



Aprendizaje profundo para la detección de malezas en sistemas agrícolas y forestales mesoamericanos*

Deep learning for weed detection in Mesoamerican agricultural and forest systems

Mario Guevara-Bonilla¹, Ramón León-González², Robin Gómez-Gómez³

* Recepción: 29 de julio, 2026. Aceptación: 21 de octubre, 2026. Este trabajo formó parte del proyecto de tesis doctoral de Mario Guevara-Bonilla, realizado en el Doctorado de Ciencias Agrícolas de la Universidad de Costa Rica.

¹ Instituto Tecnológico de Costa Rica. Cartago, Costa Rica. maguevara@itcr.ac.cr (autor para correspondencia; <https://orcid.org/0000-0003-3305-9132>).

² Universidad Estatal de Carolina del Norte. Raleigh, Estados Unidos. rleon@ncsu.edu (<https://orcid.org/0000-0002-1924-3331>).

³ Universidad de Costa Rica, Estación Experimental Agrícola Fabio Baudrit Moreno. Alajuela, Costa Rica. robin.gomezgomez@ucr.ac.cr (<https://orcid.org/0000-0002-8543-8137>)

Resumen

Introducción. El control de malezas es una actividad permanente en sistemas agrícolas y forestales tropicales. El aprendizaje profundo ofrece posibilidades para su detección, identificación y manejo. **Objetivo.** Describir los principales avances y aplicaciones del aprendizaje profundo para la identificación y localización de malezas y analizar su aplicabilidad como herramienta de apoyo para el manejo de malezas en sistemas agrícolas y forestales de Mesoamérica. **Desarrollo.** Se realizó una revisión de literatura científica mediante la búsqueda de artículos publicados entre 2016 y 2025 en Scopus, Web of Science, ScienceDirect y SpringerLink. La revisión se basó principalmente en 15 estudios que aplicaron técnicas de aprendizaje profundo y reportaron métricas de desempeño, y 15 artículos de revisión. Se identificaron los componentes del proceso y se propuso una ruta metodológica para su implementación en Mesoamérica, considerando las particularidades de los sistemas agrícolas y forestales tropicales. Los modelos de aprendizaje profundo han mostrado un desempeño favorable para el reconocimiento de malezas en diversos sistemas agrícolas. No se encontraron estudios en plantaciones forestales en la búsqueda realizada, aunque se identificó un potencial de aplicación durante las etapas iniciales de establecimiento. La hoja de ruta contempla la priorización de cultivos, especies forestales y malezas de mayor impacto, el desarrollo coordinado de bases de datos y el entrenamiento e implementación de modelos. **Conclusiones.** El aprendizaje profundo presenta avances relevantes para la identificación y localización de malezas. Su implementación en Mesoamérica podría contribuir al manejo mediante estrategias de agricultura y silvicultura de precisión, pero requiere datos representativos, validación en campo e integración con sistemas de apoyo a la toma de decisiones.

Palabras clave: agricultura de precisión, visión por computadora, redes neuronales convolucionales, clasificación de imágenes.



Abstract

Introduction. Weed control is a permanent activity in tropical agricultural and forestry systems. Deep learning offers possibilities for weed detection, identification, and management. **Objective.** To describe the main advances and applications of deep learning for weed identification and localization and analyze its applicability as a supporting tool for weed management in agricultural and forestry systems in Mesoamerica. **Development.** A scientific literature review was conducted by searching for articles published between 2016 and 2025 in Scopus, Web of Science, ScienceDirect, and SpringerLink. The review was primarily based on 15 studies that applied deep learning techniques and reported performance metrics, and 15 review articles. The components of the process were identified, and a methodological route for implementation in Mesoamerica was proposed, considering the particularities of tropical agricultural and forestry systems. Deep learning models have shown favorable performance for weed recognition in various agricultural systems. No studies in forest plantations were found in the search conducted, although potential for application during the initial stages of establishment was identified. The roadmap considers the prioritization of crops, forest species, and weeds with the greatest impact, the coordinated development of databases, and the training and implementation of models. **Conclusions.** Deep learning presents relevant advances for weed identification and localization. Its implementation in Mesoamerica could contribute to management through precision agriculture and forestry strategies but requires representative data, field validation, and integration with decision support systems.

Keywords: precision agriculture, computer vision, convolutional neural networks, image classification.

Introducción

Las malezas son plantas capaces de adaptarse exitosamente a ambientes perturbados por actividades humanas y pueden interferir con los objetivos productivos y de manejo de sistemas agrícolas y forestales (Zimdahl & Basinger, 2024). Estas especies representan uno de los principales obstáculos para el establecimiento, la supervivencia y la productividad de cultivos agrícolas y plantaciones forestales (Rubilar et al., 2018; Zimdahl & Basinger, 2024). Por ello, su identificación es un componente fundamental para orientar las decisiones de manejo y reducir la competencia por recursos.

La identificación de malezas debe complementarse con su localización dentro del área productiva para determinar dónde se concentran y orientar las acciones de manejo. Los métodos convencionales de localización, basados principalmente en la inspección visual y el muestreo directo en campo, pueden proporcionar información precisa, pero demandan tiempo y mano de obra y están sujetos a la subjetividad del observador (Murad et al., 2023). También se han utilizado dispositivos ópticos básicos (cámaras convencionales, medidores de reflectancia, entre otros) para aplicaciones agrícolas; sin embargo, su desempeño puede verse afectado por factores operativos como la intensidad de la radiación solar, la densidad de malezas, la vibración y el polvo. Adicionalmente, estas limitaciones no están asociadas únicamente al sensor, sino también a los métodos tradicionales de análisis de imágenes empleados, lo que restringe su eficacia operativa (López-Granados, 2011).

Una vez identificadas y localizadas las malezas, se pueden seleccionar las estrategias de manejo más apropiadas. Los métodos tradicionales incluyen prácticas manuales (machete o pala), mecánicas (motoguadañas o tractores con implementos de corte) y químicas mediante herbicidas (Hasan et al., 2021). Estas prácticas pueden requerir cantidades importantes de tiempo y mano de obra y, en algunos casos, implicar impactos sobre el suelo, el agua y la biodiversidad (Minogue & Osiecka, 2022). Además, con frecuencia se controla la vegetación sin distinguir entre especies, aun cuando sus características y requerimientos de manejo sean diferentes.

En los últimos años, las técnicas de visión por computadora basadas en aprendizaje profundo (AP) han surgido como herramientas prometedoras para la identificación y localización automatizadas de malezas (Kamilaris & Prenafeta-Boldú, 2018). Antes de su desarrollo, estos procesos se abordaban mediante métodos de procesamiento digital de imágenes basados en características morfológicas (área, diámetro, perímetro y convexidad) y texturales, como la rugosidad y la uniformidad (Herrera et al., 2014; Mursalin & Mesbah-Ul-Awal, 2014). A diferencia de estos enfoques, los modelos de AP no requieren la definición previa de características específicas, ya que aprenden patrones complejos que integran formas, texturas y relaciones espaciales, manteniendo un alto desempeño incluso bajo condiciones variables de iluminación y cobertura vegetal (LeCun et al., 2015).

Pese a los avances, la aplicación del AP para la detección automática de malezas en Mesoamérica sigue siendo limitada. El desarrollo de sistemas robustos enfrenta desafíos relacionados con la variabilidad ambiental y la disponibilidad de datos anotados de calidad en los sistemas tropicales de la región. Estos datos, que contienen la identificación o delimitación de las malezas en imágenes, son fundamentales para entrenar y evaluar los modelos de AP (Bah et al., 2018; Milioto et al., 2017).

Los pocos estudios disponibles que evalúan modelos de reconocimiento de malezas con datos de campo locales se enfocan en un número reducido de cultivos y especies de malezas (Garibaldi-Márquez, Valentín-Coronado et al., 2025). Adicionalmente, la mayoría de los desarrollos y validaciones de sistemas de detección automática provienen de regiones templadas, cuyas condiciones ambientales, tipos de cultivo y comunidades de malezas difieren de las presentes en Mesoamérica. Esta situación limita la transferencia directa de estos enfoques y evidencia la necesidad de analizar su aplicabilidad en los sistemas productivos de la región.

La evidencia disponible sobre AP para la detección de malezas en Mesoamérica se encuentra dispersa entre diferentes cultivos, especies, tipos de imágenes y modelos. Además, la diversidad de condiciones productivas y ambientales de la región dificulta establecer la aplicabilidad de estos enfoques a partir de estudios aislados. En este contexto, resulta pertinente integrar la evidencia disponible sobre el uso de estas tecnologías en sistemas agrícolas y forestales mesoamericanos.

Esta revisión busca identificar en qué medida los enfoques desarrollados en otras regiones pueden aplicarse en contextos tropicales y, cuando sea necesario, señalar las adaptaciones metodológicas y tecnológicas que permitirían una implementación más efectiva en la realidad local. Asimismo, se pretende ofrecer un marco conceptual y terminológico que facilite la comprensión y el uso contextualizado de estas tecnologías. Para ello, se planteó como objetivo describir los principales avances y aplicaciones del aprendizaje profundo para la identificación y localización de malezas y analizar su aplicabilidad como herramienta de apoyo para el manejo de malezas en sistemas agrícolas y forestales de Mesoamérica.

