International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(19s) 2026 pp. 522-544
https://doi.org/10.70917/ijcisim-2026-5034
Article

Machine Learning and Deep Learning Techniques for Weed Detection in Agricultural Fields: Recent Advances, Challenges, and Future Directions

Akanksha Bodhale[1], Dr. Seema Verma[2]

[1]Research Scholar, Department of Computer Science Engineering, Banasthali Vidyapith, Rajasthan, 304022, India
Email: akansha2611@gmail.com, ORCID ID- 0009-0008-6726-9641
[2]Professor, Department of Electrical & Electronics Engineering, Natinal Institute of Technical Teachers Training and Research, Bhopal 462003, India
Email: sverma@nitttrbpl.ac.in, ORCID ID- 0000-0001-6299-2576

Abstract: Weed infestation is a key biotic limitation on agricultural productivity. Weeds compete with crops for light, water, nutrients and space, while increasing production costs and, in many instances, functioning as reservoirs for pests. Traditional weed management is mainly dependent on manual, mechanical and blanket chemical control. Herbicides are effective; yet, indiscriminate use increases input cost, accelerates herbicide resistance development, and raises environment and food safety problems. Precision agriculture so increasingly relies on computer vision, machine learning (ML), deep learning (DL), unmanned aerial vehicles (UAVs) and agricultural robotics to detect weeds and to assist site-specific treatment. This critical review summarizes the evolution of handcrafted color, texture, shape, spatial and spectral features with support vector machines, random forests, k-nearest neighbors and artificial neural networks to modern convolutional neural networks, YOLO-family object detectors, encoder-decoder segmentation models, transformers and hybrid architectures. The entire operational pipeline is investigated, including sensing, picture pre-processing, augmentation, feature representation, classification, detection, segmentation, model evaluation, edge deployment, and closed-loop actuation. Recent literature demonstrates that in the real-time object identification, YOLO versions are leading the way while U-Net and other networks are still dominating in the semantic segmentation and ResNet / VGG family are commonly employed in classification. However, the outstanding performance in the datasets often deteriorates in the real field settings because of the illumination variations, occlusion, crop-weed morphological resemblance, small target size, growth stage variations, class imbalance, motion blur, and geographic or seasonal domain shift. High-priority research directions include cross-domain generalization, diversity of datasets, external validation, uncertainty-aware prediction, open-set recognition, multimodal sensing, self-supervised learning, domain adaptation, lightweight edge inference and integration with robotic or variable-rate weed control. The primary takeaway is that future advancement must go beyond incremental gains in benchmark accuracy toward robust, safe, energy-efficient, economically feasible and agronomically verified intelligent weed-management systems.
Keywords: precision agriculture, weed identification, machine learning, deep learning, computer vision, YOLO, semantic segmentation, UAV, agricultural robots, edge AI, domain adaptability, transformer

1. Introduction

Agricultural production is transitioning from uniform, input-intensive management to spatially selective and more autonomous decision systems. The move is driven by the desire to boost productivity, reduce superfluous chemical inputs, labor, environmental consequences and production costs. This transition is particularly suited to weed management since weed populations are spatially heterogeneous: the density, species composition, development stage and competitiveness of the weed population vary widely within a single area. Thus, a blanket therapy overlooks information that can be exploited to decrease inputs without reducing control efficacy.

Weeds compete directly with crops for sun radiation, moisture, nutrients and physical growing space. Severe infestations can prevent crop establishment, impede with harvesting, and cause yield losses. Historically weed management was via physical removal, tillage and mechanical cultivation, cultural methods and herbicides. Chemical weed control has lowered labor needs and protected yield but consistent treatment frequently targets areas where weeds are absent or below economically significant thresholds. This inefficiency raises the cost and selection pressure for herbicide resistant populations and leads to concerns about off-target movement, residues, biodiversity loss and contamination of soil and water.

One alternative is site-specific weed management. Intelligent systems do not treat the field as a homogeneous entity, but identify weed location, density or species and provide chemical, mechanical, electrical, thermal or laser treatment just where necessary. It is consequently the perception layer of a smart weeding system to reliably discriminate crops from weeds. The reliability of subsequent decisions to spray or remove mechanically depends on the accurate identification and localization of the plants.

The evolution of precision agriculture has led to a technical ecosystem with the potential to enable automated weed management. RGB, multispectral, hyperspectral, Depth and LIDAR sensors can be mounted on the tractor, ground robots, handheld devices or UAVs. These sensing platforms can be integrated with GNSS, embedded CPUs, actuators, and cloud/edge computing. At the same time, the developments in ML and DL have altered the agricultural image interpretation [1-3]. Previous methods used manually constructed color, shape, texture, edge and vegetation-index features with classifiers. Such pipelines can work in controlled settings but often fail when the image qualities vary. DL has minimized the need of hand-crafted feature engineering by learning hierarchical representations directly from photos [4-6].

This is a rapidly growing field as evidenced by recent polls. Hasan et al. [7-9] provided a review of deep learning methods for picture based weed detection. Hu et al. [29] calculated weed identification in-crop in big scale grain production. A 2024 survey by Bakhshipour et al. [38] gave a comprehensive overview of ML and DL techniques, and recent 2025 studies focused authentic-field validation, UAV images, transformers, lightweight architectures and real-world smart weeders [10–13]. In the 2025 authentic-farmland evaluation, 349 works were reviewed, and it was stated that YOLO variations were employed in more than half of the detection implementations and U-Net was the most used segmentation family [14-16]. These changes suggest that the research frontier is transitioning from proof-of-concept classification to deployment-oriented detection and segmentation.

Still, great reported accuracy should not be mistaken for operational readiness. Agricultural settings contain combinations of characteristics that are rarely seen together in selected benchmark datasets. Leaves overlap, weeds and crops can have similar colors and forms, lighting can change in seconds, wind causes movement, soil and residue differ from field to field, seedlings vary rapidly with stage of growth and overhead imagery has very small targets [17-19]. Therefore, it may be that a model is successful on random train-test splits of one dataset, but fails when applied to another farm, season, crop type, camera or nation.

This research thus highlights the difference between benchmark performance and deployable robustness. It also gives a critical comparison of traditional ML and modern DL, relating sensing and data quality to model behavior, and discussing classification, object detection, semantic and instance segmentation, as well as assessing emerging approaches like transformers, self-supervised learning, domain adaptation, generative data augmentation, multimodal fusion, open-set recognition and edge AI [20-22]. The final aim is to identify research priorities that allow a reliable translation from picture identification to agronomically effective weed control technologies [23-25].

2. Framework and Scope of Review

The review focuses on automatic detection, recognition, location, classification, segmentation, density estimate and mapping of weeds in crop agricultural situations. Studies that only consider crop disease diagnosis, nutritional deficit, fruit quality, livestock or non-agricultural optimization algorithms are conceptually excluded as they do not include crop-weed discrimination [26-29]. We discuss the historical literature when it defined methods or datasets, but focus more on innovations after 2015 and notably on breakthroughs from 2021-2026 in real-time detection, UAV sensing, transformers, authentic-field datasets and edge deployment [30-33].

The source manuscript searched combinations of the phrases weed detection, weed classification, weed recognition, deep learning, deep neural network, and machine learning throughout Scopus, Web of Science, IEEE Xplore, ScienceDirect, SpringerLink, Google Scholar, MDPI, and similar academic sources. The conceptual search vocabulary was updated to include semantic segmentation, instance segmentation, YOLO, transformer, UAV, agricultural robot, edge AI, domain adaptability, and precision weed management. The update reflects recent review and representative research literature through 2025 and early 2026. For this reason, the current work is presented as a structured critical review and not a falsely claimed PRISMA-compliant systematic review, as the original manuscript failed to maintain a fully reproducible record-by-record screening log.

The review is structured around six questions: Which image processing and ML strategies are applied in the crop-weed discrimination? DL architectures have changed the game in classification, detection, and segmentation. Which sensor platforms and datasets have the biggest impact on performance? Besides classification accuracy, what measures do we need to evaluate deployment? What causes performance degradation in real field conditions? Which upcoming technology are most likely to enhance reliability and scalability?

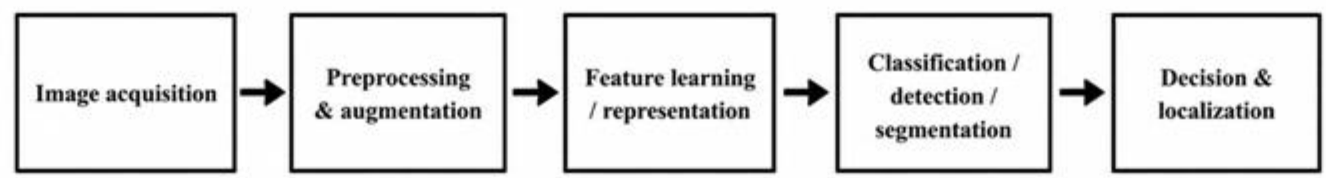

Figure 1. End-to-end intelligent weed detection and site-specific weed management system.

