Smart Agricultural Technology 12 (2025) 101147

Contents lists available at ScienceDirect

Smart Agricultural Technology

journal homepage: www.journals.elsevier.com/smart-agricultural-technology

# Weed detection based on deep learning from UAV imagery: A review

Lucía Sandoval-Pillajo [a,b,*], Iván García-Santillán [b], Marco Pusdá-Chulde [b], Adriana Giret [a]

[a] Departamento de Sistemas Informáticos y Computación, Universitat Politècnica de València, Valencia, Spain
[b] Faculty of Engineering in Applied Sciences, Universidad Técnica del Norte, Ibarra, Ecuador

ARTICLE INFO

*Keywords:*
Weed detection
Weed mapping
Deep learning
UAV images
Precision agriculture
SLR

ABSTRACT

Weeds are undesirable plants that compete with crops for essential resources such as light, soil, water, and nutrients. Additionally, they can harbor pests that reduce crop yields. In traditional agriculture, weed control is based on applying pesticides throughout the agricultural field, resulting in soil damage, environmental contamination, damage to farm products, and risks to human health. Precision agriculture (PA) has evolved in recent years thanks to sensors, hardware, software, and innovations in unmanned aerial vehicle (UAV) systems. These systems aim to improve the localized application of chemicals in weed control by using advanced image analysis techniques, computer vision, deep learning (DL), and geo-positioning (GPS) to detect and recognize weeds. This subsequently facilitates the implementation of specific control mechanisms in real environments. Recently, automatic weed detection techniques have been developed using UAV imagery. However, these face a significant challenge due to the morphological similarities between weeds and crops, such as color, shape, and texture, which makes their practical and effective differentiation and implementation difficult. This paper presents a systematic literature review (SLR) based on 77 recent and relevant studies on weed detection and classification in UAV imagery using DL architectures. The analysis focuses on key aspects such as using UAVs and sensors, image acquisition and processing, DL architecture, and evaluation metrics. The review covers publications from 2017 to June 2024 from WoS, Scopus, ScienceDirect, SpringerLink, and IEEE Xplore databases. The results allowed the identification of various limitations, trends, gaps, and opportunities for future research. In general, there is a predominant use of multirotor UAVs, particularly the DJI Phantom with RGB sensors, showing a trend towards the integration of multiple sensors (multispectral, LiDAR) operating at heights of around 10 meters, providing good spatial coverage in data acquisition. Likewise, the rapid development of deep learning architectures has driven CNN models such as ResNet for classification, YOLO for detection, U-Net for semantic segmentation, and Mask R-CNN for weed instance segmentation, with a tendency towards new Transformer-based and hybrid architectures. The most common metrics used to evaluate these models include precision, recall, F1-Score, and mAP.

## 1. Introduction

The world's population has increased rapidly in recent years. According to the Food and Agriculture Organization of the United Nations [1], world population growth is estimated to reach nine billion people by 2050, so the demand for food [2] feed, fiber, and fuels will double [3]. Agricultural production must increase by 70% to meet the growing demand [4]. However, agrarian processes are facing serious challenges such as food security liability, gradual reduction of arable land [4,5], scarcity of water resources, climate change, and the threat of diseases, pests, and weeds [3]. These challenges have triggered the indiscriminate use of synthetic fertilizers and pesticides to increase food production [6].

The noxious incidence of undesirable plants, known as weeds, is among the most important abiotic factors hindering agricultural production worldwide, constituting a permanent problem and causing 25% of crop yield losses [7].

Weeds are plants that grow and can be found scattered all over the agricultural field [8]; they are easily blown away by animals, wind, water, or carried by humans. They do not need special conditions to germinate [9], adapt quickly to any soil type, and reproduce readily. When mature, they can produce thousands to hundreds of seeds that can survive for a long time and pose a significant threat to crops [10].

* Corresponding author at: Departamento de Sistemas Informáticos y Computación, Universitat Politècnica de València, Valencia, Spain.
*E-mail addresses:* ansanpil@upv.edu.es, alsandoval@utn.edu.ec (L. Sandoval-Pillajo).

https://doi.org/10.1016/j.atech.2025.101147
Received 4 April 2025; Received in revised form 20 June 2025; Accepted 29 June 2025
Available online 30 June 2025

Weeds compete with crops for essential resources such as water, nutrients, fertilizers, growing space [11], and sunlight [8,12–14]. Also, they disrupt uniform crop growth, leading to uneven maturity and complicating harvesting processes [15], causing a drastic reduction in crop yield and quality deterioration [14,16]. Therefore, weeds are considered one of the significant obstacles in agricultural production as they lead to economic and social damage to farmers [1]. In addition, some weed species harbor pests and diseases, becoming a refuge and reservoir for pathogenic organisms that can interfere with growth, harvest, and devastate crops [17,18].

Weed control is carried out using various strategies, such as **manual** (use tools such as machetes and scythes); these are very slow, tedious, and costly processes applicable to small plots. **Cultural** (includes crop rotation, sowing distance, and short-cycle crops to reduce proliferation). **Mechanical** (uses agricultural machinery, such as tractors and weeding and hilling techniques). **Chemical** (application of herbicides, fungicides, and pesticides [19] according to weed growth stages (pre-seeding, pre-emergence, post-emergence)). Finally, **biological** (use natural enemies such as insects against certain weeds [4]).

The use of agrochemicals is the most widespread and predominant weed management method. However, their application is often made uniformly throughout the crop without considering the weeds' precise location or spatial distribution [19]. This causes serious pollution problems in soil, air, surface, and groundwater sources, affecting animals, and posing health risks to farmers [20] and threatening food security. Furthermore, indiscriminate use has favored weed resistance to regular herbicides, forcing the use of more expensive and potent compounds [21] applied at higher rates, increasing agronomic problems [15,22], and high costs.

To mitigate these problems, the European Union (EU) has proposed restricting the use of herbicides in agriculture [3] and encouraging the adoption of new technologies, such as Precision Agriculture (PA) [23]. This management strategy uses technological innovations such as sensors, hardware, software, unmanned vehicle systems (UAVs and tractors), remote sensing [24], and geopositioning (GPS) [25], allowing the localized application of chemicals or fertilizers in real-time [26,27]. Furthermore, the integration of Artificial Intelligence, Deep Learning, advanced image analysis, computer vision, and autonomous learning systems [28] facilitates the accurate identification and elimination of weeds [29,30]. This reduces herbicide use, optimizes crop yields [31], and minimizes environmental impact.

In this context, specific methods for site-specific weed management, known as Site-Specific Weed Management (SSWM), have been developed [32,33]. Accurate detection of weeds in the field is essential before applying any precision management technology, significantly when weeds overlap or intersect with the crop [29,34] and ensuring accurate individual location for targeted spraying [35]. This approach has proven effective, achieving 14-39.25% herbicide use reductions in maize fields compared to conventional spraying [36].

PA and weed control methods rely on agricultural image processing and analysis to inspect intra- and inter-crop row conditions, providing key information for efficient management. However, weed detection in outdoor crop fields represents a significant challenge due to multiple factors. The similarity of textures, shapes, and colors between crops and weeds makes their accurate identification difficult [37]. In addition, variability in the appearance of plants of the same species, changes in lighting, foliage occlusion, overlapping leaves [38], and different growth stages [39] further complicate the detection and classification process.

For various plant detection applications, having a more diverse dataset that reflects a wide range of actual field conditions is essential. This includes variations in natural lighting, different scenery, shadows, and leaf occlusion [40]. Incorporating various crops and weeds from different geographical locations, climatic conditions, and growth stages [4]. Having this data available will enable more robust and generalizable deep-learning models to be trained, improving the accuracy of plant identification and weed mapping in UAV-captured imagery [41].

However, the limitations of Deep Learning (DL) architectures, mainly traditional convolutional neural networks, pose significant challenges in precision agriculture, such as the scarcity of labeled datasets [16], the high computational cost for real-time applications, and the need for higher accuracy in weed detection. These limitations have driven the accelerated growth of research proposals on new DL architectures, highlighting the need for periodic systematic reviews to evaluate these emerging technological advances. This undoubtedly justifies the development of the present work.

Thus, this work aims to conduct a systematic literature review (SLR) to analyze the use of various DL architectures in detecting, identifying, and classifying weeds in crops based on images acquired by UAVs. Furthermore, it aims to identify trends, opportunities, challenges, and gaps and suggests directions for future research. This will contribute to developing innovative applications for detecting specific weeds based on advanced platforms and technologies in localized and efficient management. This is aligned with Sustainable Development Goals 12 (responsible production and consumption) and 13 (climate action) [42]. Finally, the most relevant results of the SLR will serve as scientific support for the proposal of more efficient and accurate models for automatic weed detection adapted to the specific conditions of Ecuador, where the adoption of agricultural technology is still limited.

To meet this objective, several research questions related to the following general topics are addressed:

- UAVs and cameras as agricultural image acquisition tools.
- UAV image processing, analysis, and annotation techniques.
- Deep Learning approaches and architectures for automatic weed detection.
- Hardware, software, and cloud services for training and deploying Deep Learning models.
- Performance metrics and graphs for validation of trained models.

The following article is organized as follows: Section 2 presents the description of the systematic SLR literature review methodology. Section 3 presents the results related to the research questions, organized as follows: section 3.1 discusses UAVs and cameras used for crop weed image acquisition. Section 3.2 indicates the processing, labeling, and annotation of UAV images. Section 3.3 reviews DL architecture, hardware, and software used in training and deployment. Section 3.4. The metrics used in the validation of results. Section 4 is the discussion of results, and Section 5 is the conclusions and future work.

## 2. SLR methodology

The present study employs the systematic literature review (SLR), following the guidelines proposed by [43,44], as a methodological framework to obtain updated and relevant information on Deep Learning architectures applied to weed detection from UAV imagery. This section is structured in four phases: (i) research questions, (ii) document search, (iii) article selection, and (iv) extraction of relevant data. Each of them is detailed below:

### 2.1. Research questions

Table 1 presents the main research questions that will guide the SLR of this study. Additionally, related sub-questions are included to provide more detailed answers to each main question, helping to cover all relevant aspects, structure the analysis, and organize the results.

### 2.2. Document search

To collect relevant articles, we applied a search string incorporating key terms related to this study, *such as* weed detection, deep learning, and UAV images. This search string was used across five bibliographic

**Table 1**
Research questions addressed in the SLR.

| Code | Main Research Questions | Sub-questions |
|---|---|---|
| RQ1 | What types of UAVs and cameras have been used for agricultural image acquisition? | What UAVs, cameras, intrinsic and extrinsic parameters, and spectral bands are used?<br>What types of crops and weeds are studied, and in which locations? |
| RQ2 | How have datasets acquired with UAVs been processed, annotated, augmented, and split? | What publicly available image databases exist for research?<br>What software tools are used for UAV image processing and annotation?<br>What strategies are used for dataset creation and augmentation?<br>What data-splitting strategies were applied? |
| RQ3 | What Deep Learning architectures and hardware/cloud services are used for automatic weed detection? | What approaches have been used for weed detection and classification?<br>What Deep Learning architectures are employed?<br>What programming languages, frameworks, hardware, or cloud services are used for model training and deployment? |
| RQ4 | What performance metrics and visualizations are used to validate Deep Learning models? | What evaluation metrics have been applied to assess Deep Learning architectures? |

databases: IEEE Xplore, ScienceDirect, SpringerLink, Scopus, and Web of Science, with slight variations depending on the database requirements.

The search covered the period from 2017 to June 2024, considering that publications in this field were limited before 2016 [4]. Table 2 presents the number of publications on UAV image-based weed detection and classification using DL, showing that ScienceDirect had the highest number of publications in this area (30%), followed by Scopus (24%), Web of Science (18%), SpringerLink (17%), and IEEE Xplore (10%).

### 2.3. Article selection

After conducting the document search in the specified databases, the following exclusion criteria were applied: duplicate articles, articles published in journals without quartiles, or those in quartiles Q3 or Q4. This SLR focuses solely on articles classified as Q1 and Q2, as these are published in higher-impact journals with greater academic rigor, ensuring a solid foundation of scientific evidence. Next, articles unrelated to the topic, objective, and research questions were discarded. Finally, articles whose titles, abstracts, and keywords were irrelevant to the research were removed, leaving 98 articles.

Fig. 1 shows the number of publications per year, where it is evident that 2023 saw a high number of studies, while only a single article was found in 2017 in this field, demonstrating a sustained increase in drone usage in various agricultural applications since 2017. Furthermore, by mid-2024, the number of publications has nearly matched that of the previous year.

Of the 95 papers that passed the document selection phase, 14 papers (15%) are literature reviews, 4 (4%) focus on image dataset preparation, and 77 (81%) are original studies that contribute to answering the research questions posed here. The complete list of reviewed articles is in the supplementary material (Appendix A1).

## 3. Results

The following section presents the results obtained for the four main research questions and the subquestions posed in this study, according to Table 1.

**Table 2**
Publications in Bibliographic Databases with Search Strings and Documents Retrieved (2017 – June 2024).

| N° | Database | Search String | # Docs | % |
|---|---|---|---|---|
| 1 | IEEE Xplore | ("All Metadata":Weed Discrimination) OR ("All Metadata":Weed Management) OR ("All Metadata":Weed Classification) OR ("All Metadata":Weed Detection) AND ("All Metadata":Deep learning) AND ("All Metadata":Unmanned Aerial Vehicle) | 70 | 10% |
| 2 | ScienceDirect | ("Weed discrimination" OR "weed management" OR "weed classification") AND ("Deep learning" OR "Vision Transformers" OR "Deep Neural Network") AND ("Drone" OR "Unmanned aerial vehicle") | 203 | 30% |
| 3 | SpringerLink | ("Weed Detection" OR "Weed Management" OR "Weed Classification" OR "Weed Discrimination") AND ("Deep Learning" OR "Convolutional Neural Networks" OR "CNN") AND ("Drone" OR "Unmanned Aerial Vehicle" OR "UAV" OR "UAS") | 118 | 17% |
| 4 | Scopus | (((weed detection) OR (weed management) OR (weed classification) OR<br>(weed discrimination)) AND ((deep learning) OR (convolutional neural networks) OR (CNN) OR (vision transformers)) AND ((drone) OR (unmanned aerial vehicle) OR (UAV) OR (UAS))) | 165 | 24% |
| 5 | Web of Science | (Weed discrimination OR weed management OR weed classification OR weed detection) AND (Deep learning OR Vision Transformers OR Deep Neural Networks OR CNN OR Convolutional Neural Networks) AND (Drone OR Unmanned aerial vehicle OR UAV OR UAS) | 119 | 18% |
| **Total 675** | | | | **100%** |

### 3.1. RQ1: UAVs and cameras used in agricultural image acquisition

Unmanned Aerial Vehicles (UAVs), also known by their acronyms Unmanned Aerial Systems (UAS or drones), Remotely Piloted Aircraft (RPA), and Remotely Piloted Aircraft Systems (RPAS), have gained popularity due to their versatility in crop monitoring and various agricultural applications [24]. Their ability to fly at low altitudes and capture high-resolution images [45–47] provides advantages over other remote sensing platforms [19], such as satellites [48]. Their low acquisition and maintenance costs, the lack of need for runways, rapid preparation time, and real-time data transmission make them the preferred option for acquiring large agricultural image datasets [49].

#### 3.1.1. What UAVs are being used in agricultural tasks?

According to the 77 selected articles, 87 UAV models have been used for agricultural image acquisition to detect crop weeds. These include multi-rotor UAVs (manufactured by companies such as DJI, 3D Robotics, Pilgrim Technology, Hylio, Onyxstar, and Microdrones) and Fixed-wing UAVs (produced by companies like senseFly, Delair, Applied Aeronautics, Wingtra, and Mugin UAV). Their usage has varied depending on availability and cost.

Among the 87 UAV models identified, their distribution by the manufacturer is as follows: 81% are from DJI, 2% from Hylio, 1% from senseFly, 1% from Pilgrim Technology, 1% from 3DR UAS, 1% from Microdrones, and 9% fall into the "other" category. This last group includes lesser-known manufacturers, custom-built drones, and 4% of drones with no brand preference.