Metodología

Esta revisión analizó y sintetizó los avances en la aplicación de técnicas de AP para la detección y clasificación de malezas, como paso previo a la toma de decisiones para su manejo, con especial énfasis en su aplicabilidad en sistemas agrícolas y forestales de la región mesoamericana. Se realizó una búsqueda de artículos científicos en Scopus, Web of Science, ScienceDirect y SpringerLink, considerando publicaciones en inglés y español publicadas entre 2016 y 2025. Se emplearon combinaciones de palabras clave relacionadas con “deep learning”, “weed detection”, “computer vision” y “precision agriculture”.

Se estableció el periodo 2016-2025 para abarcar la etapa reciente de desarrollo y aplicación del AP en la detección y clasificación de malezas. Para ello, se tomó como referencia la consolidación del AP como enfoque de inteligencia artificial (IA) para el procesamiento de imágenes, documentada por LeCun et al. (2015). Esta delimitación permitió concentrar la revisión en desarrollos recientes y aplicaciones contemporáneas.

Los criterios de inclusión consideraron estudios que: (1) aplicaran metodologías basadas en AP para la detección o clasificación de malezas, (2) reportaran métricas de desempeño del modelo, y (3) describieran condiciones de adquisición de datos o el contexto de aplicación. Se excluyeron trabajos enfocados exclusivamente en la clasificación general de cultivos o sin validación experimental clara. Las investigaciones seleccionadas fueron analizadas de forma cualitativa, considerando variables como el tipo de arquitectura utilizada (p. ej., redes convolucionales, detectores de objetos y transformadores), las métricas de evaluación reportadas, las condiciones de captura de imágenes, el tipo de cultivo y el contexto geográfico. Este análisis permitió identificar patrones en el desempeño de los modelos, sus principales limitaciones y su potencial de transferencia a sistemas tropicales.

La búsqueda bibliográfica permitió identificar 30 artículos científicos que constituyeron la base principal de la revisión. De estos, 15 correspondieron a estudios que aplicaron técnicas de AP para la detección, identificación, clasificación o segmentación de malezas, mientras que los otros 15 fueron artículos de revisión utilizados para sintetizar los principales avances, arquitecturas y aplicaciones del AP en este campo. Adicionalmente, se incorporaron referencias metodológicas y conceptuales para sustentar los fundamentos del AP, las arquitecturas empleadas, las métricas de evaluación y el contexto del manejo de malezas.

La inteligencia artificial y el aprendizaje profundo en la detección de malezas

La visión por computadora (VC) se centra en el análisis automático de imágenes y videos digitales con el objetivo de extraer información visual relevante para la identificación y localización de objetos en distintos entornos. En este contexto, el término “visión” no se refiere a un proceso biológico basado en ojos, sino al procesamiento computacional de información visual representada como matrices bidimensionales de valores de píxeles. La VC emplea algoritmos matemáticos y estadísticos para analizar las variaciones espaciales y permitir la identificación y localización de objetos animados e inanimados, como personas, plantas, árboles o infraestructuras, comúnmente mediante cuadros delimitadores y categorías predefinidas (Xiao et al., 2020).

La IA complementa y potencia las capacidades de la VC, con aplicaciones prometedoras en la agricultura y la forestería, particularmente en la identificación de malezas (Amend et al., 2019; Costello et al., 2022). La IA comprende técnicas computacionales diseñadas para resolver problemas mediante el reconocimiento de patrones y la optimización, sin reproducir procesos cognitivos o de toma de decisiones propios de los seres humanos (Russell & Norvig, 2021). Dentro de este campo, el aprendizaje automático (AA) desarrolla algoritmos capaces de aprender a partir de datos sin programación explícita (Goodfellow et al., 2016; Souza Silva et al., 2024) y ha mostrado resultados prometedores en la localización de malezas (Benos et al., 2021; Fernandes et al., 2025; Islam et al., 2021).

El AP es una subcategoría del AA basada en arquitecturas de redes neuronales, que permite a los algoritmos identificar patrones complejos a partir de grandes volúmenes de datos (Souza Silva et al., 2024), como colecciones extensas de imágenes de malezas. Entre sus enfoques más utilizados destacan las redes neuronales convolucionales (CNN, por sus siglas en inglés), fundamento de los sistemas modernos de visión por computadora. Estas redes procesan datos con estructura espacial mediante capas convolucionales que aplican filtros para detectar automáticamente patrones visuales, como bordes, texturas o formas. A través de sus capas, las CNN aprenden representaciones jerárquicas que permiten reconocer objetos o características específicas en una imagen (LeCun et al., 2015).

La aplicación de CNN ha permitido desarrollar herramientas para la identificación y localización de malezas y posee un alto potencial de integración con sistemas de AA y plataformas robóticas (Patel & Kumbhar, 2016). En este contexto, el proceso general para identificar y localizar malezas de manera automática mediante técnicas de AP suele organizarse en cuatro etapas principales: (1) adquisición de imágenes, (2) preparación del conjunto de datos, que incluye la anotación de elementos de interés, (3) extracción de características y construcción del modelo, y (4) evaluación y despliegue. Estas etapas se apoyan en herramientas algorítmicas e infraestructura computacional,

como unidades de procesamiento gráfico (GPU, por sus siglas en inglés), que permiten procesar eficientemente grandes volúmenes de datos visuales (Shanmugam et al., 2020).

Adquisición de imágenes

Las técnicas de detección y clasificación de malezas mediante AP requieren una cantidad suficiente de datos etiquetados. Estos datos pueden obtenerse mediante diferentes sensores y plataformas de captura (Hasan et al., 2021). En general, existen tres fuentes principales: imágenes capturadas directamente, colecciones de imágenes disponibles y, más recientemente, imágenes generadas de forma sintética.

La adquisición directa de imágenes se realiza principalmente mediante cámaras digitales de alta resolución, que pueden operar de forma independiente o integrarse en diversas plataformas. Entre las plataformas más empleadas se encuentran los vehículos aéreos no tripulados (VANT), los robots terrestres y los vehículos todoterreno (ATV, por sus siglas en inglés). Las cámaras RGB (bandas del espectro rojo, verde y azul) constituyen la fuente predominante de datos, aunque algunos estudios utilizan sensores de profundidad, cámaras multispectrales e hiperspectrales, o combinan varios tipos de información para mejorar la capacidad de detección y clasificación.

El segundo enfoque consiste en el uso de imágenes provenientes de bases de datos públicas para el entrenamiento de modelos de AA. Deng et al. (2024) identificaron 36 bases de datos públicas de imágenes de malezas; sin embargo, ninguna fue desarrollada específicamente en Mesoamérica y solo una correspondió a América Latina. Existen bases de datos y repositorios botánicos en Mesoamérica, pero estos recursos no necesariamente fueron desarrollados como conjuntos de imágenes anotadas específicamente para entrenar y evaluar modelos de AP. Por tanto, la disponibilidad de información botánica en la región no implica necesariamente la existencia de conjuntos de datos adecuados para el desarrollo y validación de estos modelos.

Debido a que los modelos para la detección de malezas requieren conjuntos de datos amplios y representativos, es necesario recurrir a estrategias complementarias cuando la captura directa y los repositorios públicos no aportan suficientes imágenes. Una alternativa prometedora es la generación de imágenes sintéticas con apariencia realista y ajustadas a criterios específicos del usuario (Sapkota et al., 2022). Este enfoque consiste en recortar segmentos de fotografías que contienen la maleza de interés y superponerlos sobre distintos fondos, produciendo composiciones artificiales que amplían los conjuntos de entrenamiento (Dwibedi et al., 2017; Georgakis et al., 2017). Aunque aún es incipiente, estudios recientes han demostrado su utilidad (Gao et al., 2020; Hu et al., 2021; Sapkota et al., 2022; Skovsen et al., 2019).

Preparación del conjunto de datos

Después de adquirir las imágenes, es necesario prepararlas para su uso en los modelos de AP. Este proceso incluye uniformar su formato y dimensiones, así como identificar y etiquetar los objetos de interés dentro de las imágenes. También pueden aplicarse técnicas de preprocesamiento, aumento de datos y generación de variaciones en las imágenes para ampliar y diversificar los datos disponibles para el entrenamiento (Hasan et al., 2021).

El preprocesamiento de imágenes es una operación inicial que puede consistir en una o varias etapas. Su principal función es proporcionar datos relevantes y de alta calidad para ayudar a la arquitectura de AP a extraer rasgos de una manera más precisa. Entre las operaciones principales de este proceso se encuentran la eliminación de imágenes de mala calidad (de baja resolución, borrosas o desenfocadas, entre otras), la homogeneización de la resolución y las dimensiones de las imágenes, y la eliminación del fondo de la imagen para aislar el objeto de interés (Badgujar et al., 2024).

El proceso de anotación de imágenes puede realizarse con cuatro métodos que difieren en complejidad y precisión: (1) clasificación de la imagen completa, (2) detección de objetos mediante cuadros delimitadores, (3) segmentación semántica y (4) segmentación de instancias (Figura 1). La clasificación de la imagen completa (Figura 1A) asigna una etiqueta global a la imagen sin identificar las áreas específicas que sustentan la predicción, mientras que la detección mediante cuadros delimitadores (Figura 1B) localiza los objetos sin representar su forma real. La segmentación semántica (Figura 1C) clasifica cada píxel como perteneciente a una clase específica (Coleman et al., 2022), y la segmentación de instancias (Figura 1D) diferencia cada objeto individual dentro de una misma categoría (Coleman et al., 2022).