3. Weed Detection as a Machine-Vision Problem

An automated weed-recognition system can be represented as a sequence of interconnected stages: image acquisition, annotation, preprocessing, augmentation, feature representation, model training, classification/detection/segmentation, evaluation, deployment, and weed-control action. The importance of each stage depends on the final agronomic task. A smartphone classifier labeling an isolated weed leaf is fundamentally different from a robotic sprayer needing to identify numerous partially occluded seedlings while moving through a crop row [34-36].

Classification assigns an image or patch to a category such as crop, weed, or weed species. It is appropriate when the target plant has already been isolated, but it does not provide precise spatial coordinates. Object detection simultaneously predicts class and location, generally through bounding boxes, and is therefore highly relevant to robotic spraying or mechanical weeding. Semantic segmentation assigns a class to every pixel and is useful for weed-cover estimation and precise maps. Instance segmentation additionally distinguishes individual plants belonging to the same class, which can be advantageous when neighboring weeds overlap.

These tasks impose different annotation and computational burdens. Image-level classification requires the least annotation; bounding boxes require spatial labeling; pixel-level masks are more expensive. The choice among them should therefore be based not only on achievable accuracy but on the information required by the downstream control mechanism.

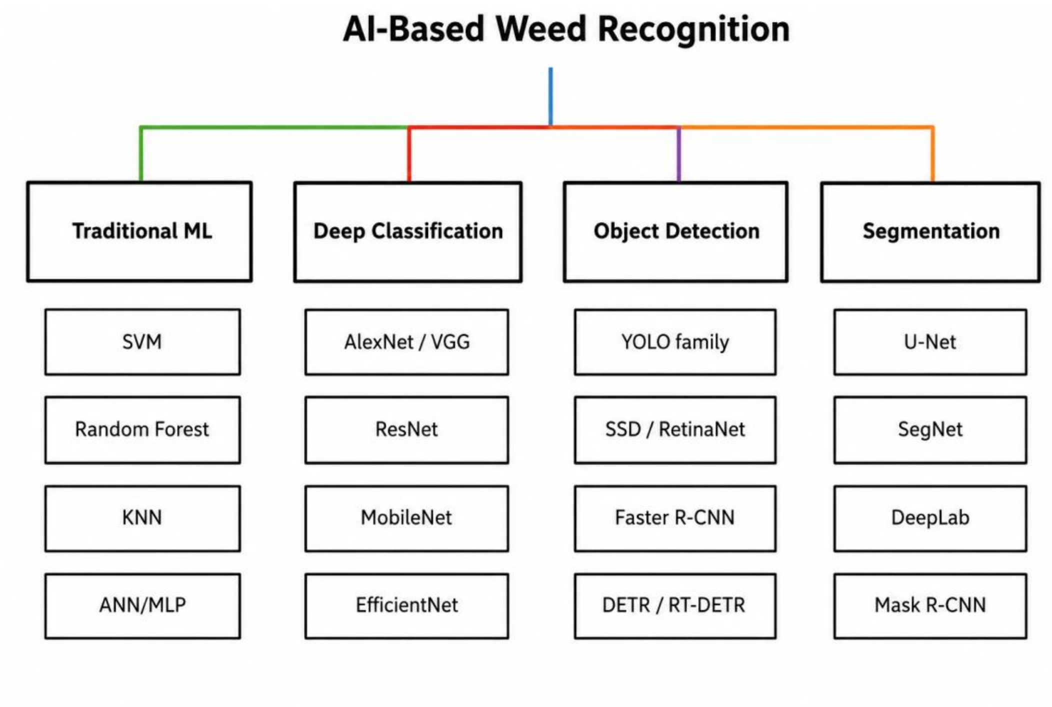

Figure 2. Taxonomy of main machine-learning and deep-learning methods for the detection of weeds.

4. Imaging and sensing platforms

4.1 Imaging in RGB

RGB cameras are the most used sensors in weed recognition investigations due to their low cost, low weight, wide availability and compatibility with embedded real-time systems. Color information is frequently adequate to identify green vegetation from soil and may also help species discrimination. The downside of RGB imaging is its sensitivity to illumination, camera white balance, shadows, reflections, and visually similar species. When the training set does not reflect true field variation, models can unknowingly acquire correlations including background or illumination conditions.

4.2 Imaging using UAVs

The UAVs provide fast spatial coverage and enable field scale weed mapping. Multirotor systems are especially popular since they give flexibility for low altitude operations. Recent evaluations show that RGB sensors Recent evaluations indicate that RGB sensors remain the most commonly used, while multispectral sensors and LiDAR integration are on the rise [40, 41]. used, with an increasing use of multispectral sensors and LiDAR integration [40, 41]. Weed detection with UAVs is a small object problem, because seedlings may only occupy a few pixels at realistic flight heights. Other factors, including motion blur, elevation changes, wind, camera angle, orthomosaic errors, and uneven illumination, make detection much more difficult [37-39].

4.3 Imaging: ground based and robot

Proximal imaging can be performed with tractors or ground robots with higher spatial resolution than aerial capture and suitable for small weeds and plant-level treatment [40]. But the camera moves through a complex canopy all the time, so latency, vibration, motion blur, calibration, and camera-to-actuator sync are all critical. The vision system should be rapid enough to provide therapy at the right physical site and not merely provide accurate offline labels [41].

4.4 Hyperspectral and Multispectral Imaging

Multispectral sensors use the reflectance beyond the typical visible bands, e.g., the near-infrared bands. Hyperspectral devices record hundreds of tiny wavelength bands and can identify differences between species that appear identical in RGB photos. A review of hyperspectral weed detection in 2024 underlined the usefulness of spectral markers for early identification but also the cost of the sensors, data volume, calibration, and processing complexity [42]. These modalities are promising when morphology alone is not sufficient but additional accuracy is needed. These modalities are promising when morphology alone is not enough but the extra accuracy needs to be justified with hardware and computational cost.

4.5 Depth, stereoscopic vision and LiDAR

Depth and stereo information can be used to discriminate plant structures from soil and to estimate height or three-dimensional layout. LiDAR can offer geometric measures and can be merged with RGB or spectral data to enhance structural understanding. Thus, multimodal fusion is intriguing for robust field perception, but sensor synchronization, calibration, cost and data-transfer needs pose great challenges [43-45].

5. Image Preprocessing, Data Augmentation and Feature Representation

5.1 Resize, Crop and Normalize

Images are scaled to meet the network input requirements and lower the processing cost. The high resolution keeps the fine details needed to detect little weeds, but it also increases memory and latency. Thus, cropping and tiling are commonly utilized for aerial photographs. Normalization normalizes the optimization process and lowers the variance of the numerical range of the sensors.

5.2 Image Denoising and Enhancement

Weak or noisy images can be improved via median and Gaussian filtering, contrast enhancement, histogram equalization, adaptive enhancement and color correction. Such modifications must be used sparingly, since excessive augmentation may produce artificial edges or alter biologically relevant color information.

5.3 Color Space Conversion

Conventional systems commonly convert RGB images into HSV, Lab, grayscale, or vegetation indices as the hue or excess-green representations can better segment vegetation. Such modifications may still be beneficial as auxiliary channels for DL, but are no longer required as deep networks can learn discriminative combinations directly from raw input [46-48].

5.4 Augmentation of Data

Rotation, flipping, scaling, translation, cropping, blur, brightness and contrast variations, noise injection, MixUp, Mosaic and synthetic picture synthesis can help reduce overfitting and imitate field variability. The augmentation must be agronomically possible. Variations in brightness, shadow, scale, direction, partial occlusion and background conditions are especially good for a model that will be used outside.

5.5 Hand-Crafted vs. Learned Features

Traditional ML involves feature engineering. First, color histograms, texture statistics, geometric descriptors, HOG, spectral signatures and crop-row position are computed and then fed to a classifier. DL substitutes much of this chain with learned hierarchical characteristics. This shift is one of the key reasons why CNNs out-perform classical approaches in the presence of complicated image variations, while hand-crafted features may still be effective for small datasets or where interpretability is necessary [49].

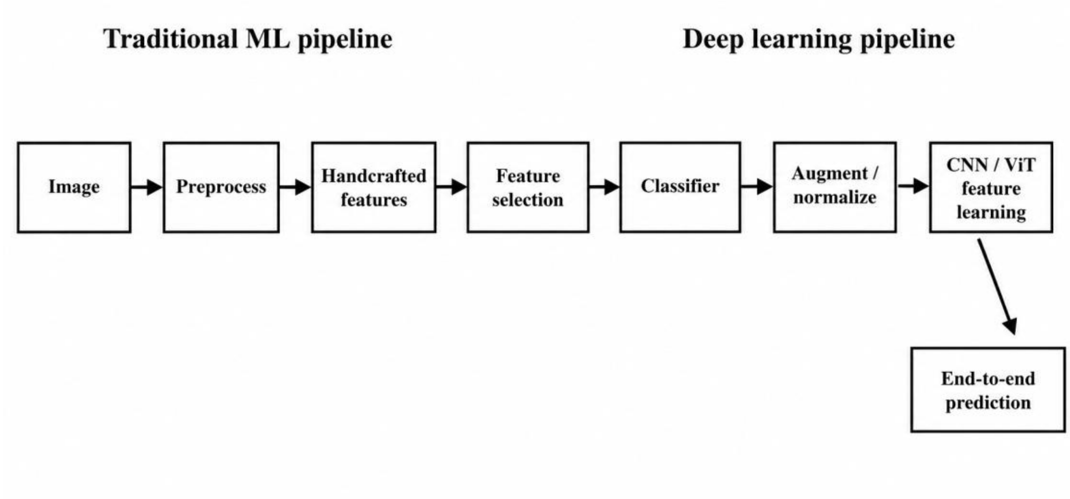

Figure 3. Traditional machine-learning and end-to-end deep-learning pipelines.