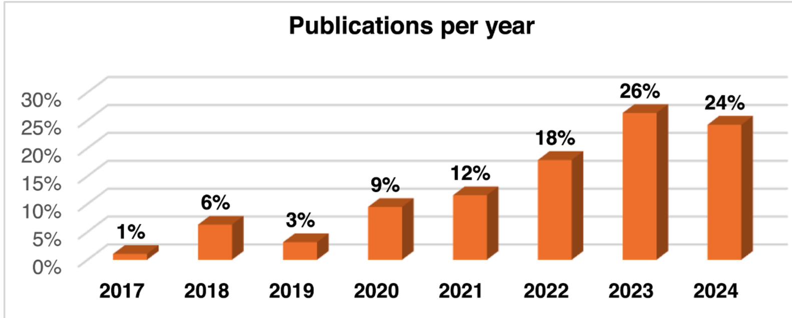


**Fig. 1.** Periodicity of Investigations from 2017 to June 2024.

DJI is the most widely used brand in agricultural research, with 70 UAVs distributed across different models: 47% Phantom, 26% Mavic, 14% Matrice, 7% Inspire, 3% Spark, and 4% without specific model details.

The reviewed sources report on the use of a wide variety of UAVs, highlighting models from the DJI Phantom brand. Although UAVs from the Parrot brand are not directly mentioned, the use of Sequoia cameras, developed by the same company, is recorded. These cameras have been integrated into platforms such as the DJI Mavic Pro [2,16,50–52], Micro Aerial Vehicles (MAV – very small-scale aerial vehicles) [53,54], and the 3DR [55]. Installing these cameras required adaptations such as 3D-printed mounts [54], and external power supplies, with no reported issues in their operation.

Regarding the most used platforms, there is a preference for multi-rotors despite their lower efficiency in long-distance flights compared to fixed-wing UAVs. This is due to the advantage of their propellers mounted around the core/body, allowing vertical takeoff and landing, which requires little space. Additionally, these UAVs are less prone to vibrations and shocks and can obtain high spatial resolution images. Furthermore, they can carry relatively higher payloads [5], such as 15 or 40 kg, in agricultural, industrial, and rescue applications.

#### *3.1.2. What cameras are used for weed detection in crop image acquisition?*

The incorporation of cameras into UAVs has significantly expanded their capabilities and applications. In the context of precision agriculture (PA), weed detection and classification require cameras with sensors that offer good spatial resolution and appropriate spectral sensitivity, particularly in bands of the visible spectrum (RGB), near-infrared (NIR), and Red Edge.

In the reviewed studies, 18 types of cameras equipped with various types of sensors were identified (see Table 3). The most used category corresponds to factory-integrated cameras in UAVs, representing 43% of the total. To a lesser extent, the Parrot Sequoia stands out with 9%, followed by MicaSense RedEdge and Hasselblad at 8%, Sony at 7%, and Zenmuse at 3%. The remaining cameras were reported with low frequencies.

Most cameras incorporated RGB sensors; to a lesser extent, multispectral sensors were used, and the use of hyperspectral cameras was limited. In some cases, these systems were complemented with LiDAR sensors, enabling the integration of spectral information with three-dimensional data of the crop and terrain.

- **RGB sensors** are the most common due to their wide availability, high accuracy, high spatial resolution, high speed, low power consumption, low cost, and ease of operation [56]. They provide images in red, green, and blue color bands. They are useful for calculating vegetation indices such as GRVI, GI, and ExG, object recognition, phenology, yield estimation, and weed mapping [57]. However, their main limitation is that their restricted spectral range prevents precise differentiation of similar materials, making them unsuitable for specialized analyses [58].

**Table 3**
Cameras and sensors are used for image acquisition with UAVs.

| Cameras / Models | Type/Sensor | Type of analysis used |
|---|---|---|
| Parrot Sequoia, MicaSense RedEdge (MX/ MX2/MX3), MAPIR Survey 3 | Multiespectral | Vegetative indices (NDVI, GNDVI), stress detection, crop mapping |
| Hasselblad (L1D-20c), Sony Alpha (6000/ 6300, NEX 5N), Sony ILCE-7R, DJI FC6310/ FC6310R, Fujifilm GFX 100, Zenmuse X5 / X5R / X7, TOP-T10X, Canon IXUS/ ELPH, GoPro Hero 7 | RGB | Photogrammetry, visual mapping, general classification by texture/ colour |
| DJI Zenmuse (P1, X7, algunos modelos) | RGB + NIR | Cover classification, vegetation detection |
| Agrocam | Multiespectral / RGB | Crop monitoring, stress identification |
| MicaSense Altum, Flir Duo R Pro | Multiespectral + Térmica | Plant health + temperature, irrigation/lack of water detection. |
| LiDAR LR1601-IRIS | 3D (Distancia / Luminancia) | 3D modelling of soil and crop structure (crop density parameters, canopy height, aerial biomass, nitrogen uptake). |
| Specim AFX VNIR, Pika L Hyperspectral | Hiperspectral | Accurate weed classification, deep spectral analysis. |

- **Multispectral sensors**, which capture a few dozen bands, allow for assessing plant health, biomass, chlorophyll content, and stress levels. These cameras typically cover the visible (VIS) and near-infrared (NIR) spectra and, in some cases, the shortwave infrared (SWIR). They are widely used to calculate vegetation indices such as NDVI, GNDVI, EVI, NDWI, NDRE, and PRI, which are crucial in agriculture [59]. Additionally, they support applications in water quality assessment and geological analysis, weed detection [54], water stress monitoring, crop growth monitoring [57], and phenotyping [60]. These sensors are more expensive than RGB cameras but still more affordable than hyperspectral sensors. While they capture more spectral bands than RGB cameras, they do not provide the detailed spectral resolution of hyperspectral sensors. However, they require specialized software to process the data and extract relevant information [58].
- **Hyperspectral sensors** capture high-resolution images across dozens or even hundreds of spectral bands. This capability allows for extracting extensive data, facilitating the development of new

vegetation indices tailored to specific crops, regions, or research objectives. These cameras provide in-depth crop analysis, enabling the detection of various pathogens and the precise classification of materials, substances, or conditions [59]. As a result, they are widely applied in geology, advanced precision agriculture, water quality assessment, remote diagnostics, and other fields. Despite their advantages, hyperspectral sensors have notable drawbacks, including high equipment and data processing costs, longer acquisition times, and significant storage requirements [58].

- **LiDAR Sensors:** Lightweight LiDAR can be integrated into small/mini-UAVs [5]. They are used, for example, in forest monitoring [61], and to create digital surface models [31].

Most 87 UAVs analyzed (71%) were equipped with RGB cameras, while 16% used multispectral (MS) sensors. Additionally, 8% carried RGB and MS sensors, and only 1% featured a hyperspectral (HS). The Phantom model was the most widely used UAV, with 56% equipped with RGB sensors, 13% carrying MS sensors, 7% integrating both MS and RGB sensors, and 1% incorporating an HS sensor.

As shown in Fig. 2 (line chart with markers), a clear trend emerges in the choice of UAV brands and camera types (MS, HS, RGB). The data confirms the dominance of DJI drones, notably the Phantom, Mavic, and Matrice models, and the prevalence of RGB sensors in agricultural imaging applications.

Although RGB cameras are popular due to their low cost and availability, there is a clear shift toward multispectral and hyperspectral sensors because of their superior ability to distinguish between weeds and similar crops, even under variable lighting conditions or different phenological stages. Near-infrared (NIR) and Red-Edge bands are particularly useful for differentiating vegetation from soil and detecting herbicide resistance [6,62].

A key trend is the fusion of data from multiple sensors (RGB, multispectral, hyperspectral, thermal, LiDAR), which enhances monitoring accuracy and robustness by combining spectral, structural, and thermal information [63,64].

However, this trend is limited by its high cost. The highly specialized nature of hyperspectral and LiDAR technologies means that their widespread practical application in farmland is still lacking [12]. Additionally, UAV payload limitations restrict the number of sensors that can be mounted simultaneously [65].

### *3.1.3. What are the optimal drone height and speed for capturing high-resolution (cm/px) images and regions of interest (ROIs)?*

The regions of interest (ROIs) are directly influenced by the height and speed of the drone, as well as the characteristics of the cameras used for agricultural image capture. Balancing these parameters (height, speed, and camera specifications) requires careful flight path planning. This involves selecting optimal [66] routes while considering key factors such as air resistance, payload weight, flight speed, and altitude, which are crucial for ensuring precision and efficiency in data acquisition.

A flight speed that is too high can result in the loss of valuable information due to motion blur [67], one of the significant drawbacks of UAVs [6]. Lower-altitude flights provide a lower Ground Sample Distance (GSD), leading to higher spatial resolution and improved accuracy in weed detection. However, flying at low altitudes also means covering smaller areas per flight [67], increased flight time, and higher battery consumption. Therefore, finding a balance between flight altitude, the area to be covered, payload weight, and the desired precision in weed detection is essential.

Spatial resolution is another critical factor, especially for object identification, as it determines the area where individual measurements can be made. It is measured by the Ground Sample Distance (GSD), which refers to the real-world area covered by a single pixel in the image [68]. The primary factors affecting spatial resolution are the type of sensor and the altitude at which the camera, UAV, or sensor operates.

The reviewed studies reported spatial resolutions ranging from 0.3 cm/px at 6m altitude [14,69,70], 0.22 cm/px at 2m [50], 0.5 cm/px at 10m [47,71,72], and 0.0274 mm/px at 4.9m [14,73]. Other resolutions included 1.46, 2, 3.3, 8.2, and 13 cm/px [74–78]. There is no single or "most common" spatial resolution, as it depends on the study objective, sensor type, and flight conditions. However, resolutions between 0.2 cm/px and 8 cm/px were considered optimal in the reviewed studies. For [75,77,78]. Resolutions of 1 to 2 cm/px were deemed sufficient for weed or individual plant detection and classification in agricultural fields.

Altitude plays a crucial role in data acquisition; choosing low altitudes increases flight time to cover agricultural areas. The predominant flight altitudes were 2m, 5m, and 10m. Specifically, 18 UAVs flew at 10m with speeds between 1 m/s and 4.8 m/s; 10 UAVs flew at 5m with speeds between 1 m/s and 3 m/s; 8 UAVs flew at 2m with a speed of 1 m/s. Lower altitudes (2m – 10m) and speeds of 1 m/s and 4.8m/s helped avoid image blur. They enabled the capture of image datasets covering areas between 600m$^2$ and 40,000m$^2$, with resolutions ranging from 2.7 mm/px to 3 cm/px, ensuring complete coverage of study areas.

To optimize UAV route planning while minimizing energy consumption and adapting to dynamic environments [66]. Researchers have explored UAV swarm systems [5] and multi-agent systems (MAS) [61].

- Swarm systems, inspired by biological behavior, excel at simple tasks like large-area exploration [67] with minimal sensors. They adapt to failures and process data using basic algorithms.
- Multi-agent systems (MAS) are more scalable, flexible, and autonomous, integrating advanced technologies like AI, sensors, and distributed communication to coordinate multiple agents, including UAVs, ground stations, robots, simulators, and virtual agents.

Although MAS offers greater robustness and scalability, they face significant challenges in agricultural applications, such as developing intelligent control schemes, ensuring efficient communication between agents, and managing large volumes of real-time data, especially in rural areas [61,66].

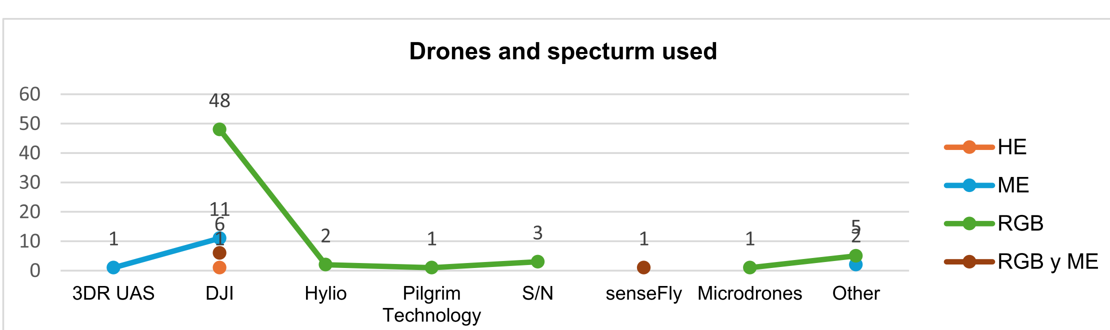


**Fig. 2.** The trend of UAVs and camera types (MS: Multispectral, HS: Hyperspectral, RGB: Red-Green-Blue) used for agricultural data acquisition.

Similarly, UAV-captured images may present variations in inclination/perspective relative to the ground (crop canopy) and different GSD values due to changes in flight altitude, both about mean sea level (MLS) and above ground level (AGL). Therefore, image preprocessing is essential with three key objectives: (1) correcting image distortion, (2) calibrating the GSD for each capture, and (3) accurately georeferencing each crop row [79].

Other factors, such as the time of day, weather, and meteorological conditions, also influence image quality and allow flight planning that minimizes shadows while improving sharpness and precision. In the reviewed studies, the optimal flight times ranged from 11:00 AM to 2:00 PM, coinciding with solar noon. This choice offers several advantages, including reducing long shadows, more uniform lighting, generally more stable atmospheric conditions, and less interference from haze or fog.

However, it is essential to adjust camera settings and carefully plan flights to mitigate potential issues related to overexposure and high temperatures. Additionally, there are freely accessible meteorological stations that provide real-time information, such as Missouri Agricultural Weather (http://agebb.missouri.edu/weather/) [79].

In a nutshell, the selection of the UAVs, sensors, and flight plan for image capture depends on several key factors, including:

- **Type of drone:** Fixed-wing or multirotor, payload capacity (kg).
- **Camera sensors:** RGB, multispectral, hyperspectral; resolution (width × height).
- **Flight plan:** Flight speed (m/s), altitude (m), GSD (cm/px, mm/px), overlap sequence, and planning software.
- **Weather conditions:** No rainfall, clear skies, soft sunlight.

Under these conditions, the images must balance these parameters and the terrain to be captured, always considering the agricultural application and budget constraints.

These limitations highlight the need for continued research and innovation to overcome current challenges, particularly in-flight path optimization and multi-objective modeling, where multiple UAVs work collaboratively to enhance agricultural data collection and monitor efficiency and effectiveness.

#### 3.1.4. *Which crops are being assessed?*

The review of the articles shows that the selection of crops in the research is based on various factors, such as their agricultural importance, economic value, susceptibility to weeds, and the need to optimize management practices [64,80].

According to the analyzed sources, 29 types of crops were evaluated. The most studied were sugar beet (12%), corn (10%), rice (9%), beans (8%), cotton (7%), wheat and grass (6%), and winter wheat (5%). Other crops, including sesame, peas, cereals, spinach, strawberries, soybean seedlings, and sorghum, were studied less frequently. Some were mentioned in only one study, such as potatoes, alfalfa, sugarcane, barley, cabbage, kimchi cabbage, lettuce, peanuts, parsley, tree pepper, pineapple, tomato, tobacco, and cassava.

- **Sugar beet** (*Beta vulgaris*): A high-value industrial crop [73]. Studies mainly focus on weed segmentation to minimize herbicide use and improve control [52].
- **Corn** (*Zea mays*): One of the most studied crops due to its high susceptibility to weed competition [62,64,70]. Research focuses on early weed detection to optimize herbicide application and reduce yield losses [74,80]. A relevant experiment analyzed five corn cycles, examining plant height dynamics and stem thickness to assess competition levels and their impact on yield parameters [81].
- **Rice** (*Oryza sativa*): As a staple food worldwide, its production is constantly threatened by weed presence [82]. Studies on this crop aim to improve early detection to minimize production losses.
- **Bean** (*Phaseolus vulgaris*): A vital source of protein and food. This crop is particularly vulnerable to weed competition, significantly when growing between rows, requiring constant monitoring [83].
- **Cotton** (*Gossypium spp.*): A key crop for textile fiber production, where weed management is crucial for yield [72,79,84]. YOLOv7 models were used to detect cotton plants and evaluate the uniformity of their distribution.
- **Wheat** (*Triticum aestivum*): This staple cereal is affected by weeds [64], and studies aim to develop methods for selective weed detection and efficient herbicide application [8].
- **Grass** (*Cynodon dactylon*): A vital forage crop for livestock, competing with perennial weeds, is not consumed by animals [85].

Despite the focus on various crops, some varieties, such as potatoes, have been the subject of only a single study. This highlights the need to diversify research and explore less-studied crops to achieve more generalized and effective weed management in precision agriculture.

