

UNIVERSIDAD MIGUEL HERNÁNDEZ DE ELCHE

ESCUELA POLITÉCNICA SUPERIOR DE ORIHUELA

**GRADO EN INGENIERÍA AGROALIMENTARIA Y
AGROAMBIENTAL**



**“INNOVACIÓN EN AGRICULTURA DE PRECISIÓN MEDIANTE DRONES
Y APRENDIZAJE PROFUNDO: REVISIÓN SISTEMÁTICA CON
ENFOQUE COMPUTACIONAL”**

TRABAJO FIN DE GRADO

Diciembre 2025

Autor: Guzmán Murcia Pérez

Tutora: Dámaris Núñez-Gómez

RESUMEN

“INNOVACIÓN EN AGRICULTURA DE PRECISIÓN MEDIANTE DRONES Y APRENDIZAJE PROFUNDO: REVISIÓN SISTEMÁTICA CON ENFOQUE COMPUTACIONAL”

Esta revisión sistemática, elaborada conforme al protocolo PRISMA 2020, analiza el procesamiento de imágenes captadas por drones mediante técnicas de *deep learning* en agricultura de precisión. Se aplicó un análisis semántico automatizado, desarrollado en un cuaderno *Jupyter*, sobre 6.612 artículos de Scopus y Web of Science (2019–2024), obteniendo una muestra filtrada de 105 documentos. De ellos, se seleccionaron 35 estudios para la revisión temática, estructurada en tres bloques: (i) detección de plagas y enfermedades, (ii) estimación de rendimiento y (iii) otras aplicaciones. El análisis destacó arquitecturas basadas en redes convolucionales como YOLO, Faster R-CNN o U-Net, junto con modelos *Transformer*, que representan un nuevo paradigma. Aunque se alcanzan altos niveles de precisión, su adopción práctica sigue limitada por barreras tecnológicas, económicas y operativas.

Palabras clave: UAV; deep learning; agricultura de precisión; detección de plagas y enfermedades; estimación de rendimiento.

ABSTRACT

“INNOVATION IN PRECISION AGRICULTURE THROUGH DRONES AND DEEP LEARNING: SYSTEMATIC REVIEW WITH A COMPUTATIONAL APPROACH”

This systematic review, conducted in accordance with the PRISMA 2020 protocol, analyses the processing of drone-acquired imagery using deep learning techniques in the context of precision agriculture. An automated semantic analysis was performed using a Jupyter notebook on 6,612 articles from Scopus and Web of Science (2019–2024), yielding a filtered sample of 105 documents. From this set, 35 studies were selected for thematic review, structured into three categories: (i) pest and disease detection, (ii) yield estimation, and (iii) other applications. The analysis highlighted architectures based on convolutional neural networks, such as YOLO, Faster R-CNN, and U-Net, as well as Transformer models, which represent a new paradigm. Although high levels of accuracy have been achieved, practical adoption remains limited by technological, economic, and operational constraints.

Keywords: UAV; deep learning; precision agriculture; pest and disease detection; yield estimation.

AGRADECIMIENTOS

Este trabajo nace de mi faceta perseverante, creativa y soñadora. Más que un simple ejercicio académico, representa el resultado de años de esfuerzo, constancia y crecimiento personal. Ha sido un camino lleno de retos y aprendizajes que me ha enseñado a superar las dificultades y a creer en mis propias capacidades. Este proyecto marca el primer escalón de muchos que aún quedan por subir, y me ha dado el impulso y la motivación necesarios para afrontar con fuerza los desafíos que vendrán.

Nada de esto habría sido posible sin mi familia, que ha sido siempre mi mayor apoyo y mi punto de referencia.

A Maite, mi madre, por su esfuerzo incansable, por ser el pilar que ha sostenido cada uno de mis pasos y por enseñarme, con el ejemplo, que la fuerza no consiste en no caer, sino en volver a levantarse cada vez con más serenidad. Gracias por creer en mí incluso en los días en los que yo no lo hacía, por tu paciencia infinita y por tu forma de acompañar sin imponer, siempre desde el cariño y la confianza. Todo lo que soy hoy lleva impreso tu esfuerzo y tu manera de entender la vida.

A Guzmán, mi padre, por todo lo que me transmitiste y por el ejemplo que dejaste. Gracias por tu trabajo, por tu entrega y por haber moldeado en mí la importancia de la honestidad y la responsabilidad. El último día que te vi te prometí que lo conseguiría, y este logro es también tuyo. Me gusta pensar que, de algún modo, sigues a nuestro lado, acompañando cada paso y compartiendo con nosotros este pequeño triunfo.

A Teresa, mi hermana, por haber sido mi compañera desde el primer instante, por esa conexión única que solo los mellizos comprendemos. Gracias por tu apoyo, por tu manera de entenderme sin palabras y por recordarme que los caminos compartidos siempre se hacen más llevaderos.

También quiero agradecer a mis amigos, que me han acompañado en este recorrido y me han ayudado a hacerlo más llevadero. Gracias por las risas, por la compañía en los días difíciles y por recordarme que, a pesar del cansancio, siempre hay motivos para seguir.

Por último, quiero dar las gracias a Dámaris, mi tutora, por haber confiado en mí y por acompañarme desde el primer momento en este proyecto. Aún recuerdo el día en que me propusiste convertir este trabajo en una publicación; fue una mezcla de emoción y vértigo que hoy se ha transformado en orgullo. Gracias por tu apoyo, tu paciencia y tu dedicación.

“El éxito no es la clave de la felicidad. La felicidad es la clave del éxito. Si amas lo que haces, tendrás éxito.” — Albert Schweitzer

ÍNDICE GENERAL

1. INTRODUCCIÓN.	7
2. OBJETIVOS.	9
2.1. Objetivo general.	9
2.2. Objetivos específicos.	9
3. METODOLOGÍA.	10
3.1. Fase exploratoria: análisis temático automatizado con cuaderno <i>Jupyter</i> .	10
3.1.1. Construcción del dataset preliminar.	10
3.1.2. Primera ronda: agrupamiento temático.	11
3.1.3. Segunda ronda: subagrupamiento del grupo sobre deep learning.	11
3.1.4. Aislamiento del subgrupo final y filtrado temático.	12
3.2. Fase sistemática: revisión mediante el protocolo PRISMA 2020.	12
3.2.1. Criterios de elegibilidad.	12
3.2.2. Fuentes de información.	13
3.2.3. Estrategia de búsqueda.	13
3.2.4. Proceso de selección.	13
3.2.5. Extracción de datos.	13
3.2.6. Variables recogidas.	13
3.2.7. Evaluación del riesgo de sesgo.	13
3.2.8. Medidas del efecto.	13
3.2.9. Síntesis de resultados.	13
3.2.10. Evaluación de certeza y sesgo por publicación.	13
4. RESULTADOS Y DISCUSIÓN.	14
4.1. Análisis temático automatizado del corpus.	14
4.1.1. Análisis cuantitativo del conjunto de artículos.	14
4.1.2. Análisis cuantitativo específico del grupo sobre deep learning.	24
4.2. Revisión sistemática.	30
4.2.1. Selección de los estudios.	30
4.2.2. Características de los estudios incluidos.	35
4.2.3. Evaluación del riesgo de sesgo.	37
4.2.4. Resultados de los estudios individuales.	37
4.2.5. Resultados de la síntesis.	37
4.2.6. Evaluación del sesgo por publicación.	37

4.2.7. <i>Certeza de la evidencia.</i>	37
4.2.8. <i>Discusión general e implicaciones.</i>	37
4.3. <i>Revisión temática.</i>	38
4.3.1. <i>Detección plagas y enfermedades.</i>	38
4.3.1.1. Modelos convolucionales clásicos.	38
4.3.1.2. Modelos de detección: comparación entre una y dos etapas.	42
4.3.1.3. Modelos combinados.	45
4.3.1.4. Modelos con mecanismos de atención.	48
4.3.1.5. Modelos especializados.	50
4.3.1.6. Modelos conceptuales.	52
4.3.2. <i>Estimación de rendimiento agrícola.</i>	54
4.3.2.1. Modelos de detección: una etapa.	54
4.3.2.2. Modelos de detección: comparación entre una y dos etapas.	60
4.3.2.3. Modelos combinados.	61
4.3.2.4. Modelos con mecanismos de atención.	64
4.3.3. <i>Otras aplicaciones agronómicas.</i>	71
4.3.3.1. Modelos de optimización operativa.	71
4.3.3.2. Modelos de monitoreo fenológico.	72
4.3.3.3. Modelos de automatización.	73
4.3.3.4. Modelos de detección de flora adventicia.	75
5. CONCLUSIÓN.	77
6. BIBLIOGRAFÍA.	78

ÍNDICE ANEXOS

Anexo I - Glosario técnico automatizado mediante inteligencia artificial.	85
Anexo II - Código fuente del cuaderno <i>Jupyter</i> para el preprocesamiento del <i>dataset</i> .	97
Anexo III - Bibliografía válida, excluida por criterios normativos.	112

1. INTRODUCCIÓN.

La agricultura atraviesa un momento decisivo. En pleno siglo XXI, el desafío de producir alimentos suficientes sin agotar los recursos naturales ni degradar los ecosistemas ha dejado de ser una preocupación técnica para convertirse en un compromiso global. La búsqueda de una agricultura más inteligente, eficiente y sostenible ha impulsado una profunda transformación en la forma de entender, observar y gestionar los sistemas agrarios. Este cambio se enmarca en el paradigma de la agricultura de precisión, definida de acuerdo con la Organización de las Naciones Unidas para la Alimentación y la Agricultura (FAO) como un sistema de gestión que utiliza tecnologías de información para observar, medir y responder a la variabilidad espacial y temporal de los cultivos, con el propósito de optimizar los recursos y reducir los impactos ambientales.

En esta nueva era de digitalización rural, la observación del territorio ha pasado de ser una práctica manual y fragmentada a convertirse en un proceso continuo, automatizado y guiado por datos. En este contexto, los vehículos aéreos no tripulados (UAV, por sus siglas en inglés) han adquirido un papel decisivo. Como señalan Velusamy et al. (2021), los drones representan una de las transformaciones más relevantes en la agricultura moderna: su capacidad para capturar imágenes de alta resolución desde baja altitud, con gran flexibilidad operativa y a un coste accesible, los convierte en aliados imprescindibles para la detección temprana de plagas, el seguimiento del crecimiento vegetal o la evaluación del estrés hídrico. Frente a las plataformas satelitales, los UAV ofrecen inmediatez, detalle y autonomía, abriendo un horizonte en el que la observación aérea deja de ser un recurso experimental para consolidarse como una herramienta operativa en la gestión agraria.

Sin embargo, disponer de más datos no implica necesariamente disponer de más conocimiento. El verdadero desafío reside en interpretar esa información masiva de forma coherente y útil. En este punto cobra protagonismo el aprendizaje profundo o *deep learning* (DL, por sus siglas en inglés), una rama avanzada del análisis computacional que, como exponen Kamilaris y Prenafeta-Boldú (2018), ha transformado el procesamiento de imágenes al permitir que los algoritmos aprendan por sí mismos las características relevantes de los datos. Las redes neuronales profundas han sustituido el antiguo trabajo manual de diseñar descriptores o índices vegetativos, logrando una comprensión automatizada del entorno agrícola. Gracias a ello, hoy es posible identificar

enfermedades foliares, estimar rendimientos o clasificar tipos de cultivo con niveles de precisión inimaginables hace apenas una década.

La unión entre visión aérea y análisis computacional avanzado marca así un nuevo punto de inflexión: los UAV capturan lo que el ojo humano no alcanza, y el aprendizaje profundo convierte esas observaciones en conocimiento agronómico aplicable. Esta sinergia no solo mejora la productividad o la eficiencia, sino que redefine el modo en que concebimos la relación entre tecnología y sostenibilidad. La agricultura deja de ser un sistema reactivo, que responde a los problemas cuando ya han ocurrido, para transformarse en un sistema predictivo y adaptativo, capaz de anticipar decisiones y optimizar recursos antes de que aparezcan las pérdidas.

Por otro lado, la rápida expansión de estas tecnologías ha generado un ecosistema científico vasto y fragmentado. Miles de estudios abordan de manera independiente las mismas preguntas, utilizando metodologías, sensores y arquitecturas de red heterogéneas. Esta dispersión dificulta identificar las tendencias reales y comprender hasta qué punto los avances descritos son transferibles al campo. En respuesta a este vacío, la presente investigación propone una revisión sistemática asistida por herramientas de ciencia de datos, que combina la metodología PRISMA 2020, Page et al. (2021) con técnicas de procesamiento de lenguaje natural (NLP, por sus siglas en inglés) implementadas en un cuaderno *Jupyter*, Álvarez-Martínez et al. (2023). Esta integración permite analizar de forma objetiva y reproducible grandes volúmenes de literatura científica, agrupar temáticas mediante *clustering* automatizado y revelar patrones de investigación difíciles de detectar mediante métodos tradicionales.

El objetivo de este trabajo no es solo sintetizar el conocimiento existente, sino reinterpretarlo desde una mirada computacional, explorando cómo la ciencia de datos puede aplicarse no solo al estudio de los cultivos, sino también al proceso de revisión científica. En este sentido, el proyecto representa una doble innovación: por un lado, en el objeto de estudio, al analizar la convergencia entre UAV y aprendizaje profundo en agricultura de precisión; y por otro, en la metodología empleada, al incorporar técnicas de análisis de datos reproducibles como instrumento de análisis académico.

En definitiva, esta memoria pretende contribuir al avance de un modelo agrícola más informado, sostenible y tecnológicamente robusto, donde los datos aéreos y el análisis computacional constituyan las nuevas raíces de una agricultura guiada por el conocimiento.

2. OBJETIVOS.

2.1. Objetivo general.

Analizar de forma crítica y estructurada las aplicaciones de modelos de *deep learning* sobre imágenes captadas por UAV en el contexto de la agricultura de precisión, valorando no solo su eficacia técnica y su aplicabilidad operativa, sino también su contribución a la sostenibilidad agronómica y su viabilidad real para ser implementadas por el sector.

2.2. Objetivos específicos.

- Revisar el estado del arte sobre el uso conjunto de UAV y técnicas de *deep learning* en el ámbito agrícola, con el fin de conocer las líneas de investigación actuales y las tendencias tecnológicas emergentes en el sector.
- Examinar críticamente los enfoques metodológicos presentes en la literatura científica, prestando atención a las estrategias de procesamiento de imágenes, entrenamiento de modelos y validación de resultados.
- Explorar las posibles aplicaciones agronómicas del uso combinado de ambas tecnologías, enmarcando su utilidad potencial en distintos contextos de producción, con atención a factores como tipo de cultivo, escala de explotación o condiciones ambientales.
- Valorar la viabilidad operativa, económica y técnica de estas soluciones, identificando los principales retos que pueden limitar su adopción por parte del sector, así como los recursos necesarios para su implementación.
- Detectar vacíos de conocimiento y proponer futuras líneas de investigación, que permitan avanzar hacia soluciones más accesibles, sostenibles y eficaces en el marco de una agricultura digital e inteligente.

3. METODOLOGÍA.

La metodología empleada en esta revisión se estructuró en dos fases complementarias: una fase exploratoria de análisis temático asistido por técnicas de NLP y una fase sistemática basada en el protocolo PRISMA 2020. Esta estrategia dual permitió abordar de forma eficaz tanto el volumen como la complejidad temática de la literatura científica relacionada con la aplicación de *deep learning* en agricultura de precisión. Para facilitar la comprensión, los tecnicismos se recogen en un glosario incluido en los anexos.

Durante la primera fase, se utilizó un cuaderno *Jupyter* adaptado a partir del propuesto por Álvarez-Martínez et al. (2023), el cual implementó técnicas de vectorización mediante una métrica de ponderación basada en la frecuencia de aparición de un término y su distribución inversa en el *corpus* (TF-IDF, por sus siglas en inglés), agrupamiento temático con el algoritmo K-means y análisis de componentes principales (PCA, por sus siglas en inglés). Esta herramienta facilitó una segmentación automatizada y objetiva de las áreas de estudio, reduciendo la carga subjetiva del investigador y mejorando la reproducibilidad del cribado inicial.

La segunda fase consistió en una revisión sistemática de un subconjunto temático específico, siguiendo las directrices del protocolo PRISMA 2020 (Page et al., 2021). Esta fase aportó rigor metodológico al proceso de selección y permitió sistematizar la extracción y síntesis de los estudios más relevantes.

3.1. Fase exploratoria: análisis temático automatizado con cuaderno *Jupyter*.

3.1.1. Construcción del dataset preliminar.

La base de datos inicial se construyó a partir de las plataformas de prestigio internacional de Scopus (www.Scopus.com) y Web of Science (www.webofscience.com) mediante la siguiente estrategia de búsqueda booleana:

("drones" OR "drone" OR "UAV" OR "Unmanned Aerial Vehicles") AND
("agriculture" OR "precision agriculture" OR "farming")

Se aplicaron filtros para limitar los resultados a artículos científicos y revisiones publicadas en el idioma inglés, restringidas al área temática de “*Agricultural and Biological Sciences*”, con el objetivo de evitar la inclusión de estudios fuera del alcance temático.

Los resultados fueron descargados en formato .xlsx y .xls, conteniendo metadatos bibliográficos como: autores, título, resumen y año de publicación. Para su tratamiento, se desarrolló un cuaderno *Jupyter* encargado de la limpieza y estandarización de los datos. Este proceso incluyó la conversión de formatos, lectura y normalización de columnas clave ("*Title*", "*Year*" y "*Abstract*"), así como la eliminación de entradas incompletas. El resultado fue un único archivo Excel estructurado y limpio, preparado para su análisis temático posterior. La herramienta puede ser consultada en el siguiente [enlace](#) o en anexo.

3.1.2. Primera ronda: agrupamiento temático.

Se fusionaron los campos "*Title*" y "*Abstract*" en un único texto que fue procesado mediante la librería NLTK de Python, aplicando tres etapas sucesivas:

- Limpieza del texto: se llevó a cabo la conversión a minúsculas y eliminación de signos de puntuación, cifras, enlaces, palabras vacías comunes en inglés y frases genéricas sin valor informativo. También se excluyeron palabras de longitud inferior a tres o superior a veinte caracteres.
- Reducción morfológica: mediante técnicas de *stemming*, se agruparon términos derivados bajo una misma raíz léxica.
- Tokenización y ponderación: los textos se segmentaron en palabras y se vectorizaron utilizando TF-IDF, que asigna un peso a cada término según su relevancia relativa.

Posteriormente, se aplicó el algoritmo K-means de la biblioteca sklearn para clasificar los textos en grupos temáticos. El número óptimo se determinó mediante el método de *Elbow* y el índice de silueta. Cada grupo se analizó cuantitativamente mediante nubes de palabras, visualización bidimensional con PCA y se generó una gráfica con la evolución temporal de los artículos por grupo.

3.1.3. Segunda ronda: subagrupamiento del grupo sobre deep learning.

Tras identificar el grupo más relevante para los objetivos del estudio (relacionado con DL), se procedió a su aislamiento y posterior subagrupamiento temático, aplicando la misma metodología de procesamiento. El cuaderno generó un nuevo *dataset* con una variable adicional "*Sub-cluster*" que indicaba el subgrupo temático correspondiente a cada documento.

3.1.4. Aislamiento del subgrupo final y filtrado temático.

Para afinar la muestra, se seleccionaron únicamente aquellos artículos que contenían el término "*deep learning*" en el título. Se estableció una ventana temporal entre 2019 y 2024 de acuerdo con los parámetros científicos aceptados. El conjunto final, compuesto por 105 artículos, fue considerado adecuado para la revisión sistemática, tanto por su relevancia como por su manejabilidad en el marco de un Trabajo de Fin de Grado.

3.2. Fase sistemática: revisión mediante el protocolo PRISMA 2020.

3.2.1. Criterios de elegibilidad.

Para asegurar la relevancia temática y la calidad metodológica de los estudios incluidos, se establecieron los siguientes criterios de elegibilidad. Se consideraron únicamente artículos originales de investigación y revisiones sistemáticas publicados en inglés y en revistas científicas revisadas por pares. La búsqueda se limitó al periodo comprendido entre 2019 y 2024, con el fin de recoger investigaciones recientes. Asimismo, se filtraron únicamente aquellos trabajos indexados en el área de "*Agricultural and Biological Sciences*", para evitar desviaciones temáticas.

En cuanto al contenido, se incluyeron estudios que abordan la aplicación de técnicas de *deep learning* sobre imágenes obtenidas mediante UAV en el ámbito de la agricultura de precisión, priorizando aquellos que mencionaban expresamente este enfoque en el título.

Finalmente, y debido a la limitación metodológica establecida por la normativa del trabajo, la cual restringe a un máximo de 35 artículos citables, se implementó una segunda fase de selección mediante un sistema de puntuación estructurado. Este procedimiento permitió priorizar de forma objetiva aquellos estudios más representativos entre los que cumplían los criterios de inclusión. El sistema de puntuación consideró cuatro dimensiones: número de citas (0-3 pts.), como indicador de impacto; diversidad temática (0-2 pts.), para garantizar líneas de investigación complementarias y no redundantes; año de publicación (0-1 pts.), priorizando trabajos recientes publicados a partir de 2021; criterio adicional (± 2 pts.), que valoró la innovación metodológica y la actualidad de estudios poco citados. Esta ponderación quedó registrada en la base de datos con una variable denominada "Intención", cuya justificación y aplicación específica se desarrolla en el apartado de resultados y discusión, dentro de la sección selección de estudios.

3.2.2. Fuentes de información.

Las plataformas utilizadas fueron Scopus y Web of Science. Ambas búsquedas se realizaron el 28 de mayo de 2025.

3.2.3. Estrategia de búsqueda.

Se empleó la misma cadena booleana descrita en la fase exploratoria, con los filtros correspondientes.

3.2.4. Proceso de selección.

El proceso de selección constó de dos fases: una segmentación automatizada mediante NLP y K-means sobre el *corpus* preliminar, y una revisión manual centrada en artículos con “*deep learning*” en el título, asegurando la coherencia temática mediante validación semántica. Se priorizó la relevancia y especificidad de los estudios seleccionados.

3.2.5. Extracción de datos.

Los artículos incluidos se registraron en una hoja Excel con las variables: título, año, número de citas, resumen, tipo de estudio, categoría temática, exclusión y motivo.

3.2.6. Variables recogidas.

Se recopilaron diversas variables tales como cultivo, tarea, modelo y arquitectura DL, tipo de sensor, resolución imagen, *dataset* empleado, métricas, limitaciones y aportes.

3.2.7. Evaluación del riesgo de sesgo.

No se aplicó ninguna herramienta formal de evaluación del sesgo, ya que el cribado fue temático y manual, y no se trataba de un metaanálisis comparativo.

3.2.8. Medidas del efecto.

Este ítem no resultó aplicable puesto que no se calcularon medidas del efecto como razón de riesgos, ya que no se realizó comparación entre grupos experimentales.

3.2.9. Síntesis de resultados.

Se desarrolló una síntesis organizada en tres bloques temáticos: (i) detección de plagas y enfermedades, (ii) estimación de rendimiento y (iii) otras aplicaciones.

3.2.10. Evaluación de certeza y sesgo por publicación.

No se evaluó el sesgo de publicación de manera formal, debido al carácter cualitativo de la revisión.

4. RESULTADOS Y DISCUSIÓN.

4.1. Análisis temático automatizado del *corpus*.

4.1.1. Análisis cuantitativo del conjunto de artículos.

El contenido lingüístico de los 1.316 artículos procedentes de Scopus y los 5.350 artículos obtenidos de Web of Science, ambos centrados en el uso de vehículos aéreos no tripulados en el ámbito de la agricultura de precisión, fue procesado y analizado empleando el cuaderno *Jupyter* alojado en Google Colab descrito previamente en el apartado de metodología.

La integración de ambas bases de datos dio lugar a un total de 6.666 artículos. No obstante, durante el acondicionamiento del *dataset* preliminar mediante la herramienta desarrollada específicamente para este estudio, se eliminaron aquellos registros incompletos que impedían su adecuado procesamiento, obteniéndose finalmente un conjunto de 6.612 artículos válidos para el análisis temático.

En el transcurso del análisis textual, se aplicó un proceso de lematización con el propósito de reducir todas las palabras a su raíz común, evitando así distorsiones asociadas a las distintas formas morfológicas de un mismo término. Paralelamente, se depuró el *corpus* eliminando aquellos términos carentes de carga semántica relevante, lo que favoreció una interpretación más precisa del conjunto textual.

Una vez finalizado el preprocesamiento del *corpus*, se identificaron los términos de mayor frecuencia, cuya representación visual (Figura 1) facilitó la delimitación de la orientación temática del conjunto documental, estrechamente vinculada al eje central de esta revisión. El término "UAV" sobresalió con claridad como el más frecuente, lo que pone de manifiesto la centralidad de los vehículos aéreos no tripulados en los estudios analizados. Le siguieron "imag", "crop" y "agricultur", términos que destacan la relevancia del procesamiento de imágenes y del contexto agrícola. Asimismo, fueron recurrentes conceptos asociados a la modelización y predicción como "model", "yield" y "detect", junto con términos técnicos como "aerial", "unman" y "vehicl". Este patrón de frecuencias confirma que el enfoque dominante en la literatura combina la adquisición remota de datos con metodologías de análisis predictivo aplicadas a cultivos y sistemas agrícolas.

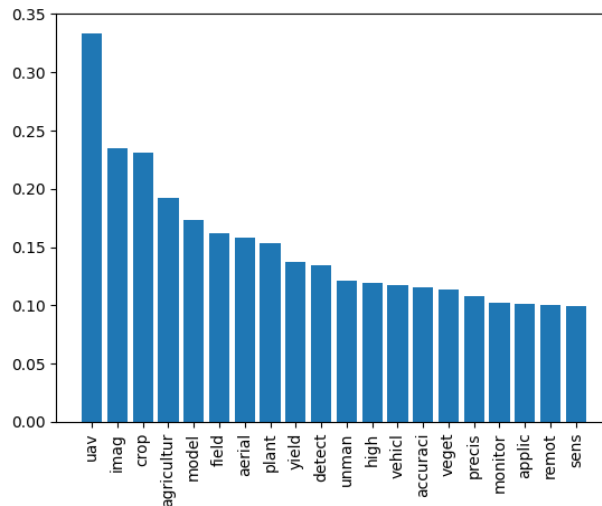


Figura 1. Gráfico de barras con las frecuencias relativas de las 20 palabras más abundantes en la bibliografía analizada.

En una representación complementaria de las raíces léxicas más frecuentes en el *corpus* se examinó la nube de palabras generada a partir de los artículos analizados (Figura 2), la cual reforzó visualmente los patrones temáticos previamente identificados, volviendo a destacar la prominencia de términos como "uav", "imag", "crop" y "agricultur". Este tipo de visualización gráfica hizo posible, además, captar de forma inmediata la riqueza y diversidad temática del *corpus*, incluyendo conceptos relacionados con el monitoreo, la predicción y evaluación de cultivos mediante sensores embarcados. La aparición destacada de términos como "model", "yield", "detect" o "accuraci" puso de manifiesto el enfoque técnico predominante en muchos de los estudios, centrado en la optimización de los procesos agrícolas a través del análisis de datos geoespaciales obtenidos por drones.

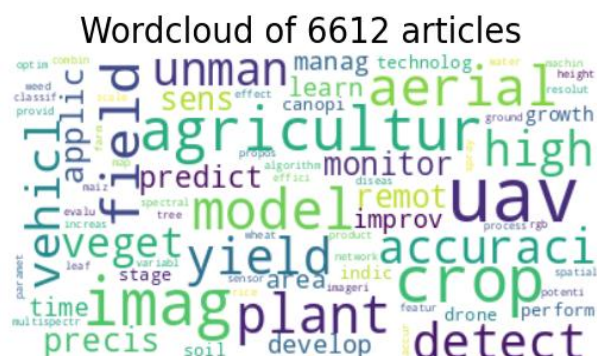


Figura 2. Nube de palabras como representación visual de las 75 palabras más comunes; un mayor tamaño de fuente indica una frecuencia relativa más alta de aparición.

Para organizar temáticamente el conjunto de artículos, se empleó un enfoque metodológico combinado que integró el método *Elbow* (Figura 3), el análisis del coeficiente de silueta (Figura 4) y los gráficos de silueta (Figura 5) con el objetivo de estimar el número óptimo de agrupaciones. El primer paso consistió en identificar visualmente el punto de inflexión más significativo en la curva de inercia del método *Elbow*, el cual coincidió con el valor $K = 4$. A continuación, se validó esta elección mediante la evaluación de la cohesión interna y la separación entre los grupos. Para dicho valor de K , se constató que todos los grupos superaban el valor medio de silueta (línea roja) y presentaban una forma y un tamaño relativamente homogéneos. Además, el valor medio de silueta alcanzó la unidad en esta configuración, lo que respaldó su idoneidad como división óptima del *corpus*.

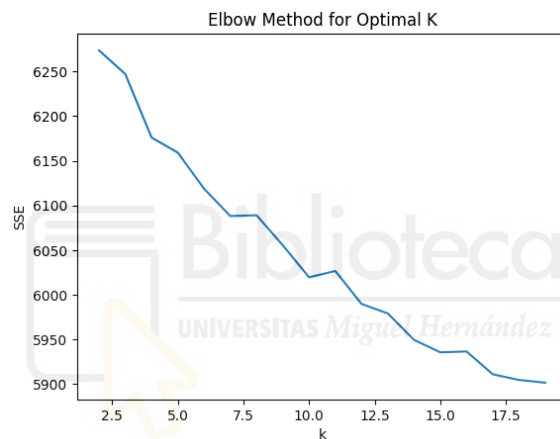


Figura 3. Gráfico del método *Elbow* para los valores de K aplicados al *dataset* preliminar.

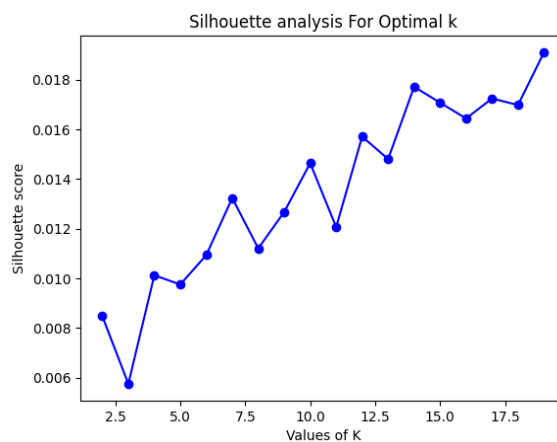


Figura 4. Gráfico del análisis de silueta para los valores de K aplicados al *dataset* preliminar.

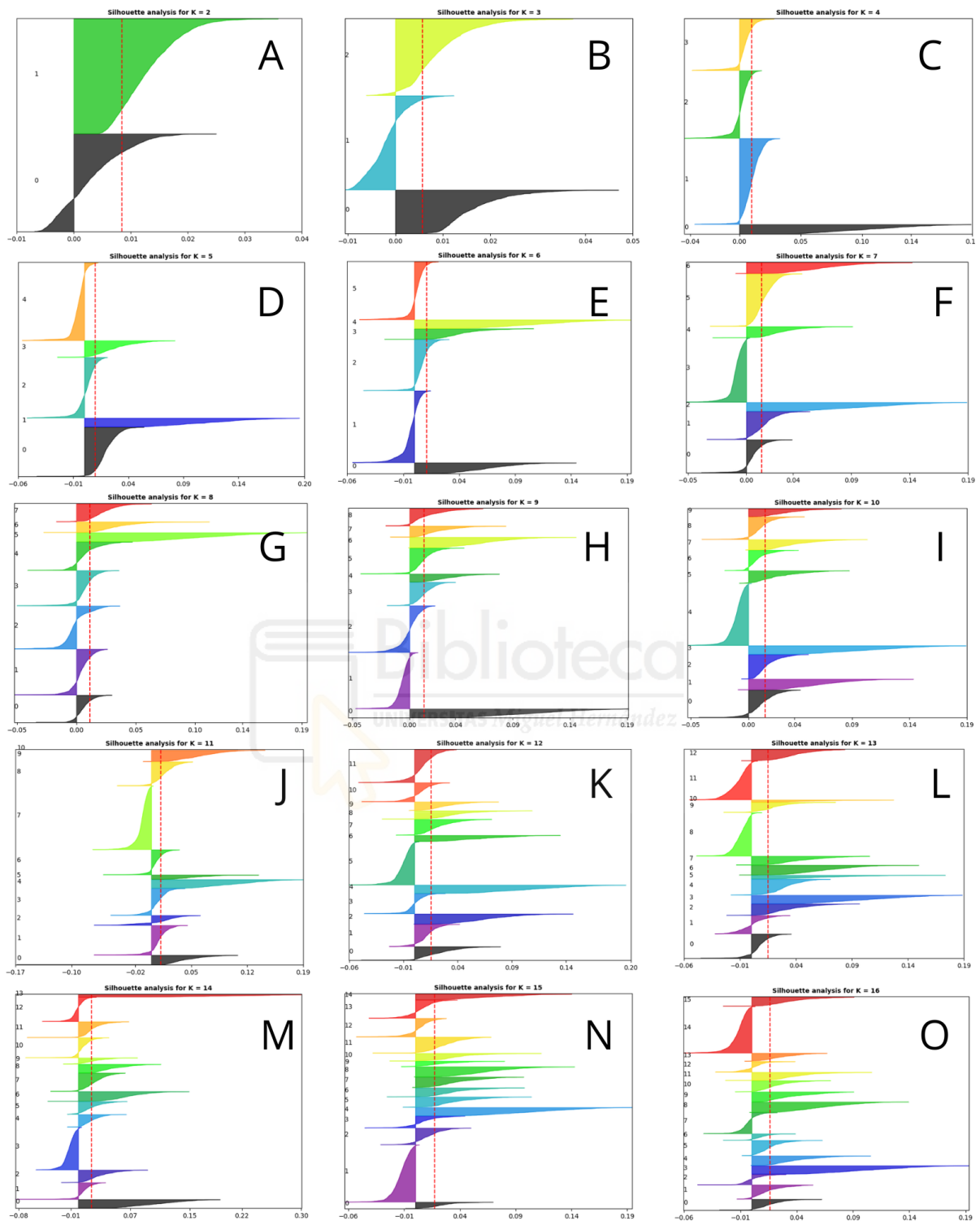


Figura 5. Gráficos de silueta para los diferentes valores de K (2-16) aplicados al *dataset* preliminar. Las subfiguras (A–O) muestran los resultados de agrupar los artículos en K grupos, indicado en la parte superior de cada gráfico.