Table 1. Main families of handmade features applied to weed recognition.

Feature family	Typical descriptors	Strength	Major limitation
Colour	RGB/HSV statistics, colour moments, excess-green indices	Low computational demand; intuitive	Sensitive to illumination and camera settings
Texture	GLCM, entropy, contrast, energy, LBP	Captures spatial leaf patterns	Depends on scale, orientation, and image quality
Shape	Area, perimeter, moments, contour, eccentricity	Biologically interpretable	Strongly affected by occlusion and overlapping leaves
Gradient	HOG, edge orientation, boundary descriptors	Captures local structure	Affected by clutter and noise
Spectral	Band ratios, NIR/hyperspectral signatures	Useful for visually similar species	Higher sensor and calibration cost
Spatial/contextual	Crop-row position, spacing, neighborhood structure	Exploits agronomic context	Less transferable between planting systems

6. Traditional Machine Learning Methods

6.1 Support Vector Machine

Support vector machines have been frequently employed for crop-weed discrimination because of their good performance in high-dimensional feature spaces and ability to deal with intermediate datasets. Kernel functions allow non-linear decision boundaries. The performance is nonetheless highly dependent on the feature design, scaling and parameter selection. In earlier field investigations, SVMs have proved competitiveness when color, shape, texture or spectral features are precisely constructed [50].

6.2 Random Forests

Random forests are an ensemble of many decision tree models and can mix heterogeneous predictors. They are appealing in precision agriculture because of their simplicity to train, interpretability of feature importance and relative resilience. In one UAV-based comparison described in the source literature, RF obtained 96% accuracy, whereas SVM and KNN produced accuracies of 94% and 63%, respectively [51]. These comparisons indicate that traditional ML can still compete when the features are correctly designed. However, the finding should not be extrapolated other datasets.

6.3 K-Nearest Neighbors

KNN is simple, non-parametric yet computationally expensive for larger reference dataset. Performance is sensitive to the distance measure, feature scaling, and choice of k. It is thus better suited as a baseline or for compact feature spaces than as the heart of a huge real-time vision system.

6.4 MLPs and Artificial Neural Networks

Shallow neural networks model nonlinear relationships between handcrafted features but do not have the spatial inductive bias of convolutional structures. MLPs have demonstrated useful performance in hyperspectral research when informative spectral representations are provided [52].

Table 2. Comparative properties of classical ML classifiers for weed detection.

Method	Advantages	Limitations	Best-use scenario
SVM	Strong with moderate data; nonlinear kernels; effective high-dimensional classification	Feature engineering and parameter tuning required	Handcrafted colour/shape/spectral features
Random Forest	Robust; interpretable feature importance; handles mixed predictors	Larger forests increase memory and latency	Structured features and moderate-size datasets
KNN	Simple; no complex training	Slow inference on large datasets; sensitive to scaling	Baseline classification with compact feature sets
ANN/MLP	Nonlinear decision boundaries; flexible	Requires engineered inputs; weaker spatial modeling than CNNs	Spectral or tabular feature classification

7. Weed Classification Using Deep Learning

7.1 Convolutional Neural Networks 7.1.1 Introduction

CNNs learn local spatial filters and visual representations in a hierarchical manner. The initial layers detect edges and textures; the deeper layers encode more abstract patterns. Architectures like AlexNet, VGG, ResNet, DenseNet, MobileNet and EfficientNet have been successfully adapted for crop-weed classification. The potential of ConvNets for soybean weed recognition was proved by Ferreira et al. [53]. Transfer learning has since become a staple across agricultural vision.

7.2 Learning Transfer

Transfer learning applies a pretrained network on a large image dataset to a weed-recognition task. It decreases training time and can improve performance when agricultural labels are scarce. However, the generic pretraining does not alleviate the agricultural domain shift. Pretrained models on natural images may not capture the subtle plant morphology, repeated row structure or sensor specific spectral characteristics. Thus, a promising next step is domain-specific pretraining on large collections of unlabeled agricultural imagery.

7.3 Light-Weight Classification Networks

The architectures MobileNet, EfficientNet, ShuffleNet and relatives reduce computation by using depthwise separable convolution, compound scaling or efficient channel operations. They are useful for embedded devices with limited power and memory Instead of simply counting parameters, their practical value should be evaluated by measuring latency and energy on target hardware.

8. Object Detection in Deep Learning

8.1 Two-Stage Detectors 8.1.1 R-CNN

Two-stage detectors, e.g. Faster R-CNN, first generate candidate regions, and then classify and refine boxes. They can achieve strong localization but generally require more processing than the one-stage detectors. They are still useful when precision is more critical than very high throughput, however real-time robotic systems frequently use quicker alternatives.

8.2 YOLO Family and Real-time Detector

YOLO has become the dominant family for real-time object identification in agriculture due to its ability to forecast bounding boxes and class labels in a single forward pass. In the 2025 authentic-farmland evaluation [54], YOLO versions were used in 53.48% of the detection methods. Successive versions have made improvements in backbone design, feature aggregation, training strategy, loss functions and small-object capacity. Weed-specific research usually incorporates attention modules, higher-resolution detecting heads, or feature-fusion blocks to tackle challenging targets.

8.3 Small object detection

Young weeds are small relative to the overall image, especially with UAV images. Progressive downsampling can be used to erase visual evidence. Typical solutions are multi-scale feature pyramids, shallow detection heads, high-resolution input, picture tiling, super-resolution, attention, and customizable losses. The goal is to keep fine detail without inference being too slow for field operations. **8.4 Transformers-Based Detectors**

DETR and its derivatives use transformer attention for modeling global interactions and simplifying portions of the detection pipeline. Marijuana detection using transformers is still less common than YOLO but gaining rapidly [55-56]. Comparative studies in precision agriculture highlight that if the computational cost and training behavior are consistent with the deployment platform, transformer detectors could be competitive [57].

9. Semantic + Instance Segmentation

9.1 U-Net and Encoder-Decoder Architectures

U-Net are an encoder decoder network with skip links securing high-resolution spatial details. Crop-weed segmentation requires accurate boundary detection and this design is good for it. U-Net is still one of the most popular segmentation families in recent assessments [58].

9.2 SegNet and DeepLab

SegNet uses encoder pooling indices for upsampling, while DeepLab uses atrous convolution and multiscale context. These techniques produce dense pixel predictions and can facilitate weed-cover estimation or selective treatment mapping.

9.3 Instance Segmentation

Generally, Mask R-CNN and its variants produce a mask each plant instance. The extra information is helpful for plant-level treatment and overlapping weeds, but requires more comprehensive annotation and computing resource than bounding-box detection.

9.4 Direct vs. Indirect Detection of Weeds

Direct detection tags the weed. Indirect detection identifies crop plants first, and all other vegetation is considered weeds. Indirect methods may reduce the number of weed species to explicitly model, but are not effective when crops are omitted or when volunteer plants and non-target vegetation are present. Strategy depends on crop geometry and control risk.

Table 3. Families of models and their applicability for intelligent weed management.

Task	Representative models	Output	Principal application	Key trade-off
Classification	ResNet, VGG, MobileNet, EfficientNet	Image/patch label	Species recognition	No direct location information
One-stage detection	YOLO, SSD, RetinaNet, EfficientDet	Boxes + labels	Real-time spraying/robotics	Speed versus small-object accuracy
Two-stage detection	Faster R-CNN	Boxes + labels	High-quality localization	Higher latency
Semantic segmentation	U-Net, SegNet, DeepLab	Pixel classes	Coverage and weed maps	High annotation cost
Instance segmentation	Mask R-CNN, instance variants	Individual masks	Plant-level treatment	High computation and annotation
Transformer detection	DETR, RT-DETR, DINO variants	Boxes + labels	Global-context detection	Training and hardware demands

10. Exemplary Evidence from the Literature

The reported literature shows that both classic ML and DL can achieve excellent performance when datasets are well matched to the training distribution. Islam et al. applied RF, SVM and KNN on UAV crop-weed images and achieved 96%, 94% and 63% accuracy correspondingly [59]. Boyina et al. used CNN/InceptionV3 features with U-Net-oriented processing and reported above 90% weed detection on aerial imagery [60]. Wang et al. demonstrated that the encoder-decoder segmentation using NIR information improved the crop-weed separation under uncontrolled illumination. Beeharry and Bassoo reported a considerable performance gap between AlexNet and a standard ANN on their UAV data set. Such results demonstrate the strength of representation learning, but they should not be directly compared unless datasets, classes, splits and metrics are the same.