#### 3.1.5. *What weeds are being researched and where?*

Various weed species are being studied due to their impact on crops and the need for more efficient management strategies. The weeds analyzed in the studies vary widely depending on geographical location, environmental conditions, and the type of affected crop.

A total of 71 weed species have been identified across different crops. Among them, 50 species belong to broadleaf plants (dicotyledons, with diverse shapes), 20 species are narrow-leaved plants (monocotyledons, with elongated leaves), and one is generally classified as "weed." The most studied weed species include *amaranthus* (5%, seven studies), *chenopodium album, cirsium arvense, grass*, and *rumex obtusifolius* (4%, five studies each). other notable species with 3% (4 studies) include *convolvulus arvensis, cyperus iria,* and *leptochloa chinensis*.

Studies highlight the viability of Vision Transformers (ViT) for weed recognition, even under significant changes in field image acquisition conditions, such as variations in resolution, lighting, and plant growth stages [86]. Additionally, there is an increasing interest in weed detection among regionally essential crops, with a significant economic and agricultural impact [40].

Research based on UAV-captured images spans 21 countries across five continents, underscoring the global nature of this field. The distribution is as follows: 42% in Europe (Germany, Switzerland, Greece, Italy, United Kingdom, Belgium, Spain, Finland, France, and Norway), 37% in Asia (China, Pakistan, India, South Korea, and Turkey**),** 23% in the Americas (United States, Mexico, Brazil, and Colombia), 5% in Oceania (Australia), 2% in Africa (Nigeria).

The countries with the highest number of studies in this research area are China (25%), Germany (20%), and the United States (17%).

This broad geographic coverage enables the development of more diverse approaches to weed detection and classification.

### 3.2. *RQ2: Processing, annotation, augmentation, and splitting of UAV datasets*

It is crucial to have a robust and diverse training dataset to apply Deep Learning architectures for automatically detecting weeds. However, these architectures have faced challenges due to the limited availability of both the quantity and variety of datasets, particularly aerial images captured by UAVs with precise annotations [87].

Various strategies have been adopted to mitigate this limitation, such as data augmentation techniques. However, even with this approach, the number of labeled images remains insufficient. Currently, dataset creation trends are shifting toward hybrid datasets (combining proprietary images with public datasets), the generation of synthetic data, and transfer learning strategies that enhance the quantity and quality of data available for training models.

#### 3.2.1. Publicly available datasets for training

Although there are online datasets on weeds, their quantity and variety remain limited. Many of these datasets lack detailed classification [72] or do not include geographical information, which affects their applicability in different agricultural environments.

According to the reviewed articles, 32 public datasets on weeds and crops obtained by UAVs have been identified [72]. Of these, 21 datasets come directly from the studies analyzed [6,8,16,36,51,54,69,80,82,94–99,104], while the remaining 11 were found through searches in databases such as Mendeley Dataset and Google Dataset using keywords like "weed and UAV images." All these cases are detailed in Appendix A4.

There are also other studies that offer their data "upon request to the corresponding author" (e.g., [14,40,41,76,88,89], while in other cases, the data are confidential [90] and therefore not publicly available.

Despite the efforts of several authors to share and make their datasets public, availability remains limited in terms of weed species diversity, crop types, growth stages [12,91], environmental conditions (light, wind) [4], and geographic locations [87]. Other limitations include an insufficient dataset size [31], lack of labeled data [72], and class imbalance [92]. As a result, manual data generation remains a critical, costly, and necessary step in most DL-based studies.

Currently, dataset creation trends are shifting toward hybrid datasets (combining proprietary and public images), synthetic data generation (e.g., generative adversarial networks – GANs) [80], data augmentation (e.g., mosaicking or image rotation, [93]. The adoption of multimodal data (integration of multiple sensor data types such as multispectral and LiDAR) [31], transfer learning [94], and semi-automatic labeling [84].

All of this highlights the need to address the limitations, trends, and challenges of public datasets and to overcome these gaps to improve existing models.

#### 3.2.2. What tools have been used for UAV image annotation?

UAV image processing employs multiple tools to facilitate image preparation and annotation, essential for automated detection and classification.

There are multiple annotation techniques, each adapted to different levels of granularity: Pixel-Level Annotation (assigns a label to each pixel); Region-Level Annotation (includes techniques such as bounding boxes annotation and polygon annotation); Image-Level Annotation (assigns a global label to the entire image) [4]; Keypoints Annotation (marks specific points in the image). In addition to these types of annotation, there are advanced approaches for structuring and generating labels, such as Synthetic Labelling (generates synthetic labels), Segmentation masks (annotations of specific regions or instances), Specific region segmentation (assigning pixel-level labels, typically performed by an architecture like U-Net) [92].

Image annotation is a labor-intensive manual process that presents a significant challenge due to its complexity, tedious nature, and high cost, in addition to requiring a substantial time investment. However, it is a crucial step in training DL models for weed identification. Annotating a single image takes approximately 10 h [100], and even for highly trained specialists, the process can be complex, leading to errors due to various factors. These include poor image quality, shadow interference when the image contains multiple classes in similar proportions [101], plant overlap, which may result in various plants being labeled as a single instance [6], or visual fatigue.

The availability of adequately annotated training data remains a bottleneck for developing and applying DL models, highlighting the need for more automated methods for generating labeled data. Table 4 presents the tools used for image annotation and their purpose.

A total of 12 tools were used for image annotation in the original studies, with LabelImg being the most frequently used (33%), followed by LabelMe (10%) and RoboFlow (10%). LabelImg is preferred due to its efficient labeling speed, zoom and panning functions, and ease of use. Additionally, it is an open-source tool compatible with multiple image formats (JPG, PNG, BMP, TIFF). It supports specific formats for DL frameworks such as TensorFlow and PyTorch, making it ideal for object detection and segmentation tasks [101].

Other tools, such as Agisoft Metashape, PIX4D Mapper, and Pynovisao, are not exclusively designed for image annotation but have been used alongside QGIS to label orthomosaics in 3D models and map cropland areas. Similarly, GIMP and Photoshop, although they do not directly generate annotation formats, have been utilized in data preparation.

On the other hand, CVAT and Roboflow are the most advanced and complex tools, but they require more structured input data. In addition to these tools, custom Python scripts (e.g., *improcessing.ipynb* or custom-written code) have also been used [80,82] to generate segmentation masks or to split images into smaller patches for processing.

Image annotation is a stage that demands considerable time and expert human effort. It also faces various challenges, such as data complexity (e.g., similarities between weeds and crops) [78,112]. The identification of very small objects [14], variability in the environment (e.g., lighting; occlusion, overlapping; growth stages) [53,113], annotation quality and consistency (e.g., human error; low-quality images) [103], and data preparation issues (e.g., image size; memory limitations; class imbalance; georeferencing) [77,114], as well as the high cost of

**Table 4**
Tools for annotation and other UAV imagery tasks.

| Type of tool or method | Main capabilities | Tool | License: Free/Commercial | References | Usage Percentage |
|---|---|---|---|---|---|
| Basic annotation with bounding boxes | Drawing of bounding boxes; object detection; export in common formats (YOLO, VOC) | LabelImg | Free and open-source. It can be integrated with MATLAB. | [8,11,40,72,77,86,89,90,100–102] | 32.5% |
| | | Makesense.ai | Proprietary, trial version available for free | [82] | 10% |
| Annotation with pixel segmentation | Manual or semi-automatic drawing of precise masks; useful in semantic segmentation | LabelMe | Free with limited options; paid plan includes advanced features and enterprise support | [103–106] | 10% |
| | | CVAT | Free and open-source (GNU GPL) | [63] | 10% |
| | | GIMP | Free and open-source | [6,74,87,107] | 10% |
| | | Photoshop | Commercial | [14] | 7.5% |
| Photogrammetric or geospatial tools | Annotations on geo-referenced imagery or 3D models; surface or vegetation classification | Agisoft Metashape | Commercial | [41,76,97,108] | 7.5% |
| | | PIX4D | Commercial | [16,31,109] | 2.5% |
| | | ENVI | Free and open-source | [75] | 2.5% |
| Hybrid tools and advanced platforms | Multiple annotation (boxes, masks, dots), integration with automatic or semi-automatic models | Roboflow | Open-source and free | [71,80,110,111] | 2.5% |
| | | Matlab | Commercial | [37,95,97] | 2.5% |
| | | CVAT | Free, but not open-source | [63] | 2.5% |

labor and time.

The time required for UAV image annotation has shown considerable variability, mainly determined by scene complexity, the type of annotation used (e.g., pixel-wise or bounding boxes), and dataset size. In the reviewed articles, one study reported an average of 60 min per multispectral image [54]. Another indicated that the manual annotation of 684 × 456-pixel subimages using Labelme took an average of 6 to 10 min per subimage. Yet another study stated that approximately 10 to 12 h were needed to annotate a couple of *Solanum* images [87]. These findings highlight the significant time and resource demands of the image annotation process.

Annotation represents a major bottleneck in the development of DL models, making the search for more efficient and less manually dependent methods imperative. In this context, current trends are shifting toward models such as the Segment Anything Model (SAM), superpixel algorithms, image synthesis pipelines, semi-supervised learning techniques, and rule-based classifiers. These approaches make the annotation process more efficient, scalable, and less prone to human error, thereby facilitating the availability of properly labeled training data.

#### 3.2.3. Strategies used to create and extend datasets

In the reviewed studies, various strategies have been employed for dataset creation and expansion to enhance image quality, increase data diversity, optimize Deep Learning model training, and address hardware limitations. The most used strategies are detailed below:

- **Data Augmentation**: A widely used set of techniques to increase the size and diversity of the training dataset from originally acquired images [36,71,115]. These techniques improve the model's generalization ability [116] and help reduce overfitting [87]. Standard methods include Geometric transformations [31,56,84,117]. Color adjustments [12,75,86,105,118]. Grid distortion [36]. Mosaic augmentation [40,72]. Negative sampling. Resampling [89].
- **Image preprocessing**: Techniques and operations applied to raw images to enhance [65,101], normalize, or transform them to be more beneficial for classification or segmentation tasks. The most used techniques include Image alignment [54], Radiometric and atmospheric correction [62,75], Image smoothing [65], Contrast enhancement [6,110,118], Vegetation index extraction (NDVI, NGRDI) [51,107], Spatial filtering and thresholding [53].
- **Image resizing and cropping**: Adjusting image size or selecting specific regions within an image [52,79,105,111]. Commonly used methods included Fixed-size resizing [8,111], Interpolation [31, 111], Distortion-free resizing [79], Cropping [78,119], orthomosaics [107].
- **Tiling (Splitting images into Sub-images)**: Segmenting large images into smaller, fixed-size fragments [100,120]. The most used methods included Patch division [14,103,105,116], Sub-image overlapping [100], Sliding windows [34,51], Polygon-based division [75,78], Spatial partitioning (based on plots or geographic areas) [78].
- **Data Synthesis**: The generation of artificial images to complement accurate data. This includes methods such as Synthetic image generation: Creating new images by combining real cropped plants with soil or other backgrounds [51,84], Generative Adversarial Networks (GANs) Used for both weed and crop image generation [107,112], 3D rendering [93], Synthetic labeling (Dynthetic labeling).

These strategies maximize model performance and improve the efficiency of available data, combining accurate and synthetic information to enhance generalization and robustness in detection systems.

#### 3.2.4. Data splitting strategies

Once the data has been prepared, it is essential to define strategies for splitting it into training, validation, and test sets. This is a widely applied practice in machine learning to ensure proper training and an objective evaluation of the model.

These strategies not only help mitigate spatial dependence in the data but also contribute to increasing the diversity and size of the training set, balancing classes when necessary, optimizing model performance, and improving generalization ability. As a result, more accurate models are developed. These strategies enhance the model's precision, robustness, and efficiency, enabling scalable and reliable weed detection in crops. Below are some of the most used methods:

- **Train-validation-test split**: This approach randomly divides the data into two or three sets (training, validation, and test) to train and evaluate the model. The most used split ratios for two sets are 70/30, 80/20, and 90/10 [121], while for three sets, the preferred configurations are 70/15/15, 80/10/10, and the most common 70/20/10 [77,78,114]. The split choice depends on the dataset size, model complexity, and computational resources.
- **Cross-validation**: It is a widely used strategy to assess a model's generalization ability, especially useful in scenarios with limited data availability. Cross-validation is particularly beneficial for simpler models and in contexts where the dataset size does not allow for setting aside a fixed subset for validation or testing [86]. However, this approach involves training the model multiple times, which can become unfeasible in terms of time and computational resources when working with complex models, such as deep neural networks applied to computer vision. In such cases, repeated training can be costly, leading to the use of alternative strategies such as fixed data partitioning (train/validation/test), early stopping techniques, or validation on reduced subsets. Therefore, although cross-validation improves the reliability of the evaluation process, its applicability is constrained by model complexity and available computational capacity [36,86]. Combining these techniques is recommended to obtain more reliable and robust machine learning model development results.

In general, combining these techniques is recommended to obtain more reliable and robust results in the development of machine learning models.

### 3.3. RQ3: deep learning architectures used in weed detection

The detection and classification of weeds in crops are challenging problems that arise at different growth stages, where crops and weeds share one or more similar spatial and morphological characteristics such as shape, size, texture, and color ("green on green"). Differentiating them is inherently complex, even for an expert human. These characteristics can also be affected by factors such as humidity, diseases, and leaf maturity [122], further complicating the discrimination between the two classes.

Additionally, real-field conditions present further challenges, such as variable lighting due to different sun angles and plant shadows caused by natural light, which create illumination and shading effects that make detection and classification even more difficult. The interweaving and overlapping of plants further complicate plant recognition, significantly reducing classification accuracy [123,124]. Other challenging factors include the non-uniform distribution of weeds in the field, geographic location, crop variations, climate, and soil conditions [4].

Given these challenges, high-precision weed management becomes imperative, especially during the early growth stages of the crop [73], when crop plants and weeds look very similar and tend to dominate the terrain [53]. To achieve this, spectral properties, morphology, texture, spatial context, and patterns in digital images can be leveraged [125]. These features enable the differentiation between crops and weeds using advanced machine learning techniques [126], including convolutional neural networks (CNNs) [127], encoder-decoder architectures, autoencoders, transformers, GANs, and hybrid models.

#### 3.3.1. What approaches were used for weed detection and classification?

Various approaches have been used for segmentation and classification in the context of weed detection in crops. These segmentation and classification approaches can be pixel-based, patch-based, or object-based. These methods are widely used in computer vision applications for agricultural image analysis. However, they are not the only ones. With technological advancements, new methods have emerged to complement or improve traditional approaches. Table 5 presents the approaches used in the reviewed studies.

Among the reviewed studies, the most used approach was Pixel-Based, with a frequency of 47%. This method classifies each pixel independently without considering its context, focusing on spectral characteristics such as color, reflectance, and indices like NDVI [75, 114]. It is useful when mixed weeds and crops are difficult to distinguish. However, their accuracy tends to decrease and is limited by several factors, including the presence of mixed pixels (which contain information from more than one class or object), morphological and spectral similarities [75], variable environmental conditions, motion blur [91], subjectivity and labeling bias [50], and noise in high-resolution imagery [128].

The Patch-Based approach, used in 28% of cases, analyzes small regions (patches) instead of individual pixels. This allows for considering each pixel's context, reducing noise, and improving accuracy in weed identification within crops. Additionally, it decreases processing time as it requires fewer trainable parameters [37,52], reducing computational complexity. While the Pixel-Based approach provides more detailed contours of crop and weed classes, it is more computationally expensive [37]. In Deep Learning, the Patch-Based approach employs CNNs such as U-Net, enabling model training and direct classification of image regions. Unlike the Object-Based Image Analysis (OBIA) approach, it does not require prior segmentation, as CNNs process patches directly.

One of the advantages of the patch-based approach is its application in selective spraying, as it allows targeting specific weed zones and prevents unnecessary spraying on soil and crop plants. In contrast, the Pixel-Based approach does not allow for targeted spraying, as the distribution of pixels classified as weeds is random, making precise localization difficult [37,52].

The OBIA approach was used in 25% of the studies. Unlike the previous methods, OBIA analyzes entire image regions instead of individual pixels or patches, considering spectral, textural, morphological, and spatial characteristics [40]. First, the image is segmented into homogeneous areas, and then objects are classified by integrating geometric and spectral information. This makes OBIA more accurate for weed detection, as it incorporates spatial context [58,129]. However, OBIA faces challenges such as object irregularities [120], dependency on expert knowledge (manual creation of rules and thresholds), and its limited applicability in early growth stages, where crops and weeds share very similar characteristics [118].