La distribución proporcional de los artículos entre los distintos grupos identificados quedó plasmada en una representación gráfica que permitió visualizar el peso relativo de cada grupo dentro del conjunto documental analizado (Figura 6). Según dicha imagen, el grupo 1 concentró el mayor porcentaje de artículos (40 %), seguido por el grupo 2 con un 32 % y el grupo 3 con un 24 %, mientras que el grupo 0 representó únicamente un 4 % del total. Esta distribución sugiere que las líneas de investigación correspondientes a los grupos 1 y 2 predominan en el panorama actual, posiblemente por tratarse de áreas consolidadas o de mayor aplicabilidad práctica.

Abstract distribution by cluster

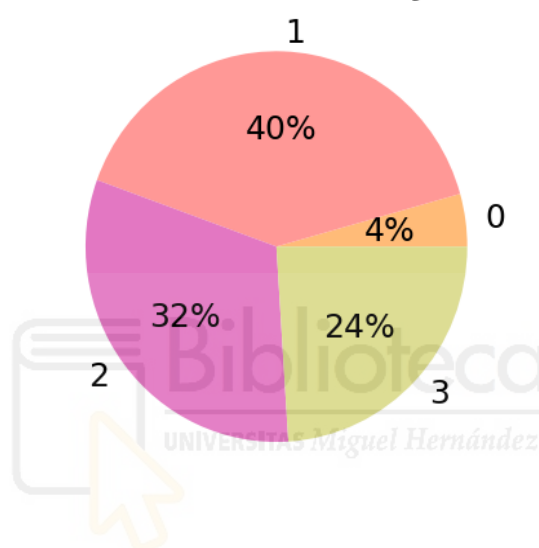


Figura 6. Diagrama de sectores con la distribución porcentual de los artículos en cada uno de los grupos.

Para visualizar la complejidad del conjunto de datos y la estructura de los grupos generados, se aplicó un análisis de componentes principales (PCA) que permitió representar en dos dimensiones la relación entre los distintos grupos (Figura 7). La proyección obtenida evidenció una segmentación clara entre ellos, aunque con diferencias notables en términos de dispersión y proximidad.

El grupo 0 se posicionó en el centro del gráfico, rodeado por el resto y con una estructura altamente compacta, lo que indicaba una fuerte cohesión interna. Su ubicación sugería una posible transversalidad temática respecto a los otros tres grupos, al mantener contacto directo con todos ellos. Esta localización estratégica lo situó como un nodo clave en el análisis, actuando como punto de conexión entre temáticas distintas pero interrelacionadas.

Por su parte, el grupo 1, localizado a la izquierda del grupo 0, presentó una dispersión moderada y una forma elongada. Su cercanía al grupo 0 y su contacto con el grupo 2 evidenciaron la existencia de similitudes metodológicas o conceptuales, si bien no mantenía conexión directa con el grupo 3.

El grupo 2, situado por debajo del grupo 0, mostró una dispersión algo mayor y se vinculó principalmente con este último, aunque también se aproximó al grupo 1 y estableció un leve contacto con el grupo 3. Esta disposición sugiere que los artículos integrados en este grupo abordan una mayor diversidad de aproximaciones técnicas y prácticas.

Finalmente, el grupo 3 apareció en la parte superior del gráfico, algo distanciado del núcleo, y destacó por su mayor dispersión en comparación con el resto. A pesar de que su conexión con otros grupos fue limitada (estableciendo un vínculo claro únicamente con el grupo 0), su amplia distribución refleja una notable heterogeneidad de enfoques o aplicaciones dentro del mismo marco temático.

En conjunto, la disposición observada en el PCA aportó evidencias tanto sobre la validez del agrupamiento como sobre las relaciones latentes entre grupos, permitiendo identificar con claridad aquellos que presentaban una mayor especificidad temática y aquellos que actuaban como vínculos metodológicos en el *corpus* analizado.

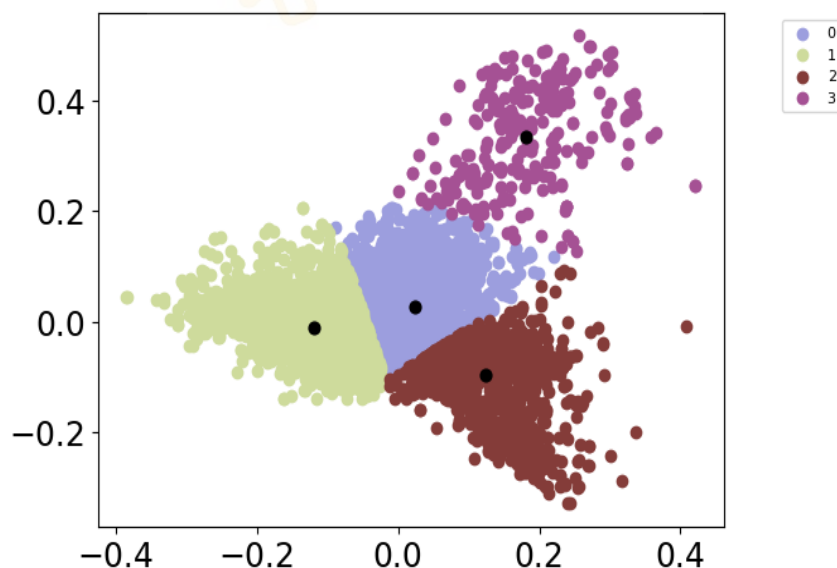


Figura 7. Gráfico de dispersión resultante del análisis PCA aplicado a los 6.612 artículos recopilados. Los colores indican la asignación de cada artículo a uno de los cuatro grupos, y los puntos negros representan los centroides.

El análisis de la distribución temporal de los artículos por grupo (Figura 8) ofreció una perspectiva complementaria que permitió contextualizar la evolución de cada línea temática a lo largo del tiempo. Durante los primeros años del periodo analizado (2000–2009), el panorama estuvo claramente dominado por el grupo 2.

A partir de 2010, comenzó a observarse una mayor diversificación temática. El grupo 1 experimentó un crecimiento progresivo y se consolidó como la línea predominante desde 2015 hasta la actualidad. De forma paralela, el grupo 0 hizo su aparición en torno a 2010, manteniendo desde entonces una presencia estable, aunque limitada.

Por otro lado, el grupo 3, cuyo interés también empezó a manifestarse a partir de 2010, registró una expansión significativa a partir de 2015. Su crecimiento sostenido hasta 2025 lo posiciona actualmente como una línea emergente en proceso de consolidación.

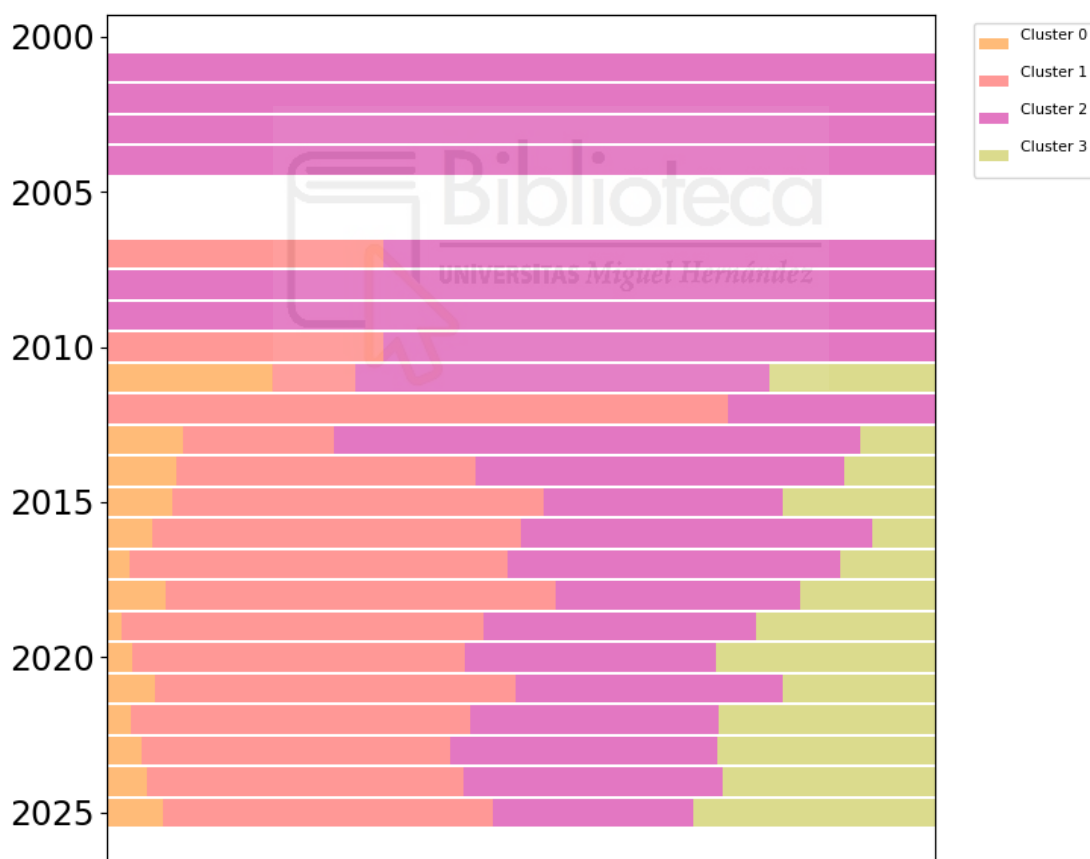


Figura 8. Gráfico de evolución del tamaño de los distintos grupos de artículos sobre la temática de la revisión entre 2001 y 2025. Se incluyen también los primeros meses de 2025.

La representación gráfica de los términos más relevantes en cada uno de los cuatro grupos identificados permitió realizar una primera aproximación a sus respectivas temáticas dominantes (Figura 9).

En términos generales, el grupo 0 se caracterizó por aglutinar estudios centrados en técnicas de aplicación localizada de herbicidas y en la detección de malas hierbas. Esta interpretación se sustentó en la alta frecuencia de términos como *“weed”*, *“herbicid”*, *“spray”*, *“detect”* y *“precis”*, lo que sugirió un enfoque claramente orientado al uso de vehículos aéreos no tripulados para la gestión de flora adventicia mediante herramientas de visión artificial y estrategias de pulverización de precisión.

En una línea diferente, el grupo 1 reunió investigaciones enfocadas principalmente en la estimación del rendimiento, el análisis del estado vegetativo de los cultivos y la modelización predictiva de variables agronómicas. La presencia destacada de palabras clave como *“yield”*, *“veget”*, *“lai”*, *“canopi”*, *“crop”*, *“predict”*, *“model”* o *“indic”* apuntó a un uso intensivo de sensores multispectrales con el objetivo de generar índices de vegetación y realizar pronósticos productivos. Estos estudios se inscribieron en un enfoque analítico orientado no solo a observar el cultivo, sino también a anticipar su comportamiento mediante modelos matemáticos y datos espectrales.

Por otra parte, el grupo 2 agrupó trabajos centrados en el desarrollo y la aplicación de tecnologías UAV en el ámbito agrícola, con especial énfasis en la optimización de sistemas de pulverización, la operativa en campo y la mejora de la precisión tecnológica. La recurrencia de términos como *“spray”*, *“farm”*, *“droplet”*, *“applic”*, *“pesticid”* o *“develop”* evidenció un interés por el diseño de nuevas plataformas UAV, así como por la evaluación de productos fitosanitarios y la innovación en procesos agronómicos.

Finalmente, el grupo 3 integró estudios de carácter más técnico y metodológico, enfocados en el uso de visión por computadora y aprendizaje profundo para el análisis automatizado de imágenes agrícolas. La aparición frecuente de términos como *“deep”*, *“segment”*, *“model”*, *“learn”*, *“detect”*, *“imag”*, *“tree”* o *“diseas”* puso de relieve que este conjunto de artículos se orientaba al desarrollo de algoritmos capaces de clasificar, contabilizar y detectar tanto enfermedades como estructuras vegetales. Este tipo de investigaciones representa una línea emergente con elevado potencial para transformar la monitorización de cultivos mediante datos visuales adquiridos por UAVs.

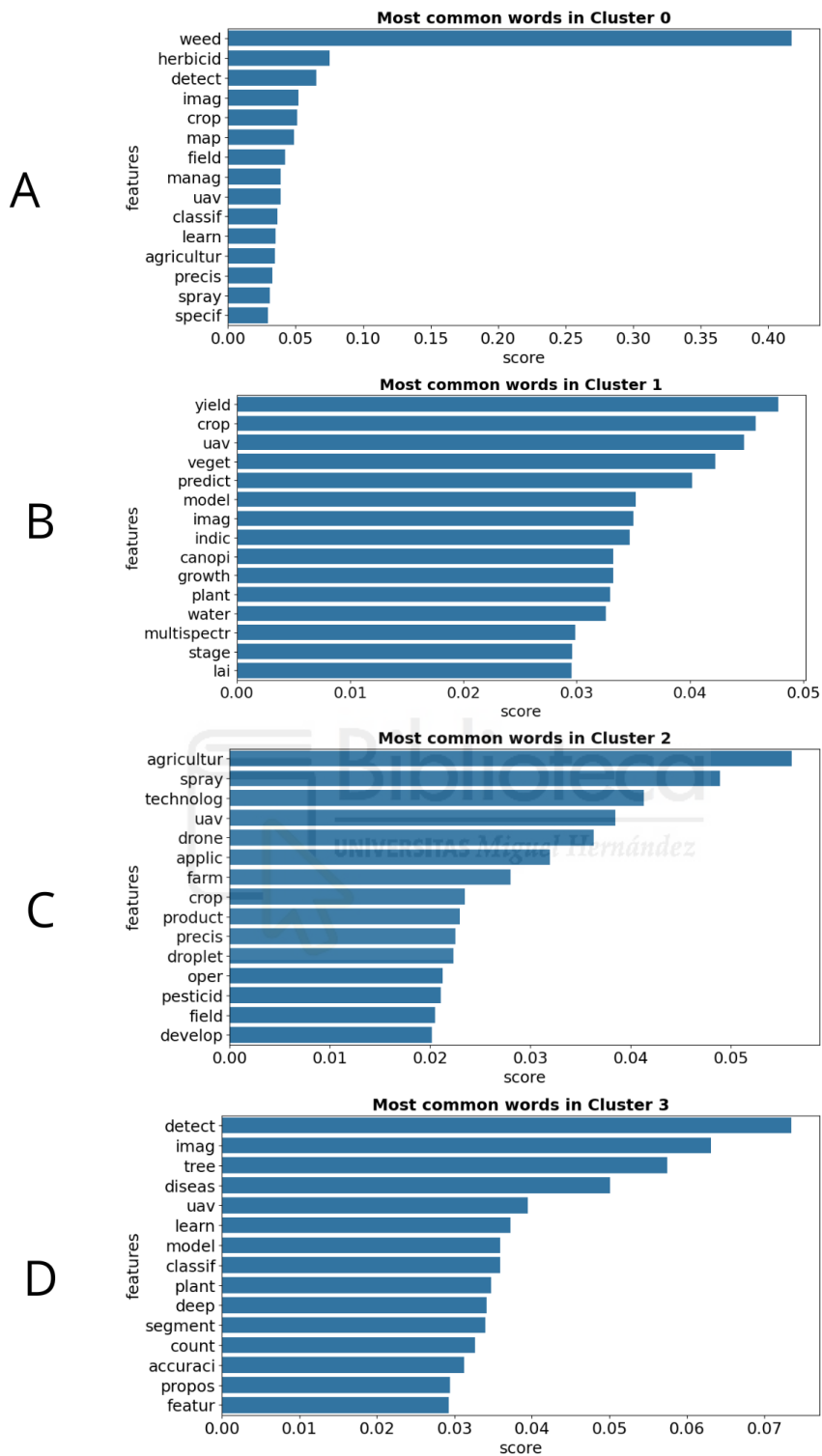


Figura 9. Representación de las quince palabras con mayor frecuencia relativa de aparición en cada uno de los cuatro grupos de artículos relacionados con la temática de la revisión. Cada uno se identifica en la parte superior de las subfiguras (A–D).

La interpretación conjunta de la distribución cuantitativa de artículos, el análisis PCA, la evolución temporal y la caracterización lingüística de los grupos permitió construir una visión integradora del panorama temático.

Por un lado, el grupo 1, que concentró el 40 % del *corpus*, representó una línea consolidada en la literatura científica, vinculada principalmente a la estimación del rendimiento de cultivos mediante índices de vegetación y modelos predictivos. Por otro lado, le siguió el grupo 2, con un 32 %, asociado al desarrollo de tecnologías UAV aplicadas a la agricultura, especialmente en tareas de pulverización y gestión operativa. Ambos grupos mostraron una dispersión moderada en el análisis PCA, lo que reforzó su carácter estructurado y especialización temática.

En cuanto al grupo 0, pese a representar solo un 4 % del total de artículos, adquirió un papel destacado por su cohesión interna, ubicación central en el espacio PCA y presencia estable desde 2010. Esta posición estratégica le permitió conectar con los otros tres grupos, reflejando su función transversal en metodologías vinculadas a la detección y gestión de malas hierbas mediante visión artificial.

Por último, el grupo 3, con un 24 % de los artículos, se configuró como el más prometedor desde una perspectiva investigadora. Aunque su volumen era menor al de los grupos 1 y 2, su elevada dispersión en el PCA apuntó a una notable diversidad metodológica y riqueza temática. Esta hipótesis se respaldó con su evolución temporal: desde 2015 mostró un crecimiento sostenido que se mantiene hasta la actualidad, posicionándolo como una línea emergente en consolidación. Este grupo se centró en estudios de aprendizaje profundo y análisis avanzado de imágenes (como segmentación, clasificación y detección), consolidándose en un ámbito en expansión dentro de la agricultura de precisión. Su conexión directa únicamente con el grupo 0 subrayó su carácter especializado, aunque compatible con enfoques transversales.

Además de estos factores, se justificó su elección como foco principal de la revisión por ser el único grupo temático que abordó explícitamente el uso de *deep learning*, en plena consonancia con la línea central de este trabajo. Esta convergencia temática y metodológica garantizó una alineación sólida entre los objetivos de la revisión y los estudios seleccionados. En consecuencia, el grupo 3 fue elegido como núcleo principal para el análisis sistemático posterior, al reunir especialización, potencial científico, trayectoria ascendente y clara delimitación temática dentro del *corpus* analizado.

4.1.2. Análisis cuantitativo específico del grupo sobre deep learning.

Dado el interés específico en el campo del aprendizaje profundo aplicado a la agricultura de precisión, se seleccionaron exclusivamente los artículos agrupados en el grupo 3, previamente identificado como el más representativo de esta temática. El cuaderno *Jupyter* empleado en el procesamiento permitió aislar dicho grupo y aplicar nuevamente el algoritmo de *clustering* sobre sus documentos, con el objetivo de lograr una segmentación temática aún más precisa y detallada.

A partir de este procedimiento, fue posible identificar subgrupos más específicos, lo que facilitó una mejor comprensión de los enfoques y metodologías predominantes en el ámbito del *deep learning*. En total, el grupo 3 quedó conformado por 1.587 artículos, que fueron sometidos al mismo tratamiento de procesamiento lingüístico y agrupación temática que el *corpus* general.

La visualización de los términos con mayor frecuencia relativa en este subconjunto temático (Figura 10) evidenció el enfoque técnico y metodológico dominante en el área. Sobresalieron términos como "*imag*", "*detect*", "*model*", "*learn*" y "*accuraci*", que destacaron la relevancia de la visión por computadora, los modelos predictivos y la evaluación del rendimiento algorítmico. Asimismo, la elevada presencia de términos como "*tree*", "*plant*", "*crop*" o "*diseas*" reforzó el estrecho vínculo con el contexto agronómico de aplicación.

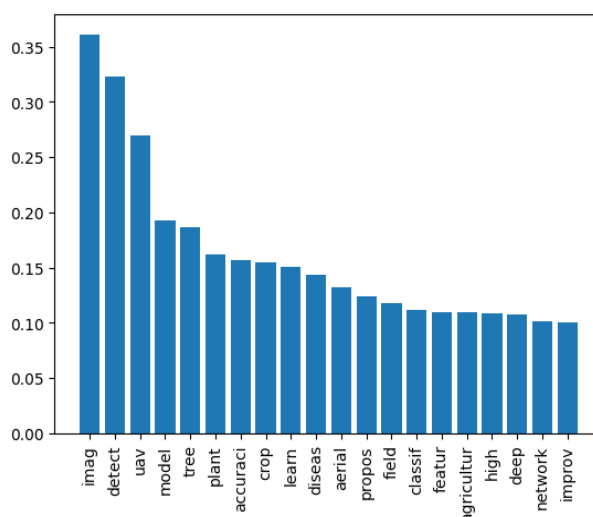


Figura 10. Gráfico de barras con las frecuencias relativas de las 20 palabras más abundantes en la bibliografía dedicada al aprendizaje profundo.

Abstract distribution by cluster

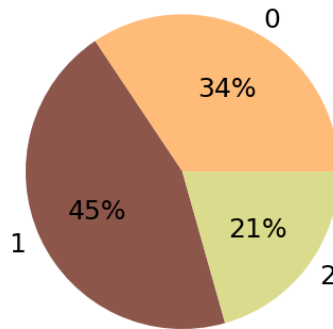


Figura 12. Diagrama de sectores con la distribución porcentual de los artículos en cada uno de los sub-grupos.

El análisis de componentes principales (PCA) permitió visualizar la distribución espacial de los subgrupos generados, facilitando la identificación de patrones temáticos y niveles de diversidad interna (Figura 13). El subgrupo 0, representado en azul, se localizó verticalmente en la zona izquierda del gráfico; el subgrupo 1, en amarillo, se extendió hacia el eje inferior central; y el subgrupo 2, en fucsia, se concentró en el cuadrante derecho. Todos presentaron una estructura definida en torno a sus respectivos centroides, lo que sugirió una coherencia interna sólida. No obstante, el subgrupo 2 evidenció una mayor dispersión, indicativa de enfoques más variados, en contraste con la compacidad de los subgrupos 0 y 1, asociada a líneas de investigación más cohesionadas.

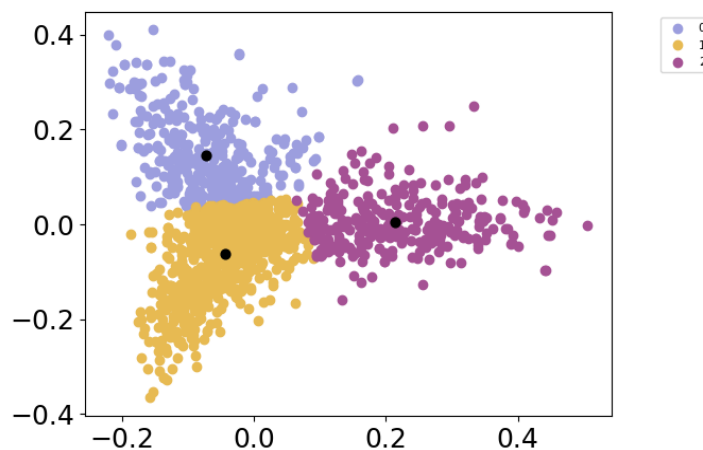


Figura 13. Gráfico de dispersión resultante del análisis PCA aplicado a los 1.587 artículos recopilados. Los colores indican la asignación de cada artículo a uno de los tres grupos, y los puntos negros representan los centroides.

La representación gráfica de los términos más relevantes en cada uno de los tres subgrupos derivados del grupo de aprendizaje profundo permitió realizar una primera aproximación a sus respectivas temáticas predominantes (Figura 14). Esta visualización resultó especialmente útil para identificar no solo las áreas de concentración investigadora, sino también las diferencias metodológicas y aplicativas entre los mismos.

Dentro del conjunto analizado, el subgrupo 0 se caracterizó por reunir estudios centrados en el reconocimiento automatizado de enfermedades y en el conteo de plantas. Esta interpretación se sustentó en la elevada frecuencia de términos como “*detect*”, “*diseas*”, “*count*” y “*plant*”, así como en la aparición destacada de palabras vinculadas al uso de modelos supervisados y redes neuronales profundas, como “*model*”, “*learn*”, “*network*” y “*deep*”. En conjunto, este vocabulario sugirió una orientación clara hacia la detección temprana de afecciones vegetales mediante técnicas avanzadas de aprendizaje profundo, con aplicaciones relevantes para la toma de decisiones agronómicas en tiempo real.

En una línea complementaria, el subgrupo 1 aglutinó investigaciones dirigidas a la segmentación y clasificación de cultivos, con un foco particular en imágenes obtenidas mediante UAVs aplicadas a especies agrícolas como el arroz y diversas plántulas. La recurrencia de términos como “*imag*”, “*crop*”, “*plant*”, “*segment*” y “*rice*”, junto con la presencia de “*classif*”, “*accuraci*” y “*seedl*”, evidenció el interés técnico por desarrollar y validar modelos de clasificación en fases tempranas del desarrollo del cultivo. Este tipo de estudios se enmarca en un enfoque orientado a optimizar la gestión agronómica mediante datos visuales recolectados en campo.

Por su parte, el subgrupo 2 integró trabajos especializados en el análisis de ecosistemas leñosos y forestales. La prominencia de términos como “*tree*”, “*forest*”, “*palm*”, “*crown*” y “*canopi*” indicó una orientación clara hacia la identificación, segmentación y caracterización de estructuras vegetales complejas, tales como copas, especies o parcelas arbóreas. Además, la aparición de palabras como “*individu*”, “*speci*” y “*orchard*” reflejó un interés por el análisis individualizado de ejemplares en entornos agroforestales, sugiriendo un enfoque más vinculado a la conservación, el inventario y la gestión precisa de masas vegetales en contextos naturales o seminaturales.

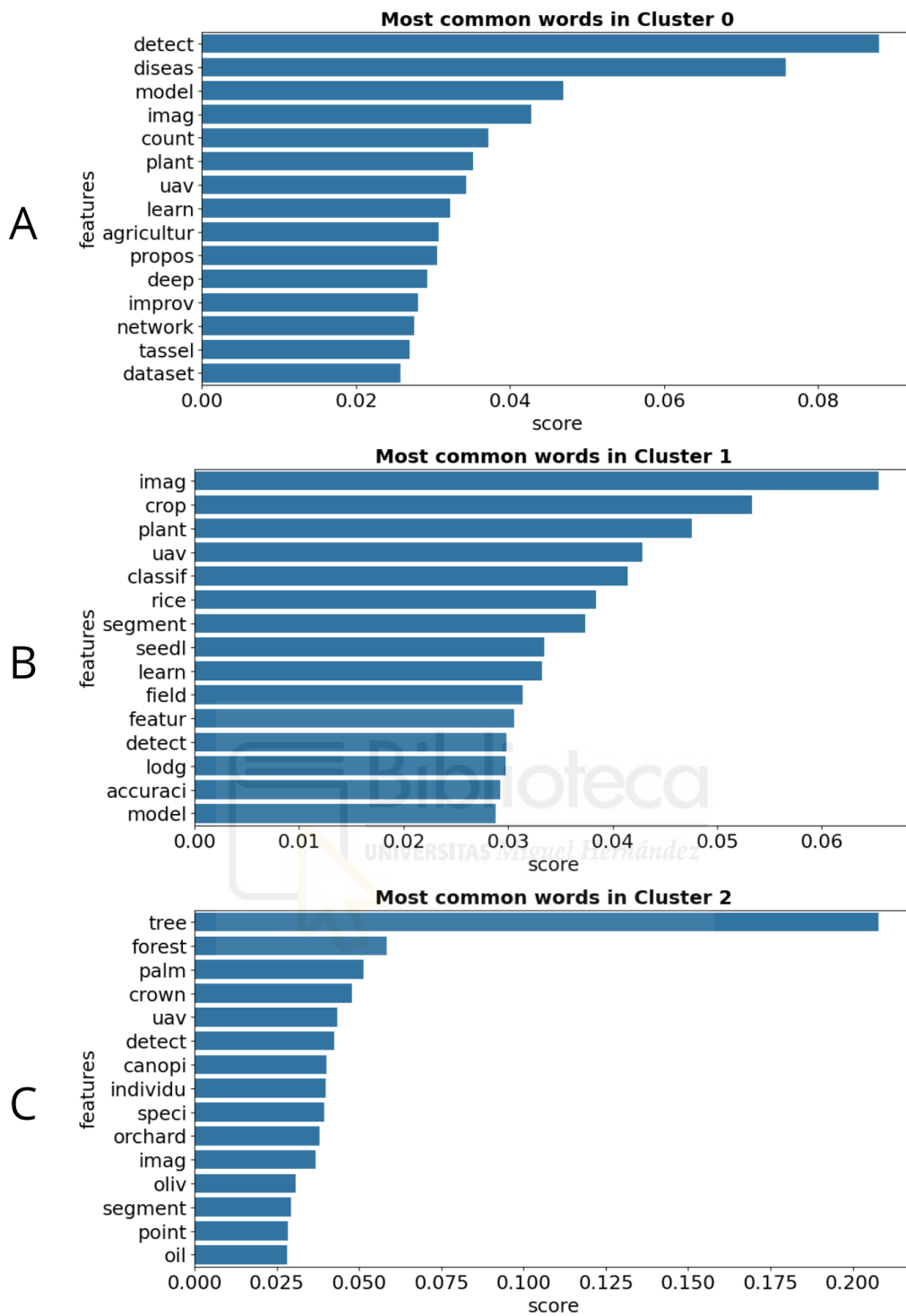


Figura 14. Quince palabras con mayor frecuencia relativa de aparición en cada uno de los tres sub-grupos de artículos relacionados con aprendizaje profundo. Cada grupo se identifica en la parte superior de las subfiguras (A–C).

La interpretación conjunta de la distribución porcentual de artículos, el análisis PCA, la representación de términos clave y la caracterización semántica de los subgrupos permitió delimitar con precisión la estructura interna del grupo temático centrado en aprendizaje profundo aplicado a la agricultura de precisión.

Dentro del análisis temático, el subgrupo 1, que concentró el mayor volumen de artículos, representó una línea de investigación consolidada en torno a la clasificación y segmentación de cultivos agrícolas mediante imágenes obtenidas por UAVs. Su compacidad en el espacio PCA y la recurrencia de términos como “*crop*”, “*segment*” o “*rice*” reflejaron una orientación técnica clara, aunque con menor especificidad hacia modelos avanzados de *deep learning*, lo que limitó su idoneidad para una revisión sistemática centrada en dicho enfoque concreto.

De manera complementaria, el subgrupo 2, menos representado cuantitativamente, mostró una elevada dispersión en el gráfico PCA y una concentración terminológica asociada a entornos forestales y leñosos, con términos como “*tree*”, “*forest*” o “*crown*”. Esta configuración temática y su naturaleza exploratoria sugirieron una línea aún incipiente, más vinculada a aplicaciones agroforestales que al objetivo específico de esta revisión.

Por último, el subgrupo 0, aunque ocupó una posición intermedia en volumen, destacó por su elevada cohesión interna, compacidad gráfica y perfil semántico bien definido. La frecuencia de términos como “*diseases*”, “*count*”, “*learn*” o “*network*” evidenció una especialización clara en la detección de enfermedades, el conteo automatizado y el uso intensivo de redes neuronales profundas. Su agrupación densa en torno al centroide en el análisis PCA reforzó esta coherencia lingüística y metodológica, posicionándolo como un bloque temático robusto y alineado con los objetivos de esta investigación.

Además de estos factores, se justificó la elección del subgrupo 0 como núcleo principal para aplicar la metodología PRISMA, al ser el único que abordó de manera explícita y especializada el uso de *deep learning* para tareas automatizadas de diagnóstico vegetal. Esta convergencia temática, junto con su densidad terminológica y delimitación precisa, garantizó la coherencia entre los objetivos de la revisión y el conjunto de artículos seleccionados. En consecuencia, el subgrupo 0 fue elegido como base para el análisis sistemático, al reunir representatividad, especialización metodológica y afinidad directa con el eje central de este trabajo.

4.2. Revisión sistemática.

4.2.1. Selección de los estudios.

El proceso de selección de estudios aplicado al *dataset* temático puede seguirse en el diagrama que resume las distintas fases de depuración y filtrado realizadas sobre el corpus inicial (Figura 15), recuperado de las bases de datos Scopus y Web of Science. Aunque la búsqueda original integraba ambas fuentes, el procesamiento automatizado unificó los resultados, por lo que el conjunto inicial se trató como una única base de datos.

