadora, como el reconocimiento de objetos, clasificación de imágenes y segmentación de imágenes, debido a su capacidad para capturar características visuales en diferentes niveles de abstracción.

5.1.1 Arquitectura de una CNN

Una CNN típica está compuesta por una secuencia de capas:

- **Capas convolucionales:** En estas capas, se aplican filtros (kernels) a la imagen de entrada para extraer características, como bordes, texturas y formas. Estos filtros permiten capturar patrones específicos que pueden estar asociados con signos de plagas, como manchas o deformaciones.
- **Capas de agrupación (pooling):** Estas capas reducen la dimensionalidad de las características extraídas, minimizando el ruido y facilitando el enfoque en los patrones más importantes de la imagen. Una técnica común es el *max-pooling*, que selecciona el valor máximo en cada región, manteniendo solo la información más relevante.
- **Capas de activación:** Funcionan como funciones de activación que introducen no linealidades en la red, permitiendo que la CNN aprenda representaciones complejas de los datos.
- **Capas completamente conectadas:** Estas capas están al final de la red y permiten que la CNN aprenda una combinación de las características extraídas para clasificar o predecir la presencia de plagas.

5.1.2 Operación de Convolución

La operación de convolución es el núcleo de las CNN. Consiste en aplicar un filtro a una región específica de la imagen de entrada, generando un mapa de características que destaca ciertas propiedades visuales en la imagen. La convolución es ideal para la detección de patrones localizados, como las manchas causadas por plagas en hojas de cultivo. Al aplicar múltiples filtros, la CNN puede capturar una variedad de detalles y texturas que son relevantes para identificar tipos específicos de plagas.

Además, las CNN aprovechan tres ideas clave:

- **Interacciones dispersas:** No todas las neuronas están conectadas entre sí, lo que permite enfocarse en regiones específicas de la imagen.
- **Compartición de parámetros:** Los mismos filtros se reutilizan en toda la imagen, reduciendo el número de parámetros y mejorando la eficiencia.
- **Representaciones equivariantes:** Los patrones detectados en una parte de la imagen pueden ser reconocidos en otras partes de la imagen, independientemente de su posición.

5.1.3 Tipos de datos

Los datos utilizados con una red convolucional suelen constar de varios canales, siendo cada canal la observación de una cantidad diferente en algún momento del espacio o del tiempo. Consulte la tabla 1 para ver ejemplos de tipos de datos con diferentes dimensionalidades y número de canales.

Los datos de entrenamiento y prueba tienen las mismas dimensiones espaciales. Una ventaja de las redes convolucionales es que también pueden procesar entradas con diferentes extensiones espaciales. Estos tipos de entrada simplemente no pueden ser representados por las redes neuronales tradicionales basadas en la multiplicación de matrices. Esto proporciona una razón convincente para usar redes convolucionales incluso cuando el costo computacional y el sobreajuste no son problemas significativos.

	Single channel	Multi-channel
1-D	Audio waveform: The axis we convolve over corresponds to time. We discretize time and measure the amplitude of the waveform once per time step.	Skeleton animation data: Animations of 3-D computer-rendered characters are generated by altering the pose of a "skeleton" over time. At each point in time, the pose of the character is described by a specification of the angles of each of the joints in the character's skeleton. Each channel in the data we feed to the convolutional model represents the angle about one axis of one joint.
2-D	Audio data that has been preprocessed with a Fourier transform: We can transform the audio waveform into a 2D tensor with different rows corresponding to different frequencies and different columns corresponding to different points in time. Using convolution in the time makes the model equivariant to shifts in time. Using convolution across the frequency axis makes the model equivariant to frequency, so that the same melody played in a different octave produces the same representation but at a different height in the network's output.	Color image data: One channel contains the red pixels, one the green pixels, and one the blue pixels. The convolution kernel moves over both the horizontal and vertical axes of the image, conferring translation equivariance in both directions.
3-D	Volumetric data: A common source of this kind of data is medical imaging technology, such as CT scans.	Color video data: One axis corresponds to time, one to the height of the video frame, and one to the width of the video frame.

Tabla 5.1.3.1: Ejemplos de diferentes formatos de datos que pueden utilizarse con redes convolucionales

5.1.4 Algoritmos de Agrupación (Pooling)

El *pooling* reduce la resolución del mapa de características, disminuyendo el número de parámetros y la carga computacional. Esta etapa contribuye a que la CNN ignore pequeñas variaciones en la posición y el tamaño de los patrones, permitiendo que el modelo sea más robusto ante cambios en la apariencia de las plagas. Los métodos de *max-pooling* y *average-pooling* son los más comunes, y cada uno tiene su propio enfoque para consolidar la información relevante de los mapas de características

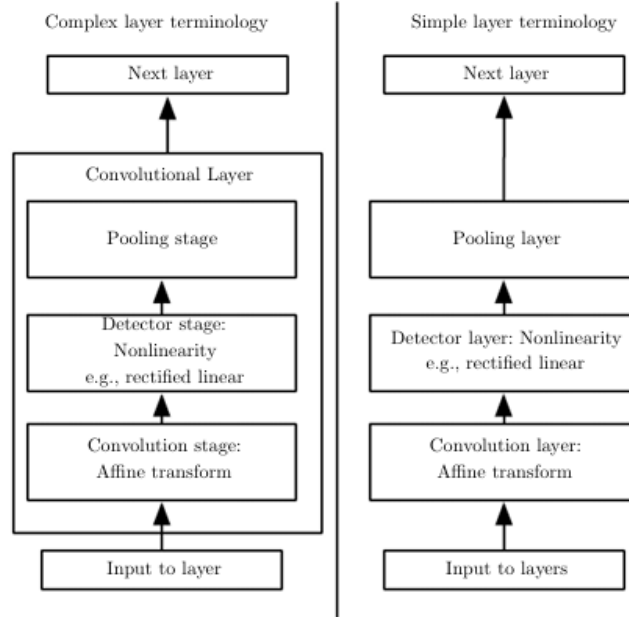


Fig. 5.1.4.1 Los componentes de una capa típica de una red neuronal convolucional cortesía de “Jacobs, R. A., Jordan, M. I., Nowlan, S. J., and Hinton, G. E. (1991). Adaptive mixtures of local experts. *Neural Computation*”

5.1.5 Entrenamiento y Optimización de la CNN

El proceso de entrenamiento de una CNN implica mostrarle a la red una gran cantidad de imágenes de hojas, tanto con plagas como sin plagas, para que aprenda a identificar patrones relevantes. Durante el entrenamiento, el modelo ajusta sus parámetros para minimizar la diferencia entre sus predicciones y la realidad (etiquetas de plagas). Esto se realiza mediante algoritmos de optimización como el *descenso de gradiente* y técnicas de regularización para evitar el sobreajuste.

Además, para mejorar la precisión y eficiencia del modelo, se emplean técnicas de *data augmentation*, como la rotación, cambio de escala y ajuste de iluminación en las imágenes de hojas. Esto permite que la CNN generalice mejor a nuevas imágenes y mejore su rendimiento en diferentes condiciones de iluminación y perspectiva.

5.1.6 Algoritmos de Convolución Eficientes

Las aplicaciones de redes convolucionales modernas a menudo involucran redes que contienen más de un millón de unidades. Las implementaciones potentes que explotan los recursos de computación paralela, son esenciales. Sin embargo, en muchos casos también es posible acelerar la convolución seleccionando un algoritmo de convolución adecuado.

Dado que el procesamiento de imágenes de alta resolución puede ser costoso en términos de recursos computacionales, la implementación de algoritmos eficientes es crucial. El uso de técnicas como la *convolución separable* permite reducir el tiempo de procesamiento al descomponer un filtro multidimensional en varios filtros unidimensionales, disminuyendo así los recursos necesarios. Estas optimizaciones son esenciales para lograr un sistema de detección de plagas en tiempo real.

5.1.7 Aplicación de CNN para la Detección de Plagas en Hojas de Cultivo

La implementación de una CNN para la detección de plagas en hojas de cultivo implica el desarrollo de un sistema que pueda analizar imágenes capturadas en campo y clasificar automáticamente las hojas según la presencia o ausencia de plagas. Este sistema puede aplicarse en dispositivos móviles o drones, permitiendo un monitoreo rápido y accesible. Al automatizar la detección de plagas, los agricultores pueden reducir la dependencia en inspecciones manuales, mejorando la eficiencia del manejo de plagas y permitiendo una respuesta más oportuna para mitigar daños en los cultivos.

6 Transfer learning

El aprendizaje por transferencia es un método de aprendizaje automático en el que un modelo desarrollado para una tarea se reutiliza como punto de partida para un modelo en una segunda tarea.

