big data and cognitive computing

MDPI

Systematic Review

Deep Learning in Farming: A Systematic Evidence-Weighted Review of Applications, Validation Gaps, and Emerging Frontiers

Vito Domenico Amodio [1], Lerina Aversano [1], Vincenzo Dentamaro [2,*] and Felice Franchini [2]

[1] Department of Agricultural Sciences, Food, Natural Resources and Engineering, University of Foggia, 71122 Foggia, Italy; vito.amodio@unifg.it (V.D.A.); lerina.aversano@unifg.it (L.A.)
[2] Department of Computer Science, University of Bari Aldo Moro, 70125 Bari, Italy; f.franchini9@phd.uniba.it
* Correspondence: vincenzo.dentamaro@uniba.it

Abstract

This article presents a systematic, evidence-weighted review of Deep Learning (DL) in farming, with a primary emphasis on the agricultural production stage and on four operational domains: precision crop management, precision livestock farming, soil and water resource management, and autonomous agricultural systems. Following a PRISMA 2020-oriented protocol, 56 primary studies (42 with quantitative results) were retained from an initial pool of 18,731 records. The reviewed literature reports applications in plant disease detection, weed recognition, yield prediction, fruit detection, livestock identification and health monitoring, soil-property estimation, crop-water-stress assessment, and robotic perception. High performance on controlled datasets, however, is frequently reported without external, temporal, or cross-site validation, making practical generalisation difficult to establish: in one widely cited benchmark, disease-classification accuracy fell from above 99% on held-out laboratory images to 31.4% on field-acquired images of the same classes. Persistent weaknesses include the limited availability of public benchmarks, inconsistent validation protocols, limited interpretability, fragmented data governance, and insufficient techno-economic analysis. The review argues that the next stage of agricultural AI should be judged less by isolated benchmark scores and more by field realism, reproducibility, deployment maturity, and practical usefulness. Unlike broad surveys that mainly catalogue architectures and applications, this review interprets the literature according to dataset representativeness, validation protocols, benchmarking transparency, deployment realism, and reproducibility, distinguishing algorithmic performance under controlled conditions from practical readiness for real farming environments. The most promising research directions include self-supervised and multimodal learning, explainable and privacy-preserving AI, edge-aware deployment, and hybrid process-informed models.

Keywords: deep learning; systematic review; PRISMA; edge AI; explainable AI; federated learning; precision livestock farming; precision agriculture; smart farming

Check for updates

Academic Editor: Shuping He

Received: 2 July 2026
Revised: 5 August 2026
Accepted: 11 August 2026
Published: 14 August 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

1.1. The Global Imperative for Agricultural Innovation

Agriculture is a central pillar of the worldwide food supply, yet it faces mounting, compounding pressures: population growth, climatic instability, resource scarcity, labour constraints, and the need to lower environmental externalities. Recent international assessments stress that agrifood systems must become more productive, resilient, inclusive and sustainable, so as to meet food demand while cutting social and environmental costs [1–3].

Rural depopulation and an ageing farming workforce compound these pressures further, making labour shortages especially acute for critical, time-sensitive operations such as harvesting, disease inspection, animal monitoring, and irrigation management [3–5]. DL-based automation, decision support, and robotic perception respond directly to this combination of demands: raising productivity while sustaining farm operational capacity under conditions of reduced and ageing rural labour.

1.2. Agriculture 4.0 and Deep Learning as Its Foundational Technology

To address these challenges, agriculture is undergoing a technological transformation often described as Agriculture 4.0 or Smart Agriculture, characterised by the integration of IoT, robotics, big-data analytics, and AI across the agricultural value chain [6–8]. Large-scale sensor deployments across fields, machinery, drones, and satellites generate heterogeneous data streams, including imagery, soil measurements, moisture data, and weather records [9,10], with the goal not merely of automating existing practices but of using these streams to support more precise interventions, better resource allocation, higher productivity, and a lower environmental impact [6,7].

Deep Learning (DL), a branch of machine learning grounded in multi-layer neural networks, has become the main technology enabling this shift [6,11], owing to its capacity to learn hierarchical representations directly from large amounts of raw, often unstructured data [12,13], precisely the kind of heterogeneous, multi-scale data that cameras, drones, satellites, animal sensors, and soil probes produce [9,11]. Unlike classical machine-learning pipelines, which rely heavily on hand-crafted features, deep neural networks learn task-relevant representations from images, time series, and multimodal records with reduced manual feature engineering [14,15]. This capacity is particularly valuable in farming, where the same biological target appears differently across crops, cultivars, growth stages, lighting conditions, soil backgrounds, disease severities, and management regimes.

High benchmark performance should not, however, be read as automatic readiness for farm deployment: models trained on curated datasets routinely overfit to narrow imaging conditions or local management practices. For this reason, the present review treats DL not as a catalogue of architectures, but as a family of methods whose value rests on validation realism, transferability, interpretability, and operational feasibility, evaluated jointly with the agronomic and socio-technical conditions in which it is meant to operate.

1.3. Scope and Structure of the Review

This review provides a systematic and analytical synthesis of deep learning in agriculture, with an emphasis on literature published between January 2020 and September 2025. The term farming is used here to delimit the review to the production-oriented part of agriculture, including crop production, livestock production, soil and water management, and autonomous field or barn operations. Broader agri-food applications such as plant breeding, biotechnology, microbiology, seed technology, industrial food processing, post-harvest packaging, retail logistics, and traceability are acknowledged as adjacent areas but are not treated as the principal object of this review, so as to preserve analytical depth and avoid merging operational farming problems with biologically and commercially distinct stages of the agri-food chain. Its objectives are to identify the main areas of progress, highlight methodological and deployment-related weaknesses, and outline the most promising directions for future research. Section 2 presents the review methodology. Section 3 synthesises the literature across four application domains and highlights cross-domain methodological patterns. Section 4 discusses emergent technical directions that may help address current limitations. Section 5 summarises the main conclusions, and Section 6 provides a roadmap for future research and practice.

1.4. Positioning with Respect to Existing Surveys and Systematic Reviews

Multiple surveys have investigated the role of artificial intelligence and deep learning in agriculture, from varied conceptual frameworks. Ref. [6] produced one of the earliest structured syntheses of deep learning in agriculture, concentrating on convolutional architectures and vision tasks documented through approximately 2017. Ref. [11] addressed farm automation using both machine learning and deep learning, emphasising robotic and sensor-equipped infrastructure, while Ref. [7] reviewed machine-learning methods for precision agriculture more broadly, positioning deep learning as one algorithmic family among others. Ref. [16] focused specifically on deep learning and agricultural advancement, identifying constraints and potential developments. Adjacent contributions have examined complementary parts of the agri-food digitalisation workflow: Ref. [17] surveyed machine-learning applications spanning yield modelling, disease identification, and supply-chain logistics; Ref. [18] examined how machine learning and, more narrowly, deep learning support sustainable supply-chain performance; and specialised reviews have covered deep learning for dense visual scene interpretation [19], satellite and airborne image classification [9,20], and mask-based segmentation for farming imagery [10]. Recently, there have also been a number of survey papers that emerged: Ref. [21] provided a survey of 129 papers in five domains of DL application in agriculture but without the realism constraint of deployment; Ref. [22] conducted a systematic review of machine learning and computer vision in precision agriculture focusing on innovative methods; Ref. [23] employed a bibliometric and keyword network approach to deep learning in crop monitoring; and Ref. [24] provided a systematic review of deep learning for lettuce farming alone. The table below contrasts this review with the aforementioned reviews.

The present review departs from these earlier contributions along four lines: it centres on the 2020–September 2025 window, during which agricultural DL shifted from mostly CNN-based classification toward object detection, segmentation, transformer architectures, self-supervised learning, edge inference, federated learning, and hybrid PBM–DL systems; it follows a PRISMA 2020-oriented protocol with explicit search blocks, screening counts, and a common critical-appraisal framework (Section 2); it tightens the object of analysis onto *farming* specifically, rather than the whole agri-food chain; and it weights the evidence explicitly, reading reported model performance through dataset representativeness, validation protocol, benchmarking transparency, deployment realism, and reproducibility, so as to separate studies showing algorithmic promise from studies giving firmer evidence of field relevance.

1.5. Main Contributions of This Review

This review makes four main contributions. First, it applies a transparent review methodology that allows readers to assess the strength of evidence of individual studies rather than treating all reported results as equivalent. Second, it synthesises the literature across four agricultural domains using common comparative dimensions. Third, it explicitly addresses the laboratory-to-field gap together with benchmarking, socio-economic, and governance challenges. Fourth, it provides a near-term research agenda, explicitly ordered by the strength of evidence currently available for each direction (Section 5.2), focused on realistic benchmarks, multimodal and self-supervised learning, system-level evaluation, and techno-economic assessment.

1.6. Analytical Framework and Originality of the Review

The originality of this review does not rest on a claim that agricultural deep learning has remained unreviewed. Section 1.4 and Table 1 show clearly that it has not. Its contribution is rather to re-read the recent literature through an evidence-weighted and

deployment-oriented framework, interpreting the empirical studies discussed along five analytical dimensions. Data realism points to the degree to which the dataset mirrors field variability.Validation strength points to the presence of external, temporal or cross-site testing. Model–task suitability asks whether the chosen architecture fits the structure of the agricultural problem. Deployment readiness looks for evidence of latency, edge feasibility, robustness and system integration. Practical value weighs agronomic interpretability, economic feasibility and usefulness to farmers or agricultural professionals. These five dimensions are not merely descriptive: Section 2 operationalises them directly as the risk-of-bias rubric in Table 2, so the same criteria introduced here are the ones actually applied to every included study. Through this framework, the review steps past a descriptive list of applications, separating studies showing narrow benchmark success from those offering sturdier evidence of operational relevance, and judging the field's recent growth in CNN-based vision, object detection, segmentation, transformers, self-supervised learning, federated learning, edge AI, and process-informed hybrid models against one practical question: are the reported systems reproducible, robust, interpretable, and economically plausible inside real farming environments?

Table 1. Positioning of the present review relative to representative prior surveys of deep learning in agriculture.

Prior Survey	Coverage Window/Scope	Explicit Search Protocol	Key Difference from This Review
Kamilaris & Prenafeta-Boldú [6]	To ca. 2017; CNN-based vision tasks across agriculture	Narrative, not PRISMA-based	Foundational but pre-dates transformers, SSL, edge/federated learning, and hybrid PBM–DL
Sharma et al. [7]	Precision agriculture; ML broadly, DL as one family	Narrative	Not DL-specific; no evidence-weighting or risk-of-bias appraisal
Albahar [16]	DL in agriculture, constraints and developments	Narrative	No explicit screening counts; whole-agriculture rather than farming-stage scope
Attri et al. [21]	129 papers, five DL application areas	Narrative	Catalogues architectures/tasks; no validation-strength or deployment-realism weighting
Kuan et al. [22]	ML/CV in precision agriculture, 2025	Systematic	Broader ML/CV scope beyond DL; not farming-stage restricted; no risk-of-bias rubric
Zhang et al. [23]	DL in crop monitoring, 2000–2024	Bibliometric	Quantifies publication/keyword trends rather than appraising individual studies' evidence strength
Qin et al. [24]	DL for lettuce cultivation specifically	Systematic	Single-crop scope; not cross-domain (livestock, soil/water, autonomy)
This review	2020–September 2025; farming stage across four domains	PRISMA 2020-oriented, screening counts reported	Explicit five-dimension evidence weighting (Table 2) linking benchmark scores to deployment realism

Table 2. Risk-of-bias/critical-appraisal rubric applied to the 56 included studies, operationalising the five analytical dimensions of Section 1.6. Authors' synthesis.

Criterion	Low Risk	Some Concerns	High Risk
Dataset representativeness	Multi-site or field-realistic data with explicit environmental variability	Single-site data with moderate variation or partially described sampling	Narrow laboratory dataset with limited description of sampling context
Validation protocol	External test set, temporal hold-out, or cross-site validation clearly reported	Internal split or *k*-fold validation reported but no external validation	Re-used training data, unclear split, or no robust validation protocol
Benchmarking transparency	Strong baseline comparisons, task-appropriate metrics, and variance reporting	Limited baselines or incomplete uncertainty reporting	Selective metrics, no meaningful baseline, or incomplete experimental details
Deployment realism	Pilot or field evaluation with operational constraints documented	Prototype deployment with partial latency or hardware information	Laboratory-only claims presented without deployment evidence
Reproducibility	Public dataset/code or sufficiently detailed acquisition and preprocessing description	Partial description of data and preprocessing pipeline	Proprietary data and insufficient methodological detail for replication

2. Materials and Methods

A systematic literature search, structured around a PRISMA 2020-oriented protocol [25], was used to identify recent empirical studies and relevant contextual reviews concerning deep learning in farming. Searches covered Scopus, Web of Science, and IEEE Xplore; Google Scholar supported backward and forward citation searching. The principal publication window was 1 January 2020 to 30 September 2025. Earlier studies were consulted only when they provided foundational technical or historical context and were not treated as part of the recent empirical evidence base. Farming and agricultural-production terms were combined with descriptors of deep learning methods. Production-stage terms such as "sowing", "harvesting", "robotic harvesting", "poultry", "precision livestock farming", and "animal welfare" were included to reduce vocabulary bias. A representative Scopus query is:

> TITLE ("deep learning" OR "CNN" OR "RNN" OR "LSTM" OR "Transformer" OR "GNN" OR "self-supervised learning" OR "neural network") AND TITLE-ABS-KEY ("agriculture" OR "crop" OR "farm" OR "livestock" OR "irrigation" OR "phenotyping" OR "robotics") AND PUBYEAR > 2019 AND PUBYEAR < 2026 AND (LIMIT-TO (SUBJAREA, "COMP") OR LIMIT-TO (SUBJAREA, "ENGI") OR LIMIT-TO (SUBJAREA, "AGRI")) AND (LIMIT-TO (DOCTYPE, "ar") OR LIMIT-TO (DOCTYPE, "cp"))

The query was adapted to the syntax and indexing fields of each database while preserving three conceptual blocks: (i) deep-learning terminology, (ii) farming and agricultural-production terminology, and (iii) task-specific terminology. Google Scholar was not used as a primary database because its ranking and coverage are difficult to reproduce; it was instead used only for backward and forward snowballing from already-eligible papers. The final search update was performed in September 2025, and online-first material was considered only when its peer-reviewed publication status could be verified by that date. The completed PRISMA 2020 Checklist is provided in the Supplementary Materials. We intentionally opted for using traditional bibliographic databases rather than a scientific

search or summarisation based on LLMs (Elicit, Consensus, and others). There were three main reasons behind our choice. First, reproducibility: Boolean search in Scopus, Web of Science, and IEEE Xplore can be repeated independently by any other researcher and will produce consistent results, while an LLM search is highly dependent on specific wording of prompts, model version used, and heuristic ranking/summarisation algorithm of the tool itself, which is all non-reproducible. Second, transparent coverage: bibliographic databases provide an explicit count of included/excluded articles at each screening step (Section 2.2), and that is what provides audibility of the PRISMA flowchart in Figure 1; no general-purpose search tool powered by LLM can provide the same kind of transparent accounting of the search process. Third, reliability: LLM-driven summarisation of scientific text comes with known risk of generating false citations or misattribution of findings - exactly the kind of mistake that a critical evidence synthesis should avoid. However, these tools are a truly valuable addition to the traditional screening process, and we will further discuss them, as well as the application of large language models in agricultural decision-making in Sections 4.4 and 6.