De los 546 documentos extraídos del subgrupo 0, se excluyeron 443 tras aplicar un filtrado temático y temporal orientado a ajustar la muestra al límite normativo de artículos citables. En concreto, se priorizaron aquellos estudios que incluían el término *deep learning* en el título y que se habían publicado entre 2019 y 2024, con el objetivo de garantizar tanto la actualidad como la pertinencia metodológica del conjunto seleccionado para la revisión sistemática.

Posteriormente, se eliminaron manualmente 17 documentos duplicados, lo que dio lugar a un total de 86 artículos válidos para la fase de cribado inicial. En esta etapa, se descartaron 31 artículos al comprobarse que no cumplían con los criterios de elegibilidad establecidos, fundamentalmente en lo relativo a su adecuación temática y metodológica, tal y como se detalla en la sección correspondiente del apartado de metodología.

Los 55 documentos restantes fueron revisados a texto completo y considerados aptos para su inclusión en la revisión. No obstante, dado que la normativa del trabajo de fin de grado limita el número de artículos citables a un máximo de 35, se aplicó un sistema de puntuación estructurado, también descrito en la metodología, con el fin de seleccionar de forma objetiva los estudios más relevantes.

Los 24 artículos excluidos por este motivo se recogen en un apartado específico, donde se citan individualmente y se detalla la razón de su baja puntuación. Por ello, en el diagrama de flujo, estos documentos aparecen bajo la categoría “*Reports excluded: Reason1 (n = 24)*”, en alusión explícita a su exclusión por puntuación insuficiente conforme al sistema de priorización establecido.

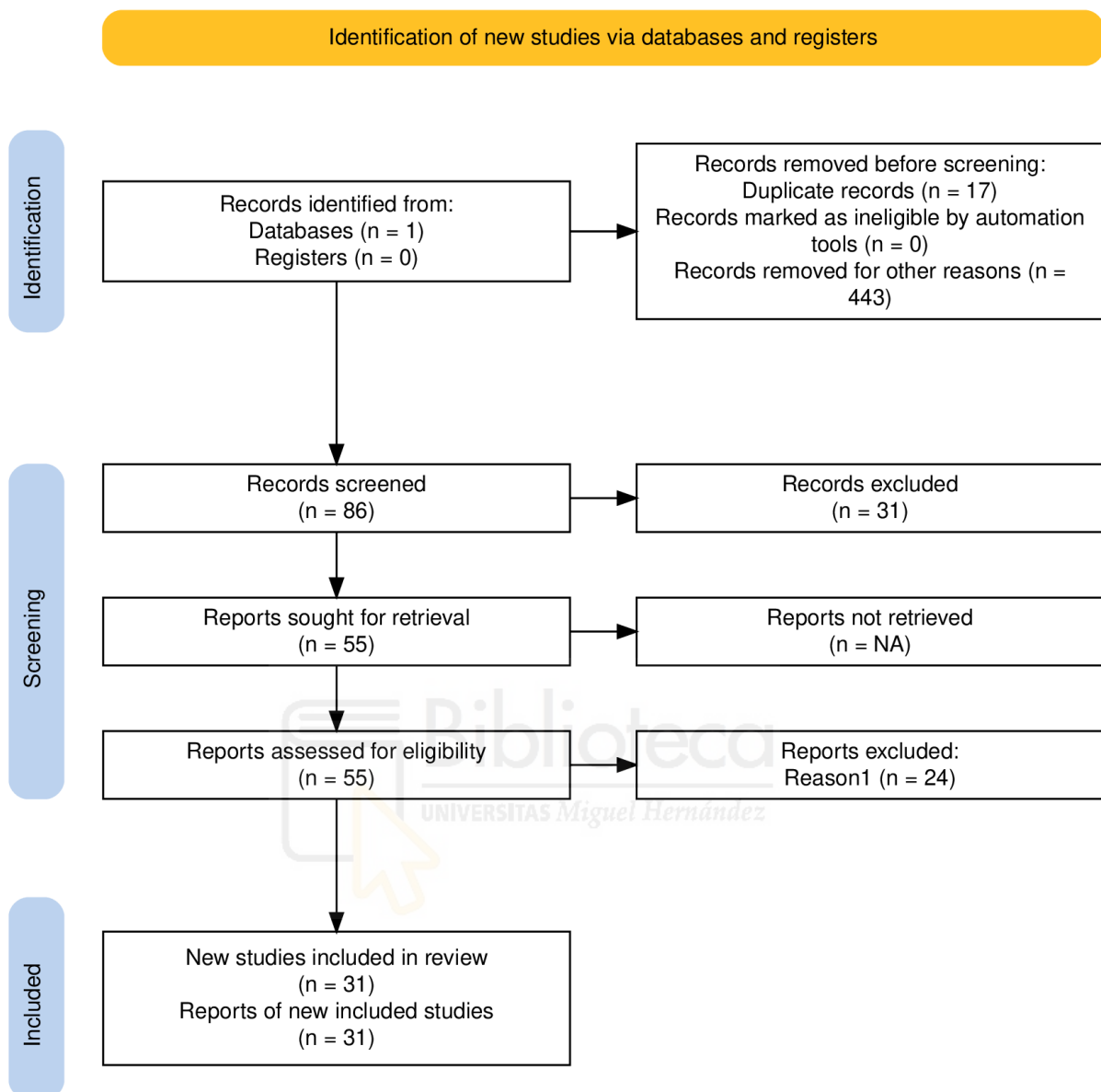


Figura 15. Diagrama de flujo del proceso de selección de estudios aplicado al *dataset* temático fusionado de Scopus y Web of Science, conforme al protocolo PRISMA 2020. Se representan las distintas fases de depuración, incluyendo el filtrado temático y temporal, la eliminación de duplicados, la aplicación de criterios de elegibilidad y la selección final de los artículos mediante un sistema de puntuación estructurado.

Tras aplicar el sistema de puntuación diseñado para priorizar los estudios, se llevó a cabo una segunda fase de cribado. De los 55 artículos inicialmente considerados aptos, se seleccionaron 31, conforme a las limitaciones establecidas por la normativa. Esta selección se basó en tres motivos de exclusión, consignados en la tabla Excel mediante la variable “Motivo”. El procedimiento respondió a la necesidad de acotar una muestra manejable y metodológicamente sólida, asegurando que los estudios incluidos representasen con precisión los objetivos del trabajo.

Bajo el criterio de baja relevancia, se descartaron los estudios con una puntuación total igual o inferior a dos puntos, al no aportar un valor significativo a los fines de la revisión. Posteriormente, se excluyeron las revisiones bibliográficas que, si bien fueron consideradas durante el análisis automatizado, no se priorizaron en la selección final debido al límite de artículos citables y a la preferencia por estudios experimentales. Por otra parte, la categoría intención agrupó estudios valorados por su carácter innovador o por su contribución metodológica, pero fueron finalmente excluidos por tratarse de trabajos más antiguos o redundantes respecto a investigaciones más actuales y robustas. Esta decisión se reflejó en las variables “Extra” y “Motivo”.

La primera tabla muestra los 57 artículos evaluados, identificados mediante su referencia en la celda correspondiente. En la columna “Total”, los estudios seleccionados aparecen resaltados en azul, mientras que los excluidos figuran en rojo, lo que permite visualizar de forma clara los resultados del cribado. Esta codificación cromática facilita el seguimiento de las decisiones adoptadas durante la depuración y mejora la trazabilidad del proceso.

En el bloque temática de detección de plagas y enfermedades, se seleccionó el estudio desarrollado por Mora et al. (2024), que destacó por su precisión y solidez metodológica. El trabajo empleó imágenes RGB y multiespectrales, integró modelos como YOLOv8 y Faster R-CNN, y consideró entornos reales con técnicas de fusión de datos. Por su parte, el artículo de Linero-Ramos et al. (2024), fue descartado por su enfoque limitado a clasificación de imágenes y su menor robustez. Esta decisión refleja la preferencia por estudios con aplicación en contextos operativos reales y con validaciones con rigor. Asimismo, el estudio de Narimani et al. (2024) fue excluido al no estar disponible en las bases de datos accesibles desde la institución académica, lo que impidió su evaluación completa, a pesar de su relevancia temática.

En la categoría de estimación de rendimiento, se priorizaron estudios con validación robusta en campo. En el caso del trigo, se seleccionó la investigación de Zhu et al. (2022), por su precisión en la detección de espigas en condiciones complejas. En maíz, el artículo de Chen et al. (2024), destacó por combinar detección precisa y estimación de rendimiento. Para el sorgo, se eligió el trabajo de Li et al. (2022a), que comparó modelos como YOLOv4, SSD y EfficientDet, obteniendo resultados consistentes. Estos estudios evidenciaron el potencial de los UAVs en la predicción precisa de variables agronómicas clave.

En el grupo de otras aplicaciones, se diferenciaron dos líneas. Por una parte, se seleccionaron dos estudios clave: Jia et al. (2024), centrado en la detección de plántulas en estadios tempranos utilizando UAVs y modelos YOLO, considerando variables como la densidad de siembra y el estadio fenológico; y Du et al. (2024), que amplió la cobertura fenológica con detección de espigas en campos de hibridación, aplicando mejoras metodológicas en entornos complejos. Ambos trabajos destacaron por su utilidad práctica en la toma de decisiones durante el ciclo fenológico del cultivo.

Por otra parte, se seleccionó el estudio de Gautam et al. (2025), valorado por su elevada precisión, aplicabilidad práctica y potencial de escalabilidad. Este se complementó con el trabajo de Yadav et al. (2023), centrado en la detección de plantas reservorio (algodón voluntario), lo que permitió avanzar hacia una gestión integrada de especies no deseadas. En conjunto, ambos estudios representaron enfoques relevantes para mejorar el control de vegetación no deseada mediante técnicas de teledetección avanzada.

En definitiva, los estudios seleccionados respondieron a criterios de relevancia temática, consistencia metodológica y aplicabilidad en contextos agronómicos reales. Se otorgó prioridad a aquellas investigaciones que, mediante el uso de UAVs e inteligencia artificial, ofrecían resultados contrastables y con potencial de transferencia al campo. Esta estrategia de selección garantizó una revisión exhaustiva y representativa de las principales líneas de avance en el ámbito de la agricultura de precisión.

Referencia	Año	N° Citas	Tipo	Nota	Citaciones	Diversidad	Año	Extra	Total	Exclusión	Motivo
Detección de Plagas y Enfermedades											
(Kerkech et al., 2020)	2020	225	Experimental	Viñedo	3	1	0	-	4	No	
(Ahmad et al., 2023)	2023	110	Revisión	Revisión	-	-	-	-	0	Sí	Revisión
(Shahi et al., 2023)	2023	93	Revisión	Revisión	-	-	-	-	0	Sí	Revisión
(Bouguettaya et al., 2023)	2023	66	Revisión	Revisión	-	-	-	-	0	Sí	Revisión
(Roosjen et al., 2020)	2020	63	Experimental	Plaga	3	2	0	-	5	No	
(Kuswidiyanto et al., 2022)	2022	39	Revisión	Revisión	-	-	-	-	0	Sí	Revisión
(Amarasingam et al., 2022)	2022	34	Experimental	Caña Azúcar	2	2	1	-	5	No	
(Ksibi et al., 2022)	2022	30	Experimental	Olivar	1	2	1	-	4	No	
(Zhao & Shi, 2022)	2022	10	Experimental	Col China	1	0	1	-	2	Sí	Baja Relevancia
(Zhu et al., 2024)	2024	8	Revisión	Revisión	-	-	-	-	0	Sí	Revisión
(Abdalla et al., 2024)	2024	8	Experimental	Desafíos	0	2	1	-	3	No	
(Lv et al., 2024)	2024	6	Experimental	Maíz	0	2	1	-	3	No	
(Kontogiannis et al., 2024)	2024	6	Experimental	Viñedo	0	1	1	-	2	Sí	Baja Relevancia
(Mora et al., 2024)	2024	4	Experimental	Banana	0	1	1	2	4	No	
(Wieme et al., 2024)	2024	3	Experimental	Patata	0	2	1	-	3	No	
(Zhang et al., 2024)	2024	2	Experimental	A.China	0	2	1	-	3	No	
(Linero-Ramos et al., 2024)	2024	2	Experimental	Banana	0	1	1	-	2	Sí	Intención
(Narimani et al., 2024)	2024	2	Experimental	Tomate	0	2	1	-2	1	Sí	Intención
Fuente no encontrada	2020	0	Experimental	Arroz	0	0	1	-	1	Sí	Baja Relevancia
(Chen et al., 2024)	2024	0	Experimental	A.China	0	0	1	-	1	Sí	Baja Relevancia
(Vijayalakshmi & Radhika, 2024)	2024	0	Experimental	Futuro	0	2	1	-	3	No	
Estimación de Rendimiento											
(Xiong et al., 2020)	2020	67	Experimental	Mango	3	2	0	-	5	No	
(Zhao et al., 2022)	2022	46	Experimental	Trigo	2	1	1	-2	2	Sí	Intención
(Sanacifar et al., 2023)	2023	39	Revisión	-	-	-	-	-	0	Sí	Revisión
(Kalantar et al., 2020)	2020	38	Experimental	Melón	2	2	0	-	4	No	
(Zhu et al., 2022)	2022	29	Experimental	Trigo	1	1	1	-	3	No	
(Alzadjali et al., 2021)	2021	28	Experimental	Maíz	1	1	1	-2	1	Sí	Intención
(Li et al., 2022a)	2022	16	Experimental	Sorgo	1	1	1	-	3	No	
(Gonzalo-Martín et al., 2021)	2021	14	Experimental	Sorgo	1	1	1	-2	1	Sí	Intención
(Liu et al., 2022)	2022	13	Experimental	Maíz	1	1	1	-2	1	Sí	Intención
(Li et al., 2023b)	2023	10	Experimental	Colza	1	2	1	-	4	No	
(Lin et al., 2022)	2022	10	Experimental	Litchi	1	2	1	-	4	No	
(Sekharamanthy et al., 2024)	2024	9	Experimental	Manzana	0	2	1	-	3	No	
(Wang et al., 2024)	2024	8	Experimental	Té	0	2	1	-	3	No	
(Melnychenko et al., 2024)	2024	8	Experimental	Innovación	0	2	1	-	3	No	
(Li et al., 2023a)	2023	6	Experimental	Longan	0	2	1	-	3	No	
(Hong-Yu et al., 2023)	2023	5	Experimental	Ortiga blanca	0	2	1	-	3	No	
(Wiggers et al., 2022)	2022	4	Experimental	Plantas	0	0	1	-	1	Sí	Baja Relevancia
(Chen et al., 2024)	2024	4	Experimental	Maíz	0	1	1	2	4	No	
(Niu et al., 2024)	2024	3	Experimental	Maíz	0	1	1	-	2	Sí	Intención
(Subeesh et al., 2024)	2024	3	Experimental	Cítricos	0	2	1	-	3	No	
Otras aplicaciones											
(Wang et al., 2020)	2020	113	Experimental	Choques	3	2	0	-	5	No	
(Menshchikov et al., 2021)	2021	38	Experimental	Invasoras	2	1	1	-2	2	Sí	Intención
(Yadav et al., 2023)	2023	26	Experimental	Reservorio	1	2	1	-	4	No	
(Li et al., 2022b)	2022	23	Experimental	Cosecha	1	2	1	-	4	No	
(Sassu et al., 2023)	2023	10	Experimental	Fitosanitarios	1	2	1	-	4	No	
(Cui et al., 2024)	2024	3	Experimental	Polinización	0	2	1	-	3	No	
(Gautam et al., 2025)	2025	2	Experimental	Invasoras	0	2	1	-	3	No	
(Jia et al., 2024)	2024	2	Experimental	Monitoreo	0	2	1	-	3	No	
(Valicharla et al., 2024)	2024	2	Experimental	Invasoras	0	1	1	-	2	Sí	Intención
(Du et al., 2024)	2024	2	Experimental	Despigado	0	2	1	-	3	No	
(Nazeer et al., 2024)	2024	1	Experimental	Flores	0	0	1	-	1	Sí	Intención
(Du et al., 2024)	2024	0	Experimental	Invasoras	0	1	1	-	2	Sí	Intención
(Slimani et al., 2024)	2024	0	Experimental	Haba	0	0	1	-	1	Sí	Intención

Tabla 1. Visualización del segundo cribado de los 55 artículos considerados aptos tras el primer filtro. Se destacan en azul aquellos seleccionados (n = 31) y en rojo los excluidos (n = 24), según la puntuación total y las variables “Motivo” y “Extra”. La columna “Referencia” permite su identificación individual. Las referencias a los artículos incluidos se encuentran en la bibliografía y las de los no incluidos se puede acceder a partir del anexo creado para esta finalidad.

4.2.2. Características de los estudios incluidos.

De acuerdo con lo establecido, se aplicó el protocolo PRISMA 2020 junto con el sistema de puntuación definido, seleccionándose 35 estudios que representaron rigurosamente el uso de UAVs y técnicas de *deep learning* en agricultura de precisión. Todos cumplieron con los criterios temáticos, metodológicos y de actualidad definidos para esta revisión.

Los cultivos analizados abarcaron especies representativas como maíz, trigo, banano, algodón, caña de azúcar, café, vid, cítricos, té y lichi, lo que evidencia una adopción creciente de estas tecnologías en sistemas agrícolas diversos y geográficamente variados.

Con el fin de facilitar el análisis comparativo, los estudios fueron organizados en tres bloques temáticos principales:

- Detección de plagas y enfermedades: investigaciones centradas en la identificación de agentes nocivos, priorizando modelos con alta precisión y aplicabilidad operativa.
- Estimación de rendimiento: estudios orientados a predecir variables productivas como la floración, el conteo de espigas o la estimación del peso de frutos, con métricas de validación en campo suficientemente sólidas.
- Otras aplicaciones: trabajos enfocados en la clasificación fenológica, la detección de flora adventicia, la automatización de tareas agronómicas o la navegación autónoma, destacando por su utilidad práctica y potencial de escalabilidad.

En el conjunto analizado predominaron modelos basados en redes convolucionales (CNN), arquitecturas YOLO (v3–v8), *Transformers*, mecanismos de atención y soluciones híbridas con IoT. Las imágenes procedieron principalmente de sensores RGB, aunque también se emplearon datos multiespectrales, hiperespectrales y cámaras de profundidad. Algunas propuestas integraron segmentación semántica, análisis multivista o estimaciones 3D, ampliando el nivel de detalle agronómico.

A continuación, se presenta la Tabla 2, en la que se recogen de forma estructurada los atributos principales de cada artículo (cultivo, aplicación, modelo y tipo de sensor), lo que permite identificar patrones metodológicos y tendencias tecnológicas en el uso combinado de UAVs y *deep learning* en la agricultura de precisión.

Referencia	Características
<u>Detección Plagas y Enfermedades</u>	
(Kerkech et al., 2020)	Aplicado a vid; detección de enfermedad (mildiu); segmentación de imágenes
(Roosjen et al., 2020)	Aplicado a frutales pequeños (<i>Drosophila suzukii</i>); detección de plaga
(Amarasingam et al., 2022)	Aplicado a caña de azúcar; detección de enfermedad de la hoja blanca; RGB
(Ksibi et al., 2022)	Aplicado a olivar; detección de enfermedades; métodos híbridos
(Abdalla et al., 2024)	Aplicado a algodón; detección de <i>Fusarium oxysporum</i> ; modelo híbrido
(Lv et al., 2024)	Aplicado a maíz; detección de roya; modelo con mecanismo de atención
(Mora et al., 2024)	Aplicado a plátano; detección de marchitez por <i>Xanthomonas</i> ;
(Wieme et al., 2024)	Aplicado a patata; detección de <i>Alternaria solani</i> ; modelo supervisado
(Zhang et al., 2024)	Aplicado a <i>zanthoxylum</i> ; detección de roya; detección objetos, imágenes H.S.
(Vijayalakshmi & Radhika, 2024)	Aplicado a cultivo no especificado; detección de plagas; IoT
<u>Estimación de Rendimiento</u>	
(Xiong et al., 2020)	Aplicado a mango; detección de frutos/estructuras
(Kalantar et al., 2020)	Aplicado a melón; estimación de peso
(Zhu et al., 2022)	Aplicado a trigo; modelo basado en transformer
(Li et al., 2022a)	Aplicado a sorgo; conteo
(Li et al., 2023b)	Aplicado a colza; conteo, imágenes RGB
(Lin et al., 2022)	Aplicado a lichi; estimación de floración
(Sekharamanthy et al., 2024)	Aplicado a manzano; detección de frutos, mecanismo de atención
(Wang et al., 2024)	Aplicado a té; estimación de rendimiento
(Melnychenko et al., 2024)	Aplicado a manzano; detección de frutos, uso de múltiples UAVs
(Li et al., 2023a)	Aplicado a longan; estimación de rendimiento
(Hong-Yu et al., 2023)	Aplicado a ramio; conteo
(Chen et al., 2024)	Aplicado a maíz; detección de espigas, imágenes RGB
(Subeesh et al., 2024)	Aplicado a cítricos; estimación de rendimiento
<u>Otras Aplicaciones</u>	
(Wang et al., 2020)	Aplicado a navegación autónoma, cámara de profundidad
(Yadav et al., 2023)	Aplicado a maíz; detección de plantas voluntarias
(Li et al., 2022b)	Aplicado a longan; localización de puntos de cosecha, localización 3D
(Sassu et al., 2023)	Aplicado a alcachofa; aplicación localizada de agroquímicos
(Cui et al., 2024)	Aplicado a cultivo no especificado; detección inteligente
(Gautam et al., 2025)	Aplicado a maleza invasora (Siam Weed); detección de especies invasoras
(Jia et al., 2024)	Aplicado a maíz; detección de plántulas
(Du et al., 2024)	Aplicado a maíz; identificación de estados de floración

Tabla 2. Clasificación temática y resumen de características de los 35 estudios incluidos en la revisión. Se especifican para cada artículo el cultivo, el objetivo agronómico y los principales aspectos metodológicos.

4.2.3. Evaluación del riesgo de sesgo.

No se aplicaron herramientas formales de sesgo por tratarse de una revisión no clínica, pero se adoptaron medidas metodológicas adecuadas (puntuación estructurada, filtrado automatizado y revisión objetiva). Aunque puede existir sesgo a favor de tecnologías emergentes, el proceso fue riguroso y acorde al enfoque del estudio.

4.2.4. Resultados de los estudios individuales.

Cada estudio fue evaluado por su enfoque técnico, objetivo agronómico y resultados. Se destacaron aquellos con alta aplicabilidad, rigor metodológico o propuestas innovadoras, seleccionándose los más representativos para análisis detallado posterior.

4.2.5. Resultados de la síntesis.

La mayoría de los estudios se basaron en datos reales de campo, lo que favorece su aplicabilidad práctica; al agruparlos por temáticas, destacaron precisiones superiores al 90 % en detección de plagas, correlaciones $R^2 > 0,95$ en estimación de rendimiento y propuestas eficientes para automatización agrícola, siendo la variabilidad observada reflejo de la complejidad agronómica y la coherencia entre métricas, objetivos y condiciones validada mediante revisión cualitativa.

4.2.6. Evaluación del sesgo por publicación.

Aunque no se analizó formalmente, se reconoció una posible sobrerrepresentación de estudios con mayor impacto. Para compensarlo, se aplicó un criterio cualitativo (“Extra”) que valoró innovación, favoreciendo la inclusión de estudios recientes y relevantes.

4.2.7. Certeza de la evidencia.

No se estimó cuantitativamente la certeza. Sin embargo, la coherencia temática, validez técnica y consistencia entre estudios otorgaron solidez a las conclusiones.

4.2.8. Discusión general e implicaciones.

Los resultados confirmaron el potencial del *deep learning* con imágenes UAV en agricultura de precisión. Pese a limitaciones como baja validación externa o escasez de datos abiertos, el enfoque metodológico fue riguroso. Se recomendó fomentar modelos replicables y *datasets* accesibles para mejorar la reproducibilidad y facilitar la transferencia tecnológica al sector agrícola.

4.3. Revisión temática.

4.3.1. Detección plagas y enfermedades.

4.3.1.1. Modelos convolucionales clásicos.

En esta categoría se agruparon estudios que emplearon arquitecturas convolucionales convencionales en tareas de segmentación y clasificación. Aunque presentan una estructura relativamente sencilla, estos modelos han mostrado un desempeño robusto en contextos agrícolas, en particular cuando son entrenados con imágenes multispectrales captadas por UAV. Su inclusión en esta revisión tiene como objetivo ofrecer un punto de referencia frente a arquitecturas de mayor complejidad.

Entre las contribuciones más representativas dentro de esta agrupación, destacó la investigación de Kerkech et al. (2020), centrada en la detección temprana de mildiu en viñedos mediante imágenes RGB y NIR obtenidas por UAV, combinando un registro multispectral optimizado y un modelo de segmentación semántica basado en *deep learning*. El objetivo fue validar un enfoque automatizado, preciso y viable para aplicaciones en viticultura de precisión.

La adquisición se realizó con un dron equipado con cámaras MAPIR Survey2 (RGB y NIR, 850 nm), volando a 25 m de altura con resolución de 1 cm/píxel. Para alinear los espectros, se empleó un algoritmo de registro basado en AKAZE con refinamiento iterativo vía RANSAC, reduciendo significativamente los errores de superposición entre sensores.

La detección se llevó a cabo mediante una red SegNet con codificador VGG-16, entrenada por separado para cada espectro. Esta arquitectura simétrica tipo *encoder-decoder* permitió segmentar a nivel de píxel mediante convoluciones, *pooling* y *upsampling* guiado, manteniendo la precisión espacial. El sistema distinguió cuatro clases (sombra, suelo, sano, sintomático), y fue entrenado con más de 100.000 parches por espectro, generados mediante aumento de datos. El etiquetado combinó predicción automática con redes LeNet5 y correcciones manuales guiadas por observaciones de campo.

SegNet mostró un rendimiento sólido: 85,13 % de precisión a nivel foliar en el espectro visible y 78,72 % en el NIR. La fusión espectral mejoró notablemente los resultados, alcanzando un 90,23 % (OR) y 95,02 % a nivel de planta. Estas cifras indicaron una

detección complementaria entre canales, especialmente útil para síntomas incipientes. El modelo también evidenció buena generalización en viñedos sin enfermedad.

En conjunto, el estudio validó a SegNet como una arquitectura eficaz y razonablemente eficiente para la segmentación multiclase en agricultura de precisión, demostrando su aplicabilidad para el monitoreo fitosanitario basado en UAVs e imágenes multispectrales, con especial potencial en contextos donde se requieren decisiones agronómicas rápidas y localizadas.

Con el objetivo de aplicar modelos convolucionales personalizados en escenarios reales de campo, Wieme et al. (2024) diseñaron un sistema para la detección automática de *Alternaria solani* en campos de patata, integrando imágenes UAV de ultra alta resolución y un modelo de *deep learning* denominado AlternarAI. El objetivo principal fue generar un sistema robusto para la identificación precoz de esta enfermedad foliar en condiciones reales de campo, minimizando el uso de productos fitosanitarios mediante mapas de tratamiento variable.

Para ello, se llevaron a cabo vuelos entre 2019 y 2022 con un UAV DJI M600 PRO equipado con una cámara Sony Alpha 7III modificada para capturar luz visible y NIR, obteniendo una resolución submilimétrica de 0,3 mm/píxel. Las imágenes se capturaron desde 10 metros de altura en campos experimentales inoculados artificialmente con diferentes cultivares y concentraciones de esporas, generando un conjunto de datos de 7.956 imágenes. Estas se dividieron en más de 22.000 parches de 256×256 píxeles, que posteriormente fueron etiquetados en cuatro clases (“*Alternaria*”, “*Healthy*”, “*Dubious*” y “*Background*”) y reducidos a dos categorías (“*Alternaria*” y “*Other*”) tras un riguroso proceso de consenso entre tres anotadores. El *dataset* final constó de 14.057 parches, con un balance aceptable entre clases.

El núcleo técnico del trabajo fue el desarrollo de AlternarAI, una CNN personalizada con cinco capas convolucionales (de 3×3 píxeles), capas de *batch normalization* y activación *Leaky ReLU*, seguidas de *max pooling* y *average pooling*. La arquitectura se completó con dos capas *fully connected* y activaciones finales *sigmoid* para la salida binaria. Se introdujeron *dropout* (p=0,3) y regularización L2 para evitar sobreajuste. Este modelo se entrenó con el optimizador Adam, tasa de aprendizaje inicial de 0.001 y un esquema multietapa de reducción del *learning rate*, alcanzando una sólida generalización.

El desempeño de AlternarAI fue evaluado frente a arquitecturas reconocidas como VGG11, ResNet50, InceptionV3, DenseNet161 y EfficientNetV2, todas ellas preentrenadas en ImageNet y aplicadas mediante técnicas de *transfer learning* (*shallow* y *deep*). En conjunto, AlternarAI obtuvo métricas comparables en F1-score (0.91) y precisión (0.92) frente a las redes más complejas, pero con un número de parámetros quince veces menor y un tiempo de inferencia notablemente inferior (1,49 ms frente a más de 20 ms en EfficientNetV2). Además, el modelo mostró una elevada capacidad de generalización al alcanzar un F1-score de 0.89 cuando se entrenó con los datos de 2019-2021 y se evaluó con un conjunto independiente de 2022. También se confirmó su robustez frente a cierto grado de desequilibrio y errores de etiquetado.

En términos aplicados, AlternarAI demostró su viabilidad operativa al generar mapas de calor de infección a nivel de campo que identificaban correctamente las zonas inoculadas incluso cinco días tras la infección, ofreciendo un soporte potencial para la aplicación variable de fitosanitarios. A pesar de las elevadas exigencias de etiquetado en *datasets* de alta resolución, este enfoque sentó una base sólida para sistemas de apoyo a la decisión en agricultura de precisión. La investigación propuso como futuras líneas la reducción del esfuerzo de etiquetado mediante aprendizaje activo o modelos semi-supervisados.

A través de una propuesta experimental centrada en la automatización del monitoreo entomológico, Roosjen et al. (2020) evaluaron la viabilidad del uso de visión artificial y modelos de *deep learning* para la detección automatizada de *Drosophila suzukii* (SWD, por sus siglas en inglés) en trampas adhesivas, con especial interés en su aplicación mediante imágenes capturadas por UAV. El objetivo principal fue evaluar la capacidad de un modelo basado en redes neuronales convolucionales para detectar y contar individuos de SWD, diferenciando entre sexos, a partir de imágenes estáticas y aéreas, como una estrategia complementaria en programas de manejo integrado de plagas (IPM, por sus siglas en inglés).

Para la adquisición de datos, se emplearon trampas adhesivas de color rojo, que fueron expuestas en una cámara de cría para capturar ejemplares de SWD, y posteriormente colocadas al aire libre. Se capturaron 249 imágenes con una cámara Sony DSC-RX100M4 (20 MP), de las cuales 173 fueron destinadas al entrenamiento, 50 a validación y 26 al test. Las imágenes se anotaron manualmente identificando machos (DSM), hembras (DSF) y otros insectos (*bycatch*), alcanzando un total de 4.753 anotaciones.

Además, se tomaron 9 imágenes adicionales con la misma cámara montada sobre un UAV RKM4x, utilizadas exclusivamente para evaluar el desempeño en vuelo.

El modelo implementado fue una red neuronal convolucional profunda basada en ResNet-18, adaptada a un enfoque de detección espacial mediante la conversión de su arquitectura en una red totalmente convolucional. Se reemplazaron las capas finales por convoluciones 1×1 , funciones *ReLU*, *dropout* (50 %) y *SoftMax* para generar un mapa de probabilidad de tres clases (fondo, macho, hembra). El entrenamiento incluyó estrategias de aumento de datos (rotaciones, volteos) y se realizó con el optimizador Adam durante 300 ciclos de entrenamiento. La red se entrenó sobre parches aleatorios de 512×512 píxeles extraídos de las imágenes, y las predicciones se reconstruyeron sobre las imágenes completas mediante técnicas de “*sliding window*” y supresión no máxima.

En imágenes estáticas, el modelo alcanzó un AUC de 0.603 para machos y 0.506 para hembras, con un valor combinado de 0.669, destacando una mayor precisión en la detección de machos debido a la presencia de marcas distintivas como los puntos negros en las alas. En cambio, las imágenes UAV mostraron una notable disminución de rendimiento (AUC de 0.284 en machos y 0.086 en hembras), atribuida a la inestabilidad del vuelo y a la menor calidad visual. Aunque el modelo no fue entrenado con imágenes aéreas, logró detectar algunas instancias, lo que validó parcialmente su capacidad de generalización.

El estudio demostró la viabilidad técnica de aplicar redes convolucionales para la detección automática de SWD a partir de imágenes de trampas, diferenciando incluso el sexo del insecto. Si bien las limitaciones en imágenes UAV (como el enfoque, la iluminación y el ángulo de captura) restringieron su aplicabilidad actual, se planteó un futuro prometedor mediante el uso de UAVs autónomos y cámaras de mayor precisión. Esta prueba de concepto representó un avance significativo hacia sistemas automatizados de monitoreo de plagas en agricultura de precisión, aunque se requiere mejorar la robustez del modelo en condiciones reales de campo para su implementación práctica.

