

UNIVERSIDAD NACIONAL JORGE BASADRE GROHMANN

Facultad de Ingeniería

Escuela Profesional de Ingeniería en Informática y Sistemas

**EVALUACIÓN COMPARATIVA DE MODELOS DE
DEEP LEARNING PARA LA CLASIFICACIÓN DE
MALEZAS EN EL CULTIVO DE OLIVO
UTILIZANDO DATOS REMOTOS UAV
EN LA YARADA-LOS PALOS
DE TACNA, 2024**

TESIS

Presentada por:

Bach. WILLIANS LUCIO MAYTA YUCRA

Para optar el Título Profesional de:

INGENIERO EN INFORMÁTICA Y SISTEMAS

TACNA – PERÚ

2025



Universidad Nacional Jorge Basadre Grohmann

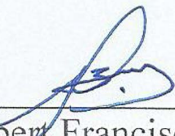
Facultad de Ingeniería

Escuela Profesional de Ingeniería en Informática y Sistemas

Certificado de similitud

Yo, Dr. Erbert Francisco Osco Mamani, en mi condición de asesor acreditado por la Resolución de Facultad N° 8503-2024-FAIN/UNJBG, de la Tesis: “Evaluación comparativa de modelos de Deep Learning para la clasificación de Malezas en el cultivo de Olivo utilizando datos remotos UAV en la Yarada - Los Palos de Tacna, 2024”, presentado por el bachiller WILLIANS LUCIO MAYTA YUCRA, para optar el Título Profesional Ingeniero en Informática y Sistemas.

Habiendo cumplido con lo establecido en el reglamento de originalidad y de similitud de trabajos de investigación y producción intelectual, considerando que según la revisión, evaluación y análisis realizada a través del software de similitud textual Turnitin, cuenta con el NIVEL DE SIMILITUD de 7%. Por lo que certifico la similitud de la tesis, que se está de acuerdo al NIVEL PERMITIDO, para continuar con los trámites correspondientes y para su publicación en el repositorio institucional. Se emite el presente certificado con fines de continuar con los trámites respectivos para su obtención del título profesional.



Dr. Erbert Francisco Osco Mamani

DNI:00409196



Bach. Willians Lucio Mayta Yucra

DNI:42947618

DEDICATORIA

*A mi Papá en el cielo
y a mi familia, por ser
quienes están a mi
lado dándome las
fuerzas necesarias.*

AGRADECIMIENTOS

En este recorrido, siempre hubo quienes me impulsaron a continuar, incluso en los instantes más desafiantes. A todas esas personas especiales, deseo manifestar mi sincero agradecimiento por su constante apoyo, por la confianza depositada en mí y por el afecto que siempre me brindaron.

De manera muy especial, agradezco a mis padres, quienes han sido un pilar fundamental durante toda esta experiencia, acompañándome en mi vida personal y universitaria, y compartiendo conmigo tanto los momentos de felicidad como los de adversidad.

A mis maestros de la Escuela Profesional de Ingeniería en Informática y Sistemas de la Universidad Jorge Basadre Grohmann por su apoyo y orientación durante mi formación académica. Al proyecto de investigación “Estimación de la productividad de olivares usando técnicas de visión artificial y aprendizaje profundo en la región de Tacna” Y por supuesto, a Dios.

ÍNDICE TEMÁTICO

DEDICATORIA	ii
AGRADECIMIENTOS	iii
ÍNDICE TEMÁTICO	iv
ÍNDICE DE TABLAS	viii
ÍNDICE DE FIGURAS	ix
RESUMEN	x
ABSTRACT	xi
INTRODUCCIÓN	12
CAPÍTULO 1. PLANTEAMIENTO DEL PROBLEMA	14
1.1 Antecedentes del problema	14
1.2 Descripción del problema	14
1.3 Formulación del problema	15
1.3.1 Problema general	15
1.3.2 Problemas específicos	15
1.4 Objetivos de la investigación	16
1.4.1 Objetivo general	16
1.4.2 Objetivos específicos	16
1.5 Justificación e importancia de la investigación	16
1.6 Limitaciones	17
1.7 Viabilidad del estudio	17
1.8 Formulación de hipótesis	17
1.8.1 Hipótesis general	17
1.8.2 Hipótesis específicas	17
1.9 Variables	18
1.9.1 Definición de la variable	18
1.10 Operacionalización de variables	18
CAPÍTULO 2. MARCO TEÓRICO	19
2.1 Antecedentes del trabajo de investigación	19
2.2 Bases teóricas	19

2.2.1	Olea europea L.	19
2.2.2	Control de malezas	20
2.2.3	Malezas	20
2.2.4	Tipos de malezas	21
2.2.5	El cultivo del olivo en la región Tacna	21
2.2.6	Inteligencia artificial	22
2.2.7	Aprendizaje automático	22
2.2.8	Aprendizaje profundo	22
2.2.9	Redes neuronales artificiales	23
2.2.10	Redes neuronales convolucionales (CNN)	23
2.2.11	Modelos de aprendizaje profundo	24
2.2.12	Arquitectura VGG16	24
2.2.13	Arquitectura MobileNet	25
2.2.14	Arquitectura MobileNetV2	25
2.2.15	Arquitectura ConvNeXtXLarge	26
2.2.16	Subajuste y sobreajuste	27
2.2.17	Aprendizaje por transferencia	27
2.2.18	Ajuste fino	28
2.2.19	Aumento de datos	29
2.2.20	Métricas de clasificación	29
2.3	Definiciones conceptuales	30
2.3.1	Matriz de confusión	30
2.3.2	Herbicidas	30
2.3.3	Pulverización uniforme de herbicidas	30
2.3.4	Precisión	30
2.3.5	Sensibilidad	31
2.3.6	Especificidad	31
2.3.7	Verdaderos positivos (TP)	31
2.3.8	Verdaderos negativos (TN)	31
2.3.9	Falso positivo (FP)	31
2.3.10	Falso negativo (FN)	31
2.3.11	Conjunto de entrenamiento	31
2.3.12	Época	31

2.3.13	UAV	31
CAPÍTULO 3. MARCO METODOLÓGICO		32
3.1	Planteamiento metodológico	32
3.1.1	Nivel	32
3.1.2	Diseño	32
3.2	Población y muestra	33
3.2.1	Población	33
3.2.2	Muestra	33
3.3	Equipos y materiales	33
3.4	Procedimiento de las pruebas experimentales	34
3.5	Técnicas de recolección de datos	34
3.6	Técnicas para el procesamiento de datos	34
3.6.1	Captura de las imágenes	36
3.6.2	Generación del ortomosaico	37
3.6.3	Recorte del conjunto de datos	37
3.6.4	Preparación de los datos	37
3.6.5	Aumento de datos	38
3.6.6	Preparar las arquitecturas	38
3.6.7	División de los datos	38
CAPÍTULO 4. RESULTADOS		39
4.1	Descripción de las pruebas experimentales	39
4.2	Presentación y análisis de los resultados	39
4.2.1	Entrenamiento: etapa de aprendizaje por transferencia	39
4.2.2	Entrenamiento: etapa de ajuste fino	41
4.2.3	Selección del modelo	42
4.3	Contrastación de hipótesis	44
4.3.1	Prueba de normalidad	44
4.3.2	Prueba de hipótesis	46
CAPÍTULO 5. DISCUSIÓN		50
5.1	Pruebas de validación	50
5.2	Contraste con trabajos de investigación similares	50
CONCLUSIONES		52
RECOMENDACIONES		53

REFERENCIAS BIBLIOGRÁFICAS	54
ANEXOS	58
ANEXO 01: MATRIZ DE CONSISTENCIA	59
ANEXO 02: DATOS OBTENIDOS DEL ENTRENAMIENTO DE LOS CUATRO MODELOS	60
ANEXO 03: PRUEBA DE NORMALIDAD DE DATOS PARA LOS INDICADORES EXACTITUD Y PÉRDIDA DE LOS MODELOS EN SPSS 25	61
ANEXO 04: PRUEBA POST HOC PARA EL INDICADOR PÉRDIDA DE LOS MODELOS EN SPSS 25	62
ANEXO 05: PRUEBA DE KRUSKAL WALLIS PARA EL INDICADOR EXACTITUD DE LOS MODELOS EN SPSS 25	63
ANEXO 06: PRUEBA DE KRUSKAL WALLIS PARA EL INDICADOR EXACTITUD PARA CADA MODELO EN SPSS 25	64
ANEXO 07: RECOLECCIÓN DE LOS DATOS EN EL FUNDO SANTA MÓNICA UNJBG	65
ANEXO 08: IMÁGENES CAPTURADAS CON DRON DJI MAVIC 3	67
ANEXO 09: ORTOFOTOS GENERADAS	68
ANEXO 10: CATEGORÍAS PARA EL ENTRENAMIENTO DE LOS MODELOS	71
ANEXO 10: CÓDIGO UTILIZADO EN PYTHON	73

ÍNDICE DE TABLAS

Tabla 1	Tabla de operacionalización de variables	18
Tabla 2	Malezas frecuentes en los cultivos de olivo	21
Tabla 3	División de las imágenes por cada clase para el entrenamiento	38
Tabla 4	Métricas sobre las repeticiones del entrenamiento en la validación	43
Tabla 5	Matriz de confusión en el conjunto de validación	44
Tabla 6	Métricas de clasificación por clase en el conjunto de validación	44
Tabla 7	Tiempo promedio de entrenamiento de modelos	44
Tabla 8	Prueba de normalidad para el indicador Exactitud	45
Tabla 9	Prueba de normalidad para el indicador Exactitud	46
Tabla 10	Prueba de Kruskal-Wallis indicador Exactitud	46
Tabla 11	Prueba de post hoc para el indicador Exactitud	47
Tabla 12	Prueba de ANOVA para el indicador Pérdida	48
Tabla 13	Prueba de post hoc para el indicador Pérdida	49

ÍNDICE DE FIGURAS

Figura 1 Aprendizaje automático	22
Figura 2 Aprendizaje automático	23
Figura 3 Arquitectura VGG16	24
Figura 4 Arquitectura MobileNet	25
Figura 5 Arquitectura MobileNetV2	26
Figura 5 Arquitectura ConvNeXtLarge	27
Figura 7 Aprendizaje por transferencia	28
Figura 8 Generación de imágenes mediante aumento de datos aleatorio	29
Figura 9 Flujo de trabajo	35
Figura 10 Lugar de toma de fotografías	36
Figura 11 Captura de imágenes mediante dron DJI MAVIC 3	37
Figura 12 Exactitud en el entrenamiento	39
Figura 13 Pérdida en el entrenamiento	40
Figura 14 Exactitud de la mejor repetición	42
Figura 15 Diagrama de cajas en la validación	42
Figura 16 Fundo Santa Monica UNJBG - Yarada -Los Palos	65
Figura 17 Olivos Yarada - Los Palos	65
Figura 18 Olivos Yarada - Los Palos	66
Figura 19 Investigadores Proyecto de investigación: Inteligencia artificial del Olivo Yarada - Los Palos	66
Figura 20 Dron	67
Figura 21 Imágenes capturadas con el Dron en la Yarada los Palos	67
Figura 22 Ortofoto generada 1	68
Figura 23 Ortofoto generada 2	69
Figura 24 Ortofoto generada 3	70
Figura 25 Olivo con presencia de Malezas con baja Infestación	71
Figura 26 Olivo con presencia de Malezas con alta Infestación	71
Figura 27 Olivo sin Presencia de Malezas	72

RESUMEN

En el ámbito agrícola, se considera malezas a aquellas plantas indeseadas que rivalizan con los cultivos por la disponibilidad de nutrientes, agua y radiación solar. Algunas de estas especies, además, pueden actuar como portadoras de plagas y enfermedades que comprometen la sanidad de las plantas cultivadas.

En el cultivo de Olivo en la Región de Tacna-Perú tiene una gran importancia debido a sus altos valores de producción, durante la siembra aparecen malezas de Hoja ancha y angosta como el Trébol blanco, Trébol amarillo, Chirillo, entre otros; debido a ello, estos son perjudiciales para el normal crecimiento del Olivo y su calidad.

Para la adquisición de las imágenes, se recolectó imágenes RGB de los cultivos de Olivo utilizando un dron modelo “DJI Mavic 3” a una distancia de 12 metros de altura y con un traslape de 75% en los cultivos del fundo “Santa Mónica” los cuales forman parte de la Universidad Nacional Jorge Basadre Grohmann de la ciudad de Tacna.

Para el procesamiento de los datos, con el conjunto de imágenes se generó 2 ortofotos utilizando el software Agisoft Metashape, seguidamente se realizó los recortes de las imágenes y los etiquetados para posteriormente realizar la evaluación y comparación del rendimiento de los modelos de clasificación con cuatro arquitecturas: ConvNeXtXLarge, VGG16, MobileNet y MobileNetv2; además, se utilizó métodos última generación. de aprendizaje como: Aprendizaje por Transferencia, Ajuste Fino y aumento de datos.

Se concluye que el modelo con mejor desempeño fue la arquitectura ConvNeXtXLarge, con un valor máximo de 98.78% en el indicador Exactitud y un valor mínimo de 9.47% en el indicador

Palabras Clave: clasificación de Malezas, aprendizaje profundo, aprendizaje por Transferencia, ajuste Fino, aumento de datos, Olivo.

ABSTRACT

In agriculture, weeds are unwanted plants that compete with crops for nutrients, water, and sunlight. Several of these weeds can carry diseases and pests that affect crop health.

In olive cultivation in the Tacna-Perú region, it is of great importance due to its high production values. During planting, broadleaf and narrowleaf weeds such as white clover, yellow clover, and Chirillo, among others, appear. These are detrimental to the normal growth and quality of the olive.

For image acquisition, RGB images of the olive crops were collected using a “DJI Mavic 3” drone at a height of 12 meters and with a 75% overlap in the crops of the “Santa Mónica” farm, which is part of the National University Jorge Basadre Grohmann in the city of Tacna.

For data processing, two orthophotos were generated from the set of images using Agisoft Metashape software. Subsequently, the images were cropped and labeled to later evaluate and compare the performance of classification models with four architectures: ConvNeXtXLarge, VGG16, MobileNet, and MobileNetv2. Additionally, state-of-the-art learning methods such as Transfer Learning, Fine-Tuning, and data augmentation were used.

It was concluded that the best-performing model was the ConvNeXtXLarge architecture, with a maximum accuracy of 98.78% and a minimum value of 9.47% in the indicator.

Keywords: Weed classification, deep learning, Transfer Learning, Fine-Tuning, data augmentation, Olive.

INTRODUCCIÓN

Tacna es considerada como la primera región en el cultivo del Olivo, lo que se ha logrado con la explotación de terrenos en las irrigaciones de la Yarada, Magollo, Ite y el Valle de Sama, donde las condiciones agro climáticas son favorables, razón por la cual, su productividad ha ido en incremento constante (Casilla García, 2011).

Los diferentes cultivos a nivel mundial presentan como problema principal las malas hierbas conocidas como malezas, que son plantas destructivas que crecen y compiten con los cultivos para poder obtener agua y nutrientes (Subeesh et al., 2022).

En la agricultura, el control de malezas es parte fundamental para poder garantizar un rendimiento óptimo en los cultivos, pero la detección y el control preciso de las malezas pueden resultar desafiantes y costosos para los agricultores.