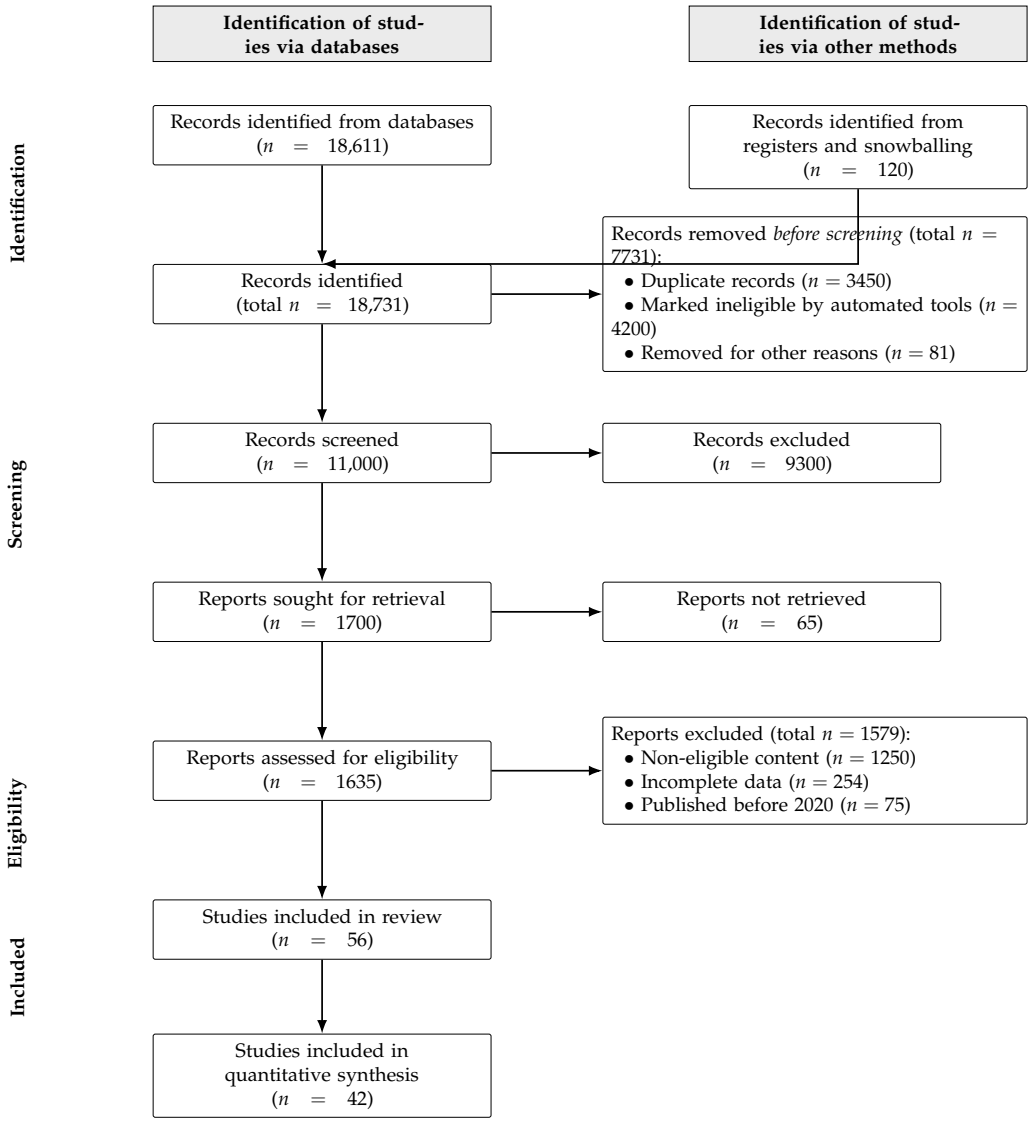


Figure 1. PRISMA 2020 flow diagram summarising the identification, screening, eligibility assessment, and inclusion of studies (2020–2025) [25]. Diagram authored natively in TikZ by the authors from the underlying screening counts, so that it can be revised as plain text rather than as a raster image.

Candidate publications were checked by title, abstract and, where necessary, full text. Bibliographic metadata were verified against the cited publication rather than inferred from citation-key names. For the evidence-weighted synthesis, the extraction fields were application domain, task, sensing modality, dataset type, architecture family, comparison models, evaluation metrics, validation protocol, code or data availability, deployment status, and stated limitations.

2.1. Inclusion and Exclusion Criteria

The core empirical evidence comprised English-language, peer-reviewed journal articles and indexed conference papers published from 2020 through September 2025 that presented and evaluated a deep-learning model for a production-oriented agricultural task. Eligible tasks included crop disease and weed detection, yield or fruit estimation, livestock identification, health and welfare monitoring, soil-property or crop-water assessment, and robotic perception or control. Empirical papers had to specify a task, a dataset or acquisition setting, a model family, and quantitative or operational performance evidence.

Review articles, reports and studies published before 2020 were used only for contextual framing and were not represented as recent empirical evidence. Papers were not used as task-level evidence when they evaluated only non-deep-learning methods, concerned a non-agricultural application, or addressed supply-chain security, industrial processing, breeding or traceability without a direct connection to farming-stage sensing, prediction or action. This distinction is important because contextual sources and primary empirical studies serve different purposes in a critical review. The reasons for full-text exclusions were categorised into three categories, as depicted in Figure 1. Non-eligible content ($n = 1250$) includes records that did not actually involve evaluating a deep-learning model upon further checking of the full text (such as opinion pieces, theoretical works only, or classical machine-learning-based papers classified due to erroneous search terms), those that discussed topics not related to agriculture even though they matched the query, or those that were about breeding, processing, or traceability but did not have any farming stage sensing/prediction/action application. Incomplete data ($n = 254$) includes records where the dataset information, model information, or results were not sufficiently available to be extracted. Publication date ($n = 75$) includes records that, although listed within the search period, were preprints, corrections, and errata relating to a major publication outside the 1 January 2020 to 30 September 2025 range.

This review has several limitations that should be considered when interpreting the findings. Its PRISMA-oriented search and eligibility rules improve reproducibility relative to an unstructured narrative review, but they do not guarantee exhaustive coverage of every relevant study, and no formal systematic-review registration (e.g., PROSPERO) was filed, since PROSPERO does not currently accept non-health protocols of this kind. The English-language emphasis may omit regional literature, while the preference for studies with sufficient methodological detail may under-represent commercial systems whose data and validation protocols are not publicly disclosed. This is a form of selection bias that is worth making explicit: proprietary systems operated by large agritechnology firms may already be deployed at a scale and reliability that our screening criteria cannot detect, because such deployments are rarely described in indexed, peer-reviewed venues. The reviewed evidence base should, therefore, be read as representative of the published, peer-reviewed state of the art rather than of the full population of deployed agricultural AI systems, and the gap between the two is itself informative about the field's publication and disclosure incentives. The literature is also highly heterogeneous in task definition, sensing modality, dataset structure, metric and validation protocol; consequently, reported results were not statistically pooled. These limitations motivate a qualitative, evidence-weighted

interpretation rather than numerical claims about the size or distribution of a definitive study corpus.

2.2. Study Selection, Screening, and Risk-of-Bias Appraisal

The first database search was conducted on 15 September 2025 and yielded 18,611 references; in addition, another 120 references were found using backward and forward snowballing and citation registers, resulting in an initial set of 18,731 references. Of those, 7731 references were filtered prior to the screening stage ($n = 3450$ duplicates or quasi-duplicates; $n = 4200$ flagged as ineligible based on automated deduplication and language/document-type filters; $n = 81$ filtered out due to other administrative reasons). Titles and abstracts of the remaining 11,000 items were independently screened by two reviewers against the eligibility criteria presented in Section 2.1, and any disagreements were resolved by discussion and, when necessary, by consultation with a third author. As a result of this stage, 9300 references were excluded as being off-scope. We obtained full reports for 1700 references; 65 references had no available full text (either paywalled without institutional access or withdrawn). Out of the remaining 1635 full texts, we assessed the eligibility using a common coding sheet that documented bibliographic metadata, application domain, task, sensory modality, type of dataset, architecture family, comparison models, metrics used, validation approach, availability of code or data, deployment status, and limitations; 1579 reports have been excluded (Section 2.1), while 56 primary studies were deemed eligible for inclusion, 42 of which presented results suitable for a descriptive synthesis. A summary of the screening process can be seen in the PRISMA 2020 flow diagram in Figure 1. Due to the heterogeneity of tasks, datasets, metrics, and validation design in the included studies, their results could only be compared descriptively and not pooled statistically, as suggested by current PRISMA guidelines for heterogeneous bodies of evidence [25].

Risk of bias in the included studies was appraised qualitatively using the criteria reported in Table 2. Each of the five criteria in the rubric was taken directly from the five analytical dimensions presented in Section 1.6 (data realism, validation robustness, task-model fit, implementation preparedness, and practicality), thereby operationalising, but not duplicating, that framework. In order to ensure the consistency of using the rubric, both reviewers first independently evaluated a common pilot set of eight studies across all four domains according to each criterion, reconciled their differences for each criterion, and then evaluated the rest of the corpus, resolving any remaining differences regarding the whole corpus through a third reviewer. These evaluations are thus better viewed as a structured qualitative evaluation than a validated inter-rater reliability measurement.

2.3. Evidence-Weighted Synthesis Methodology

The synthesis gives priority to the 56 included empirical papers that satisfy the scope and reporting criteria above. Contextual reviews and foundational studies are cited where they explain the development of an architecture or application area, but they are kept conceptually separate from recent task-level evidence. Quantitative results are compared descriptively because the tasks, datasets, metrics and validation designs are too heterogeneous for statistical pooling. In particular, a high score obtained from an internal split of a curated dataset is not treated as equivalent to performance demonstrated through temporal, cross-site or external validation. For each selected study, the first step in the review process consists of noting down the raw performance metric reported by the authors followed by classifying that performance metric in accordance with the five-dimensional framework shown in Table 2. The relative weight given to the evidence of a particular study in the domain-independent synthesis in Sections 2.4 and 3 will be based upon the latter

classification rather than the former one. This double-stage process enables the review to make a distinction between the studies, having the same level of accuracy, depending upon the validity criteria used in those studies.

2.4. Overview of the Evidence

Each of the 56 articles reviewed herein was tied to only one particular primary field of application. The field of precision crop management includes 28 articles (disease and pest detection, weed detection, yield estimation, and fruit detection or counting). The field of precision livestock farming includes 10 articles. The field of soil and water resources management includes 8 articles. The field of autonomous agriculture, which includes federated and edge learning, includes the other 10 articles. The literature discussed in this review is, therefore, dominated by crop-oriented computer vision, particularly disease detection, weed recognition, yield estimation, and fruit detection or quality assessment, while precision livestock farming, soil and water management, and autonomous agricultural systems form smaller and methodologically more heterogeneous bodies of evidence. This imbalance is itself a finding rather than an artefact to be hidden: it mirrors the relative maturity of public benchmarks and vision datasets across domains (Section 3.7), and it directly motivates the cross-domain synthesis below, which weighs each domain's evidence on its own terms rather than by study count alone.

The evidence-weighted reading changes the interpretation of this distribution. Crop-management studies are the most numerous and often report the highest benchmark scores, but many still rely on curated imagery, single-season datasets, or internal validation. Precision livestock farming shows strong practical relevance because welfare, identification, and health problems are directly linked to farm management, yet the evidence is frequently constrained by private barn-specific datasets and limited cross-farm testing. Soil and water studies are fewer, but they often address agronomically consequential variables and, therefore, require stronger uncertainty reporting than is currently common. Autonomous agricultural systems provide the clearest link between DL and operational action, but they also face the strictest deployment requirements because perception errors can affect machinery safety, crop damage, and labour workflows. Table 3 summarises this cross-domain interpretation.

Table 3. Evidence-weighted interpretation of the reviewed literature by operational domain. Authors' synthesis.

