

Article

Not peer-reviewed version

Object Detection in Agriculture: A Comprehensive Review of Methods, Applications, Challenges, and Future Directions

[Zohaib Khan](#), [Yue Shen](#)^{*}, [Hui Liu](#)

Posted Date: 8 May 2025

doi: 10.20944/preprints202505.0614.v1

Keywords: Object Detection; Deep Learning; Crop Monitoring; Autonomous Agricultural Robots; Agricultural Datasets; Smart Farming



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Article

Object Detection in Agriculture: A Comprehensive Review of Methods, Applications, Challenges, and Future Directions

Zohaib Khan, Yue Shen * and Hui Liu

School of Electrical and Information Engineering, Jiangsu University, Zhenjiang 212013, China; zai1318@foxmail.com (Z.K); amity@ujs.edu.cn (H.L.)

* Correspondence: shen@ujs.edu.cn

Abstract: Object detection has emerged as a transformative technology in precision agriculture, driving significant advancements in crop monitoring, weed management, pest detection, and autonomous field operations. This review provides a comprehensive synthesis of object detection methodologies, tracing their evolution from traditional hand-crafted feature-based approaches to modern deep learning architectures. Key agricultural applications are examined, emphasizing the role of publicly available datasets, including PlantVillage, DeepWeeds, and AgriNet, in catalyzing research progress. A comparative analysis of leading algorithms is presented, evaluating trade-offs among accuracy, inference speed, and computational efficiency within agricultural contexts. Persistent challenges are critically analyzed, including environmental variability, limited labeled data, difficulties in model generalization, real-time processing constraints, and the need for improved interpretability. Emerging research directions are also examined as potential strategies for enhancing object detection in complex agricultural environments. By bridging technical innovation with practical deployment, future object detection systems are positioned to revolutionize agricultural productivity, sustainability, and resilience on a global scale.

Keywords: object detection; deep learning; crop monitoring; autonomous agricultural robots; agricultural datasets; smart farming

1. Introduction

Over the past two decades, the field of artificial intelligence (AI) has undergone a profound paradigm shift, catalyzed by transformative advancements in machine learning and computer vision that have redefined the capabilities of automated systems [1]. Object detection, encompassing the simultaneous localization and classification of objects in images, has become an essential component of machine vision systems in agriculture, as demonstrated by its successful application in classifying apple color and deformity using CNN-based methods [2,3]. This evolution has been propelled by the transition from manually engineered feature extraction methods to data-driven approaches, culminating in the widespread adoption of deep learning techniques that exploit extensive computational resources and large-scale annotated datasets [4]. At present, object detection underpins a broad spectrum of applications, including autonomous navigation in self-driving vehicles, anomaly detection in medical imaging, quality assurance in industrial manufacturing, and precision agriculture, where its transformative potential is increasingly evident [5]. The initial reliance on conventional signal processing and static feature engineering has been progressively addressed with the emergence of deep learning architectures, particularly convolutional neural networks (CNNs), enhancing accuracy and robustness in complex, real-world perception tasks such as autonomous driving [6].

The development of object detection has progressed through distinct stages, reflecting the overall evolution of artificial intelligence in terms of innovation and adaptation. In its early stages, object detection relied on traditional computer vision methods that emphasized manual feature extraction

and object boundary definition through heuristic and statistical techniques which required extensive human intervention to define object boundaries and features [7]. Hyperspectral imaging techniques have shown strong potential for non-destructive chemical analysis in agriculture, as demonstrated by the quantitative detection of mixed pesticide residues on lettuce leaves [8,9]. Near-infrared transmission spectroscopy has been effectively applied for the non-destructive identification of pesticide residues in leafy vegetables, as demonstrated by the detection of contaminants in lettuce leaves [10]. However, these methods exhibited substantial limitations when confronted with complex scenes, occlusions, or variability in object appearance, primarily due to their dependence on hand-crafted feature descriptors. The transition toward machine learning methodologies in the late 2000s introduced trainable classifiers coupled with hand-crafted descriptors; however, performance remained constrained by a limited ability to generalize across diverse contexts [11]. The advent of deep learning, ignited by the success of AlexNet in 2012, marked a pivotal shift, as convolutional neural networks (CNNs) enabled the autonomous learning of hierarchical feature representations directly from raw pixel data, thereby eliminating the need for manual feature engineering [12]. Subsequent innovations, including Region-based CNNs (R-CNN), Faster R-CNN, SSD, and YOLO, further refined this paradigm. These models integrated region proposal, feature extraction, and classification into end-to-end trainable architectures, achieving remarkable improvements in both speed and accuracy [13,14]. These advancements have substantially enhanced the technical capabilities of object detection and expanded its practical applicability across a wide range of domains, thereby laying the foundation for its widespread adoption in real-world applications [15]. A summary of key milestones in the evolution of object detection algorithms from 1999 to 2025 is illustrated in Figure 1.

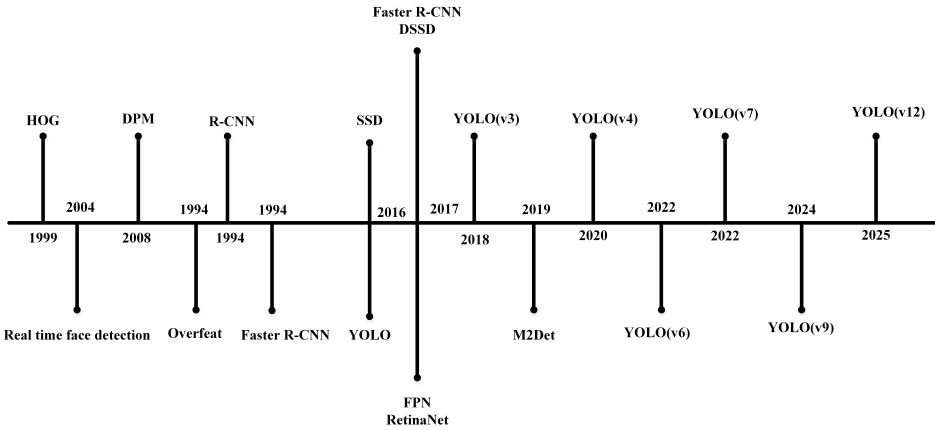


Figure 1. Timeline of Advances in Object Detection Algorithms (1999–2025).

In agriculture, the impact of AI-driven object detection systems is particularly pronounced, as they address longstanding inefficiencies inherent in traditional monitoring and management practices. Traditional agricultural tasks relied on labor-intensive manual inspections, resulting in processes that were time-consuming, prone to human error, and poorly suited to the scale and variability of modern farming operations [16]. Visual assessment of crop health across extensive fields or distinguishing weeds from crops under inconsistent lighting conditions often resulted in delayed interventions and suboptimal resource utilization [17]. In contrast, object detection algorithms powered by deep learning process imagery from drones, satellites, and ground-based sensors to deliver consistent, scalable, and real-time insights, thereby enabling precision agriculture at an unprecedented level [18]. Numerous studies have validated the effectiveness of such systems in agricultural applications, including the high-accuracy identification of plant diseases using convolutional neural networks (CNNs), the spatial mapping of weed distributions for precision herbicide deployment, and yield estimation through automated fruit counting [19,20]. Advanced techniques have further tailored object detection to the unique challenges of rural environments by enabling model adaptation to agricultural datasets and improving detection of small or occluded objects [21]. Beyond agricultural domains, the

same foundational principles enhance capabilities in other fields; in robotics [22,23], object detection facilitates grasping and navigation, in surveillance, it enables threat identification, and in healthcare, it supports diagnostic imaging by pinpointing abnormalities [24]. This versatility underscores the role of object detection as a critical discipline within AI, bridging theoretical innovation with tangible societal benefits [25]. Examples of agricultural robots employing advanced detection systems, perception modules, and actuation units are shown in Figure 2.

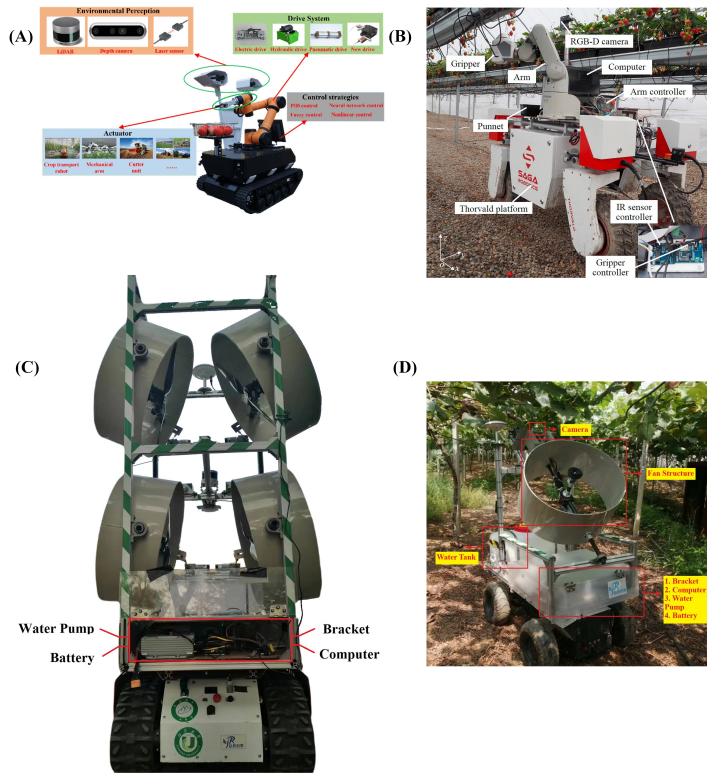


Figure 2. Agricultural robots integrated with detection systems, perception modules, and actuation units. (A) Forest mapping robot [26]; (B) Strawberry harvesting robot with multiple sensors [27]; (C) Autonomous orchard spraying robot with flexible mechanism [28]; (D) Variable spray robot for precision agriculture [29].

The significance of object detection extends beyond its immediate applications, reflecting broader trends and challenges in AI research. The transition from traditional methods to deep learning has not only enhanced performance but also introduced new complexities, including the need for large, annotated datasets and substantial computational resources, which pose barriers to deployment in resource-constrained settings [30]. In agricultural applications, domain-specific factors, including variable lighting conditions, dense foliage occlusions, and limited labeled data for rare crops or pests, further exacerbate these challenges, necessitating innovative solutions based on synthetic data generation and domain adaptation techniques [31]. Achieving real-time performance on edge devices continues to drive research into lightweight and computationally efficient architectures, particularly in tasks like field navigation where improved YOLOv8 structures have been applied successfully [32]. Across multiple domains, object detection remains intertwined with fundamental AI research questions, including the ability of models to generalize across diverse datasets, the trade-offs between accuracy and computational efficiency, and strategies for ensuring robustness in noisy, unstructured environments.

To address object detection in agriculture, this review conducts a comprehensive and critical synthesis focused on practical applications and contributions to artificial intelligence research.

(1) Object detection methodologies are reviewed, tracing their progression from classical feature-based techniques to deep learning frameworks, with emphasis on agricultural applications including crop monitoring, and weed classification. Technical foundations, including early feature extraction

methods and CNN-based models, are summarized, highlighting adaptations for agricultural environments.

(2) Object detection methods are systematically compared across agricultural tasks. Evaluation considers metrics including mean Average Precision (mAP), inference speed, and robustness to environmental noise, with additional comparisons to domains in robotic navigation and medical diagnostics to analyze trade-offs between accuracy and computational efficiency.

(3) Key deployment challenges are analyzed, covering both general AI issues and agricultural-specific complexities. Agricultural challenges include unstructured scenes, seasonal variability, and integration of multi-modal data sources (RGB, thermal, hyperspectral imagery). Solutions focus on improvements in data preprocessing, model design, and validation strategies.

(4) Future research directions are proposed, focusing on the development of lightweight, energy-efficient models for edge deployment, fusion of multi-modal sensor data to enhance detection robustness, and integration of explainable AI for system transparency [33]. Recent reviews of non-destructive vision-based assessment methods in agriculture highlight the growing role of computer vision in fruit quality evaluation, although challenges remain in developing models that generalize across environments [34]. To better illustrate the progression of major object detection frameworks, Table 1 summarizes key models, their release years, types, and distinguishing features. This comparison highlights the evolution from two-stage architectures to faster and more efficient one-stage methods.

Table 1. Overview of Major Object Detection Frameworks

Model	Year	Type	Key Features
R-CNN	2014	Two-stage	Region proposals + CNN classification [35]
Fast R-CNN	2015	Two-stage	ROI pooling, faster training [36]
Faster R-CNN	2015	Two-stage	Integrated RPN for proposal generation [37]
YOLOv1	2016	One-stage	Unified detection and classification [38]
YOLOv3	2018	One-stage	Multi-scale prediction, Darknet-53 [39]
YOLOv7	2022	One-stage	E-ELAN optimization, fast and accurate [40]
SSD	2016	One-stage	Multi-box detection with multiple feature maps [41]

2. Object Detection Fundamentals

Object detection constitutes a cornerstone of computer vision, addressing the dual challenge of identifying and localizing objects within an image or video frame by generating bounding boxes around detected entities and assigning corresponding class labels [42]. Unlike image classification, which labels an entire image, or semantic segmentation, which assigns per-pixel classes without bounding discrete objects, object detection provides localized bounding box predictions [43]. The capability to spatially localize objects has rendered object detection indispensable across a wide range of applications, including autonomous systems and precision agriculture, where the identification of crops, weeds, or pests within complex scenes drives actionable insights [1,5]. The historical evolution of object detection methodologies is examined, beginning with traditional approaches based on hand-crafted features and progressing toward the transformative influence of deep learning, with particular emphasis on their technical foundations and applicability to agricultural domains [4,11].