Recent works are progressively evaluated in terms of mAP, IoU, F1-score, model size and real-time latency. The 2026 Frontiers review highlights representative experiments in object detection, including a 2024 YOLOv8_NLB model for winter-wheat weeds with reported mAP of 88.3%. The same paper describes a rising range of object recognition, segmentation, and classification algorithms and identifies data scarcity, annotation cost, morphological variability, and deployment limits as persistent issues. The 2025 UAV study also reports YOLO for detection, U-Net for semantic segmentation, Mask R-CNN for instance segmentation, and a tendency toward transformer and hybrid architectures.

Hence, these investigations are rightly interpreted comparatively rather than absolutely. The model trained on isolated plants or highly curated patches has a much simpler task compared to that in dense field canopies. Random splits of images may also distribute near-duplicate frames or plants between training and test sets, which may lead to optimistic accuracy. Cross-field or cross-season validation is a better indicator of operational potential.

Table 4. Representative studies of weed detection and methodological focus

Reference	Crop / setting	Model / method	Reported result / emphasis	Key observation
Islam et al. (2021) [15]	Australian chilli / UAV	RF, SVM, KNN	96%, 94%, 63% accuracy	Classical ML can remain competitive with strong features
Li et al. (2021) [12]	Pasture weeds / hyperspectral	MLP + spectral features	89.1% accuracy reported	Spectral information supports fine discrimination
Boyina et al. (2021) [14]	Broad-leaf weeds / aerial images	CNN, InceptionV3, U-Net-oriented processing	>90% detection reported	Transfer learning and augmentation improve aerial recognition
Wang et al. (2020) [21]	Crop and weed / outdoor illumination	Encoder-decoder segmentation + NIR	96.1% average accuracy reported	NIR can improve segmentation under difficult illumination
Beeharry & Bassoo (2020) [22]	Soybean / UAV	ANN vs AlexNet	AlexNet 99.8% on test set	Deep features strongly outperform shallow ANN in that dataset
Li et al. (2024), summarized in [40]	Winter wheat weeds	YOLOv8_NLB	mAP 88.3%	Modern YOLO variants target real-time field detection
Allmendinger et al. (2025) [35]	Precision weed detection	YOLO and transformer detectors	Comparative real-time evaluation	Architecture choice should consider both accuracy and latency

11. Annotation and Benchmark Design of Datasets

11.1 Variety of datasets

Robust datasets should contain differences in species of crops and weeds, plant age, weed density, field location, soil texture, residue, camera systems, acquisition height, viewing angle, time of day, cloud cover, shadows, occlusion, motion blur and season. The data from one field in one day is insufficient to characterize the deployment domain. In the authentic-farmland review of 2025, 34 publicly available crop-weed datasets were identified, and the gap between controlled datasets and real field variability was highlighted [39].

11.2 Annotation Load

The cost of annotation increases from picture categorization to bounding boxes to pixel masks. Species that are morphologically similar may require expert knowledge to be properly identified. Weak supervision, semi-supervised learning, active learning, and foundation-model-assisted annotating can help to lessen this burden.

11.3 Class Imbalance Issue

Agricultural datasets are often dominated by common crops and have few examples of rare weeds. Imbalance can lead to the models being skewed toward the common classes. Typical methods include class-weighted losses, focal loss, balanced sampling, targeted data collecting, and synthetic augmentation.

11.4 Data Leakage Prevention

Images taken from the same film, adjacent frames, or repeated observations on the same plants are highly associated. If they are split arbitrarily across training and test sets, performance is exaggerated. Partitioning should be at the level of field, plot, date, or acquisition session if possible. External datasets are even more powerful in terms of generalization estimation.

12. Metrics & Reporting Expectations

Accuracy is not enough especially in the case of class imbalance. Accuracy is the number of correctly predicted weed. Recall is the number of true weeds found. These are combined taking the harmonic mean to give the F1-score. Common metrics for detection are Intersection over Union (IoU), Average Precision (AP) and mean Average Precision (mAP). For segmentation, IoU/Jaccard, Dice coefficient and mean IoU are often utilized.

Deployment evaluation must go beyond computer-vision measures. Real-time field systems should report frames per second, inference latency, model parameters, FLOPs when appropriate, memory footprint, processor resolution, device type, and energy utilization. If the detector controls a sprayer or mechanical tool, end-to-end treatment accuracy, missed-weed rate, crop false-positive rate, actuator latency, and input savings are more important than image metrics alone.

A four-level evaluation framework is recommended: within-dataset performance; cross-field/cross-season robustness; hardware deployment performance; and agronomic field outcomes. This hierarchy inhibits claims of practical preparedness based only on a high validation-set accuracy.

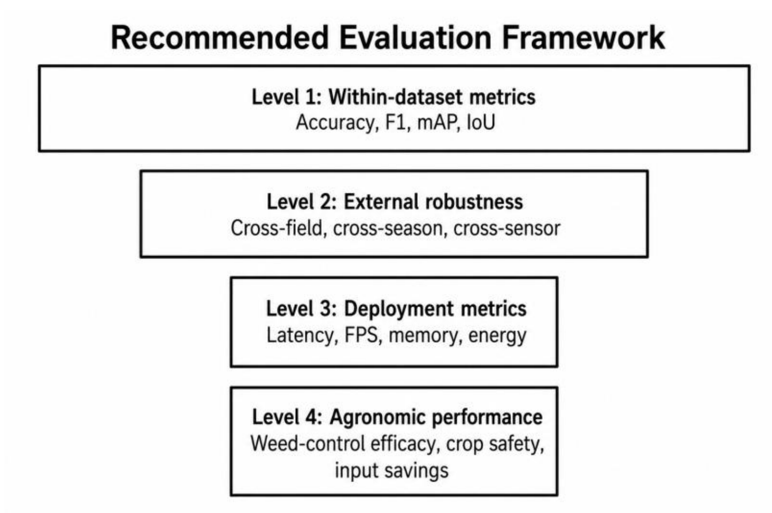

Figure 4. Proposed multilayer evaluation methodology for intelligent weed recognition systems.

Table 5. Recommended measures for performance for weed detection investigations

Evaluation level	Metrics	What it measures	Minimum reporting recommendation
Classification	Accuracy, precision, recall, F1, confusion matrix	Class discrimination	Always report class-wise precision/recall when imbalance exists
Object detection	IoU, AP, mAP@0.5, mAP@0.5:0.95	Localization and classification	Define IoU protocol and input resolution
Segmentation	IoU, mIoU, Dice, pixel accuracy	Pixel-level delineation	Report per-class IoU and boundary errors where relevant
Deployment	Latency, FPS, parameters, memory, energy	Real-time feasibility	Report actual hardware and batch size
Agronomic	Weed-control efficacy, crop injury, herbicide/energy savings	Operational value	Required for claims of field readiness

13. Major Challenges in Deployment in Real-World

13.1 Morphological Similarity between Crop and Weed

Early weed detection is agronomically advantageous, however in the seedling stage, weeds are often visually identical to crops. Separation can be improved using fine-grained morphology, spectral cues, temporal context, row geometry and metric learning, but remains a key source of mistake.

13.2 Occlusion

As the canopy develops, the leaves interlock and the outline of the plants is lost. Instance segmentation, temporal observations with a moving platform, synthetic occlusion augmentation and multi-view sensing are interesting options.

13.3 Variability of Shadow and Illumination

Sunlight, cloud, shadow and reflections change the external scene quickly. Color normalization and augmentation can diminish sensitivity, but real multi-condition training data and cross-time validation are more reliable than preprocessing alone.

13.4 Variation with growth stage

The plant is changing so fast. Models trained on mature plants may fail on seedlings and vise versa. Therefore, longitudinal data sets should sample different crop and weed growth stages.

13.5 Domain Transfer based on Geography and Season

Soil, crop cultivar, weed community, climate, and management vary from region to region. Models can learn local background correlations, and can degrade dramatically out of their training domain. Domain adaptation and domain generalization tackle the problem directly.

13.6 UAV Scale and Small Targets

Small weeds may cover a few pixels. Higher resolution gives more visibility, but costs more processing. Key decisions at the system level include multi-scale feature fusion, tiling, small-object detection heads and optimized flight height.

13.7 Vibration of the Platform and Motion Blur

Robots and UAVs acquire images while moving. Shorter exposure, stabilization, faster shutter speed, temporal fusion and motion-blur augmentation can improve robustness.

13.8 Computational and Energy Constraints

Large networks developed on desktop GPUs may be unsuitable for embedded field devices. Pruning, quantization, distillation, lightweight backbones, neural architecture search, and hardware-aware training can improve deployment efficiency.