Es un enfoque popular en el aprendizaje profundo en el que se utilizan modelos previamente entrenados como punto de partida en tareas de visión por computadora y procesamiento del lenguaje natural, dados los vastos recursos computacionales y de tiempo necesarios para desarrollar modelos de redes neuronales en estos problemas y los enormes avances en habilidades que brindan en problemas relacionados.

técnica en el campo del aprendizaje automático y la inteligencia artificial que permite reutilizar el conocimiento adquirido por un modelo entrenado en una tarea específica para resolver otra tarea relacionada. En lugar de entrenar un modelo desde cero, se aprovechan los parámetros y características aprendidas de un modelo preentrenado en un conjunto de datos grande y se adaptan para un nuevo problema, lo que ahorra tiempo y recursos.

6.1 Características Clave del Transfer Learning

- Reutilización de modelos preentrenados: Se utilizan modelos que han sido entrenados en grandes bases de datos (como ImageNet para imágenes o GPT para texto) como punto de partida.
- Entrenamiento más rápido y eficiente: Como el modelo ya contiene características útiles aprendidas de datos similares, la cantidad de datos y el tiempo necesarios para entrenar el nuevo modelo disminuyen.
- Aplicable en casos de datos limitados: Es particularmente útil cuando se tiene un conjunto de datos limitado en la nueva tarea, ya que se aprovechan las características ya aprendidas en un conjunto de datos más grande.

6.1.1 Reutilización de modelos

La reutilización de modelos preentrenados es un aspecto fundamental del Transfer Learning y se refiere a la práctica de tomar un modelo que ha sido previamente entrenado en un conjunto de datos grande y diverso y adaptarlo para una tarea específica que puede tener menos datos disponibles. Este enfoque permite aprovechar el conocimiento ya adquirido por el modelo para mejorar el rendimiento en la nueva tarea.

Cómo Funciona la Reutilización de Modelos Preentrenados

- Selección de un Modelo Preentrenado:
 - Se elige un modelo que ha sido entrenado en una tarea relacionada. Por ejemplo, en visión por computadora, se pueden usar modelos como VGG16, ResNet, Inception o EfficientNet, que han sido entrenados en datasets masivos como ImageNet.
- Extracción de Características:
 - Las capas iniciales del modelo preentrenado se utilizan para extraer características de los datos nuevos. Estas capas suelen aprender representaciones generales que son útiles para una variedad de tareas.
- Congelamiento y Ajuste de Capas:
 - Dependiendo de la cantidad de datos disponibles para la nueva tarea, algunas capas del modelo pueden "congelarse" (mantenerse sin cambios) para conservar las características aprendidas.
 - Las últimas capas del modelo se reemplazan o se ajustan (fine-tuning) para adaptarse a la nueva tarea, permitiendo que el modelo aprenda características específicas del nuevo dominio.

- Entrenamiento del Modelo:
 - El modelo se entrena utilizando los datos de la nueva tarea, ajustando solo las capas que no han sido congeladas. Esto permite que el modelo se especialice en la nueva tarea sin perder el conocimiento general adquirido previamente.

6.1.3 Tipos de modelos

- Transfer Learning Basado en Características (Feature-based Transfer Learning):

En este tipo, se utilizan las características extraídas del modelo preentrenado para resolver un nuevo problema. Las características extraídas se utilizan como entradas a un nuevo modelo.

Se puede hacer extracción de características congelando las capas del modelo preentrenado y entrenando solo el clasificador final.

- Transfer Learning Basado en Parámetros (Parameter-based Transfer Learning):

El modelo preentrenado se adapta a un nuevo dominio ajustando sus parámetros, es decir, se realiza un fine-tuning. Aquí, las capas preentrenadas no se congelan completamente, sino que se permite que los pesos del modelo se ajusten a los datos del nuevo problema.

- Transfer Learning Basado en Relación (Relational-based Transfer Learning):

Se utiliza cuando hay una relación entre los datos fuente y los datos objetivo. Aquí, el aprendizaje transferido se centra en encontrar relaciones entre instancias de diferentes dominios que puedan ayudar a realizar predicciones.

- Domain Adaptation:

Se emplea cuando el dominio de origen (datos preentrenados) y el dominio de destino (nuevo conjunto de datos) son diferentes pero se intenta que el modelo se adapte a las características del nuevo dominio.

- Multitask Learning:

Implica el entrenamiento simultáneo en múltiples tareas relacionadas para mejorar el rendimiento de todas ellas. Aquí, las tareas comparten parámetros o capas de la red para facilitar la transferencia de conocimiento.

- Zero-Shot Learning (ZSL):

Permite al modelo generalizar a clases no vistas durante el entrenamiento. Se basa en aprender a partir de atributos semánticos que describen las nuevas clases, en lugar de usar ejemplos específicos.

- Few-Shot Learning:

Se refiere a cuando el modelo es capaz de generalizar y aprender a partir de un número muy pequeño de ejemplos. Utiliza un modelo preentrenado y lo ajusta con pocos datos etiquetados del nuevo problema.

- Transductive Transfer Learning:

La transferencia se realiza con la condición de que el dominio objetivo y el dominio de origen compartan la misma distribución de características. Sin embargo, las etiquetas en el dominio de origen y destino no necesariamente son las mismas.

7 Redes Generativas Antagónicas (GANs)

Una red generativa antagónica (GAN) es un tipo de arquitectura dentro del aprendizaje profundo que trabaja haciendo que dos redes neuronales compitan entre sí. Su objetivo es crear datos nuevos que parezcan reales, utilizando como referencia un conjunto de datos existente, es decir, puede crear imágenes nuevas basadas en una colección de fotos reales, o generar música original a partir de una base de canciones ya existentes.

Se les denomina “antagónica” debido a que involucra dos redes con funciones opuestas:

- Una red, llamada generador, toma muestras de datos y las transforma para crear algo nuevo.
- La otra, llamada discriminador, intenta detectar si esos datos fueron generados o si provienen del conjunto original.

Ambas redes se entrenan juntas: el generador mejora continuamente hasta que el discriminador ya no puede distinguir lo falso de lo real. Así se logran datos artificiales altamente realistas.

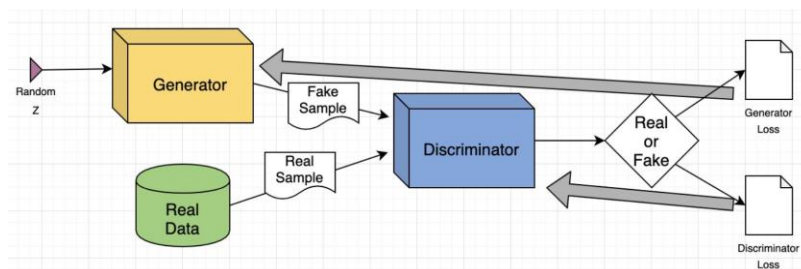


Fig. 7.1 Diagrama de bloques del funcionamiento de una GAN

7.1 Características de las GANs

Se distinguen por estar compuestas por dos redes neuronales que trabajan en conjunto, pero con objetivos opuestos, una llamada generador, encargada de crear datos falsos que imitan los reales, y otra llamada discriminador, cuya tarea es detectar si los datos que recibe son reales o han sido generados. Esta estructura de competencia entre ambas redes es lo que impulsa el aprendizaje de la GAN.

Una de las características más notables es su entrenamiento adversarial, ya que durante el proceso, como ya se mencionó anteriormente, el generador intenta engañar al discriminador con datos falsos cada vez más realistas, mientras que el discriminador se entrena para volverse mejor detectando esas falsificaciones. Esta lucha constante lleva al generador a perfeccionar su capacidad para crear datos tan realistas que incluso el discriminador no puede distinguirlos de los verdaderos.

Estas redes cuentan con la capacidad para generar datos nuevos y realistas a partir de un conjunto de entrenamiento, es decir, pueden crear imágenes de personas que no existen, generar música nueva o producir contenido que nunca se ha visto antes, pero que parece auténticos, además, las GANs suelen trabajar con datos no etiquetados, lo que significa que no necesitan ejemplos clasificados para entrenarse, a diferencia de otros modelos supervisados. Esto las hace especialmente útiles cuando se dispone de grandes volúmenes de datos sin información adicional.

El generador de una GAN toma como entrada un vector de ruido aleatorio y, a partir de él, produce datos con características similares a los del conjunto original. Esta capacidad de transformar ruido en información realista es una de las razones por las que las GANs son tan sorprendentes.