2.1. Traditional Approaches in Agriculture

Traditional object detection methods in agriculture predominantly relied on classical computer vision and machine learning techniques that utilized hand-crafted features, including Scale-Invariant Feature Transform (SIFT), Histogram of Oriented Gradients (HOG), and color histograms. These approaches typically followed a multi-stage pipeline consisting of feature extraction, classification using algorithms like Support Vector Machines (SVMs), and localization through sliding window techniques or selective search [44]. Although computationally intensive, these methods established the foundation of early agricultural automation by enabling the recognition of visual patterns in plant

leaves, fruits, and weeds under controlled conditions [45]. Support Vector Machines (SVMs) have been applied to shape- and color-based features for the detection of tea leaf diseases [46,47], while other approaches have utilized Gabor filters and texture analysis for the classification of weed species in crop fields.

Despite inherent limitations, traditional approaches demonstrated effectiveness in scenarios with minimal variability in visual inputs, including indoor farming, greenhouse environments, and early-stage laboratory datasets. Their relatively low data requirements and explainable pipelines rendered them suitable for classification and detection tasks under consistent lighting conditions [45]. Moreover, these techniques contributed to the early development of precision agriculture systems by enabling targeted spraying and automated yield monitoring. Although modern deep learning methods have largely superseded classical pipelines, the interpretability and lower computational cost associated with traditional techniques continue to offer advantages for low-power and edge computing applications in resource-limited agricultural settings [24]. The overall workflow of traditional object detection methods using SIFT/HOG features and classical classifiers is illustrated in Figure 3.

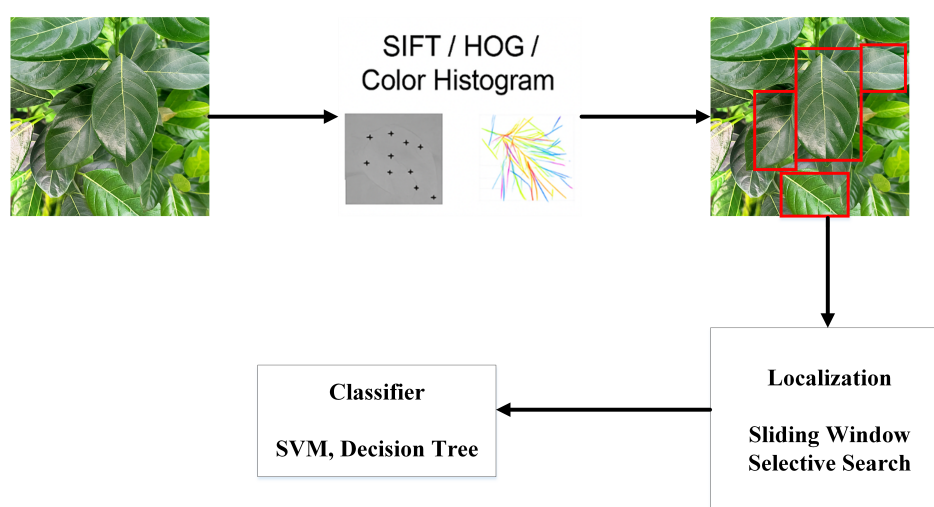


Figure 3. Object Detection Workflow Using SIFT/HOG Features and Classical Classifiers

2.2. Deep Learning-Based Methods in Agriculture

In recent years, deep learning has emerged as a transformative force in agricultural object detection, offering substantial improvements over traditional computer vision techniques in terms of accuracy, robustness, and scalability. While Convolutional Neural Networks (CNNs) have become central to agricultural AI for image-based tasks, other sensing modalities such as Near-Infrared Spectroscopy (NIRS) combined with machine learning also offer powerful alternatives for tasks like seed age discrimination [48]. This data-driven paradigm has enabled the deployment of end-to-end models capable of identifying complex agricultural objects, including fruits, leaves, pests, and weeds under varying environmental conditions [49]. YOLO-based models have been adapted for tasks including grapevine disease detection [29], fruit counting [19], and tea leaf classification [50,51], often achieving real-time inference speeds suitable for deployment on UAVs and mobile robotic platforms. Transfer learning from large-scale datasets and the use of data augmentation techniques have further improved the generalizability of deep models in agricultural domains with limited labeled data [52]. Moreover, recent innovations in lightweight architectures, including MobileNet and EfficientNet, have facilitated the integration of deep learning models into edge computing devices for on-site agricultural decision-making [53]. Collectively, these advances represent a major progression in precision farming, enabling more efficient resource utilization, early disease detection, and enhanced crop monitoring capabilities.

2.2.1. R-CNN and Fast R-CNN

Region-based Convolutional Neural Networks (R-CNN) introduced the concept of region proposals followed by CNN-based feature extraction and classification, marking a pivotal shift in object detection paradigms [35]. Although R-CNN achieved high accuracy, its multi-stage pipeline resulted in significant computational overhead. Fast R-CNN addressed these limitations by integrating feature extraction and classification within a single network using ROI pooling, thereby reducing both inference time and memory usage. In agricultural applications, both modern region proposal-based frameworks and traditional feature-based methods have demonstrated reliable performance under variable lighting conditions for tasks such as plant disease identification and fruit counting [54,55].

2.2.2. Faster R-CNN

Faster R-CNN further advanced the R-CNN family by embedding a Region Proposal Network (RPN) directly into the backbone CNN, enabling end-to-end training and significantly improving inference speed without compromising accuracy [37]. Its applications in agriculture include greenhouse detection and fruit recognition in complex orchard environments, where high detection accuracy is often prioritized over latency [56]. The ability to generate high-quality region proposals renders Faster R-CNN particularly effective for detecting dense or overlapping agricultural targets.

2.2.3. YOLO (You Only Look Once)

The YOLO series of models reformulated object detection as a single regression problem, thereby enabling real-time processing capabilities [38]. Due to its efficiency, YOLO has become particularly suitable for embedded agricultural applications, including drone-based weed monitoring and robotic fruit picking in real-time environments [57]. Successive versions, from YOLOv3 to YOLOv8, have introduced architectural enhancements that boost detection accuracy and robustness against occlusions and scale variations.

2.2.4. SSD (Single Shot MultiBox Detector)

The Single Shot MultiBox Detector (SSD) architecture achieves a balance between speed and accuracy by simultaneously predicting object classes and bounding boxes in a single forward pass through the network [41]. In contrast to YOLO, SSD utilizes multiple feature maps at different resolutions, making it particularly adept at detecting objects of varying sizes, an essential advantage in agricultural environments where pests or produce may appear at different scales. SSD-based models have been explored for agricultural applications that demand lightweight inference while maintaining acceptable detection accuracy.

The transition to deep learning methodologies has enhanced object detection capabilities and expanded their application scope. Recent architectures have refined performance by addressing challenges related to class imbalance and contextual reasoning [58]. Advancements in object detection technologies enable precision tasks including real-time identification of diseased leaves and mapping of crop distributions across extensive fields, with underlying principles that also extend to domains including robotics, surveillance, and medical imaging [59]. The structural differences and evolutionary progression of major object detection architectures, from R-CNN to SSD, are summarized in Figure 4.

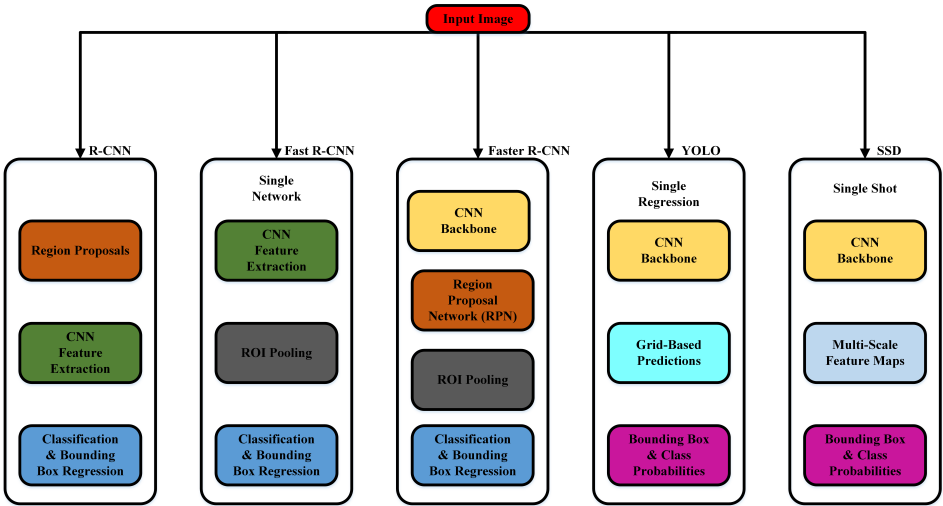


Figure 4. Evolution of Object Detection Architectures.

3. Applications in Agriculture

The transformative potential of object detection is vividly illustrated in agriculture, where it addresses a diverse spectrum of tasks critical to modern farming practices [60]. By leveraging advanced computer vision and deep learning techniques, these applications enhance efficiency, precision, and scalability, replacing labor-intensive manual methods with automated systems capable of operating under complex, real-world conditions [61]. A structured analysis of key agricultural use cases, encompassing weed detection, fruit counting and ripeness assessment, disease and pest identification, livestock and wildlife monitoring, as well as crop row and canopy detection, is presented, emphasizing the computational methods employed, datasets referenced, and domain-specific challenges addressed [20]. While grounded in agriculture, these applications reflect broader AI principles, offering insights into object detection’s adaptability and limitations across diverse domains [62]. Table 2 summarizes representative agricultural applications of object detection algorithms across various tasks in precision farming.

Table 2. Applications of Object Detection in Agriculture with Algorithmic Contributions

Task	Application Example	Reference	Algorithmic Contribution
Disease Detection and feature enhancement for orchard environments	YOLOv7 for grapevine powdery mildew detection	Sun et al. (2025)	Improved YOLOv7 with backbone pruning
Disease Detection for detecting small objects in plant parts	RetinaNet for multi-crop disease classification	Duan et al. (2024)	YOLOv8-GDCI with global detail-context interaction
Fruit Counting techniques including object-level fruit analysis	YOLOv5 applied to apple counting	Ma et al. (2024)	Reviewed deep learning maturity detection
Fruit Counting capability using multispectral image fusion	SSD for citrus fruit detection in orchards	Sa et al. (2016)	Developed SSD-based detection with real-time
Weed Detection YOLOv3 under real-world conditions	DeepWeeds dataset classification using YOLOv3	Olsen et al. (2019)	Introduced multiclass weed dataset; evaluated
Weed Detection feature fusion for dense weed clusters	Improved YOLOv8 for weed detection in crop field	Jia et al. (2024)	Enhanced YOLOv8 with attention-guided dual-layer
Spraying Robotics targeting grape leaf health conditions	Precision pesticide application in vineyards	Khan et al. (2025)	YOLOv7 improved with custom feature extractors
Spraying Robotics using refined YOLOv8 architecture	Precision pesticide application in orchards	Khan et al. (2024)	Real-time instance segmentation of canopies

3.1. Weed Detection

Effective weed management is a cornerstone of crop productivity, requiring precise discrimination between crop and weed species to enable selective herbicide application or autonomous weeding [63]. Object detection has significantly advanced this task by enabling real-time identification of invasive weed species amidst dense vegetation [57]. Datasets such as DeepWeeds, comprising thousands of annotated images from Australian rangelands, have supported the training of deep learning models including YOLOv5 and SSD, which excel at detecting weeds under varying field conditions [64]. YOLOv5, with its lightweight architecture and multi-scale prediction capabilities, achieves rapid inference speeds, while SSD's utilization of multi-layer feature maps enhances accuracy in identifying small or partially occluded weeds [57]. Studies have reported mean Average Precision (mAP) scores exceeding 0.85 for weed classification in controlled environments, although performance declines in cluttered or shadowed conditions due to visual similarities between crops and weeds [65]. Techniques such as transfer learning, which adapts pre-trained weights from datasets like ImageNet to weed-specific datasets, have been employed to address data scarcity challenges in agricultural vision applications [66]. These advancements underscore the critical role of object detection in precision agriculture, with parallels observed in fine-grained classification tasks across other fields, such as species identification in ecological studies [67]. Figure 5 showcases representative samples from the DeepWeeds dataset, illustrating the visual diversity and complexity encountered in field-based weed detection tasks.

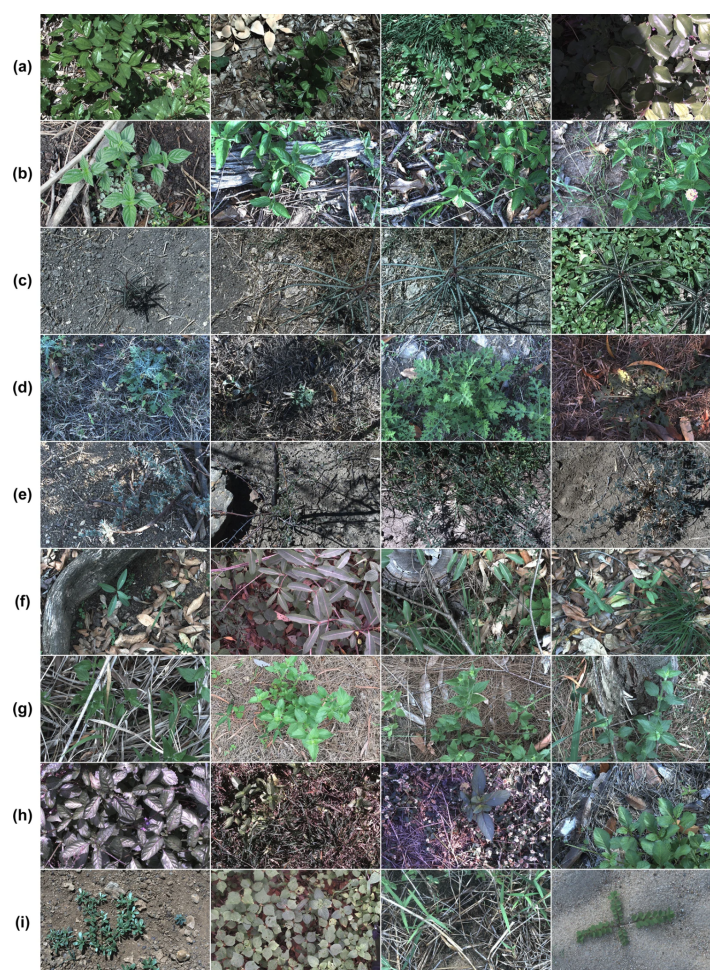


Figure 5. Representative Samples from the Weed Species Dataset [64].