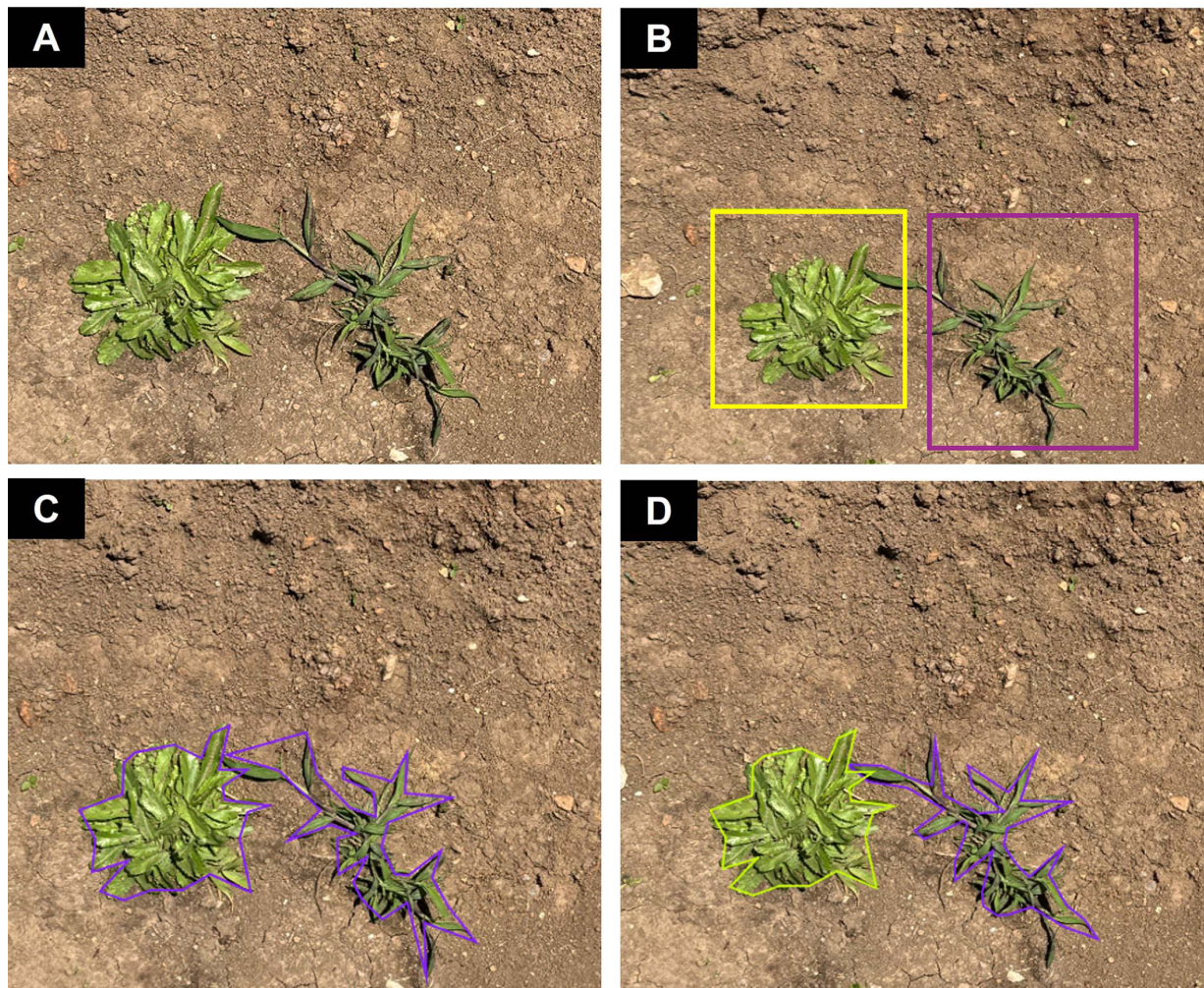


Figura 1. Posibles niveles de clasificación según diferentes arquitecturas de aprendizaje profundo: (A) clasificación de la imagen completa, donde la imagen contiene malezas; (B) detección mediante cuadros delimitadores, donde solo el área dentro de cada cuadro tiene malezas y el resto es fondo; (C) segmentación semántica, la cual es similar a B pero reduce el fondo dentro del área delimitada, y (D) segmentación de instancias, en la cual la clasificación de la maleza abarca todo el objeto y no incluye fondo ni elementos ajenos.

Figure 1. Possible classification levels according to different deep learning architectures: (A) full-image classification, where the image contains weeds; (B) object detection using bounding boxes, where only the area within each box contains weeds and the rest is background; (C) semantic segmentation, which is similar to B but reduces the background within the delimited area; and (D) instance segmentation, in which the weed classification covers the entire object and excludes background or unrelated elements.

Existen plataformas como Roboflow que simplifican y aceleran el preprocesamiento, la anotación de imágenes y los procesos de aumento de datos (Dwyer et al., 2025). Estas utilizan interfaces gráficas y herramientas de anotación asistida por IA, lo cual permite reducir el tiempo de etiquetado y mejorar la consistencia del proceso. Esta ventaja resulta crítica para entrenar modelos robustos en visión por computadora, especialmente porque la mayoría de las investigaciones emplean anotaciones manuales validadas por expertos. No obstante, su uso requiere registro de usuario o una suscripción para acceder a funciones avanzadas, además de conectividad a internet, lo que puede limitar su aplicabilidad en entornos con restricciones tecnológicas o de acceso (Ciaglia et al., 2022).

Construcción del modelo

Las imágenes anotadas se organizan en conjuntos de entrenamiento, validación y prueba. El conjunto de entrenamiento se utiliza para ajustar el modelo, mientras que los conjuntos de validación y prueba permiten evaluar su desempeño con imágenes que no fueron utilizadas para su entrenamiento. Esta separación permite determinar la capacidad del modelo para reconocer malezas en datos nuevos.

El AP puede desarrollarse mediante enfoques supervisados y no supervisados. En el aprendizaje supervisado, el modelo aprende a reconocer las malezas a partir de imágenes previamente etiquetadas, relacionando sus características visuales con las clases correspondientes (Hasan et al., 2021). En el aprendizaje no supervisado, el modelo utiliza imágenes sin etiquetas para identificar patrones o grupos con características similares, que posteriormente pueden relacionarse con diferentes tipos de vegetación o malezas. Este enfoque puede ser útil cuando se dispone de grandes cantidades de imágenes, pero pocas están etiquetadas.

El método de aprendizaje determina los datos requeridos y la estrategia de entrenamiento. La eficacia del modelo también depende de la arquitectura utilizada, porque esta define cómo se procesan las características de las imágenes y se realiza la detección de objetos (Badgujar et al., 2024). Las arquitecturas más utilizadas en visión por computadora pueden agruparse en tres categorías: detectores de dos etapas, detectores de una etapa y modelos basados en transformadores de visión (Kaur & Singh, 2023).

El enfoque de detección de objetos en dos etapas separa los procesos de localización y clasificación. Inicialmente, el modelo genera propuestas de regiones que probablemente contengan objetos; posteriormente, cada región propuesta es clasificada en una categoría específica. Entre las principales redes se encuentran las redes convolucionales computacionales (R-CNN, por sus siglas en inglés) (Girshick et al., 2014) y la familia R-CNN, que incluye Fast-R-CNN (Girshick, 2015), Faster R-CNN (Ren et al., 2017), Mask-R-CNN (He et al., 2017) y Cascade R-CNN (Cai & Vasconcelos, 2018). La principal ventaja de los detectores de objetos de dos etapas es su alta precisión de detección, mientras que su principal desventaja es su baja velocidad de detección debido a sus mayores requerimientos computacionales.

Los sistemas de detección de objetos de una sola etapa realizan simultáneamente dos tareas: localizar los objetos en una imagen y clasificarlos. Para ello, utilizan redes neuronales profundas de doble canal (D-CNN, por sus siglas en inglés), sin necesidad de dividir el proceso en pasos separados (Kaur & Singh, 2023). Estos sistemas funcionan con una sola pasada a través de la red neuronal, lo que los hace más rápidos y eficientes. Durante esta pasada, la red predice de forma directa todas las ubicaciones de los objetos (cuadros delimitadores) y las probabilidades de que pertenezcan a determinadas clases. Es decir, convierten directamente los píxeles de la imagen en coordenadas de los objetos y en etiquetas de clasificación.

Dentro de este grupo destacan DetectorNet (Sermanet et al., 2013; Szegedy et al., 2013) y los modelos YOLO (*you only look once*). Este último modelo es un detector de una sola etapa que formula la detección como un problema de regresión, prediciendo en un único paso las coordenadas y categorías de los objetos (Redmon et al., 2016). Su arquitectura permite una optimización integral y gran velocidad (Shetty & Jogi, 2019); sin embargo,

al basarse en un conjunto limitado de regiones candidatas y en características globales de la imagen, puede comprometer la precisión frente a métodos basados en regiones más detalladas (Liu et al., 2020).

Recientemente, han surgido arquitecturas basadas en transformadores de visión (ViT) como una alternativa a las CNN para la detección de objetos. A diferencia de las CNN, que extraen características locales mediante convoluciones, los transformadores de visión modelan relaciones globales en la imagen mediante mecanismos de autoatención (Reedha et al., 2022). Este enfoque tiene la capacidad de analizar características espaciales y relaciones globales dentro de las imágenes, lo que lo hace adecuado para la clasificación de malezas en entornos agrícolas (Khan et al., 2022).