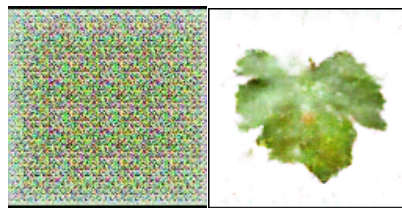


Fig. 7.1.1 imagen de ruido inicio vs imagen de procesamiento final de una GAN

Sin embargo, también presentan desafíos. Son modelos difíciles de entrenar, ya que el equilibrio entre el generador y el discriminador es delicado. Si uno aprende más rápido que el otro, el entrenamiento puede volverse inestable. Uno de los problemas más comunes es el llamado colapso de modo, donde el generador produce siempre el mismo tipo de salida sin variación.

7.2 Modelos de GANs

Existen varios tipos de Redes Generativas Antagónicas (GANs), cada uno diseñado para resolver problemas o limitaciones específicas del modelo original.

- GAN estándar (Vanilla GAN)

Es la forma básica (propuesta por Ian Goodfellow en 2014). Utiliza redes completamente conectadas y se entrena con una función de pérdida de entropía cruzada binaria. Su principal función está basada para datos simples, pero tiene problemas de estabilidad.

- Deep Convolutional GAN (DCGAN)

Utiliza redes convolucionales profundas en lugar de redes totalmente conectadas. Es ideal para generar imágenes realistas. Introduce buenas prácticas como el uso de Batch Normalization y ReLU/LeakyReLU.

- Conditional GAN (cGAN)

Permite controlar la salida del generador mediante etiquetas. Puede generar imágenes de perros o gatos dependiendo del valor de entrada. Ideal para tareas donde necesitas generar datos de una clase específica.

- CycleGAN

Diseñada para tareas de traducción de imágenes entre dominios, sin necesidad de pares exactos, como convertir fotos en pinturas o cambiar el estilo de una imagen de verano a invierno. Usa dos generadores y dos discriminadores.

- StyleGAN / StyleGAN2 / StyleGAN3

Desarrollada por NVIDIA, es una de las GANs más potentes. Puede generar rostros humanos ultra realistas. Introduce una arquitectura basada en el "estilo", lo que permite controlar aspectos específicos como expresiones, edad o peinados.

- Pix2Pix

Funciona con pares de imágenes como contornos y fotos reales. Se usa para traducción de imagen a imagen, donde hay correspondencia clara entre la entrada y la salida, como transformar un boceto en una imagen realista.

- InfoGAN

Introduce una variable latente interpretable, lo que permite controlar y entender lo que el generador está haciendo. Ayuda a descubrir estructuras internas en los datos, como rotación, escala, etc.

- WGAN (Wasserstein GAN)

Mejora la estabilidad del entrenamiento usando la distancia de Wasserstein como función de pérdida. Es menos propensa a colapsos y permite una métrica más informativa durante el entrenamiento.

- SRGAN (Super Resolution GAN)

Diseñada para aumentar la resolución de imágenes borrosas o de baja calidad, generando detalles realistas. Se utiliza en tareas de restauración de imagen y mejora visual.

8 Metodología y Resultados

8.1 Clasificación con Base de Datos

La metodología que se propone en esta sección puede ser planteada en 6 etapas: Adquisición y preprocesamiento de imágenes, Entrenamiento, Validación, Análisis comparativo de modelos, Predicción de nuevas imágenes y Optimización. El objetivo es obtener un modelo que permita clasificar entre hojas sanas y hojas con enfermedades sob las mismas plantas de vid empleando un conjunto de imágenes de hojas de vid.

8.1.1 Etapa 1: Adquisición y preprocesamiento de imágenes

El Dataset con el que se trabajo fue directamente tomado de un campo con este tipo de cultivo, en el cual se pudieron encontrar las diversas muestras sobre las que se trabajaron, este dataset está compuesto de 3 tipos de imágenes: hoja con daño Black rot. Hoja con daño de plagas vivas y hojas sanas.



Fig. 8.1.1.1 Imágenes originales del Dataset con los 3 tipos de hoja con las que se trabajan los modelos

Este Dataset contiene la cantidad de 3.085 imágenes de entrenamiento con dimensiones de 128x128 y como anteriormente se comentó están clasificadas en: “hoja black rot”, “hoja plaga viva” y “hoja sana” con 929, 984 y 1,172 imágenes por clasificación respectivamente

En esta etapa se trabajó con Redes Generativas Antagónicas, conocidas como GAN, *Generative Adversarial Networks*, con el objetivo de generar imágenes sintéticas que permitieran ampliar y fortalecer la base de datos existente. Este tipo de redes es ampliamente utilizado en tareas donde se requiere la generación de nuevos datos visuales a partir de ejemplos reales

Las GAN funcionan mediante dos redes neuronales que se entrenan simultáneamente y de forma competitiva: una red generadora y una red discriminadora.

- La red generadora se encarga de crear imágenes sintéticas a partir de ruido aleatorio, tratando de que estas imágenes se parezcan lo más posible a las reales.
- Por otro lado, la red discriminadora recibe tanto imágenes reales del conjunto de datos como las generadas por el generador, y su tarea es distinguir si una imagen es real o falsa.

Ambas redes se enfrentan en un proceso de aprendizaje continuo:

- El generador mejora sus imágenes para engañar al discriminador.
- El discriminador mejora su capacidad para identificar imágenes falsas.

Este ciclo de retroalimentación se repite hasta que el generador logra producir imágenes tan convincentes que el discriminador no puede distinguir con certeza si son reales o generadas. De esta manera se lograron agregar mas imágenes a la base de datos, a continuación, se muestran algunas de las imágenes creadas por esta red, así como el proceso que va siguiendo la misma

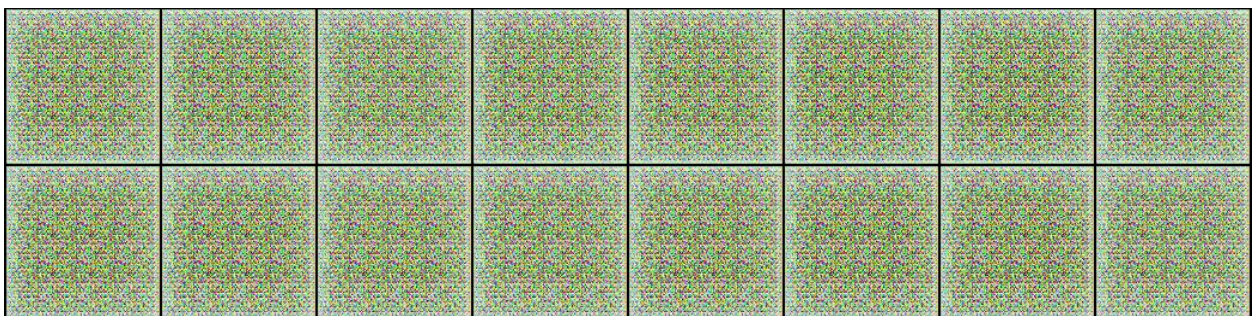


Fig. 8.1.1.2 imagen inicial de ruido que arroja la red GAN

Como se mencionó anteriormente, al momento de iniciar el entrenamiento de una Red Generativa Antagónica (GAN), tanto la red generadora como la discriminadora se

encuentran en un estado inicial sin haber aprendido aún las características relevantes del conjunto de datos. Debido a esto, en las primeras etapas del entrenamiento, las imágenes generadas por la red generadora suelen contener un alto nivel de ruido, formas indefinidas o artefactos visuales sin estructura coherente. Esto es completamente normal, ya que el generador está empezando a aprender cómo transformar un vector de ruido en una imagen que pueda parecerse a una muestra real.

Del mismo modo, la red discriminadora en estas primeras fases tampoco tiene suficiente capacidad para identificar de manera precisa si una imagen es real o generada. Por lo tanto, su criterio para discriminar es todavía rudimentario, lo que se traduce en una clasificación poco confiable tanto para imágenes reales como sintéticas.

Durante el proceso de entrenamiento, ambas redes comienzan a mejorar progresivamente tal cual podemos observar en la figura 13 que se muestra a continuación.