13.9 Generalization and External Validation

Generalization is arguably the most important unresolved problem. Many models are evaluated exclusively on held-out chunks of the same dataset. A robust examines within-dataset evaluation, then cross-field or cross-season evaluation, and lastly a totally external data or operational field deployments**.**

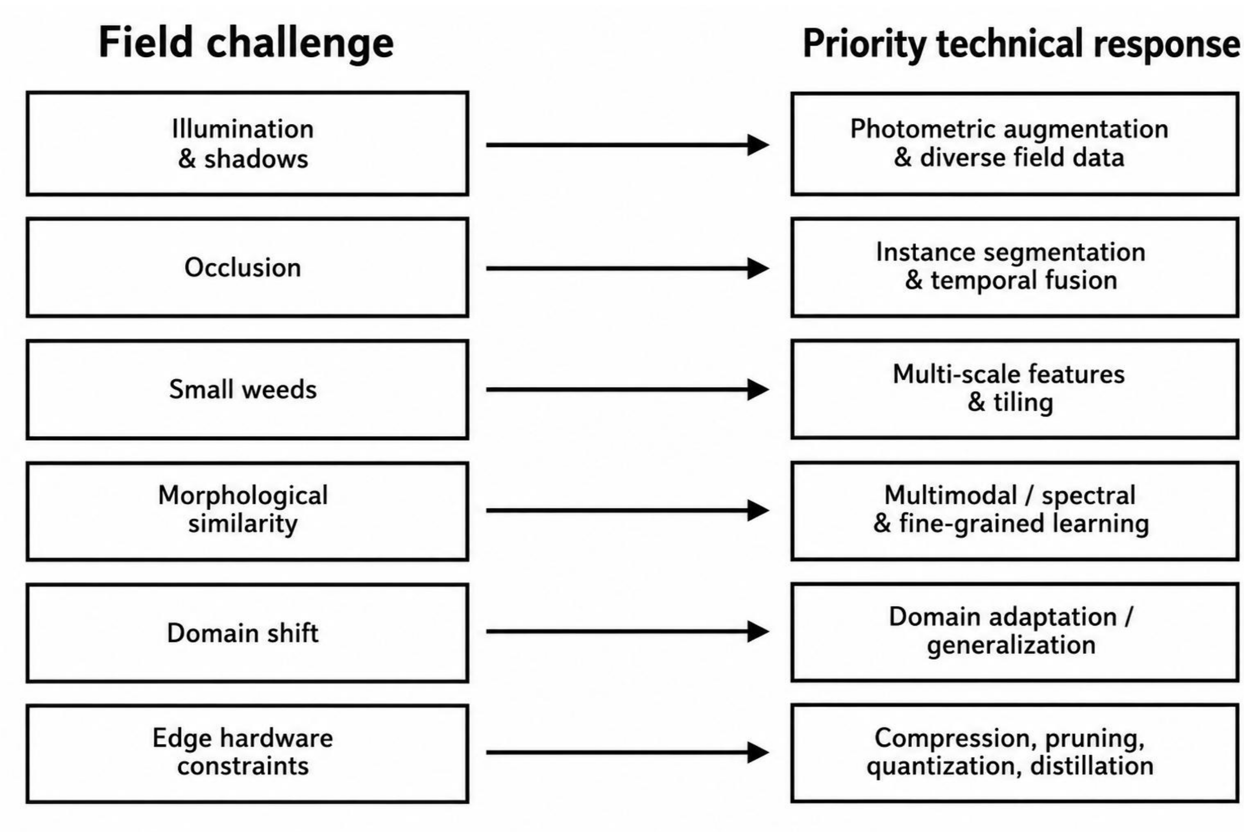

Figure 5. Main real world problems and accompanying high priority approaches remedies.

Table 6. Main challenges, implications and research objectives.

Challenge	**Consequence**	**Priority response**
Illumination/shadows	False classification and unstable colour features	Diverse field data, photometric augmentation, spectral cues
Occlusion	Missed or merged plants	Instance segmentation, multi-view/temporal fusion

Small weeds	Low recall in UAV and wide-view imagery	Multi-scale features, tiling, high-resolution detection heads
Morphological similarity	Crop-weed class confusion	Fine-grained learning, multispectral/hyperspectral fusion
Domain shift	Poor cross-farm performance	Domain adaptation/generalization, external validation
Limited labels	Overfitting and poor rare-class learning	Self-/semi-supervised learning, active learning
Edge limits	High latency and power use	Pruning, quantization, distillation, lightweight networks
Unknown species	Unsafe closed-set decisions	Open-set recognition and uncertainty calibration

14. Emerging and Future Research Directions

14.1 Vision Transformers and Hybrid Architectures

Transformers can replicate long-range interactions vs self-attention and provide worldwide crop-row and canopy context. Recent evaluations indicate an expanded use of DETR, Swin Transformer, DINO, RT-DETR and hybrid CNN-transformer models [39-44]. Hybrid systems can combine efficient local convolution with global attention but precision gains must offset increased delay.

14.2 Attention Mechanisms

Channel, spatial, coordinate, and self-attention modules are commonly incorporated into YOLO and CNN architectures to focus critical plant sections. However, tiny single-dataset improvements should not be viewed automatically as substantial originality. New modules should be assessed across external datasets and hardware metrics.

14.3 Domain Adaptation and Domain Generalization

Domain adaptation moves a model from a source field or sensor to a different target distribution, sometimes without target labels. Domain generalization instead trains across several known domains and aims to perform directly on an unseen domain. Both are highly relevant because farm-to-farm transfer remains difficult.

14.4 Self-Supervised and Semi-Supervised Learning

Agricultural platforms can collect enormous volumes of unlabeled imagery. Self-supervised pretraining can learn agricultural representations from these photos, while semi-supervised approaches only use a small labeled set and a huge unlabeled collection. These approaches are directly related to annotation cost.

14.5 Few Shot Learning

Rare weeds may be represented by a handful of specimens. Few-shot and metric-learning techniques are designed to identify novel species with few labels and could enhance transferability to spatially specific weed ecosystems.

14.6 Synthetic Data and GenAI

GANs and diffusion models can generate or edit crop-weed scenes, augment rare classes, or simulate occlusion and illumination. Visually plausible images can still contain unrealistic statistics, and thus synthetic data should be evaluated by improvement on authentic external test sets.

14.7 Edge AI, Compression and Knowledge Distillation

Local inference lowers network dependency and latency. Deployment on embedded GPUs or NPUs can be achieved through model trimming, quantisation, knowledge distillation, efficient backbones, and custom accelerators. Future studies should disclose energy per accurately recognised or treated weed, not just the model's size.

14.8 Multimodal Fusion

RGB, spectral, depth, LiDAR, and location information. The fundamental technical question is not whether more sensors improve recognition, but which minimal sensor combination provides a substantial increase in robustness per unit cost, power, and computational strain.

14.9 Explainable and Uncertainty-Aware AI

Explainability can also uncover shortcut learning; for example, if a model learns to use a soil backdrop instead of leaf structure. More importantly, uncertainty estimates can prevent dangerous actuation. A sensible control policy is as follows: treat high-confidence weeds, protect high-confidence crops, treat low-confidence weeds, protect high-confidence crops, and either defer treatment or take another observation for low-confidence cases.

14.10 Foundation Models for Open-Set Recognition

Real fields. The fields contain unknown weeds, volunteer crops, wounded plants, and unusual growth forms. Closed-set models force these objects into predefined classes. Therefore, safe autonomous treatment requires open-set and out-of-distribution recognition. Foundation vision models may speed up annotation and few-shot adaptability, but agriculture-specific pretraining and model compression will be needed for field implementation.

15. From Weed Recognition to Closed-Loop Weed Control

The ultimate goal of weed detection is not an image metric but an agronomic action. An intelligent weeding system is composed of sensing, perception, decision, localization, actuation, and verification in its entirety. A detector may correctly detect a weed, but treatment may fail because of camera-actuator misalignment, vehicle motion, nozzle delay, wind, mechanical positioning error, or confusing plant identity.

Hence, field studies should include end-to-end measures, including the proportion of weeds successfully treated, crop injury rate, herbicide-volume reduction, untreated weed rate, field capacity, energy use per hectare, and cost per hectare. This connects AI performance to farmer value and environmental impact.

This view also influences the choice of models. A detector with a slightly lower mAP but much shorter latency is useful for a moving sprayer. Similarly, a bigger model is warranted if it will reduce crop injury or safely handle new species. So system-wide optimization is more important than isolated network ranking.

16. Translational Synthesis: What Has Actually Advanced from 2021 to 2026

16.1 From Image Classification to Spatially Explicit Recognition

A significant change in recent studies is from image-level crop vs. weed classification to spatially explicit detection and segmentation. Classification helps to identify species, but a field robot needs to know where the target is, its coordinates, and geographic extent. This operational requirement: The rapid expansion of YOLO-based detection and U-Net-type segmentation mirrors this operational requirement. This operational requirement is mirrored by the rapid expansion of YOLO based detection and U-Net type segmentation.

This is important from a scientific perspective as a high classification score on cropped plant images cannot be assumed to translate into accurate localization in dense, cluttered crop scenes. Modern studies are increasingly expected to show that a network can identify small weeds, separate overlapping plants, and maintain acceptable throughput on the hardware intended for field use.

16.2 From Accuracy Maximization to Accuracy-Efficiency Trade-offs

Previous weed recognition papers often only reported a single headline accuracy. Recent work increasingly acknowledges that model quality is multi-dimensional. A robotic detector has to trade off precision, recall, localization error, latency, memory and power use. Recent YOLO modifications and lightweight CNNs thus aim at a Pareto improvement rather than a single metric: improved small-object recall without excessive model size, or reduced latency without an unacceptable fall in accuracy. This is especially true for battery-powered robots and UAVs, where the computational demand directly impacts the endurance. As such, Q1-level studies should feature direct comparisons of model size and inference speed, and should not characterize a network as 'lightweight' without measurements on a defined device.