3.2. Fruit Counting and Ripeness Detection

Fruit detection serves dual purposes in agriculture: estimating yield for harvest planning and assessing ripeness to optimize picking schedules [13]. Object detectors trained on annotated datasets,

such as MinneApple for apples, GrapeCS for grapes, and TomatoID for tomatoes, have demonstrated robust performance in localizing and classifying fruits under diverse conditions [68]. Faster R-CNN models have demonstrated strong performance in detecting fruits under occluded conditions, achieving high precision through the extraction of contextual features from deep convolutional layers [69]. EfficientDet, a more recent architecture, balances speed and accuracy through compound scaling, making it suitable for real-time ripeness assessment on mobile platforms [70]. Research has reported detection accuracies above 90% in well-lit orchard environments; however, performance degrades under low-light conditions or heavy occlusion, necessitating preprocessing techniques such as contrast enhancement or the integration of multi-modal data, including thermal imaging [71]. Ripeness detection often incorporates color-based features or temporal tracking, reflecting object detection's adaptability to task-specific cues, a principle similarly employed in industrial quality control applications [72]. Figure 6 presents detection results of grape clusters across multiple varieties, demonstrating the capability of object detection models to localize fruits under varying occlusion and illumination conditions.



Figure 6. Detection Results of Grape Clusters Across Different Varieties [68].

3.3. Disease and Pest Detection

Early detection of plant diseases and pest infestations is critical for crop protection, enabling timely interventions to minimize yield losses [73]. Object detection systems trained on datasets such as PlantVillage, which includes over 50,000 images of diseased and healthy leaves across multiple crops, have proven effective in identifying subtle symptoms of infection and the presence of insect pests [74]. YOLOv3 and RetinaNet leverage high-resolution feature maps and focal loss functions to enhance the detection of small objects, including disease spots and aphids, and to address class imbalance [75]. YOLOv7 has been deployed for the detection of powdery mildew on grapevines, achieving mAP scores exceeding 0.80, while RetinaNet's emphasis on hard examples improves recall in sparse pest distributions [76]. Key challenges include differentiating disease symptoms from natural leaf variations and managing limited training data for rare conditions, issues often addressed through data augmentation or synthetic image generation [77]. These developments parallel diagnostic imaging

applications in healthcare, where object detection is similarly employed to identify anomalies with high precision [78]. Figure 7 illustrates the experimental setup involving grape leaves marked for validating disease detection accuracy and evaluating spray coverage effectiveness.

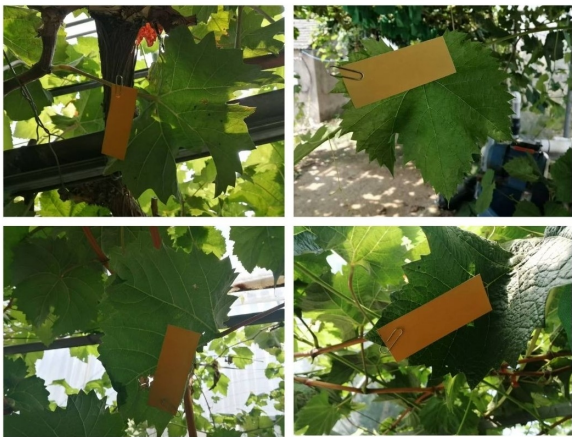


Figure 7. Setup of Grape Leaves with Markers for Experimental Validation of Detection Accuracy and Spray Coverage.

3.4. Crop Row and Canopy Detection

Autonomous agricultural vehicles, including tractors and harvesters, rely on object detection systems to navigate fields by following crop rows and mapping plant canopies [79,80]. This task integrates geometric reasoning, detecting linear row patterns, with semantic understanding of canopy boundaries, often utilizing RGB or multispectral imagery captured by onboard cameras [81]. Faster R-CNN and SSD models, supplemented by post-processing techniques like Hough transforms for line detection, facilitate precise row alignment, thereby reducing crop damage during mechanical operations [82]. The Sugar Beet Field dataset has facilitated the training of canopy detection models, with reported accuracies exceeding 95% in structured field environments, although performance declines in uneven terrains or irregular planting conditions [83]. The real-time requirements of autonomous navigation have driven the adoption of lightweight architectures such as MobileNet, reflecting a broader trend toward edge-based AI in robotics and autonomous systems [14]. These techniques extend beyond agriculture to domains such as autonomous driving, underscoring shared challenges in spatial reasoning and environmental perception. Figure 8 depicts the key components of an agricultural robot path-tracking system, integrating power, navigation, chassis, and remote control modules for autonomous field operations.

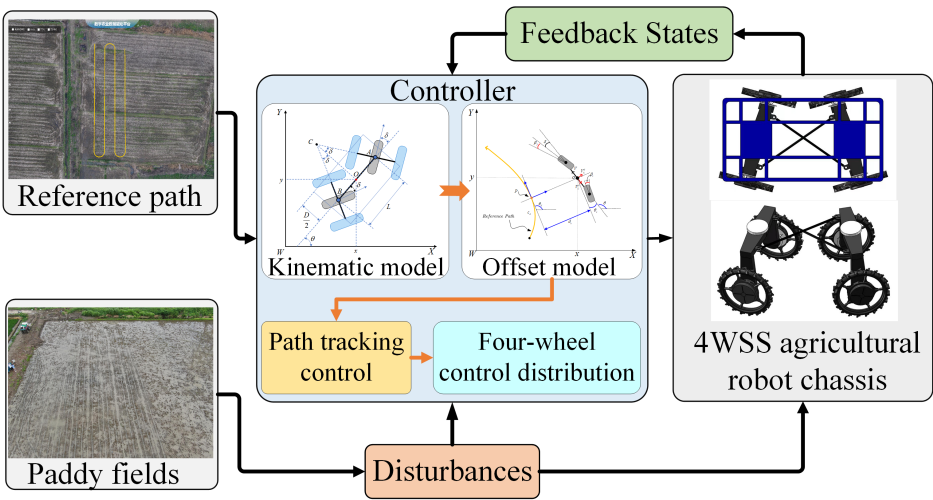


Figure 8. Agricultural robot path-tracking system [68].

4. Dataset Overview

The rapid advancement of object detection in agriculture has been strongly facilitated by the availability of publicly accessible datasets, which serve as the backbone for training, validating, and benchmarking deep learning models [84]. These datasets provide annotated visual data, including images or videos with labeled bounding boxes, class labels, or segmentation masks, enabling algorithms to learn complex agricultural patterns under diverse conditions [85]. While object detection as a broader AI discipline benefits from general-purpose datasets such as COCO and ImageNet, agricultural applications require specialized datasets that capture the unique variability of crops, weeds, pests, and environmental factors [84].

4.1. Key Public Datasets

Several datasets have emerged as foundational resources for agricultural object detection, each tailored to specific tasks and reflecting the diversity of farming environments. Table 3 provides an overview of major agricultural datasets commonly used for object detection, detailing their image counts, crop or weed types, and notable characteristics.

Table 3. Overview of Major Agricultural Datasets for Object Detection

Dataset	Images	Crop/Weed Types	Notes
PlantVillage	50,000+	38 crop-disease pairs	Controlled lab images [86]
DeepWeeds	17,509	9 weed species	Field conditions, weeds in Australia [64]
GrapeLeaf Dataset	5,000+	Grapevine diseases	Grape disease segmentation [76]
DeepFruit	35,000+	Apple, mango, citrus	Fruit detection for yield estimation [87]

4.1.1. PlantVillage

Designed for leaf disease detection, PlantVillage comprises over 54,000 images of leaves from 14 crop species, annotated with 38 disease classes and healthy states [88]. Collected under controlled greenhouse conditions, it provides high-quality RGB imagery, establishing itself as a benchmark for training models such as YOLO and RetinaNet to identify subtle disease symptoms [89]. Its extensive class coverage supports fine-grained classification, although its lack of field-based variability limits real-world generalization [85].

4.1.2. DeepWeeds

Focused on weed identification in pastures, DeepWeeds offers 17,509 images of eight weed species along with a negative (no-weed) class, captured in Australian rangelands using drones and ground cameras [64]. Annotated with bounding boxes, it has supported the development of real-time detectors such as YOLOv5 and SSD, achieving mAP scores above 0.908 for weed-crop differentiation [90]. Its emphasis on natural outdoor settings enhances its utility for precision weed management, although its regional specificity restricts broader generalization [91].

4.1.3. AppleAphid

The AppleAphid dataset targets pest detection, providing annotated images of apple leaves infested with aphids [92]. Featuring hundreds of high-resolution images with bounding box labels, it supports small-object detection tasks using models such as Faster R-CNN, which leverage multi-scale features to accurately identify tiny insects [93]. Its narrow focus on a single pest-crop pair highlights the need for more diverse pest detection datasets [94].

4.1.4. AgriNet

AgriNet is a large-scale, multi-category dataset encompassing over 100,000 images across crops, weeds, pests, and diseases, collected from various global agricultural regions [95]. Annotated with bounding boxes and class labels, it facilitates comprehensive training of models such as EfficientDet

for tasks ranging from yield estimation to pest monitoring [96]. Its scale and diversity position it as a versatile resource, although inconsistencies in annotation quality across regions present challenges [97].

4.1.5. Mini-PlantNet

Mini-PlantNet, a subset of the larger PlantNet database, focuses on plant species classification, containing thousands of images with bounding box annotations for plant identification [98]. Initially developed for species recognition, it has been adapted for object detection tasks including canopy mapping, supporting the training of lightweight models like MobileNet for edge deployment [99]. While its emphasis on botanical diversity is a strength, its limited coverage of agricultural tasks restricts its broader utility [100].

4.2. Dataset Characteristics and Contributions

These datasets exhibit substantial variation in size, annotation type, and acquisition methods, reflecting the multifaceted nature of agricultural object detection [101]. PlantVillage and AppleAphid, with their controlled conditions, provide high-quality annotations suitable for initial model training, whereas DeepWeeds and AgriNet capture real-world complexity, supporting robust testing under natural environmental conditions [20,21]. Annotation types range from bounding boxes (DeepWeeds, AgriNet) to class labels (PlantVillage), and some datasets incorporate metadata such as growth stage or environmental context, thereby enhancing model interpretability [85]. The contributions of these datasets extend beyond agriculture, as PlantVillage's fine-grained labels are analogous to datasets used in medical imaging, and DeepWeeds' outdoor imagery aligns with ecological monitoring applications [86]. By providing standardized benchmarks, these datasets have driven algorithmic advancements, including the adoption of transfer learning techniques to adapt general-purpose models pre-trained on datasets like COCO to specialized agricultural tasks, thereby mitigating the reliance on extensive labeled data [11].

4.3. Challenges in Dataset Diversity and Quality

Despite their significant contributions, agricultural datasets face notable challenges that hinder model performance and generalization, reflecting broader issues in AI data curation [21]. Seasonal variation, such as shifts in crop appearance across growth cycles or weather conditions, is poorly represented in datasets like PlantVillage, limiting models' adaptability to temporal changes [100]. Occlusions, frequently encountered in dense agricultural fields where leaves obscure fruits or pests, complicate bounding box accuracy, as observed in small-object detection tasks with datasets like AppleAphid [102]. Inconsistent annotation standards across datasets, including variations in bounding box tightness and class definitions in AgriNet, introduce noise that undermines cross-dataset compatibility [103]. Moreover, the scarcity of labeled data in underrepresented regions, particularly tropical agriculture, restricts the global applicability of trained models, a problem exacerbated by the high cost and expertise required for manual annotation [104].

4.4. Implications for Object Detection Research

The limitations of current agricultural datasets underscore the need for innovative data strategies, which represent a critical frontier in AI research [103]. Techniques such as domain adaptation, which fine-tunes models on small, region-specific datasets, and multi-modal integration, which combines different sensor modalities, are emerging to address diversity gaps [105]. Furthermore, the push for open-source, standardized datasets aligns with broader efforts to democratize AI development, enabling researchers and practitioners to address real-world deployment challenges more effectively [106].

5. Comparison of Algorithms

Object detection algorithms constitute the foundation of modern computer vision systems, distinguished by trade-offs in accuracy, inference speed, and computational demands, attributes that

critically influence their practical utility across diverse applications [107]. In agriculture, these trade-offs are particularly significant due to the need for deployment on edge devices, such as drones, robots, and handheld tools, where processing power is limited and real-time performance is often essential [108]. The comparative evaluation highlights universal design principles that extend beyond agricultural contexts, with applicability to domains including robotics, surveillance, and autonomous systems. Table 4 compares representative object detection models applied to agricultural tasks, summarizing their dataset usage, performance characteristics, and practical deployment considerations.

5.1. Algorithmic Foundations and Performance

Each model represents a distinct approach to object detection, balancing competing demands for precision and efficiency through innovative architectural designs. Table 4 compares representative object detection models applied to agricultural tasks, summarizing their dataset usage, performance characteristics, and practical deployment considerations.

Table 4. Comparison of Object Detection Algorithms for Agricultural Tasks.

Model	Dataset	Performance	Notes
YOLOv7	Grape disease detection	High accuracy, fast inference	Suitable for real-time deployment [29,76]
YOLOv3	Weed detection (DeepWeeds)	Good balance of speed and accuracy	Field condition tested [64,109–111]
Faster R-CNN	PlantVillage	High detection accuracy	Slower but more robust [37,112,113]
RetinaNet	Multi-crop disease datasets	Handles class imbalance well	Useful for rare diseases [114,115]

5.1.1. Faster R-CNN

Introduced as a two-stage detector, Faster R-CNN integrates a Region Proposal Network (RPN) with a CNN backbone (ResNet-50 or ResNet-101) to generate and classify region proposals, followed by bounding box regression and classification [37]. Its strength lies in achieving high accuracy, often exceeding mean Average Precision (mAP) scores of 0.90 on datasets such as COCO, due to its capacity to leverage deep feature hierarchies and contextual reasoning [116]. However, its inference speed remains relatively slow, typically 5–10 frames per second (FPS) on high-end GPUs, owing to the computational overhead of processing multiple stages. In agriculture, Faster R-CNN excels in offline analysis tasks requiring meticulous precision, including the detection of early disease symptoms on leaves and the identification of pests in high-resolution imagery [117]. Nevertheless, its reliance on substantial computational resources limits its applicability in real-time edge scenarios, relegating it to cloud-based or high-performance GPU deployments.