Un ejemplo reciente de este enfoque es la aplicación de transformadores de detección (DETR). Este modelo de detección de objetos utiliza transformadores de visión para identificar y localizar objetos directamente a partir de la imagen. A diferencia de los métodos tradicionales basados en CNN, que primero generan regiones de interés y luego refinan las detecciones, DETR utiliza consultas de objetos y mecanismos de atención para identificar directamente qué objetos están presentes y dónde se encuentran. De esta manera, reúne estos pasos en un solo modelo de detección.

Esta característica reduce la complejidad del diseño del sistema y facilita su interpretación conceptual. Sin embargo, estos modelos demandan una alta capacidad de cómputo y tiempos de entrenamiento prolongados, lo que limita su uso en dispositivos de bajo costo o en aplicaciones de campo con recursos computacionales restringidos (Shehzadi et al., 2025). En términos generales, los detectores de dos etapas priorizan la precisión, los de una etapa ofrecen mayor rapidez, mientras que los modelos basados en transformadores introducen un nuevo equilibrio entre simplicidad del proceso y requerimientos computacionales (Zhang et al., 2023).

Métricas de evaluación

Las métricas de evaluación permiten cuantificar el rendimiento de los modelos y comparar sus resultados de manera objetiva (Hand, 2009; Ozcift & Gulten, 2011). En la detección de malezas, son importantes porque los errores de clasificación pueden tener consecuencias agronómicas y económicas (Milioto et al., 2017). Estas métricas permiten determinar con qué grado de acierto el modelo identifica las especies presentes en las imágenes.

Las principales métricas de clasificación se obtienen a partir de la matriz de confusión, que compara las predicciones del modelo con las observaciones reales. Estas observaciones corresponden a las etiquetas de referencia asignadas previamente a las imágenes mediante la identificación de las especies presentes. Su confiabilidad depende de criterios de anotación definidos y de una adecuada identificación de las especies. Los verdaderos positivos (TP) corresponden a los casos en que el modelo identifica correctamente la especie de maleza objetivo, mientras que los falsos positivos (FP) ocurren cuando identifica como maleza una planta que no pertenece a la especie objetivo.

Los falsos negativos (FN) se presentan cuando la maleza está presente en la imagen, pero el modelo no logra detectarla. En contraste, los verdaderos negativos (TN) corresponden a las situaciones en que el modelo identifica correctamente la ausencia de la maleza objetivo. A partir de estos cuatro componentes se calculan diversas métricas que permiten evaluar el desempeño del modelo.

Entre las métricas de clasificación más utilizadas se encuentra la precisión (P), que mide la proporción de verdaderos positivos con respecto al total de predicciones positivas realizadas por el modelo (Ecuación 1). Esta métrica indica qué tan confiables son las detecciones positivas del sistema y resulta especialmente útil cuando existe un número elevado de falsos positivos. Por su parte, el *recall* (R) mide la capacidad del modelo para detectar correctamente los casos positivos presentes en el conjunto de datos (Ecuación 2).

$$P = \frac{TP}{TP + FP} \quad (1)$$

$$R = \frac{TP}{TP + FN} \quad (2)$$

El *recall* es especialmente relevante cuando se busca minimizar la cantidad de malezas que el modelo no logra detectar. De las dos métricas anteriores se deriva el puntaje F1 (F1-score). Este puntaje corresponde a la media armónica entre la precisión y el *recall*, y proporciona una medida equilibrada del desempeño del modelo (Ecuación 3).

$$F1\text{-score} = \frac{2(PR)}{P + R} \quad (3)$$

Otra métrica ampliamente utilizada es la exactitud de clasificación (CA), que representa la proporción total de predicciones correctas realizadas por el modelo (Ecuación 4). Complementariamente, la especificidad (S) mide la capacidad del modelo para identificar correctamente los casos negativos, es decir, aquellos casos en los que la maleza objetivo no está presente en la imagen (Ecuación 5). Junto con estas métricas, el coeficiente de correlación de Matthews (MCC) proporciona una medida global de la calidad de la clasificación binaria al considerar simultáneamente verdaderos positivos, verdaderos negativos, falsos positivos y falsos negativos. A diferencia de la exactitud de clasificación, el MCC proporciona una evaluación más equilibrada del desempeño en conjuntos de datos con clases desbalanceadas, donde una elevada proporción de predicciones correctas puede ocultar errores en la clase minoritaria (Chicco & Jurman, 2020) (Ecuación 6).

$$CA = \frac{TP + TN}{TP + TN + FP + FN} \quad (4)$$

$$S = \frac{TN}{TN + FP} \quad (5)$$

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \quad (6)$$

Métricas de detección de objetos

En la detección de objetos, como la identificación automática de malezas en imágenes, es necesario evaluar tanto la clasificación como la precisión espacial de las detecciones. La intersección sobre unión (IoU) mide la superposición entre la región predicha por el modelo y la región real del objeto (Ecuación 7). Valores altos de IoU indican una mayor coincidencia entre ambas regiones. El promedio de IoU entre las clases evaluadas se denomina media de la intersección sobre unión (mIoU) (Ecuación 8).

$$IoU = \frac{TP}{TP + FP + FN} \quad (7)$$

$$mIoU = \frac{1}{N} \sum_{j=1}^N IoU_j \quad (8)$$

Donde N corresponde al número total de clases evaluadas.

Una métrica ampliamente utilizada en detección de objetos es la precisión promedio media (mAP). Este indicador resume el desempeño global del modelo considerando simultáneamente la precisión y el *recall* a diferentes umbrales de detección (Ecuación 9). Por lo tanto, valores más altos de mAP reflejan un mejor equilibrio entre precisión y sensibilidad del modelo en la detección de objetos.

$$mAP = \frac{1}{n} \sum_{i=1}^n AP_i \quad (9)$$

Donde AP_i corresponde a la precisión promedio de cada clase evaluada y n representa el número total de clases del conjunto de datos.

Métricas de segmentación de imágenes

En modelos de segmentación semántica, cada píxel de la imagen se clasifica en una categoría, por lo que se requieren métricas que evalúen tanto la clasificación como la ubicación espacial de los objetos. La exactitud media por píxel (MPA) se calcula como el promedio de las proporciones de píxeles correctamente clasificados en cada clase (Ecuación 10), mientras que el coeficiente de similitud Dice (DSC) mide la coincidencia entre los píxeles predichos y los píxeles reales del objeto (Ecuación 11). Estas métricas permiten evaluar con mayor detalle qué tan bien el modelo delimita las malezas en la imagen, aspecto importante cuando se busca conocer su distribución o estimar el área que ocupan.

$$MPA = \frac{1}{k+1} \sum_{i=0}^k \frac{p_{ii}}{\sum_{j=0}^k p_{ij}} \quad (10)$$

Donde p_{ii} representa el número de píxeles correctamente clasificados para la clase i , p_{ij} representa el número de píxeles de la clase i clasificados como j , y k corresponde al número de clases.

$$DSC = \frac{2TP}{2TP + FP + FN} \quad (11)$$

Uso del aprendizaje profundo para la localización de malezas en sistemas agrícolas y forestales de Mesoamérica

Los cultivos más frecuentemente estudiados a nivel mundial con sistemas de AP para la detección de malezas son maíz (*Zea mays* L.), soya (*Glycine max* (L.) Merr.), remolacha azucarera (*Beta vulgaris* L.) y trigo (*Triticum aestivum* L.). Otros cultivos analizados con menor frecuencia incluyen papa (*Solanum tuberosum* L.), girasol (*Helianthus annuus* L.) y maní (*Arachis hypogaea* L.) (Murad et al., 2023). En cuanto a las malezas más estudiadas, los análisis mostraron una fuerte recurrencia de especies de amplia distribución agrícola, como *Chenopodium album* L. Bosc ex Moq., *Amaranthus retroflexus* L., *Digitaria sanguinalis* (L.) Scop., *Cyperus esculentus* L., *Cirsium arvense* (L.) Scop., *Sinapis arvensis* L., *Portulaca oleracea* L. y *Stellaria media* (L.) Vill. (Garibaldi-Márquez, Valentín-Coronado et al., 2025; Murad et al., 2023).

Las especies anteriores son representativas tanto de malezas de hoja ancha como de gramíneas, lo que ha permitido evaluar la eficacia de arquitecturas de una y dos etapas y de transformadores de visión convolucionales (p. ej., YOLO y Faster R-CNN) en escenarios de alta variabilidad morfológica. En términos de distribución geográfica, los países que lideraron la investigación fueron China, Estados Unidos, India, Alemania y España. Estos países concentraron la mayoría de los estudios con imágenes de campo y sistemas automatizados (Murad et al., 2023).

La literatura disponible sobre el uso de arquitecturas de AP para la detección de malezas en Latinoamérica es aún limitada y se concentra principalmente en estudios realizados en México, Brasil y Costa Rica (Cuadro 1). Estos estudios han evidenciado una amplia diversidad de arquitecturas de AP aplicadas a la detección y localización de malezas, con predominio de modelos de segmentación semántica, segmentación de instancias y detección de objetos. Entre las arquitecturas más utilizadas destacan variantes de U-Net, Mask R-CNN, YOLO, Faster R-CNN y modelos basados en transformadores visuales, como Swin-UNet, Segmenter y SegFormer. Todos los trabajos evaluados utilizaron imágenes capturadas por cámaras RGB o vehículos aéreos no tripulados (VANT).

Cuadro 1. Estudios sobre detección y segmentación de malezas en cultivos agrícolas mediante arquitecturas de aprendizaje profundo en Latinoamérica, publicados entre 2020 y 2025.