With Deep Learning, OBIA combines segmentation techniques with convolutional neural networks (CNNs) such as ResNet or YOLO to classify entire objects. While it is suitable for real-time processing, it depends on an initial segmentation (dividing the image into objects) and requires high computational capacity. According to [53] this approach is unsuitable for weed detection in the early growth stages, as crops and weeds share similar characteristics at this phase.

Both object-based and pixel-based classification have shown notable performance in identifying weeds in crops. However, several factors can significantly influence the results, including image resolution, flight altitude, sensor characteristics, camera specifications (optical parameters), spectral similarity between crops and weeds, crop classes with similar properties, dataset quality, and choice of Deep Learning architectures [67].

Additionally, 25% of the studies focused on classification-based approaches, assigning class labels to image patches or detecting weed objects using bounding boxes. The entire image is classified as weed (or specific weed types) or crop. While this approach is faster, it is less precise in determining the exact location of weeds within a crop.

Finally, several studies combined segmentation and classification to detect weeds. They first applied segmentation to separate vegetation from the background and then used classification to distinguish between crops and weeds [86]. However, the method choice depends on the images' complexity and the specific objectives of each study.

Moreover, there is a growing trend toward deep learning (DL) architectures. Models such as FCN, U-Net, and DeepLabV3+ have demonstrated superior performance for end-to-end pixel-wise classification, outperforming traditional manually designed approaches [106, 109]. In patch-based approaches, DL architectures like YOLOv7-tiny and YOLOv8n are being integrated [102]. OBIA remains a dominant method; however, the trend indicates that DL surpasses it in both accuracy and efficiency for complex mapping tasks [120].

**Table 5**
Strategies and Approaches Used in Image Segmentation and Classification.

| Strategy | Approach | Related articles | Description |
|---|---|---|---|
| **Pixel-Based** | Segmentation Pixel-Wise | [6,14,31,41,50, 53,54,70,75,78, 80,82,84,87,95, 100,104–110,113, 114,116,117] | Semantic segmentation can be based on independent pixels. |
| | Classification | [8,14,70,83,95, 103,112,114,116, 120] | Each pixel is classified individually without considering spatial context. |
| **Patch-Based** | Semantic Segmentation | [8,36,37,47,52, 63,64,74,75,79, 81,85,86,97,128] | The image is divided into small patches, each classified as a unit. Models such as U-Net or Fully Convolutional Networks (FCN) are commonly used. |
| | Classification | [2,52,78,86,94, 114] | Patch classification as independent units, using advanced CNN models |
| **Object-Based (OBIA)** | Object-Based Segmentation (Image Segmentation) | [2,40,56,80,94, 102,108,111,115, 118,121,129] | The image is segmented into complete objects before classification. Use segmentation techniques such as Mask R-CNN or U-Net. |
| | Classification | [77,120] | Object classification is based on spectral and spatial features. |

#### 3.3.2. Deep Learning architecture used in weed detection

DL architectures are specific neural network structures designed to address machine learning tasks. Various architectures have been widely used in the reviewed articles for weed detection, semantic segmentation, instance segmentation, and crop classification. The choice of architecture depends on factors such as the specific task, available computational resources, dataset size, required accuracy, efficiency, and processing speed.

Fig. 3 presents the central architecture used for each task. It was observed that for classification, detection, semantic segmentation, and instance segmentation, convolutional neural networks (CNNs) were the most employed, accounting for 91% of the total. Other techniques included unsupervised learning (4%), transformers (3%), hybrid architectures (2%), and generative models (1%), all applied to weed detection using images acquired by UAVs. These results highlight the dominance of CNNs in this field of study (Fig. 3).

Below, the most used pre-trained architectures are presented.

*3.3.2.1. Convolutional neural networks (CNNs).* CNNs are widely used in image processing for classification, segmentation, and weed detection

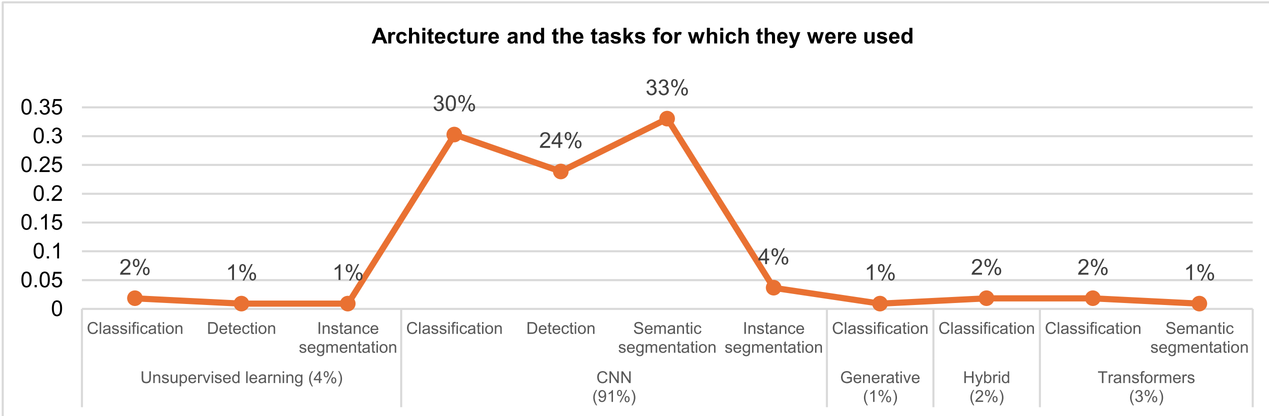


**Fig. 3.** Architecture DL used in the different agricultural tasks: classification, detection, semantic segmentation, and instance segmentation. The complete list of pre-trained architectures can be found in the supplementary material (Appendix A2).

tasks. Unlike traditional methods, these networks combine feature extraction and classification in an automated and efficient manner [4, 72]. Below are the most used pre-trained architectures for each task:

**Classification:** This task assigns a single label to the entire image. The most frequently used architectures include ResNet (various versions) with 15%, with ResNet-18 standing out as the lightest variant [83] and ResNet-101 offering higher accuracy, although increased depth led to overfitting [6,69]. VGGNet (16 and 20 layers) with 5%, although its high computational cost is limited. AlexNet had 3%, followed by less commonly used architectures such as Bionic Convolutional Neural Network, DenseNet, MobileNet, GoogLeNet, and SesameWeedNet-2CNN.

According to the reviewed studies, the innovations introduced by ResNet, such as residual blocks and skip connections, contribute to the effective construction and training of much deeper neural networks without performance degradation.

The ResNet18 model stood out for its efficiency and low computational resource requirements while maintaining high accuracy, making it particularly suitable for integration into UAV systems for real-time evaluations on resource-constrained devices. For instance, it was selected for assessing cotton emergence, demonstrating lower GPU and memory usage compared to many other popular models such as ResNet-50, AlexNet, and VGG (SD7). Another study showed that ResNet18 enabled fast weed mapping in wheat fields, achieving 94% accuracy [130].

ResNet-101 was chosen for its versatility and high accuracy as a backbone in semantic segmentation tasks (e.g., DeepLab) and object detection tasks (e.g., Mask R-CNN and Faster R-CNN) [47,84].

The review also identified several significant challenges that affect the performance of classification models. Chief among them is intra-class similarity, particularly during early growth stages when the spectral and visual characteristics of crops and weeds are nearly indistinguishable [112]. Other challenges include occlusion and overlapping with crops [12], as well as the need for large volumes of labeled data for effective training [103].

Emerging trends in this task include: the use of pretrained models on large datasets such as ImageNet or DeepWeeds, followed by fine-tuning with weed-specific datasets [83]; the application of lightweight models optimized for edge computing in terms of parameters and inference speed [12]; the integration of attention mechanisms such as CBAM and efficient architectures like MobileViT; and the combination of supervised and unsupervised learning to reduce dependence on manual labeling.

Important gaps have also been identified, such as the need for robust, usable, and diverse datasets for weed classification using DCNNs [130]. Emerging applications like the use of SGANs for classification with RGB images remain underexplored [112]. Some studies have been limited to a narrow range of crop and weed species [88].

**Detection:** This task involves identifying individual objects, classifying them, and localizing them using bounding boxes. The most utilized architecture includes YOLO (v5, v7, v8, CBAM) with 15%, with YOLOv8n as the lightest and most efficient for field applications [8]. Overall, YOLO strikes a good balance between speed and accuracy. Faster R-CNN with 4% is known for its high precision but requires more computational resources [8]. They have shown that faster R-CNN achieves better accuracy than SSD. Other less common architectures include DenseNet, Mul-3DCNN, and MulDFNet. Additionally, detection models based on ResNet have been implemented.

The YOLO series is designed for object detection and classification of weeds or crops. The YOLOv8n version is one of the models that offers superior accuracy, detection capability, and reduced computational complexity compared to models like YOLOv7 and Faster R-CNN [77]. YOLOv8n incorporates lighter structures (e.g., GSConv, C2f) and integrates modern learning mechanisms such as anchor-free detection and a decoupled head to more efficiently and precisely determine positive and negative samples. It also improves the handling of occluded, blurry, and small objects (enhanced SPPF, multiscale layers) [102]. Thanks to these improvements, a customized version BSS-YOLOv8n demonstrated superior detection performance compared to YOLOv8, achieving a precision of 91.1%, a recall of 86.7%, an mAP50 of 92.6%, an F1-score of 88.85%, and an mAP50:95 of 61.3% on a *P. umbrosa* seedling dataset.

In summary, YOLOv8n is highly useful for object-related tasks within a detection framework, standing out for its speed, detection accuracy, and classification efficiency. However, detection models still face challenges such as visual camouflage of weeds [75], morphological variability (shape, size, species), overlap with crops, low spatial resolution [77], variable environmental conditions, and the high cost of manual annotation. Additionally, achieving adequate real-time detection speeds remains a challenge [101].

Current trends in this task include: a preference for single-stage detectors like the YOLO family due to their efficiency; use of multiscale detection (e.g., YOLOv3) to handle plant morphological variability; development of customized YOLO versions (e.g., YOLOv5-T, YOLOv5-KE) with attention mechanisms, improved feature aggregation, and micro-detection layers [90]; combining CNNs with Transformers (e.g., BoTNet, MobileViT) to improve accuracy; and implementation on edge devices for in-field processing [40].

Main gaps identified: There is a lack of annotated, large-scale, and multi-regional datasets to support more robust generalization.

**Semantic segmentation:** This approach classifies pixels based on their class without distinguishing individual instances. The most frequently used architectures include U-Net, with 7%, and are noted for their skip connections, which enhance accuracy. According to [8], U-Net with ResNet-34 as a feature extractor (backbone) showed solid performance. However, some studies suggest that SegNet and RefineNet slightly outperform U-Net in accuracy [107]. Fully Convolutional Networks (FCN) with 5%, particularly the VGG16-based variant, which balances accuracy and efficiency. Other architectures include DeepLabV3+ (3%), SegNet (2%), and Modified U-Net (2%).

U-Net has been one of the most widely used convolutional neural network architectures for semantic segmentation, thanks to its encoder–decoder design with skip connections. These allow the model to simultaneously capture high-level context and fine spatial details. The skip connections are essential for preserving spatial information during the reconstruction process, which is crucial for achieving accurate segmentations. Additionally, U-Net uses progressive upsampling and is highly adaptable due to its flexibility to integrate with various backbones and its ability to handle multiscale information. These characteristics have enabled U-Net to deliver accurate and detailed results, even in scenarios with limited datasets [6,36].

A modified version of the U-Net architecture was used to solve the semantic segmentation of *Colchicum autumnale* weeds, showing satisfactory performance with a sensitivity of 88.6% [116]. Another study developed a U-Net-based convolutional neural network called DeepSolanum-Net for the recognition and pixel-wise segmentation of *Solanum rostratum Dunal* (a common invasive weed), achieving a pixel recognition accuracy of 89.95% and a recall rate of 90.3% in field tests [103].

This task involves pixel-level segmentation, which entails high computational and memory demands [53], particularly due to the manual annotation of images with small weeds [14]. Occlusion and overlap with crops, as well as spectral and morphological similarity during early growth stages, further complicate segmentation tasks [113]. Traditional architectures present limitations: for example, FCNs show poor accuracy on fine details, U-Net has high computational complexity, and SegNet is associated with high false-positive rates [14].

Observed trends include the adoption of encoder–decoder architectures such as U-Net, SegNet, and DeepLabv3+ [16]; multiscale feature extraction through ASPP or SPP modules [105]; the design of lightweight models for inference on embedded systems [119]; data augmentation strategies, image cropping, and postprocessing using CRFs [41]; the implementation of attention mechanisms such as CBAM and CnSAU; and the fusion of multispectral information to improve vegetation discrimination [62].

Among the most relevant identified gaps are challenges caused by motion blur in UAV images [91], the lack of region-sensitive metrics that more accurately evaluate efficiency in practical applications like herbicide spraying beyond standard metrics such as F1-score or IoU [2]; difficulties in adapting Vision Transformer architectures to variations in real agricultural environments [86]; and the limited real-time integration of detection models with spraying systems [119].

**Instance segmentation:** This method classifies pixels and differentiates individual objects by combining segmentation and detection, increasing complexity. According to Appendix A2, the architecture used includes Mask R-CNN with 2% and other architectures with 1% usage.

**Mask R-CNN** is an extension of the Faster R-CNN framework specifically designed for instance segmentation. Its main innovation lies in the addition of a **mask branch** that generates a pixel-wise object mask for each detected instance, in addition to predicting the bounding box [80,84]. Key innovations introduced by Mask R-CNN include a two-stage architecture, a dedicated mask branch for pixel-level segmentation, and the use of pretrained backbones (e.g., ResNet101, Inception V2) [50]. These features make Mask R-CNN well-suited for weed segmentation, as it enables precise delineation of individual targets, high accuracy in edge detection, pixel-level detail, and robust performance across diverse datasets, with strong generalization capabilities. One example of its application is the early detection of weed seedlings (cranesbill) [105].

Instance segmentation, however, presents complex challenges such as the need to differentiate weeds in densely distributed areas and the high computational cost of two-stage models, which are often unsuitable for deployment on edge devices [95,110]. Additionally, object-level annotation is required, which is time-consuming, and there are issues with overlapping bounding boxes [95].

Recent trends in this area include a transition toward single-stage architectures that integrate detection and segmentation, making them more suitable for portable devices. Preprocessing techniques such as the use of color indices (e.g., G/R) have been applied to enhance contrast. Optimized backbones like ResNet101, as well as modules such as DSASPP and attention mechanisms, are being used to improve model performance. Data augmentation techniques (e.g., rotation, cropping, flipping) and loss functions such as Cross Entropy are commonly applied [105]. For hardware deployment, models are being converted into formats like ONNX or TorchScript for edge deployment (e.g., Jetson AGX). Additionally, new masking modules such as ProtoNet and Mask DINO are being explored [110].

Among the most relevant gaps identified are the lack of robust single-stage model designs capable of providing complete information on weed growth, the need for hardware performance validation (including accuracy, memory usage, and thermal behavior), limitations in implementing U-Net on edge devices, the necessity for automated hyperparameter optimization using evolutionary algorithms, and the limited exploration of multiscale training strategies to enhance performance on images with multiple resolutions.

**3.3.2.2. Other Deep Learning architectures:** Beyond CNNs, other architectures have been explored for weed detection, such as unsupervised learning, generative models, hybrid architectures, and Transformers. Table 6 presents the pre-trained architecture used in these approaches.

**Unsupervised learning** is a machine learning paradigm in which models operate with unlabeled data. This approach has been used in 4% of cases, with the following architectures applied: 1D-CNN (DenseNet-like) and Inception for classification tasks, PC/BC-DIM for detection, and 2D-CNN for instance segmentation.

**Transformers** were used in 3% of cases. Although initially designed for natural language processing and sequential data, these architectures have been adapted for classification tasks in the reviewed studies. The most frequently applied transformer models include 1D Transformer and Vision Transformer (ViT)**,** each with a 1% usage frequency. Additionally, MiT (Mixed Transformer) has been employed for semantic segmentation.