Por lo general, los métodos de detección de malezas se basan en la observación visual, lo que puede ser ineficiente y expuesto a errores, en especial en grandes cultivos; por otra parte, la erradicación de estas malezas también se lleva a cabo mediante la aplicación de agroquímicos. Las malas técnicas de implementación de estos químicos como la pulverización uniforme de herbicidas actualmente provocan la intoxicación de cultivos, contaminación ambiental y además el alto costo del consumo de herbicidas presentan efectos secundarios negativos sobre el suelo (Sabzi et al., 2020).

En el 2014 en la ciudad de Tacna, se realizó un estudio de las malezas existentes en los cultivos del Instituto de Producción y Extensión (INPREX) y Centro Experimental Agrícola (CEA III) Los Pichones, los cuales forman parte de la Universidad Nacional Jorge Basadre Grohmann de Tacna, en donde se recolectaron 28 especies de malezas en diferentes cultivos, entre los cultivos Olivares se encontraron 5 tipos de malezas de Hoja ancha y angosta en abundancia como la: Celandilla, Trébol Blanco, Trebol amarillo, Gainsoga paniflora Cav., Foquitos y Chirillo (Zegarra & Solsol, 2015), por lo tanto; de ello existe la necesidad de clasificar automáticamente las malas hierbas en los cultivos para mejorar y evitar pérdidas de calidad y de cosecha.

En el 2023, en el área de Lapan Gwari de Minna, estado de Níger en Nigeria se realizó la adquisición de imágenes aéreas, se realizó mediante un UAV (Vehículo Aéreo no Tripulado) a 30 metros de altura para cultivos de caña de azúcar, espinacas, pimiento

y plátano; Sin embargo, para las malezas más comunes en la región de Tacna-Perú no existe un conjunto de datos para entrenar y evaluar los modelos.

De acuerdo a los últimos estudios publicados en revistas científicas, se han utilizado estrategias de inteligencia artificial (IA) y aprendizaje profundo para el manejo de malezas, la parte más importante de estos pasos es la identificación y clasificación de malezas. Los modelos basados en aprendizaje profundo se entrenan directamente a partir de los datos y, para alcanzar altos niveles de precisión, necesitan contar con extensos volúmenes de información. Adicionalmente para mejorar los resultados en el entrenamiento, se utilizan técnicas para obtener modelos con mejores resultados, como el Data Augmentation, Transfer Learning y ajuste fino.

El desarrollo de esta investigación se centra en realizar un análisis comparativo de distintos modelos de inteligencia artificial fundamentados en técnicas de aprendizaje profundo, aplicados a la clasificación de malezas presentes en los cultivos de olivo en la zona de La Yarada, Los Palos, Tacna – Perú. Se evaluará el desempeño de diversas arquitecturas, entre ellas ConvNeXtXLarge, VGG16, MobileNet y MobileNetV2.

El presente trabajo está organizado en cinco capítulos de la siguiente manera: El capítulo I trata la descripción de la problemática de la investigación, la formulación del problema, los objetivos planteados, además de la justificación y las limitaciones del estudio. El capítulo II recopila antecedentes relacionados al ámbito de la investigación, así como las bases teóricas y definiciones. En el capítulo III se detalla la metodología, procedimientos y técnicas usadas en este trabajo. En el capítulo IV se muestran los resultados, así como la contrastación estadística de las hipótesis. El capítulo V muestra las discusiones. Finalmente las conclusiones y recomendaciones.

CAPÍTULO 1.

PLANTEAMIENTO DEL PROBLEMA

1.1 Antecedentes del problema

A nivel internacional se presentan problemas en el control de malezas en los cultivos, Schrader et al. (2024) en su artículo se realizó un estudio sobre el control de Malezas en Australia y las prácticas que realizan los agricultores para minimizar los impactos que esta puede efectuar en los sistemas agrícolas. Utilizan estrategias de control de malezas como herbicidas, técnicas mecánicas, biológicas, pero a pesar de ello las malezas continúan propagándose y la cantidad de malezas resistentes a los herbicidas y su distribución.

En el artículo de Sabzi et al. (2020) los agricultores para combatir la maleza utilizan la pulverización uniforme de herbicidas, pero hoy en día es una de las causas más comunes de intoxicación en los cultivos debido al uso innecesario en zonas que no necesitan la aplicación de gran proporción.

En la región de Tacna en el artículo de (Zegarra & Solsol, 2015), se detalla las malas hierbas sobre su rusticidad, su fácil adaptación y dominancia en los cultivos, además que los animales no las consumen y esto favorece a su multiplicación. El estudio identifica, clasifica y describe las malezas existentes en los cultivos del Instituto Los Pichones, los cuales forman parte de la Universidad Nacional 'Jorge Basadre Grohmann' de Tacna, en donde se colectaron 28 especies de malezas.

De acuerdo a los estudios mencionados, es de necesidad la estimación temprana de las malas hierbas para el tratamiento adecuado.

1.2 Descripción del problema

La producción de las aceitunas es una actividad vital para la economía local, se utilizan métodos tradicionales de cultivo y cosecha, pero los tratamientos que se le dan a las malas hierbas para la aplicación de herbicidas o los métodos tradicionales que se realizan de manera manual se ven afectadas en la producción y la calidad del producto.

En la agricultura, el control de malezas es parte fundamental para poder garantizar un rendimiento óptimo en los cultivos, pero la detección y el control preciso de las malezas pueden resultar desafiantes y costosos para los agricultores. Por lo general, los

métodos de detección de malezas se basan en la observación visual, lo que puede ser ineficiente y expuesto a errores, en especial en grandes cultivos.

Por otra parte, la erradicación de estas malezas también se lleva a cabo mediante la aplicación de agroquímicos, las malas técnicas de implementación de estos químicos como la pulverización uniforme de herbicidas. Actualmente la pulverización uniforme provoca la intoxicación de cultivos, contaminación ambiental y además el alto costo del consumo de herbicidas presentan efectos secundarios negativos sobre el suelo (Sabzi et al., 2020).

En la ciudad de Tacna en el estudio de Zegarra & Solsol (2015) se colectaron 28 especies de malezas en diferentes cultivos, entre los cultivos Olivares se encontraron 5 tipos de malezas de Hoja ancha y angosta en abundancia como la: Celandilla, Trébol Blanco, Trébol amarillo, Gainsoga paniflora Cav., Foquitos y Chirillo.

Las redes neuronales artificiales, nos ofrecen nuevas oportunidades para mejorar la identificación y el control de malezas en los cultivos, debido a su capacidad de aprender patrones complejos a partir de datos de entrada para la identificación de malezas en imágenes; sin embargo, para llevar a cabo dicho análisis, no se dispone en la ciudad de Tacna del conjunto de datos requerido, lo que dificulta optimizar la eficiencia productiva y disminuir los costos vinculados a la mano de obra; por otra parte, existe una variedad de arquitecturas y técnicas de modelos para realizar una evaluación comparativa y determinar cuál de estos enfoques es el más adecuado.

1.3 Formulación del problema

1.3.1 Problema general

¿Cómo es el rendimiento de diferentes modelos de aprendizaje profundo para la clasificación de malezas en el cultivo de Olivo en la Yarada-Los Palos de Tacna?

1.3.2 Problemas específicos

- a) ¿Cómo es el rendimiento del indicador Exactitud en la validación de diferentes modelos de aprendizaje profundo para la clasificación de malezas en el cultivo de olivo en la Yarada-Los Palos de Tacna?
- b) ¿Cómo es el rendimiento del indicador Pérdida en la validación de diferentes modelos de aprendizaje profundo para la clasificación de malezas en el cultivo de Olivo en la Yarada-Los Palos de Tacna?

1.4 Objetivos de la investigación

1.4.1 Objetivo general

Comparar el rendimiento de diferentes modelos de aprendizaje profundo para la clasificación de malezas en el cultivo de Olivo en la Yarada-Los Palos de Tacna.

1.4.2 Objetivos específicos

- a) Comparar el rendimiento del indicador Exactitud en la validación de diferentes modelos de aprendizaje profundo para la clasificación de Malezas en el cultivo de Olivo en la Yarada-Los Palos de Tacna.
- b) Comparar el rendimiento del indicador Pérdida en la validación de diferentes modelos de aprendizaje profundo aplicados en la clasificación de Malezas en el cultivo de Olivo en la Yarada-Los Palos de Tacna.

1.5 Justificación e importancia de la investigación

En la actualidad existen trabajos con buenos resultados utilizando Redes Neuronales Convolucionales para la clasificación de enfermedades y malas hierbas en cultivos. En el artículo científico de Subeesh et al. (2022), se evaluó la viabilidad de técnicas basadas en aprendizaje profundo en la identificación de malezas a partir de imágenes RGB, donde se realizaron comparación de modelos de inteligencia artificial aplicados en un campo de morrones; sin embargo, hay pocos estudios relacionados a las malas hierbas en la ciudad de Tacna.

En el estudio de Ajayi et al. (2023) utilizando un DRON para la adquisición de imágenes, el clasificador fue capaz de detectar eficazmente la presencia de malezas en los cultivos como caña de azúcar, espinacas, pimiento y plátano con un alto grado de precisión; por ello, la importancia de obtener imágenes RGB de las Malezas de la ciudad de Tacna ya que no existe un conjunto de datos para entrenar los modelos para la clasificación.

Dada la diversidad de arquitecturas de redes neuronales y el constante desarrollo de nuevas técnicas de aprendizaje profundo, resulta fundamental analizar y contrastar el desempeño de distintos modelos bajo las condiciones reales en las que se encuentran los cultivos de olivo en la ciudad de Tacna.

Adicionalmente, la clasificación automatizada de Malezas permite un uso más eficiente de los recursos humanos o de la cantidad de herbicidas necesarios en lugar de aplicar tratamientos de manera indiscriminada, los agricultores podrían dirigir sus esfuerzos solo en las áreas afectadas, reduciendo así el costo y el impacto ambiental de la agricultura (Ferraj et al., 2011).

Para la calidad del producto y competitividad, la detección y el control temprano de las Malezas pueden mejorar la calidad de los productos finales al prevenir la contaminación y el deterioro de los cultivos; por ello, la presencia de malezas puede afectar negativamente la calidad del aceite de Oliva, lo que podría reducir la competitividad de los productores en el mercado; además, estas malezas podrían competir con los olivos por minerales, agua y fertilizantes sino están bien gestionados (Ferraj et al., 2011).

1.6 Limitaciones

Entre las limitaciones para el presente trabajo se consideran:

Para la muestra, solo está conformada por conjuntos de datos de fotografías tomadas por medio de un DRON en cultivos específicamente en la Yarada de Tacna para las Malezas Olivo.

Para las Malezas solo se consideró las siguientes clases: Alta infestación, Baja Infestación y Olivo Sano.

Para la captura de las imágenes RGB mediante un DRON puede variar en el tiempo por las condiciones climáticas y estacionales; también, por la iluminación y la altura en que se posicionará el DRON, esto podría afectar en los resultados y la capacidad de poder entrenar nuestros modelos por ser entornos diferentes.

1.7 Viabilidad del estudio

El estudio es viable ya que se cuenta con las herramientas y tecnología para poder entrenar los modelos de aprendizaje profundo, también con equipo computacional necesario. Finalmente, la investigación se encuentra de acuerdo a los lineamientos vigentes.

1.8 Formulación de hipótesis

1.8.1 Hipótesis general

Existe diferencia significativa en el rendimiento de diferentes modelos de aprendizaje profundo aplicados para la clasificación de malezas en el cultivo de Olivo en la Yarada-Los Palos de Tacna.

1.8.2 Hipótesis específicas

a) Existe diferencia significativa en el rendimiento del indicador Exactitud en la validación de diferentes modelos de aprendizaje profundo para la clasificación de Malezas en el cultivo de Olivo en la Yarada-Los Palos de Tacna.

- b) Existe diferencia significativa en el rendimiento del indicador Pérdida en la validación de diferentes modelos de aprendizaje profundo para la clasificación de Malezas en el cultivo de Olivo en la Yarada-Los Palos de Tacna.

1.9 Variables

De acuerdo a Hernández Sampieri et al. (2014) una variable es una propiedad que puede cambiar y cuya variación se puede medir u observar.

Para el presente trabajo las variables consideradas son: Modelos de aprendizaje profundo y clasificación de malezas en el cultivo de olivo.

1.9.1 Definición de la variable

Variable de estudio 1: Modelos de aprendizaje profundo

Variable de estudio 2: Clasificación de malezas en el cultivo de olivo

1.10 Operacionalización de variables

Tabla 1

Tabla de operacionalización de variables

Variable	Definición conceptual	Definición operacional	Indicadores	Escala
V1: Modelos de aprendizaje profundo	Es un programa que ha sido entrenado en un conjunto de datos para reconocer determinados patrones o tomar ciertas decisiones sin más intervención humana (IBM, s.f.b)	Selección de un modelo para ser entrenado y comparado	Arquitectura	Nominal
V2: Clasificación de Malezas en el cultivo de Olivo	Indicador medible para ver qué tan efectivo es el modelo de aprendizaje profundo en la clasificación de malezas en el cultivo de olivo.	Evalúa el modelo de aprendizaje profundo en su capacidad de clasificar Malezas en el cultivo de Olivo en un conjunto de imágenes.	I1: Exactitud I2: Pérdida	Decimal Decimal

CAPÍTULO 2.

MARCO TEÓRICO

2.1 Antecedentes del trabajo de investigación

En el estudio de Ajayi et al. (2023) en el área de Lapan Gwari de Minna, estado de Níger en Nigeria, la adquisición de imágenes aéreas se realizó mediante un UAV (vehículo aéreo no tripulado), donde se utilizó un mismo modelo InceptionV2, donde se demostró que el clasificador fue capaz de detectar eficazmente la presencia de malezas en los cultivos como caña de azúcar, espinacas, pimiento y plátano con un alto grado de precisión para 10 000, 20 000, 100 000, 200 000 y 242 000 épocas. Además, se observó que la precisión más alta (98,4%) del modelo de clasificación de maleza se obtuvo en 200 mil épocas antes de que la precisión comenzara a disminuir en 242 mil épocas, donde cayó al 97%.

En el estudio de Subeesh et al. (2022) se evaluó la viabilidad de técnicas basadas en aprendizaje profundo (Alexnet, GoogLeNet, InceptionV3, Xception) en la identificación de malezas a partir de imágenes RGB de un campo de pimiento (morrón). Los resultados indicaron que todos los modelos han funcionado satisfactoriamente con una precisión general que varía entre 94,5 y 97,7%. El experimento se repitió con un número de épocas 10, 20, 30 y se observó una mejora significativa en la precisión, donde el modelo InceptionV3 superó a otros en términos de precisión, exactitud y recuperación.

En el estudio de Sabzi et al. (2020), se propone un prototipo de visión artificial en este estudio basado en el procesamiento de video y clasificadores para la identificación y clasificación de malezas de la planta de papa Marfona y cinco tipos diferentes de malezas en el campo de papa. Se encontró superioridad del clasificador híbrido ANN-CA (hybrid artificial neural network - cultural algorithm) en más del 95% en sensibilidad, precisión y especificidad en comparación con los clasificadores LDA (linear discriminant analysis), RF (Random Forest) y SVM (support vector machines).

2.2 Bases teóricas

2.2.1 Olea europea L.

De acuerdo a Picornell Buendía (2013) el olivo (*Olea europea*) perteneciente a la familia de las Oleáceas, puede llegar a alcanzar una altura de hasta 15 metros,

dentro de la que se encuentra el olivo común. Es una de las 600 especies que conforman los 29 géneros en esta familia.