En conjunto, los estudios analizados demuestran que las redes convolucionales clásicas pueden constituir soluciones eficaces para la detección de agentes nocivos en cultivos mediante imágenes UAV, siempre que se adapten adecuadamente al contexto agronómico. La simplicidad estructural de estas arquitecturas favorece su implementación en entornos con recursos tecnológicos limitados, lo que las convierte en opciones accesibles para una agricultura de precisión más inclusiva. Sin embargo, su

rendimiento se muestra sensible a factores como la calidad del *dataset*, la resolución de las imágenes y el tipo de sensor empleado. Mientras que modelos optimizados, como AlternarAI, evidencian una clara viabilidad práctica, otros como ResNet-18 requieren ajustes sustanciales para lograr una aplicabilidad real en campo. Por tanto, estas soluciones resultan prometedoras cuando se integran en estrategias agronómicas coherentes y alineadas con las capacidades técnicas del usuario final.

4.3.1.2. Modelos de detección: comparación entre una y dos etapas.

Esta sección agrupó estudios que compararon modelos de detección de una etapa (YOLO) con arquitecturas de dos etapas (Faster R-CNN o DETR), integrando ambos enfoques en un mismo diseño experimental. Esta aproximación permitió explorar de forma coherente las diferencias metodológicas y operativas entre las dos estrategias, facilitando un análisis comparativo en condiciones controladas. Mientras que los modelos de una etapa se caracterizaron por su rapidez y eficiencia computacional, las arquitecturas de dos etapas ofrecieron una mayor precisión en entornos agrícolas complejos, a costa de un mayor consumo de recursos. Esta dualidad abrió interrogantes relevantes sobre la idoneidad de cada enfoque según el tipo de cultivo, la finalidad del análisis y la disponibilidad tecnológica en el campo.

Con el propósito de enfrentar las limitaciones del monitoreo tradicional en regiones de alta heterogeneidad vegetal y complejidad topográfica, Mora et al. (2024) plantearon un enfoque para la detección de la marchitez bacteriana del plátano (BXW), una enfermedad devastadora en los sistemas agrícolas de África Central, donde la diversidad vegetal y las condiciones topográficas dificultaron el monitoreo convencional. Para ello, se diseñó un sistema basado en UAV e imágenes espectrales procesadas mediante *deep learning*, utilizando como casos de prueba imágenes RGB, multispectrales (MS) y fusionadas por *pansharpening* (PAN), captadas en el noreste del Congo entre el intervalo de años de 2019 y 2022.

Se emplearon drones DJI Phantom 4 Pro equipados con sensores RGB (FC6310) y multispectrales (MicaSense RedEdge), con vuelos a 60 m y 100 m de altura. Las imágenes se ortorrectificaron y fusionaron mediante técnicas de *pansharpening*. El *dataset* incluyó más de 16.000 anotaciones manuales, divididas entre tallos sanos e infectados, categorizados mediante observación en campo. Para mitigar el desbalance de clases, se aplicó un riguroso proceso de aumentación de datos basado en seis técnicas para evitar el sobreajuste del modelo durante la experimentación (rotación, traslación,

cambio de brillo, etc.), logrando una proporción equilibrada entre clases todas las clases de la experimentación.

Se compararon dos arquitecturas:

- Faster R-CNN (dos etapas), basado en ResNet-50 y entrenado con *TensorFlow*. Su estructura se basa en una red de propuesta de regiones (RPN) y un clasificador final que refina la localización mediante *RoIAlign*.
- YOLOv8 (una etapa), implementado con la librería Ultralytics. Esta versión mejora versiones previas al introducir un backbone C2f, un módulo de fusión con FPN y PAN, y una *decoupled head* para tareas de localización y clasificación por separado. Ambas redes se entrenaron mediante aprendizaje por transferencia con pesos preentrenados en COCO y se validaron con imágenes no vistas.

Los resultados mostraron que YOLOv8 obtuvo mejores métricas globales a 100 m de altura, alcanzando un F1-score del 0.96 para tallos infectados en imágenes RGB y del 0.92 en PAN, con *MissedGTError* por debajo del 10 %. Sin embargo, su rendimiento se degradó significativamente a 60 m, con tasas de error por omisión superiores al 50 %, posiblemente debido al ruido visual y la complejidad de los detalles a corta distancia. Por el contrario, Faster R-CNN mostró mayor robustez a variaciones de altura y tipo de imagen. En particular, las imágenes PAN alcanzaron un F1-scores de 0.82 para tallos enfermos y 0.91 para sanos, mostrando una estabilidad destacable frente a cambios de resolución. Las combinaciones MS mostraron limitaciones debido a la baja resolución espacial, aunque algunos índices vegetativos (como NDVI, NPCI y NDRE) mejoraron la detección de tallos sanos.

El enfoque propuesto demostró un alto potencial para la vigilancia temprana de BXW en paisajes agrícolas complejos mediante UAVs y DL. Mientras YOLOv8 se consideró óptimo para misiones rápidas a gran altitud por su precisión y eficiencia, Faster R-CNN resultó más adecuado para entornos variables donde se requieren modelos estables y adaptables. El estudio recomendó el uso de imágenes fusionadas PAN como alternativa eficiente y accesible frente a sensores hiperespectrales costosos. A futuro, la mejora en la resolución espacial de sensores MS y la integración con métodos diagnósticos más específicos podría permitir una clasificación más precisa entre BXW y otros estreses con síntomas similares.

Por otro lado, abordando una enfermedad de alto impacto en la producción de caña de azúcar en Sri Lanka, Amarasingam et al. (2022) evaluaron el rendimiento de distintas arquitecturas de *deep learning* para la detección de la enfermedad *White Leaf Disease* (WLD), utilizando imágenes RGB capturadas por UAV. La investigación abordó una problemática de gran impacto económico en el sector azucarero, proponiendo un enfoque automatizado y eficiente frente a los métodos tradicionales, frecuentemente limitados por su subjetividad y bajo rendimiento.

La adquisición de datos se realizó mediante un dron DJI Phantom 4 con sensor RGB CMOS de 2.08 MP, operado a 20 m de altura y una resolución de 1.1 cm/píxel. Las imágenes fueron tomadas en condiciones controladas y posteriormente ortorrectificadas. El *dataset* se segmentó en teselas de 640×640 píxeles para entrenamiento (110), validación (60) y prueba (40), ampliadas mediante técnicas de aumentación de datos aplicando (rotación, brillo, desenfoque, volteo) para evitar el sobreajuste del modelo hasta alcanzar 1200, 240 y 240 imágenes, respectivamente. Las muestras se etiquetaron manualmente usando LabelImg, siguiendo el formato YOLO, y fueron validadas por expertos en fotointerpretación.

Se compararon cuatro modelos de detección de objetos: YOLOv5, YOLOR, Faster R-CNN y DETR. Todos fueron entrenados en Google Colab Pro+ empleando arquitecturas preentrenadas, adaptadas a una única clase (WLD). En cuanto a estructura, YOLOv5 destacó como un detector de una etapa con arquitectura eficiente basada en *CSPDarknet* y capas *PANet*. YOLOR introdujo un diseño híbrido que combinaba conocimiento implícito y explícito, mientras que Faster R-CNN representó un enfoque de dos etapas tradicionalmente robusto, y DETR incorporó *Transformers* para detección *set-based*, aunque con mayores requerimientos computacionales.

Los resultados mostraron un claro liderazgo del modelo YOLOv5, que alcanzó una precisión del 95 %, exhaustividad del 92 %, $mAP@0.50$ del 93 % y $mAP@0.95$ del 79 %, con tan solo 14 MB de peso y un tiempo de entrenamiento de 6 horas. YOLOR obtuvo una mayor sensibilidad (93 % exhaustividad), pero a costa de una menor precisión y un tamaño de 281 MB. DETR presentó el peor rendimiento global (77 % precisión, 69 % exhaustividad, 41 % $mAP@0.95$), y fue el modelo más costoso en entrenamiento (30 horas y 473 MB). Faster R-CNN ofreció resultados intermedios (90 % precisión, 95 % $mAP@0.50$), destacando por su velocidad de entrenamiento (3 horas).

Como conclusión aplicada, el modelo YOLOv5 se posicionó como la opción más equilibrada entre precisión, eficiencia y adaptabilidad para entornos agrícolas reales. Su bajo peso lo hizo especialmente idóneo para drones ligeros y sistemas operativos móviles. Los autores señalaron que, aunque los resultados son prometedores, existen desafíos asociados a la similitud visual entre plantas infectadas y el fondo del terreno, así como a la limitada variabilidad del *dataset*. Por lo que, se recomendó explorar futuras mejoras mediante el uso de sensores multispectrales o hiperspectrales y técnicas de enmascaramiento espectral. Así, el enfoque propuesto constituyó un avance viable y replicable hacia una agricultura de precisión más autónoma y efectiva frente a enfermedades como WLD.

Ambos estudios coinciden en señalar que los modelos de detección de una etapa ofrecen un equilibrio destacable entre precisión y eficiencia computacional, lo que los hace especialmente adecuados para su aplicación en condiciones reales de campo. Si bien las arquitecturas de dos etapas presentan una mayor robustez técnica, su implementación requiere capacidades computacionales y recursos operativos que no siempre están disponibles en el ámbito agrario. En consecuencia, la viabilidad de cada enfoque no depende únicamente de su rendimiento técnico, sino también de su compatibilidad con las limitaciones estructurales del entorno y con el perfil tecnológico de los usuarios finales. Esta realidad refuerza la necesidad de adaptar las soluciones tecnológicas a las condiciones específicas de cada explotación agrícola.

4.3.1.3. Modelos combinados.

En este apartado se analizaron propuestas metodológicas que combinaron arquitecturas convolucionales con otros tipos de redes neuronales o componentes complementarios. Estos modelos híbridos se diseñaron con el propósito de mejorar la capacidad de generalización, optimizar la eficiencia computacional o incrementar la adaptabilidad en contextos agrícolas variables. Las estrategias incluyeron la integración de redes recurrentes (RNN, por sus siglas en inglés), mecanismos de adaptación de dominio o la fusión de distintas variantes de CNN. Su presencia en esta revisión respondió al interés creciente por soluciones versátiles y eficaces, capaces de equilibrar precisión, velocidad y operatividad en condiciones reales de uso.

Como parte de los avances hacia sistemas de diagnóstico agrícola más portables y eficientes, Ksibi et al. (2022) propusieron una arquitectura híbrida de *deep learning* denominada MobiRes-Net, con el objetivo de mejorar la clasificación multiclase de

enfermedades en hojas de olivo a partir de imágenes captadas por UAV. El estudio, desarrollado en un olivar de Arabia Saudí, se enmarca en la búsqueda de herramientas automáticas, eficientes y portables para el diagnóstico temprano de enfermedades como la mancha de pavo real (*Peacock spot*), la tuberculosis del olivo (*Olive scab*) y la infección por la seta del olivo (*Omphalotus olearius*).

El conjunto de datos se construyó a partir de 5.400 imágenes obtenidas mediante un dron de ala fija equipado con una cámara RGB, volando a una altitud de 25–30 m. Las imágenes fueron tomadas en un olivar de 30 ha durante un período de varios meses. Tras su captura, se realizó una fase de preprocesamiento (reescalado a 224×224 píxeles, normalización del contraste y mejora del brillo), y se aplicaron técnicas de aumento de datos (rotación y desenfoque gaussiano) para equilibrar las clases y mejorar la robustez del modelo. Las imágenes se clasificaron en cuatro clases por un ingeniero agrónomo: “Healthy”, “*Aculus olearius*”, “*Olive scab*” y “*Peacock spot*”.

El modelo propuesto, MobiRes-Net, integró las arquitecturas preentrenadas ResNet50 y MobileNet mediante un mecanismo de concatenación de características. Se eliminaron las capas clasificadoras originales de ambas redes, sustituyéndolas por una capa de *average pooling* y una capa totalmente conectada de 1024 neuronas en cada rama. Posteriormente, se fusionaron ambas salidas mediante concatenación y se añadieron tres capas densas adicionales (1024, 512 y 256 neuronas), finalizando con una capa *Softmax* para la clasificación multiclase. Esta configuración permitió aprovechar la capacidad de ResNet50 para modelar patrones complejos, junto con la eficiencia computacional de MobileNet, especialmente adecuada para su implementación en dispositivos móviles o entornos con recursos limitados.

El modelo alcanzó una precisión global del 97,08 %, superando los rendimientos individuales de ResNet50 (94,86 %) y MobileNet (95,63 %), lo cual justificó su combinación. Las métricas por clase (no detalladas en el resumen) presentaron F1-*Scores* superiores al 96 %, lo que indicó un rendimiento equilibrado entre precisión y exhaustividad en la detección de las diferentes patologías. No obstante, los ensayos se realizaron en condiciones controladas, con fondos poco complejos, lo que limitó parcialmente la extrapolación de los resultados a escenarios reales más heterogéneos.

La propuesta representó un avance prometedor hacia la creación de sistemas portables y eficaces para el diagnóstico automático de enfermedades en olivar. La estructura híbrida MobiRes-Net ofreció un compromiso sólido entre rendimiento y eficiencia

computacional, adaptándose a dispositivos móviles de bajo coste. No obstante, el estudio reconoció que su viabilidad en condiciones reales requiere validaciones adicionales. La integración de este enfoque en explotaciones agrícolas de tamaño medio podría facilitar la toma de decisiones agronómicas precisas, reducir el uso innecesario de fitosanitarios y mejorar la sostenibilidad de los sistemas de producción olivarera.

Con el propósito de abordar la complejidad de los patrones espaciotemporales en cultivos, Abdalla et al. (2024) propusieron un enfoque híbrido de *deep learning* para estimar la severidad de la enfermedad por *Fusarium oxysporum*, a partir de imágenes RGB multitemporales captadas por drones. El objetivo fue superar los desafíos derivados de la variabilidad espaciotemporal y del cambio de dominio entre campañas agrícolas, mediante una arquitectura que combinó redes convolucionales profundas con redes recurrentes temporales.

Se recopilaron imágenes con un dron DJI Phantom 4 Pro a 25 m de altura en dos campañas (2020 y 2021), obteniéndose un total de 3.684 imágenes de parcelas experimentales etiquetadas en 5 niveles de severidad según la evaluación visual de raíces. Las imágenes fueron segmentadas utilizando un *encoder* basado en VGG16 y posteriormente etiquetadas según el grado de necrosis vascular en raíces. Para corregir el desbalance de clases, se aplicó *median frequency balancing* en la función de pérdida.

El modelo propuesto se basó en una combinación CNN–BiLSTM. Primero, se evaluaron cinco arquitecturas CNN preentrenadas para extracción de características: AlexNet, Inceptionv3, VGG, ResNet18 y ResNet101, seleccionándose esta última por su mayor capacidad discriminativa (visualizada mediante t-SNE). Los vectores extraídos se introdujeron secuencialmente a una red BiLSTM, que permitió modelar la progresión temporal de los síntomas a lo largo del ciclo del cultivo.

Para aumentar la robustez del modelo frente a diferencias entre campañas, se aplicó una técnica de adaptación de dominio basada en *Correlation Alignment* (CORAL), que ajustó las distribuciones estadísticas de las características extraídas por la CNN entre los conjuntos fuente y objetivo.

El modelo ResNet101-BiLSTM alcanzó una precisión del 89,7 % en la clasificación intranual (2020) y del 72,7 % en clasificación interanual (2021), evidenciando una adecuada capacidad de generalización tras aplicar CORAL. Frente a modelos clásicos como SVM o CNN sin componente secuencial, la BiLSTM demostró un rendimiento

claramente superior en métricas como *F1-score* (87,7 %) y *recall* (87,9 %), especialmente en tareas donde la evolución temporal de la sintomatología es clave. El modelo mostró una menor capacidad para discriminar correctamente la clase 2, correspondiente a un nivel moderado de severidad, debido a su solapamiento visual con clases adyacentes (1 y 3). Esta clase obtuvo los valores más bajos de área bajo la curva (AUC), lo que evidenció una mayor ambigüedad en su representación espectral y una dificultad añadida para su clasificación precisa.

La iniciativa mostró un elevado potencial para aplicaciones reales, al apoyarse en UAVs y sensores RGB de bajo coste, prescindiendo de plataformas hiperspectrales complejas. Su capacidad para evaluar la progresión de enfermedades radicales no visibles inicialmente en el dosel vegetal resulta crítica para decisiones agronómicas como la selección varietal o la programación de tratamientos. No obstante, el estudio señaló limitaciones como el reducido número de muestras en clases severas y sugirió futuras mejoras en técnicas de adaptación de dominio y validación en entornos más diversos.

Los estudios analizados evidencian que las arquitecturas híbridas pueden aportar soluciones más adaptables y eficaces para el monitoreo de enfermedades agrícolas. La portabilidad del modelo MobiRes-Net y la capacidad espaciotemporal de la arquitectura ResNet101-BiLSTM reflejan avances significativos hacia herramientas inteligentes capaces de responder a necesidades agronómicas concretas. Si bien es necesario validar su desempeño en entornos reales de mayor complejidad, estas propuestas destacan por su coherencia con las condiciones operativas del campo, combinando precisión técnica con factibilidad práctica. En este sentido, se consolidan como alternativas con alto potencial de transferencia al sector agrícola.

4.3.1.4. Modelos con mecanismos de atención.

Este grupo reunió estudios que incorporaron mecanismos de atención espacial, espectral o multicanal en las arquitecturas de detección. Estos enfoques buscaron mejorar la capacidad del modelo para focalizarse en regiones clave de la imagen o asignar mayor relevancia a determinadas bandas espectrales o zonas del dosel vegetal. Su aplicación resultó especialmente valiosa en escenarios donde los síntomas de plaga o enfermedad eran poco visibles, estaban dispersos o se manifestaban en fases tempranas del desarrollo del cultivo.

Con el fin de optimizar la detección de enfermedades foliares desde diferentes perspectivas aéreas, Lv et al. (2024) abordaron la identificación de la roya sureña del maíz (*Puccinia polysora*) mediante imágenes multiespectrales adquiridas con UAV desde diferentes ángulos (nadir, 15° y 30°). Con el objetivo de mejorar la monitorización en campo, los autores desarrollaron y compararon modelos de *deep learning* basados en mecanismos de atención aplicados a información espectral multiángulo.

La toma de datos se llevó a cabo en Sanya (China), zona endémica de SCR. Se empleó un dron DJI Phantom 4 Multispectral, capturando imágenes en cinco bandas (azul, verde, rojo, borde rojo y NIR) con una resolución espacial de 1,5 cm/píxel. Se obtuvieron 149 parcelas clasificadas en cuatro niveles de severidad mediante inspección visual. Las imágenes se capturaron en tres ángulos (0°, 15° y 30°), georreferenciadas y calibradas radiométricamente. La segmentación de las áreas de maíz se realizó con *Random Forest*, y posteriormente se extrajeron 15 índices espectrales, entre los cuales GPRI, NPCI, VARI y RENDVI fueron seleccionados como predictores óptimos.

Se evaluaron tres enfoques para la clasificación de la severidad de SCR. En primer lugar, se utilizó un modelo tradicional basado en *Random Forest* (RF), empleado como línea base por su eficacia en conjuntos de datos pequeños y su capacidad para identificar variables relevantes mediante importancia de características. En segundo lugar, se desarrolló una *Fully Connected Network* (FCN) de baja complejidad, que integró un módulo de atención en su capa inicial para asignar pesos adaptativos a los índices espectrales procedentes de distintas vistas. Finalmente, se diseñó una CNN con módulos CBAM de atención espacial y por canal, estructurada en tres flujos paralelos para cada ángulo de captura, cuyos resultados se fusionaron en una capa de concatenación. Ambas arquitecturas de *deep learning* fueron implementadas en *TensorFlow* y *Keras*, aplicando regularización mediante *dropout* y normalización por lotes para prevenir el sobreajuste.

La FCN con atención logró la mayor precisión (OA: 82,1 %; F1 macro: 0,82), superando tanto al modelo RF (~8 %) como a la CNN (~12 %). El análisis de pesos indicó que las vistas inclinadas (especialmente 15°) aportaron más información útil, debido al patrón de infección ascendente característico de la SCR. La CNN, a pesar de incorporar mecanismos de atención, fue menos efectiva en este caso por su menor capacidad de explotar características espectrales en contextos con muestras limitadas. Además, se observó una mejora notable (~10 %) en las métricas al combinar vistas múltiples frente a vistas nadir, confirmando la utilidad de la fusión multiángulo.

Este estudio demostró el potencial de las arquitecturas FCN con atención para monitorizar enfermedades foliares en condiciones reales. La metodología resultó especialmente relevante para patologías con patogénesis ascendente, ya que las vistas inclinadas permitieron detectar síntomas en capas inferiores del dosel vegetal. Aunque la validación externa se limitó a un único ensayo, los resultados respaldaron la viabilidad de este enfoque para optimizar decisiones agronómicas basadas en datos y reducir el uso de fitosanitarios en sistemas de agricultura inteligente.

La integración de mecanismos de atención en los modelos de *deep learning* mejora notablemente su sensibilidad ante síntomas poco visibles o distribuidos en capas inferiores del dosel vegetal. Además, la posibilidad de ajustar el ángulo de captura según la fisiología del cultivo incrementa la aplicabilidad práctica de estos enfoques en plataformas UAV convencionales. Aunque aún se requiere un mayor número de estudios para confirmar su robustez en contextos diversos, este tipo de arquitecturas representa una vía de innovación prometedora para el desarrollo de sistemas de detección temprana en agricultura de precisión.

4.3.1.5. Modelos especializados.

Dentro de este apartado se consideraron enfoques metodológicos que integraron múltiples modelos en un mismo sistema analítico, combinando tareas de detección y regresión a través de arquitecturas diferenciadas. Estas propuestas se orientaron a mejorar no solo la identificación de enfermedades, sino también la estimación cuantitativa de su severidad, lo que permitió desarrollar herramientas más completas y funcionales para la toma de decisiones agronómicas en campo.

Contribuyendo a la innovación en teledetección hiperespectral aplicada al monitoreo de enfermedades, Zhang et al. (2024) desarrollaron una metodología para la detección y cuantificación de la roya en *Zanthoxylum bungeanum* mediante imágenes hiperespectrales captadas con UAV. El estudio introduce el marco OQIF (*Object-based Quantitative Inversion Framework*), que combina técnicas de detección de objetos y regresión profunda para estimar de forma automatizada el índice de enfermedad (DI). Esta propuesta aborda un vacío crítico en la literatura, ya que representa la primera aplicación de teledetección hiperespectral UAV para esta enfermedad a nivel de planta individual, lo cual tiene implicaciones relevantes en la gestión agronómica de precisión.

Se empleó un UAV DJI M600 Pro equipado con una cámara hiperespectral S185 (450–950 nm, 125 bandas, 4 nm de resolución espectral) sobre parcelas de 8.000 m² en Chongqing (China), donde se monitorearon 2.558 plantas durante seis campañas entre 2021 y 2022. Las imágenes fueron calibradas radiométricamente y segmentadas en teselas de 512×512 píxeles, generando un *dataset* supervisado con anotaciones manuales (posición + DI). La fase de entrenamiento se estructuró en tres conjuntos (entrenamiento, validación y test), sin solapamientos entre ellos para evitar el sobreajuste del modelo experimental.

El sistema OQIF constó de dos ramas complementarias. La primera, de detección de objetos, empleó la arquitectura Fast R-CNN con backbone ResNet50 y estructura PAFPN como “*neck*”, alcanzando un AP@0.5 de 0.900. Esta rama permitió localizar automáticamente cada planta mediante *bounding boxes*.

La segunda rama, de regresión del índice de enfermedad, se basó en una versión modificada de EfficientNetV2 denominada RustEfficientNet, que incorporó un módulo PSA (*Polarized Self-Attention*) adaptado al cubo espectral. Esta red realizó regresión continua del DI a partir de los ROIs extraídos, utilizando funciones de pérdida MSE y activación *Sigmoid*. El entrenamiento se realizó con más de 100.000 muestras, obteniendo un R² de 0.90 y un RMSE global de 3.97. En plantas con DI < 10, el RMSE fue de 2.48, destacando su potencial para detección temprana.

Los resultados mostraron una alta fiabilidad tanto en la localización de plantas como en la cuantificación del daño. RustEfficientNet superó a modelos como MobileNetV3 o SwinV2, mostrando mayor robustez ante variaciones en la posición del ROI. Además, se evaluó la sensibilidad espectral mediante más de 180.000 combinaciones de índices de vegetación, confirmando la importancia de las bandas en el infrarrojo cercano. No obstante, el uso de *bounding boxes* rectangulares podría incluir información de fondo no deseada, lo cual fue reconocido como limitación metodológica por los autores.

Este enfoque híbrido demostró gran utilidad en entornos reales, lo que permitió no solo detectar plantas afectadas, sino generar mapas de severidad continuos que facilitaron la aplicación localizada de tratamientos fitosanitarios. Su capacidad de adaptación a otras variables agronómicas, como el rendimiento, reforzó su valor como herramienta escalable para agricultura de precisión. A pesar de las limitaciones en segmentación morfológica y en la interpretabilidad del modelo, OQIF representó un avance significativo hacia sistemas automatizados de diagnóstico con alto potencial operativo en campo.

Esta línea metodológica demuestra que la integración de tareas de detección y regresión en un mismo sistema puede ofrecer soluciones más completas y operativas ante los retos de la agricultura de precisión. La capacidad de generar mapas de severidad a partir de imágenes hiperespectrales obtenidas mediante UAVs facilita un monitoreo más detallado del cultivo y permite ajustar las intervenciones fitosanitarias de forma localizada y eficiente. Aunque persisten desafíos técnicos, como la mejora en la definición espacial de las regiones analizadas, este tipo de propuestas especializadas marca una dirección prometedora hacia sistemas de diagnóstico agronómico con alto valor estratégico.

4.3.1.6. Modelos conceptuales.

Como cierre de este bloque, se incluyeron propuestas metodológicas de carácter conceptual que, aunque no fueron validadas en condiciones reales de vuelo con UAVs, delinearon nuevas vías para el diseño de sistemas más eficientes en la detección de plagas y enfermedades. Estos enfoques exploratorios integraron técnicas de optimización evolutiva, visión artificial e inteligencia computacional, y apuntaron a mejorar tanto el rendimiento como la capacidad de adaptación de las redes convolucionales en escenarios agrícolas futuros.

Desde una perspectiva conceptual orientada a la integración de tecnologías emergentes, Vijayalakshmi y Radhika (2024) presentaron un enfoque innovador basado en la combinación de UAVs e IoT para la detección automatizada de plagas agrícolas mediante un modelo CNN optimizado, denominado ACBWO-CNN. El objetivo principal fue diseñar un sistema eficiente de detección temprana de plagas apoyado en visión artificial, procesamiento en la nube y técnicas de optimización.

Las imágenes fueron recolectadas mediante UAVs equipados con sensores RGB y procesadas a través de una arquitectura IoT basada en la nube. El conjunto de datos fue generado de manera artificial mediante un *script* de *scraping* en Google, incluyendo 10 clases de plagas (gusano del algodón, ácaros, oruga militar, pulgones, escarabajos, saltamontes, mosquitos, sierras y barrenadores del tallo), con 300 imágenes de entrenamiento y 50 de validación por clase, todas en formato RGB de 225×225 píxeles. El proceso incluyó una fase de segmentación (*K-means clustering*) para aislar el fondo, seguida de la extracción de descriptores LBP e *Improved LVP*.

El modelo propuesto se basó en una CNN estructurada en capas de convolución, *pooling* y clasificación totalmente conectada. Para optimizar el rendimiento, se integró el

algoritmo ACBWO (*Arithmetic Crossover based Black Widow Optimization*), una técnica evolutiva que ajusta dinámicamente los pesos de la red durante el entrenamiento. Este algoritmo mejora la convergencia del modelo mediante operadores de cruce aritmético y mecanismos de canibalismo inspirados en la biología de las arañas viudas negras, superando las limitaciones del BWO tradicional. La arquitectura también incorporó descriptores mejorados de textura (*Improved LVP*) para enriquecer la representación espacial de patrones de plaga.

El modelo ACBWO-CNN alcanzó una precisión del 94 %, con valores elevados en sensibilidad (91,9 %), especificidad (91,9 %) y F1-score (91,9 %). Superó consistentemente a otras variantes como BWO-CNN, POA-CNN, y COOT-CNN en métricas positivas y redujo significativamente errores tipo I y II (FPR y FNR). Además, se realizaron estudios de ablación que demostraron que la supresión de los módulos de optimización o de descriptores LVP reducía notablemente el rendimiento, validando su aporte. La robustez del modelo también fue confirmada mediante validación cruzada ($k=5$), manteniendo métricas superiores en todos los escenarios.

Aunque el enfoque no fue validado con imágenes UAV en condiciones reales de campo, los resultados virtuales revelaron un potencial considerable para su futura implementación. Su arquitectura optimizada permitió una ejecución eficiente en entornos con recursos computacionales limitados, característica valiosa para explotaciones agrícolas de pequeña escala. La combinación de técnicas de optimización con CNN representó una línea prometedora para el desarrollo de sistemas ligeros, adaptables y de alta precisión en tareas de detección temprana de plagas, contribuyendo así a una agricultura más sostenible y tecnológica.

Estos enfoques conceptuales abren líneas de innovación relevantes para el futuro de la detección automatizada en agricultura, al incorporar principios como la optimización evolutiva y la inteligencia computacional. Aunque aún requieren validación en condiciones reales, sus planteamientos apuntan a una mejora sustancial en la capacidad de generalización y robustez de los modelos, especialmente en contextos complejos y variables. Este tipo de propuestas anticipa el desarrollo de sistemas más adaptativos, eficientes y sostenibles, alineados con las demandas emergentes del sector agroalimentario.

4.3.2. Estimación de rendimiento agrícola.

4.3.2.1. Modelos de detección: una etapa.

La revisión de esta temática se inició con modelos de detección de una etapa (YOLO, SSD y EfficientDet), reconocidos por su arquitectura directa y elevada velocidad de inferencia. Estas arquitecturas fueron implementadas en tareas de conteo y estimación de rendimiento, mostrando una notable capacidad para identificar y cuantificar frutos u órganos vegetativos en condiciones reales de campo. Su principal ventaja radicó en la eficiencia operativa, lo que las convierte en herramientas especialmente adecuadas para contextos donde el análisis en tiempo real resulta prioritario.

Con el propósito de automatizar el conteo de frutos en condiciones reales de campo, Xiong et al. (2020) desarrollaron un enfoque basado en visión artificial y *deep learning* para la detección visual de mangos verdes, con el objetivo de estimar rápidamente el número de frutos por árbol. Para ello, utilizaron imágenes RGB adquiridas mediante un UAV comercial y entrenaron un modelo YOLOv2, optimizado para la detección rápida de objetos pequeños en imágenes aéreas.

Las imágenes se capturaron en un huerto de mango en Guangzhou (China) utilizando un dron DJI Phantom 3 Standard, equipado con una cámara de 12 MP (4000×3000 píxeles). Se realizaron vuelos a una altura de 1,5–2 m en ambos laterales del árbol y se aplicó un algoritmo de mosaico (SURF) para reconstruir vistas completas. De las 471 imágenes iniciales, se seleccionaron 360 para el *dataset* (300 entrenamiento, 60 test), etiquetadas manualmente con LabelImg. El conjunto incluyó diferentes condiciones de iluminación y niveles de oclusión.

El modelo utilizado fue YOLOv2, una red de tipo de detección de una etapa construida sobre Darknet-19 (19 capas convolucionales + 5 capas de *max pooling*), entrenada con imágenes reescaladas a 544×544 píxeles. El modelo se preentrenó con pesos oficiales y se optimizó con tasa de aprendizaje inicial de 0.01, batch size de 64 y un total de 3000 iteraciones, tras analizar la convergencia de la función de pérdida. El entrenamiento se ejecutó en un entorno Ubuntu 16.04 con GPU NVIDIA GTX 1080.

YOLOv2 alcanzó una precisión del 96.1% y un *recall* del 89.0% sobre el conjunto de prueba (60 imágenes, 438 frutos). La tasa de error en la estimación del número de frutos por árbol fue de solo 1.1% (media en 10 árboles). El rendimiento del modelo fue superior a los métodos tradicionales de segmentación (Otsu, K-means, FCM), con mejor precisión

y menor tiempo de ejecución (0.08 s/imagen). Sin embargo, la precisión disminuyó en imágenes con más de 10 frutos o bajo retroiluminación, debido a oclusiones, solapamientos y degradación del contraste.

El estudio demostró que YOLOv2 es una herramienta eficiente para la detección y cuantificación de frutos de mango mediante UAV en escenarios reales, aportando rapidez operativa y precisión suficiente para tareas de monitoreo agronómico. Como limitación, destacó su sensibilidad a condiciones de luz adversa y la necesidad de reentrenar el modelo ante variedades distintas. Aun así, representó una solución rentable y reproducible en campañas agrícolas, con potencial de extrapolación a otros cultivos arbóreos bajo condiciones similares.

En el contexto de la automatización del conteo de estructuras reproductivas en cultivos extensivos, Li et al. (2022a) llevaron a cabo un estudio comparativo para evaluar la eficacia de distintos algoritmos de detección basados en *deep learning* en el conteo automático de espigas de sorgo mediante imágenes capturadas con UAVs. El experimento se desarrolló en una parcela experimental de 2 hectáreas ubicada en el Delta del Río Amarillo (China), una región salinizada donde el sorgo se cultiva ampliamente por su tolerancia a suelos adversos.