Domain	Evidence Strength	Main Limitation	Interpretative Consequence
Precision crop management	Largest corpus; mature vision benchmarks; strong reported accuracy	Frequent dependence on curated images, internal splits, and narrow crop–disease settings	Strong algorithmic promise, but field generalisation should be treated cautiously unless external or cross-site validation is reported
Precision livestock farming	High practical relevance for identification, welfare, and behaviour monitoring	Private datasets, barn-specific conditions, and limited public reproducibility	Promising for farm management, but cross-farm transfer and data governance remain decisive barriers
Soil and water management	Agronomically important variables and increasing multimodal modelling	Small number of studies, heterogeneous metrics, and incomplete uncertainty reporting	Useful for decision support only when prediction intervals, sensor context, and spatial transfer are reported
Autonomous systems	Direct operational relevance for perception, navigation, harvesting, and edge inference	Laboratory or prototype evaluation often substitutes for long-duration field trials	Deployment readiness requires latency, safety, robustness, and end-to-end system metrics, not only detection accuracy

Figure 2 complements this table by showing which model families occur across multiple operational domains in the included corpus. This network was generated directly from the authors' extracted study-level data (Table 4) and drawn natively in TikZ; it is an architecture co-occurrence summary rather than a software-generated bibliometric map. In addition to that, we produced a true bibliometric map using software, where the lists of all authors of papers listed in our bibliography were exported as a co-authorship graph for use in VOSviewer Online, version 1.2.4 (Centre for Science and Technology Studies, Leiden University, Leiden, The Netherlands; accessed 12 August 2026; Figure 3).

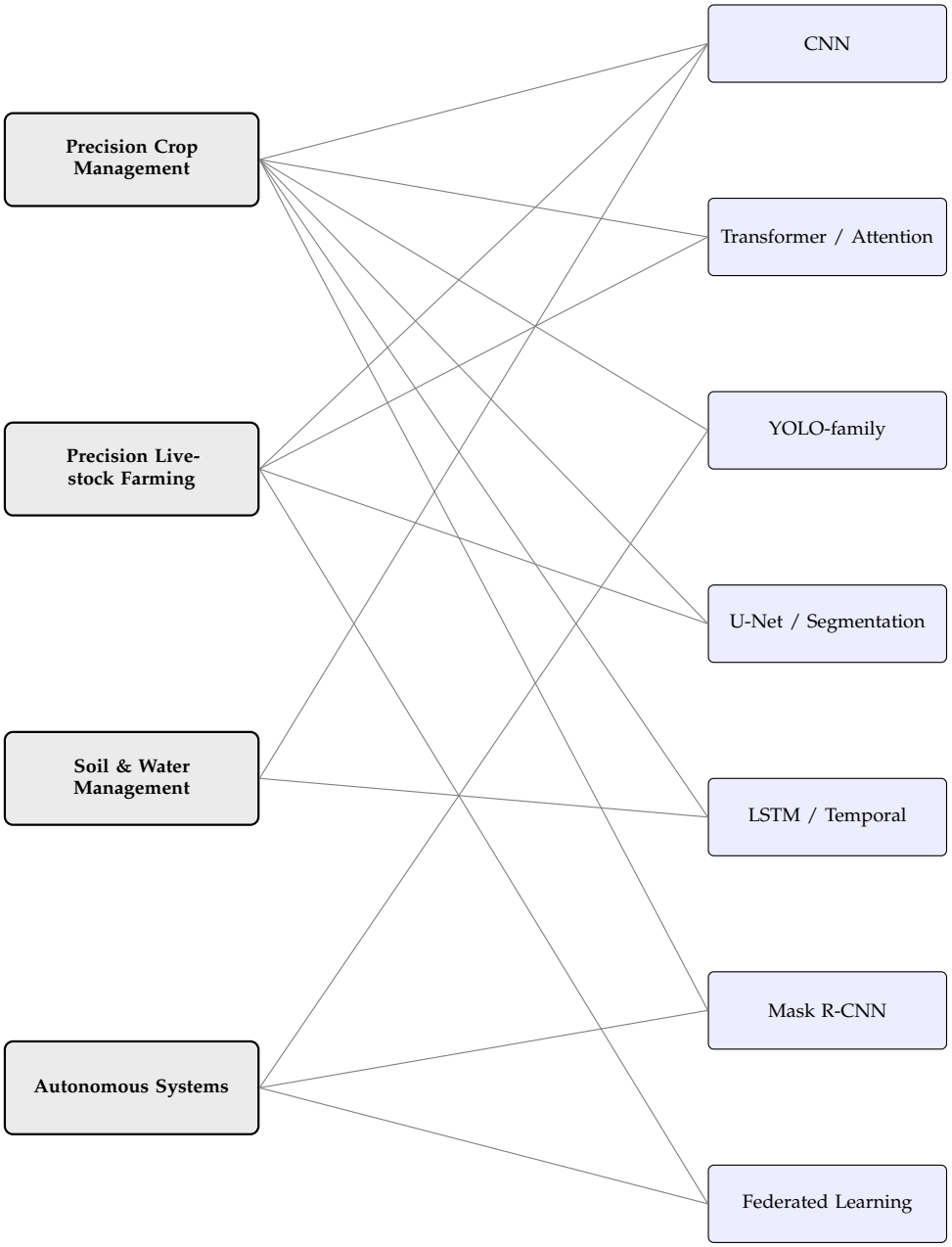


Figure 2. Domain–architecture co-occurrence network for the 56 included studies, summarising which representative DL architecture families (Table 4) recur across more than one operational domain. CNN and Transformer/Attention architectures are the most cross-cutting; Federated Learning links precision livestock farming and autonomous systems specifically, reflecting the shared distributed-deployment infrastructure discussed in Sections 3.3.3 and 3.5.2. Authors' synthesis, produced natively in TikZ from Table 4; not software-generated bibliometric output (see main text).

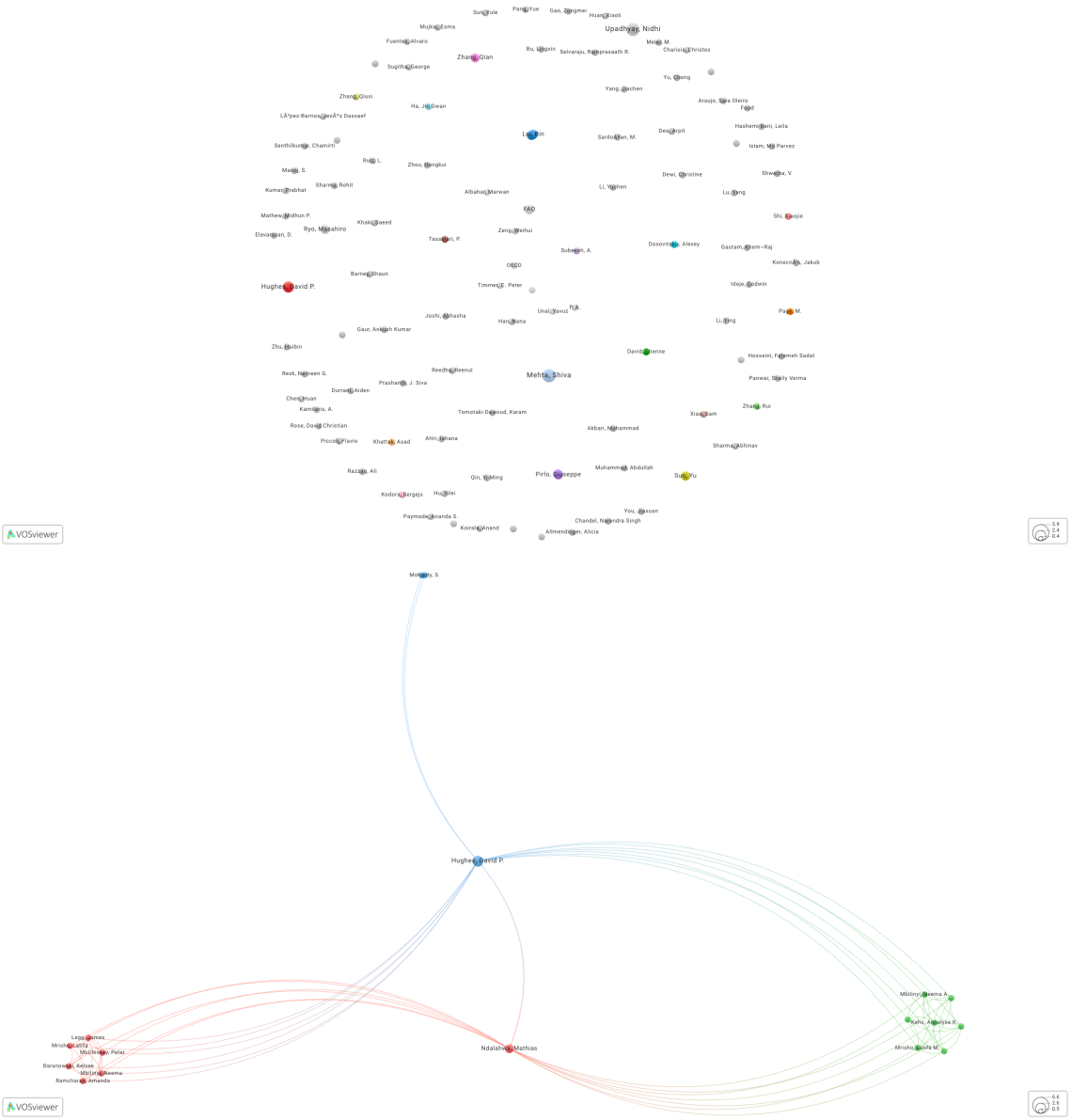


Figure 3. VOSviewer co-authorship network of the reference list underlying this review (network visualisation; node size reflects number of documents by that author within the corpus, colour reflects cluster membership). (**Top**): full network. Most authors form small, mutually disconnected clusters, reflecting a reference list drawn from many independent research groups rather than one tight sub-field, itself an informative bibliometric signal about the breadth of the reviewed evidence base. (**Bottom**): the single largest connected component, centred on Hughes, D.P. This cluster is not a bibliometric curiosity: it links the three studies that anchor the laboratory-to-field evidence discussed in Section 3.2.1: the original PlantVillage benchmark ([26], 99.35% held-out accuracy falling to 31.4% on field-like images), the mobile cassava-diagnosis system [27], and the PlantVillage Nuru field validation reporting 65–88% field accuracy [28]. Together, these show that the laboratory-to-field contrast we draw on throughout Section 3 is substantiated by a coherent, traceable body of work from a single research programme rather than by disconnected, incommensurable studies. Generated in VOSviewer from a network file the authors constructed from the manuscript's own bibliography.

Table 4. Comprehensive summary of representative deep-learning architectures reported across the 56 studies included in the review, by domain and primary task. "Representative DL Architectures" lists the principal model families used by the cited studies for that task, not an exhaustive enumeration; "Transformer-based" denotes studies using a Vision Transformer (ViT) backbone or self-/cross-attention modules, including attention blocks embedded within hybrid CNN detectors.

Domain	Primary Task	Representative DL Architectures	References
Precision Crop Management	Disease and pest detection	CNN (ResNet, ResNeXt), hybrid CNN–CRF, transformer-based (DSC-TransNet)	[26,29–36]
	Weed identification	U-Net (3SW-Net), YOLO-family detectors, Vision Transformer (ViT)	[37–41]
	Yield prediction	Deep transfer learning, deep Gaussian process, deep reinforcement learning	[42–46]
	Fruit counting, detection, and quality	Mask R-CNN, YOLO variants (e.g., MangoYOLO)	[47–51]
Precision Livestock Farming	Identification and tracking	Attentional feature-fusion CNN, ResNet-based re-identification	[52–54]
	Behaviour and welfare monitoring (video)	3D-CNN/ConvLSTM, pose estimation, spatiotemporal transformer	[55–58]
	Health monitoring and disease detection	U-Net (thermal segmentation), CNN with attention (CBAM), transfer-learning CNN	[59–63]
Soil and Water Management	Soil property estimation	1D/2D CNN, hybrid CNN–RNN	[64–67]
	Precision irrigation and water-stress detection	CNN classifiers (AlexNet, GoogLeNet, Inception), season-specific CNN	[68–70]
Autonomous Agricultural Systems	Perception, navigation, and manipulation	YOLOX/YOLOv8, Mask R-CNN, vision-guided motion planning	[39,71–74]
	Distributed coordination (federated/edge AI)	Federated CNN, privacy-encoding FL, edge scheduling	[75–83]

3. State of the Art: Deep Learning Applications in Agriculture (2020–2025)

The reviewed evidence demonstrates a significant expansion of deep learning (DL) beyond traditional image classification into spatiotemporal forecasting, dense scene segmentation, and distributed robotic control. This section provides a thematic synthesis of these applications, organised by operational domain, and supported by a comparative analysis of the underlying architecture families. Table 4 first summarises, domain by domain, the representative deep-learning architectures reported across the 56 included studies; Figure 4 then presents the corresponding taxonomy, with each branch and leaf labelled by the section number in which it is discussed below, so that the figure can be read as a visual table of contents for this section rather than as a free-standing diagram.

3.1. Comparative Reading of Architecture Families

The architectural landscape is dominated by Convolutional Neural Networks (CNNs), whose early agricultural applications established the basis for many current vision pipelines [12,26,84,85]. The 2020–2025 period also marks the growing use of Vision Transformers (ViT) [86], applied for example to UAV-based weed and crop classification [40], and of hybrid attention mechanisms embedded within convolutional backbones, such as self-attention [87] and CBAM [61,88] modules, for complex field scenes [32,33]. We use "Transformer-based" throughout this review to denote either a full ViT backbone or an attention module embedded within an otherwise convolutional detector; where a study reports a specific detector version (e.g., YOLOv8 [89]), we cite that version explicitly rather than the generic "YOLO" family name, following Table 4. For temporal or multimodal data such

as yield forecasting, recurrent and hybrid neural networks remain important [42], although comparisons across studies are limited by different data partitions and validation settings.