**Hybrid architecture** makes up 2% of applications, integrating different approaches and paradigms into a single network. In classification tasks, notable architectures include MobileViT, which integrates CNNs with ViTs [12], and ResNet-LSTM, which has been reported to achieve high weed classification accuracy [85], outperforming standalone models such as ResNet50.

**Generative models** were used in 1% of cases, primarily employing

**Table 6**

Deep Learning architecture by task in weed detection.

| Architecture | References | Pretrained architecture | Task |
|---|---|---|---|
| Unsupervised learning | [74] | 1DCNN (DenseNet-like) | Classification |
| | [128] | 2D-CNN | Instance Segmentation |
| | [71] | Inception | Classification |
| | [53] | PC/BC-DIM | Detection |
| Generative | [112] | SGAN | Classification |
| Hybrid | [12,85] | MobileVit, ResNet-LSTM | Classification |
| Transformer | [2,74,86] | 1D Transformer, Vision Transformer (ViT). | Classification |
| | | MiT (Mix Transformer) | Classification |
| | | RT-DETR | Semantic Segmentation |

SGAN (Sam-GAN). The primary function of this architecture is to generate synthetic images that preserve the structure and semantic content of authentic images [117]. This is particularly beneficial for expanding training datasets, especially when acquiring and annotating accurate data is challenging [112].

Summary of DL Architectures in Weed Detection: The reviewed studies indicate that deep learning architecture has been adapted and optimized for various tasks, including weed detection, segmentation, and classification. The most preferred architectures include FCN (various versions), U-Net, DeepLabV3+, YOLO (various versions), and ResNet (various versions). While ResNet offers higher accuracy, it comes with a higher computational cost. In contrast, MobileNet and specific versions of YOLO are better suited for real-time applications and resource-limited devices. Additionally, FCN combined with pre-trained CNN such as VGG16 and skip architecture has proven to be a highly effective and widely used approach.

Despite these advancements, deep learning techniques struggle with the trade-off between accuracy, real-time detection speed, and small weed localization. As a result, research trends are shifting toward hybrid models that combine CNNs and Transformers to optimize weed detection.

A particularly promising architecture is RT-DETR, recently adopted in the field. It integrates a CNN-based backbone with a Transformer, offering high speed and accuracy, making it an efficient alternative for real-time weed detection.

### *3.3.3. Languages, frameworks, hardware, and cloud services used for model training and deployment*

The reviewed studies have employed various programming languages, frameworks, hardware, and cloud services for training and deploying deep learning (DL) models in weed detection. The most used resources are detailed below.

**Programming languages:** Python has been the predominant language for training, evaluating, and deploying DL models [50,52,63,79,101] due to its ease of use and extensive library support. Its usage has been accompanied by various libraries such as SciKit-Learn and Keras, as well as frameworks like TensorFlow and PyTorch [6,36,52,79]. Additionally, C++ has been combined with CUDA, while JavaScript, Node.js, and Electron have been utilized to develop user interfaces in projects like Cofly.

**Frameworks: TensorFlow:** An open-source framework developed by Google that is widely used for creating, training, and deploying deep learning models. Its scalability allows execution on mobile devices, local servers, and cloud platforms. It excels in object detection tasks through its API, supporting models like Faster R-CNN and SSD [8]. Additionally, it includes Keras, a high-level API that simplifies the construction, training, and evaluation of neural networks [75]. PyTorch: An open-source framework developed by Meta (Facebook), known for its flexibility, ease of use, and strong integration with Python. It has been primarily used with pre-trained models such as ResNet, YOLOv5-T [90], and VGG. OpenMMLab [105] has also been mentioned for implementing instance and semantic segmentation pipelines.

**Hardware:** Various types of hardware, ranging from CPUs, GPUs, edge devices, and cloud servers, have been employed for implementing and testing deep learning architectures, depending on each model's computational requirements. Below is a classification of these devices based on their training power for DL. The classification is based on the specific performance of Tensor Cores, memory capacity (VRAM), memory bandwidth, and compatibility with frameworks such as TensorFlow and PyTorch. Fig. 4.

**CPUs:** Four CPUs have been used, featuring Core i5 processors at 1.8 GHz [53], an Intel i7-6600 [112], an Intel(R) Core (TM) i7-7820HQ [118], and an Intel Xeon E5-2670 [116]. The CPUs used in these studies had to be optimized, as well as image processing, to carry out weed detection. This highlighted the high computational cost of the methods employed and emphasized the need for GPU-based systems for execution and evaluation.

**Edge Devices:** Embedded systems have been used for local processing, reducing latency and bandwidth consumption. Notable examples include the NVIDIA Jetson TX2 [119], NVIDIA Jetson Xavier NX [107], NVIDIA Jetson AGX Xavier [56,100,110], and NVIDIA Jetson AGX Orin [40]. These devices enable model execution at the network edge, optimizing real-time processing while consuming less energy.

**GPUs**: Variants of NVIDIA GPUs have been used to accelerate model training and inference due to their parallel processing capability [110]. Among the reported architectures, considering their power, the following classifications can be made:

- Low-end: From the NVIDIA Quadro P2200 [96] to the NVIDIA Quadro M2000 [86].
- Mid-range: From the NVIDIA Titan X Pascal [51,54] to the NVIDIA GeForce GTX 1660 SUPER [109].
- High-end: Powerful GPUs for Deep Learning training, such as the NVIDIA A40 [6], NVIDIA RTX A5500 [74], NVIDIA GeForce RTX 3090 [6,11,72,90,91,100,102,103,106], NVIDIA GeForce RTX 3060 [40,56], and NVIDIA Quadro RTX 5000 [80].

Each of these GPUs has been selected based on the computational demands of the model, allowing for optimization in training and processing times.

**Cloud services:** The use of cloud services in the reviewed studies has been limited, as adoption is still in its early stages. Google Collaboratory [36,71,111,115] is mentioned as the most used tool, providing preferential access to GPUs such as Tesla K80 and Tesla P100-PCIE for more efficient training. Only one study reported the use of the three major cloud platforms: AWS (Amazon Web Services), Azure (Microsoft Azure), and GCP (Google Cloud) [81].

**Python** has been the most widely used programming language due to its versatility, ease of use, and extensive support community. PyTorch and TensorFlow are the predominant frameworks, with PyTorch being preferred for pre-trained models. Regarding hardware, NVIDIA GPUs have been essential for accelerating training, while Jetson devices have

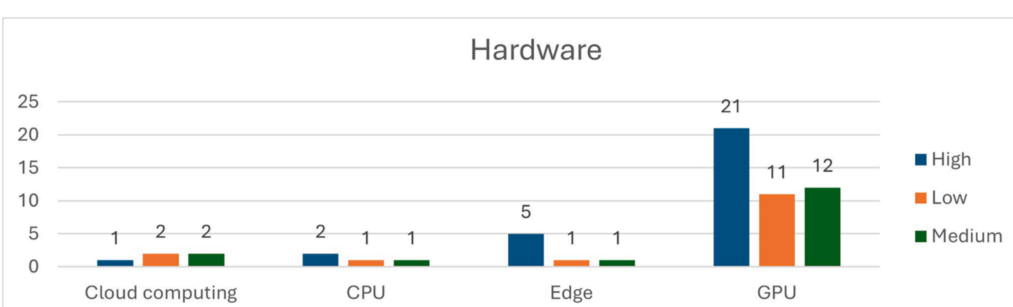


**Fig. 4.** Trend in hardware usage for training deep learning architectures.

enabled model deployment on embedded systems. Although cloud computing has been limited in the reviewed studies, Google Colaboratory is the most frequently mentioned platform.

### 3.4. RQ4: Metrics used for evaluating Deep Learning models in weed detection

Validation metrics are essential tools used to quantify and evaluate how well a DL model such as a convolutional neural network or a transformer performs in a specific task (detection, classification, semantic segmentation, or instance segmentation). These metrics help assess the model's validity, effectiveness, and reliability, support the selection of hyperparameters, and guide decision-making regarding its deployment in real-world applications in precision agriculture, such as weed management [4].

In the 77 reviewed articles, a total of 51 different metrics were identified, none of them used in isolation. Instead, studies combine multiple metrics to obtain a more comprehensive evaluation of each model's performance. The selection of these metrics largely depends on the type of problem addressed and the specific requirements of detection, segmentation, and classification tasks. Thus, the following metrics are found in the works reviewed according to the task.

In classification tasks, the most used metrics were Accuracy (20%), Recall and Precision (19% each), F1-Score (17%), Area Under the Curve (AUC) (12%), Intersection over Union (IoU) and confusion matrix (3.4%), with other metrics used less frequently. For detection tasks, the most common metrics were Precision (25%), Recall (23%), mean Average Precision (mAP) (20%), Accuracy (11%), and F1-Score (10%). In instance segmentation, Precision, Accuracy, Recall, IoU, and mAP were all used equally at 20%. For semantic segmentation, the most relevant metrics were Recall and F1-Score (both 18%), IoU (16%), and Accuracy and Precision (14%) (see Fig. 5). Additional metrics are detailed in Appendix A3.

**Classification** (Weeds vs. Crops): In this task, Accuracy measures the proportion of correct predictions, Recall indicates the proportion of weeds that were detected, Precision reflects the proportion of elements classified as "weed" that are weeds, and the F1-Score represents the harmonic mean of Precision and Recall. However, in agricultural contexts where class imbalance is common (more crops than weeds), Accuracy can be misleading. In this sense, the F1-Score is considered the most informative metric, as it balances Type I and Type II errors (false positives and false negatives), offering a more robust evaluation. The formula is:

$$F1 - Score = 2\,\frac{Precision * Recall}{Precision + Recall}$$

In this context, **Precision** represents the proportion of true positive samples among all predicted positives, while **Recall** indicates the proportion of true positive samples among all actual positives.

**Detection** (localization and classification of weeds): In this task, **Precision** measures the proportion of correct detections among all positive predictions, while **Recall** evaluates the model's ability to find all real instances. The mean Average Precision (mAP) measures performance at different overlapping thresholds (IoU) between predictions and ground truth annotations. mAP is the standard metric for this task as it integrates Precision and Recall, allowing assessment of both correct classification and precise spatial localization of weeds. Its formula is:

$$mAP = \frac{1}{N}\sum_{i-1}^{N} APi$$

Where APi is the AP for class i, and N is the number of classes.

**Instance Segmentation** (separates each individual object): This task uses metrics such as Precision, Recall, mAP, and AP equally. The goal is not only to label correctly but also to distinguish between different individual objects, even if they belong to the same class. The recommended metric is Average Precision (AP) per instance, as it simultaneously measures detection quality and segmentation accuracy for each individual object. Its formula is:

$$AP = \int_0^1 P(R)dr$$

Where the meaning of the terms is like that in classification.

**Semantic Segmentation (**pixel-wise classification): This task employs metrics such as Recall, F1-Score, IoU, Precision, and Accuracy. Among these, mean Intersection over Union (mIoU) stands out as the most comprehensive metric, as it quantifies the overlap between predicted and ground truth masks for each class. The specific formula is:

$$mIoU = \frac{1}{N}\sum_{i-1}^{N}\frac{TP_i}{TP_i + FP_i + FN_i}$$

Among them, the meanings of $TP_i$, $FN_i$, and $FP_i$ in mIoU are consistent with their definitions in IoU.

The studies were complemented with graphical evaluation metrics such as the Precision-Recall Curve, ROC Curve, and Confusion Matrix, which allow visualization of model performance from different perspectives. Additionally, key metrics related to temporal and resource efficiency were included, such as FPS (Frames Per Second), speed or inference time (ms), and training time (s). To evaluate computational consumption, indicators like memory usage (Resources in MB/GB), energy consumption (in J/Wh), and computational load in FLOPs were considered. Structural aspects of the model, such as file size (MB) and number of parameters, were also analyzed, as they are critical factors for deployment on resource-constrained devices.

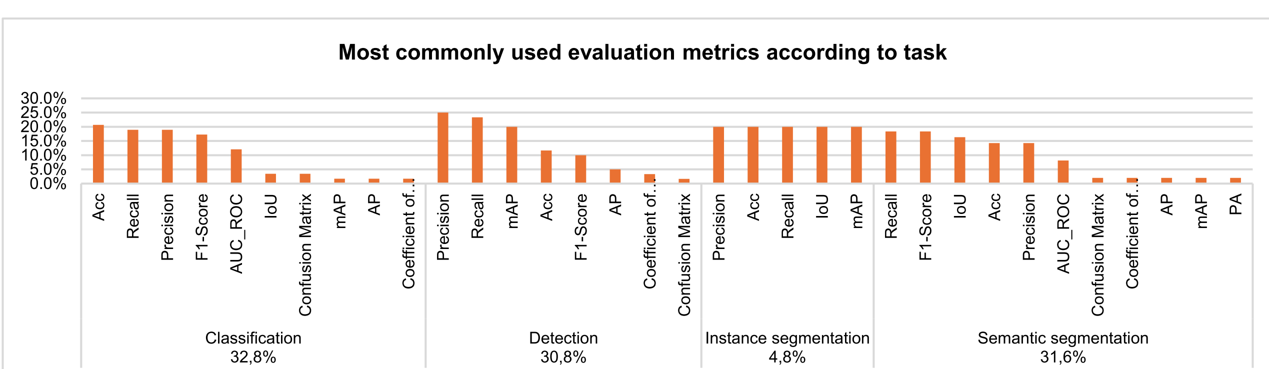


**Fig. 5.** Metrics used in the evaluation of deep learning models. The complete list can be found in the supplementary information (Appendix A3).

## 4. Discussion

In this section, a Systematic Literature Review (SLR) is conducted to analyze key topics, challenges, and opportunities in using Deep Learning architectures for weed detection and classification in images captured by UAVs. The analysis covers 77 original studies, carefully selected to address four general research questions regarding the current state of this technology. The findings provide a solid scientific basis to support the need to improve and develop more efficient and accurate models, capable of adapting to the real conditions of the agricultural environment.

Furthermore, the results synthesized in this review can guide future technological innovation initiatives with a local focus, especially in contexts where the adoption of digital tools in agriculture is still limited, such as in developing countries. A practical application of these findings is currently being developed in a parallel study focused on potato cultivation under the specific agricultural conditions of the Carchi-Imbabura region (Ecuador), where the goal is to implement an automatic detection system adapted to this reality based on the insights from this review [131].

### 4.1. UAVs in weed classification and detection

UAVs have proven to be a promising tool for acquiring high-resolution agricultural images efficiently, non-intrusively, and cost-effectively [14,36,51,83,111,113]. UAVs can integrate various types of sensors, making them ideal for weed detection and classification. However, their use presents technical, operational, and data-processing challenges.

In the analyzed studies, the most used UAVs were commercial DJI models (Phantom, Mavic, Matrice, Inspire, and Spark), favored for their availability, reliability, and ease of operation, which has facilitated their adoption in precision agriculture [14,80]. While many studies employed integrated cameras, others incorporated more specialized cameras such as Parrot Sequoia, MicaSense RedEdge, or Hasselblad, among others. Additionally, multispectral, hyperspectral, and LiDAR sensors were included, thereby expanding the capabilities for information capture.

RGB cameras are excellent for general monitoring and rapid detection tasks at low cost, especially when combined with optimized models like YOLO-Spot [56]. Multispectral sensors enable better vegetation discrimination and early stress detection thanks to bands such as NIR and Red-Edge [62,82]. Hyperspectral sensors offer greater specificity, useful for detecting resistances or similar weeds, albeit with higher cost and processing requirements. Sensor fusion (e.g., RGB + MS or MS + LiDAR) improves information quality, allowing the acquisition of both spectral and structural vegetation properties.

The quality of data obtained largely depends on flight operational conditions: altitude, spatial resolution, flight speed, and environmental conditions (sunlight, cloudiness, wind) [8,91]. Low-altitude flights improve resolution (cm/pixel) but reduce coverage area; high speeds or windy conditions can cause motion blur, affecting image quality and real-time detection [36,74].

In this context, finding a balance between altitude, resolution, and flight speed, as well as selecting sensors according to study objectives and environment, is crucial. This comprehensive approach enables improved accuracy in weed detection and classification, particularly in scenarios requiring real-time response.

### 4.2. Studied weeds and crops

Crops: Most reviewed studies identified seven main crops: sugar beet, maize, rice, bean, cotton, wheat, and grass. This selection focuses on high-economic-value crops or those with specific weed management challenges [80,132]. The choice of crops in these studies is based on factors such as the rapid spread of invasive species, herbicide resistance, and impact on yield performance.