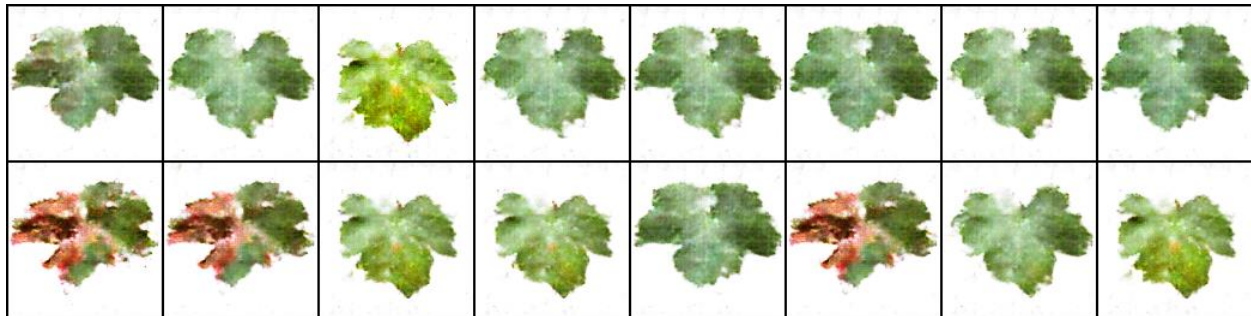


Fig- 8.1.1.3 imágenes desarrolladas por la red GAN para la base de datos

Como se puede apreciar en la imagen anterior existe una gran mejora sobre las imágenes desarrolladas por esta red, tanto así que se utilizaron en la base de datos para robustecer la misma y poder trabajar con una cantidad mas grande de imágenes.

8.1.2 Etapa 2: Entrenamiento de modelos

La arquitectura que se trabajó sobre el modelo simple de CNN está compuesta de 4 capas convolucionales seguidas de las capas MaxPooling, esto reduce las dimensiones de las características sobre las imágenes, a este paso le sigue su propia capa de activación ReLU. Estas características se aplanaron mediante una capa de aplanado Flatten.

La clasificación se realizó utilizando dos capa completamente conectadas y una capa de salida, con 512 y 3 neuronas respectivamente. Las funciones de activación fueron ReLU y softmax para la clasificación

Capa	Dimension Kernel	Pool Size	Stride	función de activación
Conv2D	3 x 3 x 32	-	-	ReLU
MaxPool2D	-	2 x 2	2	
Conv2D	3 x 3 x 64	-	-	ReLU
MaxPool2D	-	2 x 2	2	
Conv2D	3 x 3 x 128	-	-	ReLU
MaxPool2D	-	2 x 2	2	
Conv2D	3 x 3 x 256	-	-	ReLU
MaxPool2D	-	2 x 2	2	
Capa densa	512	-	-	ReLU
Capa densa	3	-	-	Softmax

Tabla 8.2: Arquitectura CNN simplificada

8.1.3 Transformaciones aplicadas a imágenes

En este punto es donde se utilizan ciertas técnicas de manipulación en las imágenes, esto se realiza para dar mayor cantidad y diversidad de datos que puedan ser disponibles para luego ser introducidos en la red. Este paso consta de modificar las imágenes aleatoriamente para así poder introducir nuevos datos.

En este punto del proceso se aplican técnicas conocidas como aumento de datos, data augmentation, las cuales consisten en realizar diversas transformaciones sobre las imágenes originales con el objetivo de aumentar la cantidad y la diversidad del conjunto de datos disponible. Este paso es especialmente importante cuando se trabaja con conjuntos de datos pequeños, ya que ayuda a mejorar la capacidad de generalización del modelo y a reducir el sobreajuste, overfitting.

Las transformaciones se aplican de manera aleatoria, y pueden incluir:

- Rotaciones
- Traslaciones
- Escalado
- Reflejos
- Cambios en el brillo, contraste o color
- Agregado de ruido:

Cada imagen transformada se considera como una nueva muestra para entrenar la red neuronal. Aunque proviene de una imagen original, las modificaciones hacen que el modelo vea cada versión como un ejemplo distinto, lo cual fortalece el aprendizaje del modelo y permite una mejor generalización al enfrentarse a datos reales no vistos durante el entrenamiento.

Como ya se mencionó anteriormente y con la finalidad de aplicar técnicas de Data Augmentation se incorporó un bloque específico en el cual se utiliza Albumentations, una

eficiente librería desarrollada en Python. Esta librería está especialmente diseñada para trabajar con imágenes en tareas de visión por computadora, y permite aplicar una amplia gama de transformaciones tanto geométricas como fotométricas de manera sencilla y rápida.

El propósito principal de usar Albumentations es aumentar artificialmente la cantidad y diversidad del conjunto de datos a partir de las imágenes originales. Esto se logra mediante la aplicación aleatoria de distintas modificaciones sobre cada imagen, tales como rotaciones, volteos horizontales o verticales, escalados, recortes aleatorios, ajustes de brillo y contraste, adición de ruido, entre muchas otras. Cada una de estas transformaciones produce una nueva versión de la imagen, lo que ayuda a que el modelo de red neuronal pueda generalizar mejor y no se sobreentrene únicamente con las imágenes originales.

Unas de las transformaciones usadas fueron las siguientes:

- A.Resize(256, 256) Escalar para permitir recortes
- A.RandomCrop(224, 224) Recorte aleatorio al tamaño final
- A.HorizontalFlip(p=0.5) Volteo horizontal aleatorio
- A.ShiftScaleRotate(shift_limit=0.05, scale_limit=0.05, rotate_limit=15, p=0.5) Transformaciones geométricas
- A.RandomBrightnessContrast(p=0.2) Cambios de brillo/contraste

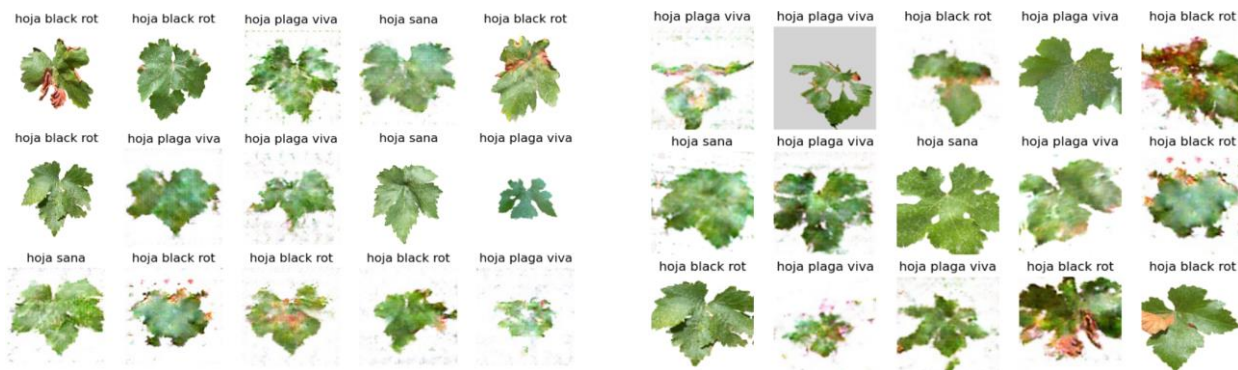


Fig. 8.1.3.4 Imágenes originales del Dataset vs imágenes con Data Augmentation

En la figura 12, de lado izquierdo se pueden observar algunas de las imágenes que son tomadas del entrenamiento para poder ser visualizadas de manera original sin ningún tipo de modificación y de lado derecho se encuentra el conjunto de imágenes con las transformaciones ya aplicadas, en algunas de ellas se puede observar como se modifican tanto en aspecto, como rotación o incluso tamaño y luminosidad de la misma imagen,

8.1.4 Transferencia de aprendizaje

Como ya se ha mencionado anteriormente en esta investigación, es una técnica en el campo del aprendizaje automático y aprendizaje profundo que consiste en aprovechar el conocimiento que un modelo ya ha aprendido en una tarea para aplicarlo a una nueva tarea relacionada.

En esta investigación se utilizaron diversas arquitecturas de redes neuronales convolucionales pre entrenadas, es decir, modelos que ya habían sido entrenados previamente con una gran base de datos general, como ImageNet, que contiene millones de imágenes anotadas de diferentes categorías. Las arquitecturas empleadas fueron:

- ResNet-18
- ResNet-50
- ResNet-101
- RegNet
- MobileNet

Estas redes se seleccionaron por su eficacia comprobada en tareas de reconocimiento de imágenes. Las variantes de ResNet, por ejemplo, son conocidas por introducir conexiones residuales que permiten el entrenamiento de redes muy profundas sin que se pierda la información durante la propagación hacia atrás, resolviendo así el problema del desvanecimiento del gradiente. Por su parte, RegNet es una arquitectura más moderna, diseñada mediante principios de búsqueda de arquitectura y escalabilidad eficiente, lo que le permite obtener un alto rendimiento con menor complejidad computacional.