Para la toma de datos, se empleó un dron DJI Phantom 4 Multispectral, que integraba sensores RGB y multiespectrales, alcanzando una resolución de 1,1 cm/píxel. Se obtuvieron un total de 928 imágenes, de las cuales se seleccionaron 480 (384 para entrenamiento y 96 para validación), y las espigas se etiquetaron utilizando una estrategia semiautomática con revisión manual a través de LabelImg. Cabe destacar que, debido al pequeño tamaño de las espigas (7×7 píxeles), se aplicó una estrategia de entrenamiento preliminar para mejorar la precisión de etiquetado.

A nivel metodológico, los autores optaron por evaluar tres algoritmos de detección one-stage: EfficientDet, SSD y YOLOv4. Estas arquitecturas, conocidas por su capacidad de detección en tiempo real, fueron seleccionadas por su potencial para equilibrar precisión y eficiencia computacional, aspecto clave en escenarios agrícolas con imágenes de alta resolución. Mientras que EfficientDet se basa en un backbone EfficientNet y una pirámide de características bidireccional (BiFPN), buscando optimizar la fusión multiescala de información semántica, SSD recurre a una estructura con base VGG-16 y pirámide de detección, diseñada para abordar objetos de diferentes tamaños mediante marcos ancla. Por su parte, YOLOv4 incorpora una arquitectura más compleja con CSPDarknet53 como

extractor de características, combinando módulos SPP y PAN para mejorar la propagación de información entre capas profundas y superficiales. Aunque sus enfoques difieren, las tres arquitecturas comparten el objetivo de localizar y contar objetos en imágenes densas, lo que las convierte en candidatas idóneas para el conteo automatizado de espigas.

Una vez entrenados los modelos bajo condiciones uniformes mediante transferencia de aprendizaje, se analizaron los resultados considerando métricas estándar como precisión, *recall*, *F1-score* y *average precision*. El modelo YOLOv4 fue el que presentó el mejor equilibrio entre rendimiento y robustez, logrando un AP del 84,51 % y un coeficiente de determinación $R^2 = 0,99$ en el conteo por imagen, lo que evidenció su alta capacidad de adaptación frente a diferentes niveles de densidad. SSD, aunque más rápido en la detección (0,016 s por imagen), mostró problemas de sobredetección en parcelas densas, especialmente al utilizar umbrales de solapamiento bajos. En contraste, EfficientDet presentó un elevado valor de precisión puntual, pero su tasa de recuperación fue muy baja, lo que limitó su utilidad práctica. A lo largo del estudio, se comprobó que un umbral de IoU y de confianza de 0,3 ofrecía un equilibrio razonable entre detección verdadera y falsos positivos para las tres arquitecturas.

En términos aplicados, los autores concluyeron que YOLOv4 es el modelo más idóneo para su implementación en tareas de agricultura de precisión relacionadas con el conteo de espigas, tanto por su precisión como por su comportamiento estable ante diferentes condiciones de campo. Si bien SSD podría ser considerado en entornos donde la velocidad de procesamiento sea prioritaria, su tendencia a la sobre-detección limita su aplicabilidad en contextos reales. Por el contrario, EfficientDet mostró limitaciones significativas frente a objetivos de tamaño reducido, lo que lo descarta como alternativa viable en este caso. En conjunto, el estudio aportó una base metodológica sólida para la selección de modelos DL en tareas de monitoreo agrícola con UAVs, y plantea un marco replicable para futuras investigaciones en detección de biomasa o estimación de rendimiento.

Con el propósito de automatizar la estimación fenológica en cultivos frutales, Lin et al. (2022) desarrollaron un sistema automático para estimar la tasa de floración en árboles de lichi mediante imágenes RGB captadas con UAV y modelos de detección basados en *deep learning*. El objetivo fue sustituir el conteo manual (costoso y subjetivo) por un método rápido y preciso, validado en un huerto real con variedad Guiwei.

Se utilizaron vuelos con un dron DJI Mavic 2 Pro, equipado con cámara Hasselblad L1D-20c (RGB), a 5,9 m de altura con ángulo nadir, durante el periodo de floración (marzo-abril). El *dataset* final consistió en 1.200 imágenes de 1.280×720 píxeles tras recorte y aumentación de datos. Se etiquetaron manualmente 47.192 objetos (38.362 inflorescencias y 8.830 brotes vegetativos) en formato VOC.

Se implementó YOLOv4, una arquitectura de una sola etapa con backbone CSPDarknet53, módulo SPP y PANet como fusión de características. El modelo se optimizó seleccionando la mejor combinación de parámetros: tamaño de entrada 640×640 píxeles, *batch size* de 4/2 (fase congelada/descongelada) y 1.000 ciclos de entrenamiento. Esta configuración alcanzó una mAP del 87,87 %, con una velocidad de detección de 0,043 s/imagen. Para estimar la tasa de floración por árbol, se ajustaron ecuaciones de regresión (binomial para inflorescencias y lineal para brotes) entre conteos reales y predichos.

El modelo YOLOv4 superó a Faster R-CNN, SSD, CenterNet y YOLOv4-tiny en precisión, velocidad y robustez, especialmente en condiciones de luz trasera y escenas con baja densidad floral. El error medio en la estimación de inflorescencias fue del 4,20 %, del 2,85 % en brotes vegetativos, y del 1,135 % en la tasa de floración; dichos valores corresponden a errores relativos absolutos respecto al conteo manual, utilizados como medida de viabilidad agronómica. Las principales limitaciones derivaron de la vista nadir (sin información lateral) y la confusión con fondos complejos, especialmente en condiciones de sobreexposición.

Este enfoque resultó altamente viable para automatizar el conteo floral en huertos de lichi, con potencial transferencia a otras especies frutales. Su utilidad agronómica radica en facilitar decisiones de poda, fertilización y regulación floral. No obstante, la futura integración de imágenes multivista o reconstrucción 3D con UAVs podría mejorar aún más la precisión, especialmente en zonas sombreadas u ocluidas.

Desarrollando una solución automatizada con múltiples UAVs coordinados, Melnychenko et al. (2024) propusieron un sistema inteligente para la detección y el conteo de frutos en huertos de manzanos Fuji, basado en la integración de aprendizaje profundo y sistemas multivista coordinados mediante varios UAVs. El estudio, buscó superar las limitaciones de los métodos convencionales, tales como la baja cobertura espacial, la oclusión de frutos o la variabilidad de condiciones lumínicas, mediante un enfoque colaborativo y automatizado con drones sincronizados en tiempo real.

La adquisición de datos se realizó en un huerto comercial de 50 hectáreas, utilizando UAVs DJI Phantom 4 RTK equipados con cámaras RGB configuradas para distintos niveles de luz y altitudes (1,2–3,2 m). Los drones se desplegaron en misiones sincronizadas, generando secuencias de vídeo continuas mediante un sistema propio de fusión y alineamiento de flujos de vídeo multicanal. Las imágenes, capturadas bajo condiciones lumínicas variables (soleado, nublado, atardecer), fueron aumentadas con técnicas de procesado (CLAID.AI) y anotadas con Amazon SageMaker, dando lugar a un conjunto de entrenamiento de 16.224 imágenes.

El sistema empleó una arquitectura personalizada denominada YOLOv5-v1, desarrollada a partir de la versión original de YOLOv5 e incorporando mejoras clave orientadas a entornos agrícolas: se optimizaron las capas CSP con módulos SE para mejorar la extracción de características, se rediseñó el módulo *Focus* para reducir parámetros y acelerar el entrenamiento, y se modificaron los esquemas de fusión de mapas de características para integrar múltiples escalas espaciales. Además, se ajustaron los tamaños de *anchors* y *bounding boxes* para mejorar la detección de frutos en primer plano y evitar falsas detecciones en segundo plano. El entrenamiento del modelo se realizó durante 30 iteraciones completas sobre el conjunto con el optimizador Adam, alcanzando una convergencia estable sin signos de sobreajuste.

Los resultados demostraron una media de precisión (mAP) del 86,8 %, además de tasas de error relativamente bajas frente al conteo manual: 14,7 % de falsos positivos (FPR) y 18,3 % de falsos negativos (FNR). Estas métricas reflejan una alta fiabilidad del sistema, incluso en condiciones de nubosidad o iluminación cambiante. La arquitectura YOLOv5-v1 superó a modelos previos como Faster R-CNN, *Suppression Mask* R-CNN o YOLOv5-CS, especialmente en cuanto a adaptabilidad operativa en campo y eficiencia en el procesamiento en tiempo real.

El enfoque propuesto demostró ser una solución robusta y escalable para explotaciones comerciales, permitiendo el conteo automatizado de frutos con precisión elevada y bajo coste computacional. A pesar de ciertas limitaciones (como la sensibilidad a la oclusión o la dependencia de parámetros de confianza bien calibrados) el sistema mostró un alto potencial de transferencia a otras tareas agrícolas, consolidándose como una herramienta alineada con los principios de digitalización y automatización.

En línea con los avances en agricultura digital aplicada a cultivos frutales, Subeesh et al. (2024) desarrollaron una metodología integral para estimar el rendimiento en cítricos

mediante imágenes UAV y *deep learning*, integradas en una aplicación web llamada *DeepYield*. El objetivo fue automatizar la detección de frutos y la predicción del peso para facilitar la planificación agrícola en campo.

Se empleó un dron DJI Phantom 4 Pro V2.0 con cámara RGB (20 MP, CMOS 1/2.3") para capturar 800 imágenes de cítricos (*Citrus reticulata*) en dos campañas. Las imágenes se etiquetaron manualmente en dos clases (“*harvest*” y “*unripe*”) y se dividieron en entrenamiento (70 %), validación (15 %) y test (15 %). Se aplicaron técnicas de aumento de datos y se emplearon diferentes resoluciones según el modelo.

Se compararon cinco arquitecturas: SSD, Faster R-CNN, YOLOv4, YOLOv5 y YOLOv7. El modelo YOLOv7 demostró el mejor rendimiento (mAP = 86.48 %; *precision* = 88.54 %; *recall* = 83.66 %; F1-Scores = 86.03 %) con un entrenamiento de 200 iteraciones sobre el conjunto y *batch size* 32. Su estructura basada en E-ELAN y eficiencia en el procesamiento lo hizo idóneo para su integración en la aplicación *DeepYield*. Además, se entrenó una versión segmentadora para estimar el tamaño del fruto a partir del eje mayor, correlacionado con el peso real.

YOLOv7 superó al resto de arquitecturas tanto en precisión como en velocidad de inferencia (15 ms). Se observó un error promedio de 6 frutos respecto al conteo manual para frutos maduros. El modelo mostró dificultades con frutos inmaduros debido a la similitud cromática con el follaje. Aun así, al aplicar regresión lineal para ajustar los conteos y pesos, la estimación final presentó un error promedio de solo ± 1.01 kg por árbol para frutos listos, frente a ± 1.92 kg si se incluyen los inmaduros.

DeepYield se diseñó con Django y es accesible desde navegadores web. Ofrece una solución práctica, precisa y ligera, capaz de integrarse en flujos de trabajo agrícolas reales con dispositivos de bajo coste. Se validó su utilidad cuando se aportan imágenes UAV del 20 % de los árboles del huerto, obteniendo errores de predicción <11 %. La herramienta representa un avance notable en la planificación de cosechas y gestión de la producción frutal.

Los modelos de detección de una etapa presentan un notable potencial para la estimación de rendimiento agrícola, gracias a su velocidad de inferencia, adaptabilidad a distintas condiciones de vuelo y una precisión operativa aceptable. Si bien su sensibilidad frente a problemas como la oclusión o la variabilidad lumínica continúa siendo un desafío, los avances recientes en diseño arquitectónico y ajuste de hiperparámetros han mejorado

significativamente su desempeño en campo. En este sentido, YOLOv4 ha demostrado ser especialmente robusto en escenarios con alta densidad de objetos, mientras que YOLOv7 se consolida como una opción versátil al integrar procesos de segmentación y estimación directa, lo que refuerza su aplicabilidad en contextos agrícolas reales.

4.3.2.2. Modelos de detección: comparación entre una y dos etapas.

Dentro de esta categoría se integraron estudios que pusieron en contraste modelos de una etapa (YOLO o FCOS) frente a arquitecturas de dos etapas (Faster R-CNN), aplicados a la estimación de rendimiento en entornos agrícolas. El propósito de estas comparativas fue determinar qué enfoque resultaba más eficaz en función de la precisión, la velocidad de inferencia y la capacidad de adaptación a distintos tipos de cultivo y condiciones de vuelo. Los resultados permitieron delimitar fortalezas relativas según el escenario de aplicación.

Contribuyendo al desarrollo de sistemas no destructivos para el fenotipado de especies vegetales, Hong-Yu et al. (2023) evaluaron la eficacia de tres arquitecturas de *deep learning* (FCOS, YOLOv5 y Faster R-CNN) para el conteo automatizado de plantas de ramio (*Boehmeria nivea*) en condiciones de alta densidad, empleando imágenes RGB captadas por drones. El objetivo principal fue analizar el impacto de la resolución espacial (GSD, por sus siglas en inglés), la diversidad de muestras y el aumento de datos (DA) en la precisión del conteo, contribuyendo al desarrollo de sistemas no destructivos de fenotipado vegetal.

La adquisición de datos se realizó en dos campos experimentales en Changsha (China), entre 2019 y 2020, mediante un dron Inspire 2 equipado con cámara RGB de alta resolución. Las imágenes se capturaron a 15 m de altura y velocidad de vuelo constante, bajo condiciones óptimas de iluminación. En total, se trabajó con más de 700 imágenes distribuidas en tres conjuntos: V1 (comparación de algoritmos), V2 (evaluación temporal interanual) y V3 (análisis del efecto del GSD). Se empleó anotación manual mediante LabelImg y técnicas de *Mosaic* para DA, así como filtros gaussianos para simular resoluciones media y baja.

Desde el punto de vista técnico, los modelos se entrenaron en un entorno con GPU RTX5000. FCOS (*Fully Convolutional One-Stage Object Detection*), respaldado por ResNet101 y FPN, mostró un rendimiento especialmente destacado en detección de objetos pequeños y solapados. Faster R-CNN (con ResNet101) y YOLOv5 (con

CSPDarknet) también fueron implementados para la comparación. Las métricas utilizadas incluyeron *Precision*, *Recall*, *RMSE* y R^2 , permitiendo evaluar tanto la precisión como la robustez general de los modelos.

Los resultados indicaron que FCOS alcanzó la mejor combinación de precisión y estabilidad en contextos de alta densidad ($Precision = 0,819$; $Recall = 0,892$; $RMSE = 0,089$), superando ligeramente a Faster R-CNN y YOLOv5, cuyos errores fueron más frecuentes debido a múltiples detecciones sobre una misma planta. Además, el uso de DA mejoró sustancialmente el rendimiento del modelo (R^2 pasó de 0,03 a 0,728), evidenciando su utilidad en contextos de escasez de muestras. En cuanto a la resolución espacial, el modelo fue sensible a variaciones del GSD: el mejor rendimiento se obtuvo con imágenes de resolución media ($Precision = 0,822$; $Recall = 0,894$), lo que sugiere que una resolución moderada puede ser más eficaz que la alta, al facilitar el enfoque sobre patrones relevantes.

La elaborada investigación, demostró que el uso de UAV junto con modelos de *deep learning* como FCOS es altamente efectivo para el monitoreo de cultivos densos como el ramio. Este enfoque no solo redujo la necesidad de conteo manual, sino que además ofreció resultados precisos con menor coste computacional.

Los estudios comparativos entre modelos de una y dos etapas ofrecen una perspectiva matizada sobre su aplicabilidad en tareas de estimación de rendimiento agrícola. Las arquitecturas de una etapa destacan por su eficiencia computacional y velocidad de procesamiento, mientras que las de dos etapas aportan una mayor precisión y capacidad para detectar objetos complejos, aunque con un coste computacional más elevado. Tal como se observó en líneas temáticas previas, la elección del enfoque más adecuado depende del contexto agronómico, la resolución requerida y los recursos técnicos disponibles. En este escenario, el modelo FCOS emerge como una alternativa intermedia prometedora, al combinar precisión y eficiencia en aplicaciones de monitoreo extensivo, especialmente en cultivos de alta densidad.

4.3.2.3. Modelos combinados.

La presente categoría incluyó estudios que extendieron los modelos de detección mediante su combinación con otras tareas, como la regresión de variables continuas (por ejemplo, peso o volumen) o el seguimiento temporal de objetos (*tracking*). Estas propuestas ofrecieron una visión más integral del rendimiento, al vincular la

identificación de estructuras vegetales con métricas derivadas o dinámicas que enriquecieron el análisis agronómico. Su aplicación supuso un avance hacia estimaciones más realistas y contextualizadas en condiciones de campo.

Aportando una mejora sustancial a las metodologías convencionales de estimación frutal, Kalantar et al. (2020) propusieron una solución innovadora para la estimación automatizada del rendimiento individual y total de melones mediante imágenes RGB captadas por un dron Phantom 4 Pro a 15 m de altitud. El objetivo fue superar las limitaciones de trabajos previos centrados únicamente en el conteo de frutos, introduciendo un sistema de visión artificial basado en *deep learning* capaz de predecir tanto el número como el peso de cada fruto en campo, incluso bajo condiciones heterogéneas de iluminación, oclusión y distribución.

La metodología se estructuró en tres fases: detección, extracción de características geométricas y estimación de peso. La adquisición de datos incluyó más de 6.000 melones anotados manualmente en ocho imágenes obtenidas durante tres campañas. El sensor embarcado fue una cámara RGB DJI FC6310 con resolución de 5472×3648 píxeles, y se utilizó etiquetado manual con el software LabelImg. El conjunto se dividió en 800 subimágenes para entrenar y validar el modelo, que fue ampliado mediante técnicas de aumentación de datos. Se aplicó transferencia de aprendizaje sobre RetinaNet con backbone ResNet50, adaptando los *anchor boxes* para mejorar la detección de objetos pequeños. Se implementó una segmentación precisa por contorno activo (Chan–Vese) seguida de un ajuste elíptico mediante PCA para cada fruto, obteniendo parámetros geométricos claves (ejes mayor/menor).

La arquitectura de detección RetinaNet se seleccionó tras comparar con alternativas como Faster R-CNN, SSD y YOLO, superándolas por su precisión en objetos pequeños y escasamente distribuidos. Esta se entrenó con *focal loss* y alcanzó una precisión promedio (AP) de 0.92 y F1-score superior al 0.9, validado en distintos entornos agrícolas. La estimación de peso se realizó mediante un modelo de regresión lineal multivariable basado en las dimensiones elipsoidales extraídas, obteniendo un R^2 ajustado de 0.94 y un error medio absoluto (MAPE) del 16 %, reduciéndose al 12 % en frutos con buena segmentación. La estimación del rendimiento total presentó solo un 3 % de subestimación respecto al peso real medido, lo que refuerza su aplicabilidad.

Este estudio representó un avance relevante al integrar detección y regresión en un único *pipeline* de bajo coste y alta fiabilidad, abriendo la posibilidad de replicar esta

metodología en cultivos frutales similares. Sin embargo, los autores reconocieron limitaciones vinculadas a la sensibilidad de la segmentación elíptica, la precisión del GSD y la necesidad de etiquetado manual, proponiendo como mejora futura el uso de segmentación basada en U-Net o Mask R-CNN para una mayor precisión geométrica a nivel de píxel.

En respuesta a la necesidad de herramientas precisas para el monitoreo de producción en campo, Li et al. (2023a) diseñaron un enfoque dual de estimación del rendimiento en longan mediante imágenes RGB captadas por UAV y modelos de *deep learning*, en respuesta a la ineficiencia del conteo manual en huertos comerciales. El objetivo principal fue desarrollar un sistema automatizado capaz de detectar y cuantificar frutos en condiciones reales de campo, aplicando arquitecturas YOLO mejoradas combinadas con algoritmos de seguimiento para abordar la complejidad visual de los entornos naturales.

La adquisición de datos se realizó utilizando un dron DJI Mavic 2 Pro equipado con una cámara RGB de alta resolución (12 MP, 1280×720 píxeles), durante los meses de julio de 2021 y 2022. Se recopilaron un total de 1.100 imágenes y 16 videos de copa completa en distintas condiciones de luz (mañana/tarde, días soleados y nublados), incluyendo dos variedades de longan (Shixia y Chuliang). Posteriormente, se construyeron dos conjuntos de datos etiquetados manualmente: uno para detección de racimos (*Cluster_Fruit*) y otro para frutos individuales (*Single_Fruit*), divididos en entrenamiento, validación y test en proporción 8:1:1. Además, se aplicaron técnicas de aumento de datos y normalización espacial para mejorar la robustez del aprendizaje.

El sistema propuesto se estructuró en dos modelos complementarios: CF-YD y SF-YD. El primero integra YOLOv5s optimizado y DeepSort para la detección y seguimiento de racimos. Su arquitectura incluye un backbone con módulos Focus, CBH y CSP, así como una red FPN+PAN para detección multiescala, combinando eficiencia computacional y precisión. El modelo SF-YD, orientado a la detección de frutos individuales dentro de los racimos, emplea YOLOv7, que incorporó la red E-ELAN y una cabeza con PAFPN para mejorar la fusión de características. Ambos modelos se entrenaron sobre PyTorch, aplicando técnicas como el ajuste de anclas mediante K-means y algoritmos genéticos, y fueron validados con métricas estándar.

Los resultados obtenidos reflejaron una alta precisión de los modelos propuestos. El modelo CF-YD, enfocado en la detección de racimos de frutos, alcanzó valores de *precision*, *recall* y *F1-score* del 83,5 %, 85,7 % y 0,85 respectivamente, mientras que el

modelo SF-YD, orientado a la identificación de frutos individuales dentro de cada racimo, obtuvo métricas aún más destacadas, con un *F1-score* de 0,93 y un AP del 97,1 %. Ambos modelos demostraron además un excelente rendimiento en tareas de seguimiento multiobjeto, con valores de MOTA y MOTP superiores al 95 % y 92 %, respectivamente, así como una alta eficiencia operativa, con tiempos de inferencia inferiores a 20 ms por imagen. La validación frente al conteo manual reveló errores relativos mínimos: 2,66 % para los racimos y 2,99 % para los frutos individuales, lo que confirma la robustez del sistema para la estimación automatizada del rendimiento en condiciones reales de campo.

Por ende, el enfoque CF-YD junto con SF-YD constituyó una solución eficaz y precisa para la estimación automatizada de cosechas en frutales complejos, especialmente útil para optimizar la planificación de la recolección y mejorar la eficiencia operativa en la agricultura de precisión. No obstante, los autores reconocieron limitaciones asociadas a la visibilidad parcial del dosel desde el ángulo cenital del UAV, proponiendo para futuras investigaciones la integración de rutas de vuelo automatizadas y análisis en tiempo real.

Los modelos combinados constituyen un avance significativo en la integración de tareas de visión artificial, al permitir estimaciones más precisas y funcionales del rendimiento agrícola. La incorporación de componentes de regresión o seguimiento favorece el cálculo de variables clave como el peso, la cantidad de frutos o las tasas de desarrollo, lo que amplía su utilidad en escenarios de gestión agronómica avanzada. Si bien la complejidad de estas arquitecturas puede dificultar su adopción en explotaciones con recursos técnicos limitados, los resultados obtenidos en condiciones reales evidencian su potencial para entornos que requieren un monitoreo detallado o una planificación logística más precisa.

4.3.2.4. Modelos con mecanismos de atención.

Como cierre de esta línea temática, se incorporaron modelos que integraron mecanismos de atención (espacial, espectral o multicanal) con el fin de optimizar la detección en escenarios visualmente complejos. Estos módulos permitieron a las redes focalizarse en las regiones más informativas de la imagen, mejorando la precisión en situaciones problemáticas como la oclusión parcial, la superposición de clases o las variaciones en las condiciones de iluminación. Su implementación reforzó la robustez del sistema ante la heterogeneidad típica de los entornos agrícolas reales.

En el marco de la adopción de redes distintas a las convolucionales en visión artificial aplicada a la agricultura, Zhu et al. (2022) propusieron un enfoque avanzado para la detección de espigas de trigo mediante UAV e imágenes de luz visible, incorporando arquitecturas basadas en Transformer para superar las limitaciones de los modelos CNN tradicionales. El objetivo principal fue mejorar la precisión en el conteo de espigas en condiciones reales de campo, caracterizadas por alta densidad, oclusión y variabilidad fenotípica.

La metodología se basó en vuelos realizados con el UAV DJI Matrice M300 RTK, equipado con el sensor Zenmuse H20T (RGB), sobre parcelas experimentales con 305 variedades de trigo. Las imágenes, capturadas a 10 m de altura con resolución de 0,06 cm/píxeles, se procesaron para construir el nuevo conjunto UWHD, compuesto por 550 imágenes anotadas manualmente y aumentadas hasta un total de 2.750. Además, se aplicaron tres modelos sobre *datasets* adicionales (GWHD, SPIKE y ACID) para validar su generalización.

El enfoque técnico se centró en el diseño de tres modelos de detección con arquitectura Transformer como backbone: FR-Transformer (basado en Faster R-CNN), R-Transformer (RetinaNet) y Y-Transformer (YOLOv3). Estas arquitecturas integraron bloques Swin Transformer con mecanismos W-MSA y SW-MSA, permitiendo el análisis global de la imagen sin perder precisión espacial. El modelo se entrenó con AdamW, empleando imágenes redimensionadas a 448×448 píxeles y un total de 100 ciclos de entrenamiento.

Los resultados mostraron un rendimiento sobresaliente del modelo FR-Transformer, alcanzando un AP50 del 88,3 %, superando a enfoques previos como Cascade R-CNN o DETR. Su rendimiento en AP75 (38,5 %) y en objetos grandes (APL = 54,1 %) confirmó su robustez ante condiciones complejas como solapamiento o variaciones de iluminación. Aunque la precisión para objetos pequeños (APS = 6,4 %) fue baja, el modelo demostró mayor eficiencia computacional (Params y FLOPs reducidos) frente a otras arquitecturas con similares niveles de exactitud. Se identificaron errores derivados de falsos positivos en suelos iluminados o detecciones sin anotación manual previa, que redujeron marginalmente las métricas.

La propuesta fue exitosa, ya que se consiguió integrar eficazmente UAVs, visión artificial y *Transformers*, ofreciendo una solución viable para el fenotipado de trigo en campo. Su capacidad para combinar precisión, rapidez y eficiencia computacional la convierte en

una alternativa prometedora para la estimación de rendimiento y selección varietal, aunque futuras investigaciones podrían centrarse en mejorar la detección de espigas pequeñas y minimizar las confusiones espectrales con el fondo del suelo.

Contribuyendo al avance del fenotipado automatizado en programas de mejora genética, Li et al. (2023b) propusieron un método automatizado para el recuento de inflorescencias de *Brassica napus L.* mediante imágenes RGB capturadas por UAV y *deep learning*, con el objetivo de mejorar el fenotipado en programas de mejora varietal. El recuento se abordó como un problema de detección de objetos, desarrollando una versión optimizada del modelo YOLOv5 con un módulo de atención CBAM.

El estudio se llevó a cabo en la región del Yangtsé (China), donde se tomaron imágenes RGB de parcelas experimentales durante el periodo de floración (2021–2022), usando un dron DJI Phantom 4 Pro V2.0 (20 MP, 5.472×3.648 píxeles). Las imágenes se ortorrectificaron con Agisoft PhotoScan y se construyó un *dataset* propio, RIB, compuesto por 165 imágenes de parcelas con más de 60.000 inflorescencias anotadas manualmente. Se evaluó la robustez del modelo en condiciones de diferente densidad floral y en años distintos.

Se empleó la arquitectura YOLOv5x mejorada con el módulo CBAM, el cual integra atención espacial y por canal. Esta combinación permitió potenciar la detección en imágenes con solapamiento, oclusión y alta densidad floral. La estructura incluyó CSPDarknet53 como backbone, FPN y PAN para fusión multiescala, y una capa de detección con función de pérdida GIoU. El CBAM se integró entre los bloques C3 y CBS, lo que mejoró la extracción de características sin penalizar el tiempo de inferencia.

El modelo YOLOv5x-CBAM alcanzó una *precision* del 90,1 %, un F1-*score* del 88,7 % y un mAP del 93,6 %, superando a arquitecturas clásicas como Faster R-CNN, YOLOv4 y TasselNetV2+. En términos de conteo, se obtuvo un coeficiente de determinación $R^2 = 0,966$ y un RMSE de 52,1 respecto al conteo manual. La mejora fue especialmente notable frente a métodos que no conservaban la información espacial (ej. TasselNetV2+) o que sufrían pérdidas por solapamientos (ej. YOLOv4). La incorporación de CBAM fue clave para reducir errores de omisión en zonas densas, sin comprometer el rendimiento computacional.

El enfoque propuesto permitió detectar y cuantificar inflorescencias de colza de forma precisa, eficiente y con alta transferibilidad a condiciones reales de campo. Su capacidad

para detectar correlaciones significativas entre el número de inflorescencias y el rendimiento por parcela lo posicionó como una herramienta prometedora para la predicción de rendimiento y la selección varietal. Sin embargo, los autores reconocieron limitaciones en escenarios de alta densidad y propusieron como mejora futura el aumento de la frecuencia temporal de muestreo y la incorporación de imágenes multispectrales para un análisis fenológico más completo.

Frente a las limitaciones de visibilidad y segmentación en huertos frutales densos, Sekharamanthy et al. (2024) desarrollaron un sistema integral basado en *deep learning* para la detección, estimación de profundidad y seguimiento de manzanas en huertos, utilizando imágenes captadas por drones. El objetivo fue superar los retos asociados a la oclusión, la variabilidad lumínica y el solapamiento de frutos en entornos reales, integrando para ello la arquitectura YOLOv7 mejorada con un mecanismo de atención multi-cabeza (MAM) y el algoritmo de seguimiento ByteTrack.

La adquisición de datos se realizó en huertos del Valle di Non (Trento, Italia), mediante un dron DJI Mavic Mini 3 equipado con cámaras estéreo ZED 2iW. Las imágenes (con resolución 2688×1520 píxeles) fueron captadas a distancias de 30–60 cm sin iluminación artificial, y anotadas manualmente con etiquetas de profundidad para manzanas maduras. El *dataset* se dividió en subconjuntos de entrenamiento (60%), validación (10%) y test (30%), y se aplicó aumentación de datos para robustecer el aprendizaje frente a variaciones ambientales.

La arquitectura base fue YOLOv7, optimizada con mecanismos de reparametrización, asignación dinámica de etiquetas y estructura en pirámide para detección multiescala. Sobre esta base se integró un mecanismo de atención multi-cabeza (MAM) que mejora la capacidad del modelo para capturar relaciones espaciales y estimar la profundidad mediante imágenes estéreo. La atención se aplicó sobre mapas de características, codificados con posiciones relativas, y se entrenó con una función de pérdida compuesta (clasificación y regresión), incluyendo un término de pérdida L1 para la predicción de *bounding boxes* y profundidad. La etapa final incorporó ByteTrack, un algoritmo de seguimiento que conserva detecciones de baja confianza y mejora la continuidad temporal, mitigando errores como el cambio de ID.

El modelo YOLOv7 junto con MAM alcanzó un *F1-Score* de 0,95 (*Precision*: 0,92; *Recall*: 0,96) y un error medio absoluto en el conteo (MAPE) del 2,7%, superando a variantes previas como YOLOv5, Faster R-CNN y AlexNet + RCNN. En condiciones

lumínicas adversas, mantuvo una robustez destacable (*F1-Score*: 0,89). Además, el sistema demostró una capacidad eficaz para estimar la distancia relativa fruto-dron, lo que habilita aplicaciones como estimación de rendimiento y planificación de cosechas. La incorporación de ByteTrack evitó recuentos duplicados al seguir individualmente cada manzana a lo largo del vídeo.

En conjunto, la iniciativa de combinar YOLOv7 junto con MAM constituyó una solución eficaz y escalable para la automatización de tareas en huertos frutales, como la cosecha asistida y la monitorización en tiempo real. Su alta precisión, combinada con la estimación de profundidad y seguimiento individualizado, la hicieron viable para aplicaciones comerciales en agricultura de precisión. Sin embargo, los autores reconocieron la necesidad de seguir optimizando el modelo para condiciones meteorológicas extremas, así como explorar su transferencia a otros cultivos y tareas agronómicas.

Avanzando hacia la digitalización del rendimiento agrícola en cultivos no convencionales, Wang et al. (2024) tuvieron como objetivo diseñar un sistema automático para la estimación de rendimiento de té mediante imágenes RGB obtenidas por UAV y técnicas avanzadas de *deep learning*. Concretamente, se centró en la detección de yemas de té mediante modelos de detección basados en YOLOv5. La investigación se llevó a cabo en diferentes campos de Zhejiang (China), evaluando el rendimiento antes y después de la recolección en distintas estaciones del año.

Para la adquisición de datos, se utilizó un dron DJI MAVIC AIR 2 con cámara RGB, capturando un total de 5.899 imágenes, en las que se etiquetaron manualmente 29.958 yemas de té. Estas imágenes fueron segmentadas en chips de 800×600 píxeles con un solapamiento del 25 %, y posteriormente se aplicó un algoritmo de supresión no máxima (NMS) doble con un umbral ajustado a 0.20 para reducir detecciones duplicadas. Se emplearon diversas técnicas de aumento de datos (mosaico, giro y escalado) para robustecer el entrenamiento del modelo.