Weed recognition has advanced towards developing more precise, efficient, and sustainable solutions, driven by the need to improve the management of these invasive species. The studies identified a wide variety of significant weeds, including *Amaranthus, Chenopodium album, Cirsium arvense, Grass, Rumex obtusifolius, Convolvulus arvensis, Cyperus iria*, and *Leptochloa chinensis*. These species present specific challenges due to their high competitiveness and adaptability in different agricultural environments.

### 4.3. Labeled data and their availability

Weed datasets obtained through UAVs are essential for training computer vision models in precision agriculture. Although some public datasets exist (Appendix A4), their availability is limited in terms of crop diversity, weed types, and environmental conditions. As a result, many researchers opt to generate their own datasets through manual collection and annotation a process that ensures quality but is time- and resource-intensive.

The most used annotation tools (such as LabelImg, Labelme, and GIMP) facilitate labeling but present limitations, including systematic errors and subjective biases. To reduce this burden, strategies such as semi/automatic annotation (e.g., Segment Anything Model), unsupervised learning, data augmentation, and transfer learning have been explored.

Currently, the trend is to combine high-quality manual annotation with automatic and synthetic techniques to efficiently scale dataset creation and reduce dependence on labor-intensive human annotation.

### 4.4. CNN architectures in weed classification and detection and their metrics

**Classification:** CNN-based architectures were the primary choice due to their ability to learn feature representations directly from image datasets, eliminating the need for manually designed filters [78]. This trend was reported in 14 review articles [5,57,61]. Models such as ResNet, VGGNet, and AlexNet were widely used to classify each pixel as crop, weed, or soil [50,120].

Additionally, Vision Transformers (ViT), based on attention mechanisms, have gained traction for weed identification, even in real-time applications (e.g., RT-DETR). Although its adoption remains limited, ViT presents a viable alternative to CNNs, particularly for smaller datasets and when combined with transfer learning [86]. Furthermore, there has been a rise in hybrid architectures (CNN+Transformer), where CNNs extract local features while Transformers model global dependencies, improving classification accuracy.

**Detection**: The YOLO architecture was the most widely adopted due to its balance between efficiency and accuracy. YOLO is faster and computationally less complex than two-stage architecture such as Faster R-CNN [90]. Variants such as YOLOv5, YOLOv7, and YOLOv8 were extensively used (Table 6). Additionally, YOLO-Spot was developed based on YOLOv7 tiny to enhance efficiency in resource-limited devices [110]. Recent versions, such as YOLOv11 (nano, small, medium, large, extra-large), aim to improve accuracy and speed, mitigating motion blur effects under adverse conditions [11].

**Semantic segmentation:** The Fully Convolutional Network (FCN) was the base architecture for segmentation, with variants like FCN-4s and FCN-8s. However, U-Net was the most widely used, aligning with prior research [5,60,133]. Other architectures, such as SegNet, DeepLabV3, encoder-decoder networks, and Transformers, were employed. According to [134], Segmentation Transformer (SETR) and TransUNet are being explored as alternatives to convolution-based approaches.

**Instance segmentation:** The most used architecture was Mask R-CNN, an extension of Faster R-CNN that integrates detection and instance segmentation. Encoder-encoder architectures were employed to generate high-resolution segmentation maps. Additionally, multi-task models, clustering-based transformations [134], and other emerging approaches were investigated for this task.

**Performance metrics:** Various metrics were used to evaluate model performance, including precision, recall, F1-score, accuracy, ROC curves, AU-ROC, and Average Precision (mAP) (Fig. 5). Metric selection depends on the specific task and model requirements.

For imbalanced datasets, precision, recall, F1-score, and mAP were the most used metrics to assess model discrimination and balance [77, 78,110]. The review highlighted the importance of multiple evaluation metrics to assess performance aspects, including real-time speed, to optimize classification, detection, and segmentation.

In general, the use of deep learning (DL) architectures for weed management using UAV imagery across tasks such as classification, detection, semantic segmentation, and instance segmentation faces challenges such as spectral and visual similarity between weeds and crops, occlusion and overlap, and limited availability of labeled data. Current trends point toward the adoption of lightweight and efficient models, edge deployment, the use of attention mechanisms (e.g., Transformers), and unsupervised learning. However, critical gaps remain, including the scarcity of diverse and annotated datasets, performance limitations under real-world conditions, and the lack of specific metrics for agri-intelligent tasks such as herbicide spraying.

**Hardware:** A strong trend was observed toward the use of high-end GPUs, such as the NVIDIA GeForce RTX 3090, as well as the NVIDIA A40 [6] and NVIDIA RTX A5500, among others (Fig. 4). These GPUs have proven to be an efficient option, as they are designed to handle large volumes of data and perform intensive computations thanks to their parallel processing capabilities, enabling the real-time implementation of models. Despite their high performance, they also present limitations such as high cost, high energy consumption, efficient heat dissipation, and specialized infrastructure requirements.

There was also a growing interest in using high-end embedded devices for edge AI, such as NVIDIA Jetson AGX, NVIDIA Jetson AGX Xavier, and NVIDIA Jetson AGX Orin [40]. These devices have been used for on-field processing, reducing latency and enabling faster decision-making in mapping and selective spraying tasks.

However, while these solutions provide greater autonomy and efficiency, they also have limitations regarding computing power, storage capacity, and thermal management.

A recurring challenge is the need for more powerful GPUs, such as the NVIDIA A100, which is specifically designed for AI workloads, featuring third-generation Tensor Cores, high memory capacity, and extreme bandwidth. However, its high-cost limits are adopted in resource-constrained environments, highlighting the importance of model optimization and compression techniques to enhance efficiency without significantly increasing computational demand.

**Frameworks and software:** TensorFlow and PyTorch were the most widely used deep learning frameworks, which were crucial in developing weed detection models. These frameworks facilitated the processing of large datasets and the implementation of complex models, promoting reproducibility and scalability.

Summary of main limitations, challenges, trends, and gaps identified in the literature

1. **High computational cost and hardware requirements**
   · **Limitations:** Semantic and instance segmentation tasks require high GPU capacity, memory, and training time, which limits their implementation in low-resource agricultural settings.
   · **Challenges:** Running heavy models like Transformers in real time is difficult due to their high computational demands. High-performance GPUs (e.g., Nvidia A100) are required, which are costly.
   · **Trends:** Use of lightweight architecture, model quantization, pruning, and compression techniques. Exploration of efficient Transformers with reduced computational load.
   · **Gaps:** Although progress has been made in the use of embedded devices such as Jetson AGX for edge computing, they still face constraints in terms of capacity, storage, and thermal management. Further optimization is needed for field scenarios.
2. **Availability and cost of labeled data**
   · **Limitations:** Manual labeling, especially pixel-wise, is time- and resource-intensive.
   · **Challenges:** Lack of UAV-labeled datasets for specific crops (e.g., potato) and limited diversity in species, environmental conditions, and geographic regions.
   · **Trends:** Increase in synthetic data, transfer learning, semi-/unsupervised learning, and deep clustering techniques.
   · **Gaps:** Persistent lack of large, multi-regional, standardized (built under consistent protocols for acquisition, annotation, and partitioning), and open-access databases for benchmarking. Public availability with metadata and licenses is essential for reproducibility and model comparison.
3. **Biological and visual variability of crops and weeds**
   · **Limitations:** High intra-class similarity in early growth stages, camouflage, crop occlusion, and morphological variability (color, shape, texture).
   · **Challenges:** Reduced model accuracy under phenological or spectral variation and uncontrolled environmental conditions.
   · **Trends:** Incorporation of spectral indices (e.g., G/R, NDVI), multi-scale models, and attention mechanisms for more robust feature extraction.
   · **Gaps:** Lack of robustness to intra/inter-crop variation and limited generalization beyond the training domain.
4. **Operational conditions of UAVs**
   · **Limitations:** Image quality depends on flight altitude, speed, lighting, wind, and spatial resolution. Lower altitudes increase resolution but reduce coverage.
   · **Challenges:** Movement, blur, and reduced quality affect real-time detection.
   · **Trends:** Use of more stable multirotor UAVs and exploration of multi-agent systems to increase coverage and autonomy.
   · **Gaps:** Need to adapt models to UAV image variability and explore motion compensation or multimodal fusion strategies.
5. **Efficiency and adaptability of DL models at the edge**
   · **Limitations:** Traditional DL models are not optimized for low-resource hardware (memory, energy, processing).
   · **Challenges:** Difficulty in making in-flight decisions or performing real-time intelligent spraying.
   · **Trends:** Development of optimized models (e.g., MobileNet, Tiny-YOLO, RT-DETR), conversion to ONNX/TorchScript formats, and deployment on Jetson AGX/Orin.
   · **Gaps:** Need for more field validation, real hardware integration, and automatic hyperparameter optimization (e.g., evolutionary algorithms).
6. **Task specificity and architecture**
   · **Classification:**
     - **Trends:** Use of pre-trained models, attention mechanisms (CBAM), CNN+Transformer integration.
     - **Gaps:** Limited exploration of methods such as SGAN; datasets restricted in species and conditions.
   · **Detection:**
     - **Trends:** One-stage models (YOLO), multi-scale detection, custom variants with attention and specialized layers.
     - **Gaps:** Need for well-annotated, generalizable datasets.
   · **Semantic segmentation:**
     - **Trends:** Optimized encoder-decoder architectures (U-Net, DeepLabv3+), multi-scale techniques (ASPP, SPP), attention mechanisms, spectral fusion.
     - **Gaps:** Lack of task-specific metrics (e.g., spraying precision), sensitivity to UAV-induced blur, limited integration with smart sprayers.
   · **Instance segmentation:**

- **Trends:** One-stage architecture, spectral indices, ProtoNet/Mask DINO masking modules, Jetson deployment.
- **Gaps:** Limited exploration of multi-scale training, insufficient real-world performance validation, need for integration into autonomous decision-making systems.

Given these findings, conducting regular, systematic literature reviews is crucial to staying at the forefront of technological advancements and identifying critical research areas in this rapidly evolving field.

## 5. Conclusion

This Systematic Literature Review (SLR) has enabled an analysis of the current state, challenges, and opportunities in using Deep Learning architectures for weed detection and classification in images captured by UAVs. Through the review of 77 original studies, key research trends and the main limitations that persist in this field have been identified.

The results highlight the need to develop more efficient and accurate models that optimize the automatic detection of weeds, especially in potato crops. In particular, the importance of adapting these solutions to specific conditions where the adoption of agricultural technology remains limited is emphasized. These findings provide a solid scientific foundation for future research and applications in the agricultural sector, driving the development of innovative tools for more sustainable and efficient crop management.

UAVs have revolutionized precision agriculture, particularly crop monitoring, weed detection, and classification. This is due to their ability to capture high-resolution spatial images efficiently and non-intrusively, operating at low altitudes and incorporating specialized sensors. These tools enable tasks ranging from monitoring plant health to applying targeted treatments on-site, supported by artificial intelligence and Deep Learning architecture. However, challenges remain, such as limited flight autonomy, payload capacity, and the variability of weather conditions. Research on multi-agent systems is emerging as a potential solution to these operational challenges.

Deep learning architectures have proven essential for weed detection, classification, and segmentation in crops using UAV imagery. Notably, models such as YOLO, U-Net, and the more recent Transformers stand out for their deep learning capabilities and effectiveness in real-time applications. Hybrid architecture combining CNNs and Transformers are evolving to better meet agricultural environment requirements, seeking a balance between speed, accuracy, and efficiency on resource-constrained devices. However, challenges remain, including the visual similarity between weeds and crops, overlap and occlusion, and limited availability of labeled data. Performance evaluation has relied on metrics such as Precision, Recall, F1-Score, and mAP, with the choice depending on the task type and dataset characteristics, prioritizing those that address class imbalance and computational efficiency.

Nonetheless, there is still a need to develop specific metrics to assess practical impact in applications like selective herbicide spraying. Despite advances, significant gaps persist related to the scarcity of diverse annotated datasets, validation under real-world conditions, and the effective integration of models into autonomous agricultural systems.

The choice of hardware and software for weed detection largely depends on the specific application needs, budget, and required accuracy. A trend is to use high-end GPUs, such as NVIDIA A40 and NVIDIA RTX A5500, to handle large data volumes and perform intensive computations. Embedded devices, such as NVIDIA Jetson AGX Xavier and NVIDIA Jetson AGX Orin, have been widely adopted for real-time, field-based applications.

Future work can focus on the following areas: improving model robustness; optimizing for real-time applications and deployment on edge computing devices with robotic integration; creating larger, more diverse, and cost-effective datasets; advancing deep learning architectures and imaging techniques by exploring new architectures such as Vision Transformers (ViT) and multimodal integration; and integrating with precision spraying systems and overall field management.

In summary, weed detection in crops is a complex problem involving multiple challenges related to UAVs (flight time, payload capacity, and weather conditions), hardware (sensors, processing, and edge devices), software (labeling, generalization, variability, and the balance between accuracy and speed), architectures (model complexity, feature extraction, speed-accuracy trade-off, and pixel segmentation), and evaluation metrics (class imbalance handling and effectiveness).

To overcome these challenges, it is essential to promote continuous research and explore new techniques that enable progress toward robust and functional applications. In this regard, future work should focus on reducing the limitations associated with current datasets, optimizing model performance under real-world conditions, and facilitating the implementation of autonomous precision agriculture systems for more efficient and sustainable weed management.

## Ethics statement

Not applicable: This manuscript does not include human or animal research.

## Note

While preparing this work, the authors used ChatGPT to improve the writing and Grammarly to translate the document. After using these tools/services, the authors reviewed and edited the content as needed and assumed full responsibility for the publication's content.

## CRediT authorship contribution statement

**Lucía Sandoval-Pillajo:** Writing – original draft, Visualization, Investigation, Conceptualization. **Iván García-Santillán:** Validation, Supervision, Conceptualization. **Marco Pusdá-Chulde:** Writing – review & editing. **Adriana Giret:** Writing – review & editing.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.atech.2025.101147.

## Data availability

No data was used for the research described in the article.

## References

[1] FAO, Recomendaciones Para el Manejo de Malezas, 1, FAO (Organización de las Naciones Unidas para la Alimentación y la Agricultura), 2008, pp. 1–61. Accessed: Mar. 27, 2020. [Online]. Available: http://www.fao.org/3/a0884s/a0884s.pdf.

[2] G. Castellano, P. De Marinis, G. Vessio, Weed mapping in multispectral drone imagery using lightweight vision transformers, Neurocomputing 562 (August) (2023) 126914, https://doi.org/10.1016/j.neucom.2023.126914.

[3] A. Wang, W. Zhang, X. Wei, A review on weed detection using ground-based machine vision and image processing techniques, Comput. Electron. Agric. 158 (November 2018) (2019) 226–240, https://doi.org/10.1016/j.compag.2019.02.005.

[4] A. S. M. M. Hasan, F. Sohel, D. Diepeveen, H. Laga, and M. G. K. Jones, "A survey of deep learning techniques for weed detection from images," 2021. doi: 10.1016/j.compag.2021.106067.