Se realizaron dos enfoques principales con estas arquitecturas:

1. Entrenamiento desde cero:

En este enfoque, se inicializaron los modelos con pesos aleatorios, sin usar conocimiento previo. Esto permite al modelo aprender directamente desde los datos específicos del problema, aunque requiere más datos, mayor capacidad de cómputo y un mayor tiempo de entrenamiento.

2. Transfer Learning + Fine-Tuning

- Primero se congelaron las capas iniciales del modelo (que detectan patrones generales como bordes, formas y texturas).

- Posteriormente se ajustaron las últimas capas del modelo (fine-tuning), permitiéndoles aprender representaciones más específicas.

Este enfoque de Transfer Learning combinado con Fine-Tuning permite reducir significativamente el tiempo de entrenamiento, mejorar la precisión del modelo, y aprovechar al máximo las capacidades de las arquitecturas seleccionadas, especialmente cuando se cuenta con un conjunto de datos limitado.

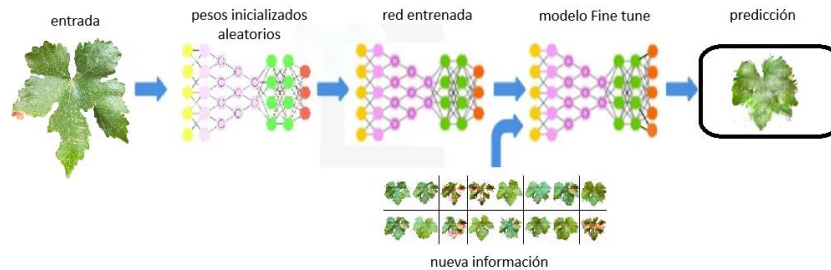


Fig. 8.1.4.5 Modelo con Transfer learning

8.1.5 Ajuste fino (Fine tuning)

Una vez que se ha aplicado Transfer Learning, es decir, cuando se ha cargado un modelo pre entrenado y se ha reemplazado su capa final para ajustarla a una nueva tarea, se suele congelar el resto de la red para conservar los pesos previamente aprendidos. Durante esta etapa, solo se entrena la nueva capa final, lo cual permite que el modelo comience a adaptarse sin modificar el conocimiento general adquirido.

El fine-tuning entra en acción cuando se “descongelan” algunas o todas las capas de la red pre entrenada. Esto significa que ahora esas capas también pueden actualizar sus pesos durante el entrenamiento, no solo la capa final. Al permitir pequeñas modificaciones en estas capas, el modelo puede refinar mejor sus representaciones internas y ajustarse con mayor precisión a las características específicas de la nueva tarea.

Este ajuste se realiza con una tasa de aprendizaje baja, para no alterar drásticamente los conocimientos previamente adquiridos, sino afinarlos sutilmente. El resultado es un modelo que mantiene la robustez del entrenamiento previo, pero ahora está mejor adaptado a los nuevos datos, lo que suele traducirse en un mejor rendimiento y una mayor capacidad de generalización.

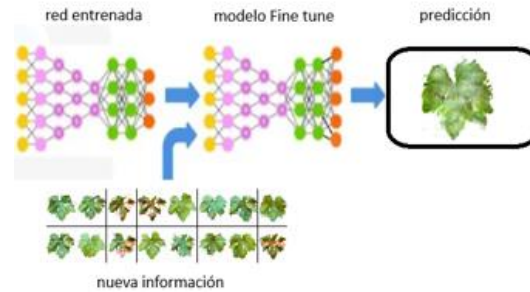


Fig. 8.1.5.6 Modelo con Fine tuning

8.2 Configuración de hiperparámetros

Para este paso se realizaron las configuraciones necesarias sobre estos puntos. Los modelos se entrenaron durante 30 épocas con el 80 % del conjunto de datos. El optimizador utilizado en todos los casos fue RMSProp con tasa de aprendizaje de 0.000001.

Estas configuraciones se tomaron en cuenta debido a que se lograron ver procesos de sobreaprendizaje en algunos modelos lograndose observar épocas con estancamientos al momento de utilizar estas configuraciones, así como también algunos pasos de programación para ir reduciendo la tasa de aprendizaje.

8.3 Validación de modelos

Para determinar la habilidad de los modelos al clasificar correctamente cada uno de los tres tipos de hoja: hoja sana, plaga viva y black rot, se evaluaron los modelos en el 20% del conjunto de datos total.

Estos datos no fueron utilizados en el proceso de entrenamiento, debido a que se hace una segmentación de las imágenes contenidas en la base de datos dejando el 80% del total para entrenamiento y el 20% restante para la evaluación de los modelos. Se obtuvo la Matriz de Confusión por cada modelo, además de un grupo de métricas derivadas de la misma.

8.4 Análisis comparativo de modelos

en este último paso de la metodología, se comparan los resultados obtenidos para los 7 modelos de Deep Learning que se codificaron. Con este fin se implementaron 7 métricas derivadas de la Matriz de Confusión, el valor de la función de pérdida, del AUC. También se analizaron los gráficos de la función de pérdida y la curva ROC. Los resultados se muestran en la Tabla 5. Tabla 5. Resultados de las métricas de desempeño. Precisión Global o Accuracy (PG), Precisión Positiva o Sensibilidad (PP), Precisión Negativa o Especificidad (PN), Falsos Positivos (PFP), Falsos Negativos (PFN), Asertividad Negativa (AN).

Modelo	PG	EG	PP	PN	FP	FN	AP	AN
CNN NORMAL	.7435	.2565	.7438	.8181	.1810	.1810	.7490	.8395
CNN NORMAL FT	.8275	.1725	.8335	.9155	.1725	.1725	.8275	.9138
RESNET 18	.9822	.0178	.9816	.9914	.0178	.0178	.9822	.9911
RESNET 50	.9829	.0171	.9821	.9917	.0171	.0171	.9829	.9914
RESNET 101	.9849	.0151	.9844	.9926	.0151	.0151	.9849	.9924
RESREG	.9756	.0244	.9749	.9881	.0244	.0244	.9756	.9878
MOBILE NET	.9855	.0145	.9846	.9927	.0145	.0145	.9855	.9928

Tabla 8.4 Comparación de modelos

Como se evidencia en los valores reportados, todos los modelos exhiben resultados por encima del 70%, siendo este el más bajo para la CNN Normal sin modificación, los demás modelos ya cuentan con TL y FT, lo cual nos ayuda a denotar en la gráfica como los niveles de exactitud por modelo llegaron casi al 99% elevando así una exactitud de manera considerable. De igual modo, en la Figura 17 se graficaron los valores de Accuracy en función de las épocas de entrenamiento. Los modelos RESNET 101 y MobileNET alcanzaron los mejores resultados. Se puede apreciar que desde las primeras épocas los valores de precisión están por encima del 80% lo cual indica que los modelos captan mayor cantidad de características en períodos tempranos del entrenamiento. Por otra parte, no se observaron muestras de sobreajuste para estos modelos pues su tendencia es ascendente en el caso de la precisión. Al analizar el comportamiento de la función de pérdida (Figura 18) se puede apreciar una tendencia descendente en sus valores los cuales se encuentran en rangos inferiores al 5%.