2.2.2 Control de malezas

El control de malezas consiste en cambiar el equilibrio cultivo-maleza de modo que los rendimientos de los cultivos no se reduzcan económicamente (Altieri, M, 1999).

Según el Ingeniero agrónomo Ospina (2002), hace mención de cuatro tipos de sistemas de control:

a) Control Cultural

Este método agrupa prácticas específicas, que logran hacer que el cultivo sea más competitivo que las malezas, como la selección adecuada del cultivar, el uso de “semilla” o estacas de buena calidad y la protección del cultivo son las prácticas agronómicas más destacadas dentro de este sistema de control.

b) Control Manual

En este control como consecuencia del lento crecimiento inicial del cultivo, es necesario realizar varios deshierbes con implementos manuales, hasta que el cultivo se cierre completamente.

c) Control Mecánico

En este control se utiliza en combinación con el control manual o químico, que consiste en la utilización de herramientas cultivadoras, rotativas, ganchos, en este caso tiradas por tractores o animales que pasan entre las hileras y caballones; este caso se da entre 15 y 30 días de plantado el cultivo.

d) Control Químico

En este control se utilizan herbicidas, los cuales evitan el crecimiento de las malezas por un periodo dependiendo del cultivo.

2.2.3 Malezas

Según Cárdenas G. (1987), define maleza como cualquier planta que crece fuera de lugar y este aprovecha los distintos elementos nutrientes, espacio vital, agua, luz, anhídrido carbónico, que el cultivo requiere para una mayor y mejor producción. Desde el punto de vista agronómico, una planta es maleza cuando dificulta el crecimiento de las plantas deseables que se cultivaron en un momento dado.

Para Zegarra (2014), la maleza tiene muchas definiciones: planta que crece sin haberla sembrado, se propaga naturalmente y causa daño; planta que crece donde no es

deseada; planta que produce más daños que beneficios; planta que llega a ser perjudicial o indeseable en determinado lugar y cierto tiempo.

2.2.4 Tipos de malezas

En el estudio de Zegarra & Solsol (2015) se encontraron Malezas frecuentes en los cultivos de Olivo como se muestra a continuación.

Tabla 2

Malezas frecuentes en los cultivos de olivo

Tipo de Malezas	Nombre Científico
Hoja angosta (Monocotiledóneas)	- Brown catbaNiturVahl "Celandilla"
Hoja ancha (Dicotiledóneas)	- Me Notas albur Desr. "Trébol blanco", "Trébol de olor", "Trébol dulce" - Meadt s indicas (L.) All. "Trébol amarillo" - Gainsoga paniflora Cav. - Conyza bonariensis (L.) Cronq. "Foquitos" - BidetupiksaL "Chiriro, amor seco"

Nota: Malezas dominantes en los cultivos del instituto de investigación, producción y extensión agraria (Inprex) y del centro experimental agrícola (CEA) III "Los pichones" Tacna 2014

2.2.5 El cultivo del olivo en la región Tacna

En la región Tacna, se sabe que el crecimiento del área de olivos se incrementó, esto se debe a las características edafo-climáticas favorables y la existencia de mercado para el producto final. La producción nacional es exportada a diferentes países, que entre julio y enero son los meses de más exportación que coinciden con época post procesamiento (Casilla García, 2011).

Tacna, es considerada como la primera región en el cultivo del olivo, lo que se ha logrado con la explotación de terrenos en las irrigaciones de la Varada, Magollo, Ite y el Valle de Sama, donde las condiciones agro climáticas han sido favorables, razón por la cual, su productividad ha ido en incremento constante. Se observa, que menos del 10% se dedica a la producción de aceite de oliva, sin embargo, con esta información no se considera la producción artesanal, que se da en los lugares de producción. los mercados no controlados, es la salida de aceituna al departamento de Puno, Arequipa, Cusco, Lima y al país vecino de Bolivia (Casilla García, 2011).

2.2.6 Inteligencia artificial

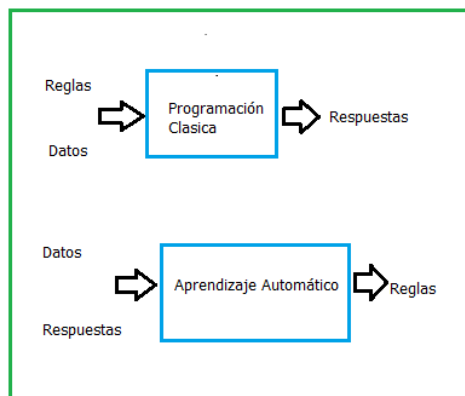
François Chollet define a la Inteligencia artificial como el esfuerzo por automatizar tareas intelectuales que normalmente realizan los humanos, además la IA es un campo muy general que abarca el aprendizaje automático y el aprendizaje profundo. (Chollet, 2017, p.2)

2.2.7 Aprendizaje automático

SGD por sus siglas en inglés, es un método de optimización ampliamente utilizado en el entrenamiento de redes neuronales. A diferencia del método de descenso de gradiente clásico, que calcula el gradiente de la función de pérdida con respecto a todos los datos de entrenamiento en cada actualización, el SGD realiza actualizaciones para cada muestra de datos de entrenamiento de manera individual. Esta característica hace que el SGD sea particularmente eficiente para grandes conjuntos de datos, ya que reduce significativamente el tiempo de cómputo por actualización (Goodfellow et al., 2016).

Figura 1

Aprendizaje automático



2.2.8 Aprendizaje profundo

El aprendizaje profundo es un subcampo específico del aprendizaje automático, una nueva forma de aprender representaciones a partir de datos que pone énfasis en aprender "capas" sucesivas de representaciones cada vez más significativas.

En el aprendizaje profundo, estas representaciones en capas se aprenden (casi siempre) a través de modelos llamados "redes neuronales", estructuradas en capas literales apiladas una tras otra, las capas más profundas pueden reconocer patrones complejos como rostros o formas de animales (Chollet, 2017).

2.2.9 Redes neuronales artificiales

Las redes pueden dividir y clasificar datos, encontrando relaciones entre ellos. A medida que el sistema o la máquina aprende de forma autónoma, utiliza esta información en otros conjuntos de datos. Cuanto más grande sea el conjunto de datos con el que trabaja el sistema, más precisas serán sus predicciones (Chollet, 2017, p.50).

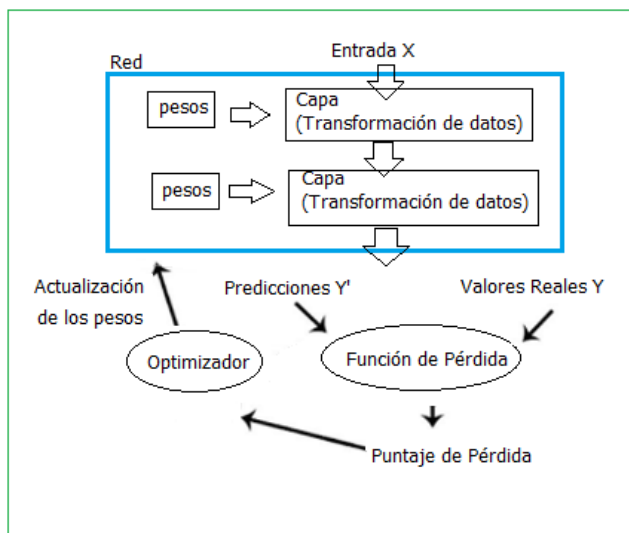
Las redes neuronales aprenden encontrando una combinación de parámetros del modelo que minimice una función de pérdida para un conjunto dado de muestras de datos de entrenamiento y sus objetivos correspondientes (Chollet, 2017)

Para realizar su funcionamiento y aprendizaje son los siguientes:

- Capas: Realizan las transformaciones de los datos indicadas por sus pesos.
- Datos de entrada y los objetivos correspondientes.
- Función de pérdida: Define el valor que se utiliza como retroalimentación para la ganancia.
- Optimizador: Determina cómo se desarrolla el aprendizaje.

Figura 2

Aprendizaje automático



2.2.10 Redes neuronales convolucionales (CNN)

Las redes neuronales convolucionales (CNN) están diseñadas para procesar diversos tipos de datos, especialmente imágenes bidimensionales, y se inspiran en la corteza visual del cerebro. En esta corteza, existe una jerarquía de células básicas: simples y complejas (Min et al., 2016). Finalmente, al añadir una red neuronal al final, se puede realizar la tarea de clasificación.

Chollet, (2017) define a las redes neuronales convolucionales como su capacidad de operar "de manera convolucional", extrayendo características de parches de entrada locales, lo que permite la modularidad de la representación y la eficiencia de los datos.

2.2.11 Modelos de aprendizaje profundo

Un modelo de inteligencia artificial es un programa que ha sido entrenado en un conjunto de datos para reconocer determinados patrones o tomar ciertas decisiones sin más intervención humana (IBM, s.f.b).

2.2.12 Arquitectura VGG16

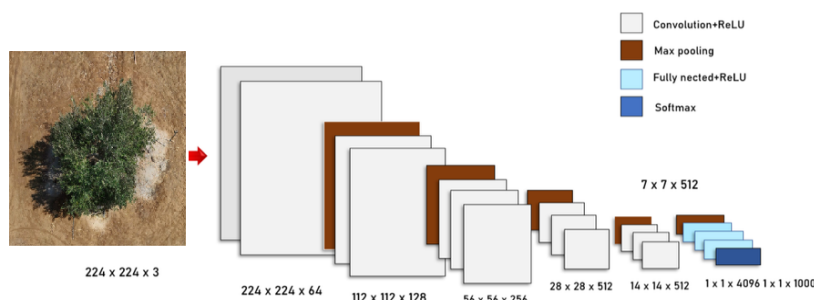
Simonyan et al. (2015) desarrollaron la arquitectura VGG16, que incluye 13 capas convolucionales para la extracción de características, como se ilustra en la Figura 00. Este modelo de clasificación puede categorizar 14 millones de imágenes en 1000 categorías diferentes (conjunto de datos Imagenet - ILSVRC 2014) con una precisión del 92.7%. Aunque existen modelos más recientes, la simplicidad de VGG16 la hace ideal para técnicas como el Aprendizaje por Transferencia y el Ajuste Fino.

El modelo y sus pesos entrenados están disponibles en Keras. Las convoluciones utilizan núcleos de 3x3, seguidas de una función de activación ReLU. Cada grupo de capas convolucionales es seguido por una capa de max-pooling, que se realiza con núcleos de 2x2, reduciendo a la mitad el tamaño del mapa de características. La arquitectura original de VGG16 incluye tres capas totalmente conectadas y una capa softmax como clasificador final, de ahí el nombre VGG16 (Simonyan, 2015).

Las 13 capas convolucionales de VGG16 son convoluciones estándar, que son fundamentales en el Deep Learning para el procesamiento de imágenes. Estas convoluciones implican deslizar un filtro (núcleo) sobre la imagen de entrada para realizar multiplicaciones y sumas, generando mapas de características que capturan diferentes patrones y jerarquías espaciales dentro de la imagen.

Figura 3

Arquitectura VGG16



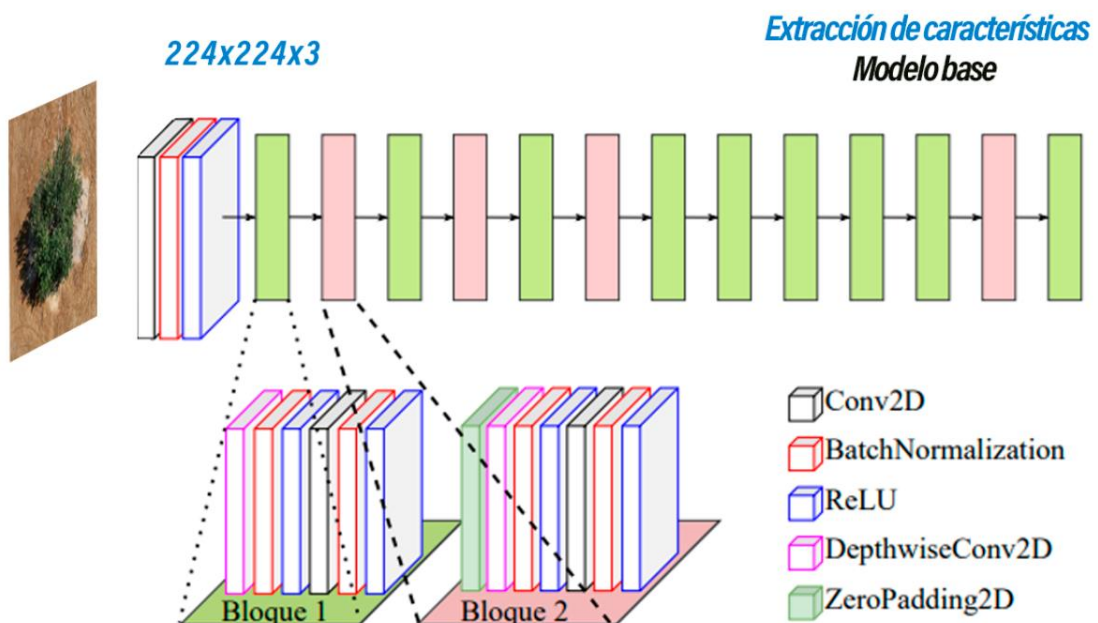
2.2.13 Arquitectura MobileNet

En 2017, ingenieros de Google presentaron la arquitectura MobileNet (Howard et al., 2017), antes de esta propuesta, las redes neuronales solían ser cada vez más profundas y complejas, incrementando el número de parámetros para mejorar la precisión. Sin embargo, esto no era adecuado para aplicaciones móviles e integradas, donde la eficiencia es crucial.

MobileNet aborda este desafío mediante la introducción de convoluciones separables en profundidad, que dividen la operación convolucional en dos etapas: convolución en profundidad y convolución puntual, manteniendo la precisión.

Figura 4

Arquitectura MobileNet



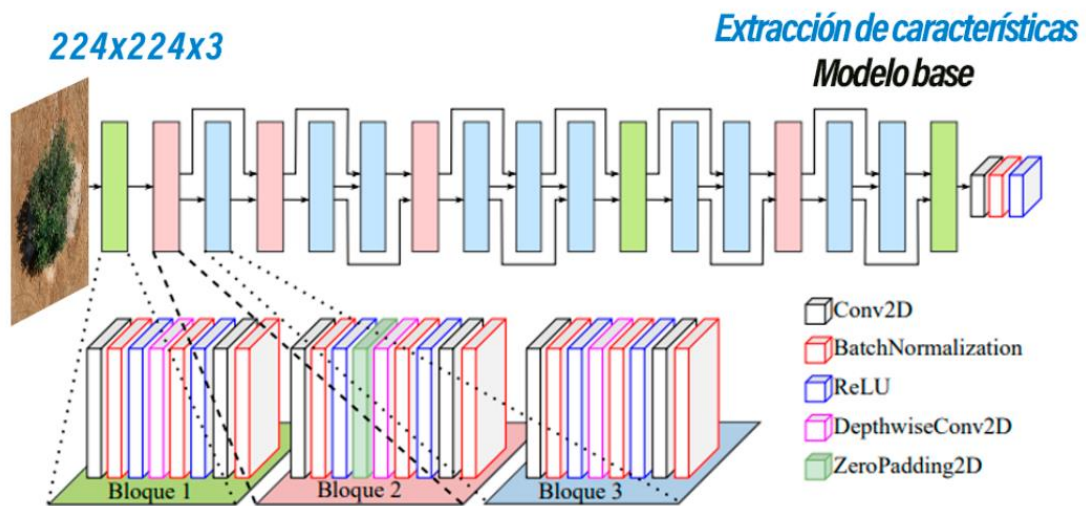
2.2.14 Arquitectura MobileNetV2

El modelo MobileNetV2, presentado en 2018, es una versión mejorada de la arquitectura MobileNet original. Esta versión incluye optimizaciones como los bloques residuales invertidos y los cuellos de botella lineales, que mejoran la capacidad de representación de características (Sandler et al., 2019).