[5] M.A. Istiak, et al., Adoption of Unmanned Aerial Vehicle (UAV) imagery in agricultural management: a systematic literature review, Ecol. Inform. 78 (September) (2023) 102305, https://doi.org/10.1016/j.ecoinf.2023.102305 no.
[6] N. Genze, R. Ajekwe, Z. Güreli, F. Haselbeck, M. Grieb, D.G. Grimm, Deep learning-based early weed segmentation using motion blurred UAV images of sorghum fields, Comput. Electron. Agric. 202 (September) (2022) 107388, https://doi.org/10.1016/j.compag.2022.107388.
[7] M. Esposito, M. Crimaldi, V. Cirillo, F. Sarghini, A. Maggio, Drone and sensor technology for sustainable weed management: a review, Chem. Biol. Technol. Agric. 8 (1) (2021) 1–11, https://doi.org/10.1186/s40538-021-00217-8.
[8] A.N.V. Sivakumar, et al., Comparison of object detection and patch-based classification deep learning models on mid-to late-season weed detection in UAV imagery, Remote Sens. 12 (13) (2020), https://doi.org/10.3390/rs12132136.
[9] A. & Soria, M., Taylor, U., Tye and S. R. Wilkinson, Identificación y manejo de malezas en las islas galápagos. Galápagos, 2002.
[10] W.C. Liang, Y.J. Yang, C.M. Chao, Low-cost weed identification system using drones, in: Proceedings of the 7th International Symposium on Computing and Networking Workshops, CANDARW 2019, 2019, pp. 260–263, https://doi.org/10.1109/CANDARW.2019.00052.
[11] Y. Lin, et al., Quick and accurate monitoring peanut seedlings emergence rate through UAV video and deep learning, Comput. Electron. Agric. 197 (November 2021) (2022) 106938, https://doi.org/10.1016/j.compag.2022.106938.
[12] P. Alirezazadeh, M. Schirrmann, F. Stolzenburg, A comparative analysis of deep learning methods for weed classification of high-resolution UAV images, J. Plant Dis. Prot. 131 (1) (2024) 227–236, https://doi.org/10.1007/s41348-023-00814-9.
[13] S. Ivić, A. Andrejčuk, S. Družeta, Autonomous control for multi-agent non-uniform spraying, Appl. Soft Comput. J. 80 (2019) 742–760, https://doi.org/10.1016/j.asoc.2019.05.001.
[14] K. Zou, X. Chen, F. Zhang, H. Zhou, C. Zhang, A field weed density evaluation method based on uav imaging and modified u-net, Remote Sens. 13 (2) (2021) 1–19, https://doi.org/10.3390/rs13020310.
[15] M. Vasileiou, et al., Transforming weed management in sustainable agriculture with artificial intelligence: a systematic literature review towards weed identification and deep learning, Crop Prot. 176 (November 2023) (2024) 106522, https://doi.org/10.1016/j.cropro.2023.106522.
[16] M. Khoshboresh-Masouleh, M. Akhoondzadeh, Improving weed segmentation in sugar beet fields using potentials of multispectral unmanned aerial vehicle images and lightweight deep learning, J. Appl. Remote Sens. 15 (03) (2021) 1–17, https://doi.org/10.1117/1.jrs.15.034510.
[17] N.N. Ahmad Loti, M.R. Mohd Noor, S.W. Chang, Integrated analysis of machine learning and deep learning in chili pest and disease identification, J. Sci. Food Agric. 101 (9) (2021) 3582–3594, https://doi.org/10.1002/jsfa.10987.
[18] INIAP, "Papa – INIAP E.V.A." Accessed: Apr. 06, 2020. [Online]. Available: http://eva.iniap.gob.ec/web/papa/.
[19] S. Khan, M. Tufail, M.T. Khan, Z.A. Khan, S. Anwar, Deep learning-based identification system of weeds and crops in strawberry and pea fields for a precision agriculture sprayer, Precis. Agric. 22 (6) (2021) 1711–1727, https://doi.org/10.1007/s11119-021-09808-9.
[20] J.G. Speight, Sources and types of organic pollutants. Environmental Organic Chemistry for Engineers, Elsevier, 2017, pp. 153–201, https://doi.org/10.1016/b978-0-12-804492-6.00004-6.
[21] N.J. Hawkins, C. Bass, A. Dixon, P. Neve, The evolutionary origins of pesticide resistance, Biol. Rev. 94 (1) (2019) 135–155, https://doi.org/10.1111/brv.12440.
[22] M. Tufail, J. Iqbal, M.I. Tiwana, M.S. Alam, Z.A. Khan, M.T. Khan, Identification of Tobacco Crop Based on Machine Learning for a Precision Agricultural Sprayer, IEEE Access 9 (2021) 23814–23825, https://doi.org/10.1109/ACCESS.2021.3056577.
[23] M.T. Linaza, et al., Data-driven artificial intelligence applications for sustainable precision agriculture, Agronomy 11 (6) (2021), https://doi.org/10.3390/agronomy11061227.
[24] J. Díaz, Estudio de Índices de vegetación a partir de imágenes aéreas tomadas desde UAS/RPAS y aplicaciones de estos a la agricultura de precisión, Madr. Esp. (2015) [Online]. Available: http://eprints.ucm.es/31423/1/TFM_Juan_Diaz_Cervignon.pdf.
[25] G. Pajares, et al., Machine-vision systems selection for agricultural vehicles: a guide, J. Imaging 2 (4) (2016), https://doi.org/10.3390/jimaging2040034.
[26] R. Bongiovanni, E. Chartuni, S. Best, and Á. Roel, Agricultura de presición: integrando conocimentos para una agricultura moderna y sustentable, vol. 10. 2006.
[27] J.M. Guerrero, Sistema de visión para agricultura de precisión, Ing. Softw. (2015) 244. Online]. Available: https://eprints.ucm.es/30145/1/T36062.pdf.
[28] I.D. García-Santillán, M. Montalvo, J.M. Guerrero, G. Pajares, Automatic detection of curved and straight crop rows from images in maize fields, Biosyst. Eng. 156 (2017) 61–79, https://doi.org/10.1016/j.biosystemseng.2017.01.013.
[29] I. García-Santillán, G. Pajares, On-line crop/weed discrimination through the Mahalanobis distance from images in maize fields, Biosyst. Eng. 166 (2018) 28–43, https://doi.org/10.1016/j.biosystemseng.2017.11.003.
[30] S. Sabzi, Y. Abbaspour-Gilandeh, G. García-Mateos, A fast and accurate expert system for weed identification in potato crops using metaheuristic algorithms, Comput. Ind. 98 (2018) 80–89, https://doi.org/10.1016/j.compind.2018.03.001.
[31] L.F. Thomas, M. Änäkkälä, A. Lajunen, Weakly Supervised perennial weed detection in a barley field, Remote Sens. 15 (11) (2023), https://doi.org/10.3390/rs15112877.
[32] M. Perez-Ortiz, P.A. Gutierrez, J.M. Pena, J. Torres-Sanchez, F. Lopez-Granados, C. Hervas-Martinez, Machine learning paradigms for weed mapping via unmanned aerial vehicles, in: Proceedings of the IEEE Symposium Series on Computational Intelligence, SSCI 2016, Institute of Electrical and Electronics Engineers Inc., 2017, https://doi.org/10.1109/SSCI.2016.7849987.
[33] J. Su, et al., Spectral analysis and mapping of blackgrass weed by leveraging machine learning and UAV multispectral imagery, Comput. Electron. Agric. 192 (December 2021) (2022) 106621, https://doi.org/10.1016/j.compag.2021.106621 p.
[34] A. Bakhshipour, Cascading feature filtering and boosting algorithm for plant type classification based on image features, IEEE Access (2021), https://doi.org/10.1109/ACCESS.2021.3086269.
[35] M. Das, A. Bais, DeepVeg: deep learning model for segmentation of weed, canola, and canola flea beetle damage, IEEE Access 9 (2021) 119367–119380, https://doi.org/10.1109/ACCESS.2021.3108003.
[36] T.B. Shahi, S. Dahal, C. Sitaula, A. Neupane, W. Guo, Deep learning-based weed detection using UAV images: a comparative study, Drones 7 (10) (2023) 1–18, https://doi.org/10.3390/drones7100624.
[37] S.I. Moazzam, et al., Patch-wise weed coarse segmentation mask from aerial imagery of sesame crop, Comput. Electron. Agric. 203 (June 2021) (2022) 107458, https://doi.org/10.1016/j.compag.2022.107458.
[38] A.H. Al-Badri, et al., Classification of weed using machine learning techniques: a review–challenges, current and future potential techniques, J. Plant Dis. Prot. 129 (4) (2022) 745–768, https://doi.org/10.1007/s41348-022-00612-9.
[39] J. Gao, A. P. French, M. P. Pound, Y. He, T. P. Pridmore, and J. G. Pieters, "Deep convolutional neural networks for image-based Convolvulus sepium detection in sugar beet fields," 2020. doi: 10.1186/s13007-020-00570-z.
[40] E.C. Nnadozie, O.N. Iloanusi, O.A. Ani, K. Yu, Detecting cassava plants under different field conditions using UAV-based RGB images and deep learning models, Remote Sens. 15 (9) (2023) 1–18, https://doi.org/10.3390/rs15092322.
[41] J. Gao, et al., Cross-domain transfer learning for weed segmentation and mapping in precision farming using ground and UAV images, Expert Syst. Appl. 246 (December 2023) (2024) 122980, https://doi.org/10.1016/j.eswa.2023.122980.
[42] N. Unidas, "La Agenda 2030 y los Objetivos de Desarrollo Sostenible: una oportunidad para América Latina y el Caribe," 2030. [Online]. Available: www.issuu.com/publicacionescepal/stacks.
[43] A. Fernandez, E. Insfran, S. Abrahão, Usability evaluation methods for the web: a systematic mapping study, Inf. Softw. Technol. (2011) 789–817, https://doi.org/10.1016/j.infsof.2011.02.007.
[44] B. Kitchenham and S. Charters, Guidelines for performing systematic literature reviews in software engineering. 2007.
[45] R. Díaz-Varela, R. de la Rosa, L. León, P. Zarco-Tejada, High-resolution airborne UAV imagery to assess olive tree crown parameters using 3D Photo reconstruction: application in breeding trials, Remote Sens. 7 (4) (2015) 4213–4232, https://doi.org/10.3390/rs70404213.
[46] A. González, G. Amarillo, M. Amarillo, F. Sarmiento, Drones aplicados a la agricultura de precisión Drones applied to precision agriculture, Rev. Espec. Ing. 10 (2017) 23–37, 24.
[47] H. Huang, et al., Accurate weed mapping and prescription map generation based on fully convolutional networks using UAV imagery, Sensors 18 (10) (2018), https://doi.org/10.3390/s18103299.
[48] A. Saddik, R. Latif, M. Elhoseny, A. El Ouardi, Real-time evaluation of different indexes in precision agriculture using a heterogeneous embedded system, Sustain. Comput. Inform. Syst. 30 (December 2019) (2021) 100506, https://doi.org/10.1016/j.suscom.2020.100506.
[49] L. Malambo, S. Popescu, N.W. Ku, W. Rooney, T. Zhou, S. Moore, A deep learning semantic segmentation-based approach for field-level sorghum panicle counting, Remote Sens. 11 (24) (2019) 2939, https://doi.org/10.3390/rs11242939.
[50] K. Osorio, A. Puerto, C. Pedraza, D. Jamaica, L. Rodríguez, A deep learning approach for weed detection in lettuce crops using multispectral images, Agriengineering 2 (3) (2020) 471–488, https://doi.org/10.3390/agriengineering2030032.
[51] I. Sa, et al., WeedMap: a large-scale semantic weed mapping framework using aerial multispectral imaging and deep neural network for precision farming, Remote Sens. 10 (9) (2018), https://doi.org/10.3390/rs10091423.
[52] S.I. Moazzam, et al., A patch-image based classification approach for detection of weeds in sugar beet crop, IEEE Access 9 (2021) 121698–121715, https://doi.org/10.1109/ACCESS.2021.3109015.
[53] A. Naveed, et al., Saliency-based semantic weeds detection and classification using UAV multispectral imaging, IEEE Access 11 (January) (2023) 11991–12003, https://doi.org/10.1109/ACCESS.2023.3242604.
[54] I. Sa, et al., WeedNet: dense semantic weed classification using multispectral images and MAV for smart farming, IEEE Robot. Autom. Lett. 3 (1) (2018) 588–595, https://doi.org/10.1109/LRA.2017.2774979.
[55] J. P. T. Lambert, D. Z. Childs, and R. P. Freckleton, "Testing the ability of unmanned aerial systems and machine learning to map weeds at subfield scales: a test with the weed Alopecurus myosuroides (Huds)," 2019. doi: 10.1002/ps.5444.
[56] N. Rai, Y. Zhang, M. Villamil, K. Howatt, M. Ostlie, X. Sun, Agricultural weed identification in images and videos by integrating optimized deep learning architecture on an edge computing technology, Comput. Electron. Agric. 216 (June 2023) (2024) 108442, https://doi.org/10.1016/j.compag.2023.108442.
[57] J. Zhang, F. Yu, Q. Zhang, M. Wang, J. Yu, Y. Tan, Advancements of UAV and deep learning technologies for weed management in farmland, Multidiscip. Digit. Publ. Inst. (MDPI) (2024), https://doi.org/10.3390/agronomy14030494.
[58] L. Feng, S. Chen, C. Zhang, Y. Zhang, Y. He, A comprehensive review on recent applications of unmanned aerial vehicle remote sensing with various sensors for