5.1.2. YOLO

The You Only Look Once (YOLO) family represents a single-stage detection paradigm, reframing object detection as a regression problem by predicting bounding boxes and class probabilities in a single forward pass across a grid of image cells [118]. Optimized for speed and efficiency, YOLOv5 achieves frame rates exceeding 50 FPS on mid-tier GPUs (NVIDIA GTX 1660) while maintaining mAP scores between 0.85 and 0.90, depending on the variant (YOLOv5s for lightweight applications, YOLOv5x for higher accuracy) [119]. Its architecture incorporates enhancements such as anchor box optimization, multi-scale predictions via Feature Pyramid Networks (FPN), and a lightweight backbone (CSPDarknet53), making it highly suitable for real-time agricultural applications [120]. In field settings, YOLOv5 supports tasks like weed detection via drones, fruit counting in orchards, and pest monitoring through mobile devices, where rapid decision-making is essential [90]. In dense and complex field environments, such as those encountered during broccoli head detection, accuracy can be challenged by occlusions and object clustering, necessitating design trade-offs in model structure and resolution settings [121].

5.1.3. SSD

The Single Shot MultiBox Detector (SSD) adopts a single-stage approach, utilizing a VGG-16 backbone to extract features and predict object classes and bounding boxes across multiple scales

from different convolutional layers [41]. This design yields a balanced trade-off, with inference speeds ranging from 20 to 40 FPS and mAP scores typically between 0.75 and 0.85, depending on input resolution and hardware capabilities [107,122]. SSD's capacity to detect objects at varying scales renders it well-suited for embedded systems, such as mobile robots engaged in weed or fruit detection in agricultural fields [123]. In practice, it provides sufficient accuracy for tasks where moderate precision aligns with operational needs, while its lower computational footprint compared to two-stage models supports deployment on resource-constrained devices [124]. Limitations include reduced performance in detecting small objects or handling highly cluttered scenes, which are common challenges in dense crop environments.

5.1.4. EfficientDet

EfficientDet represents a scalable, state-of-the-art model, employing compound scaling to simultaneously adjust network depth, width, and resolution for optimizing both accuracy and efficiency [70]. Built on an EfficientNet backbone and enhanced with a Bidirectional Feature Pyramid Network (BiFPN), it achieves mAP scores exceeding 0.90 while maintaining inference speeds of 30–50 FPS on edge TPUs or high-end GPUs [125,126]. Its scalability across D0 to D7 variants enables customization for a wide range of hardware, from lightweight edge devices to powerful servers, supporting diverse agricultural tasks such as crop row mapping, livestock monitoring, and multi-object field analysis [127]. In agricultural applications, EfficientDet's robust feature fusion excels under variable conditions, including partial occlusions and low-light environments, although its complexity necessitates careful optimization for real-time edge deployment [128].

5.2. Comparative Analysis in Agricultural Contexts

The suitability of these algorithms for agricultural tasks hinges on their performance metrics and deployment constraints [129]. Faster R-CNN's high accuracy (mAP approximately 0.92) makes it suitable for offline tasks involving disease detection from high-resolution drone imagery, although its low speed (5–10 FPS) precludes real-time use [107]. YOLOv5's balanced profile (mAP approximately 0.88, 50+ FPS) dominates real-time applications, including weed detection on drones and pest tracking on edge AI devices [130]. SSD's moderate accuracy (mAP approximately 0.80) and speed (20–40 FPS) make it viable for embedded systems, particularly for fruit and weed detection where hardware limitations are significant [131]. EfficientDet's high accuracy (mAP approximately 0.91) and efficiency (30–50 FPS) position it as a versatile general-purpose solution, excelling in tasks requiring both precision and real-time operation. Hardware compatibility further shapes algorithm selection, with edge devices (NVIDIA Jetson) favoring lightweight models, while cloud or server-based setups accommodate more computationally intensive models [132]. Beyond agriculture, these dynamics similarly inform model selection in fields such as robotics and surveillance, where accuracy-speed balances are critical [133].

5.3. Broader Implications and Trends

This comparative analysis highlights a continuum of design philosophies: two-stage models such as Faster R-CNN prioritize accuracy at the expense of speed, whereas single-stage models such as YOLOv5 and SSD emphasize efficiency [70]. EfficientDet bridges this gap through scalable architecture. In agriculture, the need for edge deployment amplifies the demand for lightweight, real-time solutions, driving innovations such as model pruning and quantization to reduce memory footprints without compromising performance [134]. Similar optimization strategies are observed across the broader field of artificial intelligence, particularly in resource-constrained environments found in IoT systems and mobile robotic platforms [135]. Furthermore, the integration of agricultural benchmarks into model evaluation, moving beyond general-purpose datasets like COCO, reflects a growing emphasis on domain-specific metrics in AI research [136].

6. Challenges and Open Problems

Despite the remarkable strides made in object detection, its application in agriculture remains hindered by a constellation of challenges spanning environmental, data-related, computational, and interpretative dimensions [137]. These obstacles reflect not only the unique complexities of agricultural environments but also broader open problems in AI, where robustness, scalability, and usability are perennial concerns [138]. Environmental variability, data scarcity, model generalization, real-time constraints, and explainability represent enduring challenges in agricultural object detection. Their implications, potential mitigation strategies, and broader relevance to artificial intelligence research are examined [139]. These issues underscore the necessity for continued interdisciplinary efforts to advance both precision agriculture and object detection frameworks [140]. Table 5 summarizes major challenges in agricultural object detection alongside key research contributions that aim to address these issues.

Table 5. Challenges and Future Directions in Agricultural Object Detection.

Challenge	Key Contributions
Tiny Object Detection (aphids, mildew spots)	Focal Loss to address class imbalance [114]
Domain Shift (lab to field conditions)	Domain adaptation techniques for agriculture [77]
Limited Labeled Data	Semi-supervised learning for crop disease detection [78]
Explaining Model Decisions (Explainability)	Visualization methods for deep model decisions [78]
Lighting and Background Variations	Robust early disease detection in varying environments [73]
Real-Time Deployment on Edge Devices	Lightweight CNN design for embedded detection [74]

6.1. Environmental Variability

Agricultural scenes are inherently dynamic, and environmental variability poses significant challenges to the robustness of object detection models [137]. Lighting conditions fluctuate dramatically, from harsh midday sunlight to dawn shadows and overcast skies, altering object appearance and reducing model accuracy [141]. YOLOv5’s performance in weed detection has been observed to decline significantly under low-light conditions, with mAP decreasing by up to 20% due to poor contrast [142]. Occlusions, such as overlapping leaves obscuring fruits or pests, complicate bounding box predictions and increase false negatives [143]. Background clutter, including soil textures, plant debris, or mixed vegetation, further confounds detectors by introducing visual noise, as seen in SSD’s difficulties isolating small pests amidst dense foliage [144]. Addressing these challenges requires robust preprocessing strategies, with methods such as histogram equalization for lighting correction and the integration of multi-modal inputs like thermal imaging to mitigate occlusions, although these approaches often result in increased computational overhead [141]. Beyond agriculture, environmental variability parallels challenges faced in outdoor robotics and autonomous driving [145].

6.2. Model Generalization

The ability of object detection models to generalize across diverse agricultural contexts, including different crop types, geographic regions, and seasonal variations, remains a major challenge, often resulting in overfitting to specific training conditions [146]. A Faster R-CNN model trained on apple orchards in temperate climates may fail to detect mangoes in tropical environments due to differences in fruit shape, color, or background. Seasonal variability in crop appearance across growth stages can lead to significant domain shifts, which negatively impact detection performance. This effect has been demonstrated in the case of SSD, where accuracy declines when models are applied across visually distinct phases of the same crop [147]. This lack of generalization stems from dataset biases and insufficient feature invariance, further exacerbated by the limited diversity of agricultural datasets [146]. Transfer learning, wherein models pre-trained on large datasets like ImageNet are fine-tuned on agricultural data, provides partial mitigation; however, significant domain gaps persist. This issue mirrors broader AI challenges related to cross-domain adaptation in fields such as robotics and

healthcare, where models must transcend training biases to perform reliably in varied real-world conditions [148].

6.3. Real-Time Constraints

Real-time performance is a critical requirement for agricultural edge devices, including drones, robots, and handheld tools, where limited processing power presents a formidable constraint [149]. Despite their high accuracy, models with inference times of 100–200 milliseconds per frame remain unsuitable for dynamic agricultural tasks that demand real-time processing, with operations like autonomous weeding requiring frame rates of at least 30 FPS [107]. Lightweight models demonstrate favorable frame rates on edge devices, typically reaching 20–50 FPS, but this often comes at the cost of reduced precision, especially in scenarios involving small object detection [150]. These trade-offs are intensified by the memory and power limitations inherent to edge hardware, prompting research into model compression techniques, including pruning, quantization, and knowledge distillation, to reduce model sizes (from 50 MB to 10 MB) while preserving inference performance [134]. Real-time constraints constitute a widespread challenge in artificial intelligence and are equally critical in domains demanding high edge efficiency, notably autonomous driving and the Internet of Things (IoT) [132].

7. Future Directions

Environmental variability, data scarcity, generalization, real-time constraints, and explainability, underscore the necessity for innovative approaches to enhance the utility of object detection in agriculture [151]. As precision agriculture evolves toward smarter and more sustainable systems, emerging AI paradigms offer promising avenues to bridge these gaps, thereby improving scalability and practical impact [152]. Future advancements are expected to center around explainable AI and the seamless integration of edge AI, multimodal sensing, and data-efficient learning into comprehensive detection frameworks [153]. The integration of these advancements has the potential to overcome existing limitations in object detection, thereby enabling the development of autonomous, adaptable, and reliable systems for agricultural and broader real-world applications [154].

7.1. Explainable AI (XAI)

Enhancing model interpretability is essential for building user trust and supporting decision-making, particularly for agricultural object detection [155]. Attention mechanisms such as Grad-CAM highlight image regions that influence model predictions, while feature attribution methods like SHAP quantify each pixel's contribution to the output [156]. Traditional neural networks, such as BP-based models, have also shown promise in early-stage agricultural disease identification tasks, such as mildew detection in aeroponically propagated mulberry cuttings [157]. Post-hoc explanation tools, including decision trees that approximate CNN behavior, offer simplified interpretative pathways for non-expert users, although real-time deployment of these explanations remains computationally demanding [158]. The rise of XAI reflects a broader movement within AI toward ethical and accountable systems, paralleling efforts in domains such as finance and medical diagnostics, where interpretability is a critical requirement [159].

Figure 9 illustrates this concept with an agricultural example. The original image (left) shows grape leaves in a vineyard setting. The middle and right panels display attention heatmaps from improved YOLOv10 and YOLOv10n models, respectively. These highlight the regions most influential in driving the model's pest or disease detection decisions. Such interpretability tools enable researchers to verify whether models attend to biologically meaningful areas, thereby supporting transparent and accountable agricultural AI systems.

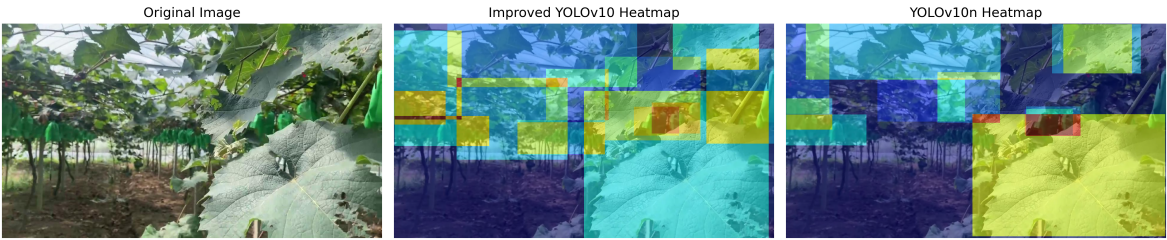


Figure 9. Visual comparison of attention heatmaps generated by improved YOLOv10 and YOLOv10n models on grape leaf imagery [160].

7.2. Few-Shot and Self-Supervised Learning

Addressing the challenge of limited annotated datasets is essential for advancing agricultural computer vision, particularly in scenarios involving rare crop varieties, pest infestations, or plant diseases for which labeled imagery is scarce. Few-shot learning (FSL) presents a viable solution by enabling models to generalize effectively from only a few labeled instances, typically between 5 to 10 samples per class, through techniques such as meta-learning or metric-based methods. Notably, prototypical networks have demonstrated efficacy by comparing query samples to learned class prototypes, facilitating classification with minimal supervision [1,29]. This concept is illustrated in Figure 10, where embedding networks process both support and query sets to compute similarity scores. Within agricultural domains, this capability is instrumental in detecting novel pests across heterogeneous field conditions using only limited training data, thereby significantly reducing the manual labeling burden [60].

In parallel, self-supervised learning (SSL) offers an alternative paradigm by harnessing the abundance of unlabeled agricultural imagery, frequently captured via drones or stationary cameras, to learn useful visual representations. Pretext tasks, such as predicting image rotations or spatial relationships between patches, enable robust pre-training before downstream fine-tuning on limited labeled data [5,161]. SSL pre-training on unlabeled video frames from agricultural fields can substantially enhance the performance of object detectors like YOLOv5 in weed identification tasks where annotations are sparse [76,162]. These methods, initially developed for data-constrained fields such as medical imaging and natural language processing, are now increasingly adopted in agricultural applications to improve generalizability and data efficiency [154,163].