Los bloques residuales invertidos, una variante de los bloques residuales tradicionales de ResNet, primero reducen la dimensionalidad antes de aumentarla, lo que permite capturar información detallada de manera más eficiente (Howard et al., 2017).

Figura 5

Arquitectura MobileNetV2



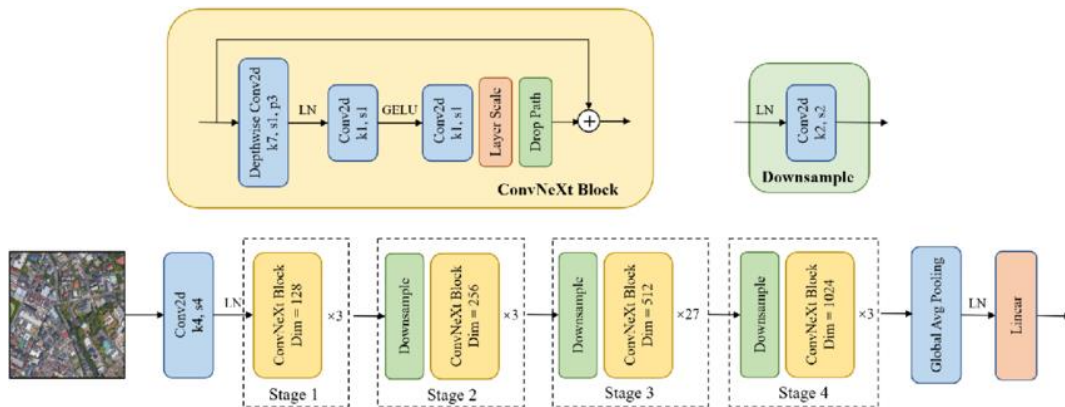
Los cuellos de botella lineales facilitan el control de la información transmitida entre capas, resultando en una arquitectura más eficiente. Estas mejoras, junto con las convoluciones separadas en profundidad, incrementaron la eficacia y la capacidad de representación de características del modelo.

2.2.15 Arquitectura ConvNeXtXLarge

ConvNeXtXLarge es una arquitectura de red neuronal convolucional diseñada para mejorar la eficiencia en tareas de visión por computadora. Basada en la simplicidad de las arquitecturas convolucionales tradicionales, incorpora mejoras en la normalización, disposición de bloques y diseño de capas para optimizar el rendimiento en problemas de clasificación de imágenes a gran escala (Liu et al., 2022). A diferencia de otras arquitecturas, como los Vision Transformers (ViTs), ConvNeXtXLarge logra un equilibrio entre precisión y eficiencia computacional, destacando en benchmarks como ImageNet (Zhuang et al., 2022). Este modelo ofrece una alternativa poderosa para la investigación en visión artificial y aprendizaje profundo.

Figura 5

Arquitectura ConvNeXtLarge



Nota: Chen, S., Ogawa, Y., Zhao, C., & Sekimoto, Y. (2023). Large-scale individual building extraction from open-source satellite imagery via super-resolution-based instance segmentation approach. *ISPRS Journal of Photogrammetry and Remote Sensing*, 195, 129-152.

2.2.16 Subajuste y sobreajuste

El subajuste ocurre cuando un modelo no logra ajustarse adecuadamente a los datos de entrenamiento, resultando en una alta tasa de error. Esto sucede durante el entrenamiento a lo largo de varias épocas cuando el modelo no puede capturar correctamente la relación entre las entradas y las salidas.

Cuando un modelo muestra buen rendimiento con los datos de entrenamiento, es crucial usar un conjunto de validación para asegurarse de que también funcione bien con datos no vistos previamente. Si el modelo tiene un buen desempeño con los datos de entrenamiento pero falla con los datos de validación, se dice que está sobreajustado, lo que indica una falta de generalización (Chollet, 2017, p.96).

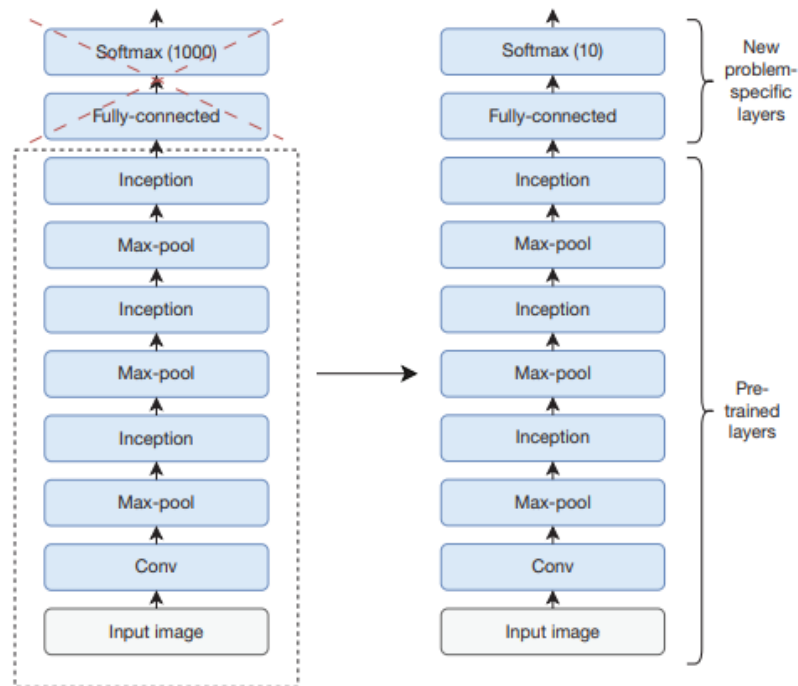
2.2.17 Aprendizaje por transferencia

De acuerdo a Ekman (2021), el aprendizaje por transferencia es una técnica de aprendizaje automático que utiliza un modelo previamente entrenado para una tarea diferente (226 p.). Esta técnica se basa en la suposición de que los datos utilizados para entrenar inicialmente el modelo están en el mismo espacio de características y tienen la misma distribución que los nuevos datos. Cuando esto es cierto, se puede utilizar la base de extracción de características congeladas de un modelo preentrenado y conectarla a una nueva red de clasificación (con pesos inicializados aleatoriamente) para entrenar solo esta última y obtener buenos resultados (Ekman, 2021).

Este procedimiento facilita que el modelo alcance la convergencia con mayor rapidez, ya sea por las limitaciones impuestas a las características congeladas o por la adecuada correspondencia entre los dominios de los conjuntos de datos.

Figura 7

Aprendizaje por transferencia



Nota: Libro Deep Learning del autor (Ekman, 2021, p.227)

En la Figura 7 en la izquierda una red inicial. Derecha: Una red nueva basada en capas preentrenadas de la red inicial, pero con las últimas dos capas reemplazadas por capas nuevas que están entrenadas para el nuevo problema.

2.2.18 Ajuste fino

Ekman nos define que el ajuste fino como una técnica aplicada en el aprendizaje por transferencia para adaptar el modelo al nuevo conjunto de datos. Este proceso implica descongelar las capas de extracción de características y entrenar el modelo completo. Es crucial realizar este procedimiento después de entrenar las capas de clasificación del modelo y seleccionar la mejor época para el ajuste fino (Ekman, 2021).

El ajuste fino mejora la extracción de características del modelo, lo que resulta en un mejor desempeño en la tarea de clasificación. La necesidad de utilizar el ajuste fino indica que el espacio de características y la distribución de los conjuntos de datos (preentrenados y nuevos) no son idénticos. La distancia entre el espacio de características y la distribución de los conjuntos de datos puede reflejarse en la mejora del modelo (Chollet, 2017, p.140).

2.2.19 Aumento de datos

De acuerdo a Chollet (2017) adopta el enfoque de generar más datos de entrenamiento a partir de muestras existentes, “aumentando” las muestras mediante una serie de transformaciones aleatorias que producen imágenes realistas. El propósito es que a lo largo del entrenamiento, el modelo no procese la misma imagen dos veces, lo cual permitirá al modelo a estar expuesto a mayor variedad de los datos y a generalizar mejor.

Figura 8

Generación de imágenes mediante aumento de datos aleatorio



Nota: Libro Deep Learning with Python p.128 de Chollet (2017)

2.2.20 Métricas de clasificación

2.2.20.1 Exactitud (Accuracy)

Número de predicciones correctas sobre el número total de predicciones.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

2.2.20.2 Pérdida (Loss)

Es una medida de cuánto error se está cometiendo en la predicción. En concreto, este valor será 0 cuando la predicción sea igual a la salida deseada. Para la clasificación multiclase, se suele utilizar la pérdida de entropía cruzada categórica, cuya definición es se muestre en la Ecuación 2 (Padilla et al., 2020).

$$Cross\ Entropy\ Loss = -\sum_i y_i \times \log(\hat{y}_i) \quad (2)$$

2.2.20.3 Precisión (Precision)

Intente responder a la siguiente pregunta: ¿Qué proporción de identificaciones positivas fue correcta? En el caso de la clasificación multiclase, se obtiene una métrica por clase (Padilla et al., 2020). Se define en la Ecuación 3.

$$Precision = \frac{TP}{TP+FP} \quad (3)$$

2.2.20.4 Sensibilidad (Recall)

Intente responder a la siguiente pregunta: ¿Qué proporción de positivos reales se identificó correctamente? En el caso de la clasificación multiclase, se obtiene una métrica por clase (Padilla et al., 2020). Se define en la Ecuación 4.

$$Recall = \frac{TP}{TP+FN} \quad (4)$$

2.2.20.5 F1-Score

Se trata de una única puntuación definida como la media armónica de la precisión y la sensibilidad. En el caso de la clasificación multiclase, se obtiene una métrica por clase (Padilla et al., 2020). Se define en la Ecuación 5.

$$F1\ Score = \frac{2 \times Precision \times Recall}{Precision+Recall} \quad (5)$$

2.3 Definiciones conceptuales

2.3.1 Matriz de confusión

Una matriz de confusión es una tabla NxN que resume el número de predicciones correctas e incorrectas que hizo un modelo de clasificación, donde N es el número de clases. Se calculan los siguientes valores: TP son verdaderos positivos, FP falsos positivos, FN falsos negativos y TN son verdaderos negativos. Teniendo en cuenta estos valores, se pueden calcular diferentes métricas de clasificación. (Giese & Seitzer, 2011).

2.3.2 Herbicidas

De acuerdo a Carod Benedico (2002) los herbicidas son compuestos químicos utilizados por el hombre para luchar contra la maleza o las “malas hierbas”.

2.3.3 Pulverización uniforme de herbicidas

La pulverización uniforme de herbicidas se refiere al proceso de aplicación de herbicidas de manera consistente y homogénea sobre un área de tratamiento específica (Sabzi et al., 2020).

2.3.4 Precisión

La precisión se define como el conjunto (tanto TP como TN) clasificado correctamente (Sabzi et al., 2020).

2.3.5 Sensibilidad

Es la fracción de muestras correctamente clasificadas que pertenecen a una determinada clase. (Sabzi et al., 2020).

2.3.6 Especificidad

Sabzi et al. (2020) indica que es la fracción de muestras correctamente clasificadas que no pertenecen a una determinada clase.

2.3.7 Verdaderos positivos (TP)

Ajayi et al. (2023) especifica el grado de veces que el clasificador clasificó correctamente una muestra positiva como positiva.

2.3.8 Verdaderos negativos (TN)

Ajayi et al. (2023) especifica el grado de veces que el clasificador clasificó con precisión una muestra Negativa como negativa.

2.3.9 Falso positivo (FP)

Ajayi et al. (2023) especifica el grado de veces que el clasificador clasificó erróneamente una muestra negativa como una muestra positiva

2.3.10 Falso negativo (FN)

Ajayi et al. (2023) especifica que implica el número de veces que el modelo clasificó incorrectamente una muestra positiva como negativa.

2.3.11 Conjunto de entrenamiento

Colección de datos utilizada para ejecutar el algoritmo de aprendizaje.

2.3.12 Época

Cada iteración sobre todos los datos de entrenamiento se denomina época. En cada época, la red calculará los gradientes de los pesos con respecto a la pérdida en el lote, y actualizará los pesos en consecuencia (Chollet, 2017, p. 53).

2.3.13 UAV

Vehículos aéreos no tripulados (Mourtzis et al., 2022).

CAPÍTULO 3.

MARCO METODOLÓGICO

3.1 Planteamiento metodológico

Según Carrasco, (2006) la investigación básica no tiene propósitos aplicativos inmediatos, solo busca ampliar y profundizar el caudal de conocimientos científicos existentes acerca de la realidad (p.43).

La presente investigación es de tipo básica, debido a que se amplían los conocimientos acerca de comparación de rendimientos de diferentes modelos en la clasificación de Malezas en el cultivo de Olivo.

De acuerdo a Hernández Sampieri et al. (2014), el enfoque cuantitativo se basa en el razonamiento deductivo o la lógica, empieza con la teoría y donde el investigador deriva expresiones lógicas llamadas "hipótesis", de acuerdo a ello, el presente trabajo toma el enfoque cuantitativo.

3.1.1 Nivel

Según Horna (2012), el propósito de la investigación descriptivo-comparativa es reconocer las variaciones o similitudes de un suceso entre dos o más conjuntos (p. 209). El presente trabajo de investigación es descriptivo-comparativo, debido a que se realizó la comparación del rendimiento en los modelos ConvNexXtXLarge, VGG16, mobileNet y mobileNetv2; a fin de reconocer las variaciones o similitudes de los modelos.

3.1.2 Diseño

La investigación no experimental lo que se hace en este tipo de investigación es observar fenómenos tal y como se dan en su contexto natural, para después analizarlos (Sampieri et al., 2014).

La presente investigación es no experimental, ya que no se realiza la manipulación de las variables; se realiza la observación de las imágenes y como se dan según la fotografías realizadas en la chacra de la Yarada - Los Palos.

Las investigaciones no experimentales se clasifican en dos categorías: transaccionales y longitudinales. En el caso de los diseños transeccionales, la recolección de datos se realiza en un único momento o punto temporal. Su propósito es describir variables y analizar su incidencia e interrelación en un momento dado (Sampieri, R et al., 2014).

El presente estudio corresponde al transaccional, porque la recolección de las fotografías se realizó en un solo periodo de tiempo correspondiente en el mes de junio del 2024 en la Yarada - Los Palos.

3.2 Población y muestra

3.2.1 Población

De acuerdo a Hernández Sampieri, R et al. (2014) “Una población es el conjunto de todos los casos que concuerdan con una serie de especificaciones” (p.174). por tal razón la población es el conjunto de recortes de imágenes de Malezas presentes en el cultivo de Olivo que se capturaron mediante un DRON.