En cuanto al modelo de *deep learning*, se utilizó YOLOv5, un detector de una etapa basado en la arquitectura CSPDarknet53. Esta estructura incluye un módulo Focus para reducción de resolución, bloques ResBlock con capas convolucionales y activación SiLU, así como mecanismos FPN y PAN para fusión multiescala de características. Además, se integraron bloques de atención Squeeze-and-Excitation (SE) en distintas configuraciones (1SE, 2SE y BSE), siendo el modelo 2SE-YOLOv5 el seleccionado por su mejor

equilibrio entre precisión (84,96 %) y estabilidad, alcanzando un mAP del 85,64 % y un *F1-score* del 78 %.

El modelo fue aplicado a campos reales para estimar el rendimiento potencial. En primavera, se alcanzó una media de 1.359 yemas/m² y un rendimiento estimado de 1.223 kg/ha, en concordancia con datos previos de la literatura. En verano, el rendimiento fue menor debido a la obstrucción visual causada por hojas maduras. No obstante, el sistema logró detectar de forma robusta las yemas en distintos tipos de cultivares y escenarios, estimando correctamente la producción antes y después de la recolección.

Es por ello, que el enfoque 2SE-YOLOv5 demostró ser eficaz para la estimación no destructiva del rendimiento del té a partir de imágenes RGB aéreas, ofreciendo una solución viable y escalable para la gestión inteligente de cultivos perennes. Si bien la ocultación de yemas bajo el dosel limitó parcialmente la precisión, los autores sugirieron que la integración de visión estereoscópica y vuelos multiángulo podría mitigar esta limitación. Además, esta experimentación abrió la puerta a futuras aplicaciones del modelo en otras especies agrícolas de hoja perenne con morfologías similares.

En el contexto de la estimación temprana del rendimiento en cereales, Chen et al. (2024) plantaron una solución innovadora para la detección y el conteo automatizado de espigas de maíz (*Zea mays* L.) durante el periodo de floración, empleando imágenes RGB obtenidas mediante drones. La investigación se enmarcó en un contexto donde la detección temprana de estructuras reproductivas es clave para estimar el rendimiento y planificar estrategias de manejo agronómico. Para ello, los autores propusieron una red neuronal mejorada, denominada RESAM-YOLOv8n, que incorporó un módulo de atención espacial residual (RESAM) sobre la arquitectura base YOLOv8n.

Desde el punto de vista metodológico, se empleó un UAV DJI Mavic 2 equipado con cámara RGB, realizando vuelos manuales a 10 m de altura durante el periodo de floración (julio-agosto de 2022). La resolución espacial alcanzada fue superior a 1 cm/píxel. Las imágenes se capturaron en distintos horarios (mañana, mediodía y tarde) y en 20 parcelas experimentales con distintos tratamientos de fertilización. El conjunto de datos final (FMTS *dataset*) consistió en 985 imágenes etiquetadas manualmente y divididas en subconjuntos de entrenamiento (689), validación (99) y prueba (197), tras un proceso de recorte a 1024×1024 píxeles y filtrado semántico para garantizar la presencia de espigas.

La arquitectura propuesta, RESAM-YOLOv8n, mantuvo la estructura general de YOLOv8n (Backbone, Neck y Head), pero introdujo mejoras críticas: el aumento del tamaño de entrada a 1024×1024 y la incorporación del módulo RESAM en la etapa Neck. RESAM combinó un módulo de atención espacial (SAM) con una conexión residual tipo ResNet, lo que permitió focalizar la atención en zonas relevantes del campo visual y reducir el sobreajuste. El entrenamiento se realizó con Pytorch durante 300 ciclos de entrenamiento, usando imágenes aumentadas con técnica Mosaic y con pesos preentrenados en COCO como punto de partida.

Los resultados evidenciaron una mejora sustancial del modelo respecto a las versiones base y a otras arquitecturas comparadas (Faster R-CNN, SSD, RetinaNet, YOLOv3, YOLOv5s y YOLOv8n). RESAM-YOLOv8n alcanzó un $mAP@0.5$ de 95,74 %, $mAP@0.75$ de 66,70 %, *precision* del 95,59 %, *recall* del 89,28 % y *F1-score* de 92 %. En tareas de conteo, el modelo obtuvo un coeficiente de determinación $R^2 = 0,976$ y un RMSE de solo 1,56 espigas, lo que indicó una excelente correlación con el conteo manual. La red demostró especial robustez en escenarios complejos como iluminación intensa, solapamiento de espigas y detección de objetos pequeños, superando significativamente a los modelos base.

Desde una perspectiva aplicada, el modelo RESAM-YOLOv8n mostró gran potencial para la automatización de tareas fenológicas en agricultura de precisión. A pesar de que su velocidad de procesamiento fue ligeramente inferior a YOLOv8n (86,5 FPS frente a 96,6), el impacto en detección de campo fue aceptable dada la mejora en precisión. Como limitación, los autores reconocieron una menor capacidad para detectar espigas con coloraciones muy similares al fondo, así como la falta de validación cruzada con otros *datasets* públicos. En conjunto, el enfoque RESAM-YOLOv8n representó un avance sólido en la detección automática de estructuras reproductivas en maíz, con implicaciones directas en estimación de rendimiento y gestión agronómica de cultivos extensivos.

En escenarios donde la complejidad visual limita el desempeño de las arquitecturas convencionales, los modelos con mecanismos de atención ofrecen una respuesta técnica notable. La incorporación de atención espacial, espectral o multicanal permite focalizar el análisis en las regiones más relevantes de la imagen, mejorando la detección en condiciones de oclusión, elevada densidad o iluminación cambiante. Según evidencian los estudios revisados, estos enfoques no solo incrementan la precisión cuantitativa, sino que también aportan mayor estabilidad operativa en campo, lo que los posiciona como

alternativas especialmente adecuadas para tareas de monitoreo exigente en agricultura de precisión.

4.3.3. Otras aplicaciones agronómicas.

4.3.3.1. Modelos de optimización operativa.

Este nuevo bloque temático dio comienzo con un estudio centrado en la optimización operativa del vuelo, que, si bien no abordó directamente la estimación de variables agronómicas, representó un avance clave en la mejora de la autonomía y eficiencia de los UAVs en agricultura de precisión. En concreto, el trabajo se orientó a la detección y evitación de obstáculos, una funcionalidad que no solo incrementó la seguridad del vuelo, sino que también contribuyó a preservar la calidad de los datos captados, especialmente en entornos agrícolas no estructurados y de alta variabilidad.

Con el objetivo de mejorar la autonomía operativa de los UAVs en entornos agrícolas complejos, Wang et al. (2020) desarrollaron una solución integral que combinó aprendizaje profundo (YOLOv3) y cámaras RGB-D (Intel RealSense D435) para dotar a los UAVs de percepción ambiental y capacidad autónoma de evitación de colisiones. La propuesta permitió detectar no solo la presencia de obstáculos, sino también estimar su categoría, perfil y posición tridimensional. El sistema se basó en una arquitectura CNN entrenada con un conjunto de datos personalizado (3.700 muestras), alcanzando una precisión media de detección del 75,4% y un tiempo medio de procesamiento por imagen de 53,33 ms.

Mediante la fusión de información RGB y de profundidad, el modelo generó trayectorias óptimas de evitación en función de las características del obstáculo, integrando estos resultados en un sistema de control de tareas (TCS) completamente autónomo. Las pruebas de campo, que simularon distintos escenarios con árboles reales, confirmaron que el sistema fue capaz de planificar y ejecutar maniobras de esquivas (cambio de altitud, giro lateral o avance vertical) sin intervención humana.

Entre sus limitaciones, destacó que la precisión en la detección de perfiles y ubicación se vio afectada por la distancia entre el dron y el obstáculo, con mejores resultados en el rango de 4,5 a 8,0 metros. Aun así, el estudio demostró la viabilidad técnica de integrar percepción tridimensional y detección por aprendizaje profundo en UAVs para la navegación segura en entornos agrícolas reales.

Este estudio constituye un avance estratégico hacia la autonomía operativa de los UAVs en contextos agrícolas, al abordar una de las principales limitaciones actuales: la ausencia de sistemas fiables para la detección y esquiwa de obstáculos. Si bien su aplicación se centró en maniobras evasivas, su integración en flujos de trabajo de teledetección aérea permitiría optimizar tanto la seguridad de vuelo como la calidad de los datos recolectados. Este tipo de soluciones resulta especialmente relevante en parcelas con vegetación dispersa, árboles aislados o infraestructuras verticales no cartografiadas, donde la navegación automática presenta mayores desafíos.

4.3.3.2. Modelos de monitoreo fenológico.

En esta sección se incluyeron estudios orientados al monitoreo fenológico del cultivo de maíz mediante imágenes obtenidas con UAVs y técnicas de *deep learning*, con el propósito de optimizar la toma de decisiones en sistemas agrícolas de alta precisión. Los trabajos analizados abordaron tanto la cuantificación de plántulas en estadios vegetativos como la identificación del estado reproductivo a partir de la detección de espigas, ofreciendo soluciones aplicables a la gestión de la densidad de siembra y a la evaluación del desespigado en parcelas de híbridos.

Ante la necesidad de una detección precisa en etapas iniciales del desarrollo del cultivo, Jia et al. (2024) se centraron en la detección y cuantificación de plántulas de maíz en estadios tempranos (V2 a V6) mediante UAVs equipados con cámaras RGB y distintos modelos de detección de objetos basados en *deep learning*. El objetivo fue comparar el rendimiento de nueve arquitecturas, entre ellas YOLOv8n, YOLOv5n, Deformable DETR y Faster R-CNN, bajo diferentes condiciones de densidad de siembra, altura de vuelo y estadio fenológico. Las imágenes fueron adquiridas en campos reales bajo un diseño experimental riguroso, y se construyó un *dataset* de 16.200 imágenes tras un proceso intensivo de aumentación. Los resultados mostraron que YOLOv8n alcanzó la mayor precisión ($AP50 = 0.976$) y robustez en campo, con rRMSE inferior al 1.64% incluso en parcelas de alto rendimiento. No obstante, se observó una disminución significativa en la precisión conforme aumentaba la densidad de plantas y el estadio fenológico, atribuida al solapamiento foliar y a la mayor complejidad visual. Se concluyó que los estadios V2–V3 son óptimos para la detección, y que la altura de vuelo se debía mantener baja para preservar la resolución espacial. El estudio destacó la viabilidad de

implementar sistemas UAV-DL en operaciones reales, aunque remarca la necesidad de ajustar los modelos a las condiciones agronómicas específicas.

Por su parte, Du et al. (2024) propusieron un enfoque avanzado para la identificación del estado de las espigas en maíz durante el proceso de desespigado manual en campos de hibridación. A partir de un conjunto de imágenes RGB de altísima resolución (8.192×5.460 píxeles) captadas con un DJI Matrice 300 RTK, se construyó un *dataset* etiquetado en tres categorías fenológicas: *Tassel-N* (sin espiga), *Tassel-S* (espiga en formación) y *Tassel-L* (espiga emergida). Mediante un proceso iterativo de optimización de los tamaños de las cajas de anotación y técnicas de augmentación, se entrenaron modelos como YOLOv5, CenterNet, RTMDet y Faster R-CNN, alcanzando una precisión promedio de hasta el 94,5 % con NSL-B y del 98 % al aplicar estrategias de división por bloques de imagen. La arquitectura RTMDet destacó por su equilibrio entre precisión y eficiencia computacional. Entre las limitaciones se señalaron los desafíos en la anotación de espigas incipientes (*Tassel-N/S*) y los efectos de distorsión por perspectiva en imágenes UAV de gran formato. El estudio evidenció la aplicabilidad real del modelo para evaluar la eficacia del desespigado, constituyéndose en una herramienta prometedora para la mejora de los procesos de selección genética en programas de mejora de maíz.

Los trabajos analizados demuestran que la integración de UAVs con modelos de *deep learning* permite captar información fenológica clave a distintas escalas temporales y espaciales. Mientras uno de los estudios se centró en el monitoreo temprano del establecimiento del cultivo y la densidad efectiva de plántulas, el otro abordó la evaluación de eventos reproductivos críticos en programas de hibridación, como el estado de las espigas. Ambos enfoques refuerzan el potencial de la agricultura de precisión basada en visión artificial, al ofrecer soluciones robustas, replicables y adaptables a condiciones agronómicas reales. No obstante, también ponen de manifiesto la necesidad de seguir desarrollando estrategias que mejoren la detección en escenarios con alta complejidad visual o estructural.

4.3.3.3. Modelos de automatización.

Por su parte, se revisaron estudios que propusieron aplicaciones innovadoras de UAVs en tareas tradicionalmente realizadas por maquinaria terrestre o mediante intervención humana directa, como la aplicación de productos fitosanitarios, la recolección de frutos o

la polinización asistida. Estas investigaciones expandieron el potencial operativo de los UAVs, orientándolos hacia una agricultura más automatizada, precisa y capaz de adaptarse a entornos complejos o de difícil acceso.

Uno de los avances más representativos en esta línea es el desarrollado por Sassu et al. (2023), quienes propusieron un sistema de detección automática de plantas de alcachofa orientado a futuras aplicaciones de pulverización sitio-específica mediante UAVs. El estudio integró dos enfoques de visión artificial una *Feature Pyramid Network* personalizada y la arquitectura YOLOv5 entrenados sobre ortomosaicos RGB obtenidos en siete fechas a lo largo del ciclo del cultivo. Ambas redes permitieron localizar con precisión cada planta del campo, alcanzando niveles elevados de *precision*, *recall* y cobertura foliar (SR y SP), siendo YOLOv5 la que mostró un rendimiento más robusto y consistente a lo largo del periodo vegetativo. Además, los autores desarrollaron un procedimiento multitemporal capaz de reconstruir la evolución espacial y morfológica de cada individuo, proporcionando información clave para planificar rutas óptimas de aplicación y reducir el desperdicio de agroquímicos en operaciones sitio-específicas. Como limitaciones, la detección se vio afectada por el solape entre plantas en fases avanzadas y por variaciones de iluminación estacional. Aun así, el estudio evidenció el potencial de los UAVs combinados con deep learning para avanzar hacia sistemas de gestión más precisos y eficientes en cultivos hortícolas.

En el ámbito de la recolección, Li et al. (2022b) propusieron un sistema integrado para la cosecha de longan mediante UAVs, combinando una versión mejorada de YOLOv5s con DeepLabv3+ y fusión de imágenes RGB-D. El sistema permitió la detección de racimos y ramas principales, seguida de la segmentación semántica y la localización tridimensional de los puntos de corte para la recolección. La precisión alcanzada fue del 85,50% en detección y del 94,52% en segmentación, con una latencia de procesamiento de 0,58 segundos por imagen. La principal fortaleza radicó en su capacidad para operar en entornos naturales complejos, como huertos montañosos con árboles de gran altura. Como limitación, los autores indicaron que el sistema dependía de cámaras RGB-D de alto coste y de condiciones de visibilidad adecuadas. Esta investigación consolidó un avance significativo hacia la recolección aérea automatizada en cultivos frutales.

En cuanto a la polinización, Cui et al. (2024) desarrollaron un modelo experimental de dron polinizador para cultivos hortícolas en invernadero, combinando visión artificial basada en CNN y algoritmos de vuelo de precisión. El dispositivo identificaba flores

abiertas de tomate y ejecutaba trayectorias de vuelo optimizadas para agitar los estambres mediante ráfagas de aire dirigidas. El modelo alcanzó tasas de polinización equivalentes al método manual (82%), con una eficiencia operativa tres veces superior. La principal limitación fue la necesidad de recalibrar el sistema en función de la morfología floral específica y condiciones ambientales como humedad o viento. A pesar de su carácter preliminar, este enfoque sienta las bases para futuras aplicaciones de robótica aérea en tareas reproductivas, especialmente en escenarios donde la mano de obra es limitada.

La automatización de tareas agrícolas mediante UAVs constituye una línea emergente con alto potencial transformador en el contexto de la agricultura de precisión. Los estudios revisados evidencian la viabilidad de adaptar drones a funciones complejas como la aplicación dirigida de fitosanitarios, la recolección selectiva de frutos o la polinización asistida, mediante la integración de sensores visuales, algoritmos de aprendizaje profundo y sistemas de navegación inteligente. Si bien aún persisten desafíos técnicos (como la variabilidad ambiental, la precisión operativa en entornos no estructurados o el coste de los componentes especializados), los resultados obtenidos reflejan una clara evolución hacia sistemas más autónomos, sostenibles y eficientes, especialmente valiosos en contextos donde la mecanización tradicional resulta limitada o inviable.

4.3.3.4. Modelos de detección de flora adventicia.

Finalmente, se incorporaron estudios centrados en la detección y gestión de flora adventicia con implicaciones agronómicas y ecológicas, a partir de imágenes captadas por UAVs y procesadas mediante algoritmos de *deep learning*. Estas contribuciones representaron un avance significativo en el uso de tecnologías aéreas para identificar especies no deseadas que afectan tanto a los cultivos como a los ecosistemas circundantes, favoreciendo intervenciones más precisas, focalizadas y sostenibles.

Una de las aplicaciones más destacadas en el ámbito agrícola es la propuesta por Yadav et al. (2023), quienes desarrollaron un sistema de detección de plantas voluntarias de algodón (*volunteer cotton*) en campos de maíz utilizando UAVs equipados con cámaras RGB y el modelo YOLOv3. Estas plantas representan un riesgo fitosanitario al servir de hospedadoras para el picudo del algodonero (*Anthonomus grandis*), una plaga de alto impacto económico. El sistema logró una precisión promedio del 80,38 % (mAP) y un F1-score del 78,5 %, manteniendo su eficacia en diferentes resoluciones de entrada (320×320, 416×416, y 512×512 píxeles), lo cual sugiere robustez frente a cambios en el

tamaño de imagen. Las pruebas se realizaron en campos reales del sur de Texas, en el marco del programa de erradicación de esta plaga. Como aportes relevantes, el modelo demostró potencial para aplicaciones en tiempo real con integración en drones pulverizadores, optimizando el uso de herbicidas y reduciendo el riesgo ambiental. No obstante, el sistema se diseñó con un umbral de IoU bajo (0,3) para priorizar la sensibilidad, lo que podría comprometer la precisión espacial en escenarios de alta densidad de malezas o cultivos mixtos.

En entornos naturales complejos, Gautam et al. (2025) abordaron la detección de la especie invasora *Siam weed* (*Chromolaena odorata*) en el norte de Australia mediante imágenes UAV RGB y el modelo YOLOv5. La investigación evaluó el impacto de la iluminación solar, el número de imágenes de entrenamiento y la complejidad del modelo sobre el rendimiento de la detección. Con tan solo 1000 imágenes y utilizando la versión pequeña del modelo (YOLOv5s), se alcanzó un F1-score de 0,88, mientras que las versiones más complejas (YOLOv5x) superaron el 0,91. La precisión se vio fuertemente influida por las condiciones lumínicas: las imágenes capturadas en días nublados redujeron las falsas detecciones hasta en un 80 %, mejorando la diferenciación del *Siam weed* respecto a la vegetación nativa. Como limitación, se identificó que los reflejos del follaje y corteza generaban falsos positivos en condiciones soleadas, aunque estos se mitigaron con una adecuada calibración del modelo y la calidad de las imágenes. El estudio demostró la viabilidad de reemplazar la vigilancia aérea con helicópteros por sistemas UAV autónomos, mejorando la seguridad operativa y reduciendo costes logísticos en la gestión de malezas a escala de cuenca.

Los estudios analizados evidencian la capacidad de los modelos YOLO aplicados sobre imágenes UAV para abordar de forma eficaz los desafíos asociados a la detección de flora adventicia en entornos tanto agrícolas como naturales. A pesar de las limitaciones derivadas de las condiciones ambientales y de la necesidad de una curación minuciosa del dataset, los resultados obtenidos validan la aplicabilidad de estas herramientas en escenarios reales. En conjunto, estas propuestas apuntan hacia una gestión de malezas más automatizada, localizada y sostenible, alineada con los principios de la agricultura de precisión.

5. CONCLUSIÓN.

El presente trabajo ha permitido establecer una visión actualizada sobre el papel del *deep learning* aplicado a imágenes UAV en la agricultura de precisión. Se identifican tres principales ejes de aplicación: la detección de plagas y enfermedades, donde arquitecturas CNN clásicas, híbridas y con mecanismos de atención han mostrado un rendimiento notable en la identificación temprana de síntomas; la estimación del rendimiento agrícola, donde modelos de una etapa como YOLOv4 y YOLOv7, combinados con técnicas de regresión o atención, destacan por su eficiencia y aplicabilidad en campo; y otras aplicaciones emergentes, como el monitoreo fenológico, la automatización de labores y la detección de flora adventicia, que posicionan a la robótica aérea y la inteligencia artificial como motores de una agricultura más autónoma y sostenible.

Sin embargo, también se detectaron limitaciones comunes: escasa validación en condiciones reales, alta dependencia de sensores avanzados y UAVs costosos, carencia de *datasets* abiertos que obstaculiza la replicabilidad, sensibilidad ambiental de los modelos y escasa evaluación económica frente a métodos tradicionales.

Ante ello, se propone impulsar soluciones más accesibles y ligeras, fomentar la creación de bases de datos abiertas y representativas, facilitar su uso mediante plataformas intuitivas para técnicos no especializados, promover servicios bajo demanda (UAV-*as-a-service*), y ampliar estudios comparativos que analicen su impacto en sostenibilidad y rentabilidad.

Con base en los resultados, se puede concluir que, aunque las tecnologías analizadas muestran un alto potencial, su adopción real por parte del agricultor medio dependerá no solo de su rendimiento técnico, sino de su integración sencilla y relación coste-efectividad en la práctica agronómica. Así, esta revisión cumple con los objetivos planteados, aportando una visión crítica y aplicada que puede orientar futuras investigaciones, estrategias de transferencia tecnológica y políticas de digitalización agrícola sostenible.

6. BIBLIOGRAFÍA.

La presente sección recoge todas las fuentes utilizadas a lo largo del trabajo, seleccionadas mediante un proceso sistemático y verificable conforme al protocolo PRISMA 2020 y al análisis temático automatizado descrito en la metodología. Las referencias incluyen estudios científicos de alta relevancia obtenidos de Scopus y Web of Science entre 2019 y 2024, seleccionados por su rigor metodológico, actualidad e impacto en el ámbito del aprendizaje profundo aplicado a la agricultura de precisión. Su integración garantiza la trazabilidad, la coherencia y la solidez científica del marco teórico y de los resultados presentados.

Abdalla, A., Wheeler, T. A., Dever, J., Lin, Z., Arce, J., & Guo, W. (2024). Assessing fusarium oxysporum disease severity in cotton using unmanned aerial system images and a hybrid domain adaptation deep learning time series model.

Biosystems Engineering, 237, 220-231.

<https://doi.org/10.1016/J.BIOSYSTEMSENG.2023.12.014>

Álvarez-Martínez, F. J., Borrás-Rocher, F., Micol, V., & Barrajon-Catalán, E. (2023).

Artificial Intelligence Applied to Improve Scientific Reviews: The Antibacterial Activity of Cistus Plants as Proof of Concept. *Antibiotics*, 12(2), 327-327.

<https://doi.org/10.3390/ANTIBIOTICS12020327>

Amarasingam, N., Gonzalez, F., Salgadoe, A. S. A., Sandino, J., & Powell, K. (2022).

Detection of White Leaf Disease in Sugarcane Crops Using UAV-Derived RGB Imagery with Existing Deep Learning Models. *Remote Sensing 2022, Vol. 14*,

Page 6137, 14(23), 6137-6137. <https://doi.org/10.3390/RS14236137>

Chen, J., Fu, Y., Guo, Y., Xu, Y., Zhang, X., & Hao, F. (2024). An improved deep

learning approach for detection of maize tassels using UAV-based RGB images.

International Journal of Applied Earth Observation and Geoinformation, 130,

103922-103922. <https://doi.org/10.1016/J.JAG.2024.103922>

- Cui, B., Liang, L., Ji, B., Zhang, L., Zhao, L., Zhang, K., Shi, F., & Creput, J. C. (2024). Exploring the YOLO-FT Deep Learning Algorithm for UAV-Based Smart Agriculture Detection in Communication Networks. *IEEE Transactions on Network and Service Management*.
<https://doi.org/10.1109/TNSM.2024.3424232>
- Du, J., Li, J., Fan, J., Gu, S., Guo, X., & Zhao, C. (2024). Detection and Identification of Tassel States at Different Maize Tasseling Stages Using UAV Imagery and Deep Learning. *Plant Phenomics*, 6, 188-188.
<https://doi.org/10.34133/PLANTPHENOMICS.0188>
- Gautam, D., Mawardi, Z., Elliott, L., Loewensteiner, D., Whiteside, T., & Brooks, S. (2025). Detection of Invasive Species (Siam Weed) Using Drone-Based Imaging and YOLO Deep Learning Model. *Remote Sensing 2025, Vol. 17, Page 120, 17(1)*, 120-120. <https://doi.org/10.3390/RS17010120>
- Hong-Yu, F., Yun-Kai, Y., Wei, W., Ao, L., Ming-Zhi, X., Xihong, G., Wei, S., & Guo-Xian, C. (2023). Ramie Plant Counting Based on UAV Remote Sensing Technology and Deep Learning. *Journal of Natural Fibers*, 20(1).
<https://doi.org/10.1080/15440478.2022.2159610>
- Jia, Z., Zhang, X., Yang, H., Lu, Y., Liu, J., Yu, X., Feng, D., Gao, K., Xue, J., Ming, B., Nie, C., & Li, S. (2024). Comparison and Optimal Method of Detecting the Number of Maize Seedlings Based on Deep Learning. *Drones 2024, Vol. 8, Page 175, 8(5)*, 175-175. <https://doi.org/10.3390/DRONES8050175>
- Kalantar, A., Edan, Y., Gur, A., & Klapp, I. (2020). A deep learning system for single and overall weight estimation of melons using unmanned aerial vehicle images. *Computers and Electronics in Agriculture*, 178, 105748-105748.
<https://doi.org/10.1016/J.COMPAG.2020.105748>

- Kamilaris, A., & Prenafeta-Boldú, F. X. (2018). Deep learning in agriculture: A survey. *Computers and Electronics in Agriculture*, 147, 70-90.
<https://doi.org/10.1016/J.COMPAG.2018.02.016>
- Kerkech, M., Hafiane, A., & Canals, R. (2020). Vine disease detection in UAV multispectral images using optimized image registration and deep learning segmentation approach. *Computers and Electronics in Agriculture*, 174, 105446-105446. <https://doi.org/10.1016/J.COMPAG.2020.105446>
- Ksibi, A., Ayadi, M., Soufiene, B. O., Jamjoom, M. M., & Ullah, Z. (2022). MobiRes-Net: A Hybrid Deep Learning Model for Detecting and Classifying Olive Leaf Diseases. *Applied Sciences* 2022, Vol. 12, Page 10278, 12(20), 10278-10278.
<https://doi.org/10.3390/APP122010278>
- Yadav, P. K., Thomasson, J. A., Hardin, R., Searcy, S. W., Braga-Neto, U., Popescu, S. C., Martin, D. E., Rodriguez, R., Meza, K., Enciso, J., Solorzano Diaz, J., & Wang, T. (2023). Detecting volunteer cotton plants in a corn field with deep learning on UAV remote-sensing imagery. *Computers and Electronics in Agriculture*, 204, 107551-107551.
<https://doi.org/10.1016/J.COMPAG.2022.107551>
- Li, D., Sun, X., Jia, Y., Yao, Z., Lin, P., Chen, Y., Zhou, H., Zhou, Z., Wu, K., Shi, L., & Li, J. (2023a). A longan yield estimation approach based on UAV images and deep learning. *Frontiers in Plant Science*, 14, 1132909-1132909.
<https://doi.org/10.3389/FPLS.2023.1132909>
- Li, D., Sun, X., Lv, S., Elkhouchlaa, H., Jia, Y., Yao, Z., Lin, P., Zhou, H., Zhou, Z., Shen, J., & Li, J. (2022b). A novel approach for the 3D localization of branch picking points based on deep learning applied to longan harvesting UAVs.

Computers and Electronics in Agriculture, 199, 107191-107191.

<https://doi.org/10.1016/J.COMPAG.2022.107191>

Li, H. ;, Wang, P. ;, Huang, C., Li, H., Wang, P., & Huang, C. (2022a). Comparison of Deep Learning Methods for Detecting and Counting Sorghum Heads in UAV Imagery. *Remote Sensing 2022, Vol. 14, Page 3143, 14(13)*, 3143-3143.

<https://doi.org/10.3390/RS14133143>

Li, J., Li, Y., Qiao, J., Li, L., Wang, X., Yao, J., & Liao, G. (2023b). Automatic counting of rapeseed inflorescences using deep learning method and UAV RGB imagery. *Frontiers in Plant Science*, 14, 1101143-1101143.

<https://doi.org/10.3389/FPLS.2023.1101143>

Lin, P., Li, D., Jia, Y., Chen, Y., Huang, G., Elkhouchlaa, H., Yao, Z., Zhou, Z., Zhou, H., Li, J., & Lu, H. (2022). A novel approach for estimating the flowering rate of litchi based on deep learning and UAV images. *Frontiers in Plant Science*, 13, 966639-966639. <https://doi.org/10.3389/FPLS.2022.966639>

Lv, Z., Xu, B., Zhong, L., Chen, G., Huang, Z., Sun, R., Huang, W., Zhao, F., & Meng, R. (2024). Improved monitoring of southern corn rust using UAV-based multi-view imagery and an attention-based deep learning method. *Computers and Electronics in Agriculture*, 224, 109232-109232.

<https://doi.org/10.1016/J.COMPAG.2024.109232>

Sekharamanthy, P. K., Melgani, F., Malacarne, J., Ricci, R., Almeida Silva, R., & Marcato, J. (2024). A seamless deep learning approach for Apple detection, depth estimation, and tracking using YOLO models enhanced by Multi-Head Attention mechanism. *Computers*, 13(3), 83.

<https://doi.org/10.3390/computers13030083>

- Melnychenko, O., Scislo, L., Savenko, O., Sachenko, A., & Radiuk, P. (2024). Intelligent Integrated System for Fruit Detection Using Multi-UAV Imaging and Deep Learning. *Sensors* 2024, Vol. 24, Page 1913, 24(6), 1913-1913. <https://doi.org/10.3390/S24061913>
- Mora, J. J., Selvaraj, M. G., Alvarez, C. I., Safari, N., & Blomme, G. (2024). From pixels to plant health: Accurate detection of banana Xanthomonas wilt in complex African landscapes using high-resolution UAV images and deep learning. *Discover Applied Sciences*, 6(7), 1-29. <https://doi.org/10.1007/S42452-024-06073-Z>
- Page, M. J., Moher, D., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., Mcdonald, S., ... Mckenzie, J. E. (2021). PRISMA 2020 explanation and elaboration: Updated guidance and exemplars for reporting systematic reviews. *BMJ*, 372. <https://doi.org/10.1136/BMJ.N160>
- Roosjen, P. P., Kellenberger, B., Kooistra, L., Green, D. R., & Fahrenttrapp, J. (2020). Deep learning for automated detection of *Drosophila suzukii*: Potential for UAV-based monitoring. *Pest Management Science*, 76(9), 2994-3002. <https://doi.org/10.1002/PS.5845>
- Sassu, A., Motta, J., Deidda, A., Ghiani, L., Carlevaro, A., Garibotto, G., & Gambella, F. (2023). Artichoke deep learning detection network for site-specific agrochemicals UAS spraying. *Computers and Electronics in Agriculture*, 213, 108185-108185. <https://doi.org/10.1016/J.COMPAG.2023.108185>
- Subeesh, A., Prakash Kumar, S., Kumar Chakraborty, S., Upendar, K., Singh Chandel, N., Jat, D., Dubey, K., Modi, R. U., & Mazhar Khan, M. (2024). UAV imagery

- coupled deep learning approach for the development of an adaptive in-house web-based application for yield estimation in citrus orchard. *Measurement*, 234, 114786-114786. <https://doi.org/10.1016/J.MEASUREMENT.2024.114786>
- Velusamy, P., Rajendran, S., Mahendran, R. K., Naseer, S., Shafiq, M., & Choi, J. G. (2021). Unmanned Aerial Vehicles (UAV) in Precision Agriculture: Applications and Challenges. *Energies* 2022, Vol. 15, Page 217, 15(1), 217-217. <https://doi.org/10.3390/EN15010217>
- Vijayalakshmi, G., & Radhika, Y. (2024). Internet of Things assisted Unmanned Aerial Vehicle for Pest Detection with Optimized Deep Learning Model. *Web Intelligence*, 22(2), 269-290. <https://doi.org/10.3233/WEB-230062>
- Wang, D., Li, W., Liu, X., Li, N., & Zhang, C. (2020). UAV environmental perception and autonomous obstacle avoidance: A deep learning and depth camera combined solution. *Computers and Electronics in Agriculture*, 175, 105523-105523. <https://doi.org/10.1016/J.COMPAG.2020.105523>
- Wang, S. M., Yu, C. P., Ma, J. H., Ouyang, J. X., Zhao, Z. M., Xuan, Y. M., Fan, D. M., Yu, J. F., Wang, X. C., & Zheng, X. Q. (2024). Tea yield estimation using UAV images and deep learning. *Industrial Crops and Products*, 212, 118358-118358. <https://doi.org/10.1016/J.INDCROP.2024.118358>
- Wieme, J., Leroux, S., Cool, S. R., Van Beek, J., Pieters, J. G., & Maes, W. H. (2024). Ultra-high-resolution UAV-imaging and supervised deep learning for accurate detection of *Alternaria solani* in potato fields. *Frontiers in Plant Science*, 15, 1206998-1206998. <https://doi.org/10.3389/FPLS.2024.1206998>
- Xiong, J., Liu, Z., Chen, S., Liu, B., Zheng, Z., Zhong, Z., Yang, Z., & Peng, H. (2020). Visual detection of green mangoes by an unmanned aerial vehicle in orchards

based on a deep learning method. *Biosystems Engineering*, 194, 261-272.