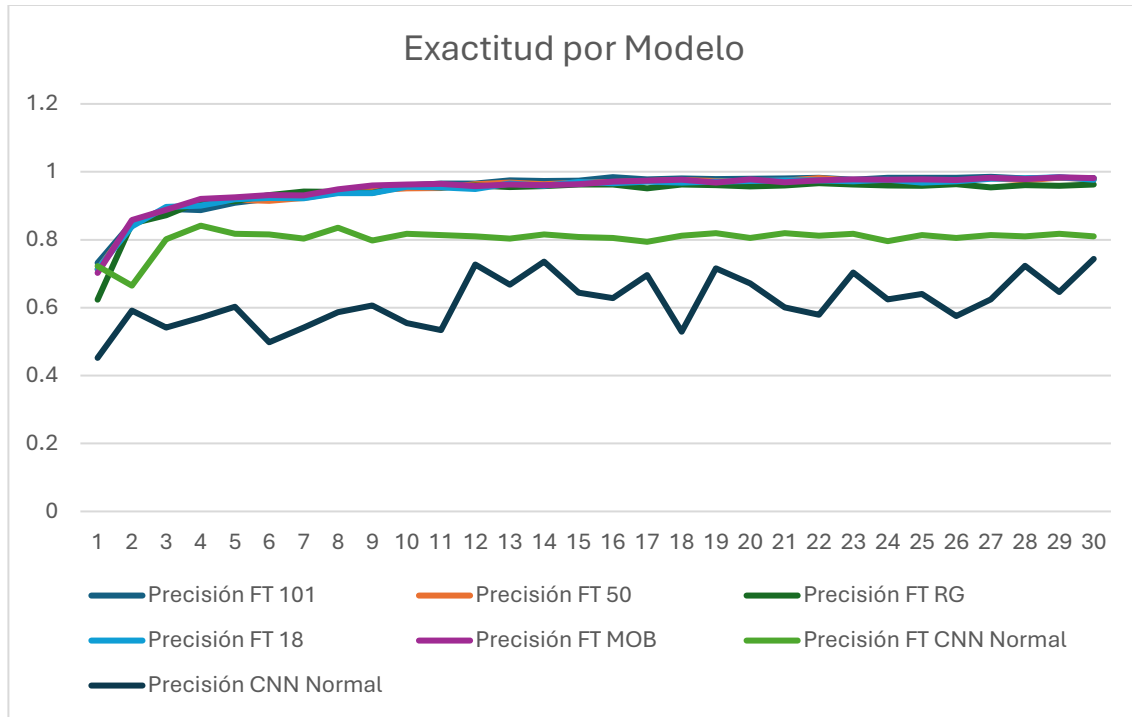


Fig. 8.4.1 Exactitud global por modelo

Figura 8.4.1. Precisión Global de los modelos. CNN: Red Neuronal Convolutiva Simple, AlexNet: Basado en AlexNet, DA: Aumento de datos, FTResNet18 Extracción de características con modelo preentrenado FEResNet18, FTResNet50: Extracción de características con modelo preentrenado, ResNet50: Extracción de características con modelo preentrenado FTResNet 101, FT: Ajuste Fino con modelo preentrenado FEResNet101. Extracción de características con modelo preentrenado ResReg, FT: Ajuste Fino con modelo preentrenado ResReg. Extracción de características con modelo preentrenado MobileNet, FT: Ajuste Fino con modelo preentrenado MobileNet.

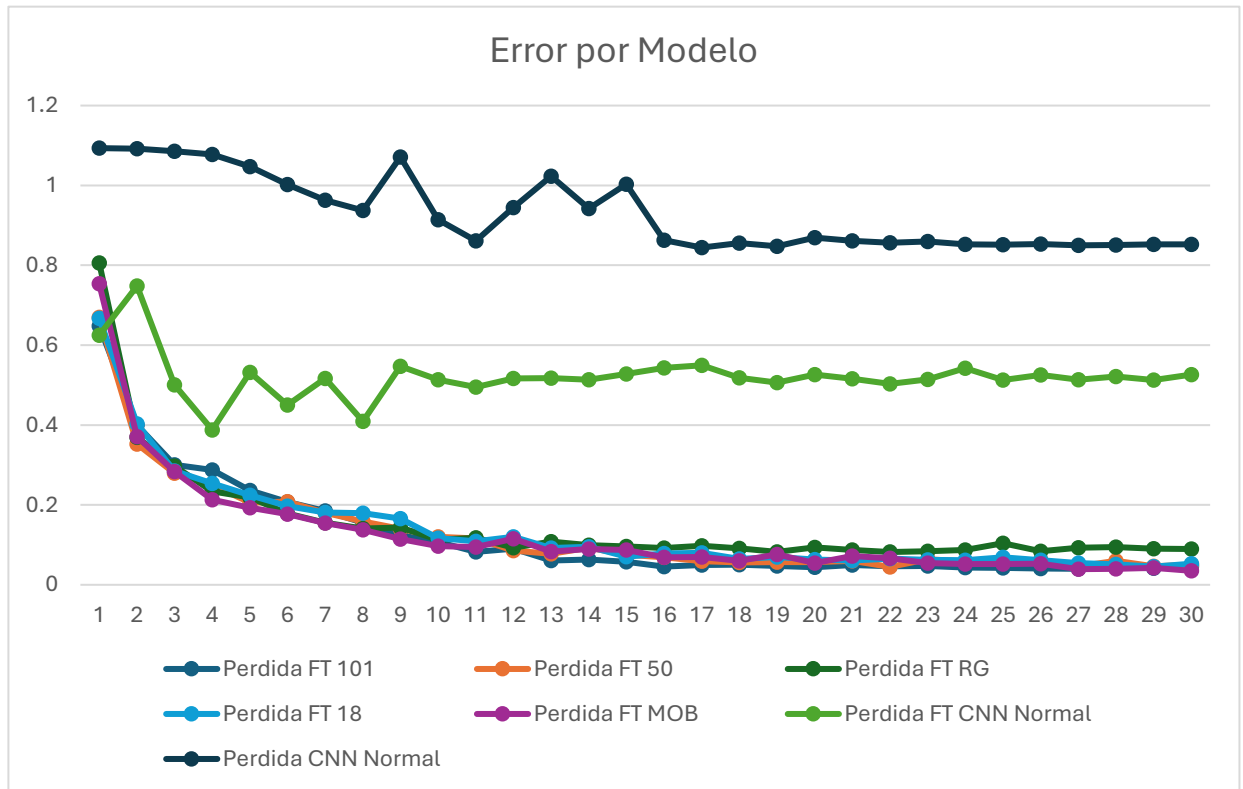


Fig. 8.4.2 Error global por modelo

Figura 8.4.2 FUNCION DE PERDIDA DE LOS MODELOS. CNN: Red Neuronal Convolutacional Simple, AlexNet: Basado en AlexNet, DA: Aumento de datos, FTResNet18 Extracción de características con modelo preentrenado FEResNet18, FTResNet50: Extracción de características con modelo preentrenado, ResNet50: Extracción de características con modelo preentrenado FTResNet 101, FT: Ajuste Fino con modelo preentrenado FEResNet101. Extracción de características con modelo preentrenado ResReg, FT: Ajuste Fino con modelo preentrenado ResReg. Extracción de características con modelo preentrenado MobileNet, FT: Ajuste Fino con modelo preentrenado MobileNet.

8.5 Predicción de nuevas imágenes

Con el objetivo de comprobar si los modelos generalizan bien ante nuevas entradas se realiza la clasificación de imágenes de: hoja sana, plaga viva y black rot, sobre imágenes del subconjunto de prueba (3% del total) que no fueron consideradas durante las etapas de entrenamiento y validación.



Fig. 8. 5.1 Predicción de nuevas imágenes con modelo MOBILENET V3

En la Figura 8.5.1 se visualizan las predicciones obtenidas utilizando el modelo MobileNet ya que fue el que mejores resultados arrojó según la sección anterior. El proceso de predicción ocurre al exponer el modelo entrenado ante nuevos datos con el objetivo de comprobar si realmente ha aprendido a detectar patrones relevantes en las imágenes. En este caso, se puede apreciar que se procesan un total de 9 imágenes (3 correspondientes hoja sana, 3 de plaga viva y 3 black rot) que no fueron utilizadas en las fases previas de entrenamiento y validación. Todas fueron clasificadas correctamente lo cual valida los resultados de las métricas mencionadas anteriormente.

8.6 Optimización del modelo

En la etapa de Optimización se compararon los tiempos de ejecución de los modelos de CNN normal FTy MobileNet, estos fueron ejecutados en un ordenador con CPU de 14

núcleos (20 procesadores lógicos) y GPU NVIDIA GeForce RTX 3050 TI. Sus características se muestran a continuación:

12th Gen Intel(R) Core(TM) i7-12700H		NVIDIA GeForce RTX 3050 Ti Laptop GPU	
Velocidad de base:	2.30 GHz	Versión del controlador:	31.0.15.5176
Sockets:	1	Fecha del controlador:	01/03/2024
Núcleos:	14	Versión de DirectX:	12 (FL 12.1)
Procesadores lógicos:	20	Ubicación física:	Bus PCI 1, dispositivo 0, función 0
Virtualización:	Habilitado	Memoria reservada para el hardware:	134 MB
Caché L1:	1.2 MB		
Caché L2:	11.5 MB		
Caché L3:	24.0 MB		

Fig. 8. 6.1 Características del ordenador

Para mejorar la capacidad de generalización del modelo y prevenir el sobreajuste (overfitting), se incorporaron diversos mecanismos dentro del proceso de entrenamiento. En primer lugar, se aplicó una regularización de tipo L2 mediante el parámetro weight decay en el optimizador, con el propósito de penalizar los pesos excesivamente grandes y promover una mayor simplicidad en la representación del modelo. Asimismo, se implementó un scheduler del tipo ReduceLROnPlateau, encargado de disminuir la tasa de aprendizaje (learning rate) cuando la precisión de validación (validation accuracy) deja de mejorar después de un número determinado de épocas, evitando así el estancamiento del aprendizaje y favoreciendo una convergencia más estable. Adicionalmente, se integró un mecanismo de guardado del mejor modelo basado en el desempeño de validación, lo cual actúa como una forma de early stopping implícito, asegurando que se conserve únicamente el modelo con mejor capacidad de generalización. En conjunto, estas estrategias contribuyen a mitigar el sobreajuste y a optimizar el proceso de entrenamiento de la red neuronal.