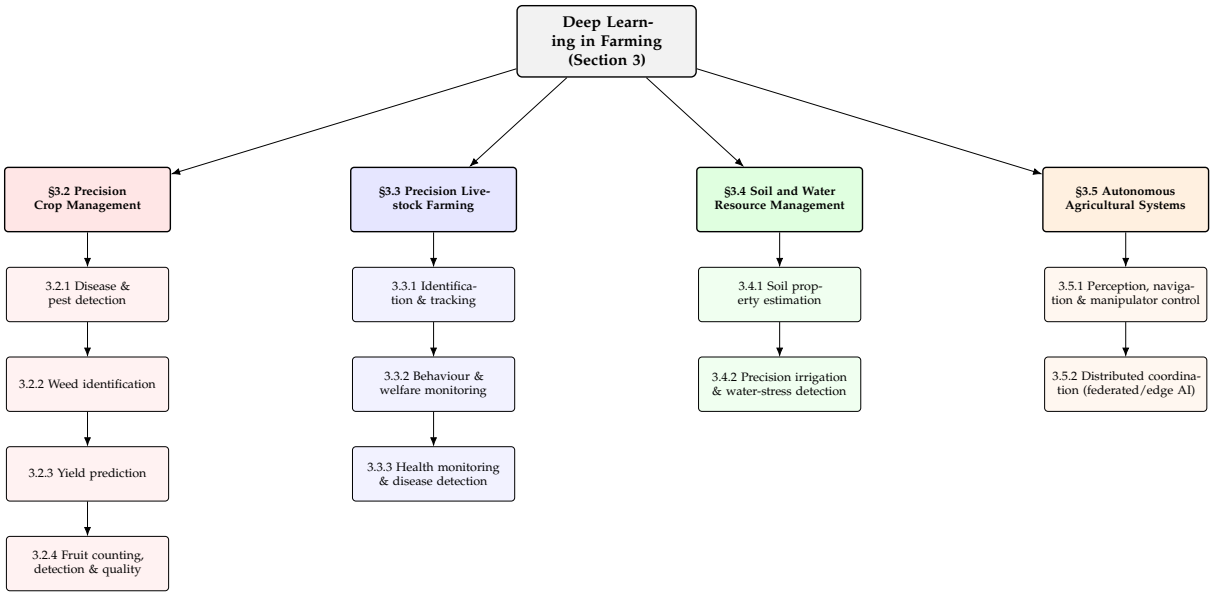


Figure 4. Taxonomy of uses of deep learning in agriculture (2020–2025). The branches and leaves are annotated with the numbers of the sections in which they are covered, such that the taxonomy reflects the structure of Section 3 and not an independent categorisation scheme. The first three categories (crop, livestock, soil/water) are categorised according to agronomic goal; Autonomous Agricultural Systems, however, are categorised according to operational mode—the closed perception-to-control loop with physical effects (navigation, manipulation, coordination of devices), since this operational mode introduces safety, latency, and end-to-end evaluation requirements that cut across the other three domains rather than belonging exclusively to any one of them (see Section 3.5). Cross-branch dependencies (e.g., the use of crop-management vision algorithms for perception by a harvesting robot) are discussed qualitatively in the text rather than encoded as additional edges in the diagram. Authors' original figure, created with native TikZ.

3.2. Precision Crop Management

Crop management represents the most mature domain of DL application in the reviewed literature, especially in plant pathology, weed control, and yield estimation.

3.2.1. Plant Disease and Pest Detection

DL has enabled the real-time, automated diagnosis of plant stressors using high-resolution leaf imagery. Models based on ResNet and ResNeXt have achieved classification accuracies exceeding 95% on public datasets like PlantVillage [26,31–35], and YOLO-family detectors, one of the most established architectures for object detection, have likewise been adapted specifically for plant disease detection: an optimised, lightweight YOLOv5 model achieved real-time detection and the classification of multiple crop diseases while remaining deployable on resource-constrained hardware [90]. Recent advances include the use of self-supervised learning to improve detection in low-resource settings [36,91] and the deployment of lightweight models on mobile devices for field use [29,33]. However, the controlled benchmark scores must also be interpreted in relation to their field counterparts: in the original PlantVillage study, the same CNN that achieved 99.35% accuracy on a held-out portion of the curated benchmark data achieved 31.4% accuracy on disease images that had been acquired in an ordinary online/field-like fashion, as opposed to a controlled acquisition procedure. The reason for the drop in performance lies in a domain shift problem that recurs in several studies: while curated benchmark data tends to consist of single leaf, well-lit images on a uniform background acquired via a standardised protocol,

field data is characterised by non-uniform backgrounds, overlapping leaves, non-standard lighting and shadows, dirt and sensor noise, as well as stressors that occur simultaneously (nutrient deficiency, multiple pathogens). This means that a classifier could inadvertently use the background and acquisition procedure cues and fail to generalise beyond them. This particular type of failure does not happen universally, however: several smartphone systems trained and evaluated on field-acquired data have reported significantly better field accuracy rates (65–88%), depending on the number of leaves per plant [27,28]. The point is that the gap between laboratory and field performance is mostly a function of the realism of the training data and evaluation methodology and not the inherent inability of DL as a technique, which explains why both the dataset and validation are considered separately in Table 2.

3.2.2. Weed Identification and Management

Precise weed identification is critical for site-specific herbicide application. Studies in this period have transitioned from binary crop/weed classification to the semantic segmentation of multiple weed species in dense canopy environments, using both YOLO-family and transformer-based object detectors [37,38]. Vision Transformers have also been applied directly to UAV-acquired weed and crop imagery, outperforming convolutional baselines such as ResNet and EfficientNet in some settings [40], while convolutional pipelines remain competitive for controlled-environment weed detection, for example in polyhouse-grown bell peppers [41]. Dense prediction and object-detection architectures provide spatial outputs that can support targeted laser or mechanical weeding [39].

3.2.3. Crop Yield Prediction and Forecasting

Yield prediction models have evolved from simple regression to spatiotemporal frameworks combining satellite imagery, weather records, and soil data [42], including deep Gaussian process formulations that provide calibrated uncertainty alongside point forecasts [44]. Deep Reinforcement Learning has also been explored for optimising yield-related decisions under varying environmental constraints [43]. More recent work combines UAV-based multispectral imagery with deep learning for field-scale yield prediction in breeding trials [45] and applies deep-transfer-learning strategies to satellite image time series so that models trained in data-rich regions can be adapted to regions with limited historical labels [46]. Multi-institution field campaigns such as the Global Wheat Head Detection dataset, assembling over 190,000 labelled wheat heads from field-acquired imagery across seven countries and acquisition platforms, illustrate that large, field-realistic training resources are achievable when data collection is explicitly coordinated across sites [92], and they directly counter the view that agricultural DL training data is necessarily laboratory-bound.

3.2.4. Fruit Counting, Detection, and Quality Assessment

Automated harvesting and grading rely on robust fruit detection under occlusion and varying lighting. Mask R-CNN and related detectors have been evaluated for sweet-pepper fruit and peduncle detection, while disease-oriented networks have also analysed fruit and leaf symptoms [47,48]. Crop-specific YOLO variants have also been benchmarked directly in the field: for example, 'MangoYOLO' was developed and validated for real-time fruit detection and orchard load estimation on in-field imagery [49], and hyperspectral imaging combined with deep learning has been used for real-time, in-field strawberry ripeness estimation [51], alongside multi-stage networks for maturity staging in wild blueberry fields [50]. These results are promising, but they should not be generalised beyond the crop, acquisition system and environmental conditions represented in each dataset.

Table 5 summarises the representative precision-crop-management applications discussed above.

Table 5. Summary of representative deep learning applications in precision crop management (2020–2025). "Representative DL Architectures" names the principal model family reported by each cited study; version numbers (e.g., YOLOX, YOLOv8) are given only where the corresponding study explicitly reports that version. Authors' synthesis.

Sub-Domain	Primary Task	Representative DL Architectures	References
Disease	Detection/Classif.	CNN (ResNet, ResNeXt), YOLOv5, hybrid CNN–CRF	[29–32,90]
Weeds	Segmentation	U-Net (3SW-Net), YOLOX, Transformer/ViT	[37–40]
Yield	Regression/Forecast	Deep transfer learning, deep Gaussian process, DRL	[42–44]
Fruit	Detection/Quality	Mask R-CNN, YOLO (e.g., MangoYOLO)	[47–49]

3.3. Precision Livestock Farming (PLF)

DL in PLF focuses on individual animal identification, behaviour and welfare monitoring, and health and disease detection to support large-herd management.

3.3.1. Individual Animal Identification and Tracking

Non-invasive identification using facial recognition or coat patterns has become feasible through convolutional and attentional feature-fusion architectures [52,53]. Computer-vision systems have also been evaluated for detecting dairy cows in free-stall barns [54]. These studies demonstrate task feasibility, although cross-farm transfer remains difficult to assess when data are collected in a single herd or barn.

3.3.2. Behaviour Monitoring and Welfare Analysis

Behaviour and welfare monitoring require temporal evidence rather than isolated-image accuracy. Recent works, starting from 2020 onwards, deal with the issue explicitly in the context of video-based models, rather than single frame-based methods. 3D convolution, along with ConvLSTM-based spatiotemporal networks, have been used to identify the behaviour of cows like feeding, exploring, grooming, and walking from continuous video clips [55], and action recognition pipelines involving multiple cameras have been validated to monitor the behaviour of individual cattle in closed barns [12]. Another increasingly relevant area is video-based multi-animal pose estimation, in which anatomical keypoints are detected and tracked for multiple cows from videos followed by gait/posture classification has been used for automated detection of lameness with an accuracy range of 94–100% across cows of varying levels of lameness [56]. For beef cattle, an improved YOLOv8-based CBR-YOLO model has been used for behaviour recognition under multiple weather conditions [58]; for poultry, a spatiotemporal transformer architecture has been proposed for recognising and localising chicken behaviour within continuous video segments [57]. Public, multi-farm video datasets nonetheless remain scarce [17], so claims about longitudinal welfare monitoring across farms should still be distinguished from the more established evidence for single-site animal identification and detection.

3.3.3. Health Monitoring and Disease Detection

Another important subdomain of application where deep learning methods are increasingly being employed is livestock and poultry health, which can be distinguished from the monitoring of livestock behaviour/welfare since it aims at identifying disease status (presence or severity of disease) rather than some kind of behaviour or management characteristic. A recent PRISMA-based systematic review identified 39 papers with application of deep learning in veterinary diagnostic imaging and animal health screening. It found that the performance of deep learning models is good when working on retrospectively collected clinical data, while prospective validation is still lacking [63]. Some examples of application of deep learning models for health outcomes are: U-Net based segmentation for automated detection of dairy cow mastitis using thermal infrared images of udders [59],

comparing different pretrained CNN backbones for early image-based detection of lumpy skin disease in cattle [60], and YOLOv5 object detector enhanced by CBAM attention module for real-time detection of broilers on litter floor as a prerequisite to flock-level health and welfare monitoring [61]. Another systematic review dedicated to machine learning and deep learning techniques for health monitoring and disease management of poultry also finds the results promising, but mostly based on retrospective studies of single facilities [62]. Environmental control studies, i.e., machine learning-based optimisation of thermal conditions in ventilated animal buildings, can be considered a related topic since it focuses on animal housing conditions and, however, is agronomically connected to health outcomes through exposure to heat stress [93]. Just like in crop disease detection (Section 3.2.1), the main question for this subdomain is not whether deep learning methods are capable of detecting disease signals, but whether their accuracies generalise across farms, breeds, camera placement and disease prevalence.

Table 6 summarises the representative precision-livestock-farming applications discussed above.

Table 6. Summary of representative deep learning applications in precision livestock farming. Authors' synthesis.

Task	Data Modality	DL Architecture	References
Identification & detection	RGB images (face/coat/barn)	Attentional feature fusion, CNN	[52–54]
Behaviour & welfare	Video/image sequences	3D-CNN/ConvLSTM, pose estimation, spatiotemporal transformer	[55–58]
Health & disease	Thermal/RGB imagery	U-Net, CNN with CBAM attention, transfer-learning CNN	[59–61]

3.4. Resource Management: Soil and Water

DL applications in resource management aim to optimise input use and environmental sustainability through precise soil property mapping and irrigation scheduling.

3.4.1. Soil Property Estimation and Health Monitoring

Recent soil-property studies use deep neural networks to process laboratory spectra, simulated hyperspectral-satellite measurements, optical imagery, topographic variables and climatic data. A scalable convolutional architecture evaluated on the LUCAS spectral library estimated several soil properties jointly, while a hybrid CNN–RNN study used satellite-derived and environmental predictors for the spatial mapping of physical and chemical properties [64,65]. Digital soil mapping with machine-learning algorithms has also been evaluated under genuinely heterogeneous field topography, including a tropical mountainous region where terrain complexity itself is a major source of prediction error [66], and data-driven soil analysis pipelines have been proposed specifically for smart-farming decision support [67]. Both studies also illustrate an important limitation: strong performance on a spectral library or a single study region does not establish transferability to soils acquired with different sensors or under different field conditions.

3.4.2. Precision Irrigation and Water Stress Detection

Image-based deep networks have been evaluated for crop-water-stress classification in maize, okra, soybean and pot-cultivated peach plants [68,69]. Season-specific CNN models combined with thermal vegetation-dryness indices have similarly been used for soil-moisture and irrigation monitoring at a district scale [70]. These studies demonstrate that visible symptoms can support stress classification, but they do not by themselves establish irrigation scheduling performance. Operational irrigation support requires connection to

water-balance variables, temporal forecasts, decision thresholds and field trials measuring water use and crop response.