<https://doi.org/10.1016/J.BIOSYSTEMSENG.2020.04.006>

Zhang, K., Deng, J., Zhou, C., Liu, J., Lv, X., Wang, Y., Sun, E., Liu, Y., Ma, Z., &

Shang, J. (2024). Using UAV hyperspectral imagery and deep learning for Object-Based quantitative inversion of *Zanthoxylum* rust disease index.

International Journal of Applied Earth Observation and Geoinformation, 135,

104262-104262. <https://doi.org/10.1016/J.JAG.2024.104262>

Zhu, J., Yang, G., Feng, X., Li, X., Fang, H., Zhang, J., Bai, X., Tao, M., & He, Y.

(2022). Detecting Wheat Heads from UAV Low-Altitude Remote Sensing

Images Using Deep Learning Based on Transformer. *Remote Sensing 2022*, Vol.

14, Page 5141, 14(20), 5141-5141. <https://doi.org/10.3390/RS14205141>



ANEXOS

Anexo I - Glosario técnico automatizado mediante inteligencia artificial.

Este glosario recoge los principales términos técnicos identificados a lo largo del trabajo, generados a partir de inteligencia artificial generativa. Su objetivo es facilitar la comprensión de los conceptos especializados vinculados a la agricultura de precisión, los UAVs y las arquitecturas de *deep learning* tratadas en el estudio.

A

Algoritmo K-means: Método de agrupamiento no supervisado que organiza un conjunto de datos en K grupos o grupos según su similitud. Asigna cada elemento al grupo con el centroide más cercano y ajusta iterativamente los centroides para minimizar la variación interna de cada grupo.

Aprendizaje automático (*Machine Learning*): Disciplina de la inteligencia artificial que desarrolla algoritmos capaces de aprender automáticamente a partir de datos. En lugar de seguir reglas programadas explícitas, los modelos de aprendizaje automático mejoran su rendimiento mediante la experiencia obtenida de ejemplos. Tiene aplicaciones en agricultura de precisión, desde la predicción de rendimientos hasta la detección automatizada de plagas, y sienta las bases de técnicas avanzadas como el aprendizaje profundo.

Arquitectura *encoder-decoder*: Tipo de modelo de red neuronal compuesto por dos partes: un codificador (*encoder*) que reduce progresivamente la dimensionalidad de la entrada para extraer una representación latente compacta, y un decodificador (*decoder*) que reconstruye una salida detallada a partir de esa representación. Esta arquitectura simétrica permite identificar simultáneamente qué objetos están presentes y dónde, y es típica en tareas de segmentación de imágenes a nivel de píxel (por ejemplo, crear máscaras de cultivos o enfermedades en imágenes UAV).

Atención espacial: Mecanismo de inteligencia artificial que permite a una red neuronal enfocar su procesamiento en regiones geográficas relevantes de una imagen. La atención espacial pondera los píxeles o características espaciales más importantes para la tarea, resaltando zonas de interés (por ejemplo, áreas con síntomas en un cultivo) e ignorando fondos o elementos irrelevantes. Su incorporación en modelos de visión por computadora

mejora la precisión en detección de objetos y segmentación en imágenes de UAV al destacar las regiones clave.

B

Biblioteca Sklearn: Librería de Python especializada en aprendizaje automático. Ofrece herramientas eficientes y accesibles para tareas como clasificación, regresión, agrupamiento (*clustering*), reducción de dimensiones y validación de modelos. Es ampliamente utilizada en proyectos científicos y de análisis de datos por su integración con otras bibliotecas como NumPy y pandas.

C

Cadena booleana: Expresión lógica compuesta por palabras clave combinadas mediante operadores booleanos (como AND, OR, NOT) utilizada para realizar búsquedas avanzadas en bases de datos. Permite refinar y personalizar la recuperación de información, facilitando la localización de documentos relevantes en función de criterios específicos.

Cámaras de profundidad: Dispositivos de captura de imagen que, además de registrar información visual, miden la distancia entre el sensor y los objetos de la escena. Utilizan tecnologías como luz estructurada, tiempo de vuelo (ToF) o estéreo para generar mapas tridimensionales. En agricultura, se emplean para estimar volumen, altura de plantas, biomasa o estructura del dosel vegetal.

Centroides: Puntos que representan la posición media de todos los datos dentro de un grupo en el espacio reducido generado por el PCA. Actúan como referencia central para visualizar la distribución de los elementos agrupados, facilitando la interpretación de las relaciones entre grupos en representaciones bidimensionales o tridimensionales.

Clasificación binaria: Problema de clasificación en el que únicamente existen dos clases posibles para cada instancia de datos (por ejemplo, presencia o ausencia de una plaga en una imagen). Los modelos entrenados bajo este esquema distinguen entre ambas categorías, resultando útiles en agricultura de precisión para detecciones dicotómicas (como "enfermo" vs. "sano") con un enfoque supervisado simple pero eficaz.

Clasificación multiclase: Problema de clasificación donde una instancia de datos puede pertenecer a más de dos categorías distintas. En el contexto agrícola, esto permite identificar múltiples clases de interés (por ejemplo, varios tipos de enfermedades foliares o diferentes cultivos) en lugar de solo distinguir entre dos estados. Aunque más compleja que la clasificación binaria, provee un análisis más rico al diferenciar diversos tipos de objetos o condiciones en las muestras.

Clústeres: Grupos de elementos con características similares que se forman como resultado de un proceso de agrupamiento (*clustering*). En análisis de datos, los grupos permiten identificar patrones, segmentar información o descubrir relaciones ocultas dentro de grandes conjuntos de datos sin necesidad de etiquetas previas.

Conteo automatizado: Procedimiento de cuantificación de objetos o elementos mediante métodos computacionales, sin intervención manual. En agricultura de precisión, el conteo automatizado se aplica, por ejemplo, para determinar la cantidad de frutos o plantas detectados en imágenes aéreas capturadas por drones, reemplazando el conteo manual por una solución rápida, precisa y objetiva basada en visión artificial.

Cuaderno Jupyter: Entorno interactivo de programación que permite combinar código, texto, fórmulas matemáticas y visualizaciones en un único documento. Es ampliamente utilizado en investigación científica y análisis de datos por su capacidad para documentar y ejecutar procesos de forma reproducible y ordenada.

D

Dataset: Colección estructurada de datos organizada generalmente en forma de tablas, que sirve como base para el análisis, entrenamiento de modelos o extracción de información. En investigación, un *dataset* puede estar compuesto por textos, imágenes, mediciones u otras variables recopiladas de forma sistemática.

Deep learning: Rama del aprendizaje automático basada en redes neuronales artificiales con múltiples capas que imitan el funcionamiento del cerebro humano. Permite el reconocimiento avanzado de patrones en grandes volúmenes de datos, siendo especialmente eficaz en tareas como el análisis de imágenes, voz o texto.

E

EfficientDet: Modelo avanzado de detección de objetos que combina eficiencia computacional y alta precisión. Se basa en la arquitectura EfficientNet como red base y en un módulo de fusión de características llamado BiFPN (*Bidirectional Feature Pyramid Network*). Gracias a su diseño optimizado, logra un buen equilibrio entre velocidad y rendimiento, siendo una alternativa eficiente a modelos como YOLO o SSD en tareas de visión por computadora, incluida la agricultura de precisión.

Etapas fenológicas: Fase específica del desarrollo de una planta dentro de su ciclo de vida, que puede incluir la germinación, floración, fructificación o senescencia, entre otras. El seguimiento de las etapas fenológicas permite optimizar prácticas agronómicas como el riego, la fertilización o la aplicación de tratamientos fitosanitarios, siendo clave en agricultura de precisión y estudios de teledetección.

F

F1-score: Métrica utilizada para evaluar el rendimiento de modelos de clasificación, especialmente en contextos con clases desbalanceadas. Representa la media armónica entre la **precisión** (*precision*) y la **sensibilidad** (*recall*), combinando ambos factores en una única medida. Su valor oscila entre 0 (rendimiento nulo) y 1 (rendimiento perfecto).

Faster R-CNN: Modelo avanzado de detección de objetos que combina la generación de regiones de interés con redes neuronales convolucionales para identificar y clasificar objetos dentro de una imagen. Introduce una red de propuestas de regiones (RPN) que mejora significativamente la velocidad respecto a versiones anteriores como R-CNN y Fast R-CNN. Es ampliamente utilizado en visión por computadora para tareas de precisión, como la detección de enfermedades en cultivos mediante imágenes UAV.

Fusión de bandas espectrales: Procedimiento en teledetección que integra múltiples bandas del espectro electromagnético (por ejemplo, combinando imágenes de diferentes longitudes de onda) para obtener información más completa de una escena. En agricultura de precisión, la fusión de bandas espectrales puede mejorar la detección de características agronómicas al aprovechar la complementariedad entre imágenes RGB, infrarrojas, multiespectrales o hiperespectrales, realizando detalles que no serían evidentes en una única banda.

G

GRADE: Metodología estandarizada para evaluar la calidad de la evidencia científica y la fuerza de las recomendaciones en revisiones sistemáticas y guías clínicas. Clasifica la certeza de la evidencia en cuatro niveles (alta, moderada, baja o muy baja) considerando factores como el diseño del estudio, el riesgo de sesgo, la consistencia y la precisión de los resultados.

H

Hibridación: Proceso mediante el cual se cruzan dos organismos de diferente genética para obtener una descendencia con características deseadas de ambos progenitores. En agricultura, la hibridación se emplea para mejorar variedades vegetales en términos de rendimiento, resistencia a enfermedades, tolerancia al estrés o calidad del producto.

I

Imágenes RGB: Representaciones digitales que combinan tres canales de color —rojo (*Red*), verde (*Green*) y azul (*Blue*)— para formar imágenes a todo color, simulando la percepción visual humana. En teledetección y agricultura de precisión, se utilizan para captar información visual básica del cultivo y realizar análisis mediante visión artificial.

Indexados: Término que se refiere a los documentos, artículos o revistas que han sido incorporados en bases de datos científicas reconocidas. Estar indexado implica que el contenido cumple ciertos estándares de calidad, visibilidad y accesibilidad, lo que favorece su consulta y citación en el ámbito académico.

Índice de silueta: Métrica utilizada para evaluar la calidad de un agrupamiento (*clustering*). Mide qué tan similar es un objeto a los elementos de su propio grupo en comparación con los de otros grupos. Su valor oscila entre -1 y 1, siendo más cercano a 1 cuando la asignación es adecuada, lo que indica grupos bien definidos.

Índices de vegetación: Indicadores numéricos calculados a partir de combinaciones de bandas espectrales reflejadas por la vegetación, captadas mediante sensores remotos. Permiten evaluar el estado fisiológico de las plantas, su cobertura, vigor o nivel de estrés, siendo herramientas clave en agricultura de precisión y teledetección.

IoT: Red de dispositivos físicos interconectados que recopilan, envían y reciben datos a través de Internet. En el ámbito agrícola, el IoT permite integrar sensores, estaciones meteorológicas, drones y maquinaria para monitorizar y automatizar procesos en tiempo real, mejorando la eficiencia, la trazabilidad y la toma de decisiones en agricultura de precisión.

IoU (*Intersection over Union*): Métrica de superposición utilizada en visión por computadora para cuantificar la precisión espacial en la detección de objetos. Se calcula como el cociente entre el área de intersección y el área de unión entre la región predicha por el modelo y la región real del objeto. Un valor de IoU elevado (cercano a 1) indica que la predicción del modelo coincide casi perfectamente con la ubicación real del objeto, mientras que valores bajos implican poca superposición.

L

Librería NLTK: Biblioteca de Python especializada en el procesamiento de lenguaje natural. Proporciona herramientas y recursos para tareas como *tokenización*, análisis sintáctico, lematización, extracción de entidades y clasificación de texto, siendo ampliamente utilizada en investigación y proyectos de análisis textual.

M

mAP_{0.5}: Métrica estándar utilizada para evaluar la precisión de modelos de detección de objetos. Representa el promedio de la precisión obtenida para todas las clases cuando la superposición entre la predicción y el objeto real (IoU) es al menos del 50%. Un valor más alto de mAP@0.5 indica una mayor capacidad del modelo para localizar y clasificar correctamente los objetos en una imagen.

Metaanálisis: Técnica estadística que combina los resultados cuantitativos de múltiples estudios independientes sobre una misma pregunta de investigación. Su objetivo es obtener una estimación global más precisa del efecto estudiado, aumentando la potencia estadística y la generalización de los hallazgos.

Metadatos: Información descriptiva que acompaña a un conjunto de datos y facilita su organización, búsqueda e interpretación. Incluye detalles como el autor, la fecha, el

formato, la fuente o el contenido temático, siendo fundamentales para garantizar la trazabilidad y reutilización de los datos.

Método de *Elbow*: Técnica gráfica utilizada para determinar el número óptimo de grupos en un análisis de agrupamiento. Consiste en representar la variación explicada (inercia) frente al número de grupos y localizar el punto en el que la mejora se vuelve marginal, formando un “codo” en la curva. Este punto indica un equilibrio entre simplicidad y precisión.

Modelización predictiva: Proceso analítico que utiliza datos históricos y algoritmos estadísticos o de aprendizaje automático para construir modelos capaces de anticipar resultados futuros. En el ámbito agrícola, permite prever variables como el rendimiento de cultivos, la aparición de enfermedades o la evolución de condiciones ambientales.

Modelos supervisados: Algoritmos de aprendizaje automático entrenados con datos etiquetados, es decir, con ejemplos que ya contienen la respuesta esperada. El modelo aprende a relacionar las entradas con sus salidas correspondientes, permitiendo predecir resultados sobre nuevos datos. Se utilizan en tareas como clasificación, regresión o detección de anomalías.

Multiespectrales: Capturas obtenidas mediante sensores que registran la reflectancia de la superficie terrestre en múltiples bandas del espectro electromagnético, más allá del rango visible. Estas imágenes permiten analizar características fisiológicas de la vegetación, identificar tipos de cultivo o detectar estrés hídrico, siendo fundamentales en teledetección agrícola.

P

PCA: Técnica estadística de reducción de dimensiones que transforma un conjunto de variables correlacionadas en un número menor de componentes principales. Estos componentes capturan la mayor parte de la variabilidad de los datos, facilitando su visualización e interpretación sin perder información relevante.

Precision agriculture: Enfoque de gestión agrícola basado en la observación, medición y respuesta a la variabilidad de los cultivos mediante el uso de tecnologías avanzadas. La agricultura de precisión integra herramientas como sensores, drones, sistemas GPS y

modelos predictivos para optimizar el uso de recursos, mejorar el rendimiento y reducir el impacto ambiental.

PRISMA: Guía metodológica utilizada para mejorar la transparencia y la calidad en la elaboración de revisiones sistemáticas. Establece criterios estandarizados para la selección, evaluación y presentación de los estudios incluidos, permitiendo una mayor trazabilidad y reproducibilidad del proceso de revisión.

Procesamiento de lenguaje natural: Conjunto de técnicas de inteligencia artificial que permiten a los ordenadores comprender, interpretar y generar lenguaje humano de forma automatizada. Se utiliza en tareas como el análisis de textos, la clasificación de documentos o la extracción de información relevante.

R

R²: Indicador estadístico que mide el grado de ajuste de un modelo predictivo a los datos observados. En estimaciones de rendimiento agrícola, refleja qué proporción de la variabilidad del rendimiento real es explicada por el modelo. Su valor oscila entre 0 y 1, siendo 1 un ajuste perfecto, mientras que valores cercanos a 0 indican bajo poder explicativo.

Red neuronal convolucional (CNN): Tipo de red neuronal profunda especialmente eficaz en el procesamiento de imágenes, gracias a su uso de capas convolucionales para extraer características locales y espaciales. Las CNN han sido ampliamente empleadas en visión por computadora (por ejemplo, para detectar plagas o clasificar el estado de cultivos en imágenes de drones) debido a su capacidad para reconocer patrones visuales con alta precisión.

Red neuronal recurrente (RNN): Tipo de red neuronal diseñada para procesar secuencias de datos mediante una memoria interna que retiene información de entradas anteriores. Las RNN pueden modelar dependencias temporales o secuenciales, siendo útiles para analizar series de tiempo u ordenamientos de datos (por ejemplo, secuencias temporales de imágenes o datos meteorológicos). Aunque menos comunes que las redes convolucionales en visión agrícola, pueden incorporarse para capturar dinámicas temporales en modelos de predicción agronómica.

Representación espacial de precisión: Visualización georreferenciada detallada de datos agronómicos clave que refleja con exactitud la variabilidad espacial de un cultivo o fenómeno. Estas representaciones (por ejemplo, mapas de rendimiento o de índices de vegetación) son fundamentales en agricultura de precisión, ya que permiten identificar heterogeneidades dentro del campo y tomar decisiones localizadas basadas en información espacial precisa.

Revisión sistemática: Proceso metodológico estructurado que busca identificar, seleccionar, evaluar y sintetizar de forma rigurosa toda la evidencia científica relevante sobre una pregunta de investigación específica. Su objetivo es ofrecer conclusiones objetivas y reproducibles a partir de criterios previamente definidos.

RoB 2.0: Herramienta actualizada desarrollada por Cochrane para evaluar el riesgo de sesgo en ensayos clínicos aleatorizados. Analiza aspectos como la generación de la secuencia aleatoria, el ocultamiento de la asignación, el cegamiento, los datos incompletos y la notificación selectiva, proporcionando una valoración estructurada sobre la fiabilidad de los resultados.

RMSE (*Root Mean Square Error*): Medida de error utilizada en modelos predictivos que se calcula como la raíz cuadrada de la media de los errores al cuadrado entre valores predichos y observados. Indica la magnitud promedio de las diferencias entre las predicciones del modelo y los valores reales, expresada en las mismas unidades de la variable objetivo. En contextos agrícolas, como la predicción de rendimiento, un RMSE bajo implica que el modelo está estimando con alta precisión.

S

SSD: Modelo de detección de objetos en tiempo real basado en redes neuronales convolucionales. A diferencia de otros enfoques que requieren múltiples etapas, SSD realiza la localización y clasificación de objetos en una sola pasada (*single shot*), lo que lo hace rápido y eficiente. Es ampliamente utilizado en aplicaciones de visión por computadora agrícola, similar a modelos como YOLO.

Segmentación a nivel de píxel: Modalidad de análisis de imágenes en la que se clasifica cada píxel según la categoría a la que pertenece. Esta segmentación de alta granularidad permite delinear con precisión regiones de interés dentro de una imagen (por ejemplo, diferenciar píxeles de cultivo, maleza y suelo). En imágenes UAV, la segmentación a

nivel de píxel proporciona mapas detallados de la distribución espacial de distintos elementos o estados del cultivo.

Segmentación multicategoría: Procedimiento de segmentación de imágenes en el cual se asigna a cada píxel una etiqueta entre más de dos categorías posibles. A diferencia de la segmentación binaria, la segmentación multicategoría permite distinguir múltiples clases en la misma imagen (por ejemplo, diferentes tipos de enfermedades o varias especies vegetales). Este enfoque aporta un análisis más completo en agricultura de precisión cuando es necesario identificar diversos objetos o condiciones en las imágenes de los cultivos.

Segmentación semántica: Tipo de segmentación de imágenes en la que cada píxel es etiquetado con una categoría semántica correspondiente a un objeto o región. En la segmentación semántica, todos los píxeles pertenecientes a la misma clase (por ejemplo, "vegetación" vs. "suelo") se marcan igual, sin distinguir instancias individuales. Se emplea en visión por computadora agrícola para generar mapas temáticos (como zonas enfermas versus sanas) a partir de imágenes de drones.

Stemming: Técnica del procesamiento de lenguaje natural que consiste en reducir las palabras a su raíz o forma base, eliminando sufijos y terminaciones. Su objetivo es unificar variantes de una misma palabra para facilitar el análisis de texto, aunque puede generar formas no reconocibles en el lenguaje natural.

Subcluster: Conjunto de elementos que forman parte de un grupo más amplio y que comparten características aún más específicas entre sí. Los *subclusters* permiten afinar el análisis dentro de un agrupamiento general, revelando patrones o temáticas diferenciadas que no eran evidentes a nivel global.

Subagrupamiento: Proceso de división de grupos principales en grupos más pequeños y homogéneos para revelar estructuras internas o patrones específicos dentro de los datos. Se utiliza cuando un grupo inicial presenta alta diversidad temática o cuando se busca un análisis más detallado y granular.

T

TF-IDF: Métrica utilizada en el análisis de textos para evaluar la importancia de una palabra dentro de un documento y en un conjunto de documentos. Combina la frecuencia

de un término en un texto con su rareza en el *corpus*, ayudando a destacar palabras clave relevantes y a filtrar términos comunes.

Tokenización: Proceso fundamental en el análisis de textos que consiste en dividir una secuencia de texto en unidades más pequeñas llamadas tokens, como palabras, frases o signos de puntuación. Es un paso previo esencial para aplicar técnicas de procesamiento de lenguaje natural.

Transformers: Arquitectura de aprendizaje profundo diseñada para procesar secuencias de datos, especialmente eficaz en tareas de procesamiento de lenguaje natural y, más recientemente, en visión por computadora. Se basa en mecanismos de atención que permiten al modelo centrarse en las partes más relevantes de la entrada. Su eficiencia manejando contextos largos ha llevado a su adopción en modelos avanzados como BERT, GPT o *Vision Transformer* (ViT).

U

Unmanned Aerial Vehicles: Dispositivos voladores que operan sin piloto a bordo, conocidos en español como Vehículos Aéreos no Tripulados. Pueden ser controlados de forma remota o mediante sistemas autónomos. En el ámbito agrícola, los UAV se utilizan para capturar imágenes aéreas y recopilar datos mediante sensores embarcados, facilitando el monitoreo de cultivos y apoyando la agricultura de precisión.

V

Validación cruzada: Técnica de evaluación de modelos predictivos que consiste en dividir los datos disponibles en varios subconjuntos (*folds*) y entrenar y probar el modelo repetidamente intercambiando los conjuntos de entrenamiento y validación. Permite estimar la capacidad de generalización del modelo de forma más robusta que una única partición fija, reduciendo el sobreajuste y aprovechando mejor los datos disponibles.

Validación semántica: Proceso mediante el cual se verifica que la agrupación o categorización de datos textuales refleja coherencia temática y sentido lógico en relación con su contenido. Se utiliza para asegurar que los resultados de técnicas como el *clustering* o la clasificación automática sean interpretables y conceptualmente consistentes.

Visión Transformer (ViT): Modelo de *Transformer* aplicado a visión por computadora. En un ViT, la imagen se divide en parches que son procesados como *tokens* por capas de autoatención, permitiendo al modelo capturar relaciones globales dentro de la imagen. Los *Vision Transformers* han logrado resultados competitivos en clasificación y detección de objetos, y comienzan a emplearse en el análisis de imágenes UAV como alternativa a las CNN tradicionales por su capacidad de modelar características visuales de forma eficiente.

Visión artificial: Disciplina de la inteligencia artificial que permite a los sistemas informáticos interpretar y analizar imágenes o secuencias visuales del entorno. En agricultura, se aplica al reconocimiento de patrones, detección de objetos o medición de parámetros a partir de imágenes captadas por cámaras o sensores, como los embarcados en UAV.

Visión por computadora: Rama de la inteligencia artificial que se centra en permitir que los ordenadores analicen, comprendan y extraigan información útil a partir de imágenes o vídeos. Emplea algoritmos avanzados para tareas como detección de objetos, clasificación, segmentación o seguimiento, y tiene aplicaciones clave en áreas como la agricultura de precisión, la robótica o la seguridad.

Y

YOLOv8: Algoritmo de visión por computadora basado en redes neuronales profundas, diseñado para la detección de objetos en tiempo real con alta precisión y velocidad. Es la octava versión de la arquitectura YOLO, y mejora el rendimiento respecto a versiones anteriores mediante optimizaciones en la segmentación, clasificación y seguimiento de objetos. Se emplea en aplicaciones como el análisis de imágenes UAV en agricultura de precisión.

Anexo II - Código fuente del cuaderno *Jupyter* para el preprocesamiento del *dataset*.

El presente cuaderno *Jupyter* ha sido diseñado como paso previo a la aplicación de técnicas de procesamiento de lenguaje natural (NLP) y análisis de agrupamiento temático. Su finalidad es facilitar la carga, unificación y limpieza de la base de datos inicial que alimentará el análisis exploratorio automatizado.

Para ello, el cuaderno incluye:

- Un módulo interactivo que permite importar archivos .xlsx almacenados en Google Drive, utilizando enlaces públicos y sin necesidad de descargarlos manualmente.
- Un bloque de procesamiento que combina automáticamente todos los archivos cargados, siempre que contengan las columnas clave para el estudio: “*Title*”, “*Year*” y “*Abstract*”.
- Una etapa de depuración que elimina registros incompletos, garantizando así que el análisis semántico posterior se base únicamente en entradas con contenido textual suficiente y metadatos válidos.

La motivación principal de este cuaderno es asegurar la coherencia estructural y la calidad mínima de los datos textuales desde el inicio del proceso metodológico, evitando errores en el modelado posterior y manteniendo un flujo de trabajo reproducible y transparente.

1. Carga, tratamiento y exportación del dataset preliminar.

Este primer bloque de código se encarga de automatizar la carga, conversión, limpieza y consolidación de los archivos que componen el dataset preliminar a partir de distintas fuentes almacenadas en Google Drive. El proceso parte del montaje del entorno colaborativo de Google Colab con el sistema de archivos del usuario, permitiendo acceder de forma dinámica a la carpeta de trabajo donde se localizan los archivos Excel originales.

En primer lugar, se implementa una rutina para detectar y convertir automáticamente los archivos en formato .xls a .xlsx, lo cual garantiza su compatibilidad con las bibliotecas modernas de procesamiento de datos. A continuación, el código identifica las columnas clave necesarias para el análisis posterior (“*Title*”, “*Year*” y “*Abstract*”) y aplica un mapeo flexible para unificar nombres de columnas heterogéneos, permitiendo una integración robusta de múltiples documentos con estructuras ligeramente diferentes.

Posteriormente, los archivos válidos son leídos, filtrados y concatenados en un único *dataset*, del cual se eliminan los registros con datos incompletos. Además, se realiza una conversión explícita del campo “Year” a un formato numérico estándar, asegurando la calidad de los metadatos. Este preprocesamiento inicial garantiza que únicamente los registros completos y estructurados pasen a la siguiente fase del análisis temático.

Finalmente, el dataset consolidado es exportado y almacenado automáticamente en una subcarpeta de destino dentro de Google Drive, cerrando el ciclo de preparación de datos con una trazabilidad total y sin necesidad de intervención manual. Esta automatización no solo mejora la reproducibilidad del flujo de trabajo, sino que también reduce significativamente la posibilidad de errores humanos en fases críticas de la recolección de datos.

```
import pandas as pd
import os
from google.colab import drive

# Montar Google Drive
print("Montando Google Drive...")
drive.mount('/content/drive')
print("Google Drive montado.")

# Especifica la ruta a tu carpeta de Google Drive
# Reemplaza 'Mi Drive/NombreDeTuCarpeta' con la ruta real a tu carpeta
carpeta_drive = '/content/drive/MyDrive/AÑO 2024-25/2° CUATRIMESTRE/TFG 2.0/Python/1. Dataset/1. INICIO' #@param {type:"string"}

# Verificar si la carpeta existe
if not os.path.exists(carpeta_drive):
    print(f"Error: La carpeta '{carpeta_drive}' no fue encontrada.")
else:
    print(f"Accediendo a la carpeta: {carpeta_drive}")

# Listar todos los archivos en la carpeta
archivos_en_carpeta = os.listdir(carpeta_drive)

# Convertir archivos .xls a .xlsx
for archivo in archivos_en_carpeta:
    if archivo.endswith('.xls') and not archivo.startswith('~$'): # Ignorar archivos temporales de Excel
        ruta_completa_xls = os.path.join(carpeta_drive, archivo)
        nombre_base, extension = os.path.splitext(archivo)
        nombre_archivo_xlsx = f"{nombre_base}.xlsx"
        ruta_completa_xlsx = os.path.join(carpeta_drive, nombre_archivo_xlsx)

# Verificar si el archivo .xlsx ya existe para evitar sobrescribir
if not os.path.exists(ruta_completa_xlsx):
    try:
        print(f"Convirtiendo '{archivo}' a '{nombre_archivo_xlsx}'...")
        df_temp = pd.read_excel(ruta_completa_xls)
        df_temp.to_excel(ruta_completa_xlsx, index=False, engine='openpyxl')
```

```

print(f"Conversión de '{archivo}' completada.")
except Exception as e:
print(f"Error al convertir el archivo '{archivo}': {e}")
else:
print(f"El archivo '{nombre_archivo_xlsx}' ya existe. Saltando conversión para '{archivo}'.")

# Ahora procedemos a leer los archivos .xlsx (incluyendo los recién convertidos)
archivos_xlsx = [f for f in os.listdir(carpeta_drive) if f.endswith('.xlsx')]

if not archivos_xlsx:
print("No se encontraron archivos .xlsx en la carpeta especificada.")
else:
print(f"Archivos .xlsx encontrados (incluyendo los convertidos): {archivos_xlsx}")

# Diccionario de mapeo para los nombres de columnas alternativos
column_mapping = {
'Article Title': 'Title',
'Publication Year': 'Year',
'Abstract': 'Abstract',
}

# Leer cada archivo .xlsx, renombrar columnas y seleccionar solo las necesarias
lista_dataframes = []
required_columns = ['Title', 'Year', 'Abstract']
for archivo in archivos_xlsx:
ruta_completa_archivo = os.path.join(carpeta_drive, archivo)
try:
df_temp = pd.read_excel(ruta_completa_archivo)
print(f"'{archivo}' leído exitosamente.")

rename_dict = {col: column_mapping[col] for col in column_mapping if col in df_temp.columns}
if rename_dict:
df_temp.rename(columns=rename_dict, inplace=True)
print(f" Columnas renombradas en '{archivo}': {rename_dict}")

present_required_columns = [col for col in required_columns if col in df_temp.columns]

if len(present_required_columns) == len(required_columns):
df_temp_selected = df_temp[required_columns]
lista_dataframes.append(df_temp_selected)
print(f" Columnas {required_columns} seleccionadas y añadidas a la lista desde '{archivo}'.")
else:
missing = [col for col in required_columns if col not in df_temp.columns]
print(f" Advertencia: Faltan columnas requeridas ({missing}) en '{archivo}' después de intentar renombrar. Este archivo no se incluirá en la combinación.")

except Exception as e:
print(f"Error al leer o procesar el archivo '{archivo}': {e}")

# Combinar múltiples datasets y limpiar registros con campos vacíos
if 'lista_dataframes' in locals() and lista_dataframes:
combined_df = pd.concat(lista_dataframes, ignore_index=True)

print(f"\nNúmero de filas antes de eliminar nulos: {len(combined_df)}")

required_columns = ['Title', 'Year', 'Abstract']
combined_df.dropna(subset=required_columns, inplace=True)

```

```

print(f"Número de filas después de eliminar nulos: {len(combined_df)}")

if 'Year' in combined_df.columns:
    combined_df['Year'] = pd.to_numeric(combined_df['Year'], errors='coerce').astype('Int64')

print("\nTipos de datos después de convertir 'Year' a int (si 'Year' existe):")
print(combined_df.dtypes)

print("\nPrimeras filas del DataFrame combinado y limpio:")
display(combined_df.head(-10))

else:
    print("Error: lista dataframes está vacía o no disponible. No se cargaron DataFrames o ninguno
    contenía las columnas requeridas.")
    print("Por favor, asegúrese de que la celda anterior se ejecutó correctamente y encontró archivos Excel
    con las columnas necesarias (posiblemente con nombres alternativos).")

# Descarga de los archivos a una carpeta de GoogleDrive.
import os

if 'combined_df' in locals() and not combined_df.empty:
    carpeta_destino_drive = '/content/drive/MyDrive/AÑO 2024-25/2º CUATRIMESTRE/TFG
    2.0/Python/1. Dataset/2. FINAL'
    if not os.path.exists(carpeta_destino_drive):
        os.makedirs(carpeta_destino_drive)
    print(f"Carpeta '{carpeta_destino_drive}' creada en Google Drive.")

    nombre_archivo_salida = 'dataset_tratado_drive.xlsx'
    ruta_guardado = os.path.join(carpeta_destino_drive, nombre_archivo_salida)

    try:
        combined_df.to_excel(ruta_guardado, index=False)
        print(f"\nDataFrame guardado exitosamente en Google Drive: {ruta_guardado}")
    except Exception as e:
        print(f"\nError al guardar el archivo en Google Drive: {e}")
    else:
        print("\nNo hay DataFrame combinado para guardar en Drive.")

```

2. Preparación del *dataset* específico para el subagrupamiento temático.

Este segundo bloque del cuaderno se centra en aislar y preparar un subconjunto específico del *corpus* documental previamente clasificado, con el objetivo de aplicarle un análisis más detallado en fases posteriores. Para ello, se parte del archivo “abstract_by_cluster.xlsx”, previamente generado en el análisis exploratorio, el cual contiene todos los registros agrupados por clúster temático.