Para las imágenes de cultivo de Olivo, solo se consideró las imágenes RGB, se obtuvieron un total de 1501 imágenes, entre estas imágenes hay 519 fotografías que no presentan Maleza y 982 fotografías que presentan malezas.

3.2.2 Muestra

De acuerdo a Carrasco (2006), “Fragmento representativo de la población, que debe poseer las mismas propiedades y características de ella” (p.238).

La muestra para el presente trabajo fue de 416 imágenes recortadas de malezas de olivo

3.2.2.1 Muestras no probabilísticas

Según Carrasco (2006), “En este tipo de muestras, no todos los elementos de la población tienen la probabilidad de ser elegidos para formar parte de la muestra” (p.243), por ello no son tan representativos.

3.2.2.2 Muestras intencionadas

Según Carrasco (2006), “El investigador procede a seleccionar la muestra en forma intencional, eligiendo aquellos elementos que considera convenientes y cree que son los más representativos” (p.243).

3.3 Equipos y materiales

Para la captura de imágenes, se utilizó por medio de una Cámara fotográfica (Dron: DJI Mavic 3 Multispectral) que cuentan con una cámara RGB de 20 megapíxeles, obturador mecánico y cámara multispectral.

Se empleó una estación de trabajo de alto rendimiento equipada con un procesador Intel® Xeon® Silver 4214 a 2.20 GHz, memoria RAM DDR4 de 64 GB, unidad de almacenamiento SSD de 1 TB y una tarjeta gráfica NVIDIA RTX A5000 con 24 GB de memoria.

Para la generación de la ortofoto se utilizó el software Agisoft Metashape Profesional V. 2.1.2 build 18146(64 bit), para los recortes de las imágenes el software Adobe Photoshop V.23.4.1. Para el entrenamiento se utilizó Python 3.10.12, Keras v2.17.0 y Tensorflow v2.17.0.

3.4 Procedimiento de las pruebas experimentales

El presente trabajo es una investigación descriptiva-comparativa, por tal motivo no es necesario realizar pruebas experimentales.

3.5 Técnicas de recolección de datos

De acuerdo a Hernández Sampieri et al. (2014) la observación “es un método de recolección de datos consiste en el registro sistemático, válido y confiable de comportamientos y situaciones observables” (p.252), ello mediante un conjunto de categorías y subcategorías.

Para este trabajo de investigación se utiliza la observación como instrumento, debido a que nos permite registrar sistemáticamente los resultados del indicador Exactitud en la validación de los modelos ConvNexXtXLarge, VGG16, MobileNet y MobileNetv2.

3.6 Técnicas para el procesamiento de datos

En la Figura 9, se observa nuestro flujo de trabajo que se ha realizado en esta investigación, estas comprenden las etapas que siguieron para poder comparar los modelos basados en la validación.

Las imágenes se obtuvieron en: CEA II LA YARADA - Fundo Santa Mónica - Los Olivos - Yarada baja POZO 17 en la ciudad de Tacna, ubicada con coordenadas 18°12'46.6"S 70°31'14.7"W.

Figura 9

Flujo de trabajo

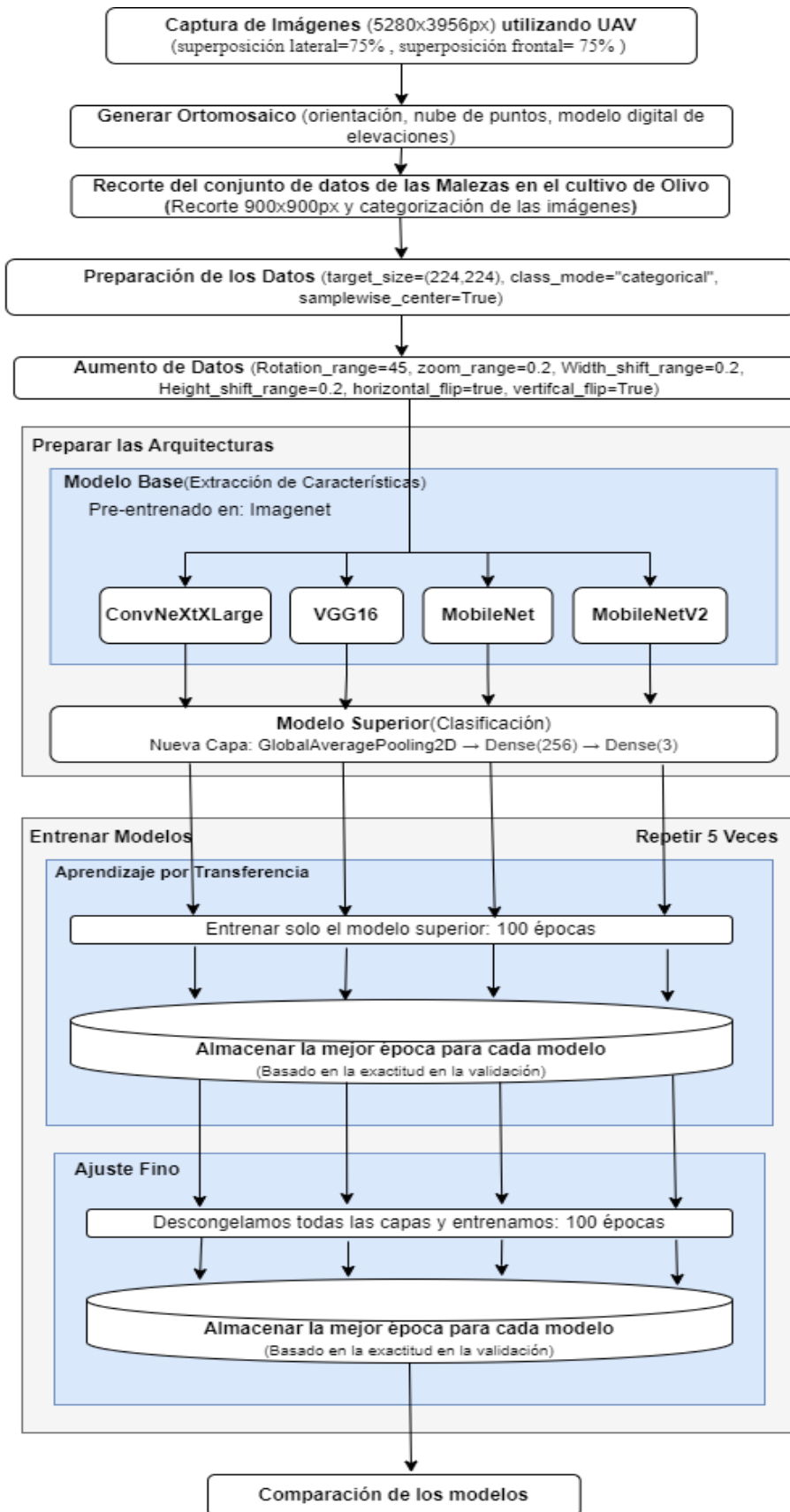


Figura 10

Lugar de toma de fotografías



Nota: Yarada los Palos, 23000: <https://maps.app.goo.gl/3yK7pP9WSddGW4dC7>, tomadas de Google Maps.

3.6.1 Captura de las imágenes

Las imágenes de las malezas en el cultivo de Olivo se tomaron en cultivos de la Yarada-Santa Mónica en la región de Tacna en Perú y fueron recolectados en abril de 2024.

A una altura de vuelo de 12 metros sobre el nivel del suelo, se desplegó un UAV DJI Mavic 3 Multispectral equipado con una cámara RGB integrada con una resolución de cámara de 20 megapíxeles para la adquisición de las imágenes. La misión de vuelo se llevó a cabo aproximadamente a las 9 a.m. para proporcionar una iluminación adecuada.

Figura 11

Captura de imágenes mediante dron DJI MAVIC 3



Para las imágenes de cultivo de Olivo que no presentan maleza se tomó un conjunto de 519 fotografías y para las imágenes de cultivo de Olivo que presentan malezas se tomó un conjunto de 982 fotografías, para la configuración del dron se utilizó una superposición lateral del 75%.

3.6.2 Generación del ortomosaico

Cada imagen originalmente se capturó por medio de un dron que tiene una resolución de 5280x3956 píxeles, seguidamente se generaron 3 Ortomosaicos utilizando el Agisoft Metashape.

En el anexo 09, se muestra el ortomosaico generado 1 y 2 que corresponden en su mayoría al cultivo de Olivo con presencia de Malezas, el ortomosaico 3 corresponde a Olivos que se encuentran libres de estas malas hierbas. Finalmente se procedió a exportar las ortofotos en formato TIF.

3.6.3 Recorte del conjunto de datos

Utilizando el Software Photoshop se realizó los recortes de cada árbol de Olivo a 900x900 píxeles y para la selección de las categorías se realizó con el apoyo de un ingeniero agrónomo como se puede observar en el anexo 10.

3.6.4 Preparación de los datos

Las imágenes RGB en el conjunto de datos fueron ajustado al tamaño de una entrada de 224x224p. y cada muestra se escaló para que su media fuera 0 para normalizar los datos y tengan una escala similar; adicional, las etiquetas (clases) de cada imagen se transformaron en variables dummy (variable binaria 0 o 1) que representa la presencia o ausencia de una categoría específica.

3.6.5 Aumento de datos

Para el aumento de datos que nos permitirá realizar unas transformaciones a las imágenes es la siguiente: 0 a 45 grados de rotación aleatoria, 0,8 a 1,2 de zoom, 0 a 0.2 de desplazamiento de altura total, 0 a 0.2 de desplazamiento de su ancho total, su giro horizontal y su giro vertical.

3.6.6 Preparar las arquitecturas

Se emplearon modelos base de ImageNet preentrenados, disponibles en Keras, con el objetivo de optimizar los resultados durante las fases de Transfer Learning y Fine Tuning. Para cada una de las arquitecturas evaluadas, se enlazó el modelo base encargado de la extracción de características con una capa de agrupación promedio global en 2D, destinada a disminuir la dimensionalidad de las características obtenidas. Se incorporó de manera secuencial una capa densa de 256 neuronas, seguida de una función de activación ReLU y, finalmente, una capa densa de 3 neuronas correspondiente a las clases definidas: Olivo Sano, Alta Infestación y Baja Infestación.

3.6.7 División de los datos

Para la división de los datos: todo el conjunto de datos (416 imágenes) se dividió en dos subconjuntos de 80% y 20%, cada uno de los cuales representa los conjuntos de datos de entrenamiento y validación para evaluar objetivamente qué tan bien se ajusta un modelo final al modelo de clasificación. A continuación, en la Tabla 3 se muestra la división de las imágenes obtenidas por cada clase.

Tabla 3

División de las imágenes por cada clase para el entrenamiento

Categoría	Entrenamiento	Validación	Total
Olivo sano	108	27	135
Alta infestación	140	35	175
Baja infestación	86	20	106
Total	334	82	416

CAPÍTULO 4. RESULTADOS

4.1 Descripción de las pruebas experimentales

Para los modelos de entrenamiento de los modelos, la pérdida fue Categorical Cross Entropy y el optimizador se utilizó ADAM. Para obtener resultados confiables, se llevó a cabo el entrenamiento durante 100 épocas para Transfer Learning y 100 épocas para Fine Tuning; adicional, el experimento se realizó 5 repeticiones y se seleccionó la mejor época de cada repetición para la fase de Transfer Learning y Fine Tuning. Finalmente se realizó la comparación de los modelos basados en la Exactitud en la validación.

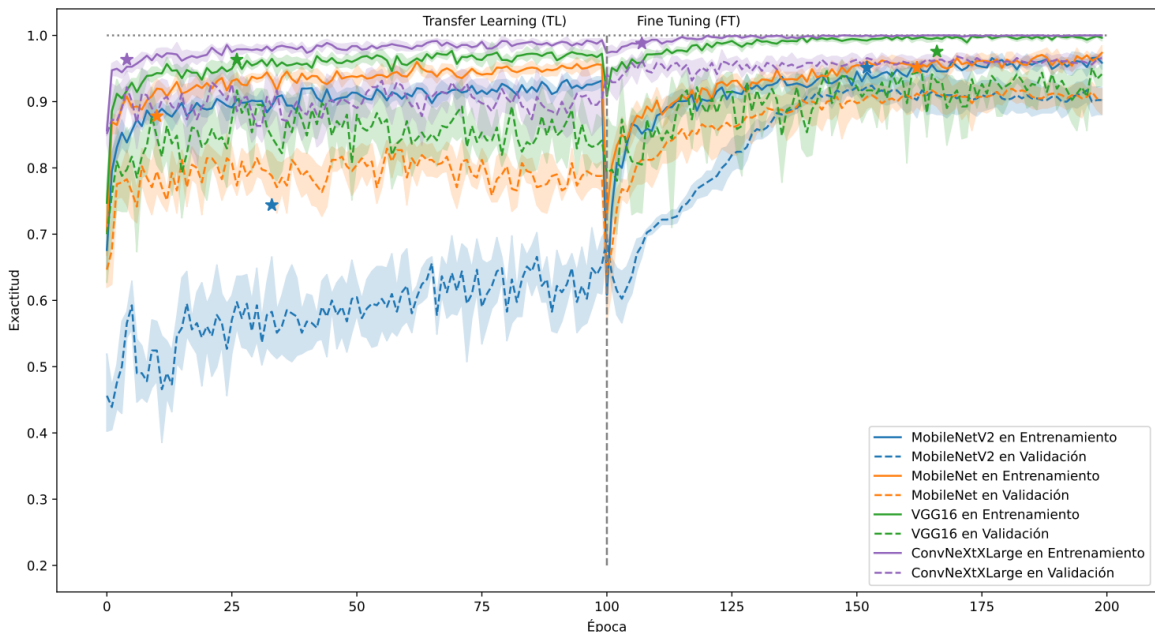
El tiempo que tomó los experimentos de las arquitecturas ConvNeXtXLarge, VGG16, MobileNet y MobileNetV2 fueron aproximadamente ≈ 8 horas 25 minutos, 2 horas 18 minutos, 2 horas 14 minutos y 2 horas con 32 minutos por cada modelo de entrenamiento. Para 5 repeticiones el tiempo total de entrenamiento fue de ≈ 15 horas, 31 minutos con 40 segundos.