Table 7 summarises the representative soil- and water-management applications discussed above.

Table 7. Summary of representative deep learning applications in soil and water management. Authors' synthesis.

Sub-domain	Primary Task	Representative DL Architectures	References
Soil properties	Joint regression and spatial mapping	CNN, RNN, hybrid CNN–RNN	[64–67]
Water stress & irrigation	Image classification/moisture regression	AlexNet, GoogLeNet, Inception, season-specific CNN	[68–70]

3.5. Autonomous Agricultural Systems

Autonomous systems integrate DL for robotic perception, field navigation, and collaborative task execution (Edge/Federated AI). As opposed to the first three domains, which revolve around an agronomic target (crop, animal, soil or water resource), this domain revolves around an operating mode: a closed loop, where the output of a perception model controls an actuator—steering the machine, activating the spray nozzle, actuating the picking arm—with implications for machinery safety, damage to crops and labour workflow that go beyond a mere perceptual model (say, a disease classifier, used by a human operator). That is why, for instance, a weeding robot's crop-weed classifier and a diagnostic weed-identification model from Section 3.2.2 can be built on the same architecture and yet be discussed in separate subsections here—with the distinction being not in the computer vision task per se, but in who receives its output (a human or an algorithm). Using the operating mode perspective, we classify the evidence according to two capabilities—perception along with navigation and manipulation and device-to-device coordination—rather than in terms of "tasks" ("weeding," "harvesting"), which are also included in the crop management domain, from the perceptual angle.

3.5.1. Perception, Navigation, and Manipulator Control

Real-time perception for weeding and harvesting increasingly uses object-detection and segmentation pipelines whose output feeds directly into a navigation or manipulator-control loop. Field studies include a YOLOX-based blue-laser weeding robot, which couples weed detection to laser-actuator targeting, and an apple-harvesting system evaluated with optimised picking-arm patterns [39,71]. Broader surveys of agricultural robotics similarly frame precision farming and autonomous harvesting as an integrated perception-to-actuation pipeline rather than as a perception task alone [72], and a direct comparison of YOLOv8 against Mask R-CNN for instance segmentation in complex orchard environments illustrates how detector choice interacts with the downstream manipulation task (e.g., branch and fruit boundary precision for a picking arm) [73]; a recent review specifically surveys object-detection and localisation techniques for fruit-harvesting robots along these lines [74]. Such work should be assessed with end-to-end measures—including cycle time, treatment precision, crop damage and failure recovery—rather than perception accuracy alone.

3.5.2. Distributed Coordination: Federated and Edge AI

Federated and edge-learning studies examine how agricultural devices or institutions can share model updates without centralising raw data [75–77]. Energy-aware device scheduling and communication constraints are key concerns for these systems [78]. Concrete federated-CNN deployments have been reported for several crop-disease clas-

sification tasks—sweet potato [81], banana leaf [82], and strawberry [83]—as well as for smart-farm crop classification more broadly [80], showing that the approach is being tested across multiple crops rather than as isolated proofs of concept. Most of this evidence concerns enabling infrastructure, single-institution pilots, or simulated networks rather than long-duration fleets of autonomous farm robots, so claims of coordinated field autonomy remain prospective.

Table 8 summarises the representative autonomous-system applications and enabling infrastructure discussed above.

Table 8. Summary of representative deep learning applications and enabling infrastructure for autonomous agricultural systems. Authors' synthesis.

Application	Target System	DL Framework	References
Weeding	Laser-weeding robot	YOLOX-based perception and actuation	[39]
Harvesting	Apple/orchard harvester	Vision-guided picking; YOLOv8 vs. Mask R-CNN	[71,73]
Distributed learning	Agricultural edge/IoT nodes	Federated CNN, privacy encoding, device scheduling	[75–78,80–83]

3.6. Cross-Domain Synthesis: What the Four Domains Reveal When Read Together

Reading sequentially, Sections 3.2–3.5 might well be thought of as four separate literatures. Reading comparatively against the same five-dimension rubric (Table 2), one finds the same underlying structural bias repeated across all four, caused by the particular needs of each domain. In each domain, the highest reported benchmark numbers come from studies employing the least robust validation method—curated crop-disease datasets (Section 3.2.1), identification datasets for single barns (Section 3.3.1), single-region soil spectral libraries (Section 3.4.1), and laboratory-based robotic trials (Section 3.5)—while more infrequently published studies with stronger validation methods report comparatively weaker but more believable numbers. This is not an accident in four disparate fields, but a result of the publishing incentives that reward reporting the highest possible internal-split score.

The fields differ, though, in *why* validation is weak, and this is important in recommending solutions (Section 5.2): in crop management, the main reason for weak validation is a dataset-curation decision, since public benchmark datasets like PlantVillage are available but are not representative of the field (Section 3.2.1); the solution here is a data solution. In livestock farming, the main reason for weak validation is a data-*access* issue, since video data at the barn or farm level is commercial property and is only rarely published (Section 3.3.2); the solution here is a governance and incentive-based solution (Section 4.8). In soil/water management, the main reason for weak validation is the fundamental cost of ground-truthing (soil samples and laboratory analyses take much time and money relative to labelling); here the solution is a resource-based solution. In automation, the main reason for weak validation is the cost of extended field trials using physical robots; here the solution is methodological: end-to-end multi-trial evaluation procedures, not single-run demonstrations (Section 3.5). Because the underlying reason differs between fields, a universal recommendation ("more field data") is insufficient; Section 5.2, therefore, grounds each roadmap priority in the specific evidentiary gap that motivates it.

3.7. Critical Analysis: Limitations and Shortcomings of Current Deep Learning Approaches

Despite the significant advances described above, several persistent weaknesses remain. The laboratory-to-field performance gap is the most prominent: models trained on public or curated datasets often fail to generalise to the variability of actual farm environ-

ments, including changing illumination, occlusion, mixed stress factors, different cultivars, seasonal drift, sensor ageing, and management heterogeneity. This gap is especially problematic when studies report high accuracy without external validation, temporal hold-out testing, or cross-site evaluation. As detailed in Section 3.2.1, this is not merely an observed correlation but a mechanistically explicable one: curated benchmark datasets systematically under-represent the visual variability of field conditions, so a model can achieve a high score by exploiting acquisition-protocol cues that are absent in the field. Critically, this gap is a property of how a given study was designed, not an inherent ceiling on deep learning as a method—field-trained and field-validated systems in the same application area (plant disease diagnosis) have reported field accuracies in the 65–88% range [27,28], and multi-institution field-acquired benchmarks now exist for at least one major crop-phenotyping task [92]. The weakness, therefore, has not been "overcome" by the existing literature—most of the corpus we reviewed still lacks external or cross-site validation (Table 2)—but the evidence base does contain a credible, if still minority, existence proof of what closing it requires: training data acquired under field conditions, evaluation on genuinely held-out sites or seasons, and reporting of the resulting field accuracy alongside the laboratory figure. Sections 4 and 6 return to this point when prioritising research directions by the strength of evidence that currently supports them.

A second weakness is the lack of standardised validation protocols. Accuracy, F_1, mAP, IoU, RMSE, and R^2 are often reported in ways that are not directly comparable across studies, and confidence intervals or variance across runs are rarely provided. Third, the dominance of private datasets limits reproducibility and makes it difficult to determine whether reported improvements reflect genuinely better models or dataset-specific advantages. Fourth, interpretability and economic feasibility remain underdeveloped. A model that detects disease symptoms or predicts water stress is useful to farmers only if its output can be trusted, explained, integrated into management decisions, and justified relative to sensor, maintenance, training, and operational costs. These issues motivate the emergent technologies and recommendations discussed in Sections 4–6.

4. Emergent Technologies and Future Frontiers

The contemporary state of deep learning technology presented above has triggered a new wave of interest in technologies and approaches that can continue to accelerate this field. The years 2020–2025 have seen several new approaches being pursued to increase trust, preserve privacy, overcome the issue of limited data, allow immediate implementation, and combine knowledge with machine learning. Specifically, the transition toward Agriculture 5.0 and Society 5.0 emphasises a human-centric, sustainable, and resilient approach to agricultural intelligence, where AI serves not only productivity but also social and environmental well-being [2,8]. Particularly important emerging frontiers are self-supervised representation learning, multimodal foundation models, and governance-aware deployment frameworks, because they directly address the current bottlenecks of annotation scarcity, weak transfer, and uneven access. Table 9 summarises these emergent technologies and their key benefits.

4.1. Enhancing Trust and Transparency with Explainable AI (XAI)

A key barrier to large-scale adoption of DL-based systems in agricultural contexts where outcomes are highly consequential remains the so-called "black box" issue [13]. Agricultural professionals need to understand the basis of consequential recommendations, and explainable AI seeks to translate model behaviour into human-interpretable representations [94]. Agricultural scenarios increasingly require XAI methods that provide actionable information alongside predictions. SHapley Additive exPlanations (SHAP), Lo-

cal Interpretable Model-Agnostic Explanations (LIME) [94], and Gradient-weighted Class Activation Mapping (Grad-CAM) [95] are among the commonly discussed approaches. SHAP and LIME can rank the influence of tabular or multimodal variables, such as soil quality, rainfall, temperature, or management inputs. Grad-CAM is particularly useful for convolutional models because it produces saliency maps that identify image regions contributing strongly to a prediction.

However, generic saliency maps are not sufficient on their own. Explanations must be aligned with agronomic reasoning and with the decisions that practitioners actually make. For example, a disease detector should not only highlight a leaf region, but should ideally relate the prediction to symptom patterns recognised by plant pathologists. Similarly, a water-stress model should connect its output to interpretable indicators such as soil moisture deficit, evapotranspiration, or irrigation timing. Counterfactual explanations may also be valuable when they clarify how a recommendation would change under different management actions, such as altered irrigation or nitrogen application. XAI is, therefore, not only a matter of computational transparency; it is a prerequisite for trust, adoption, regulatory accountability, and farmer-facing usefulness.

4.2. Preserving Privacy with Federated Learning (FL)

The development of robust DL models in agriculture often requires data from multiple farms, institutions, or regions, yet many stakeholders are reluctant to share raw data because of privacy, competition, and regulatory concerns [75]. Federated Learning (FL) has, therefore, emerged as a promising approach in which models are trained collaboratively while raw data remain at their source [96].

In a typical FL pipeline, a shared model is distributed to participating nodes, each node trains locally on its own data, and only model updates are transmitted back to a coordinating server for aggregation [76]. This cycle is repeated until convergence. Because only parameter updates are exchanged, FL can reduce direct exposure of farm-level data while still benefiting from geographically distributed evidence.

Agricultural applications of FL already include disease recognition and yield-related prediction tasks. For example, [79] showed that FL can support wheat disease identification across multiple institutions with performance close to centralised training, while [75] argued that FL may help overcome regulatory barriers that currently prevent access to distributed agricultural datasets.

However, FL still faces important challenges, including non-IID farm data, limited rural connectivity, and privacy risks linked to information leakage from model updates. Current work, therefore, focuses on communication-efficient training, more robust aggregation strategies, and combinations of FL with privacy-preserving techniques such as differential privacy [76]. FL also raises broader governance questions, including data ownership, fairness, and the concentration of agricultural intelligence in a small number of platforms. For these reasons, FL should be viewed as a promising infrastructure for collaborative agricultural AI rather than as a fully solved deployment strategy.

4.3. Attention Mechanisms for Enhanced Convolutional Architectures

In addition to the full Vision Transformer backbones (Section 3.1), the second approach, gaining traction since 2020, involves the embedding of a light attention module into an otherwise convolutional neural network so that the model could learn to give higher weights to the informative regions/feature channels rather than uninformative channels. It is different from using the entire Transformer architecture in terms of its design principles and there is consistent evidence that it can help in improving the accuracy of analysis of

crowded or low-contrast agricultural images without the additional computational cost of the whole backbone being transformed.

A number of cases can be found among those domains listed in Section 3. A self-attention module embedded in the base CNN helped in improving crop leaf disease recognition accuracy to 95–98% by helping the network to localise small and low-contrast lesions against a background [87]. The Squeeze-and-Excitation and Convolutional Block Attention Module (CBAM) blocks, which recalibrate the importance of channels/space-wise features, have been used for calibrating pretrained backbone networks (DenseNet121, MobileNetV2) for grapevine and other leaf-disease recognition, thus increasing the accuracy by several percentage points [88]. In the case of precision livestock farming, the same mechanism of CBAM helped in improving real-time broiler detection on crowded litter-floor backgrounds (Section 3.3.3) [61]. Increasingly, weed-detection papers benchmark the YOLO- and Transformer-based (attention-based) detectors against each other [37]. All in all, the presented evidence shows that attention augmentation should be interpreted as an architectural improvement that is transferable to multiple domains rather than something that applies only to a certain application.

4.4. Overcoming Data Scarcity with Generative AI and the Prospective Role of Large Language Models

The lack of comprehensively annotated datasets across agricultural domains is one of the main constraints on robust DL development [13]. Generative AI spans a broader family of methods than is sometimes acknowledged: Generative Adversarial Networks (GANs), autoencoders, and, more recently, diffusion models, may support carefully controlled data augmentation, although synthetic data must be checked for biological plausibility and distributional artefacts.