Table 1. Studies on weed detection and segmentation in agricultural crops using deep learning architectures in Latin America, published between 2020 and 2025.

Cultivo	País del estudio	Malezas objetivo	Arquitectura de AP	Fuente de imágenes	Métricas de evaluación	Arquitectura de mejor desempeño	Nivel de precisión	Referencia
Maíz	México	Varias	Mask R-CNN Mask R-CNN-ASPP Residual U-Net	Cámara RGB	DSC mIoU	Residual U-Net	93 % 87 %	Garibaldi-Márquez, Flores, & Valentín-Coronado (2025)
Maíz	México	Varias	Swin-UNet Segmenter SegFormer	Cámara RGB	Precisión DSC mIoU	SegFormer	94 % 95 % 91 %	Garibaldi-Márquez, Martínez-Barba et al. (2025)
Maíz	Brasil	Varias	Faster R-CNN YOLOv3 SABL FoveaBox	VANT	Recall F1-score	Faster R-CNN	79 % 76 %	Tetila et al. (2025)
Soya	Brasil	<i>Conyza bonariensis</i> (L.) Cronquist <i>Conyza sumatrensis</i> (Retz.) E. Walker <i>Digitaria insularis</i> (L.) Mez ex Ekman	YOLOv5s6 YOLOv5m6 YOLOv5l6 YOLOv5x6	VANT	MAE RMSE R ²	YOLOv5s6	1,14 1,67 0,93	Tetila et al. (2024)
Frijol	Brasil	Varias	YOLOv8 Mask R-CNN U-Net	VANT	mAP50 Precisión Recall	YOLOv8s	97 % 99,7 % 99 %	Souza Silva et al. (2024)
Maíz y frijol	Costa Rica*	<i>Brassica nigra</i> (L.) W.D.J. Koch <i>Matricaria chamomilla</i> L. <i>Lolium perenne</i> L. <i>Chenopodium album</i> L.	Mask R-CNN	Cámara RGB	Precisión promedio de MSCOCO	Mask R-CNN	50 %	Mora-Fallas et al. (2020)

* Estudio liderado por autores costarricenses; sin embargo, la metodología no especifica el origen geográfico de la base de datos ni si las imágenes fueron obtenidas en cultivos establecidos en Costa Rica. **AP:** Aprendizaje profundo. **Cámara RGB:** Cámara que capta las bandas roja, verde y azul. **VANT:** Vehículo aéreo no tripulado. **mAP:** Precisión promedio media. **mIoU:** Media de la intersección sobre unión. **DSC:** Coeficiente de similitud Dice. **RMSE:** Raíz del error cuadrático medio. **MAE:** Error absoluto medio. **R²:** Coeficiente de determinación. / * Study led by Costa Rican authors; however, the methodology does not specify the geographical origin of the dataset or whether the images were acquired from crops grown in Costa Rica. **AP:** Deep learning. **Cámara RGB:** Camera that captures red, green, and blue bands. **VANT:** Unmanned aerial vehicle. **mAP:** Mean average precision. **mIoU:** Mean intersection over union. **DSC:** Dice similarity coefficient. **RMSE:** Root mean square error. **MAE:** Mean absolute error. **R²:** Coefficient of determination.

Los modelos orientados a la segmentación, como Residual U-Net y SegFormer, han mostrado un buen desempeño en cultivos de maíz en México, con valores de DSC superiores al 90 % y de mIoU entre el 87 % y el 91 % (Garibaldi-Márquez, Flores & Valentín-Coronado, 2025; Garibaldi-Márquez, Martínez-Barba et al., 2025). Por otra parte, las investigaciones que utilizaron modelos de detección como YOLOv8 y YOLOv5 han reportado precisiones superiores al 97 % y coeficientes de determinación cercanos a 0,93 en sistemas agrícolas de Brasil, particularmente con imágenes obtenidas mediante VANT (Souza Silva et al., 2024; Tetila et al., 2024). En contraste, Faster R-CNN presentó valores de *recall* y F1-score cercanos al 79 % y 76 %, respectivamente, en escenarios con múltiples especies de malezas y condiciones de campo más complejas (Tetila et al., 2025).

Las métricas de evaluación empleadas variaron según el tipo de arquitectura y el objetivo del estudio. Se incluyeron indicadores de segmentación (DSC y mIoU), detección de objetos (mAP50, precisión y *recall*) y desempeño predictivo (MAE, RMSE y R^2). En términos generales, los resultados sugieren que las arquitecturas más recientes, particularmente aquellas basadas en transformadores visuales y versiones optimizadas de YOLO, ofrecen ventajas importantes en precisión y capacidad de generalización, aunque su desempeño depende en gran medida de la calidad de las imágenes, el tipo de cultivo, la complejidad estructural de las malezas y las condiciones de adquisición de datos.

La mayoría de estos estudios fueron realizados en 2024 y 2025, lo que refleja un creciente interés en aplicar técnicas de AP para el manejo automatizado de malezas en la región. Sin embargo, aún existe una escasa diversidad de cultivos y contextos geográficos analizados y no se encontraron estudios en plantaciones forestales. Las investigaciones se concentran en condiciones controladas o en etapas iniciales de validación, lo cual evidencia la necesidad de ampliar las investigaciones hacia otros sistemas agrícolas y forestales representativos de la agricultura y forestería tropical mesoamericana.

Perspectivas del uso del AP para la detección de malezas en Mesoamérica

En Mesoamérica, el AP para identificar y localizar malezas puede ser útil en sistemas donde su manejo representa una parte importante de los costos de producción. En agricultura, su aplicación es de interés en cultivos de alto valor y manejo intensivo, así como en sistemas extensivos con arreglos espaciales definidos. La identificación y localización automatizadas podrían facilitar el manejo localizado de malezas y un monitoreo más eficiente de las áreas productivas.

En el sector forestal, la aplicación de modelos de AP podría ser especialmente útil durante las primeras etapas de establecimiento de plantaciones comerciales, cuando la competencia con malezas afecta la supervivencia de los árboles y su crecimiento inicial (Murillo-Gamboa et al., 2024). Su implementación sería más factible en plantaciones de baja densidad (es decir, menos de 1000 árboles por hectárea) y con espaciamientos amplios entre hileras. Estas condiciones facilitan la captura de imágenes y el monitoreo mediante plataformas terrestres o aéreas.

La viabilidad del uso del AP para la identificación y localización de malezas dependerá también de las condiciones productivas de la región. El tamaño y la fragmentación de las fincas, la topografía, el nivel de mecanización, la accesibilidad y la infraestructura tecnológica pueden limitar la adopción de sistemas automatizados. En fincas pequeñas o medianas, las soluciones basadas en teléfonos móviles y cámaras convencionales podrían ser más accesibles y prácticas que sistemas de mayor complejidad.

La mayoría de los modelos de AP para la detección de malezas se ha desarrollado a partir de imágenes RGB obtenidas mediante VANT o plataformas terrestres (Murad et al., 2023). Estos datos son particularmente útiles en sistemas de dosel abierto, como cultivos anuales y plantaciones forestales jóvenes, donde la captura aérea puede ser eficiente y relativamente sencilla (Guevara-Bonilla et al., 2020). En sistemas con dosel cerrado, como plantaciones forestales maduras o cultivos perennes bajo sombra, las imágenes aéreas presentan mayores limitaciones y pueden requerirse plataformas terrestres o portátiles.

La selección del modelo de AP debe considerar los recursos computacionales disponibles y las condiciones de uso. Los modelos rápidos y de menor tamaño pueden facilitar aplicaciones cercanas al tiempo real en contextos con infraestructura limitada (Badgujar et al., 2024). Otros modelos, como U-Net y Mask R-CNN, han alcanzado precisiones superiores al 90-95 % en cultivos agrícolas a partir de imágenes obtenidas con VANT y sensores multispectrales (Souza Silva et al., 2024), aunque la falta de datos locales y de personal capacitado sigue siendo una limitación para su aplicación en Mesoamérica.

Limitaciones del aprendizaje profundo

A pesar de los avances recientes, el desarrollo de aplicaciones de AP para la localización de malezas continúa enfrentando limitaciones relacionadas con la diversidad de especies y la disponibilidad de datos adecuados para el entrenamiento de los modelos. La generación de bases de datos anotadas requiere una correcta identificación de las especies presentes en las imágenes y, por tanto, puede demandar la participación de especialistas, así como tiempo y coordinación interdisciplinaria, particularmente con botánicos y taxónomos. Esta tarea puede ser especialmente compleja en sistemas con una alta diversidad de malezas y cuando existen especies con características morfológicas similares. En este contexto, se han documentado discrepancias entre especialistas durante la anotación de imágenes (Dyrmann et al., 2016).

El desempeño de los sistemas de AP también está condicionado por las características morfológicas de las malezas objetivo. La identificación puede resultar particularmente difícil en especies con características visuales similares, especialmente en las gramíneas durante los primeros estadios de desarrollo, debido a la similitud de la arquitectura foliar entre especies. Esta dificultad puede incrementar la confusión entre clases y afectar el desempeño de los modelos, por lo que la diversidad de especies presentes en los sistemas agrícolas y forestales constituye un aspecto importante para el desarrollo y la evaluación de aplicaciones de AP.

La complejidad del reconocimiento puede aumentar en especies rastreras como los bejucos, debido a la superposición de tallos, el crecimiento horizontal y la fragmentación visual de los individuos, condiciones que dificultan la delimitación de los objetos y su correcta anotación. Estas características pueden ser relevantes en sistemas tropicales y agroforestales, donde se presentan diferentes formas de crecimiento y comunidades de malezas. Asimismo, la variabilidad fenológica y la coexistencia de múltiples especies por unidad de área pueden incrementar la ambigüedad visual y afectar la capacidad de generalización de los modelos.