4.2 Presentación y análisis de los resultados

4.2.1 Entrenamiento: etapa de aprendizaje por transferencia

Figura 12

Exactitud en el entrenamiento



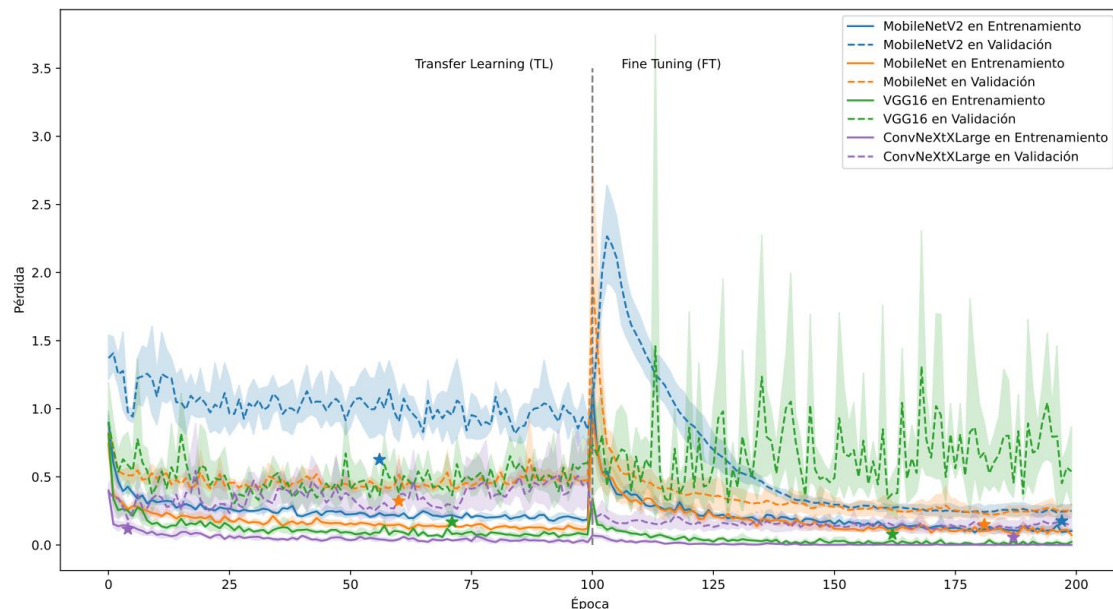
La Figura 12 nos muestra la exactitud media en los conjuntos de entrenamiento y validación donde las primeras 100 épocas representan la etapa de aprendizaje por transferencia. Además, se observa el valor de la exactitud más alta que se obtuvo en la validación.

Durante la etapa de aprendizaje por transferencia, el modelo ConvNeXtXLarge alcanzó en el conjunto de validación una precisión máxima de 0,963415 en la época 4. El modelo VGG16 obtuvo el mismo valor de exactitud (0,963415) en la época 26. En cuanto a MobileNet, su mejor desempeño fue de 0,878049 en la época 10, mientras que MobileNetV2 registró un valor máximo de 0,743902 en la época 33.

Se evidencia que, en la fase de aprendizaje por transferencia, los modelos ConvNeXtXLarge y VGG16 alcanzaron la mayor precisión en el conjunto de validación. En contraste, el modelo MobileNetV2 presentó el desempeño más bajo.

Figura 13

Pérdida en el entrenamiento



La figura 13 nos muestra la pérdida media en los conjuntos de entrenamiento y validación, donde las 100 épocas representan la etapa de aprendizaje por transferencia. Además, se observa el valor de la pérdida más baja que se obtuvo en la validación.

La pérdida en el conjunto de validación más baja que se obtuvo en la etapa de aprendizaje por transferencia utilizando el modelo ConvNeXtXLarge es de 0,116407 en la época 4, para el modelo VGG16 fue de 0,168353 en la época 71, para el modelo MobileNet fue de 0,321803 en la época 60, y utilizando MobileNetV2 fue de 0,626189 en la época 56.

Se observa que el modelo ConvNeXtXLarge se obtuvo la mejor pérdida en el conjunto de validación en la etapa de aprendizaje por transferencia. Además, se observa que el que dio un rendimiento bajo es el modelo MobileNetV2.

Se observó que los rendimientos son diferentes, siendo el mejor el modelo ConvNeXtXLarge, seguido del modelo VGG16, seguido del modelo MobileNet y por último el modelo MobileNetV2.

4.2.2 Entrenamiento: etapa de ajuste fino

En la figura 13, se observa también se nos muestra las 100 épocas que representan el entrenamiento en la etapa de ajuste fino para cada experimento.

La mayor exactitud obtenida en el conjunto de validación en la etapa de ajuste fino utilizando el modelo ConvNeXtXLarge es de 0,987805 en la época 107, para el modelo VGG16 fue de 0,975609 en la época 166, para el modelo MobileNet fue de 0,951219 en la época 162, y utilizando MobileNetV2 fue de 0,951219 en la época 152.

Se observa que el modelo ConvNeXtXLarge obtuvo la mejor exactitud en el conjunto de validación en la etapa de ajuste fino. También se observa que obtuvo un mejor rendimiento que en la etapa de aprendizaje por transferencia. Además, un rendimiento bajo en el modelo MobileNet y MobileNetV2.

La pérdida en el conjunto de validación más baja que se obtuvo en la etapa de ajuste fino utilizando el modelo ConvNeXtXLarge es de 0,056650 en la época 187, para el modelo VGG16 fue de 0,078202 en la época 162, para el modelo MobileNet fue de 0,150381 en la época 181, y utilizando MobileNetV2 fue de 0,175640 en la época 197.

Se observa que el modelo ConvNeXtXLarge se obtuvo la mejor pérdida en el conjunto de validación en la etapa de ajuste fino. Además, el modelo que dio un rendimiento bajo es MobileNetV2.

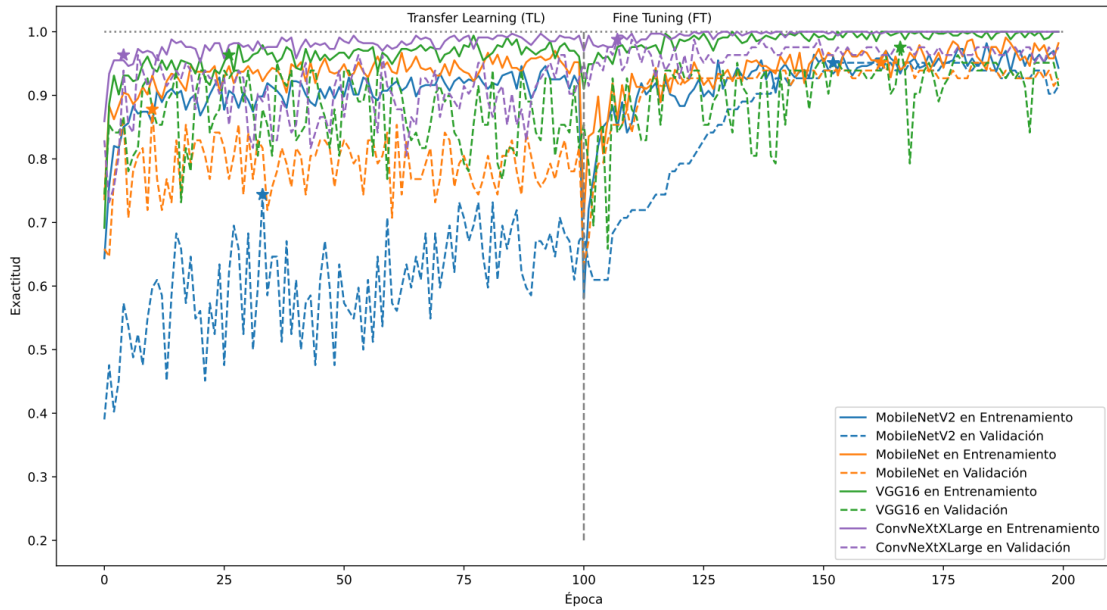
Se observó que los rendimientos son diferentes y de similar manera en la etapa de aprendizaje por transferencia, siendo el mejor el modelo ConvNeXtXLarge, seguido del modelo VGG16, seguido del modelo MobileNet y por último el modelo MobileNetV2.

Se puede observar que la etapa de ajuste fino para los 4 modelos ayuda a poder mejorar la exactitud y los valores de pérdida en el conjunto de validación.

4.2.3 Selección del modelo

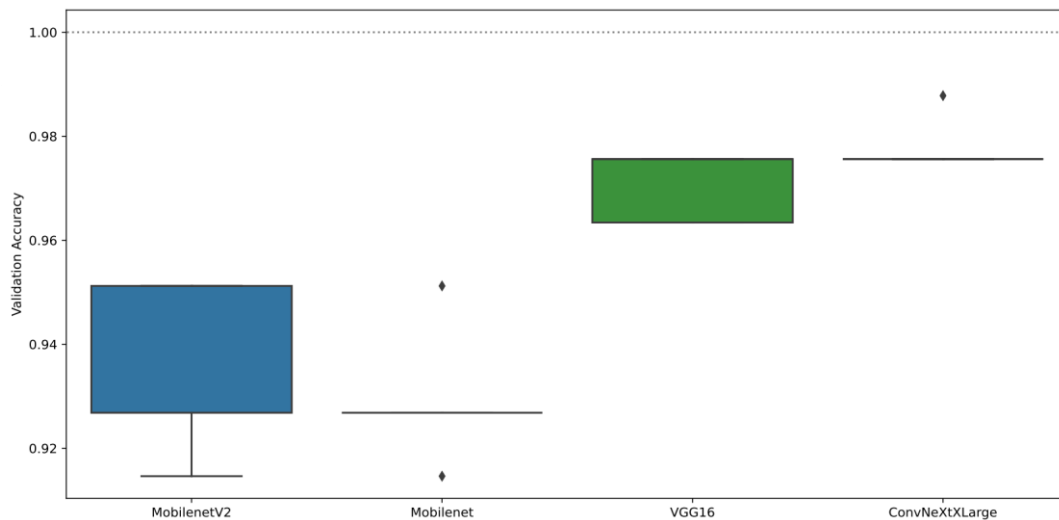
Figura 14

Exactitud de la mejor repetición



La Figura 14 presenta la mejor repetición, a partir de ello se seleccionó el mejor modelo tomando la mejor época basado en la mayor exactitud de validación; en consecuencia, se optó por el modelo ConvNeXtXLarge correspondiente a la repetición 1, época 37, el cual presentó una pérdida de 0,007454 y una exactitud de 0,997005 en el conjunto de entrenamiento, así como una pérdida de 0,072490 y una exactitud de 0,9878048 en el conjunto de validación, resultados obtenidos tras las fases de aprendizaje por transferencia y ajuste fino.

Figura 15



En la Figura 15, se muestran los valores de la métrica Exactitud de las cinco repeticiones mediante un gráfico de cajas. Se observa que el rango de valores de Exactitud del modelo ConvNeXtXLarge es poco superior a VGG16, pero muy superior a MobileNet y MobilenetV2. Además, ConvNeXtXLarge y MobileNet presenta una dispersión mínima de los datos en comparación del modelo MobilenetV2 y VGG16.

Tabla 4

Métricas sobre las repeticiones del entrenamiento en la validación

Datos	Métrica	Valor	ConvNeXtXLarge	VGG16	MobileNet	MobileNetV2
Conjunto de validación	Pérdida	Mean	0.137212	0.178765	0.242121	0.281173
		Median	0.135268	0.162672	0.236748	0.284176
	Pérdida	Min	0.094731	0.088716	0.157558	0.207756
		max	0.168376	0.307919	0.309934	0.344452
	Pérdida	iqr	0.048679	0.119324	0.112543	0.101093
		Mean	0.977642	0.970867	0.928862	0.934282
	Exactitud	Median	0.975610	0.975610	0.926829	0.926829
		Min	0.975610	0.963415	0.914634	0.914634
		max	0.987805	0.975610	0.951219	0.951219
		iqr	0.006098	0.012195	0.018293	0.030488

La Tabla 4 se visualiza los valores: máximo, mínimo, mediana, media y el rango intercuartil (iqr) de exactitud y pérdida en el conjunto de entrenamiento y validación de todas las repeticiones realizadas. Los mejores valores se encuentran marcadas con negrita, donde el conjunto de validación se observa que el modelo ConvNeXtXLarge obtiene los mejores valores de exactitud y pérdida, seguido del modelo VGG16.

En la tabla 5, se muestra los resultados de la matriz de confusión, donde se observa 1 error en la clase de alta infestación.

Tabla 5*Matriz de confusión en el conjunto de validación*

Conjunto de validación			
Clase	Alta infestación	Olivo sano	Baja infestación
Alta infestacion	34	0	1
Olivo sano	0	27	0
Baja infestacion	0	0	20

Las métricas de clasificación como: la sensibilidad, la precisión y la puntuación F1, para la evaluación posterior del modelo seleccionado se muestran a continuación

Tabla 6*Métricas de clasificación por clase en el conjunto de validación*

Conjunto de validación			
Clase	Precision	Sensibilidad	F1-Score
Alta infestacion	1.0000	0.9714	0.9855
Olivo sano	1.0000	1.0000	1.0000
Baja infestacion	0.9524	1.0000	0.9756

En la tabla 7 se muestra el tiempo promedio de las cinco repeticiones para cada modelo:

Tabla 7*Tiempo promedio de entrenamiento de modelos*

Modelos	Tiempo promedio de entrenamiento
ConvNeXtXLarge	8 horas, 25 minutos y 40 segundos
VGG16	2 horas, 18 minutos y 45 segundos
MobileNet	2 horas, 14 minutos y 30 segundos
MobileNetV2	2 horas, 32 minutos y 45 segundos

4.3 Contrastación de hipótesis

4.3.1 Prueba de normalidad

De acuerdo a González (2011), para realizar la prueba de normalidad con datos menores a treinta ($n < 30$) se sugiere usar la prueba Shapiro-Wilk; por tanto, se procedió a realizar la prueba de normalidad Shapiro-Wilk, debido a que el número de datos por cada modelo es menor a treinta. Se usó un nivel de confianza de 95% y una significancia α de 5%.

4.3.1.1 Prueba de normalidad para el indicador Exactitud

Para realizar la prueba de normalidad para ambos indicadores se utilizó los mejores desempeños de los cuatro modelos para las cinco repeticiones de cada una de ellas(anexo02).

Para el análisis de la normalidad de los datos del indicador Exactitud de los cuatro modelos, se utilizó una prueba de normalidad con un nivel de significancia de 0.05 y los resultados obtenidos fueron los siguientes:

H0: Los datos del indicador Exactitud de los modelos provienen de una distribución normal.

H1: Los datos del indicador Exactitud de los modelos no provienen de una distribución normal.

Tabla 8

Prueba de normalidad para el indicador Exactitud

Pruebas de normalidad				
Indicador	Modelo	Shapiro-Wilk		
		Estadístico	gl	Sig.
Exactitud	ConvNeXtXLarge	,861	20	,008
	VGG16			
	Mobilenet			
	MobilenetV2			

Debido a que los valores de significancia para los modelos ConvNeXtXLarge y VGG16 son menores que 0.05, se rechaza la hipótesis nula (H0) para los dos modelos, concluyendo que los datos del indicador Exactitud de estos dos modelos no provienen de una distribución normal.

4.3.1.2 Prueba de normalidad para el Pérdida

H0: Los datos del indicador Pérdida de los modelos provienen de una distribución normal.

H1: Los datos del indicador Pérdida de los modelos no provienen de una distribución normal.

Tabla 9*Prueba de normalidad para el indicador Pérdida*

Pruebas de normalidad				
Indicador	Modelo	Shapiro-Wilk		
		Estadístico	gl	Sig.
Pérdida	ConvNeXtXLarge			
	VGG16			
	Mobilenet	,942	20	,226
	MobilenetV2			

De acuerdo a la Tabla 9, el p-valor obtenido para el indicador Pérdida en los modelos es superior al nivel de significancia del 5%. En consecuencia se acepta la hipótesis nula; por tanto, se concluye que los datos siguen una distribución normal.

4.3.2 Prueba de hipótesis

4.3.2.1 Prueba de hipótesis para el indicador Exactitud

H0: No existe diferencia significativa en el rendimiento del indicador Exactitud en la validación de diferentes modelos de aprendizaje profundo para la clasificación de Malezas en el cultivo de Olivo en la Yarada- Los Palos de Tacna.