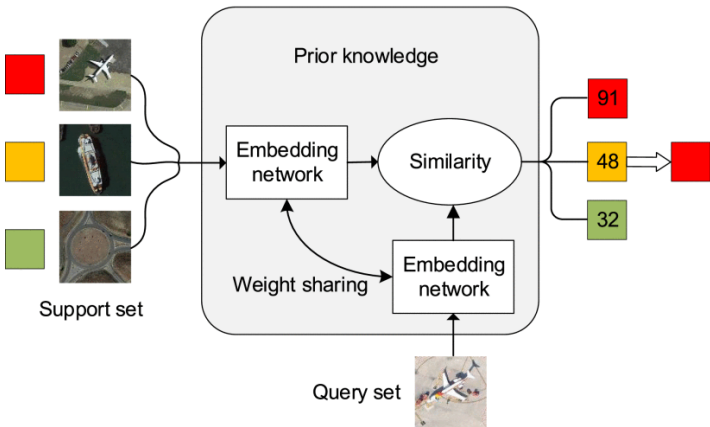


Figure 10. Illustration of a metric-based few-shot learning framework using shared embedding networks to compare a query image against a limited labeled support set. The model assigns the class based on similarity scores to prototype representations. This approach is applicable in agriculture for pest or disease classification with limited annotated data [164].

7.3. Multimodal Approaches

Integrating RGB imagery with complementary sensing modalities, such as hyperspectral, thermal, and LiDAR imaging, has emerged as a powerful strategy for mitigating the effects of environmental

variability in agricultural object detection [163]. Hyperspectral imaging, by capturing reflectance patterns across a wide range of wavelengths beyond the visible spectrum, facilitates the differentiation of crop health conditions, enabling the early identification of stress symptoms or diseases even under low illumination [5]. Thermal imaging, which detects infrared radiation, provides the advantage of penetrating partial occlusions such as dense foliage and enhances nighttime detection of pests and wildlife by leveraging heat signatures [165]. Meanwhile, LiDAR technology contributes high-resolution 3D structural information that supports precise tasks such as crop row localization and terrain mapping for autonomous machinery [80].

The fusion of such heterogeneous modalities, whether through multi-stream convolutional neural networks or attention-based architectures like Transformers, has been shown to significantly improve detection accuracy. Multimodal frameworks combining RGB and thermal imagery have achieved performance gains of 10–15% in mean Average Precision (mAP) in complex agricultural environments [14,166]. Recent architectures also utilize generative adversarial networks (GANs) to perform high-resolution reconstruction by fusing RGB and thermal inputs, further enhancing the utility of multimodal sensing in complex agricultural scenarios [11]. These multimodal approaches align with broader trends in robotics and autonomous driving, where sensor fusion is leveraged to enhance perception robustness and contextual awareness [13].

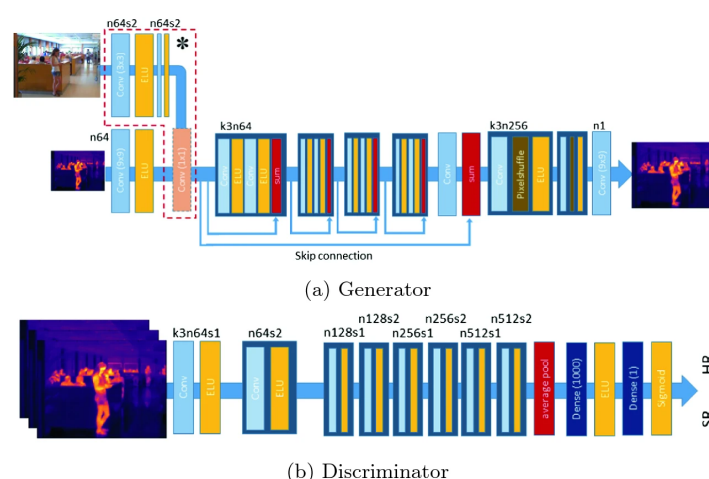


Figure 11. Example of a multimodal fusion network using RGB and thermal imagery. The generator fuses both modalities to reconstruct high-resolution thermal outputs, while the discriminator evaluates output quality. Such architectures improve the fidelity and robustness of downstream tasks like detection or mapping [167].

7.4. Federated Learning

Federated learning (FL) offers a privacy-preserving framework for collaborative model training across decentralized sources, addressing the dual challenges of data scarcity and confidentiality in agriculture [163]. In FL, local models, typically deployed on edge devices at individual farms, train on private datasets (e.g., drone-captured crop imagery) and communicate only model parameters or gradients to a central aggregator, which synthesizes a global model without accessing raw data [168]. This process is illustrated in Figure 12, where local models update independently and only share model parameters, not raw data, preserving privacy throughout the learning cycle. Models such as attention-enhanced CNNs, which have been proven effective in tasks like tomato leaf disease diagnosis [169,170], can serve as local learners in federated pipelines.

In agricultural object detection, federated versions of deep models like YOLOv5 have demonstrated promising results. Studies show that FL can yield up to a 20% improvement in mean Average Precision (mAP) over isolated, locally-trained models when applied to rare pest or disease detection across farms with limited annotations [171,172]. This framework parallels developments in domains such as healthcare and smart cities, where federated learning supports secure and collaborative AI development under strict data protection requirements [12].

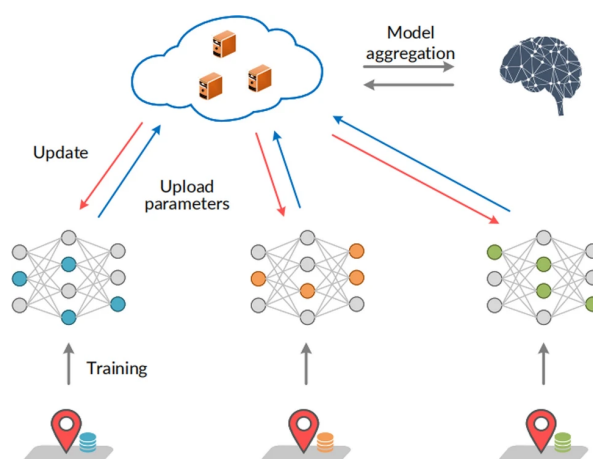


Figure 12. Illustration of federated learning architecture. Local devices train models on private data and share only model parameters with a central server, which aggregates them to update a global model. This decentralized paradigm enables collaborative learning without compromising data privacy [173].

7.5. Edge AI Optimization

Real-time agricultural operations, ranging from on-site disease detection to autonomous robot navigation, demand efficient deep learning models deployable on edge devices with constrained resources. Edge AI optimization techniques, such as model pruning, quantization, and knowledge distillation, are instrumental in reducing model size and computation without sacrificing accuracy [1]. Pruning YOLOv5 can reduce its size from 50 MB to 10 MB, while maintaining accuracy and achieving frame rates exceeding 30 FPS on lightweight platforms like the Jetson Nano, significantly outperforming heavier architectures such as Faster R-CNN [76,162].

Lightweight neural network architectures, including MobileNet, YOLO-NAS, and TinyML frameworks, are increasingly adopted for tasks like fruit counting, weed detection, and livestock tracking on handheld or mobile platforms [174,175]. These strategies support low-latency and energy-efficient inference, which is vital for large-scale deployment in remote or infrastructure-limited agricultural regions. This direction aligns with broader AI and IoT trends that emphasize distributed intelligence and real-time responsiveness in resource-constrained environments [172].

8. Conclusions

Object detection has emerged as a transformative technology within precision agriculture, enabling significant advancements in crop monitoring, weed management, pest detection, and autonomous field operations. The evolution from traditional hand-crafted feature-based methods to modern deep learning architectures, including Faster R-CNN, YOLO, SSD, and EfficientDet, has substantially improved detection accuracy, scalability, and real-time applicability. The availability of agricultural-specific datasets, such as PlantVillage, DeepWeeds, and AgriNet, has further catalyzed progress, though challenges related to dataset diversity, annotation quality, and domain adaptation persist.

Despite substantial progress, object detection in agriculture faces persistent challenges, including environmental variability, limited labeled data, generalization difficulties, real-time constraints, and the need for interpretability. Addressing these issues requires integrating emerging AI paradigms such as few-shot learning, edge optimization, multimodal sensing, federated learning, and explainable AI. Future advancements in robust, efficient, and interpretable models will be essential to support smart farming practices, enhance agricultural productivity, and contribute valuable insights to broader AI research.

Author Contributions: Conceptualization, Y.S., Z.K., H.L.; methodology, Z.K.; validation, Y.S., Z.K., H.L.; formal analysis, Y.S., H.L.; investigation, Z.K.; resources, Y.S., H.L.; writing—original draft preparation, Z.K.; writing—

review and editing, Z.K., H.L.; supervision, Y.S., H.L.; project administration, Y.S., H.L.; funding acquisition, H.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by National Natural Science Foundation of China, grant number 32171908.

Institutional Review Board Statement: Not applicable

Informed Consent Statement: Not applicable

Data Availability Statement: Not applicable

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Alif, M.A.R.; Hussain, M. YOLOv1 to YOLOv10: A comprehensive review of YOLO variants and their application in the agricultural domain. *arXiv preprint arXiv:2406.10139* **2024**.
2. Qiu, D.; Guo, T.; Yu, S.; Liu, W.; Li, L.; Sun, Z.; Hu, D. Classification of Apple Color and Deformity Using Machine Vision Combined with CNN. *Agriculture* **2024**, *14*, 978.
3. Ji, W.; Zhai, K.; Xu, B.; Wu, J. Green Apple Detection Method Based on Multidimensional Feature Extraction Network Model and Transformer Module. *Journal of Food Protection* **2025**, *88*, 100397.
4. Suganthi, S.U.; Prinslin, L.; Selvi, R.; Prabha, R. Generative AI in Agri: Sustainability in Smart Precision Farming Yield Prediction Mapping System Based on GIS Using Deep Learning and GPS. *Procedia Computer Science* **2025**, *252*, 365–380.
5. Zhou, Z.; Majeed, Y.; Naranjo, G.D.; Gambacorta, E.M. Assessment for crop water stress with infrared thermal imagery in precision agriculture: A review and future prospects for deep learning applications. *Computers and Electronics in Agriculture* **2021**, *182*, 106019.
6. Wang, H.; Li, J.; Dong, H. A Review of Vision-Based Multi-Task Perception Research Methods for Autonomous Vehicles. *Sensors* **2025**, *25*, 2611. <https://doi.org/10.3390/s25082611>.
7. Wang, H.; Gu, J.; Wang, M. A review on the application of computer vision and machine learning in the tea industry. *Frontiers In Sustainable Food Systems* **2023**, *7*, n/a.
8. Sun, J.; Cong, S.; Mao, H.; Wu, X.; Yang, N. Quantitative Detection of Mixed Pesticide Residue of Lettuce Leaves Based on Hyperspectral Technique. *Journal of Food Process Engineering* **2018**, *41*, e12654.
9. Wu, M.; Sun, J.; Lu, B.; Ge, X.; Zhou, X.; Zou, M. Application of Deep Brief Network in Transmission Spectroscopy Detection of Pesticide Residues in Lettuce Leaves. *Journal of Food Process Engineering* **2019**, *42*, e13005.
10. Sun, J.; Ge, X.; Wu, X.; Dai, C.; Yang, N. Identification of Pesticide Residues in Lettuce Leaves Based on Near Infrared Transmission Spectroscopy. *Journal of Food Process Engineering* **2018**, *41*, e12816.
11. Sharma, A.; Jain, A.; Gupta, P.; Chowdary, V. Machine learning applications for precision agriculture: A comprehensive review. *IEEE Access* **2020**, *9*, 4843–4873.
12. Akhter, R.; Sofi, S.A. Precision agriculture using IoT data analytics and machine learning. *Journal of King Saud University-Computer and Information Sciences* **2022**, *34*, 5602–5618.
13. Zhang, G.; Tian, Y.; Yin, W.; Zheng, C. An apple detection and localization method for automated harvesting under adverse light conditions. *Agriculture* **2024**, *14*, 485.
14. Zhao, H.; Tang, Z.; Li, Z.; Dong, Y.; Si, Y.; Lu, M.; Panoutsos, G. Real-time object detection and robotic manipulation for agriculture using a YOLO-based learning approach. *2024 IEEE International Conference on Industrial Technology (ICIT)* **2024**, pp. 1–6.
15. Kashyap, P.K.; Kumar, S.; Jaiswal, A.; Prasad, M.; Gandomi, A.H. Towards precision agriculture: IoT-enabled intelligent irrigation systems using deep learning neural network. *IEEE Sensors Journal* **2021**, *21*, 17479–17491.
16. Zhang, F.; Chen, Z.; Ali, S.; Yang, N.; Fu, S.; Zhang, Y. Multi-class detection of cherry tomatoes using improved YOLOv4-Tiny. *International Journal of Agricultural and Biological Engineering* **2023**, *16*, 225–231.
17. Zhou, X.; Jun, S.; Yan, T.; Bing, L.; Hang, Y.; Quansheng, C. Hyperspectral technique combined with deep learning algorithm for detection of compound heavy metals in lettuce. *Food Chemistry* **2020**, *321*, n/a.
18. Sabir, R.M.; Mehmood, K.; Sarwar, A.; Safdar, M.; Muhammad, N.E.; Gul, N.; Akram, H.M.B. Remote Sensing and Precision Agriculture: A Sustainable Future. In *Transforming Agricultural Management for a Sustainable Future: Climate Change and Machine Learning Perspectives*; Springer Nature Switzerland, 2024; pp. 75–103. https://doi.org/10.1007/978-3-031-63430-7_4.