Muchos estudios utilizan imágenes cenitales capturadas por VANT, pero esta perspectiva puede limitar la identificación de especies con estructuras complejas o que crecen bajo el dosel. Por ello, los conjuntos de imágenes deben representar la diversidad de especies, formas de crecimiento, estadios de desarrollo y condiciones de captura presentes en el sistema de estudio. Las imágenes a nivel del suelo, oblicuas o multiescala pueden aportar información complementaria, mientras que las técnicas de generación sintética, como las redes generativas antagónicas (GAN, por sus siglas en inglés), pueden ampliar la diversidad de los datos de entrenamiento (Ledig et al., 2017).

Retos operativos tras la identificación automática de malezas

Incluso cuando se dispone de sistemas de identificación con alta exactitud y precisión, persisten retos importantes para su uso operativo en los sistemas productivos mesoamericanos. Factores como la topografía irregular, la fragmentación de las fincas, la diversidad de escalas productivas y la limitada mecanización condicionan la integración de estas tecnologías en campo. En agricultura, la variabilidad en los tamaños de las parcelas y los niveles de tecnificación influye en la viabilidad de aplicaciones sitio-específicas, mientras que en el sector forestal los retos se intensifican debido a menores niveles de inversión, menor frecuencia de intervenciones y una ventana de aplicación concentrada en los primeros años del cultivo.

La traducción de mapas de detección en acciones de manejo requiere sistemas de apoyo a la toma de decisiones que consideren umbrales económicos, estrategias de control y compatibilidad con prácticas locales. Sin esta integración, el valor del reconocimiento automático se limita a un ejercicio diagnóstico, lo que reduce su impacto real sobre el manejo integrado de malezas.

Estrategia y hoja de ruta para la adopción del aprendizaje profundo en Mesoamérica

Para avanzar del desarrollo experimental a la aplicación del AP en sistemas productivos mesoamericanos, se requiere una hoja de ruta que integre aspectos técnicos, institucionales y operativos. Además del desempeño de los modelos, su implementación depende de la disponibilidad de datos, infraestructura computacional y personal capacitado, así como de la coordinación entre instituciones. Con base en la evidencia revisada, se propone una hoja de ruta para orientar la incorporación del AP en el reconocimiento de malezas en sistemas agrícolas y forestales de Mesoamérica.

Las dos primeras etapas de la hoja de ruta consisten en la priorización de cultivos y especies de malezas de mayor impacto económico y operativo. En sistemas agrícolas, esta priorización puede orientarse a cultivos extensivos, de alta importancia regional y altamente mecanizados, como café, banano, caña de azúcar y piña. Por otra parte, en sistemas forestales, se debe priorizar la selección de especies de alto valor comercial, como *Tectona grandis* L. f. (teca), *Gmelina arborea* Roxb. (melina) y especies particulares de los géneros *Pinus* y *Eucalyptus*. Adicionalmente, se debe iniciar la implementación en plantaciones recién establecidas, cuando la competencia con malezas es más crítica y las intervenciones son más frecuentes. Este enfoque selectivo permite optimizar recursos y reducir la complejidad del reconocimiento visual.

Estudios realizados en distintos sistemas productivos, incluyendo café (Mora-Delgado & Acosta Arce, 2001; Ricci et al., 2008), banano (Agüero-Alvarado et al., 2018; Rodríguez & Agüero, 2000) y caña de azúcar (León et al., 2017; Rojas et al., 2003), han evidenciado una alta diversidad de arvenses, dentro de la cual múltiples especies pueden comportarse como malezas. Esta heterogeneidad limita la aplicación de enfoques generalistas y resalta la importancia de considerar las particularidades de cada sistema productivo. En este contexto, la segunda etapa del proceso se orienta a la identificación y priorización de las malezas más problemáticas para cada cultivo seleccionado.

La determinación de las especies de malezas más problemáticas puede realizarse mediante encuestas a productores, técnicos y agentes de extensión, así como a través de muestreos sistemáticos de campo, lo cual permite capturar tanto la percepción de importancia agronómica como la diversidad real de las comunidades de malezas (Sarangi & Jhala, 2018). La cooperación interinstitucional es clave para estandarizar criterios y asegurar representatividad regional. Esta definición de especies objetivo constituye el insumo técnico fundamental para orientar de manera eficiente la generación de datos en la etapa posterior.

La tercera etapa corresponde a la generación coordinada de bases de datos de imágenes. Se propone que organismos regionales e internacionales lideren el financiamiento para la creación y mantenimiento de repositorios abiertos de imágenes, así como la provisión de equipamiento básico. Los ministerios del sector agropecuario, institutos forestales y universidades aportarían el personal técnico para la colecta estandarizada de imágenes en condiciones reales de campo, incorporando la variabilidad ambiental, estructural y fenológica propia de los sistemas tropicales. La estandarización de protocolos y la centralización del procesamiento facilitarían el desarrollo de modelos más robustos y su transferencia a distintos contextos productivos.

En una cuarta etapa, la hoja de ruta plantea la centralización del procesamiento y la anotación de imágenes, así como del entrenamiento de modelos, en universidades y centros de investigación, donde se concentran las capacidades técnicas y los recursos computacionales necesarios. Estos espacios deberían conformar grupos interdisciplinarios integrados por botánicos, agrónomos, ingenieros en computación y especialistas en educación, para favorecer tanto la precisión taxonómica como la transferencia efectiva del conocimiento hacia los usuarios finales.

La etapa final se enfoca en la implementación operativa de los sistemas de reconocimiento. En agricultura, los modelos pueden integrarse a plataformas de agricultura de precisión y aplicaciones sitio-específicas. En el sector forestal, donde la mecanización y el financiamiento son más limitados, se prioriza el uso de sensores simples y estrategias de monitoreo de bajo costo. En ambos casos, la validación en campo y la cooperación regional son esenciales para asegurar la sostenibilidad y adopción de estas tecnologías.

Conclusiones

Las técnicas de aprendizaje profundo representan una herramienta con alto potencial para la detección y manejo de malezas en sistemas agrícolas y forestales, debido a su capacidad para automatizar el reconocimiento de especies y procesar grandes volúmenes de imágenes con elevada precisión. En Mesoamérica, estas tecnologías podrían contribuir al desarrollo de estrategias de manejo más eficientes y sostenibles, particularmente mediante aplicaciones localizadas de herbicidas y sistemas de monitoreo más precisos.

La efectividad de los sistemas de reconocimiento está fuertemente influenciada por la calidad y representatividad de los conjuntos de datos utilizados para el entrenamiento. La escasez de bases de datos anotadas para especies tropicales, la alta diversidad florística y la complejidad estructural de los sistemas agrícolas y forestales mesoamericanos continúan siendo limitaciones importantes para el desarrollo de modelos robustos y generalizables. Estas restricciones evidencian la necesidad de fortalecer la colaboración interdisciplinaria entre ingenieros agrónomos, ingenieros forestales, profesionales en computación y especialistas en taxonomía vegetal, así como de promover el desarrollo de repositorios regionales abiertos y estandarizados.

El impacto real del aprendizaje profundo en la agricultura y silvicultura de precisión dependerá no solo de la capacidad de detectar malezas con alta exactitud, sino también de integrar esta información en sistemas de apoyo a la toma de decisiones y estrategias de manejo adaptadas a las condiciones locales. En consecuencia, futuras investigaciones deberían priorizar la validación de modelos en escenarios reales de campo, el desarrollo de soluciones accesibles para productores y la integración operativa entre detección, monitoreo y manejo sitio-específico de malezas.

Agradecimientos

Los autores agradecen al Posgrado en Ciencias Agrícolas y Recursos Naturales del Sistema de Estudios de Posgrado de la Universidad de Costa Rica por el apoyo económico brindado, y al grupo de Silvicultura Intensiva del Instituto Tecnológico de Costa Rica por colaborar en la elaboración de las figuras.

Conflicto de intereses

Los autores declaran que no existen conflictos de intereses relacionados con este estudio.