Generative models can produce candidate training images for rare disease or stress classes and may help explore hypothetical weather or pest scenarios. Autoencoders contribute a complementary, non-generative use case: trained to reconstruct normal patterns, they support unsupervised anomaly detection where reconstruction error flags out-of-distribution inputs, which has been demonstrated for smart-farming sensor ecosystems [97] and, in a closely related agricultural-vehicle setting, for detecting unknown obstacles that a closed-set object detector would otherwise miss [98]. Diffusion models are a more recent addition to the agricultural toolkit: denoising diffusion probabilistic models have been used to synthesise additional plant-disease images for training-set augmentation [99], and, combined with transfer learning, to expand small citrus-disease datasets [100]. Because diffusion-based augmentation for agriculture is a very recent development, the evidence base for it remains limited to a handful of studies, and its field-deployment value is not yet established. Such augmentation is useful only when synthetic and authentic samples are kept distinguishable during evaluation, leakage is prevented, and agronomic experts verify that generated symptoms are plausible. A similar but different line of work focuses on the use of large language models (LLMs) and retrieval-augmented generation (RAG) in agriculture and decision support, which has developed sufficiently that a brief discussion of it seems warranted at this point. Domain-adapted agricultural LLMs incorporating external knowledge bases have been proposed to ensure that responses are grounded in verified agronomic data rather than the unfettered parametric knowledge of the model itself [101], and RAG-based conversational frameworks for advising have been compared to human-curated agronomic question sets and were found to yield accuracy up to 93% and a hallucination rate as low as 3% for the most successful configuration [102]. Specifically in precision livestock farming, a recent narrative review of computer-vision and large-language-model applications for animal farming considers the possibility of injecting species- and herd-specific knowledge into RAG to address a gap in the general-purpose

LLMs, while pointing out the need for high factual reliability when producing veterinary and welfare-related recommendations [103]. These examples show the potential for rapid development in this field, but the provenance of the retrieved information, factual verification, and the ability of the model to abstain (i.e., choose not to answer rather than make up an agronomic recommendation) remain important prerequisites that need to be addressed before deploying the LLMs unsupervised to farmers: a model hallucinating a dosage of a pesticide is a different kind of mistake from a vision model missing a leaf lesion. Because this field (LLMs, RAG, and multi-agent decision support systems for agriculture) is sufficiently big and rapidly evolving, it deserves its own dedicated review; curious readers are pointed toward these systems as a starting point, and we return to their prospective application together with LLM-based literature search tools described in Section 2 in the roadmap in Section 6. Presently, however, these uses of language models and retrieval-augmented generation are still emerging lines of research rather than demonstrated evidence of field benefit.

Table 9. Emergent technologies addressing key challenges in agricultural AI (2020–2025). Authors' synthesis.

Core Challenge	Emergent Solution	Key Benefits & Features	Representative References
Model Opacity	**Explainable AI (XAI)**	• Tools (LIME, SHAP, **Grad-CAM**) for translating DL outputs into readable interpretations • Saliency maps for visual justification of classifications	[13,94,95]
Data Privacy	**Federated Learning (FL)**	• Collaborative training without exchanging raw data • Maintains privacy while benefiting from distributed evidence	[75,76,79,96]
Low-Contrast/ Cluttered Scenes	**Attention-Augmented CNNs**	• Self-attention, SE, and CBAM modules embedded in CNN backbones • Improves localisation of small/low-contrast targets at lower cost than a full transformer	[61,87,88]
Data Scarcity	**Generative AI**	• GANs, autoencoders, and diffusion models generate or reconstruct synthetic candidates for controlled augmentation • Requires plausibility and leakage checks	[13,97,99]
Knowledge Access & Decision Support	**LLMs & Retrieval-Augmented Generation**	• Grounds farmer-facing advisory answers in verifiable sources • Provenance, factual verification, and abstention remain open problems	[101–103]
Latency	**Edge AI**	• On-device inference on drones/robots • Reduces reliance on cloud and improves privacy	[16]
Annotation Cost	**Self-Supervised Learning**	• Learns representations from unlabeled corpora • Reduces dependence on expert labels	[36,91]
Robustness	**Hybrid PBM-DL Models**	• Merges mechanistic models with neural networks • Constrains predictions through physical principles	Emerging direction; limited field evidence

4.5. Enabling Real-Time Action with Edge AI

Many agricultural applications require immediate analysis and decision making in environments with limited connectivity, strict latency constraints, or privacy concerns [16].

Edge AI addresses this need by running models directly on devices such as IoT nodes, drones, robots, and farm machinery rather than relying entirely on remote cloud processing [16]. Recent work has, therefore, focused on on-device inference strategies that preserve acceptable performance while reducing computational cost.

Common optimisation techniques include quantisation, pruning, and knowledge distillation, all of which can compress deep models for low-power hardware. Interest in TinyML further reflects the need to operate on microcontrollers and energy-constrained embedded systems. For instance, a low-power insect monitoring device could use a compact CNN to detect characteristic flight patterns or sounds and trigger an alert in real time. In field robotics, optimised MobileNet or YOLO variants can support immediate weed or fruit detection directly on the platform.

Edge AI is especially attractive because it reduces dependence on network connectivity, lowers transmission costs, and can improve privacy by keeping raw data local. As embedded hardware continues to improve, edge deployment is likely to become increasingly important for practical agricultural AI, particularly in remote or infrastructure-limited settings.

4.6. Self-Supervised and Multimodal Representation Learning

Self-supervised learning (SSL) aims to reduce reliance on expensive expert labels by pretraining models on large collections of unlabeled agricultural imagery or spatiotemporal sensor streams before adapting them to downstream tasks. Recent studies suggest that SSL can remain competitive while using substantially fewer labels, which is particularly relevant in agriculture because annotation is costly and environmental variability is high [36,91].

SSL is increasingly combined with multimodal representation learning, where image, sensor, weather, and management data are aligned in a shared latent space. In principle, this strategy offers improved robustness to missing modalities, better transfer across sensing platforms, and stronger support for forecasting and decision-support tasks. However, the evidence base remains early-stage: most results are still reported on benchmark datasets or institutional collections, and only a limited number of studies assess whether these gains persist under real-world deployment conditions. SSL should, therefore, be regarded as a promising frontier rather than a mature solution.

4.7. Integrating Mechanistic and Data-Driven Models: Hybrid PBM-DL Systems

Another important development is the integration of DL with mechanistic or process-based models (PBMs), often described as hybrid modelling or knowledge-informed AI. Agriculture already has a long tradition of PBMs that simulate plant growth, water balance, soil processes, pest dynamics, and animal production using equations grounded in biological or physical principles. Hybrid PBM–DL approaches attempt to combine mechanistic constraints with the pattern-recognition capacity of neural networks, although the reviewed literature provides less deployment evidence for these hybrids than for vision-based applications.

This integration can occur in several ways. A DL yield model may include biologically plausible constraints so that predictions remain within feasible harvest limits. A neural network may use PBM outputs, such as simulated biomass, evapotranspiration, or soil water deficit, as additional inputs alongside satellite or sensor data. Conversely, a neural component may replace a poorly specified sub-module inside a larger mechanistic model when observational data are sufficient. In irrigation management, for example, water-balance equations can be combined with LSTM or GRU models to improve short-term water-demand forecasts. In pest management, a PBM can describe expected population growth while a neural model updates this estimate using real-time monitoring data.

The advantage of hybrid PBM–DL systems is not simply higher accuracy. Their deeper value is that they can reduce implausible predictions, improve extrapolation under data scarcity, and make model behaviour easier to interpret for agronomists and farm advisors. The main challenge is that these systems require close collaboration across computer science, agronomy, soil science, hydrology, and animal science. For this reason, hybrid modelling should be treated as a promising but demanding research direction rather than as a simple add-on to existing DL workflows.

4.8. Ethics, Data Governance, and Equity

In addition to architectural design, large-scale agricultural AI deployment will raise questions of governance related to data ownership in farming, bargaining power in platform-based ecosystems, fairness in farm size and geographic location, and the sustainability of the energy costs of sensing and autonomy. Recent analyses of the adoption of Agriculture 4.0 technologies stress the need for data governance to include data consent, access control, reuse, and value sharing. These issues cannot be ignored. Indeed, if the data used for training is dominated by large and well-capitalised farms, the resulting AI may not function well in smaller and less connected farms, thus exacerbating existing inequalities.

In parallel with technical benchmarking, ethical analysis is also necessary. At the very least, subsequent studies need to discuss data provenance, distributional performance across relevant groups of users, fallback mechanisms in the event of critical failures, and the sustainability of the energy costs of sensing and computation. In the case of autonomous vehicles, data governance will also overlap with certification and liability. It is not sufficient to discuss the accuracy of perception in the field if the deployment is related to worker safety, animal welfare, or irreversible agronomic actions.

4.9. The Transition Toward Agriculture 5.0 and Society 5.0

The emergent technologies discussed in this section should be interpreted within the wider transition from Agriculture 4.0, which emphasises automation, efficiency, and data-driven optimisation, toward the more human-centred vision of Agriculture 5.0 and Society 5.0 [2,8]. While Agriculture 4.0 focuses mainly on digital tools for improving output and efficiency, Agriculture 5.0 frames AI and robotics as instruments for resilience, sustainability, and well-being, with human decision makers remaining central.

Within this paradigm, deep learning is not an end in itself. It should support farmers, veterinarians, and agronomists rather than replace their judgement. This changes how DL systems should be evaluated: beyond benchmark accuracy, they should be assessed for interpretability, their ability to strengthen rather than erode farmer agency, and their contribution to equitable access to agricultural innovation. Small-scale and resource-constrained producers should be able to benefit from AI tools, rather than merely supplying data for systems designed primarily around large agribusiness needs.

The Society 5.0 framework demands, in addition, that agricultural AI confronts the environmental and social externalities. Energy-efficient architectures, low-carbon sensing pipelines, governance structures for data sovereignty, and transparent mechanisms of accountability do not represent optional additions but constitute core requirements. Future research programmes should, therefore, integrate the sociological, ethical, and economic evaluation alongside the technical benchmarking, so the next generation of agricultural AI delivers a value which is measurable not only in yield or efficiency, but in the quality of livelihood, the health of the ecosystem, and the long-term resilience of the food-system.

5. Conclusions and Outlook

5.1. Synthesis of Key Advances and Persistent Challenges

The evidence synthesised in this review indicates that deep learning has moved agriculture beyond isolated proof-of-concept demonstrations, but not yet into uniformly mature large-scale deployment. The strongest body of evidence remains concentrated in crop-management tasks such as disease recognition, weed detection, fruit counting, and yield estimation, where public datasets and computer-vision benchmarks are relatively more developed. Livestock monitoring, soil and water management, and autonomous operations have also advanced substantially, yet these domains still rely more heavily on private datasets, single-site studies, or partial system evaluations. The central conclusion is, therefore, twofold: deep learning is already a high-value methodological asset for agricultural intelligence, but the maturity of the evidence base remains uneven across domains and far more limited in field deployment than headline benchmark scores might imply.

The most consequential weakness across the literature is the mismatch between controlled evaluation and operational performance. Models that appear highly accurate on narrow benchmarks often degrade under domain shift, mixed stress conditions, longitudinal deployment, and resource-constrained hardware. This laboratory-to-field mismatch is compounded by several related weaknesses: incomplete benchmarking standards, weak external validation, limited interpretability, fragmented data governance, and insufficient techno-economic analysis. As a result, the most publishable future work in this area will likely be that which demonstrates not only algorithmic novelty, but also reproducible superiority under realistic field conditions, transparent reporting, and credible integration into farmer-facing workflows.

5.2. A Roadmap for Future Research

To address these challenges and unlock the full potential of artificial intelligence in agriculture, future research studies must focus on more inclusive and interwoven solutions rather than working towards individual model extensions. Based on an analysis of achievements and weaknesses conducted in this literature review, below are the priority research directions, ordered from the direction with the strongest existing evidentiary support in our corpus to the direction that is most prospective. This ordering reflects how consistently the reviewed studies already document each problem or partial solution (Table 2 and Section 3.7), not a judgement about the ultimate scientific importance.

1. Development of robust benchmarks and public datasets (strongest evidence basis: this is the single most consistently documented gap across all four domains in Section 3, and Section 3.2.1 shows concretely what closing it looks like): A major priority is the creation of realistic benchmarks that reflect actual deployment conditions. Public datasets should expand in coverage across crops, regions, sensing setups, and difficult edge cases such as heavy rainfall, strong occlusion, or low disease prevalence. Community challenges, similar to those used in broader computer-vision research, could help establish shared benchmarks for tasks such as disease detection, yield forecasting, and robotic perception. It may also be useful to structure benchmark design across pre-harvest, harvest, and post-harvest phases, since each stage presents distinct technical and adoption barriers [104].

2. Modelling of multimodal and temporal data (moderate-to-strong evidence basis: self-supervised pretraining already shows competitive results with far fewer labels in the studies we reviewed, Section 4.6, though multimodal fusion specifically remains comparatively less tested): Agricultural datasets are becoming larger, but their greater scientific value lies in their increasing multimodal and temporal structure. Future work

should develop models able to integrate heterogeneous information from images, sensors, weather records, soil measurements, management logs, and textual advisory data. Attention mechanisms (Section 4.3) and graph neural networks provide useful tools for assigning different weights to inputs depending on time, location, and agronomic context. Self-supervised pretraining and multimodal representation learning should become central design choices, because they offer one of the most plausible ways to reduce dependence on expert annotation while improving transfer across farms, seasons, crops, and sensing platforms. Few-shot learning and domain adaptation will also be important for adapting models to new crops or regions with limited labelled data.

However, multimodal systems create economic and computational trade-offs. Adding more sensors can improve robustness, but it also increases installation cost, maintenance burden, energy use, bandwidth demand, and data-governance complexity. The real-time processing of multimodal data may also exceed the capacity of edge hardware used on drones, tractors, or low-cost IoT nodes. Future studies should, therefore, report whether the marginal gain obtained from additional modalities justifies the additional cost and complexity, or whether sufficiently reliable performance can be achieved with a smaller and more accessible set of inputs.