high-throughput plant phenotyping, Comput. Electron. Agric. 182 (August 2020) (2021) 106033, https://doi.org/10.1016/j.compag.2021.106033.
[59] M. Canicattì, M. Vallone, Drones in vegetable crops: a systematic literature review, Smart Agric. Technol. 7 (September 2023) (2024) 100396, https://doi.org/10.1016/j.atech.2024.100396.
[60] A.T. Seiche, L. Wittstruck, T. Jarmer, Weed detection from unmanned aerial vehicle imagery using deep learning–a comparison between high-end and low-cost multispectral sensors, Sensors 24 (5) (2024), https://doi.org/10.3390/s24051544. Mar.
[61] S.K. Phang, T.H.A. Chiang, A. Happonen, M.M.L. Chang, From satellite to UAV-based remote sensing: a review on precision agriculture, IEEE Access 11 (November) (2023) 127057–127076, https://doi.org/10.1109/ACCESS.2023.3330886.
[62] F. Xia, L. Quan, Z. Lou, D. Sun, H. Li, X. Lv, Identification and comprehensive evaluation of resistant weeds using unmanned aerial vehicle-based multispectral imagery, Front. Plant Sci. 13 (July) (2022) 1–15, https://doi.org/10.3389/fpls.2022.938604.
[63] P. Fraccaro, et al., A deep learning application to map weed spatial extent from unmanned aerial vehicles imagery, Remote Sens. 14 (17) (2022) 1–15, https://doi.org/10.3390/rs14174197.
[64] Z. Lou, L. Quan, D. Sun, F. Xia, H. Li, Z. Guo, Multimodal deep fusion model based on transformer and multi-layer residuals for assessing the competitiveness of weeds in farmland ecosystems, Int. J. Appl. Earth Obs. Geoinf. 127 (January) (2024) 103681, https://doi.org/10.1016/j.jag.2024.103681.
[65] N. Amarasingam, et al., Bitou bush detection and mapping using UAV-based multispectral and hyperspectral imagery and artificial intelligence, Remote Sens. Appl. 34 (2024) 101151, https://doi.org/10.1016/j.rsase.2024.101151.
[66] B. Cetinsaya, D. Reiners, C. Cruz-Neira, From PID to swarms: a decade of advancements in drone control and path planning - a systematic review (2013–2023), Swarm Evol. Comput. 89 (December 2023) (2024) 101626, https://doi.org/10.1016/j.swevo.2024.101626.
[67] A. Bouguettaya, H. Zarzour, A. Kechida, A.M. Taberkit, Deep learning techniques to classify agricultural crops through UAV imagery: a review, Neural Comput. Appl. 34 (12) (2022) 9511–9536, https://doi.org/10.1007/s00521-022-07104-9.
[68] B. Costello, et al., Detection of parthenium weed (Parthenium hysterophorus L.) and its growth stages using artificial intelligence, Agriculture 12 (11) (2022) 1–23, https://doi.org/10.3390/agriculture12111838.
[69] H. Huang, J. Deng, Y. Lan, A. Yang, X. Deng, and L. Zhang, "A fully convolutional network for weed mapping of unmanned aerial vehicle (UAV) imagery," 2018. doi: 10.1371/journal.pone.0196302.
[70] H. Huang et al., "A semantic labeling approach for accurate weed mapping of high resolution UAV imagery," 2018. doi: 10.3390/s18072113.
[71] P.K. Yadav, et al., Assessing the performance of YOLOv5 algorithm for detecting volunteer cotton plants in corn fields at three different growth stages, Artif. Intell. Agric. 6 (2022) 292–303, https://doi.org/10.1016/j.aiia.2022.11.005.
[72] A.M. Almalky, K.R. Ahmed, Deep learning for detecting and classifying the growth stages of consolida regalis weeds on fields, Agronomy 13 (3) (2023), https://doi.org/10.3390/agronomy13030934.
[73] R. Sapkota, J. Stenger, M. Ostlie, P. Flores, Towards reducing chemical usage for weed control in agriculture using UAS imagery analysis and computer vision techniques, Sci. Rep. 13 (1) (2023) 1–14, https://doi.org/10.1038/s41598-023-33042-0.
[74] J. Zhao, T.W. Berge, J. Geipel, Transformer in UAV image-based weed mapping, Remote Sens. 15 (21) (2023), https://doi.org/10.3390/rs15215165.
[75] V. Singh, D. Singh, H. Kumar, Efficient application of deep neural networks for identifying small and multiple weed patches using drone images, IEEE Access 12 (2024) 71982–71996, https://doi.org/10.1109/ACCESS.2024.3402213.
[76] M. Pöttker, K. Kiehl, T. Jarmer, D. Trautz, Convolutional neural network maps plant communities in semi-natural grasslands using multispectral unmanned aerial vehicle imagery, Remote Sens. 15 (7) (2023), https://doi.org/10.3390/rs15071945.
[77] Z. Mawardi, D. Gautam, T.G. Whiteside, Utilisation of remote sensing dataset and a deep learning object detection model to map siam weed infestations, IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. (2024), https://doi.org/10.1109/JSTARS.2024.3465554.
[78] I.L. Bretas, et al., Detection and mapping of Amaranthus spinosus L. in bermudagrass pastures using drone imagery and deep learning for a site-specific weed management, Agron. J. 116 (3) (2024) 990–1002, https://doi.org/10.1002/agj2.21545.
[79] A. Feng, J. Zhou, E. Vories, K.A. Sudduth, Evaluation of cotton emergence using UAV-based imagery and deep learning, Comput. Electron. Agric. 177 (May) (2020) 105711, https://doi.org/10.1016/j.compag.2020.105711.
[80] I. Gallo, A.U. Rehman, R.H. Dehkordi, N. Landro, R.La Grassa, M. Boschetti, Deep object detection of crop weeds: performance of YOLOv7 on a real case dataset from UAV images, Remote Sens. 15 (2) (2023) 1–17, https://doi.org/10.3390/rs15020539.
[81] L. Quan, et al., Multimodal remote sensing application for weed competition time series analysis in maize farmland ecosystems, J. Environ. Manag. 344 (June) (2023) 118376, https://doi.org/10.1016/j.jenvman.2023.118376.
[82] M.F. Danilevicz, et al., Segmentation of sandplain lupin weeds from morphologically similar narrow-leafed lupins in the field, Remote Sens. 15 (7) (2023) 1–15, https://doi.org/10.3390/rs15071817.
[83] M.D. Bah, A. Hafiane, R. Canals, Deep learning with unsupervised data labeling for weed detection in line crops in UAV images, Remote Sens. 10 (11) (2018) 1690, https://doi.org/10.3390/rs10111690.
[84] B.B. Sapkota, et al., Use of synthetic images for training a deep learning model for weed detection and biomass estimation in cotton, Sci. Rep. 12 (1) (2022) 1–18, https://doi.org/10.1038/s41598-022-23399-z.
[85] Y. Guo, Y. Zhao, T.A. Rothfus, A.S. Avalos, A novel invasive plant detection approach using time series images from unmanned aerial systems based on convolutional and recurrent neural networks, Neural Comput. Appl. 34 (22) (2022) 20135–20147, https://doi.org/10.1007/s00521-022-07560-3.
[86] R. Reedha, E. Dericquebourg, R. Canals, A. Hafiane, Transformer neural network for weed and crop classification of high resolution UAV images, Remote Sens. 14 (3) (2022), https://doi.org/10.3390/RS14030592.
[87] C. Hu, J.A. Thomasson, M.V. Bagavathiannan, A powerful image synthesis and semi-supervised learning pipeline for site-specific weed detection, Comput. Electron. Agric. 190 (September) (2021) 106423, https://doi.org/10.1016/j.compag.2021.106423.
[88] S. Wang et al., "Weed density extraction based on few-shot learning through UAV remote sensing RGB and multispectral images in ecological irrigation area," 2022. doi: 10.3389/fpls.2021.735230.
[89] C. Xiaoming, et al., An improved algorithm based on YOLOv5 for detecting Ambrosia trifida in UAV images, Front. Plant Sci. 15 (2024), https://doi.org/10.3389/fpls.2024.1360419.
[90] R. Gao, Y. Jin, X. Tian, Z. Ma, S. Liu, Z. Su, YOLOv5-T: a precise real-time detection method for maize tassels based on UAV low altitude remote sensing images, Comput. Electron. Agric. 221 (2024), https://doi.org/10.1016/j.compag.2024.108991.
[91] N. Genze, M. Wirth, C. Schreiner, R. Ajekwe, M. Grieb, D.G. Grimm, Improved weed segmentation in UAV imagery of sorghum fields with a combined deblurring segmentation model, Plant Methods 19 (1) (2023), https://doi.org/10.1186/s13007-023-01060-8.
[92] W. Hu, et al., Review of deep learning-based weed identification in crop fields, Int. J. Agric. Biol. Eng. 16 (4) (2023), https://doi.org/10.25165/j.ijabe.20231604.8364.
[93] B. Sapkota, V. Singh, C. Neely, N. Rajan, M. Bagavathiannan, Detection of Italian ryegrass in wheat and prediction of competitive interactions using remote-sensing and machine-learning techniques, Remote Sens. 12 (18) (2020), https://doi.org/10.3390/RS12182977.
[94] A. dos Santos Ferreira, D.M. Freitas, G.G. da Silva, H. Pistori, M.T. Folhes, Unsupervised deep learning and semi-automatic data labeling in weed discrimination, Comput. Electron. Agric. 165 (July) (2019) 104963, https://doi.org/10.1016/j.compag.2019.104963.
[95] S.I. Moazzam, T. Nawaz, W.S. Qureshi, U.S. Khan, M.I. Tiwana, A W-shaped convolutional network for robust crop and weed classification in agriculture, Precis. Agric. 24 (5) (2023) 2002–2018, https://doi.org/10.1007/s11119-023-10027-7.
[96] M. Krestenitis, et al., CoFly-WeedDB: a UAV image dataset for weed detection and species identification, Data Brief 45 (2022) 108575, https://doi.org/10.1016/j.dib.2022.108575.
[97] J. Valente, S. Hiremath, M. Ariza-Sentís, M. Doldersum, L. Kooistra, Mapping of Rumex obtusifolius in nature conservation areas using very high resolution UAV imagery and deep learning, Int. J. Appl. Earth Obs. Geoinf. 112 (June) (2022) 0–6, https://doi.org/10.1016/j.jag.2022.102864.
[98] N. Rai, et al., Multi-format open-source weed image dataset for real-time weed identification in precision agriculture, Data Brief 51 (2023), https://doi.org/10.1016/j.dib.2023.109691.
[99] M. Jonak, J. Mucha, S. Jezek, D. Kovac, K. Cziria, SPAGRI-AI: smart precision agriculture dataset of aerial images at different heights for crop and weed detection using super-resolution, Agric. Syst. 216 (2024), https://doi.org/10.1016/j.agsy.2024.103876.
[100] Q. Wang, M. Cheng, S. Huang, Z. Cai, J. Zhang, H. Yuan, A deep learning approach incorporating YOLO v5 and attention mechanisms for field real-time detection of the invasive weed Solanum rostratum Dunal seedlings, Comput. Electron. Agric. 199 (January) (2022) 107194, https://doi.org/10.1016/j.compag.2022.107194.
[101] A. Etienne, A. Ahmad, V. Aggarwal, D. Saraswat, Deep learning-based object detection system for identifying weeds using uas imagery, Remote Sens. 13 (24) (2021) 1–22, https://doi.org/10.3390/rs13245182.
[102] B. Guo, S. Ling, H. Tan, S. Wang, C. Wu, D. Yang, Detection of the grassland weed phlomoides umbrosa using multi-source imagery and an improved YOLOv8 network, Agronomy 13 (12) (2023) 1–18, https://doi.org/10.3390/agronomy13123001.
[103] Q. Wang, et al., An image segmentation method based on deep learning for damage assessment of the invasive weed Solanum rostratum Dunal, Comput. Electron. Agric. 188 (June) (2021) 106320, https://doi.org/10.1016/j.compag.2021.106320.
[104] E.K. Raptis, et al., CoFly: an automated, AI-based open-source platform for UAV precision agriculture applications, SoftwareX 23 (2023) 101414, https://doi.org/10.1016/j.softx.2023.101414.
[105] B. Xu, J. Fan, J. Chao, N. Arsenijevic, R. Werle, Z. Zhang, Instance segmentation method for weed detection using UAV imagery in soybean fields, Comput. Electron. Agric. 211 (June) (2023) 107994, https://doi.org/10.1016/j.compag.2023.107994.
[106] Y. Cai, et al., Attention-aided semantic segmentation network for weed identification in pineapple field, Comput. Electron. Agric. 210 (June) (2023) 107881, https://doi.org/10.1016/j.compag.2023.107881.
[107] I. Koshelev, M. Savinov, A. Menshchikov, A. Somov, Drone-aided detection of weeds: transfer learning for embedded image processing, IEEE J. Sel. Top. Appl.

Earth Obs. Remote Sens. 16 (2023) 102–111, https://doi.org/10.1109/JSTARS.2022.3224657.

[108] J. Valente, M. Doldersum, C. Roers, L. Kooistra, R. Sensing, and I. T. Group, "Detecting rumex obtusifolius weed plants in grasslands from UAV," vol. IV, no. June, pp. 10–14, 2019.

[109] L. Gutiérrez-Lazcano, C.J. Camacho-Bello, E. Cornejo-Velazquez, J.H. Arroyo-Núñez, M. Clavel-Maqueda, Cuscuta spp. segmentation based on unmanned aerial vehicles (UAVs) and orthomasaics using a U-Net xception-style model, Remote Sens. 14 (17) (2022) 1–12, https://doi.org/10.3390/rs14174315.

[110] N. Rai, X. Sun, WeedVision: a single-stage deep learning architecture to perform weed detection and segmentation using drone-acquired images, Comput. Electron. Agric. 219 (2024), https://doi.org/10.1016/j.compag.2024.108792.

[111] F. Tian, C.J. Ransom, J. Zhou, B. Wilson, K.A. Sudduth, Assessing the impact of soil and field conditions on cotton crop emergence using UAV-based imagery, Comput. Electron. Agric. 218 (November 2023) (2024) 108738, https://doi.org/10.1016/j.compag.2024.108738.

[112] S. Khan, M. Tufail, M.T. Khan, Z.A. Khan, J. Iqbal, M. Alam, A novel semi-supervised framework for UAV based crop/weed classification, PLoS ONE 16 (2021) 2021, https://doi.org/10.1371/journal.pone.0251008.

[113] S.D. Khan, S. Basalamah, A. Lbath, Weed–crop segmentation in drone images with a novel encoder–decoder framework enhanced via attention modules, Remote Sens. 15 (23) (2023) 1–18, https://doi.org/10.3390/rs15235615.

[114] N. Belissent, J.M. Peña, G.A. Mesías-Ruiz, J. Shawe-Taylor, M. Pérez-Ortiz, Transfer and zero-shot learning for scalable weed detection and classification in UAV images, Knowl. Based Syst. 292 (2024), https://doi.org/10.1016/j.knosys.2024.111586.

[115] C. Lu, E. Nnadozie, M.P. Camenzind, Y. Hu, K. Yu, Maize plant detection using UAV-based RGB imaging and YOLOv5, Front. Plant Sci. 14 (January) (2023) 1–13, https://doi.org/10.3389/fpls.2023.1274813.

[116] L. Petrich et al., "Detection of Colchicum autumnale in drone images, using a machine-learning approach," 2020. doi: 10.1007/s11119-020-09721-7.

[117] J. Weyler, T. Labe, F. Magistri, J. Behley, C. Stachniss, Towards domain generalization in crop and weed segmentation for precision farming robots, IEEE Robot. Autom. Lett. 8 (6) (2023) 3310–3317, https://doi.org/10.1109/LRA.2023.3262417.

[118] O.H.Y. Lam, et al., An open source workflow for weed mapping in native grassland using unmanned aerial vehicle: using Rumex obtusifolius as a case study, Eur. J. Remote Sens. 54 (supp 1) (2021) 71–88, https://doi.org/10.1080/22797254.2020.1793687.

[119] J. Deng, Z. Zhong, H. Huang, Y. Lan, Y. Han, Y. Zhang, Lightweight semantic segmentation network for real-time weed mapping using unmanned aerial vehicles, Appl. Sci. 10 (20) (2020) 1–11, https://doi.org/10.3390/app10207132.

[120] H. Huang, Y. Lan, A. Yang, Y. Zhang, S. Wen, J. Deng, Deep learning versus object-based image analysis (OBIA) in weed mapping of UAV imagery, Int. J. Remote Sens. 41 (9) (2020) 3446–3479, https://doi.org/10.1080/01431161.2019.1706112.

[121] A. Pandey, K. Jain, An intelligent system for crop identification and classification from UAV images using conjugated dense convolutional neural network, Comput. Electron. Agric. 192 (November 2021) (2022) 106543, https://doi.org/10.1016/j.compag.2021.106543.

[122] J. Cervantes, J. Taltempa, F. García-Lamont, J.S.R. Castilla, A.Y. Rendon, L. D. Jalili, Análisis Comparativo de las técnicas utilizadas en un Sistema de Reconocimiento de Hojas de Planta, RIAI - Rev. Iberoam. Autom. Inform. Ind. 14 (1) (2017) 104–114, https://doi.org/10.1016/j.riai.2016.09.005.

[123] I. García-Santillán, M.R. Pusdá-Chulde, G. Pajares, Identificación automática de vegetación utilizando imágenes agrícolas: una revisión de métodos, Tecnol. Apl. Ing. (June 2017) (2016) 155–161.

[124] V. Nguyen Thanh Le, B. Apopei, K. Alameh, Effective plant discrimination based on the combination of local binary pattern operators and multiclass support vector machine methods, Inf. Process. Agric. 6 (1) (2019) 116–131, https://doi.org/10.1016/j.inpa.2018.08.002.

[125] B. Liu, R. Bruch, Weed detection for selective spraying: a review, Curr. Robot. Rep. 1 (1) (2020) 19–26, https://doi.org/10.1007/s43154-020-00001-w.

[126] V.N.T. Le, S. Ahderom, K. Alameh, Performances of the lbp based algorithm over cnn models for detecting crops and weeds with similar morphologies, Sensors 20 (8) (2020), https://doi.org/10.3390/s20082193.

[127] A. Bojeri, F. Melgani, G. Giamotta, G. Ristorto, G. Guglieri, J. Marcato Junior, Automatic crop rows segmentation for multispectral aerial imagery, in: Proceedings of the IEEE Mediterranean and Middle-East Geoscience and Remote Sensing Symposium (M2GARSS), 2022, pp. 78–81.

[128] S. Park, N.W. Park, Effects of class purity of training patch on classification performance of crop classification with convolutional neural network, Appl. Sci. 10 (11) (2020), https://doi.org/10.3390/app10113773.

[129] A.K. Dutta, Y. Albagory, A.R.W. Sait, I.M. Keshta, Autonomous unmanned aerial vehicles based decision support system for weed management, Comput. Mater. Contin. 73 (1) (2022) 899–915, https://doi.org/10.32604/cmc.2022.026783.

[130] T. de Camargo, M. Schirrmann, N. Landwehr, K.H. Dammer, M. Pflanz, Optimized deep learning model as a basis for fast UAV mapping of weed species in winter wheat crops, Remote Sens. 13 (9) (2021), https://doi.org/10.3390/rs13091704.

[131] K. Vinueza, L. Sandoval-Pillajo, A. Giret-Boggino, D. Trejo-España, M. Pusdá-Chulde, I. García-Santillán, Automatic weed quantification in potato crops based on a modified convolutional neural network using drone images, Data Metadata 4 (2025), https://doi.org/10.56294/DM2025194.

[132] A. dos Santos Ferreira, D. Matte Freitas, G. Gonçalves da Silva, H. Pistori, M. T. Folhes, Weed detection in soybean crops using ConvNets, Comput. Electron. Agric. 143 (October) (2017) 314–324, https://doi.org/10.1016/j.compag.2017.10.027.

[133] O.L. García-Navarrete, A. Correa-Guimaraes, L.M. Navas-Gracia, Application of Convolutional Neural Networks in Weed Detection and Identification: A Systematic Review, Multidisciplinary Digital Publishing Institute (MDPI), 2024, https://doi.org/10.3390/agriculture14040568.

[134] L. Lei, Q. Yang, L. Yang, T. Shen, R. Wang, C. Fu, Deep learning implementation of image segmentation in agricultural applications: a comprehensive review, Artif. Intell. Rev. 57 (6) (2024), https://doi.org/10.1007/s10462-024-10775-6.