H1: Existe diferencia significativa en el rendimiento del indicador Exactitud en la validación de diferentes modelos de aprendizaje profundo para la clasificación de Malezas en el cultivo de Olivo en la Yarada- Los Palos de Tacna.

Tabla 10*Prueba de Kruskal-Wallis indicador Exactitud*

	Exactitud
H de Kruskal-Wallis	15,861
gl	3
Sig. asintótica	,001

Con base en los resultados de la Tabla 10, se obtuvo un p-valor unilateral de 0,001, el cual es menor al nivel de significancia de 0,05. Por ello, se rechaza la hipótesis nula y se concluye que existe una diferencia significativa en el desempeño de los modelos respecto al indicador de exactitud.

4.3.2.2 Prueba de post hoc para el indicador Exactitud

Así mismo en la tabla 11, se puede observar que la comparación de los modelos Mobilenet-MobilenetV2, Mobilenet-VGG16, MobilenetV2-VGG16 y VGG16-ConvNeXtXLarge son los grupos que no existe una diferencia significativa.

Tabla 11

Prueba de post hoc para el indicador Exactitud

Muestra 1 - Muestra 2	Estadístico de contraste	Error Error	Desv. Estadístico de contraste	Sig.	Sig. Ajust.
Mobilenet-MobilenetV2	-0,800	3,624	-0,221	0,825	1,000
Mobilenet-VGG16	9,100	3,624	2,511	0,012	0,072
Mobilenet-ConvNeXtXLarge	11,700	3,624	3,228	0,001	0,007
MobilenetV2-VGG16	8,300	3,624	2,290	0,022	0,132
MobilenetV2-ConvNeXtXLarge	10,900	3,624	3,007	0,003	0,016
VGG16-ConvNeXtXLarge	2,600	3,624	0,717	0,473	1,000

4.3.2.3 Prueba de hipótesis para el indicador Pérdida

De acuerdo a González (2011), el análisis de varianza (ANOVA) es la prueba paramétrica más comúnmente utilizada para determinar si las medias de dos o más muestras proceden de poblaciones idénticas; por tanto, para el indicador Pérdida utilizaremos una prueba paramétrica como el ANOVA.

H0: No existe diferencia significativa en el rendimiento del indicador Pérdida en la validación de diferentes modelos de aprendizaje profundo para la clasificación de Malezas en el cultivo de Olivo en la Yarada- Los Palos de Tacna.

H1: Existe diferencia significativa en el rendimiento del indicador Pérdida en la validación de diferentes modelos de aprendizaje profundo para la clasificación de Malezas en el cultivo de Olivo en la Yarada- Los Palos de Tacna.

Tabla 12*Prueba de ANOVA para el indicador Pérdida*

ANOVA						
		Suma de cuadrados	gl	Media cuadrática	F	Sig.
	Entre grupos	,061	3	,020	5,968	,006
Pérdida	Dentro de grupos	,055	16	,003		
	Total	,116	19			

Según los resultados presentados en la Tabla 12, el p-valor unilateral obtenido fue de 0,006, valor inferior al nivel de significancia de 0,05. En consecuencia, se rechaza la hipótesis nula y se establece que existe una diferencia significativa en el rendimiento de los modelos en relación con el indicador de pérdida.

4.3.2.4 Prueba de post hoc (Tukey HSD) para el indicador Pérdida

La prueba de Tukey HSD es una técnica estadística utilizada como un procedimiento post hoc en el análisis de varianza (ANOVA). Se debería aplicar cuando hay más de dos niveles en el factor de interés, permitiéndonos identificar las diferencias entre cada par de niveles (González, 2011).

Tabla 13*Prueba de post hoc para el indicador Pérdida*

Comparaciones múltiples						
(I) Modelo	(J) Modelo	Diferencia de medias (I-J)	Desv. Error	Sig.	Intervalo de confianza al 95%	
					Límite inferior	Límite superior
ConvNeXtX Large	VGG16	-0.044074144	0.036930974	0.640	-0.14973439	0.06158611
	Mobilenet	-0.104637207	0.036930974	0.053	-0.21029746	0.00102304
	MobilenetV2	-,144020440*	0.036930974	0.006	-0.24968069	-0.03836019
VGG16	ConvNeXtX Large	0.044074144	0.036930974	0.640	-0.06158611	0.14973439
	Mobilenet	-0.060563064	0.036930974	0.386	-0.16622331	0.04509719
	MobilenetV2	-0.099946296	0.036930974	0.067	-0.20560655	0.00571395
Mobilenet	ConvNeXtX Large	0.104637207	0.036930974	0.053	-0.00102304	0.21029746
	VGG16	0.060563064	0.036930974	0.386	-0.04509719	0.16622331
	MobilenetV2	-0.039383233	0.036930974	0.714	-0.14504348	0.06627702
MobilenetV2	ConvNeXtX Large	-,144020440*	0.036930974	0.006	0.03836019	0.24968069
	VGG16	0.099946296	0.036930974	0.067	-0.00571395	0.20560655
	Mobilenet	0.039383233	0.036930974	0.714	-0.06627702	0.14504348

CAPÍTULO 5.

DISCUSIÓN

5.1 Pruebas de validación

Los indicadores Exactitud y Pérdida presentados en este trabajo se tomaron en referencia a trabajos y artículos que abordaron la comparación de modelos de clasificación, a continuación, se presentan los siguientes trabajos donde se utilizan las métricas vistas:

El indicador Exactitud para la comparación de modelos clasificación, fue usado en el trabajo de Subeesh et al. (2022).

El indicador Pérdida para la comparación de modelos clasificación, fue usado en el trabajo de Subeesh et al. (2022).

5.2 Contraste con trabajos de investigación similares

Después del entrenamiento de los cuatro modelos de aprendizaje profundo para la clasificación de malezas en el cultivo de olivo, se obtuvieron los siguientes valores, tanto para el indicador Exactitud y Pérdida. Para el modelo ConvNeXtXLarge se obtuvo un valor máximo de 0.987805 en el indicador Exactitud y 0.094730 en el indicador Pérdida, VGG16 tuvo un valor máximo de 0.975610 en el indicador Exactitud y 0.088716 en el indicador Pérdida, MobileNet obtuvo un valor máximo de 0.951219 en el indicador Exactitud y 0.157558 en el indicador Pérdida, por último, mobileNetv2 tuvo un valor máximo de 0.951219 en el indicador Exactitud y 0.207756 en el indicador Pérdida.

Es importante mencionar que, si bien se hizo referencia a trabajos como el de Osco-Mamani & Chaparro-Cruz (2023), la comparación de los indicadores exactitud y pérdida debe interpretarse con cautela, esto debido a que el estudio de Osco-Mamani se realizó con un conjunto de datos de hojas de Olivo obtenidas en un entorno controlado (cada una de las hojas se colocaron sobre un fondo blanco), lo que difiere de las condiciones de captura de las imágenes de malezas en el cultivo de Olivo que fueron adquiridas directamente en el campo (entorno no controlado).

Dichas diferencias afectan de forma importante los valores métricos, ya que condiciones de iluminación, fondo y ruido en el dataset influyen en el desempeño observado de los modelos; por tanto, para poder afirmar mejoras en la eficacia del modelo sería necesario aplicar en un mismo entorno comparativo, ya sea utilizando el dataset de malezas para evaluar también los modelos de Osco-Mamani o entrenar los modelos en un conjunto de datos homogéneo.

En comparación con el estudio de Subeesh et al. (2022), en el cual se analizaron modelos como AlexNet, GoogLeNet, InceptionV3 y Xception para la detección de malezas en cultivos de pimientos, donde nuestros modelos ConvNeXtXLarge y VGG16 presentan altas métricas de exactitud, es importante subrayar que la comparación directa se ve dificultada por la diferencia en arquitecturas.

Finalmente, en el estudio de Ajayi et al. (2023) se centró en la detección de malezas en cultivos variados mediante imágenes aéreas adquiridas a través de UAV, utilizando el modelo InceptionV2. En ese trabajo se alcanzó una precisión del 98,4% obtenida a 200,000 épocas, lo cual evidencia que diferentes técnicas de adquisición aéreas versus in situ y el tipo de cultivo influyen notablemente en las métricas de desempeño. Esta situación refuerza la idea de que los indicadores de exactitud y pérdida dependen fuertemente del contexto experimental y del conjunto de datos empleado.

Si bien cada estudio reporta muy buenos resultados en su propio marco experimental, la comparación directa entre ellos se ve limitada por diferencias metodológicas y de condiciones de adquisición de imágenes.

CONCLUSIONES

Primero

Los resultados de los diferentes modelos de aprendizaje profundo para la clasificación de Malezas en el cultivo de Olivo para el primer indicador, el modelo con mejores resultados fue ConvNeXtXLarge, que obtuvo un valor del 98.78% de Exactitud, seguido del modelo VGG16 que se obtuvo el 97.56% de Exactitud, seguido del modelo mobileNet que obtuvo un 95.12% de Exactitud y el modelo mobileNetv2, donde se obtuvo un valor máximo del 95.12% de Exactitud, todo ello en sus 5 repeticiones para cada modelo utilizando las etapas de aprendizaje por transferencia y ajuste fino.

Segundo

Los diferentes modelos de aprendizaje profundo para la clasificación de Malezas en el cultivo de Olivo para el segundo indicador, el modelo con mejores resultados fue VGG16, que obtuvo un valor del 8.87% en el indicador Pérdida, seguido del modelo ConvNeXtXLarge que se obtuvo un 9.47% el indicador Pérdida, seguido del modelo mobileNet que obtuvo un 15.75% el indicador Pérdida, por último mobileNetv2 que se obtuvo un valor del 20.77% en el indicador Pérdida, todo ello en sus 5 repeticiones para cada modelo utilizando las etapas de aprendizaje por transferencia y ajuste fino.

Tercero

De los resultados se concluye que en la prueba de hipótesis hay una diferencia significativa en los resultados para los indicadores de Exactitud y Pérdida. El valor p en la prueba Kruskal-Wallis fue 0.001 para Exactitud y la prueba de ANOVA fue de 0.006; por tanto, ambas pruebas indican que existe una diferencia significativa en el rendimiento de los diferentes modelos de aprendizaje profundo para la clasificación de Malezas en el cultivo de Olivo en La Yarada-Los Palos – Tacna.

RECOMENDACIONES

Primero

Para futuras investigaciones se sugiere tener en cuenta la optimización de la altura del Dron, para el presente trabajo se realizó la captura de imágenes a una altura de 12 metros, se sugiere realizar pruebas a distintas alturas para determinar si a 12 metros es la altura óptima para obtener imágenes de calidad y la diferencia en los resultados en su rendimiento para la clasificación de Malezas.

Segundo

Se sugiere también tener en consideración la hora en que se realizará las capturas, ello debido al exceso de sombra de los árboles de Olivo que puede presentarse en las imágenes.

Tercero

Comparar los modelos con mejor rendimiento (ConvNeXtXLarge y VGG16) con otros modelos, para identificar cuales ofrecen mejores resultados en la clasificación de Malezas.

Cuarto

Se sugiere realizar las pruebas en diferentes campos de cultivo de Olivo para evaluar la generalización del modelo y su aplicabilidad en condiciones reales.

REFERENCIAS BIBLIOGRÁFICAS

- Ajayi, O. G., & Ashi, J. (2023). Effect of varying training epochs of a Faster Region-Based Convolutional Neural Network on the Accuracy of an Automatic Weed Classification Scheme. *Smart Agricultural Technology*, 3, 100128.
<https://doi.org/10.1016/j.atech.2022.100128>
- Altieri, M. M. (1999). *AGROECOLOGIA Bases científicas para una agricultura sustentable*.
- Cárdenas G., J. (1987). *Manual De Control De Malezas En Papa* (9.^a ed.). INIAP Archivo Historico.
- Carod Benedico, E. (2002). Herbicidas, ¿qué debemos saber los profesionales de Atención Primaria? *SEMERGEN - Medicina de Familia*, 28(8), 424-428.
[https://doi.org/10.1016/S1138-3593\(02\)74099-9](https://doi.org/10.1016/S1138-3593(02)74099-9)
- Casilla García, M. E. (2011). Análisis de los factores que influyen en la vejería del olivo (*Olea europea L.*) en la Región Tacna. *Universidad Nacional Jorge Basadre Grohmann*. <https://renati.sunedu.gob.pe/handle/sunedu/2803072>
- Chen, S., Ogawa, Y., Zhao, C., & Sekimoto, Y. (2023). Large-scale individual building extraction from open-source satellite imagery via super-resolution-based instance segmentation approach. *ISPRS Journal of Photogrammetry and Remote Sensing*, 195, 129-152.
- Chollet, F. (2017). *Deep Learning with Python* (6.^a ed.). Manning Publications.
- Díaz S., C. (2006). *METODOLOGÍA DE LA INVESTIGACIÓN CIENTÍFICA: pausas metodológicas para diseñar y elaborar el proyecto de investigación*.

- Ekman, M. (2021). *Learning Deep Learning: THEORY AND PRACTICE OF NEURAL NETWORKS, COMPUTER VISION, NATURAL LANGUAGE PROCESSING, AND TRANSFORMERS USING TENSORFLOW*. Addison-Wesley Professional.
- Ferraj, B., Teqja, Z., Susaj, L., Fasllia, N., Gjeta, Z., Vata, N., & Balliu, A. (2011). Effects of different soil management practices on production and quality of olive groves in Southern Albania. *Journal of Food, Agriculture & Environment*, 99, 430-433.
- Giese, A., & Seitzer, J. (2011). *Using a Genetic Algorithm to Evolve a D* Search Heuristic*. (p. 72).
- Guisande González, C. (2011). *Tratamiento de Datos con R, Statistica y SPSS*. Ediciones Diaz de Santos S.A.
- Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., & Adam, H. (2017). *MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications* (arXiv:1704.04861). arXiv. <http://arxiv.org/abs/1704.04861>
- Liu, Z., Mao, H., Wu, C., Feichtenhofer, C., Darrell, T., & Xie, S. (2022). A ConvNet for the 2020s. arXiv preprint arXiv:2201.03545.
- Min, S., Lee, B., & Yoon, S. (2016). Deep learning in bioinformatics. *Briefings in Bioinformatics*, bbw068. <https://doi.org/10.1093/bib/bbw068>
- Mourtzis, D., Angelopoulos, J., & Panopoulos, N. (2022). Unmanned Aerial Vehicle (UAV) manipulation assisted by Augmented Reality (AR): The case of a drone. *IFAC-PapersOnLine*, 55(10), 983-988. <https://doi.org/10.1016/j.ifacol.2022.09.483>
- Osco-Mamani, E. F., & Chaparro-Cruz, I. N. (2023). Highly accurate deep learning model for olive leaf disease classification: A study in tacna-peru. *International*

Journal of Advanced Computer Science and Applications, 14(4).