3. Modernising crop- and field-scale monitoring with UAV, satellite, and geospatial foundation models (moderate evidence basis: individual UAV/satellite studies already report strong task-level results, but few evaluate multi-season or cross-region transfer): UAV- and satellite-based sensing is already a mature data-acquisition channel for several tasks in this review (Sections 3.2.3 and 3.5.1), and the near-term opportunity is less about acquiring more imagery than about modernising the models applied to it. Deep-transfer-learning strategies that adapt satellite time-series models from data-rich to data-scarce regions [46], and the fine-tuning of large pretrained geospatial foundation models (e.g., NASA/IBM's Prithvi) for crop-type and crop-health mapping in smallholder settings [105], both point toward a shift from bespoke, per-study CNNs trained from scratch toward adapting large pretrained models to local conditions with comparatively little labelled data. This direction is particularly relevant for smallholder and resource-constrained producers (Section 4.9), who are least able to fund a bespoke acquisition and modelling pipeline of their own.

4. Synergistic integration of emerging methodologies (emerging evidence basis: XAI, FL, and Edge AI are each individually evidenced in Section 4, but studies combining them are still scarce): Future agricultural AI systems will likely combine multiple complementary paradigms rather than rely on a single modelling strategy. For example, plant-disease platforms may need *XAI* for transparent decisions, *FL* for privacy-preserving collaboration, and *Edge AI* for near-real-time field inference. Research should, therefore, focus on how to integrate these elements coherently, including federated setups with communication constraints and model-compression strategies that preserve both predictive quality and interpretability.

5. Encouraging inter-disciplinary cooperation and user-centric design (least directly evidenced by our corpus, since few included studies report co-design or adoption data; included as a normative priority motivated by the deployment gaps documented throughout Section 3): To address these challenges, AI in agriculture will need continuous cooperation among computer scientists, agronomists, soil scientists, horticultural scientists, robotic engineers, social scientists, and economists. Human-centred design should promote not only accuracy but also usability and relevance to agricultural practice [11]. Participatory research can help align interfaces and decision-support outputs with practitioner workflows, while socio-economic studies can identify adoption barriers and the conditions under which agricultural AI creates practical value.

5.3. Concluding Reflections on the Trajectory of AI in Sustainable Agriculture

The trajectory of deep learning in agriculture is best understood not simply as a sequence of accuracy improvements but as a broader transition from narrow perception tasks to integrated socio-technical systems. The field is increasingly judged not only by whether models can classify images or predict agronomic variables, but by whether they can do so reliably across seasons, locations, user groups, and hardware constraints while remaining explainable, economically justifiable, and governable. This shift matters because it changes the standard of scientific maturity: the strongest contributions are those that connect model innovation with deployment realism and transparent failure analysis.

Final Assessment

In sum, the literature reviewed here supports a strong but qualified conclusion: deep learning is already reshaping agricultural sensing, prediction, and automation, but the field remains methodologically immature in several of the dimensions that matter most for adoption. The most useful contribution of this review is, therefore, not to celebrate benchmark gains, but to clarify where the evidence is strong, where it is fragile, and what must change for agricultural AI to become scientifically robust and practically credible.

6. Recommendations for Future Research and Practice

On the basis of the evidence and critical considerations that have been presented in Sections 3–5, this section provides a set of explicit recommendations that may constitute a research and innovation agenda for deep learning applications in the agricultural domain. These recommendations are organised according to four complementary dimensions: data and benchmarking, methodological advances, system-level deployment and socio-technical governance.

6.1. Data, Benchmarks, and Experimental Protocols

- **Develop open, field-realistic benchmark suites.** It is important that future research efforts prioritise the construction of multi-institutional benchmark datasets that are openly accessible and capable of capturing realistic field variability for each principal task including plant disease detection, weed segmentation, yield prediction, animal behaviour analysis, soil and water monitoring and robotic perception. Such benchmarks should intentionally incorporate challenging conditions like occlusions, extreme illumination and mixed stress factors, and they must be accompanied by comprehensive metadata regarding agronomic context, sensor specifications and management practices.
- **Standardise evaluation metrics and reporting.** At the present time, research studies report heterogeneous metrics such as accuracy, F_1, mAP, IoU, R^2 and RMSE using diverse validation protocols. For each task category, it is necessary to establish community-level agreements concerning primary and secondary metrics, cross-validation schemes and reporting standards that include confidence intervals and ablation results. This standardisation would render quantitative comparisons across different methods more meaningful and would enable genuine progress tracking.
- **Systematic inclusion of negative and borderline results.** The current evidence base is likely affected by publication bias, with an under-representation of studies in which deep learning does not outperform simpler baselines or fails to generalise under realistic field conditions. Scientific journals and conferences should explicitly encourage the publication of methodologically sound negative or borderline findings, particularly for domain shift and long-term deployment studies.

- **Modernise UAV- and satellite-based monitoring with pretrained geospatial models.** Rather than continuing to train bespoke CNNs from scratch for each new UAV or satellite monitoring task, future benchmark efforts should systematically evaluate deep-transfer-learning strategies for satellite image time series [46] and the fine-tuning of large pretrained geospatial foundation models for crop-type and crop-health mapping [105], reporting how much labelled local data is required to reach usable accuracy in a new region relative to a from-scratch baseline.

6.2. Methodological Priorities for Robust and Interpretable Models

- **Robustness across domains, seasons, and management regimes.** Methodological research should concentrate on architectures and training regimes that explicitly address domain shift phenomena including cross-region, cross-season, cross-cultivar and cross-management generalisation. Approaches such as domain adaptation, continual learning, invariant representation learning and meta-learning should be systematically benchmarked on agricultural tasks where environmental drift and management changes represent intrinsic characteristics.
- **Integration of multimodal and mechanistic information.** Future deep learning models should combine heterogeneous data sources comprising spectral imagery, time-series sensor streams, textual records and genomic data together with mechanistic crop, soil and animal models. Self-supervised pretraining across these heterogeneous modalities should be prioritised in such sparse-label settings so that the downstream models do not rely solely on the narrow and manually curated test cases. Hybrid PBM-DL architectures should not just be developed as proofs of concepts but as useful tools with well-specified protocols for the calibration and integration of the process-based and the DL-based approaches.
- **Task-specific explainability that reflects agronomic reasoning.** Explainable AI needs to progress from generic saliency maps towards task-specific domain-aligned explanations. For instance, disease diagnosis models should explain predictions in terms of symptom patterns that correspond to the descriptors employed by pathologists, whilst water stress models should relate inferred states to interpretable water balance descriptors. The development of evaluation protocols for explanation quality rather than prediction accuracy alone represents a crucial research direction.

6.3. From Prototypes to Integrated, Edge-Ready Systems

- **Design for edge deployment from the outset.** Rather than adapting large models after their development, future work should adopt an "edge-by-design" philosophy whereby model search and architecture design incorporate constraints on latency, memory, energy consumption and on-board learning capabilities. Neural architecture search constrained by edge profiles, compact transformer variants and hardware-aware pruning and distillation represent promising research avenues.
- **Holistic evaluation of end-to-end systems.** Many current studies optimise isolated components such as weed detection in static images; however, agricultural decision support and robotics require complete pipelines encompassing sensing, perception, decision and actuation. Research should report system-level metrics including hectares treated per hour, reduction in chemical inputs, labour savings and safety incidents in addition to pixel-level or object-level metrics. This approach will align evaluation more closely with stakeholder objectives.
- **Federated and collaborative learning infrastructures.** To realise the potential of cross-farm learning under privacy and sovereignty constraints, there is a need for applied research on federated learning architectures that are robust to heterogeneous

data, unreliable connectivity and asynchronous participation. Research should address realistic deployment scenarios involving cooperatives, retailers and veterinary networks, co-designing aggregation schemes and incentive mechanisms together with stakeholders.

6.4. Socio-Technical, Economic, and Governance Dimensions

- **Human-centred design and co-creation with practitioners.** Future research programmes should involve farmers, agronomists, veterinarians and farm workers as co-designers rather than treating them merely as end-users. This approach implies iterative prototyping, field trials and usability studies to ensure that interfaces, alerting thresholds and visualisations are aligned with existing workflows and cognitive models.
- **Rigorous techno-economic assessment.** Alongside technical evaluation, each proposed solution should be accompanied by an assessment of costs including sensors, computing, maintenance and training, together with economic benefits such as yield improvements, input savings and quality premiums, as well as risk profiles stratified by farm size and geographical region. This consideration is particularly important to avoid widening the digital divide between well-capitalised and resource-constrained producers.
- **Towards governance frameworks and certification schemes.** As deep-learning-based decision support and autonomous systems become increasingly pervasive, regulatory and certification frameworks will be required to address safety, data governance, accountability and transparency. Future work should contribute to the conceptual and empirical foundations for such frameworks, for example by defining test suites for safety-critical perception tasks and documenting failure modes in realistic edge cases.
- **Establish verification and abstention standards for LLM- and RAG-based advisory systems.** As large language model and retrieval-augmented decision-support tools for agriculture mature beyond the early systems discussed in Section 4.4, future work should define domain-specific benchmarks for factual accuracy, source provenance, and abstention behaviour (the system declining to answer rather than fabricating a recommendation) before such tools are deployed in farmer-facing advisory roles, given that an erroneous agronomic recommendation (e.g., a pesticide dose) can carry consequences beyond a typical conversational-AI error.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/bdcc10080273/s1.

Author Contributions: Conceptualisation, V.D.A., L.A., V.D. and F.F.; methodology, V.D.A., L.A., V.D. and F.F.; software, V.D. and F.F.; validation, V.D.A., L.A., V.D. and F.F.; formal analysis, V.D.A. and L.A.; investigation, V.D. and F.F.; resources, V.D.A. and L.A.; data curation, V.D. and F.F.; writing—original draft preparation, V.D.A., L.A., V.D. and F.F.; writing—review and editing, V.D.A., L.A., V.D. and F.F.; visualisation, V.D. and F.F.; supervision, L.A. and V.D.; project administration, V.D.A.; funding acquisition, L.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: No new dataset was created for this review. The publications on which the synthesis is based are identified in the reference list.

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

References

1. Food and Agriculture Organization of the United Nations. *World Food and Agriculture—Statistical Yearbook 2024*; FAO: Rome, Italy, 2024. [CrossRef]
2. FAO. *The State of Food and Agriculture 2024—Value-Driven Transformation of Agrifood Systems*; Technical Report; Food and Agriculture Organization of the United Nations: Rome, Italy, 2024. [CrossRef]
3. FAO. *The State of Food and Agriculture 2022. Leveraging Automation in Agriculture for Transforming Agrifood Systems*; Technical Report; Food and Agriculture Organization of the United Nations: Rome, Italy, 2022. [CrossRef]
4. OECD. *Reinforcing Rural Resilience* In *OECD Rural Studies*; OECD Publishing: Paris, France, 2025. [CrossRef]
5. Li, T.; Lu, H.; Luo, Q.; Li, G.; Gao, M. The Impact of Rural Population Aging on Agricultural Cropping Structure: Evidence from China's Provinces. *Agriculture* **2024**, *14*, 586. [CrossRef]
6. Kamilaris, A.; Prenafeta-Boldú, F. Deep Learning in Agriculture: A Survey. *Comput. Electron. Agric.* **2018**, *147*, 70–90. [CrossRef]
7. Sharma, A.; Jain, A.; Gupta, P.; Chowdary, V. Machine Learning Applications for Precision Agriculture: A Comprehensive Review. *IEEE Access* **2021**, *9*, 4843–4873. [CrossRef]
8. Rose, D.C.; Chilvers, J. Agriculture 4.0: Broadening Responsible Innovation in an Era of Smart Farming. *Front. Sustain. Food Syst.* **2018**, *2*, 87. [CrossRef]
9. Hashemi-Beni, L.; Gebrehiwot, A.A. Deep Learning for Remote Sensing Image Classification in Agriculture Applications. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2020**, *XLIV-M-2-2020*, 51–54. [CrossRef]
10. Charisis, C.; Argyropoulos, D. Deep Learning-Based Instance Segmentation Architectures in Agriculture: A Review of the Scopes and Challenges. *Smart Agric. Technol.* **2024**, *8*, 100448. [CrossRef]
11. Prashanth, J.S.; Krishna, G.B.; Prasad, A.V.K.; Rao, P.R. Smart Farming Revolution: A Cutting-Edge Review of Deep Learning and IoT Innovations in Agriculture. *Oper. Res. Forum* **2025**, *6*, 35. [CrossRef]
12. Fuentes, A.; Yoon, S.; Kim, S.C.; Park, D.S. A Robust Deep-Learning-Based Detector for Real-Time Tomato Plant Diseases and Pests Recognition. *Sensors* **2017**, *17*, 2022. [CrossRef] [PubMed]
13. Sarker, I.H. Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions. *SN Comput. Sci.* **2021**, *2*, 420. [CrossRef] [PubMed]
14. Sardoğan, M.; Tuncer, A.; Ozen, Y. Plant Leaf Disease Detection and Classification Based on CNN with LVQ Algorithm. In Proceedings of the *2018 3rd International Conference on Computer Science and Engineering (UBMK)*, Sarajevo, Bosnia and Herzegovina, 20–23 September 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 382–385. [CrossRef]
15. Ruiz, L.; Recio, J.; Fernández-Sarría, A.; Hermosilla, T. A Feature Extraction Software Tool for Agricultural Object-Based Image Analysis. *Comput. Electron. Agric.* **2011**, *76*, 284–296. [CrossRef]
16. Albahar, M. A Survey on Deep Learning and Its Impact on Agriculture: Challenges and Opportunities. *Agriculture* **2023**, *13*, 540. [CrossRef]
17. Araujo, S.O.; Peres, R.S.; Ramalho, J.C.; Lidon, F.; Barata, J. Machine Learning Applications in Agriculture: Current Trends, Challenges, and Future Perspectives. *Agronomy* **2023**, *13*, 2976. [CrossRef]
18. Sharma, R.; Kamble, S.S.; Gunasekaran, A.; Kumar, V.; Kumar, A. A Systematic Literature Review on Machine Learning Applications for Sustainable Agriculture Supply Chain Performance. *Comput. Oper. Res.* **2020**, *119*, 104926. [CrossRef]
19. Zhang, Q.; Liu, Y.; Gong, C.; Chen, Y.; Yu, H. Applications of Deep Learning for Dense Scenes Analysis in Agriculture: A Review. *Sensors* **2020**, *20*, 1520. [CrossRef] [PubMed]
20. Li, Y.; Zhang, H.; Xue, X.; Jiang, Y.; Shen, Q. Deep Learning for Remote Sensing Image Classification: A Survey. *WIREs Data Min. Knowl. Discov.* **2018**, *8*, e1264. [CrossRef]
21. Attri, I.; Awasthi, L.K.; Sharma, T.P.; Rathee, P. A Review of Deep Learning Techniques Used in Agriculture. *Ecol. Inform.* **2023**, *77*, 102217. [CrossRef]
22. Kuan, Y.N.; Goh, K.M.; Lim, L.L. Systematic Review on Machine Learning and Computer Vision in Precision Agriculture: Applications, Trends, and Emerging Techniques. *Eng. Appl. Artif. Intell.* **2025**, *148*, 110401. [CrossRef]
23. Zhang, R.; Wu, X.; Li, J.; Zhao, P.; Zhang, Q.; Wuri, L.; Zhang, D.; Zhang, Z.; Yang, L. A Bibliometric Review of Deep Learning in Crop Monitoring: Trends, Challenges, and Future Perspectives. *Front. Artif. Intell.* **2025**, *8*, 1636898. [CrossRef] [PubMed]
24. 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. [CrossRef]
25. Page, M.; McKenzie, J.; Bossuyt, P.; Boutron, I.; Hoffmann, T.; Mulrow, C.; Shamseer, L.; Tetzlaff, J.; Akl, E.; Brennan, S.; et al. The PRISMA 2020 Statement: An Updated Guideline for Reporting Systematic Reviews. *PLoS Med.* **2021**, *18*, e1003583. [CrossRef] [PubMed]
26. Mohanty, S.; Hughes, D.P.; Salathé, M. Using Deep Learning for Image-Based Plant Disease Detection. *Front. Plant Sci.* **2016**, *7*, 1419. [CrossRef] [PubMed]
27. Ramcharan, A.; McCloskey, P.; Baranowski, K.; Mbilinyi, N.; Mrisho, L.; Ndalahwa, M.; Legg, J.; Hughes, D.P. A Mobile-Based Deep Learning Model for Cassava Disease Diagnosis. *Front. Plant Sci.* **2019**, *10*, 272. [CrossRef] [PubMed]