19. Xiuyan, G.A.O.; ZHANG, Y. Detection of Fruit using YOLOv8-based Single Stage Detectors. *International Journal of Advanced Computer Science & Applications* **2023**, *14*, n/a.
20. Zheng, Y.Y.; Kong, J.L.; Jin, X.B.; Wang, X.Y.; Su, T.L.; Zuo, M. CropDeep: The crop vision dataset for deep-learning-based classification and detection in precision agriculture. *Sensors* **2019**, *19*, 1058.
21. Li, Z.; Wang, D.; Zhu, T.; Tao, Y.; Ni, C. Review of deep learning-based methods for non-destructive evaluation of agricultural products. *Biosystems Engineering* **2024**, *245*, 56–83.
22. Ahmed, S.; Qiu, B.; Ahmad, F.; Kong, C.W.; Xin, H. A State-of-the-Art Analysis of Obstacle Avoidance Methods from the Perspective of an Agricultural Sprayer UAV's Operation Scenario. *Agronomy* **2021**, *11*, 1069.
23. Liu, H.; Zhu, H. Evaluation of a Laser Scanning Sensor in Detection of Complex-Shaped Targets for Variable-Rate Sprayer Development. *Transactions of the ASABE* **2016**, *59*, 1181–1192.
24. Myers, V.I.; Allen, W.A. Electrooptical remote sensing methods as nondestructive testing and measuring techniques in agriculture. *Applied Optics* **1968**, *7*, 1819–1838.
25. Hu, T.; Wang, W.; Gu, J.; Xia, Z.; Zhang, J.; Wang, B. Research on Apple Object Detection and Localization Method Based on Improved YOLOX and RGB-D Images. *Agronomy* **2023**, *13*, 1816. <https://doi.org/10.3390/agronomy13071816>.
26. Xie, D.; Chen, L.; Liu, L.; Chen, L.; Wang, H. Actuators and sensors for application in agricultural robots: A review. *Machines* **2022**, *10*, 913. <https://doi.org/10.3390/machines10100913>.
27. Xiong, Y.; Peng, C.; Grimstad, L.; From, P.J.; Isler, V. Development and field evaluation of a strawberry harvesting robot with a cable-driven gripper. *Computers and Electronics in Agriculture* **2019**, *157*, 392–402. <https://doi.org/10.1016/j.compag.2019.01.006>.
28. Khan, Z.; Liu, H.; Shen, Y.; Zeng, X. Deep learning improved YOLOv8 algorithm: Real-time precise instance segmentation of crown region orchard canopies in natural environment. *Computers and Electronics in Agriculture* **2024**, *224*, 109168. <https://doi.org/10.1016/j.compag.2024.109168>.
29. Khan, Z.; Liu, H.; Shen, Y.; Yang, Z.; Zhang, L.; Yang, F. Optimizing precision agriculture: A real-time detection approach for grape vineyard unhealthy leaves using deep learning improved YOLOv7 with feature extraction capabilities. *Computers and Electronics in Agriculture* **2025**, *231*, 109969.
30. Chen, C.; Zhang, P.; Zhang, H.; Dai, J.; Yi, Y.; Zhang, H.; Zhang, Y. Deep Learning on Computational-Resource-Limited Platforms: A Survey. *Mobile Information Systems* **2020**, *2020*, 1–19. <https://doi.org/10.1155/2020/8454327>.
31. Qin, Y.M.; Tu, Y.H.; Li, T.; Ni, Y.; Wang, R.F.; Wang, H. Deep Learning for Sustainable Agriculture: A Systematic Review on Applications in Lettuce Cultivation. *Sustainability* **2025**, *17*, 3190. <https://doi.org/10.3390/su17073190>.
32. Lv, R.; Hu, J.; Zhang, T.; Chen, X.; Liu, W. Crop-Free-Ridge Navigation Line Recognition Based on the Lightweight Structure Improvement of YOLOv8. *Agriculture* **2025**, *15*, 942. <https://doi.org/10.3390/agriculture15090942>.
33. Joshi, H. Edge-AI for Agriculture: Lightweight Vision Models for Disease Detection in Resource-Limited Settings. *arXiv preprint arXiv:2412.18635* **2024**.
34. Yu, K.; Zhong, M.; Zhu, W.; Rashid, A.; Han, R.; Virk, M.S.; Ren, X. Advances in Computer Vision and Spectroscopy Techniques for Non-Destructive Quality Assessment of Citrus Fruits: A Comprehensive Review. *Foods* **2025**, *14*, 386. <https://doi.org/10.3390/foods14030386>.
35. Girshick, R.; Donahue, J.; Darrell, T.; Malik, J. Rich feature hierarchies for accurate object detection and semantic segmentation. In Proceedings of the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2014, pp. 580–587.
36. Girshick, R. Fast R-CNN. In Proceedings of the Proceedings of the IEEE International Conference on Computer Vision (ICCV). IEEE, 2015, pp. 1440–1448. <https://doi.org/10.1109/ICCV.2015.169>.
37. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. In Proceedings of the Advances in Neural Information Processing Systems, 2015, Vol. 28.
38. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You Only Look Once: Unified, Real-Time Object Detection. In Proceedings of the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 779–788. <https://doi.org/10.1109/CVPR.2016.91>.
39. Redmon, J.; Farhadi, A. YOLOv3: An Incremental Improvement, 2018, [arXiv:cs.CV/1804.02767].
40. Wang, C.Y.; Bochkovskiy, A.; Liao, H.Y.M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors, 2022, [arXiv:cs.CV/2207.02696].

41. Liu, W.; Anguelov, D.; Erhan, D.; Szegedy, C.; Reed, S.; Fu, C.Y.; Berg, A.C. SSD: Single Shot MultiBox Detector. In Proceedings of the European Conference on Computer Vision. Springer, 2016, pp. 21–37. https://doi.org/10.1007/978-3-319-46448-0_2.
42. Ji, W.; Pan, Y.; Xu, B.; Wang, J. A Real-Time Apple Targets Detection Method for Picking Robot Based on ShuffleNetV2-YOLOX. *Agriculture* **2022**, *12*, 856. <https://doi.org/10.3390/agriculture12060856>.
43. Liu, X.; Jia, W.; Ruan, C.; Zhao, D.; Gu, Y.; Chen, W. The recognition of apple fruits in plastic bags based on block classification. *Precision Agriculture* **2018**, *19*, 735–749.
44. Uijlings, J.R.; van de Sande, K.E.; Gevers, T.; Smeulders, A.W. Selective Search for Object Recognition. *International Journal of Computer Vision* **2013**, *104*, 154–171.
45. Sladojevic, S.; Arsenovic, M.; Anderla, A.; Culibrk, D.; Stefanovic, D. Deep Neural Networks Based Recognition of Plant Diseases by Leaf Image Classification. *Computational Intelligence and Neuroscience* **2016**, *2016*, 3289801. <https://doi.org/10.1155/2016/3289801>.
46. Hossain, S.; Mou, R.M.; Hasan, M.M.; Chakraborty, S.; Razzak, M.A. Recognition and detection of tea leaf's diseases using support vector machine. In Proceedings of the 2018 IEEE 14th International Colloquium on Signal Processing & Its Applications (CSPA), 2018, pp. 150–154. <https://doi.org/10.1109/CSPA.2018.8368748>.
47. Tang, L.; Tian, L.; Steward, B.L. Classification of broadleaf and grass weeds using Gabor wavelets and an artificial neural network. *Transactions of the ASAE* **2003**, *46*, 1247–1254. <https://doi.org/10.13031/2013.13944>.
48. Zhu, Y.; Fan, S.; Zuo, M.; Zhang, B.; Zhu, Q.; Kong, J. Discrimination of New and Aged Seeds Based on On-Line Near-Infrared Spectroscopy Technology Combined with Machine Learning. *Foods* **2024**, *13*, 1570.
49. Jin, X. Development status and trend of agricultural robot technology. *International Journal of Agricultural and Biological Engineering* **2021**, *14*, 1–14. <https://doi.org/10.25165/j.ijabe.20211404.6821>.
50. Zhang, Z.; Lu, Y.; Zhao, Y.; Pan, Q.; Jin, K.; Xu, G.; Hu, Y. TS-YOLO: An all-day and lightweight tea canopy shoots detection model. *Agronomy* **2023**, *13*, 1411. <https://doi.org/10.3390/agronomy13051411>.
51. Ge, X.; Sun, J.; Lu, B.; Chen, Q.; Xun, W.; Jin, Y. Classification of Oolong Tea Varieties Based on Hyperspectral Imaging Technology and BOSS-LightGBM Model. *Journal of Food Process Engineering* **2019**, *42*, e13289.
52. Paymode, A.S.; Malode, V.B. Transfer learning for multi-crop leaf disease image classification using convolutional neural network VGG. *Artificial Intelligence in Agriculture* **2022**, *6*, 23–33. <https://doi.org/10.1016/j.aiia.2021.12.002>.
53. Peng, Y.; Zhao, S.; Liu, J. Fused-Deep-Features Based Grape Leaf Disease Diagnosis. *Agronomy* **2021**, *11*, 2234. <https://doi.org/10.3390/agronomy11112234>.
54. Xu, C.; Lu, C.; Piao, J.; Wang, Y.; Zhou, Y.; Li, S. Rice virus release from the planthopper salivary gland is independent of plant tissue recognition by the stylet. *Pest Management Science* **2020**, *76*, 3208–3216. <https://doi.org/10.1002/ps.5876>.
55. Yang, N.; Qian, Y.; EL-Mesery, H.S.; Zhang, R.; Wang, A.; Tang, J. Rapid Detection of Rice Disease Using Microscopy Image Identification Based on the Synergistic Judgment of Texture and Shape Features and Decision Tree–Confusion Matrix Method. *Journal of the Science of Food and Agriculture* **2019**, *99*, 6589–6600.
56. Viveros Escamilla, L.D.; Gómez-Espinosa, A.; Escobedo Cabello, J.A.; Cantoral-Ceballos, J.A. Maturity recognition and fruit counting for sweet peppers in greenhouses using deep learning neural networks. *Agriculture* **2024**, *14*, 331. <https://doi.org/10.3390/agriculture14030331>.
57. Fatima, H.S.; ul Hassan, I.; Hasan, S.; Khurram, M.; Stricker, D.; Afzal, M.Z. Formation of a Lightweight, Deep Learning-Based Weed Detection System for a Commercial Autonomous Laser Weeding Robot. *Applied Sciences* **2023**, *13*, 3997. <https://doi.org/10.3390/app13063997>.
58. Amjoud, A.B.; Amrouch, M. Object Detection Using Deep Learning, CNNs and Vision Transformers: A Review. *IEEE Access* **2023**, *11*, 35479–35516. <https://doi.org/10.1109/ACCESS.2023.3266093>.
59. Sun, T.; Zhang, W.; Miao, Z.; Zhang, Z.; Li, N. Object localization methodology in occluded agricultural environments through deep learning and active sensing. *Computers and Electronics in Agriculture* **2023**, *212*, 108141.
60. Saranya, T.; Deisy, C.; Sridevi, S.; Anbananthan, K.S.M. A comparative study of deep learning and Internet of Things for precision agriculture. *Engineering Applications of Artificial Intelligence* **2023**, *122*, 106034.
61. Wang, Y.; Han, Y.; Wang, C.; Song, S.; Tian, Q.; Huang, G. Computation-efficient Deep Learning for Computer Vision: A Survey. *Cybernetics and Intelligence* **2024**. <https://doi.org/10.48550/arXiv.2308.13998>.
62. Ariza-Sentís, M.; Vélez, S.; Martínez-Peña, R.; Baja, H.; Valente, J. Object detection and tracking in Precision Farming: A systematic review. *Computers and Electronics in Agriculture* **2024**, *219*, 108757. <https://doi.org/10.1016/j.compag.2024.108757>.