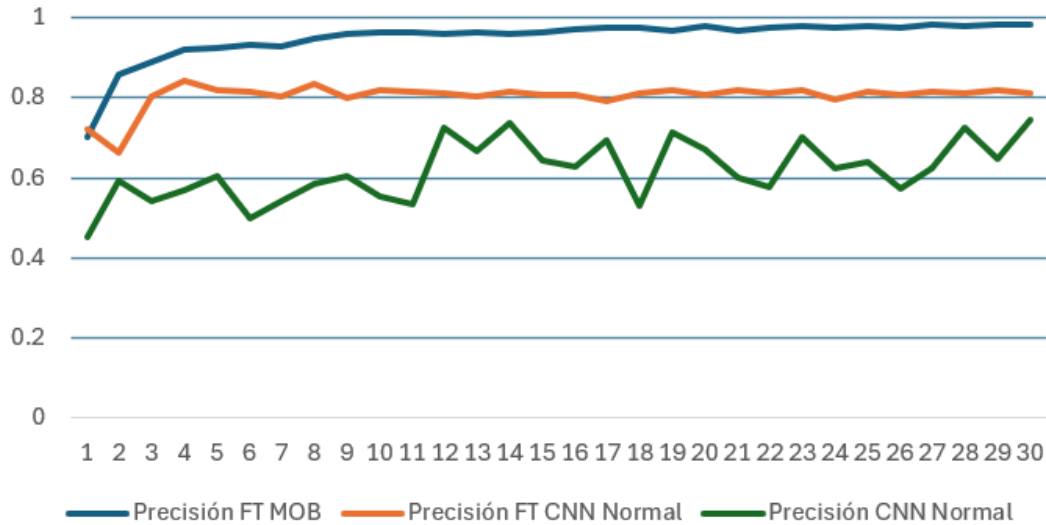


Figura 8.6.2 Comportamiento de la precisión durante el entrenamiento de los modelos propuestos.

En la figura 8.6.2 se observa la evolución de la precisión de validación para los distintos modelos implementados a lo largo de 30 épocas. El modelo FT MOB (línea azul), correspondiente a una arquitectura con fine-tuning sobre una red preentrenada tipo MobileNet, muestra un desempeño consistentemente superior, alcanzando una precisión cercana al 99% desde las primeras épocas y manteniendo una estabilidad notable durante todo el entrenamiento. Por otro lado, el modelo FT CNN Normal (línea naranja), que también emplea fine-tuning pero sobre una arquitectura convolucional tradicional, presenta un comportamiento más variable, estabilizándose alrededor del 80%. Finalmente, el modelo CNN Normal (línea verde), entrenado desde cero sin transferencia de aprendizaje, evidencia una menor capacidad de generalización y una alta fluctuación en su precisión, manteniéndose en un rango aproximado del 55–70%.

Estos resultados confirman la ventaja del aprendizaje por transferencia, ya que los modelos que emplean fine-tuning logran una convergencia más rápida, mayor estabilidad y mejor desempeño general en comparación con los entrenamientos realizados desde cero.

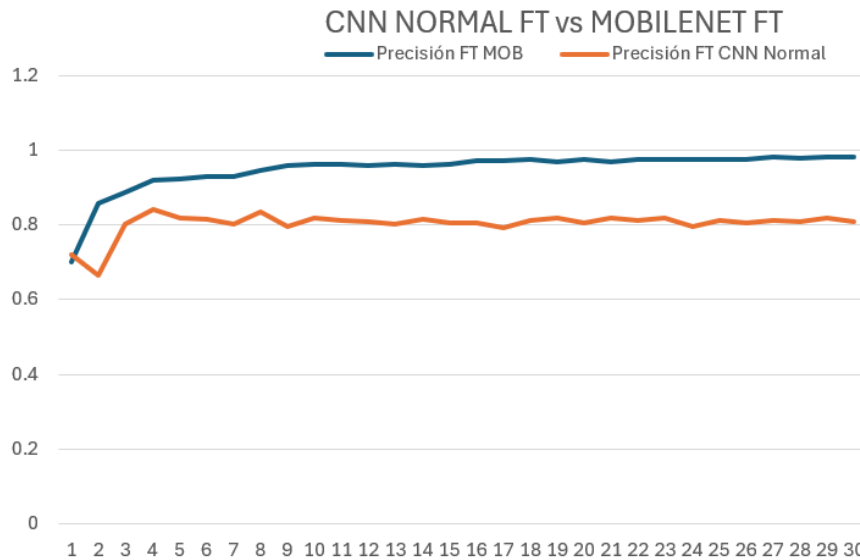


Figura 8.6.3 Comparativa de desempeño entre los modelos CNN Normal FT y MobileNet FT.

En la figura 8.6.3 se presenta la comparación directa entre los modelos CNN Normal FT y MobileNet FT, ambos entrenados bajo el enfoque de fine-tuning. Se observa que el modelo basado en MobileNet (línea azul) alcanza un desempeño superior y más estable a lo largo de las 30 épocas, con una precisión cercana al 98%. En contraste, el modelo CNN Normal FT (línea naranja) mantiene una precisión promedio en torno al 80%, mostrando ligeras fluctuaciones a lo largo del entrenamiento.

Estos resultados evidencian que la utilización de una arquitectura preentrenada optimizada, como MobileNet, permite aprovechar de manera más eficiente las características previamente aprendidas en grandes conjuntos de datos, lo que se traduce en una mejor capacidad de generalización y convergencia más rápida frente a arquitecturas convolucionales tradicionales, incluso cuando ambas emplean fine-tuning.

La comparación final confirma que el modelo MobileNet FT presenta un rendimiento significativamente superior, consolidándose como la opción más eficiente para el problema abordado.

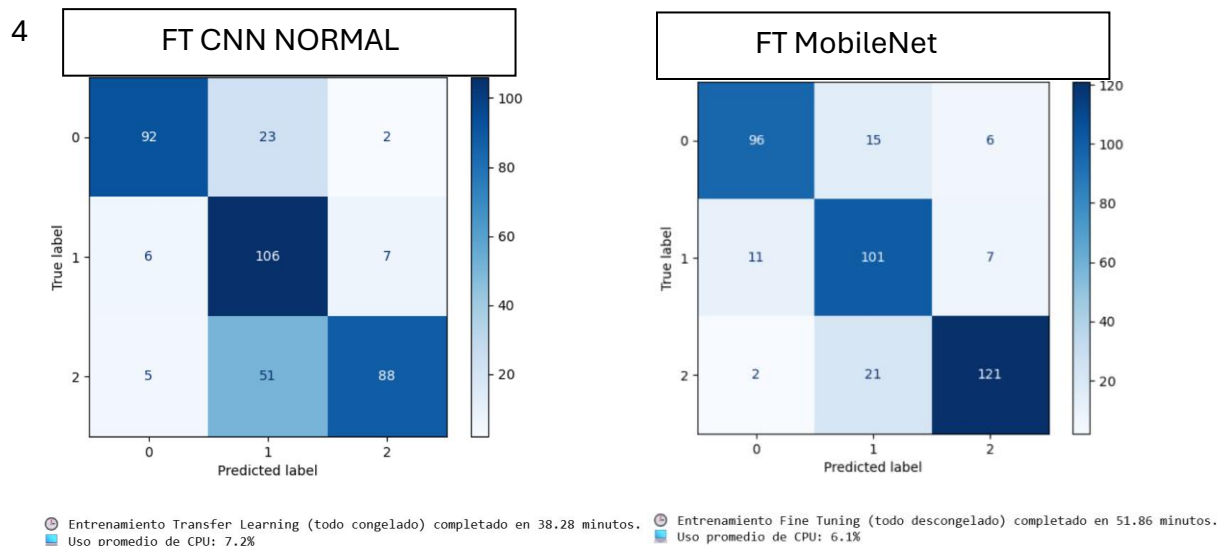


Figura 8.6.4. Comparativa de matrices de confusión entre los modelos CNN Normal con Transfer Learning y MobileNet con Fine Tuning.