<http://dx.doi.org/10.14569/IJACSA.2023.0140494>

- Ospina, B. (2002). *La Yuca en el Tercer Milenio: Sistemas Modernos de Producción, Procesamiento, Utilización y Comercialización*. CIAT.
- Picornell Buendía, M. R., & Melero Martínez, J. M. (2013). Historia del cultivo del olivo y el aceite; su expresión en la biblia. *Ensayos: Revista de la Facultad de Educación de Albacete*, 28, 155-181.
- Sabzi, S., Abbaspour-Gilandeh, Y., & Arribas, J. I. (2020). An automatic visible-range video weed detection, segmentation and classification prototype in potato field. *Heliyon*, 6(5), e03685. <https://doi.org/10.1016/j.heliyon.2020.e03685>
- Sampieri, R. H., Collado, F., & Baptista, L. (2014). *Metodología de la Investigación* (6.^a ed.).
- Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L.-C. (2019). *MobileNetV2: Inverted Residuals and Linear Bottlenecks* (arXiv:1801.04381). arXiv. <https://doi.org/10.48550/arXiv.1801.04381>
- Schrader, S., Graham, S., Campbell, R., Height, K., & Hawkes, G. (2024). Grower attitudes and practices toward area-wide management of cropping weeds in Australia. *Land Use Policy*, 137, 107001. <https://doi.org/10.1016/j.landusepol.2023.107001>
- Simonyan, K., & Zisserman, A. (2015). *Very Deep Convolutional Networks for Large-Scale Image Recognition* (arXiv:1409.1556). arXiv. <https://doi.org/10.48550/arXiv.1409.1556>
- Subeesh, A., Bhole, S., Singh, K., Chandel, N. S., Rajwade, Y. A., Rao, K. V. R., Kumar, S. P., & Jat, D. (2022). Deep convolutional neural network models for weed detection in polyhouse grown bell peppers. *Artificial Intelligence in Agriculture*, 6, 47-54. Bell pepper

Vara Horna, A. (2012). *Desde La Idea hasta la sustentación: Siete pasos para una tesis exitosa*. Instituto de Investigación de la Facultad de Ciencias Administrativas y Recursos Humanos.

Zegarra, R. Z., & Solsol, N. A. (2015). MALEZAS DOMINANTES EN LOS CULTIVOS DEL INSTITUTO DE INVESTIGACIÓN, PRODUCCIÓN Y EXTENSIÓN AGRARIA (INPREX) Y DEL CENTRO EXPERIMENTAL AGRÍCOLA (CEA) III «LOS PICHONES» TACNA - 2014. *Ciencia & Desarrollo*, 20, Article 20. <https://doi.org/10.33326/26176033.2015.20.512>

Zhuang, B., Liu, L., Cao, Y., Zhai, X., & Lin, Z. (2022). Vision Transformers vs. Convolutional Networks: A Year Later. arXiv preprint arXiv:2201.13430.

ANEXOS

ANEXO 01: MATRIZ DE CONSISTENCIA

Título: Evaluación comparativa de modelos de Deep Learning para la clasificación de Malezas en el cultivo de Olivo utilizando datos remotos UAV en la Yarada-Los Palos de Tacna, 2024.

Problema	Objetivo	Hipótesis	Variables e indicadores
Problema General	Objetivo general	Hipótesis General	VARIABLE 1: Modelos Aprendizaje profundo
¿Cómo es el rendimiento de diferentes modelos de aprendizaje profundo para la clasificación de malezas en el cultivo de olivo en la Yarada-Los Palos de Tacna?	Comparar el rendimiento de diferentes modelos de aprendizaje profundo para la clasificación de malezas en el cultivo de Olivo en la Yarada-Los Palos de Tacna.	No existe diferencia significativa en el rendimiento de diferentes modelos de aprendizaje profundo para la clasificación de Malezas en el cultivo de Olivo en la Yarada- Los Palos de Tacna.	
Problemas específicos	Objetivos Específicos	Hipótesis Específica	VARIABLE 2: Clasificación de malezas en el cultivo de Olivo. Indicadores: I1: Exactitud I2: Pérdida
PE1: ¿Cómo es el rendimiento del indicador Exactitud en la validación de diferentes modelos de aprendizaje profundo para la clasificación de malezas en el cultivo de Olivo en la Yarada-Los Palos de Tacna?	OE1: Comparar el rendimiento del indicador Exactitud en la validación de diferentes modelos de aprendizaje profundo para la clasificación de Malezas en el cultivo de Olivo en la Yarada- Los Palos de Tacna.	H0: No existe diferencia significativa en el rendimiento del indicador Exactitud en la validación de diferentes modelos de aprendizaje profundo para la clasificación de Malezas en el cultivo de Olivo en la Yarada- Los Palos de Tacna.	
PE2: ¿Cómo es el rendimiento del indicador Pérdida en la validación de diferentes modelos de aprendizaje profundo para la clasificación de malezas en el cultivo de Olivo en la Yarada-Los Palos de Tacna?	OE2: Comparar el rendimiento del indicador Pérdida en la validación de diferentes modelos de aprendizaje profundo para la clasificación de Malezas en el cultivo de Olivo en la Yarada-Los Palos de Tacna.	H0: No existe diferencia significativa en el rendimiento del indicador Pérdida en la validación de diferentes modelos de aprendizaje profundo para la clasificación de Malezas en el cultivo de Olivo en la Yarada-Los Palos de Tacna.	
Población: 1501 imágenes en el cultivo de Olivo utilizando UAV		Tipo de investigación: básica Nivel de investigación: descriptivo-comparativo Diseño de investigación: no experimental	
Muestra: 416 imágenes recortadas			

**ANEXO 02: DATOS OBTENIDOS DEL ENTRENAMIENTO DE LOS CUATRO
MODELOS**

N°	Modelo	Val_accuracy	Val_Loss	Época	Repetición	Etapa
1	ConvNeXtXLarge	0.9756098	0.1352675	3	1	ft
2	ConvNeXtXLarge	0.9878049	0.0947307	7	2	ft
3	ConvNeXtXLarge	0.9756098	0.1683756	52	3	ft
4	ConvNeXtXLarge	0.9756098	0.1542842	32	4	ft
5	ConvNeXtXLarge	0.9756098	0.1305720	23	5	ft
6	VGG16	0.9756098	0.1626718	66	1	ft
7	VGG16	0.9756098	0.1624244	47	2	ft
8	VGG16	0.9634146	0.3079186	59	3	ft
9	VGG16	0.9634146	0.0887156	7	4	ft
10	VGG16	0.9756098	0.1818703	16	5	ft
11	MobileNet	0.9268293	0.2147326	76	1	ft
12	MobileNet	0.9268293	0.3099338	39	2	ft
13	MobileNet	0.9268293	0.2367481	46	3	ft
14	MobileNet	0.9146342	0.2874436	80	4	ft
15	MobileNet	0.9512195	0.1575579	62	5	ft
16	MobileNetV2	0.9512195	0.2841761	52	1	ft
17	MobileNetV2	0.9268293	0.2077563	98	2	ft
18	MobileNetV2	0.9146342	0.3162197	47	3	ft
19	MobileNetV2	0.9512195	0.2507283	65	4	ft
20	MobileNetV2	0.9268293	0.3444518	49	5	ft

**ANEXO 03: PRUEBA DE NORMALIDAD DE DATOS PARA LOS
INDICADORES EXACTITUD Y PÉRDIDA DE LOS MODELOS EN
SPSS 25**

Pruebas de normalidad						
	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Estadístico	gl	Sig.	Estadístico	gl	Sig.
Exactitud	,222	20	,011	,861	20	,008
Pérdida	,152	20	,200*	,942	20	,266

*. Esto es un límite inferior de la significación verdadera.
a. Corrección de significación de Lilliefors

**ANEXO 04: PRUEBA POST HOC PARA EL INDICADOR PÉRDIDA DE
LOS MODELOS EN SPSS 25**

Comparaciones múltiples						
Variable dependiente: Pérdida						
HSD Tukey						
(I) Modelo	(J) Modelo	Diferencia de medias (I-J)	Desv. Error	Sig.	Intervalo de confianza al 95%	
					Límite inferior	Límite superior
ConvNeXtLarge	VGG16	-,044074144	,036930974	,640	-,14973439	,06158611
	Mobilenet	-,104637207	,036930974	,053	-,21029746	,00102304
	MobilenetV2	-,144020440*	,036930974	,006	-,24968069	-,03836019
VGG16	ConvNeXtLarge	,044074144	,036930974	,640	-,06158611	,14973439
	Mobilenet	-,060563064	,036930974	,386	-,16622331	,04509719
	MobilenetV2	-,099946296	,036930974	,067	-,20560655	,00571395
Mobilenet	ConvNeXtLarge	,104637207	,036930974	,053	-,00102304	,21029746
	VGG16	,060563064	,036930974	,386	-,04509719	,16622331
	MobilenetV2	-,039383233	,036930974	,714	-,14504348	,06627702
MobilenetV2	ConvNeXtLarge	,144020440*	,036930974	,006	,03836019	,24968069
	VGG16	,099946296	,036930974	,067	-,00571395	,20560655
	Mobilenet	,039383233	,036930974	,714	-,06627702	,14504348

*. La diferencia de medias es significativa en el nivel 0.05.

**ANEXO 05: PRUEBA DE KRUSKAL WALLIS PARA EL INDICADOR
EXACTITUD DE LOS MODELOS EN SPSS 25**

Prueba de Kruskal-Wallis

Rangos

	Modelo	N	Rango promedio
Exactitud	ConvNeXtLarge	5	16,80
	VGG16	5	14,20
	Mobilenet	5	5,10
	MobilenetV2	5	5,90
	Total	20	

Estadísticos de prueba^a

	Exactitud
H de Kruskal-Wallis	15,788
gl	3
Sig. asintótica	,001

a. Prueba de Kruskal Wallis

b. Variable de agrupación:
Modelo

ANEXO 06: PRUEBA DE KRUSKAL WALLIS PARA EL INDICADOR EXACTITUD PARA CADA MODELO EN SPSS 25

Resumen de prueba de hipótesis

	Hipótesis nula	Prueba	Sig.	Decisión
1	La distribución de Exactitud es la misma entre las categorías de Modelo.	Prueba de Kruskal-Wallis para muestras independientes	,001	Rechazar la hipótesis nula.

Se muestran significaciones asintóticas. El nivel de significación es de ,05.

Cada nodo muestra el rango promedio de muestras de Modelo.

Muestra 1-Muestra 2	Estadístico de contraste	Error Error	Desv. Estadístico de contraste	Sig.	Sig. ajust.
Mobilenet-MobilenetV2	-,800	3,624	-,221	,825	1,000
Mobilenet-VGG16	9,100	3,624	2,511	,012	,072
Mobilenet-ConvNeXtXLarge	11,700	3,624	3,228	,001	,007
MobilenetV2-VGG16	8,300	3,624	2,290	,022	,132
MobilenetV2-ConvNeXtXLarge	10,900	3,624	3,007	,003	,016
VGG16-ConvNeXtXLarge	2,600	3,624	,717	,473	1,000

Cada fila prueba la hipótesis nula de que las distribuciones de la Muestra 1 y la Muestra 2 son las mismas.
 Se muestran las significaciones asintóticas (pruebas bilaterales). El nivel de significación es ,05.
 Los valores de significación se han ajustado mediante la corrección de Bonferroni para varias pruebas.

ANEXO 07: RECOLECCIÓN DE LOS DATOS EN EL FUNDO SANTA MÓNICA UNJBG

Figura 16

Fundo Santa Monica UNJBG - Yarada -Los Palos



Figura 17

Olivos Yarada - Los Palos



Figura 18

Olivos Yarada - Los Palos



Figura 19

Investigadores Proyecto de investigación: Inteligencia artificial del Olivo Yarada - Los Palos



ANEXO 08: IMÁGENES CAPTURADAS CON DRON DJI MAVIC 3

Figura 20

Dron



Figura 21

Imágenes capturadas con el Dron en la Yarada los Palos



ANEXO 09: ORTOFOTOS GENERADAS

Figura 22

Ortofoto generada 1



Figura 23

Ortofoto generada 2



Figura 24

Ortofoto generada 3



**ANEXO 10: CATEGORÍAS PARA EL ENTRENAMIENTO DE LOS
MODELOS**

Figura 25

Olivo con presencia de Malezas con baja Infestación



Figura 26

Olivo con presencia de Malezas con alta Infestación



Figura 27

Olivo sin Presencia de Malezas



ANEXO 10: CÓDIGO UTILIZADO EN PYTHON

```
#data aummentation
datagen_train = keras.preprocessing.image.ImageDataGenerator(
    samplewise_center=True,
    rotation_range=45,
    zoom_range=0.2,
    width_shift_range=0.2,
    height_shift_range=0.2,
    horizontal_flip=True,
    vertical_flip=True,
)
# No se realiza aumento de datos en el conjunto de validación, solo
reescalamiento
datagen_valid =
keras.preprocessing.image.ImageDataGenerator(samplewise_center=True)
#datos con dataaugmentation que se usaran en el entrenamiento
train_it = datagen_train.flow_from_directory(
    "/home/proyectoiaolivo/notebooks/google colab/malezasf5/train",
    target_size=(224, 224),
    color_mode="rgb",
    class_mode="categorical",
    batch_size=8,
)
```

```
valid_it = datagen_valid.flow_from_directory(
    "/home/proyectoiaolivo/notebooks/google colab/malezasf5/valid",
    target_size=(224, 224),
    color_mode="rgb",
    class_mode="categorical",
    batch_size=8,
)
```

```
epocas = 100
inicio = 1
fin = 5
for i in range(inicio, fin + 1):
    # base_model = applications.convnext.ConvNeXtXLarge
    # base_model = keras.applications.MobileNetV2
    # base_model = keras.applications.VGG16
    base_model = keras.applications.MobileNet(
        weights='imagenet', # Load weights pre-trained on ImageNet.
        input_shape=(224, 224, 3),
        include_top=False)
```

```

x = base_model.output
x = GlobalAveragePooling2D()(x)
x = Dense(256, activation='relu')(x)
predictions = Dense(3, activation='softmax')(x)

model = Model(inputs=base_model.input, outputs=predictions)
model.compile(optimizer='adam', loss='categorical_crossentropy',
metrics=['accuracy'])

model_checkpoint_callback = keras.callbacks.ModelCheckpoint(
    filepath= str(nombre_carpeta)+'/MobileNet_tl_'+str(i)+'_keras',
    monitor='val_accuracy',
    mode='max',
    save_best_only=True)

history = model.fit(train_it,
    validation_data=valid_it,
    epochs=epocas,
    callbacks=[model_checkpoint_callback])

hist_df = pd.DataFrame(history.history)
hist_csv_file = str(nombre_carpeta)+'/history_MobileNet_tl_'+str(i)+'_csv'
with open(hist_csv_file, mode='w') as f:
    hist_df.to_csv(f)

```

```

model=tf.keras.models.load_model(str(nombre_carpeta)+'/'+'MobileNet_tl_'+str(i)+'.keras')
    for layer in model.layers:
        layer.trainable = True
model.compile(optimizer=keras.optimizers.RMSprop(learning_rate = .00001), #
Very low learning rate
                loss="categorical_crossentropy", metrics=["accuracy"])

model_checkpoint_callback = keras.callbacks.ModelCheckpoint(
    filepath=str(nombre_carpeta)+'/'+'MobileNet_ft_'+str(i)+'.keras',
    monitor='val_accuracy',
    mode='max',
    save_best_only=True)

history = model.fit(train_it,
                    validation_data=valid_it,
                    epochs=epocas,
                    callbacks=[model_checkpoint_callback])

hist_df = pd.DataFrame(history.history)
hist_csv_file = str(nombre_carpeta)+'/'+'history_MobileNet_ft_'+str(i)+'.csv'
with open(hist_csv_file, mode='w') as f:
    hist_df.to_csv(f)

```