63. Zoubek, T.; Bumbálek, R.; Ufitikirezi, J.D.D.M.; Strob, M.; Filip, M.; Špalek, F.; Bartoš, P. Advancing precision agriculture with computer vision: A comparative study of YOLO models for weed and crop recognition. *Crop Protection* **2025**, *190*, 107076.
64. Olsen, A.; Kononov, D.A.; Philippa, B.; Ridd, P.; Wood, J.C.; Johns, J.; Banks, W.; Girgenti, B.; Kenny, O.; Whinney, J.; et al. DeepWeeds: A Multiclass Weed Species Image Dataset for Deep Learning. *Scientific Reports* **2019**, *9*, 2058. <https://doi.org/10.1038/s41598-018-38343-3>.
65. Saleem, M.H.; Velayudhan, K.K.; Potgieter, J.; Arif, K.M. Weed Identification by Single-Stage and Two-Stage Neural Networks: A Study on the Impact of Image Resizers and Weights Optimization Algorithms. *Frontiers in Plant Science* **2022**, *13*, 850666. <https://doi.org/10.3389/fpls.2022.850666>.
66. Zhu, W.; Sun, J.; Wang, S.; Shen, J.; Yang, K.; Zhou, X. Identifying Field Crop Diseases Using Transformer-Embedded Convolutional Neural Network. *Agriculture* **2022**, *12*, 1083. <https://doi.org/10.3390/agriculture12081083>.
67. Tang, S.; Xia, Z.; Gu, J.; Wang, W.; Huang, Z.; Zhang, W. High-precision apple recognition and localization method based on RGB-D and improved SOLOv2 instance segmentation. *Frontiers in Sustainable Food Systems* **2024**, *8*, 1403872. <https://doi.org/10.3389/fsufs.2024.1403872>.
68. Shen, L.; Su, J.; Huang, R.; Quan, W.; Song, Y.; Fang, Y.; Su, B. Fusing attention mechanism with Mask R-CNN for instance segmentation of grape cluster in the field. *Frontiers in Plant Science* **2022**, *13*, 934450. <https://doi.org/10.3389/fpls.2022.934450>.
69. Yang, J.; Han, M.; He, J.; Wen, J.; Chen, D.; Wang, Y. Object detection and localization algorithm in agricultural scenes based on YOLOv5. *Journal of Electronic Imaging* **2023**, *32*, 052402.
70. Tan, M.; Pang, R.; Le, Q.V. EfficientDet: Scalable and Efficient Object Detection. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* **2020**, pp. 10781–10790. <https://doi.org/10.1109/CVPR42600.2020.01080>.
71. Chu, P.; Li, Z.; Zhang, K.; Chen, D.; Lammers, K.; Lu, R. O2RNet: Occluder-Occludee Relational Network for Robust Apple Detection in Clustered Orchard Environments. *Smart Agricultural Technology* **2023**, *5*, 100284. <https://doi.org/10.1016/j.atech.2023.100284>.
72. Sun, J.; He, X.; Ge, X.; Wu, X.; Shen, J.; Song, Y. Detection of Key Organs in Tomato Based on Deep Migration Learning in a Complex Background. *Agriculture* **2018**, *8*, 196. <https://doi.org/10.3390/agriculture8120196>.
73. Wang, A.; Gao, B.; Cao, H.; Wang, P.; Zhang, T.; Wei, X. Early detection of Sclerotinia sclerotiorum on oilseed rape leaves based on optical properties. *Biosystems Engineering* **2022**, *224*, 80–91. <https://doi.org/10.1016/j.biosystemseng.2022.09.005>.
74. Jasim, M.; Al-Tuwaijari, A. Detection and identification of plant leaf diseases using YOLOv4. *PLOS ONE* **2023**, *18*, e0284567. <https://doi.org/10.1371/journal.pone.0284567>.
75. Duan, Y.; Han, W.; Guo, P.; Wei, X. YOLOv8-GDCI: Research on the Phytophthora Blight Detection Method of Different Parts of Chili Based on Improved YOLOv8 Model. *Agronomy* **2024**, *14*, 2734. <https://doi.org/10.3390/agronomy14112734>.
76. Sun, F.; Lv, Q.; Bian, Y.; He, R.; Lv, D.; Gao, L.; Li, X. Grape Target Detection Method in Orchard Environment Based on Improved YOLOv7. *Agronomy* **2025**, *15*, 42. <https://doi.org/10.3390/agronomy15010042>.
77. Muhammad, A.; Salman, Z.; Lee, K.; Han, D. Harnessing the power of diffusion models for plant disease image augmentation. *Frontiers in Plant Science* **2023**, *14*, 1280496. <https://doi.org/10.3389/fpls.2023.1280496>.
78. Yang, R.; Yu, Y. Artificial convolutional neural network in object detection and semantic segmentation for medical imaging analysis. *Frontiers in Oncology* **2021**, *11*, 638182. <https://doi.org/10.3389/fonc.2021.638182>.
79. Huang, S.; Wu, S.; Sun, C.; Ma, X.; Jiang, Y.; Qi, L. Deep localization model for intra-row crop detection in paddy field. *Computers and Electronics in Agriculture* **2020**, *169*, 105203.
80. Ding, H.; Zhang, B.; Zhou, J.; Yan, Y.; Tian, G.; Gu, B. Recent developments and applications of simultaneous localization and mapping in agriculture. *Journal of Field Robotics* **2022**, *39*, 956–983.
81. Milioto, A.; Lottes, P.; Stachniss, C. Real-time semantic segmentation of crop and weed for precision agriculture robots leveraging background knowledge in CNNs. In Proceedings of the 2018 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2018, pp. 2229–2235. <https://doi.org/10.1109/ICRA.2018.8460968>.
82. Pang, Y.; Shi, Y.; Gao, S.; Jiang, F.; Veeranampalayam-Sivakumar, A.N.; Thompson, L.; Luck, J.; Liu, C. Improved crop row detection with deep neural network for early-season maize stand count in UAV imagery. *Computers and Electronics in Agriculture* **2020**, *178*, 105766. <https://doi.org/10.1016/j.compag.2020.105766>.

83. Milioto, A.; Lottes, P.; Stachniss, C. Real-time blob-wise sugar beets vs weeds classification for monitoring fields using convolutional neural networks. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **2017**, *4*, 41–48. <https://doi.org/10.5194/isprs-annals-IV-2-W3-41-2017>.
84. Chiu, M.T.; Xu, X.; Wei, Y.; Huang, Z.; Schwing, A.G.; Brunner, R.; Khachatryan, H.; Karapetyan, H.; Dozier, I.; Rose, G.; et al. Agriculture-Vision: A Large Aerial Image Database for Agricultural Pattern Analysis. In Proceedings of the Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020, pp. 2828–2838. <https://doi.org/10.1109/CVPR42600.2020.00290>.
85. Lu, Y.; Young, S. A survey of public datasets for computer vision tasks in precision agriculture. *Computers and Electronics in Agriculture* **2020**, *178*, 105760. <https://doi.org/10.1016/j.compag.2020.105760>.
86. Hughes, D.P.; Salathé, M. An open access repository of images on plant health to enable the development of mobile disease diagnostics. *arXiv preprint arXiv:1511.08060* **2015**.
87. Rahnemoonfar, M.; Sheppard, C. Deep count: fruit counting based on deep simulated learning. *Sensors* **2017**, *17*, 905. <https://doi.org/10.3390/s17050905>.
88. Liu, J.; Abbas, I.; Noor, R.S. Development of Deep Learning-Based Variable Rate Agrochemical Spraying System for Targeted Weeds Control in Strawberry Crop. *Agronomy* **2021**, *11*, 1480. <https://doi.org/10.3390/agronomy11081480>.
89. Kang, R.; Huang, J.; Zhou, X.; Ren, N.; Sun, S. Toward Real Scenery: A Lightweight Tomato Growth Inspection Algorithm for Leaf Disease Detection and Fruit Counting. *Plant Phenomics* **2024**, *6*, 0174. <https://doi.org/10.34133/plantphenomics.0174>.
90. Tao, T.; Wei, X. STBNA-YOLOv5: An Improved YOLOv5 Network for Weed Detection in Rapeseed Field. *Agriculture* **2025**, *15*, 22. <https://doi.org/10.3390/agriculture15010022>.
91. Gerhards, R.; Andujar Sanchez, D.; Hamouz, P.; Peteinatos, G.G.; Christensen, S.; Fernandez-Quintanilla, C. Advances in site-specific weed management in agriculture—A review. *Weed Research* **2022**, *62*, 123–133. <https://doi.org/10.1111/wre.12515>.
92. Zhang, X.; Li, H.; Sun, S.; Zhang, W.; Shi, F.; Zhang, R.; Liu, Q. Classification and identification of apple leaf diseases and insect pests based on improved ResNet-50 model. *Horticulturae* **2023**, *9*, 1046. <https://doi.org/10.3390/horticulturae9091046>.
93. Bjerger, K.; Frigaard, C.E.; Karstoft, H. Object Detection of Small Insects in Time-Lapse Camera Recordings. *Sensors* **2023**, *23*, 7242. <https://doi.org/10.3390/s23167242>.
94. Venkateswara, S.M.; Padmanabhan, J. Deep Learning Based Agricultural Pest Monitoring and Classification. *Scientific Reports* **2025**, *15*, 8684. <https://doi.org/10.1038/s41598-025-92659-5>.
95. Al Sahili, Z.; Awad, M. The Power of Transfer Learning in Agricultural Applications: AgriNet. *Frontiers in Plant Science* **2022**, *13*, 992700. <https://doi.org/10.3389/fpls.2022.992700>.
96. Čirjak, D.; Aleksi, I.; Lemic, D.; Pajač Živković, I. EfficientDet-4 Deep Neural Network-Based Remote Monitoring of Codling Moth Population for Early Damage Detection in Apple Orchard. *Agriculture* **2023**, *13*, 961. <https://doi.org/10.3390/agriculture13050961>.
97. McKechnie, I.; Raymond, K.; Stacey, D. Identifying Inconsistencies in Data Quality Between FAOSTAT, WOA, UN Agriculture Census, and National Data. *Data Science Journal* **2024**, *23*. <https://doi.org/10.5334/dsj-2024-001>.
98. Garcin, C.; Joly, A.; Bonnet, P.; Lombardo, J.C.; Affouard, A.; Chouet, M.; Servajean, M.; Lorieul, T.; Salmon, J. Pl@ntNet-300K: a plant image dataset with high label ambiguity and a long-tailed distribution. In Proceedings of the NeurIPS Datasets and Benchmarks 2021, 2021.
99. Alibabaei, K.; Assunção, E.; Gaspar, P.D.; Soares, V.N.G.J.; Caldeira, J.M.L.P. Real-Time Detection of Vine Trunk for Robot Localization Using Deep Learning Models Developed for Edge TPU Devices. *Future Internet* **2022**, *14*, 199. <https://doi.org/10.3390/fi14070199>.
100. Noyan, M.A. Uncovering Bias in the PlantVillage Dataset. *arXiv preprint arXiv:2206.04374* **2022**. <https://doi.org/10.48550/arXiv.2206.04374>.
101. Lu, D.; Wang, Y. MAR-YOLOv9: A Multi-Dataset Object Detection Method for Agricultural Fields Based on YOLOv9. *PLOS ONE* **2024**, *19*, e0307643. <https://doi.org/10.1371/journal.pone.0307643>.
102. Li, T.; Feng, Q.; Qiu, Q.; Xie, F.; Zhao, C. Occluded Apple Fruit Detection and Localization with a Frustum-Based Point-Cloud-Processing Approach for Robotic Harvesting. *Remote Sensing* **2022**, *14*, 482. <https://doi.org/10.3390/rs14030482>.
103. Cravero, A.; Pardo, S.; Sepúlveda, S.; Muñoz, L. Challenges to Use Machine Learning in Agricultural Big Data: A Systematic Literature Review. *Agronomy* **2022**, *12*, 748. <https://doi.org/10.3390/agronomy12030748>.

104. Rufin, P.; Wang, S.; Lisboa, S.N.; Hemmerling, J.; Tulbure, M.G.; Meyfroidt, P. Taking it further: Leveraging pseudo-labels for field delineation across label-scarce smallholder regions. *International Journal of Applied Earth Observation and Geoinformation* **2024**, *134*, 104149. <https://doi.org/10.1016/j.jag.2024.104149>.
105. Li, L.; Xie, S.; Ning, J.; Chen, Q.; Zhang, Z. Evaluating green tea quality based on multisensor data fusion combining hyperspectral imaging and olfactory visualization systems. *Journal of the Science of Food and Agriculture* **2019**, *99*, 1787–1794. <https://doi.org/10.1002/jsfa.9371>.
106. Schuhmann, C.; Beaumont, R.; Vencu, R.; Gordon, C.; Wightman, R.; Cherti, M.; Coombes, T.; Katta, A.; Mullis, C.; Wortsman, M.; et al. LAION-5B: An open large-scale dataset for training next generation image-text models. *Advances in Neural Information Processing Systems* **2022**, *35*, 25278–25294.
107. Huang, J.; Rathod, V.; Sun, C.; Zhu, M.; Korattikara, A.; Fathi, A.; Fischer, I.; Wojna, Z.; Song, Y.; Guadarrama, S.; et al. Speed/Accuracy Trade-Offs for Modern Convolutional Object Detectors. In Proceedings of the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 7310–7311. <https://doi.org/10.1109/CVPR.2017.351>.
108. Li, A.; Wang, C.; Ji, T.; Wang, Q.; Zhang, T. D3-YOLOv10: Improved YOLOv10-Based Lightweight Tomato Detection Algorithm Under Facility Scenario. *Agriculture* **2024**, *14*, 2268. <https://doi.org/10.3390/agriculture14122268>.
109. Dandekar, Y.; Shinde, K.; Gangan, J.; Firdausi, S.; Bharne, S. Weed Plant Detection from Agricultural Field Images using YOLOv3 Algorithm. In Proceedings of the 2022 6th International Conference on Computing, Communication, Control and Automation (ICCUBEA). IEEE, 2022, pp. 1–6.
110. Jin, X.; Sun, Y.; Che, J.; Bagavathiannan, M.; Yu, J.; Chen, Y. A novel deep learning-based method for detection of weeds in vegetables. *Pest Management Science* **2022**, *78*, 1861–1869.
111. Deng, L.; Miao, Z.; Zhao, X.; Yang, S.; Gao, Y.; Zhai, C.; Zhao, C. HAD-YOLO: An Accurate and Effective Weed Detection Model Based on Improved YOLOv5 Network. *Agronomy* **2025**, *15*, 57.
112. Mohanty, S.P.; Hughes, D.P.; Salathé, M. Using Deep Learning for Image-Based Plant Disease Detection. *Frontiers in Plant Science* **2016**, *7*, 1419. <https://doi.org/10.3389/fpls.2016.01419>.
113. Waheed, A.; Goyal, M.; Gupta, D.; Khanna, A.; Hassanien, A.E.; Pandey, H.M. An optimized dense convolutional neural network model for disease recognition and classification in corn leaf. *Computers and Electronics in Agriculture* **2020**, *175*, 105456. <https://doi.org/10.1016/j.compag.2020.105456>.
114. Lin, T.Y.; Goyal, P.; Girshick, R.; He, K.; Dollár, P. Focal Loss for Dense Object Detection. In Proceedings of the Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2017, pp. 2980–2988.
115. Li, M.; Zhang, Z.; Lei, L.; Wang, X.; Guo, X. Agricultural greenhouses detection in high-resolution satellite images based on convolutional neural networks: Comparison of Faster R-CNN, YOLO v3 and SSD. *Sensors* **2020**, *20*, 4938. <https://doi.org/10.3390/s20174938>.
116. Liu, S.; Li, Z.; Sun, J. Self-EMD: Self-Supervised Object Detection without ImageNet. *arXiv preprint arXiv:2011.13677* **2020**.
117. Gunay, M.; Koseoglu, M. Detection of circuit components on hand-drawn circuit images by using faster R-CNN method. *International Journal of Advanced Computer Science and Applications* **2021**, *12*, 1–7.
118. Štancel, M.; Hulič, M. An Introduction to Image Classification and Object Detection Using YOLO Detector. In Proceedings of the CEUR Workshop Proceedings, 2019, Vol. 2403, pp. 1–8.
119. Roboflow. YOLOv5 is Here: State-of-the-Art Object Detection at 140 FPS, 2020.
120. Xu, B.; Cui, X.; Ji, W.; Yuan, H.; Wang, J. Apple Grading Method Design and Implementation for Automatic Grader Based on Improved YOLOv5. *Agriculture* **2023**, *13*, 124. <https://doi.org/10.3390/agriculture13010124>.
121. Zuo, Z.; Gao, S.; Peng, H.; Xue, Y.; Han, L.; Ma, G.; Mao, H. Lightweight Detection of Broccoli Heads in Complex Field Environments Based on LBDC-YOLO. *Agronomy* **2024**, *14*, 2359. <https://doi.org/10.3390/agronomy14102359>.
122. Kulhandjian, H.; Yang, Y.; Amely, N. Design and Implementation of a Smart Agricultural Robot bull-DOG (SARDOG). In Proceedings of the 2024 International Conference on Computing, Networking and Communications (ICNC). IEEE, February 2024, pp. 767–771.
123. Fu, C.Y.; Liu, W.; Ranga, A.; Tyagi, A.; Berg, A.C. DSSD: Deconvolutional Single Shot Detector. *arXiv preprint arXiv:1701.06659* **2017**.
124. Zhang, H.; Hong, X.; Zhu, L. Detecting Small Objects in Thermal Images Using Single-Shot Detector. *arXiv preprint arXiv:2108.11101* **2021**.
125. Tan, M.; Le, Q.V. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. *arXiv preprint arXiv:1905.11946* **2019**.