En la figura 8.6.4 se presentan las matrices de confusión correspondientes a los modelos CNN Normal con Transfer Learning (todo congelado) y MobileNet con Fine Tuning (todas las capas descongeladas). La primera matriz (izquierda) corresponde al entrenamiento mediante transferencia de aprendizaje con FT de la CNN Normal, donde se observa un desempeño aceptable, aunque con una cantidad considerable de errores de clasificación en la tercera clase, reflejando cierta limitación para generalizar características complejas del conjunto de datos.

Por otro lado, la matriz de la derecha representa el modelo MobileNet FT, en el cual se aplicó fine tuning total, es decir, se permitió el ajuste de todos los parámetros de la red preentrenada. Este modelo muestra una mejora notable en la precisión global, especialmente en la tercera clase, donde se reducen significativamente los falsos negativos. El modelo evidencia una mayor capacidad para discriminar entre categorías y una distribución más equilibrada de las predicciones correctas.

En términos de desempeño computacional, el modelo CNN Normal TL completó su entrenamiento en 38.28 minutos con un uso promedio de CPU de 7.2%, mientras que el modelo MobileNet FT requirió 51.86 minutos con un uso promedio de 6.1%. Aunque el tiempo de entrenamiento del modelo MobileNet es ligeramente mayor debido al ajuste completo de sus capas, el incremento en la precisión y la estabilidad del modelo justifican el costo computacional adicional.

Estos resultados confirman que el aprendizaje por transferencia con ajuste fino (Fine Tuning) ofrece un mejor equilibrio entre precisión y generalización, superando al enfoque

con capas congeladas, especialmente en tareas de clasificación con mayor complejidad visual.

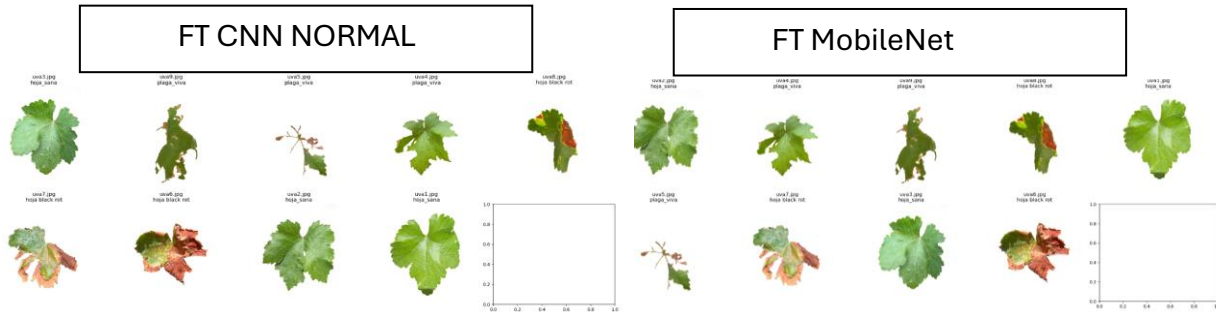


Figura 8.6.5 Comparativa de predicciones entre los modelos CNN Normal con Transfer Learning FT y MobileNet con Fine Tuning.

Como se puede observar en la figura 8.6.5, las predicciones realizadas por los modelos son correctas en su mayoría. En cada caso se utilizaron nueve imágenes de prueba, tanto para el modelo CNN Normal con Fine Tuning como para el MobileNet con Fine Tuning, logrando una clasificación adecuada en todas las categorías.

Estos resultados reflejan que ambos modelos fueron capaces de identificar correctamente los patrones visuales presentes en las imágenes, evidenciando una buena capacidad de generalización y un comportamiento estable tras el proceso de entrenamiento. La correcta categorización en los dos modelos respalda la efectividad de las estrategias implementadas para evitar el sobreajuste y optimizar el rendimiento en la etapa de inferencia.

8.7 Conclusiones parciales

En el presente capítulo se desarrollaron varios modelos de Deep Learning tanto

En este capítulo se desarrollaron diversos modelos de Deep Learning, tanto supervisados como no supervisados, para la clasificación de plagas en hojas de uva. La necesidad de contar con métodos de diagnóstico efectivos impulsa la investigación científica y la exploración de enfoques alternativos. En este contexto, se empleó una base de datos compuesta por muestras de cultivos de uva sanos y enfermos, las cuales fueron analizadas mediante técnicas de Deep Learning.

En total, se implementaron seis modelos: CNN con Transfer Learning, ResNet18, ResNet50, ResNet101, RegNet y MobileNet V3. Para evaluar el desempeño de cada modelo, se calcularon métricas derivadas de la matriz de confusión, destacando especialmente la Precisión Global, la Especificidad y la Sensibilidad. Asimismo, se analizaron las gráficas de las funciones de pérdida para identificar posibles problemas de sobreajuste.

Finalmente, se concluyó que el modelo más robusto fue MobileNet V3, con el cual se alcanzó una Precisión Global del 98.55% y una Sensibilidad del 97%.

8.8 Conclusiones generales

El presente estudio evidencia que las técnicas de Deep Learning constituyen una herramienta altamente efectiva para la detección y clasificación de plagas en hojas de uva, mostrando un gran potencial para su aplicación en el ámbito agrícola. La implementación de seis modelos distintos, que incluyen tanto enfoques supervisados como no supervisados, permitió realizar una comparación integral de su desempeño, lo que facilitó identificar fortalezas y limitaciones de cada enfoque en el reconocimiento de patrones complejos en imágenes de hojas de uva sanas y afectadas por plagas.

La evaluación de los modelos se realizó mediante métricas cuantitativas derivadas de la matriz de confusión, entre las cuales destacan la Precisión Global, la Sensibilidad y la Especificidad, que permiten medir la exactitud y la capacidad de los modelos para diferenciar correctamente entre hojas sanas y enfermas. Además, el análisis de las funciones de pérdida durante el entrenamiento fue clave para detectar posibles problemas de sobreajuste, asegurando que los modelos aprendieran patrones generalizables y no solo memorizaran los datos de entrenamiento.

Entre los modelos evaluados, MobileNet V3 se consolidó como el más robusto y confiable, alcanzando una Precisión Global del 98.55% y una Sensibilidad del 97%, lo que indica su elevada capacidad para identificar correctamente las plagas sin generar un número significativo de falsos positivos o negativos. Este resultado resalta la relevancia de seleccionar arquitecturas de Deep Learning optimizadas para tareas específicas de clasificación de imágenes, donde la eficiencia computacional y la precisión deben equilibrarse.

En términos prácticos, los hallazgos de este estudio demuestran que el uso de Deep Learning puede convertirse en una herramienta complementaria para los agricultores y profesionales del sector vitivinícola, facilitando un monitoreo más rápido, preciso y automatizado de la salud de los cultivos. La adopción de este tipo de tecnologías no solo optimiza la toma de decisiones en el manejo de plagas, sino que también contribuye a reducir pérdidas económicas y promover prácticas agrícolas más sostenibles.

En conclusión, los resultados obtenidos respaldan la viabilidad y efectividad del Deep Learning como un método confiable para la clasificación de plagas en hojas de uva, abriendo la puerta a su integración en sistemas de diagnóstico agrícola de próxima generación y al desarrollo de soluciones tecnológicas que potencien la productividad y la calidad de los cultivos.

8.9 Referencias

Bibliografía

- Abelardo Gutiérrez Suanzes, V. R. (2008). *APLICACIÓN DE MÉTODOS BIO-ACÚSTICOS PARA LA DETECCIÓN TEMPRANA DEL PICUDO ROJO*. Valencia.
- Barboni, L. (2014). *Desarrollo de un prototipo de nodo WSN con capacidad visual para la detección temprana de plagas*. uruguay.
- Berzal, F. (2018). *REDES NEURONALES & DEEP LEARNING*. Granada.
- Hinton. (1990). *Mapping part-whole hierarchies into connectionist networks*. *Artificial Intelligence*. G.E.
- Hinton, G. (2012). *Neural networks for machine learning*. .
- Hinton, G. V. (2015). *Distilling the knowledge in a neural network*.
- Huang, P.-S. H. (2013). *Learning deep structured semantic models for web search using clickthrough data*.
- Hutter, F. H.-B. (2011). *Sequential model-based optimization for general algorithm configuration*.
- Jain, V. M. (2007). *Supervised learning of image restoration with convolutional networks*.
- Johan Sebastián Penagos Figueredo, S. V. (2023). *Diseño de Sistema Automatizado de Innovación en los Cultivos Hidropónicos Mediante Tecnologías 4.0*. Bucaramanga, Colombia: Universidad de Santander.
- Rouhiainen, L. (2018). *Inteligencia artificial*. Alienta Editorial.