Referencias

- Agüero-Alvarado, R., Rodríguez-Ruiz, A. M., González-Lutz, M. I., Portuguese-García, P., & Brenes-Prendas, S. (2018). Abundancia y cobertura de arvenses bajo manejo convencional y orgánico de café y banano. *Agronomía Mesoamericana*, 29(1), 85-93. <https://doi.org/10.15517/ma.v29i1.28053>

- Amend, S., Brandt, D., Di Marco, D., Dipper, T., Gässler, G., Höferlin, M., Gohlke, M., Kesenheimer, K., Lindner, P., Leidenfrost, R., Michaels, A., Mugele, T., Müller, A., Riffel, T., Sampangi, Y., & Winkler, J. (2019). Weed management of the future. *KI - Künstliche Intelligenz*, 33(4), 411-415. <https://doi.org/10.1007/s13218-019-00617-x>
- Badgujar, C. M., Poulouse, A., & Gan, H. (2024). Agricultural object detection with You Only Look Once (YOLO) algorithm: a bibliometric and systematic literature review. *Computers and Electronics in Agriculture*, 223, Artículo 109090. <https://doi.org/10.1016/j.compag.2024.109090>
- Bah, M. D., Hafiane, A., & Canals, R. (2018). Deep learning with unsupervised data labeling for weed detection in line crops in UAV images. *Remote Sensing*, 10(11), Artículo 1690. <https://doi.org/10.3390/rs10111690>
- Benos, L., Tagarakis, A. C., Dolias, G., Berruto, R., Kateris, D., & Bochtis, D. (2021). Machine learning in agriculture: a comprehensive updated review. *Sensors*, 21(11), Artículo 3758. <https://doi.org/10.3390/s21113758>
- Cai, Z., & Vasconcelos, N. (2018). Cascade R-CNN: delving into high quality object detection. En D. Forsyth, I. Laptev, A. Oliva, & D. Ramanan (Eds.), *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 6154-6162). IEEE. <https://doi.org/10.1109/CVPR.2018.00644>
- Chicco, D., & Jurman, G. (2020). The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genomics*, 21(1), Artículo 6. <https://doi.org/10.1186/s12864-019-6413-7>
- Ciaglia, F., Zuppichini, F. S., Guerrie, P., McQuade, M., & Solawetz, J. (2022). *Roboflow 100: a rich, multi-domain object detection benchmark* [Preprint]. arXiv. <https://doi.org/10.48550/arXiv.2211.13523>
- Coleman, G. R., Bender, A., Hu, K., Sharpe, S. M., Schumann, A. W., Wang, Z., Bagavathiannan, M. V., Boyd, N. S., & Walsh, M. J. (2022). Weed detection to weed recognition: reviewing 50 years of research to identify constraints and opportunities for large-scale cropping systems. *Weed Technology*, 36(6), 741-757. <https://doi.org/10.1017/wet.2022.84>
- Costello, B., Osunkoya, O. O., Sandino, J., Marinic, W., Trotter, P., Shi, B., Gonzalez, F., & Dhileepan, K. (2022). Detection of parthenium weed (*Parthenium hysterophorus* L.) and its growth stages using artificial intelligence. *Agriculture*, 12(11), Artículo 1838. <https://doi.org/10.3390/agriculture12111838>
- Deng, B., Lu, Y., & Xu, J. (2024). Weed database development: an updated survey of public weed datasets and cross-season weed detection adaptation. *Ecological Informatics*, 81, Artículo 102546. <https://doi.org/10.1016/j.ecoinf.2024.102546>
- Dwibedi, D., Misra, I., & Hebert, M. (2017). Cut, paste and learn: surprisingly easy synthesis for instance detection. En K. Ikeuchi, G. Medioni, M. Pelillo, & E. Mortensen (Eds.), *2017 IEEE International Conference on Computer Vision (ICCV)* (pp. 1310-1319). IEEE. <https://doi.org/10.1109/ICCV.2017.146>
- Dwyer, B., Nelson, J., & Hansen, T. (2025). *Roboflow* (Versión 1.0) [Software]. <https://roboflow.com>
- Dyrmann, M., Karstoft, H., & Midtiby, H. S. (2016). Plant species classification using deep convolutional neural network. *Biosystems Engineering*, 151, 72-80. <https://doi.org/10.1016/j.biosystemseng.2016.08.024>
- Fernandes, P., Pereira da Silva, S. D., Fantinel, R. A., Junior, N. B., de Sousa-Baracho, I. P., Schuh, M. S., Pereira, R. S., & Eugenio, F. C. (2025). Use of hyperspectral sensor and random forest for broadleaf and narrowleaf weed differentiation. *New Forests*, 56(5), Artículo 46. <https://doi.org/10.1007/s11056-025-10113-w>
- Gao, J., French, A. P., Pound, M. P., He, Y., Pridmore, T. P., & Pieters, J. G. (2020). Deep convolutional neural networks for image-based *Convolvulus sepium* detection in sugar beet fields. *Plant Methods*, 16, Artículo 29. <https://doi.org/10.1186/s13007-020-00570-z>

- Garibaldi-Márquez, F., Flores, G., & Valentín-Coronado, L. M. (2025). Leveraging deep semantic segmentation for assisted weed detection. *Journal of Agricultural Engineering*, 56(2), Artículo 1741. <https://doi.org/10.4081/jae.2025.1741>
- Garibaldi-Márquez, F., Martínez-Barba, D. A., Montañez-Franco, L. E., Flores, G., & Valentín-Coronado, L. M. (2025). Enhancing site-specific weed detection using deep learning transformer architectures. *Crop Protection*, 190, Artículo 107075. <https://doi.org/10.1016/j.cropro.2024.107075>
- Garibaldi-Márquez, F., Valentín-Coronado, L. M., Díaz-Ponce, A., Servín-Palestina, M., García-Hernández, R. V., & Ramos-Cantú, L. (2025). Advances on deep learning for proximal image-based weed recognition and control under authentic farmlands: a state-of-the-art review. *Smart Agricultural Technology*, 12, Artículo 101405. <https://doi.org/10.1016/j.atech.2025.101405>
- Georgakis, G., Mousavian, A., Berg, A. C., & Kosecka, J. (2017). Synthesizing training data for object detection in indoor scenes. En N. Amato, S. Srinivasa, N. Ayanian, & S. Kuindersma (Eds.), *Robotics: Science and Systems XIII*. Robotics: Science and Systems Foundation. <https://doi.org/10.15607/RSS.2017.XIII.043>
- Girshick, R. (2015). Fast R-CNN. En R. Bajcsy, G. Hager, Y. Ma, E. Mortensen, & S. Fidler (Eds.), *2015 IEEE International Conference on Computer Vision (ICCV)* (pp. 1440-1448). IEEE. <https://doi.org/10.1109/ICCV.2015.169>
- Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. En R. Basri, C. Fermuller, A. Martinez, & R. Vidal (Eds.), *2014 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 580-587). IEEE. <https://doi.org/10.1109/CVPR.2014.81>
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press. <https://www.deeplearningbook.org/>
- Guevara-Bonilla, M., Meza-Leandro, A. S., Esquivel-Segura, E. A., Arias-Aguilar, D., Tapia-Arenas, A., & Masís-Meléndez, F. (2020). Uso de vehículos aéreos no tripulados (VANTS) para el monitoreo y manejo de los recursos naturales: una síntesis. *Revista Tecnología en Marcha*, 33(4), 77-88. <https://doi.org/10.18845/tm.v33i4.4528>
- Hand, D. J. (2009). Measuring classifier performance: a coherent alternative to the area under the ROC curve. *Machine Learning*, 77(1), 103-123. <https://doi.org/10.1007/s10994-009-5119-5>
- Hasan, A. M., Sohel, F., Diepeveen, D., Laga, H., & Jones, M. G. (2021). A survey of deep learning techniques for weed detection from images. *Computers and Electronics in Agriculture*, 184, Artículo 106067. <https://doi.org/10.1016/j.compag.2021.106067>
- He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask R-CNN. En K. Ikeuchi, G. Medioni, M. Pelillo, & E. Mortensen (Eds.), *2017 IEEE International Conference on Computer Vision (ICCV)* (pp. 2961-2969). IEEE. <https://doi.org/10.1109/ICCV.2017.322>
- Herrera, P. J., Dorado, J., & Ribeiro, A. (2014). A novel approach for weed type classification based on shape descriptors and a fuzzy decision-making method. *Sensors*, 14(8), 15304-15324. <https://doi.org/10.3390/s140815304>
- Hu, C., Sapkota, B. B., Thomasson, J. A., & Bagavathiannan, M. V. (2021). Influence of image quality and light consistency on the performance of convolutional neural networks for weed mapping. *Remote Sensing*, 13(11), Artículo 2140. <https://doi.org/10.3390/rs13112140>
- Islam, N., Rashid, M. M., Wibowo, S., Xu, C.-Y., Morshed, A., Wasimi, S. A., Moore, S., & Rahman, S. M. (2021). Early weed detection using image processing and machine learning techniques in an Australian chilli farm. *Agriculture*, 11(5), Artículo 387. <https://doi.org/10.3390/agriculture11050387>