126. Ultralytics. EfficientDet vs RTDETRv2: A Technical Comparison for Object Detection, 2023.
127. Wang, Y.; Qin, Y.; Cui, J. Occlusion Robust Wheat Ear Counting Algorithm Based on Deep Learning. *Frontiers in Plant Science* **2021**, *12*, 645899. <https://doi.org/10.3389/fpls.2021.645899>.
128. Rayhan, A. Artificial Intelligence in Robotics: From Automation to Autonomous Systems. *ResearchGate* **2023**. <https://doi.org/10.13140/RG.2.2.15540.42889>.
129. Xiao, Y.; Tian, Z.; Yu, J.; Zhang, Y.; Liu, S.; Du, S.; Lan, X. A review of object detection based on deep learning. *Multimedia Tools and Applications* **2020**, *79*, 23729–23791.
130. Wang, A.; Peng, T.; Cao, H.; Xu, Y.; Wei, X.; Cui, B. TIA-YOLOv5: An improved YOLOv5 network for real-time detection of crop and weed in the field. *Frontiers in Plant Science* **2022**, *13*, 1091655. <https://doi.org/10.3389/fpls.2022.1091655>.
131. Moreira, G.; Magalhães, S.A.; Pinho, T.M.; Cunha, M. Evaluating the Single-Shot MultiBox Detector and YOLO Deep Learning Models for the Detection of Tomatoes in a Greenhouse. *Sensors* **2021**, *21*, 3569. <https://doi.org/10.3390/s21123569>.
132. Li, E.; Zhou, Z.; Chen, X. Edge intelligence: On-demand deep learning model co-inference with device-edge synergy. *Proceedings of the 2018 Workshop on Mobile Edge Communications* **2018**, pp. 31–36. <https://doi.org/10.1145/3242487.3242490>.
133. Grigorescu, S.; Trasnea, B.; Cocias, T.; Macesanu, G. A survey of deep learning techniques for autonomous driving. *Journal of Field Robotics* **2020**, *37*, 362–386. <https://doi.org/10.1002/rob.21918>.
134. Cheng, H.; Zhang, M.; Shi, J.Q. A Survey on Deep Neural Network Pruning: Taxonomy, Comparison, Analysis, and Recommendations. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **2024**. <https://doi.org/10.1109/TPAMI.2024.3447085>.
135. Zhu, M.; Gupta, S. To Prune, or Not to Prune: Exploring the Efficacy of Pruning for Model Compression. *arXiv preprint arXiv:1710.01878* **2017**. <https://doi.org/10.48550/arXiv.1710.01878>.
136. Wang, R.; Liu, L.; Xie, C.; Yang, P.; Li, R.; Zhou, M. AgriPest: A Large-Scale Domain-Specific Benchmark Dataset for Practical Agricultural Pest Detection in the Wild. *Sensors* **2021**, *21*, 1601. <https://doi.org/10.3390/s21051601>.
137. Badgular, C.M.; Poulose, A.; Gan, H. Agricultural Object Detection with You Look Only Once (YOLO) Algorithm: A Bibliometric and Systematic Literature Review. *arXiv preprint arXiv:2401.10379* **2024**.
138. Shi, Y.; Han, L.; Zhang, X.; Sobeih, T.; Gaiser, T.; Thuy, N.H.; Behrend, D.; Srivastava, A.K.; Halder, K.; Ewert, F. Deep Learning Meets Process-Based Models: A Hybrid Approach to Agricultural Challenges. *arXiv preprint arXiv:2504.16141* **2025**.
139. Dong, M.; Yu, H.; Sun, Z.; Zhang, L.; Sui, Y.; Zhao, R. Research on Agricultural Environmental Monitoring Internet of Things Based on Edge Computing and Deep Learning. *Journal of Intelligent Systems* **2024**, *33*, 20230114. <https://doi.org/10.1515/jisys-2023-0114>.
140. Chiu, M.T.; Xu, X.; Wei, Y.; Huang, Z.; Schwing, A.; Brunner, R.; Khachatryan, H.; Karapetyan, H.; Dozier, I.; Rose, G.; et al. Agriculture-Vision: A Large Aerial Image Database for Agricultural Pattern Analysis. *arXiv preprint arXiv:2001.01306* **2020**.
141. Wang, T.s.; Kim, G.T.; Shin, J.; Jang, S.W. Hierarchical Image Quality Improvement Based on Illumination, Resolution, and Noise Factors for Improving Object Detection. *Electronics* **2024**, *13*, 4438. <https://doi.org/10.3390/electronics13224438>.
142. Li, Z.; Xiang, J.; Duan, J. A low illumination target detection method based on a dynamic gradient gain allocation strategy. *Scientific Reports* **2024**, *14*, 29058. <https://doi.org/10.1038/s41598-024-80265-w>.
143. Beldek, C.; Cunningham, J.; Aydin, M.; Sariyildiz, E.; Phung, S.L.; Alici, G. Sensing-based Robustness Challenges in Agricultural Robotic Harvesting. *arXiv preprint arXiv:2502.12403* **2025**.
144. Lyu, Z.; Jin, H.; Zhen, T.; Sun, F.; Xu, H. Small object recognition algorithm of grain pests based on SSD feature fusion. *IEEE Access* **2021**, *9*, 43202–43213. <https://doi.org/10.1109/ACCESS.2021.3066510>.
145. Silwal, A.; Parhar, T.; Yandun, F.; Kantor, G. A Robust Illumination-Invariant Camera System for Agricultural Applications. *arXiv preprint arXiv:2101.02190* **2021**.
146. Bargoti, S.; Underwood, J. Deep Fruit Detection in Orchards. *arXiv preprint arXiv:1610.03677* **2016**.
147. Liu, S.; Peng, D.; Zhang, B.; Chen, Z.; Yu, L.; Chen, J.; Yang, S. The Accuracy of Winter Wheat Identification at Different Growth Stages Using Remote Sensing. *Remote Sensing* **2022**, *14*, 893.
148. Kamath, U.; Liu, J.; Whitaker, J. Transfer Learning: Domain Adaptation. In *Deep Learning for NLP and Speech Recognition*; Springer, 2019; pp. 495–535. https://doi.org/10.1007/978-3-030-14596-5_11.
149. Kamilaris, A.; Prenafeta-Boldú, F.X. Deep learning in agriculture: A survey. *Computers and Electronics in Agriculture* **2018**, *147*, 70–90. <https://doi.org/10.1016/j.compag.2018.02.016>.

150. Migneco, P. Traffic sign recognition algorithm: a deep comparison between YOLOv5 and SSD Mobilenet. Doctoral dissertation, Politecnico di Torino, 2024.
151. Albahar, M. A survey on deep learning and its impact on agriculture: challenges and opportunities. *Agriculture* **2023**, *13*, 540. <https://doi.org/10.3390/agriculture13030540>.
152. Vincent, D.R.; Deepa, N.; Elavarasan, D.; Srinivasan, K.; Chauhdary, S.H.; Iwendi, C. Sensors driven AI-based agriculture recommendation model for assessing land suitability. *Sensors* **2019**, *19*, 3667. <https://doi.org/10.3390/s19173667>.
153. Yang, J.; Guo, X.; Li, Y.; Marinello, F.; Ercisli, S.; Zhang, Z. A survey of few-shot learning in smart agriculture: developments, applications, and challenges. *Plant Methods* **2022**, *18*, 1–15. <https://doi.org/10.1186/s13007-022-00866-2>.
154. Dhanya, V.; Subeesh, A.; Kushwaha, N.; Vishwakarma, D.; Kumar, T.; Ritika, G.; Singh, A. Deep learning based computer vision approaches for smart agricultural applications. *Artificial Intelligence in Agriculture* **2022**, *6*, 211–229. <https://doi.org/10.1016/j.aiia.2022.09.007>.
155. Hrast Essenfelder, A.; Toreti, A.; Seguini, L. Expert-driven explainable artificial intelligence models can detect multiple climate hazards relevant for agriculture. *Communications Earth & Environment* **2025**, *6*. <https://doi.org/10.1038/s43247-024-01987-3>.
156. Bhattacharya, A. *Applied Machine Learning Explainability Techniques: Make ML models explainable and trustworthy for practical applications using LIME, SHAP, and more*; Packt Publishing Ltd, 2022.
157. Guo, Y.; Gao, J.; Tunio, M.H.; Wang, L. Study on the Identification of Mildew Disease of Cuttings at the Base of Mulberry Cuttings by Aeroponics Rapid Propagation Based on a BP Neural Network. *Agronomy* **2022**, *13*, 106.
158. Kawakura, S.; Hirafuji, M.; Ninomiya, S.; Shibasaki, R. Adaptations of Explainable Artificial Intelligence (XAI) to Agricultural Data Models with ELI5, PDPbox, and Skater using Diverse Agricultural Worker Data. *European Journal of Artificial Intelligence* **2022**, *3*, 14.
159. Dara, R.; Hazrati Fard, S.M.; Kaur, J. Recommendations for ethical and responsible use of artificial intelligence in digital agriculture. *Frontiers in Artificial Intelligence* **2022**, *5*, 884192. <https://doi.org/10.3389/frai.2022.884192>.
160. Shen, Y.; Khan, Z.; Liu, H.; Yang, Z.; Hussain, I. YOLO Optimization for Small Object Detection: DyFAM, EFRAdaptiveBlock, and Bayesian Tuning in Precision Agriculture. *SSRN Electronic Journal*, 2025. <https://doi.org/10.2139/ssrn.5231695>.
161. Liu, H.; Zeng, X.; Shen, Y.; Xu, J.; Khan, Z. A Single-Stage Navigation Path Extraction Network for agricultural robots in orchards. *Computers and Electronics in Agriculture* **2025**, *229*, 109687.
162. Xiang, W.; Wu, D.; Wang, J. Enhancing stem localization in precision agriculture: A Two-Stage approach combining YOLOv5 with EffiStemNet. *Computers and Electronics in Agriculture* **2025**, *231*, 109914.
163. Coulibaly, S.; Kamsu-Foguem, B.; Kamissoko, D.; Traore, D. Deep learning for precision agriculture: A bibliometric analysis. *Intelligent Systems with Applications* **2022**, *16*, 200102.
164. Sun, X.; Wang, B.; Wang, Z.; Fu, K. Research Progress on Few-Shot Learning for Remote Sensing Image Interpretation. *Remote Sensing* **2021**, *13*, 678.
165. Chen, Z.; Feng, J.; Yang, Z.; Wang, Y.; Ren, M. YOLOv8-ACCW: Lightweight grape leaf disease detection method based on improved YOLOv8. *IEEE Access* **2024**.
166. Chen, J.W.; Lin, W.J.; Cheng, H.J.; Hung, C.L.; Lin, C.Y.; Chen, S.P. A smartphone-based application for scale pest detection using multiple-object detection methods. *Electronics* **2021**, *10*, 372.
167. Almasri, F.; Debeir, O. Multimodal Sensor Fusion in Single Thermal Image Super-Resolution. In *Proceedings of the Computer Vision – ACCV 2018 Workshops*. Springer, 2019, pp. 418–433. https://doi.org/10.1007/978-3-030-21074-8_34.
168. Padhiary, M.; Hoque, A.; Prasad, G.; Kumar, K.; Sahu, B. Precision Agriculture and AI-Driven Resource Optimization for Sustainable Land and Resource Management. In *Smart Water Technology for Sustainable Management in Modern Cities*; IGI Global, 2025; pp. 197–232.
169. Zhao, S.; Peng, Y.; Liu, J.; Wu, S. Tomato Leaf Disease Diagnosis Based on Improved Convolution Neural Network by Attention Module. *Agriculture* **2021**, *11*, 651.
170. Shen, Y.; Yang, Z.; Khan, Z.; Liu, H.; Chen, W.; Duan, S. Optimization of Improved YOLOv8 for Precision Tomato Leaf Disease Detection in Sustainable Agriculture. *Sensors* **2025**, *25*, 1398.
171. Dhanya, K.; Gopal, P.; Srinivasan, V. Deep learning in agriculture: challenges and future directions. *Artificial Intelligence in Agriculture* **2022**, *6*, 1–11.

172. Padhiary, M.; Hoque, A.; Prasad, G.; Kumar, K.; Sahu, B. The Convergence of Deep Learning, IoT, Sensors, and Farm Machinery in Agriculture. In *Designing Sustainable Internet of Things Solutions for Smart Industries*; IGI Global, 2025; pp. 109–142.
173. Zheng, W.; Cao, Y.; Tan, H. Secure sharing of industrial IoT data based on distributed trust management and trusted execution environments: a federated learning approach. *Neural Computing and Applications* **2023**, *35*, 21499–21509. <https://doi.org/10.1007/s00521-023-08663-9>.
174. Kumar, Y.; Kumar, P. Comparative study of YOLOv8 and YOLO-NAS for agriculture application. In *2024 11th International Conference on Signal Processing and Integrated Networks (SPIN)* (pp. 72-77) **2024**.
175. Padhiary, M.; Kumar, R. Enhancing Agriculture Through AI Vision and Machine Learning: The Evolution of Smart Farming. In *Advancements in Intelligent Process Automation*; IGI Global, 2025; pp. 295–324.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.