16.3 Transition from Managed Data to Actual Farmland

The key difference recently is not architecture but methodology. Researchers pay closer attention to real agricultural. In the 2025 state-of-the-art review on proximal imagery, it was revealed that while hundreds of model-development studies have used intelligent weeders, only a subset of the broader literature has actually applied intelligent weeders in real fields [39]. This finding points to a translational gap. Controlled datasets are useful for reproducibility but flatten the variability that determines success in agriculture. Real validation has mixed stages of growth, partial visibility, residue, soil heterogeneity, fluctuating sunshine, wind, and plant density. A mature literature should therefore distinguish between controlled benchmark performance and field deployment, rather than treat them as identical evidence.

16.4 Shift toward UAV-Robot Complementarity

UAVs and ground strong should not be considered competing sensing systems. UAVs are efficient for rapid field-scale scouting, infestation mapping, and georeferenced prioritizing, while ground platforms provide the spatial precision and actuator closeness required for plant-level intervention. A hierarchical systems design is promising. UAV photography identifies weed hotspots and provides prescription maps, and then ground robots or variable-rate sprayers execute high-resolution local recognition and treatment. Such a strategy can minimize the quantity of ground traversal and focus computing where it is of greatest agronomic value. Compared to standalone platform studies, coordinated aerial-ground systems have not been studied as extensively and present a significant possibility for investigation.

16.5 Shift toward Multimodal and Context-Aware Perception

Although cost and convenience still make RGB imaging the dominating source, challenging instances are increasingly pushing the uses of multispectral, hyperspectral, depth, LiDAR, or contextual data. Interest in spectrum information for early weed recognition. The 2024 hyperspectral study [42] demonstrates interest in using spectrum information for early weed identification. in the 2024 hyperspectral study [42]. But multimodal systems must be constructed with economic discipline. A sensor that only improves recognition by a small amount but doubles the hardware cost or power demand may not be practical for commercial deployment. Hence, future study should not only evaluate fusion accuracy but also incremental value: the gain derived by adding more sensors compared to cost, calibration work, processing load, and robustness.

16.6 Shift toward Safer Autonomous Decisions

Autonomous therapy alters the risk profile of weed recognition. In a passive monitoring system, an incorrect prediction leads to a lousy map, but in an autonomous sprayer or laser weeder it can harm the crop. This points out the importance of uncertainty calibration, open-set recognition, and conservative actuation. A high confidence threshold is not sufficient if confidence is improperly measured. Models should be evaluated on plants that are out of distribution and in atypical field circumstances. High ambiguity may mean that the safest action is to withhold treatment, re-observe the plant from another vantage point, or flag the situation for human evaluation. This safety perspective is mainly missing in the previous literature and should be a key necessity of future autonomous systems.

16.7 Reproducibility and Comparability as Research Infrastructure

Other big need is the standardized reporting of experiments. Two published mAP values are not always comparable due to variations in image quality, augmentation, data division, IoU threshold, hardware, and field circumstances. Public datasets and open code enhance reproducibility, but benchmark techniques are just as crucial. Field independent splits Multiple domains, challenging negative instances Publish both computer vision and deployment measurements. A good future benchmark should contain. Providing challenge datasets where the concealed test set comes from farms or seasons not represented in training would be a service to the community. Like procedures would reward true generalization (not memorizing dataset-specific backgrounds).

16.8 Economic and Environmental Evidence

The third translational challenge is whether intelligent weed identification provides measurable benefit to agriculture. An often-cited environmental benefit is the reduction of blanket spraying. But the extent of that benefit is dependent on infestation density, detector recall, false-positive treatment, machinery energy uses, and the control method. The economic analysis should consider camera and processor expenses, depreciation of the robot or UAV, maintenance, labeling and model update costs, labor savings, herbicide reduction, field capacity, and crop loss saving.

Environmental assessment should consider chemical reduction together with power or fuel use. Integrating these dimensions will help distinguish technically impressive prototypes from technologies capable of sustained adoption by farmers and service providers.

Table 7. Research focus changes of intelligent weed recognition.

Earlier emphasis	Current / emerging emphasis	Reason for the transition
Cropped image classification	Detection and instance/semantic segmentation	Robotic treatment requires spatial target information
Single accuracy metric	mAP/F1 plus latency, memory and energy	Real-time hardware imposes operational constraints
Random within-dataset splits	Cross-field, cross-season and external validation	Measures domain generalization rather than dataset memorization
Ground or UAV systems in isolation	Hierarchical UAV + ground-robot workflows	Combines field coverage with plant-level precision
RGB-only perception	Selective multimodal sensing and contextual fusion	Supports difficult crop-weed discrimination
Closed-set forced classification	Open-set and uncertainty-aware decisions	Reduces risk of autonomous crop injury
Model-centric novelty	Data, deployment and agronomic novelty	Better alignment with practical scientific value

17. Research Gaps in the Consolidated

1. Reliance on curated or single-field datasets overly restricts evidence of broad generalization.

2. Real-world testing of integrated intelligent weeders is still far less prevalent than offline model validation.

3. Benchmark comparisons are hindered by differing crops, weed species, sensor resolutions, data splits, metrics and technology.

4. Geographic and seasonal variety is lacking in many public databases.

5. Morphological similarity at the most desirable time for treatment makes early stage crop-weed differentiation still challenging.

6. Open-set capacity and uncertainty calibration are seldom integrated into autonomous treatment decisions.

7. Multimodal sensing is intriguing, but there is no uniform cost-benefit assessment.

8. Energy, memory and end-to-end latency are commonly neglected in edge deployment studies.

9. Agronomic validation—weed control efficacy, crop safety, herbicide savings and economics—is underreported relative to vision metrics.

10. Incremental network alterations are generally tested on one dataset only, providing insufficient proof that the proposed novelty is generalizable.

18. Research Roadmap Proposed

Stage 1—Better data: Build datasets that are multi-location, multi-season, multi-growth-stage, and multi-sensor, with realistic weed density, shadows, residue, occlusion, and motion.

Stage 2—Improved validation: moving from exclusive reliance on random image splits to plot-independent, field-independent, season-independent and external-dataset testing.

Stage 3—Generalizable learning: Focus on self-supervised learning, domain adaptation, domain generalization, few-shot learning, open-set recognition, and agriculture-specific foundation representations.

Stage 4--Efficient deployment: use distillation, pruning, quantization, efficient transformers, model compression, and hardware-aware training to maximize accuracy, latency, memory, and energy simultaneously.

Stage 5—Closed-loop agronomic intelligence: link perception to treatment and assess whole systems with computer-vision, operational, environmental, and economic parameters.

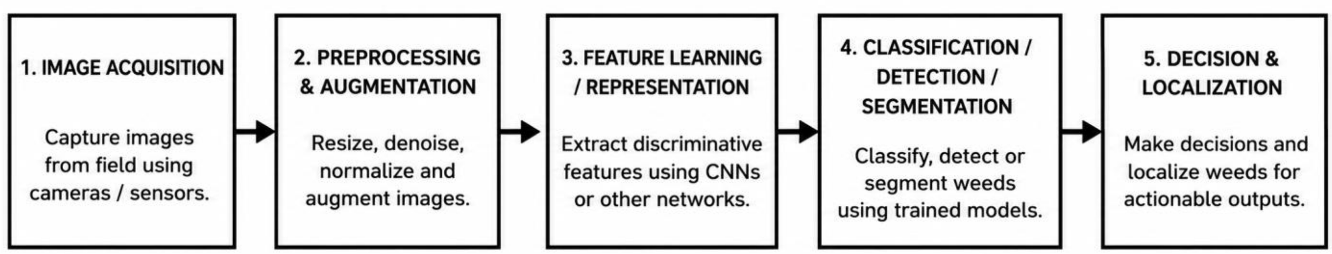

Fig. 6. Proposed five-stage roadmap toward the next generation of smart weed management

Table 8 Future research path for strong and deployable weed recognition systems.

Priority	**Current limitation**	**Recommended objective**
Dataset development	Narrow field and seasonal diversity	Multi-country, multi-season, multi-sensor benchmarks
Evaluation	Random train/test splits dominate	External domain and field-level testing
Learning paradigm	Heavy supervised dependence	Self-/semi-supervised and few-shot learning
Unknown species	Closed-set classifiers	Open-set and out-of-distribution recognition
Reliability	Uncalibrated confidence	Uncertainty-aware decision policies

Sensors	Predominantly RGB	Cost-effective multimodal fusion
Hardware	Desktop-GPU benchmarking	Edge deployment with energy reporting
Application	Offline detection	Closed-loop robotic or variable-rate control
Impact	Accuracy/mAP emphasis	Agronomic, environmental, and economic evaluation

19. Discussion

We can look at the development of automated weed identification as four technological steps. The first one was based on image processing and handcrafted segmentation of vegetation. The second introduced statistical and ML classifiers including SVM, RF, KNN and shallow ANN models. The third, which is now prevailing, learns representations directly from images using CNNs, one-stage object detectors and encoder-decoder segmentation networks. We are witnessing the emergence of a fourth phase, where the primary goal is not deeper models but deployable, generalizable agricultural intelligence.