- 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>
- Kaur, R., & Singh, S. (2023). A comprehensive review of object detection with deep learning. *Digital Signal Processing*, 132, Artículo 103812. <https://doi.org/10.1016/j.dsp.2022.103812>
- Khan, S., Naseer, M., Hayat, M., Zamir, S. W., Khan, F. S., & Shah, M. (2022). Transformers in vision: a survey. *ACM Computing Surveys*, 54(10s), Artículo 200. <https://doi.org/10.1145/3505244>
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436-444. <https://doi.org/10.1038/nature14539>
- Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z., & Shi, W. (2017). Photo-realistic single image super-resolution using a generative adversarial network. En R. Chellappa, Z. Zhang, & A. Hoogs (Eds.), *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 105-114). IEEE. <https://doi.org/10.1109/CVPR.2017.19>
- León, R. G., Agüero, R., & Calderón, D. (2017). Diversity and spatial heterogeneity of weed communities in a sugarcane cropping system in the dry tropics of Costa Rica. *Weed Science*, 65(2), 206-212. <https://doi.org/10.1017/wsc.2016.27>
- Liu, L., Ouyang, W., Wang, X., Fieguth, P., Chen, J., Liu, X., & Pietikäinen, M. (2020). Deep learning for generic object detection: a survey. *International Journal of Computer Vision*, 128(2), 261-318. <https://doi.org/10.1007/s11263-019-01247-4>
- López-Granados, F. (2011). Weed detection for site-specific weed management: mapping and real-time approaches. *Weed Research*, 51(1), 1-11. <https://doi.org/10.1111/j.1365-3180.2010.00829.x>
- Milioto, A., Lottes, P., & Stachniss, C. (2017). Real-time blob-wise sugar beets vs weeds classification for monitoring fields using convolutional neural networks. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, IV-2/W3, 41-48. <https://doi.org/10.5194/isprs-annals-IV-2-W3-41-2017>
- Minogue, P. J., & Osiecka, A. (2022). *Herbicides for weed control in Eucalyptus culture* (FOR310). University of Florida Institute of Food and Agricultural Sciences Extension. <https://doi.org/10.32473/edis-fr378-2013>
- Mora-Delgado, J., & Acosta Arce, L. (2001). Uso, clasificación y manejo de la vegetación asociada al cultivo de café (*Coffea arabica*) desde la percepción campesina en Costa Rica. *Agroforestería en las Américas*, 8(32), 20-27.
- Mora-Fallas, A., Goëau, H., Joly, A., Bonnet, P., & Mata-Montero, E. (2020). Instance segmentation for automated weeds and crops detection in farmlands. *Tecnología en Marcha*, 33(5), 13-17. <https://doi.org/10.18845/tm.v33i5.5069>
- Murad, N. Y., Mahmood, T., Forkan, A. R. M., Morshed, A., Jayaraman, P. P., & Siddiqui, M. S. (2023). Weed detection using deep learning: a systematic literature review. *Sensors*, 23(7), Artículo 3670. <https://doi.org/10.3390/s23073670>
- Murillo-Gamboa, O., Guevara-Bonilla, M., Badilla-Valverde, Y., & Esquivel-Segura, E. (2024). *Manual de cultivo de melina (Gmelina arborea Roxb.) en ciclo corto*. Instituto Tecnológico de Costa Rica.
- Mursalin, M., & Mesbah-Ul-Awal, M. (2014). Towards classification of weeds through digital image. In *2014 Fourth International Conference on Advanced Computing & Communication Technologies* (pp. 1-4). IEEE. <https://doi.org/10.1109/ACCT.2014.101>
- Ozcift, A., & Gulten, A. (2011). Classifier ensemble construction with rotation forest to improve medical diagnosis performance of machine learning algorithms. *Computer Methods and Programs in Biomedicine*, 104(3), 443-451. <https://doi.org/10.1016/j.cmpb.2011.03.018>

- Patel, D., & Kumbhar, B. (2016). Weed and its management: a major threat to crop economy. *Journal of Pharmaceutical Science and Bioscientific Research*, 6(6), 753-758. https://www.jpsbr.org/volume_6/JPSBR_Vol_6_Issue_1_html_files/JPSBR16RS6010.pdf
- Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: unified, real-time object detection. En R. Bajcsy, F.-F. Li, & T. Tuytelaars (Eds.), *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 779-788). IEEE Computer Society. <https://doi.org/10.1109/CVPR.2016.91>
- Reedha, R., Dericquebourg, E., Canals, R., & Hafiane, A. (2022). Transformer neural network for weed and crop classification of high resolution UAV images. *Remote Sensing*, 14(3), Artículo 592. <https://doi.org/10.3390/rs14030592>
- Ren, S., He, K., Girshick, R., & Sun, J. (2017). Faster R-CNN: towards real-time object detection with region proposal networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(6), 1137-1149. <https://doi.org/10.1109/TPAMI.2016.2577031>
- Ricci, M. D. S. F., Virgínio Filho, E. D. M., & Costa, J. R. (2008). Diversidade da comunidade de plantas invasoras em sistemas agroflorestais com café em Turrialba, Costa Rica. *Pesquisa Agropecuária Brasileira*, 43(7), 825-834. <https://doi.org/10.1590/S0100-204X2008000700006>
- Rodríguez, A. M., & Agüero, R. (2000). Identificación de malezas trepadoras del banano (*Musa* sp.) en la zona Caribe de Costa Rica. *Agronomía Mesoamericana*, 11(1), 123-125. <https://doi.org/10.15517/am.v11i1.17395>
- Rojas, L. A., Rodríguez, J. M., Villalobos, H., Arias, M., & Méndez, E. (2003). Malezas asociadas al cultivo de la caña de azúcar. *Tecnología en Marcha*, 16(1), 66-79. https://revistas.tec.ac.cr/index.php/tec_marcha/article/view/1457
- Rubilar, R. A., Allen, H. L., Fox, T. R., Cook, R. L., Albaugh, T. J., & Campoe, O. C. (2018). Advances in silviculture of intensively managed plantations. *Current Forestry Reports*, 4(1), 23-34. <https://doi.org/10.1007/s40725-018-0072-9>
- Russell, S., & Norvig, P. (2021). *Intelligence artificielle: une approche moderne* (L. Miclet, F. Popineau, & C. Cadet, Trans.; 4.^a ed.). Pearson. (Trabajo original publicado en 2020).
- Sapkota, B. B., Popescu, S., Rajan, N., Leon, R. G., Reberg-Horton, C., Mirsky, S., & Bagavathiannan, M. V. (2022). Use of synthetic images for training a deep learning model for weed detection and biomass estimation in cotton. *Scientific Reports*, 12(1), Artículo 19580. <https://doi.org/10.1038/s41598-022-23399-z>
- Sarangi, D., & Jhala, A. J. (2018). A statewide survey of stakeholders to assess the problem weeds and weed management practices in Nebraska. *Weed Technology*, 32(5), 642-655. <https://doi.org/10.1017/wet.2018.35>
- Sermanet, P., Eigen, D., Zhang, X., Mathieu, M., Fergus, R., & LeCun, Y. (2013). *OverFeat: integrated recognition, localization and detection using convolutional networks* [Preprint]. arXiv. <https://doi.org/10.48550/arXiv.1312.6229>
- Shanmugam, S., Assunção, E., Mesquita, R., Veiros, A., & Gaspar, P. D. (2020). Automated weed detection systems: a review. *KnE Engineering*, 5(6), 271-284. <https://doi.org/10.18502/keg.v5i6.7046>
- Shehzadi, T., Hashmi, K. A., Liwicki, M., Stricker, D., & Afzal, M. Z. (2025). Object detection with transformers: a review. *Sensors*, 25(19), Artículo 6025. <https://doi.org/10.3390/s25196025>
- Shetty, J., & Jogi, P. S. (2019). Study on different region-based object detection models applied to live video stream and images using deep learning. En D. Pandian, X. Fernando, Z. Baig, & F. Shi (Eds.), *Proceedings of the International Conference on ISMAC in Computational Vision and Bio-Engineering 2018 (ISMAC-CVB)* (pp. 51-60). Springer. https://doi.org/10.1007/978-3-030-00665-5_6

- Skovsen, S., Dyrmann, M., Mortensen, A. K., Laursen, M. S., Gislum, R., Eriksen, J., Farkhani, S., Karstoft, H., & Jørgensen, R. N. (2019). The GrassClover image dataset for semantic and hierarchical species understanding in agriculture. En L. Davis, P. Torr, S.-C. Zhu, W. Brendel, & M. R. Amer (Eds.), *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)* (pp. 2676-2684). IEEE Computer Society. <https://doi.org/10.1109/CVPRW.2019.00325>
- Souza Silva, J. A. O., Soares de Siqueira, V., Mesquita, M., Rodrigues Vale, L. S., Borges Marques, T. do N., Bezerra da Silva, J. L., da Silva, M. V., Nunes Lacerda, L., de Oliveira-Júnior, J. F., de Lima, J. L. M. P., & Fonseca Elias de Oliveira, H. (2024). Deep learning for weed detection and segmentation in agricultural crops using images captured by an unmanned aerial vehicle. *Remote Sensing*, *16*(23), Artículo 4394. <https://doi.org/10.3390/rs16234394>
- Szegedy, C., Toshev, A., & Erhan, D. (2013). Deep neural networks for object detection. En C. J. C. Burges, L. Bottou, M. Welling, Z. Ghahramani, & K. Q. Weinberger (Eds.), *Advances in neural information processing systems* *26* (pp. 2553-2561). Curran Associates, Inc. https://proceedings.neurips.cc/paper_files/paper/2013/hash/f7cade80b7cc92b991cf4d2806d6bd78-Abstract.html
- Tetila, E. C., Junior, G. W., Higa, G. T. H., da Costa, A. B., Amorim, W. P., Pistori, H., & Barbedo, J. G. A. (2025). Deep learning models for detection and recognition of weed species in corn crop. *Crop Protection*, *195*, Artículo 107237. <https://doi.org/10.1016/j.cropro.2025.107237>
- Tetila, E. C., Moro, B. L., Astolfi, G., da Costa, A. B., Amorim, W. P., de Souza Belete, N. A., & Barbedo, J. G. A. (2024). Real-time detection of weeds by species in soybean using UAV images. *Crop Protection*, *184*, Artículo 106846. <https://doi.org/10.1016/j.cropro.2024.106846>
- Xiao, Y., Tian, Z., Yu, J., Zhang, Y., Liu, S., Du, S., & Lan, X. (2020). A review of object detection based on deep learning. *Multimedia Tools and Applications*, *79*(33-34), 23729-23791. <https://doi.org/10.1007/s11042-020-08976-6>
- Zhang, Y., Wang, M., Zhao, D., Liu, C., & Liu, Z. (2023). Early weed identification based on deep learning: a review. *Smart Agricultural Technology*, *3*, Artículo 100123. <https://doi.org/10.1016/j.atech.2022.100123>
- Zimdahl, R. L., & Basinger, N. T. (2024). *Fundamentals of weed science* (6.^a ed.). Academic Press.