El *script* automatiza la carga de dicho archivo desde Google Drive, establece la columna “Cluster” como índice del dataset y permite acceder directamente a los registros pertenecientes al clúster 3, considerado de interés por su coherencia semántica y

relevancia investigadora. Esta operación de filtrado posibilita aislar únicamente los artículos asociados a una determinada temática emergente dentro del conjunto global.

Una vez seleccionadas las filas correspondientes, el código verifica la presencia de las columnas clave (“Title”, “Year” y “Abstract”) y extrae únicamente la información necesaria para el posterior subagrupamiento y análisis cualitativo. El subconjunto final es exportado a una carpeta específica de Google Drive en formato .xlsx, manteniendo la integridad del flujo de trabajo y garantizando una estructura de datos limpia, manejable y lista para el análisis semántico refinado.

Este paso resulta fundamental para aumentar la precisión del estudio temático, permitiendo aplicar técnicas más específicas sobre subconjuntos bien delimitados y evitando la contaminación cruzada entre líneas de investigación diferentes.

```
from IPython import get_ipython
from IPython.display import display
import pandas as pd
import os
from google.colab import drive

# Montar Google Drive
print("Montando Google Drive...")
drive.mount('/content/drive')
print("Google Drive montado.")

# Especifica la ruta a la carpeta de Google Drive que contiene el archivo .xlsx
# Reemplaza 'Mi Drive/NombreDeTuCarpeta' con la ruta real a tu carpeta
# Asegúrate de que la carpeta y el archivo tengan permisos de lectura
carpeta_drive_lectura = '/content/drive/MyDrive/AÑO 2024-25/2° CUATRIMESTRE/TFG
2.0/Python/1. Dataset/3. CLUSTER' #@param {type:"string"}

# Especifica el nombre del archivo .xlsx a leer
nombre_archivo_entrada = 'abstract_by_cluster.xlsx' #@param {type:"string"}
ruta_completa_archivo_entrada = os.path.join(carpeta_drive_lectura, nombre_archivo_entrada)

# Verificar si la carpeta y el archivo existen
if not os.path.exists(ruta_completa_archivo_entrada):
    print(f"Error: El archivo '{ruta_completa_archivo_entrada}' no fue encontrado.")
else:
    print(f"Accediendo al archivo: {ruta_completa_archivo_entrada}")
    try:
        # Leer el archivo .xlsx
        df = pd.read_excel(ruta_completa_archivo_entrada)
        print(f"'{nombre_archivo_entrada}' leído exitosamente.")

        # Mostrar las primeras filas y columnas para inspección
        print("\nPrimeras filas del DataFrame leído:")
        display(df.head())
        print("\nColumnas del DataFrame leído:")
        print(df.columns)

        # Establecer la columna "Cluster" como índice
```

```

if 'Cluster' in df.columns:
    df.set_index('Cluster', inplace=True)
    print("\nColumna 'Cluster' establecida como índice.")
    print("\nDataFrame con 'Cluster' como índice:")
    display(df.head())

# Filtrar las filas donde el índice 'Cluster' es igual a 5
# Usamos .loc para filtrar por el valor del índice
df_cluster_5 = df.loc[3]

# Seleccionar solo las columnas "Title", "Year" y "Abstract"
columnas_deseadas = ['Title', 'Year', 'Abstract']
# Asegurarse de que las columnas deseadas existan en el DataFrame filtrado
columnas_existentes = [col for col in columnas_deseadas if col in df_cluster_5.columns]

if columnas_existentes:
    df_cluster_5_seleccionado = df_cluster_5[columnas_existentes]
    print("\nFilas del Cluster 5 con columnas seleccionadas ('Title', 'Year', 'Abstract'):")
    display(df_cluster_5_seleccionado.head(-10))

# Especifica la carpeta de destino para el archivo de salida
carpeta_destino_drive = '/content/drive/MyDrive/AÑO 2024-25/2º CUATRIMESTRE/TFG
2.0/Python/1. Dataset/3. CLUSTER' #@param {type:"string"}

# Crear la carpeta de destino si no existe
if not os.path.exists(carpeta_destino_drive):
    os.makedirs(carpeta_destino_drive)
    print(f"\nCarpeta '{carpeta_destino_drive}' creada en Google Drive.")

# Especifica el nombre del archivo de salida
nombre_archivo_salida = 'cluster_tratado.xlsx' #@param {type:"string"}
ruta_guardado = os.path.join(carpeta_destino_drive, nombre_archivo_salida)

try:
    # Exportar el DataFrame filtrado y seleccionado a un archivo .xlsx
    # No incluimos el índice en el archivo de salida a menos que sea necesario
    df_cluster_5_seleccionado.to_excel(ruta_guardado, index=False)
    print(f"\nDataFrame del Cluster 5 guardado exitosamente en Google Drive:
{ruta_guardado}")
except Exception as e:
    print(f"\nError al guardar el archivo en Google Drive: {e}")
else:
    print("Error: Ninguna de las columnas 'Title', 'Year', 'Abstract' se encontró en el DataFrame
después de filtrar por Cluster 5.")

else:
    print("Error: La columna 'Cluster' no se encontró en el DataFrame leído.")

except Exception as e:
    print(f"Error al leer o procesar el archivo '{nombre_archivo_entrada}': {e}")

```

3. Extracción del subconjunto *deep learning* a partir del subagrupamiento.

Este bloque de código automatiza la extracción de un subconjunto altamente específico del *corpus* documental, compuesto por aquellos artículos del subgrupo 0 que comparten

una orientación metodológica centrada en técnicas de aprendizaje profundo. Esta operación se enmarca dentro de una fase avanzada del análisis exploratorio, donde los grupos previos han sido subdivididos para identificar líneas temáticas más refinadas.

El *script* accede al archivo “abstract_by_SUB-cluster.xlsx”, ubicado en Google Drive, y establece la columna “Cluster” como índice del *dataset*, permitiendo localizar con precisión el subgrupo de interés. En este caso, se extraen exclusivamente las entradas pertenecientes al subgrupo 0, que previamente fue identificado como núcleo central de la investigación en aprendizaje profundo aplicado a agricultura de precisión.

Una vez localizado el grupo, se seleccionan las columnas fundamentales (“Title”, “Year” y “Abstract”) y se exporta el conjunto de datos resultante en formato .xlsx bajo el nombre “DeepLearning.xlsx”. Este archivo servirá como base estructurada para realizar la revisión sistemática final, con criterios específicos centrados en redes neuronales y técnicas avanzadas de visión artificial.

Este paso asegura que el análisis posterior se construya sobre un corpus temáticamente consistente, filtrado no solo por su afinidad lingüística sino también por su enfoque metodológico, optimizando así la precisión, trazabilidad y profundidad del estudio.

```
from IPython import get_ipython
from IPython.display import display
import pandas as pd
import os
from google.colab import drive

# Montar Google Drive
print("Montando Google Drive...")
drive.mount('/content/drive')
print("Google Drive montado.")

# Especifica la ruta a la carpeta de Google Drive que contiene el archivo .xlsx
# Reemplaza 'Mi Drive/NombreDeTuCarpeta' con la ruta real a tu carpeta
# Asegúrate de que la carpeta y el archivo tengan permisos de lectura
carpeta_drive_lectura = '/content/drive/MyDrive/AÑO 2024-25/2º CUATRIMESTRE/TFG
2.0/Python/1. Dataset/4. SUB-CLUSTER' #@param {type:"string"}

# Especifica el nombre del archivo .xlsx a leer
nombre_archivo_entrada = 'abstract_by_SUB-cluster.xlsx' #@param {type:"string"}
ruta_completa_archivo_entrada = os.path.join(carpeta_drive_lectura, nombre_archivo_entrada)

# Verificar si la carpeta y el archivo existen
if not os.path.exists(ruta_completa_archivo_entrada):
    print(f"Error: El archivo '{ruta_completa_archivo_entrada}' no fue encontrado.")
else:
    print(f"Accediendo al archivo: {ruta_completa_archivo_entrada}")
```

```

try:
    # Leer el archivo .xlsx
    df = pd.read_excel(ruta_completa_archivo_entrada)
    print(f"El archivo '{nombre_archivo_entrada}' leído exitosamente.")

    # Mostrar las primeras filas y columnas para inspección
    print("\nPrimeras filas del DataFrame leído:")
    display(df.head())
    print("\nColumnas del DataFrame leído:")
    print(df.columns)

    # Establecer la columna "Cluster" como índice
    if 'Cluster' in df.columns:
        df.set_index('Cluster', inplace=True)
        print("\nColumna 'Cluster' establecida como índice.")
        print("\nDataFrame con 'Cluster' como índice:")
        display(df.head())

    # Filtrar las filas donde el índice 'Cluster' es igual a 5
    # Usamos .loc para filtrar por el valor del índice
    df_cluster_5 = df.loc[5]

    # Seleccionar solo las columnas "Title", "Year" y "Abstract"
    columnas_deseadas = ['Title', 'Year', 'Abstract']
    # Asegurarse de que las columnas deseadas existan en el DataFrame filtrado
    columnas_existentes = [col for col in columnas_deseadas if col in df_cluster_5.columns]

    if columnas_existentes:
        df_cluster_5_seleccionado = df_cluster_5[columnas_existentes]
        print("\nFilas del Cluster 5 con columnas seleccionadas ('Title', 'Year', 'Abstract'):")
        display(df_cluster_5_seleccionado.head(-10))

    # Especifica la carpeta de destino para el archivo de salida
    carpeta_destino_drive = '/content/drive/MyDrive/AÑO 2024-25/2º CUATRIMESTRE/TFG
2.0/Python/1. Dataset/4. SUB-CLUSTER' #@param {type:"string"}

    # Crear la carpeta de destino si no existe
    if not os.path.exists(carpeta_destino_drive):
        os.makedirs(carpeta_destino_drive)
        print(f"La carpeta '{carpeta_destino_drive}' creada en Google Drive.")

    # Especifica el nombre del archivo de salida
    nombre_archivo_salida = 'DeepLearning.xlsx' #@param {type:"string"}
    ruta_guardado = os.path.join(carpeta_destino_drive, nombre_archivo_salida)

    try:
        # Exportar el DataFrame filtrado y seleccionado a un archivo .xlsx
        # No incluimos el índice en el archivo de salida a menos que sea necesario
        df_cluster_5_seleccionado.to_excel(ruta_guardado, index=False)
        print(f"El DataFrame del Cluster 5 guardado exitosamente en Google Drive:
{ruta_guardado}")
    except Exception as e:
        print(f"Error al guardar el archivo en Google Drive: {e}")
    else:
        print("Error: Ninguna de las columnas 'Title', 'Year', 'Abstract' se encontró en el DataFrame
después de filtrar por Cluster 5.")

else:

```

```
print("Error: La columna 'Cluster' no se encontró en el DataFrame leído.")  
  
except Exception as e:  
    print(f"Error al leer o procesar el archivo '{nombre_archivo_entrada}': {e}")
```

4. Preparación del *dataset* para la aplicación del protocolo PRISMA.

Este último bloque del cuaderno tiene como objetivo la depuración final del *corpus* específico vinculado al *deep learning*, con el fin de adaptarlo a los requisitos metodológicos del protocolo PRISMA 2020. Este paso marca la transición entre la fase exploratoria y el inicio de la revisión sistemática formal, donde los criterios de inclusión y exclusión se aplican de forma explícita y reproducible.

El procedimiento comienza con la carga del archivo “DeepLearning.xlsx”, que contiene los registros previamente filtrados por su afinidad temática. A partir de este conjunto inicial, el código implementa una serie de transformaciones y filtros estructurados:

- Limpieza textual: se eliminan espacios en blanco residuales en campos clave como “*Title*” y “*Abstract*”, evitando errores en los procesos posteriores de análisis lingüístico.
- Eliminación de duplicados: se identifican y descartan entradas repetidas para asegurar la unicidad de cada registro dentro del corpus final.
- Filtrado temporal: se seleccionan únicamente los estudios publicados en los últimos cinco años, lo que garantiza la actualidad del estado del arte.
- Filtrado semántico: se aplica una búsqueda del término “*Deep Learning*” en los títulos de los artículos, restringiendo el conjunto a investigaciones que aborden esta metodología de forma explícita.

Una vez aplicados estos criterios, el conjunto resultante es exportado en formato .xlsx a una carpeta independiente, ya preparado para ser sometido al análisis cualitativo bajo el protocolo PRISMA. Este procedimiento asegura que el corpus final cumpla con estándares de calidad, pertinencia temática y actualidad, y actúa como punto de partida sólido para la posterior revisión manual y extracción de información clave.

```

from IPython import get_ipython
from IPython.display import display
import pandas as pd
import os
from google.colab import drive
from datetime import datetime

# Montar Google Drive
print("Montando Google Drive...")
drive.mount('/content/drive')
print("Google Drive montado.")

# --- Configuración de rutas y nombres de archivo ---

# Especifica la ruta a la carpeta de Google Drive que contiene el archivo de entrada
carpeta_drive_lectura = '/content/drive/MyDrive/AÑO 2024-25/2º CUATRIMESTRE/TFG
2.0/Python/1. Dataset/4. SUB-CLUSTER'
nombre_archivo_entrada = 'DeepLearning.xlsx'
ruta_completa_archivo_entrada = os.path.join(carpeta_drive_lectura, nombre_archivo_entrada)

# Especifica la carpeta de destino para el archivo de salida
carpeta_destino_drive = '/content/drive/MyDrive/AÑO 2024-25/2º CUATRIMESTRE/TFG
2.0/Python/1. Dataset/5. PRISMA'
nombre_archivo_salida = 'dataset_prisma.xlsx'
ruta_guardado = os.path.join(carpeta_destino_drive, nombre_archivo_salida)

# --- Procesamiento del DataFrame ---

# Verificar si el archivo de entrada existe
if not os.path.exists(ruta_completa_archivo_entrada):
    print(f'Error: El archivo '{ruta_completa_archivo_entrada}' no fue encontrado.")
else:
    print(f'Accediendo al archivo: {ruta_completa_archivo_entrada}')
    try:
        # Leer el archivo .xlsx
        df = pd.read_excel(ruta_completa_archivo_entrada)
        print(f"'{nombre_archivo_entrada}' leído exitosamente.")

        # Mostrar información inicial del DataFrame
        print("\nInformación inicial del DataFrame:")
        print(f"Número de filas: {len(df)}")
        print(f"Columnas: {df.columns.tolist()}")

        # 1. Eliminar espacios en blanco de las columnas relevantes (Title, Year, Abstract)
        columns_to_strip = ['Title', 'Abstract']
        for col in columns_to_strip:
            if col in df.columns:
                df[col] = df[col].astype(str).strip()
                print(f'Espacios en blanco eliminados en la columna '{col}'.')
            else:
                print(f'Advertencia: La columna '{col}' no se encontró para eliminar espacios en blanco.")

        # 2. Eliminar filas duplicadas
        filas_antes_duplicados = len(df)
        df.drop_duplicates(inplace=True)
        filas_despues_duplicados = len(df)
        print(f"Número de filas antes de eliminar duplicados: {filas_antes_duplicados}")
        print(f"Número de filas después de eliminar duplicados: {filas_despues_duplicados}")
        print(f"Filas duplicadas eliminadas: {filas_antes_duplicados - filas_despues_duplicados}")

```

```

# 3. Filtrar por los últimos 5 años
if 'Year' in df.columns:
    df['Year'] = pd.to_numeric(df['Year'], errors='coerce')
    df.dropna(subset=['Year'], inplace=True)
    df['Year'] = df['Year'].astype(int)

    año_actual = datetime.now().year
    año_inicio_periodo = año_actual - 5

    filas_antes_filtro_años = len(df)
    df_filtrado_años = df[df['Year'] >= año_inicio_periodo].copy()
    filas_despues_filtro_años = len(df_filtrado_años)

    print(f"\nNúmero de filas antes de filtrar por año: {filas_antes_filtro_años}")
    print(f"Año de inicio del periodo (últimos 5 años): {año_inicio_periodo}")
    print(f"Número de filas después de filtrar por los últimos 5 años: {filas_despues_filtro_años}")
    print(f"Filas eliminadas por el filtro de año: {filas_antes_filtro_años -
filas_despues_filtro_años}")

    df_procesado = df_filtrado_años
else:
    print("Advertencia: La columna 'Year' no se encontró para aplicar el filtro de años.")
    df_procesado = df

# 4. Filtrar por el término "Deep Learning" en el Abstract
if 'Title' in df_procesado.columns:
    filas_antes_filtro_abstract = len(df_procesado)
    # Asegurarse de que la columna 'Abstract' sea de tipo string para usar .str.contains
    df_procesado['Title'] = df_procesado['Title'].astype(str)
    df_filtrado_abstract = df_procesado[df_procesado['Title'].str.contains('Deep Learning',
case=False, na=False)].copy()
    filas_despues_filtro_abstract = len(df_filtrado_abstract)

    print(f"\nNúmero de filas antes de filtrar por 'Deep Learning' en Abstract:
{filas_antes_filtro_abstract}")
    print(f"Número de filas después de filtrar por 'Deep Learning' en Abstract:
{filas_despues_filtro_abstract}")
    print(f"Filas eliminadas por el filtro de Abstract: {filas_antes_filtro_abstract -
filas_despues_filtro_abstract}")

    df_final = df_filtrado_abstract
else:
    print("Advertencia: La columna 'Abstract' no se encontró para aplicar el filtro de término.")
    df_final = df_procesado # Si no hay columna 'Abstract', no se aplica este filtro

# --- Exportar el DataFrame procesado ---

if 'df_final' in locals() and not df_final.empty:
    # Crear la carpeta de destino si no existe
    if not os.path.exists(carpeta_destino_drive):
        os.makedirs(carpeta_destino_drive)
        print(f"\nCarpeta '{carpeta_destino_drive}' creada en Google Drive.")

    try:
        # Exportar el DataFrame a un archivo .xlsx
        df_final.to_excel(ruta_guardado, index=False)
        print(f"\nDataFrame procesado guardado exitosamente en Google Drive: {ruta_guardado}")
    except Exception as e:
        print(f"\nError al guardar el archivo en Google Drive: {e}")
else:

```

```
print("\nEl DataFrame procesado está vacío o no contiene abstracts con 'Deep Learning'. No
hay datos para guardar.")

except Exception as e:
    print(f"Error al leer o procesar el archivo '{nombre_archivo_entrada}': {e}")
```

5. Enriquecimiento del dataset con el número de citas.

Como complemento al preprocesamiento exigido por el protocolo PRISMA, se incorporó un subbloque adicional con el propósito de añadir una columna clave al conjunto de datos: el número de citas por artículo. Esta variable bibliométrica resulta esencial para aplicar el sistema de puntuación personalizado desarrollado en el apartado metodológico, donde se valoró la relevancia científica de los estudios en función de su impacto académico.

El *script* automatiza la lectura e integración de los archivos Excel iniciales que, además de contener metadatos básicos, incluyen la métrica “*Cited by*” correspondiente al número de citas totales por parte de otras publicaciones. Para ello, se implementó un sistema de renombrado flexible que permite unificar nomenclaturas heterogéneas como “Times Cited, All Databases”, garantizando la estandarización del campo.

Una vez localizadas las columnas necesarias, el *script* concatena los distintos archivos, elimina registros incompletos, asegura la correcta tipificación del campo “*Year*” y conserva únicamente aquellos documentos que disponen de todos los metadatos relevantes, incluida la cita. El resultado final es un conjunto enriquecido y estructurado que se exporta como “Número_Citados.xlsx” a una carpeta independiente, destinado a alimentar el sistema de puntuación de la revisión sistemática.

Este paso otorga solidez adicional al análisis comparativo, permitiendo ponderar no solo la adecuación temática de los artículos, sino también su peso relativo en la literatura científica, integrando así una dimensión cuantitativa relevante para la toma de decisiones en la selección final.

```
import pandas as pd
import os
from google.colab import drive

# Montar Google Drive
print("Montando Google Drive...")
drive.mount('/content/drive')
print("Google Drive montado.")
```

```

# Especifica la ruta a tu carpeta de Google Drive
# Reemplaza 'Mi Drive/NombreDeTuCarpeta' con la ruta real a tu carpeta
carpeta_drive = '/content/drive/MyDrive/AÑO 2024-25/2º CUATRIMESTRE/TFG 2.0/Python/1.
Dataset/1. INICIO' #@param {type:"string"}

# Verificar si la carpeta existe
if not os.path.exists(carpeta_drive):
    print(f'Error: La carpeta '{carpeta_drive}' no fue encontrada.")
else:
    print(f'Accediendo a la carpeta: {carpeta_drive}")

# Listar todos los archivos en la carpeta
archivos_en_carpeta = os.listdir(carpeta_drive)

# Convertir archivos .xls a .xlsx
for archivo in archivos_en_carpeta:
    if archivo.endswith('.xls') and not archivo.startswith('~$'): # Ignorar archivos temporales de Excel
        ruta_completa_xls = os.path.join(carpeta_drive, archivo)
        nombre_base, extension = os.path.splitext(archivo)
        nombre_archivo_xlsx = f'{nombre_base}.xlsx'
        ruta_completa_xlsx = os.path.join(carpeta_drive, nombre_archivo_xlsx)

        # Verificar si el archivo .xlsx ya existe para evitar sobrescribir
        if not os.path.exists(ruta_completa_xlsx):
            try:
                print(f'Convirtiendo '{archivo}' a '{nombre_archivo_xlsx}'...")
                df_temp = pd.read_excel(ruta_completa_xls)
                df_temp.to_excel(ruta_completa_xlsx, index=False, engine='openpyxl')
                print(f'Conversión de '{archivo}' completada.")
            except Exception as e:
                print(f'Error al convertir el archivo '{archivo}': {e}")
            else:
                print(f'El archivo '{nombre_archivo_xlsx}' ya existe. Saltando conversión para '{archivo}'.")

# Ahora procedemos a leer los archivos .xlsx (incluyendo los recién convertidos)
archivos_xlsx = [f for f in os.listdir(carpeta_drive) if f.endswith('.xlsx')]

if not archivos_xlsx:
    print("No se encontraron archivos .xlsx en la carpeta especificada.")
else:
    print(f'Archivos .xlsx encontrados (incluyendo los convertidos): {archivos_xlsx}")

# Diccionario de mapeo para los nombres de columnas alternativos
column_mapping = {
    'Article Title': 'Title',
    'Publication Year': 'Year',
    'Abstract': 'Abstract',
    'Times Cited, All Databases': 'Cited by'
}

# Leer cada archivo .xlsx, renombrar columnas y seleccionar solo las necesarias
lista_dataframes = []
required_columns = ['Title', 'Year', 'Abstract', 'Cited by']
for archivo in archivos_xlsx:
    ruta_completa_archivo = os.path.join(carpeta_drive, archivo)
    try:
        df_temp = pd.read_excel(ruta_completa_archivo)
        print(f"'{archivo}' leído exitosamente.")

```

```

        rename_dict = {col: column_mapping[col] for col in column_mapping if col in
df_temp.columns}
        if rename_dict:
            df_temp.rename(columns=rename_dict, inplace=True)
            print(f" Columnas renombradas en '{archivo}': {rename_dict}")

        present_required_columns = [col for col in required_columns if col in df_temp.columns]

        if len(present_required_columns) == len(required_columns):
            df_temp_selected = df_temp[required_columns]
            lista_dataframes.append(df_temp_selected)
            print(f" Columnas {required_columns} seleccionadas y añadidas a la lista desde
'{archivo}'.")
        else:
            missing = [col for col in required_columns if col not in df_temp.columns]
            print(f" Advertencia: Faltan columnas requeridas ({missing}) en '{archivo}' después de
intentar renombrar. Este archivo no se incluirá en la combinación.")

    except Exception as e:
        print(f"Error al leer o procesar el archivo '{archivo}': {e}")

# Combinar múltiples datasets y limpiar registros con campos vacíos
if 'lista_dataframes' in locals() and lista_dataframes:
    combined_df = pd.concat(lista_dataframes, ignore_index=True)

    print(f"\nNúmero de filas antes de eliminar nulos: {len(combined_df)}")

    required_columns = ['Title', 'Year', 'Abstract', 'Cited by']
    combined_df.dropna(subset=required_columns, inplace=True)

    print(f"Número de filas después de eliminar nulos: {len(combined_df)}")

    if 'Year' in combined_df.columns:
        combined_df['Year'] = pd.to_numeric(combined_df['Year'], errors='coerce').astype('Int64')

    print("\nTipos de datos después de convertir 'Year' a int (si 'Year' existe):")
    print(combined_df.dtypes)

    print("\nPrimeras filas del DataFrame combinado y limpio:")
    display(combined_df.head(-10))

else:
    print("Error: lista_dataframes está vacía o no disponible. No se cargaron DataFrames o ninguno
contenía las columnas requeridas.")
    print("Por favor, asegúrese de que la celda anterior se ejecutó correctamente y encontró archivos
Excel con las columnas necesarias (posiblemente con nombres alternativos).")

# Descarga de los archivos a una carpeta de GoogleDrive.
import os

if 'combined_df' in locals() and not combined_df.empty:
    carpeta_destino_drive = '/content/drive/MyDrive/AÑO 2024-25/2º CUATRIMESTRE/TFG
2.0/Python/1. Dataset/6. CITADOS'
    if not os.path.exists(carpeta_destino_drive):
        os.makedirs(carpeta_destino_drive)
        print(f"Carpeta '{carpeta_destino_drive}' creada en Google Drive.")

    nombre_archivo_salida = 'Número_Citados.xlsx'
    ruta_guardado = os.path.join(carpeta_destino_drive, nombre_archivo_salida)

```

```
try:
    combined_df.to_excel(ruta_guardado, index=False)
    print(f"\nDataFrame guardado exitosamente en Google Drive: {ruta_guardado}")
except Exception as e:
    print(f"\nError al guardar el archivo en Google Drive: {e}")
else:
    print("\nNo hay DataFrame combinado para guardar en Drive.")
```



Anexo III - Bibliografía válida, excluida por criterios normativos.

Este anexo reúne los estudios que, pese a haber superado los criterios de inclusión metodológica establecidos en la revisión sistemática, fueron excluidos del corpus final debido a la limitación normativa que restringe a 35 el número máximo de fuentes principales en el Trabajo de Fin de Grado. Todos los trabajos listados aquí fueron evaluados como relevantes y rigurosos.

- Ahmad, A., Saraswat, D., & El Gamal, A. (2023). A survey on using deep learning techniques for plant disease diagnosis and recommendations for development of appropriate tools. *Smart Agricultural Technology*, 3, 100083-100083. <https://doi.org/10.1016/J.ATECH.2022.100083>
- Alzadjali, A., Alali, M. H., Veeranampalayam Sivakumar, A. N., Deogun, J. S., Scott, S., Schnable, J. C., & Shi, Y. (2021). Maize Tassel Detection From UAV Imagery Using Deep Learning. *Frontiers in Robotics and AI*, 8, 600410-600410. <https://doi.org/10.3389/FROBT.2021.600410>
- Bouguettaya, A., Zarzour, H., Kechida, A., & Taberkit, A. M. (2023). A survey on deep learning-based identification of plant and crop diseases from UAV-based aerial images. *Cluster Computing*, 26(2), 1297-1317. <https://doi.org/10.1007/S10586-022-03627-X>
- Chen, J. S., Liu, M., Li, J. S., & Chen, J. X. (2024). LFA-YOLO: DETECTION OF LYCHEE FRUIT ANTHRACNOSE BASED ON UAV IMAGES AND DEEP LEARNING. *Applied Engineering in Agriculture*, 40(5), 515-523. <https://doi.org/10.13031/AEA.16017>
- Du, S., Yang, Y., Yuan, H., & Cheng, M. (2024). Application of deep learning for real-time detection, localization, and counting of the malignant invasive weed *Solanum rostratum* Dunal. *Frontiers in Plant Science*, 15, 1486929-1486929. <https://doi.org/10.3389/FPLS.2024.1486929>
- Gonzalo-Martín, C., García-Pedrero, A., & Lillo-Saavedra, M. (2021). Improving deep learning sorghum head detection through test time augmentation. *Computers and Electronics in Agriculture*, 186, 106179-106179. <https://doi.org/10.1016/J.COMPAG.2021.106179>
- Kontogiannis, S., Konstantinidou, M., Tsioukas, V., & Pikridas, C. (2024). A Cloud-Based Deep Learning Framework for Downy Mildew Detection in Viticulture Using Real-Time Image Acquisition from Embedded Devices and Drones. *Information 2024, Vol. 15, Page 178, 15(4)*, 178-178. <https://doi.org/10.3390/INFO15040178>

- Kuswidiyanto, L. W., Noh, H. H., & Han, X. (2022). Plant Disease Diagnosis Using Deep Learning Based on Aerial Hyperspectral Images: A Review. *Remote Sensing* 2022, Vol. 14, Page 6031, 14(23), 6031-6031. <https://doi.org/10.3390/RS14236031>
- Linero-Ramos, R., Parra-Rodríguez, C., Espinosa-Valdez, A., Gómez-Rojas, J., & Gongora, M. (2024). Assessment of Dataset Scalability for Classification of Black Sigatoka in Banana Crops Using UAV-Based Multispectral Images and Deep Learning Techniques. *Drones* 2024, Vol. 8, Page 503, 8(9), 503-503. <https://doi.org/10.3390/DRONES8090503>
- Liu, W., Zhou, J., Wang, B., Costa, M., Kaeppler, S. M., & Zhang, Z. (2022). IntegrateNet: A Deep Learning Network for Maize Stand Counting From UAV Imagery by Integrating Density and Local Count Maps. *IEEE Geoscience and Remote Sensing Letters*, 19. <https://doi.org/10.1109/LGRS.2022.3186544>
- Menshchikov, A., Shadrin, D., Prutyaynov, V., Lopatkin, D., Sosnin, S., Tsykunov, E., Iakovlev, E., & Somov, A. (2021). Real-Time Detection of Hogweed: UAV Platform Empowered by Deep Learning. *IEEE Transactions on Computers*, 70(8), 1175-1188. <https://doi.org/10.1109/TC.2021.3059819>
- Narimani, M., Pourreza, A., moghimi, A., B Mesgaran, M., Farajpoor, P., & jafarbiglu, H. (2024). Drone-based multispectral imaging and deep learning for timely detection of branched broomrape in tomato farms., *13053*, 16-25. <https://doi.org/10.1117/12.3021219>
- Nazeer, I., Umer, S., Rout, R. K., & Tanveer, M. (2024). Artificial intelligence-based smart agricultural systems for saffron cultivation with integration of Unmanned Aerial Vehicle imagery and deep learning approaches. *Computers and Electrical Engineering*, 119, 109542-109542. <https://doi.org/10.1016/J.COMPELECENG.2024.109542>
- Niu, S., Nie, Z., Li, G., & Zhu, W. (2024). Multi-Altitude Corn Tassel Detection and Counting Based on UAV RGB Imagery and Deep Learning. *Drones* 2024, Vol. 8, Page 198, 8(5), 198-198. <https://doi.org/10.3390/DRONES8050198>
- Sanaeifar, A., Guindo, M. L., Bakhshipour, A., Fazayeli, H., Li, X., & Yang, C. (2023). Advancing precision agriculture: The potential of deep learning for cereal plant head detection. *Computers and Electronics in Agriculture*, 209, 107875-107875. <https://doi.org/10.1016/J.COMPAG.2023.107875>
- Shahi, T. B., Xu, C. Y., Neupane, A., & Guo, W. (2023). Recent Advances in Crop Disease Detection Using UAV and Deep Learning Techniques. *Remote Sensing* 2023, Vol. 15, Page 2450, 15(9), 2450-2450. <https://doi.org/10.3390/RS15092450>
- Slimani, H., Mhamdi, J. E., & Jilbab, A. (2024). Deep Learning Structure for Real-time Crop Monitoring Based on Neural Architecture Search and UAV. *Brazilian*

Archives of Biology and Technology, 67, e24231141-e24231141.
<https://doi.org/10.1590/1678-4324-2024231141>

Valicharla, S. K., Karimzadeh, R., Naharki, K., Li, X., & Park, Y. L. (2024). Detection and Multi-Class Classification of Invasive Knotweeds with Drones and Deep Learning Models. *Drones*, 8(7), 293-293.

<https://doi.org/10.3390/DRONES8070293>

Wiggers, K. L., Pohlod, C. D., Orlovski, R., Ferreira, R., & Santos, T. A. (2022). Detection and counting of plants via deep learning using images collected by RPA. *Revista Brasileirade Ciencias Agrarias*, 17(2).

<https://doi.org/10.5039/AGRARIA.V17I2A1353>

Zhao, J., Yan, J., Xue, T., Wang, S., Qiu, X., Yao, X., Tian, Y., Zhu, Y., Cao, W., & Zhang, X. (2022). A deep learning method for oriented and small wheat spike detection (OSWSDet) in UAV images. *Computers and Electronics in Agriculture*, 198, 107087-107087.

<https://doi.org/10.1016/J.COMPAG.2022.107087>

Zhao, R., & Shi, F. (2022). A novel strategy for pest disease detection of Brassica chinensis based on UAV imagery and deep learning. *International Journal of Remote Sensing*, 43(19-24), 7083-7103.

<https://doi.org/10.1080/01431161.2022.2155082>

Zhu, H., Lin, C., Liu, G., Wang, D., Qin, S., Li, A., Xu, J. L., & He, Y. (2024). Intelligent agriculture: Deep learning in UAV-based remote sensing imagery for crop diseases and pests detection. *Frontiers in Plant Science*, 15, 1435016-1435016.

<https://doi.org/10.3389/FPLS.2024.1435016>