YOLO is still the leader for rapid detection, and U-Net is in a strong position for segmentation in this growth period. Transformers, attention, self-supervised learning, domain adaptability, generative data, and multimodal sensing are opening up design space. But architectural newness will not in itself resolve the fundamental challenge of the field. Many performance concerns are really data and assessment errors: lack of diversity, coupled train-test splits, poor external validation, and inadequate alignment of image metrics with agronomic outputs.

Q1 research should therefore identify architectural originality, data novelty, and deployment novelty. A carefully curated independent field trial or a well-organized cross-country weed dataset is more scientifically useful than a small attention block that improves mAP by less than one point on a single benchmark. For example, a system that reliably minimizes pesticide use at operating speed may be more important than a larger network that achieves the best laboratory score.

The findings also signal a change from model-centric optimization to system-centric robustness. The proper unit of study is not the neural network in isolation but the sensing-perception-decision-actuation chain. The sensor selection, camera height, image resolution, model delay, vehicle speed, treatment technique, and safety policy are related. Thus, the ideal architecture for weed recognition is application-dependent.

"Claims of sustainability need to be proven, not assumed." Precision detection may reduce unnecessary herbicide use, but the environmental value is contingent upon the energy cost and accuracy of the whole platform. Future work needs to assess input savings, treatment efficacy, crop damage, battery or fuel use, and operating expenses. The holistic assessment will facilitate the transfer of intelligent weed management from promising prototypes to ordinary agricultural practice.

20. Future Studies Minimum Reporting Recommendations

1. Description of source of dataset. Origin, crop and weed species, development stage, sensor, acquisition height, image resolution, environmental conditions, and annotation procedure.

2. Split training, validation and test data by field, plot, date or acquisition session if possible to avoid leakage.

3. Report precision, recall, F1 , confusion matrices, mAP or mIoU if applicable instead of only accuracy.

4. For real-time systems, specify input resolution, hardware, latency, FPS, parameters, memory and energy/power if applicable.

5. Error analysis per weed species, growth stage, object size, severity of illumination and occlusion.

6. Offer external, cross-domain validation when claims of generality are made.

7. consideration of uncertainty and unknown classes before autonomous chemical or mechanic treatment.

8. Share code, trained weights, datasets and annotation techniques where legal and ethical limitations permit.

9. Connect computer vision metrics to agronomics and treatment efficiency, crop damage, input savings and cost.

21. Conclusion

ML and DL have transformed automated weed detection from a primarily manual image-processing task into a more advanced component of precision agriculture. Traditional ML algorithms using color, texture, shape, and spectral and spatial aspects are still effective when data and computational resources are constrained. But DL has dominated as CNNs learn hierarchical representations directly from imagery; transfer learning reduces reliance on very large labeled agricultural datasets; YOLO-family detectors provide real-time localization; and encoder-decoder networks such as U-Net allow for pixel-level weed mapping. Ground robots and UAVs have significantly improved weed recognition, evolving from static image categorization to more dynamic, spatially explicit systems capable of potentially autonomous management. However, despite these advancements, achieving reliable perception in real-world agricultural settings continues to be a challenge. Some of the factors that often lead to degraded performance outside the training distribution are occlusion, lighting, morphological resemblance, growth-stage variation, small targets, class imbalance, motion blur, shortage of dataset, geographical domain shift, and embedded hardware constraints. Thus, high scores on random splits of curated datasets are not sufficient to indicate practical readiness. The next research step should be on cross-domain generalization, external validation, self-supervised learning, domain adaptation and generalization, open-set recognition, uncertainty-aware prediction, multimodal sensing, transformers, foundation models, synthetic data, model compression, and edge AI. Weed recognition must be viewed as part of a closed-loop agricultural system, where perception is integrated with decision-making, actuation, and verification. Success should be measured not only by accuracy, mAP or IoU, but also by weed-control efficacy, crop safety, savings in herbicide or energy, speed of operation, and economic feasibility. The key science challenge is not just to detect weeds in photos but to build intelligent weed management systems that generalize across fields, interpret ambiguity, run effectively on field hardware, and translate reliable perception into sustainable agronomic action.

REFERENCES

1. Wiles, L. J. (2009). Beyond patch spraying: Site-specific weed management with several herbicides. *Precision Agriculture*. https://doi.org/10.1007/s11119-008-9097-6
2. Asad, M. H., & Bais, A. (2020). Weed detection in canola fields using maximum likelihood classification and deep convolutional neural network. *Information Processing in Agriculture, 7*(4), 535–545. https://doi.org/10.1016/j.inpa.2019.12.002
3. Bah, M. D., Hafiane, A., Canals, R., & Emile, B. (2019). Deep features and one-class classification with unsupervised data for weed detection in UAV images. *2019 Ninth International Conference on Image Processing Theory, Tools and Applications (IPTA)*, 1–5. https://doi.org/10.1109/IPTA.2019.8936091
4. Al-Badri, M., et al. (2022). Classification of weed using machine learning techniques: A review—Challenges, current and future potential techniques. *Journal of Plant Diseases and Protection*. https://doi.org/10.1007/s41348-022-00612-9
5. Hasan, A. S. M. M., Sohel, F., Diepeveen, D., Laga, H., & Jones, M. G. K. (2021). A survey of deep learning techniques for weed detection from images. *Computers and Electronics in Agriculture, 184*, 106067. https://doi.org/10.1016/j.compag.2021.106067
6. Toğaçar, M. (2022). Using DarkNet models and metaheuristic optimization methods together to detect weeds growing along with seedlings. *Ecological Informatics, 68*, 101519. https://doi.org/10.1016/j.ecoinf.2021.101519
7. Du, Y., Zhang, G., Tsang, D., & Jawed, M. K. (2022). Deep-CNN based robotic multi-class under-canopy weed control in precision farming. *2022 IEEE International Conference on Robotics and Automation (ICRA)*, 2273–2279. https://doi.org/10.1109/ICRA46639.2022.9812240
8. Abouzahir, S., Sadik, M., & Sabir, E. (2022). Lightweight computer vision system for automated weed mapping. *2022 IEEE 12th Annual Computing and Communication Workshop and Conference*, 372–376. https://doi.org/10.1109/CCWC54503.2022.9720800
9. Jin, X., Che, J., & Chen, Y. (2021). Weed identification using deep learning and image processing in vegetable plantation. *IEEE Access, 9*, 10940–10950. https://doi.org/10.1109/ACCESS.2021.3050296
10. Moazzam, S. I., et al. (2021). A patch-image based classification approach for detection of weeds in sugar beet crop. *IEEE Access, 9*, 121698–121715. https://doi.org/10.1109/ACCESS.2021.3109015

11. Dyrmann, M., Jørgensen, R. N., & Midtiby, H. S. (2017). RoboWeedSupport—Detection of weed locations in leaf-occluded cereal crops using a fully convolutional neural network. *Advances in Animal Biosciences, 8*(2), 842–847.
12. Li, Y., et al. (2021). Identification of weeds based on hyperspectral imaging and machine learning. *Frontiers in Plant Science, 11*, 611622. https://doi.org/10.3389/fpls.2020.611622
13. Armstrong, D. R., Götz, M., Nardelli, V. C., & Gomes, V. E. O. (2021). Machine learning for identification and classification of crops and weeds. *2021 XI Brazilian Symposium on Computing Systems Engineering*, 1–7. https://doi.org/10.1109/SBESC53686.2021.9628296
14. Boyina, L., Sandhya, G., Vasavi, S., Koneru, L., & Koushik, V. (2021). Weed detection in broad leaves using invariant U-Net model. *International Conference on Communication, Control and Information Sciences*, 1–4. https://doi.org/10.1109/ICCISc52257.2021.9485001
15. Islam, N., Rashid, M., Wibowo, S., Xu, C., & Morshed, A. (2021). Early weed detection using image processing and machine learning techniques in an Australian chilli farm.
16. Siddiqui, S. A., Fatima, N., & Ahmad, A. (2021). Neural network based smart weed detection system. *International Conference on Communication, Control and Information Sciences*, 1–5. https://doi.org/10.1109/ICCISc52257.2021.9484925
17. Junior, L. C. M., & Ulson, J. A. C. (2021). Real-time weed detection using computer vision and deep learning. *2021 14th IEEE International Conference on Industry Applications (INDUSCON)*, 1131–1137. https://doi.org/10.1109/INDUSCON51756.2021.9529761
18. Kamath, R., Balachandra, M., & Prabhu, S. (2021). Classification of crop and weed from digital images: A review. *2021 IEEE International Conference on Distributed Computing, VLSI, Electrical Circuits and Robotics*, 12–17. https://doi.org/10.1109/DISCOVER52564.2021.9663729
19. Selvi, C. T., Sankara Subramanian, R. S., & Ramachandran, R. (2021). Weed detection in agricultural fields using deep learning process. *2021 7th International Conference on Advanced Computing and Communication Systems*, 1470–1473. https://doi.org/10.1109/ICACCS51430.2021.9441683
20. Alam, M., Alam, M. S., Roman, M., Tufail, M., Khan, M. U., & Khan, M. T. (2020). Real-time machine-learning based crop/weed detection and classification for variable-rate spraying in precision agriculture. *2020 7th International Conference on Electrical and Electronics Engineering*, 273–280. https://doi.org/10.1109/ICEEE49618.2020.9102505
21. Wang, A., Xu, Y., Wei, X., & Cui, B. (2020). Semantic segmentation of crop and weed using an encoder-decoder network and image enhancement method under uncontrolled outdoor illumination. *IEEE Access, 8*. https://doi.org/10.1109/ACCESS.2020.2991354
22. Beeharry, Y., & Bassoo, V. (2020). Performance of ANN and AlexNet for weed detection using UAV-based images. *2020 3rd International Conference on Emerging Trends in Electrical, Electronic and Communications Engineering*, 163–167. https://doi.org/10.1109/ELECOM49001.2020.9296994
23. Forero, M. G., Herrera-Rivera, S., Ávila-Navarro, J., Franco, C. A., Rasmussen, J., & Nielsen, J. (2019). Color classification methods for perennial weed detection in cereal crops. In *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications* (pp. 117–123).
24. Shorewala, S., Ashfaque, A., Sidharth, R., & Verma, U. (2021). Weed density and distribution estimation for precision agriculture using semi-supervised learning. *International Journal of Information Technology*. The work was originally released as a semi-supervised weed-density study using Crop Weed Field Image Dataset and sugar-beet imagery.
25. dos Santos Ferreira, A., Matte Freitas, D., Gonçalves da Silva, G., Pistori, H., & Theophilo Folhes, M. (2017). Weed detection in soybean crops using ConvNets. *Computers and Electronics in Agriculture, 143*, 314–324. https://doi.org/10.1016/j.compag.2017.10.027
26. Rai, N., Zhang, Y., Ram, B. G., Schumacher, L., Yellavajjala, R. K., Bajwa, S., & Sun, X. (2023). Applications of deep learning in precision weed management: A review. *Computers and Electronics in Agriculture, 206*, 107698. https://doi.org/10.1016/j.compag.2023.107698
27. McCool, C., et al. (2017). Improved vision-based robotic weed detection through transfer learning. *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*.
28. Lottes, P., Khanna, R., Pfeifer, J., Siegwart, R., & Stachniss, C. (2017). UAV-based crop and weed classification for smart farming. *IEEE International Conference on Robotics and Automation*.
29. Hu, K., Wang, Z., Coleman, G., Bender, A., Yao, T., Zeng, S., Song, D., Schumann, A., & Walsh, M. (2024). Deep learning techniques for in-crop weed recognition in large-scale grain production systems: A review. *Precision Agriculture, 25*, 1–29. https://doi.org/10.1007/s11119-023-10073-1
30. Adhinata, F. D., Wahyono, & Sumiharto, R. (2024). A comprehensive survey on weed and crop classification using machine learning and deep learning. *Artificial Intelligence in Agriculture, 13*, 45–63. https://doi.org/10.1016/j.aiia.2024.06.005
31. Garibaldi-Márquez, F., Valentín-Coronado, L. M., Díaz-Ponce, A., Servín-Palestina, M., García-Hernández, R. V., & Ramos-Cantú, L. (2025). Advances on deep learning for proximal image-based weed recognition and control under authentic farmlands: A state-of-the-art review. *Smart Agricultural Technology, 12*, 101405. https://doi.org/10.1016/j.atech.2025.101405
32. *Weed detection based on deep learning from UAV imagery: A review.* (2025). *Smart Agricultural Technology, 12*, 101147. https://doi.org/10.1016/j.atech.2025.101147

33. Saini, P., & Nagesh, D. S. (2025). A review of deep learning applications in weed detection: UAV and robotic approaches for precision agriculture. *European Journal of Agronomy, 168*, 127652.
34. *Advances in weed identification using hyperspectral imaging: A comprehensive review of platform sensors and deep learning techniques.* (2024). *Journal of Agriculture and Food Research, 18*, 101388. https://doi.org/10.1016/j.jafr.2024.101388
35. Allmendinger, A., Saltık, A. O., Peteinatos, G. G., Stein, A., & Gerhards, R. (2025). Assessing the capability of YOLO- and transformer-based object detectors for real-time weed detection. *Precision Agriculture, 26*, 52. https://doi.org/10.1007/s11119-025-10246-0
36. Darbyshire, M., Coutts, S., Bosilj, P., Sklar, E., & Parsons, S. (2024). Review of weed recognition: A global agriculture perspective. *Computers and Electronics in Agriculture, 227*, 109499. https://doi.org/10.1016/j.compag.2024.109499
37. Pai, D. G., Kamath, R., & Balachandra, M. (2024). Deep learning techniques for weed detection in agricultural environments: A comprehensive review. *IEEE Access, 12*, 113193–113214. https://doi.org/10.1109/ACCESS.2024.3418454
38. Badgujar, C. M., Poulose, A., & Gan, H. (2024). Agricultural object detection with You Only Look Once (YOLO) algorithm: A bibliometric and systematic literature review. *Computers and Electronics in Agriculture, 223*, 109090. https://doi.org/10.1016/j.compag.2024.109090
39. *Transforming weed management in sustainable agriculture with artificial intelligence: A systematic literature review towards weed identification and deep learning.* (2024). *Crop Protection, 176*, 106522. https://doi.org/10.1016/j.cropro.2023.106522
40. Li, Y., Guo, Z., Sun, Y., Chen, X., & Cao, Y. (2024). Weed detection algorithms in rice fields based on improved YOLOv10n. *Agriculture, 14*(11), 2066. https://doi.org/10.3390/agriculture14112066
41. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems, 30*.
42. Carion, N., Massa, F., Synnaeve, G., Usunier, N., Kirillov, A., & Zagoruyko, S. (2020). End-to-end object detection with transformers. *European Conference on Computer Vision*, 213–229.
43. Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 234–241.
44. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 770–778.
45. Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., & Adam, H. (2017). MobileNets: Efficient convolutional neural networks for mobile vision applications. *arXiv preprint arXiv:1704.04861*.
46. Tan, M., & Le, Q. V. (2019). EfficientNet: Rethinking model scaling for convolutional neural networks. *Proceedings of the International Conference on Machine Learning*, 6105–6114.
47. Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster R-CNN: Towards real-time object detection with region proposal networks. *Advances in Neural Information Processing Systems, 28*.
48. He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask R-CNN. *Proceedings of the IEEE International Conference on Computer Vision*, 2961–2969.
49. Chen, L.-C., Papandreou, G., Kokkinos, I., Murphy, K., & Yuille, A. L. (2018). DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 40*(4), 834–848.
50. Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 779–788.
51. Lin, T.-Y., Goyal, P., Girshick, R., He, K., & Dollár, P. (2017). Focal loss for dense object detection. *Proceedings of the IEEE International Conference on Computer Vision*, 2980–2988.
52. Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A. C., Lo, W.-Y., Dollár, P., & Girshick, R. (2023). Segment Anything. *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 4015–4026.
53. Badrinarayanan, V., Kendall, A., & Cipolla, R. (2017). SegNet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 39*(12), 2481–2495. https://doi.org/10.1109/TPAMI.2016.2644615
54. Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., Lin, S., & Guo, B. (2021). Swin Transformer: Hierarchical vision transformer using shifted windows. *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 10012–10022. https://doi.org/10.1109/ICCV48922.2021.00986
55. Jiang, H., Zhang, C., Qiao, Y., Zhang, Z., Zhang, W., & Song, C. (2020). CNN feature based graph convolutional network for weed and crop recognition in smart farming. *Computers and Electronics in Agriculture, 174*, 105450. https://doi.org/10.1016/j.compag.2020.105450
56. You, J., Liu, W., & Lee, J. (2020). A DNN-based semantic segmentation for detecting weed and crop. *Computers and Electronics in Agriculture, 178*, 105750. https://doi.org/10.1016/j.compag.2020.105750
57. Espejo-Garcia, B., Mylonas, N., Athanasakos, L., Fountas, S., & Vasilakoglou, I. (2020). Towards weeds identification assistance through transfer learning. *Computers and Electronics in Agriculture, 171*, 105306. https://doi.org/10.1016/j.compag.2020.105306

58. *Graph weeds net: A graph-based deep learning method for weed recognition.* (2020). *Computers and Electronics in Agriculture, 174*, 105520. https://doi.org/10.1016/j.compag.2020.105520
59. *Application of convolutional neural networks in weed detection and identification: A systematic review.* (2024). *Agriculture, 14*(4), 568. https://doi.org/10.3390/agriculture14040568
60. Murad, N. Y., Mahmood, T., Forkan, A. R. M., Morshed, M., Jayaraman, P. P., & Siddiqui, M. S. (2024). Weed detection using deep learning: A systematic literature review. Ninth International Conference on Image Processing Theory, Tools and Applications, 1-5.