28. Mrisho, L.M.; Mbilinyi, N.A.; Ndalahwa, M.; Ramcharan, A.M.; Kehs, A.K.; McCloskey, P.C.; Murithi, H.; Hughes, D.P.; Legg, J.P. Accuracy of a Smartphone-Based Object Detection Model, PlantVillage Nuru, in Identifying the Foliar Symptoms of the Viral Diseases of Cassava—CMD and CBSD. *Front. Plant Sci.* **2020**, *11*, 590889. [CrossRef] [PubMed]
29. Xie, X.; Ma, Y.; Liu, B.; He, J.; Li, S.; Wang, H. A Deep-Learning-Based Real-Time Detector for Grape Leaf Diseases Using Improved Convolutional Neural Networks. *Front. Plant Sci.* **2020**, *11*, 751. [CrossRef] [PubMed]
30. Rezk, N.G.; Attia, A.F.; El-Rashidy, M.A.; El-Sayed, A.; Hemdan, E.E. An Efficient Plant Disease Recognition System Using Hybrid Convolutional Neural Networks (CNNs) and Conditional Random Fields (CRFs) for Smart IoT Applications in Agriculture. *Int. J. Comput. Intell. Syst.* **2022**, *15*, 65. [CrossRef]
31. Upadhyay, N.; Gupta, N. Detecting Fungi-Affected Multi-Crop Diseases on Heterogeneous Region Datasets Using a Modified ResNeXt Approach. *Environ. Monit. Assess.* **2024**, *196*. [CrossRef] [PubMed]
32. Upadhyay, N.; Gupta, N. SegLearner: A Segmentation-Based Approach for Predicting Disease Severity in Infected Leaves. *Multimed. Tools Appl.* **2025**, *84*, 42523–42546. [CrossRef]
33. Mathew, M.P.; Elayidom, S.; Raj, V.P.J.; Abubeker, K.M. Development of a Handheld GPU-Assisted DSC-TransNet Model for the Real-Time Classification of Plant Leaf Disease Using Deep Learning Approach. *Sci. Rep.* **2025**, *15*, 3579. [CrossRef] [PubMed]
34. Liu, B.; Zhang, Y.; He, D.; Li, Y. Identification of Apple Leaf Diseases Based on Deep Convolutional Neural Networks. *Symmetry* **2017**, *10*, 11. [CrossRef]
35. Wang, G.; Sun, Y.; Wang, J. Automatic Image-Based Plant Disease Severity Estimation Using Deep Learning. *Comput. Intell. Neurosci.* **2017**, *2017*, 1–8. [CrossRef] [PubMed]
36. Huan, X.; Chen, B.; Zhou, H. A Unified Self-Supervised Framework for Plant Disease Detection on Laboratory and In-Field Images. *Electronics* **2025**, *14*, 3410. [CrossRef]
37. Allmendinger, A.; Saltik, A.; Peteinatos, G.; Stein, A.; Gerhards, R. Assessing the Capability of YOLO- and Transformer-Based Object Detectors for Real-Time Weed Detection. *Precis. Agric.* **2025**, *26*, 52. [CrossRef]
38. Upadhyay, N.; Sharma, D.K.; Bhargava, A. 3SW-Net: A Feature Fusion Network for Semantic Weed Detection in Precision Agriculture. *Food Anal. Methods* **2025**, *18*, 2241–2257. [CrossRef]
39. Zhu, H.; Zhang, Y.; Mu, D.; Bai, L.; Hao, Z.; Li, H. YOLOX-Based Blue Laser Weeding Robot in Corn Field. *Front. Plant Sci.* **2022**, *13*, 1017803. [CrossRef] [PubMed]
40. Reedha, R.; Dericquebourg, E.; Canals, R.; Hafiane, A. Transformer Neural Network for Weed and Crop Classification of High Resolution UAV Images. *Remote Sens.* **2022**, *14*, 592. [CrossRef]
41. Subeesh, A.; Bhole, S.; Singh, K.; Chandel, N.S.; Rajwade, Y.A.; Rao, K.V.R.; Kumar, S.P.; Jat, D. Deep Convolutional Neural Network Models for Weed Detection in Polyhouse Grown Bell Peppers. *Artif. Intell. Agric.* **2022**, *6*, 47–54. [CrossRef]
42. Khaki, S.; Pham, H.; Wang, L. Simultaneous Corn and Soybean Yield Prediction from Remote Sensing Data Using Deep Transfer Learning. *Sci. Rep.* **2020**, *11*, 11132. [CrossRef] [PubMed]
43. Elavarasan, D.; Member, I.P.M.D.V.; Durairaj, P.M. Crop Yield Prediction Using Deep Reinforcement Learning Model for Sustainable Agrarian Applications. *IEEE Access* **2020**, *8*, 86886–86901. [CrossRef]
44. You, J.; Li, X.; Low, M.; Lobell, D.; Ermon, S. Deep Gaussian Process for Crop Yield Prediction Based on Remote Sensing Data. In *Proceedings of the AAAI Conference on Artificial Intelligence*; Association for the Advancement of Artificial Intelligence (AAAI): Palo Alto, CA, USA, 2017; Volume 31, pp. 4559–4566. [CrossRef]
45. Zhou, H.; Huang, F.; Lou, W.; Gu, Q.; Ye, Z.; Hu, H.; Zhang, X. Yield Prediction through UAV-Based Multispectral Imaging and Deep Learning in Rice Breeding Trials. *Agric. Syst.* **2025**, *223*, 104214. [CrossRef]
46. Joshi, A.; Pradhan, B.; Chakraborty, S.; Varatharajoo, R.; Gite, S.; Alamri, A. Deep-Transfer-Learning Strategies for Crop Yield Prediction Using Climate Records and Satellite Image Time-Series Data. *Remote Sens.* **2024**, *16*, 4804. [CrossRef]
47. López-Barrios, J.D.; Cabello, J.A.E.; Gómez-Espinosa, A.; Montoya-Cavero, L.E. Green Sweet Pepper Fruit and Peduncle Detection Using Mask R-CNN in Greenhouses. *Appl. Sci.* **2023**, *13*, 6296. [CrossRef]
48. Khattak, A.; Asghar, M.U.; Batool, U.; Asghar, M.Z.; Ullah, H.; Al-Rakhami, M.; Gumaei, A. Automatic Detection of Citrus Fruit and Leaves Diseases Using Deep Neural Network Model. *IEEE Access* **2021**, *9*, 112942–112954. [CrossRef]
49. Koirala, A.; Walsh, K.B.; Wang, Z.; McCarthy, C. Deep Learning for Real-Time Fruit Detection and Orchard Fruit Load Estimation: Benchmarking of 'MangoYOLO. *Precis. Agric.* **2019**, *20*, 1107–1135. [CrossRef]
50. Schumann, A.W.; Mood, N.S.; Mungofa, P.D.; MacEachern, C.; Zaman, Q.; Esau, T. Detection of Three Fruit Maturity Stages in Wild Blueberry Fields Using Deep Learning Artificial Neural Networks. In *2019 ASABE Annual International Meeting Boston, MA, USA, 7–10 July 2019*; American Society of Agricultural and Biological Engineers: St. Joseph, MI, USA, 2019; p. 1. [CrossRef]
51. Gao, Z.; Shao, Y.; Xuan, G.; Wang, Y.; Liu, Y.; Han, X. Real-Time Hyperspectral Imaging for the In-Field Estimation of Strawberry Ripeness with Deep Learning. *Artif. Intell. Agric.* **2020**, *4*, 31–38. [CrossRef]
52. Manoj, S.; S, R.; V, K. Identification of Cattle Breed Using the Convolutional Neural Network. In *2021 3rd International Conference on Signal Processing and Communication (ICPSC)*; IEEE: Piscataway, NJ, USA, 2021; pp. 503–507. [CrossRef]

53. Pang, Y.; Yu, W.; Zhang, Y.; Xuan, C.; Wu, P. An Attentional Residual Feature Fusion Mechanism for Sheep Face Recognition. *Sci. Rep.* **2023**, *13*, 17128. [CrossRef] [PubMed]
54. Tassinari, P.; Bovo, M.; Benni, S.; Franzoni, S.; Poggi, M.; Mammi, L.; Mattoccia, S.; Stefano, L.D.; Bonora, F.; Barbaresi, A.; et al. A computer vision approach based on deep learning for the detection of dairy cows in free stall barn. *Comput. Electron. Agric.* **2021**, *182*, 106030. [CrossRef]
55. Qiao, Y.; Guo, Y.; Yu, K.; He, D. C3D-ConvLSTM Based Cow Behaviour Classification Using Video Data for Precision Livestock Farming. *Comput. Electron. Agric.* **2022**, *193*, 106650. [CrossRef]
56. Barney, S.; Dlay, S.; Crowe, A.; Kyriazakis, I.; Leach, M. Deep Learning Pose Estimation for Multi-Cattle Lameness Detection. *Sci. Rep.* **2023**, *13*, 4499. [CrossRef] [PubMed]
57. Hu, Y.; Xu, J.; Gou, Z.; Cui, D. Recognizing and Localizing Chicken Behaviors in Videos Based on Spatiotemporal Feature Learning. *Artif. Intell. Agric.* **2025**, *15*, 816–828. [CrossRef]
58. Mu, Y.; Hu, J.; Wang, H.; Li, S.; Zhu, H.; Luo, L.; Wei, J.; Ni, L.; Chao, H.; Hu, T.; et al. Research on the Behavior Recognition of Beef Cattle Based on the Improved Lightweight CBR-YOLO Model Based on YOLOv8 in Multi-Scene Weather. *Animals* **2024**, *14*, 2800. [CrossRef] [PubMed]
59. Zhang, Q.; Yang, Y.; Liu, G.; Ning, Y.; Li, J. Dairy Cow Mastitis Detection by Thermal Infrared Images Based on CLE-UNet. *Animals* **2023**, *13*, 2211. [CrossRef] [PubMed]
60. Senthilkumar, C.; Sindhu, C.; Vadivu, G.; Neethirajan, S. Early Detection of Lumpy Skin Disease in Cattle Using Deep Learning—A Comparative Analysis of Pretrained Models. *Vet. Sci.* **2024**, *11*, 510. [CrossRef] [PubMed]
61. Guo, Y.; Aggrey, S.E.; Yang, X.; Oladeinde, A.; Qiao, Y.; Chai, L. Detecting Broiler Chickens on Litter Floor with the YOLOv5-CBAM Deep Learning Model. *Artif. Intell. Agric.* **2023**, *9*, 36–45. [CrossRef]
62. Shwetha, V.; Maddodi, B.S.; Laxmi, V.; Kumar, A.; Shrivastava, S. Latest Trend and Challenges in Machine Learning- and Deep Learning-Based Computational Techniques in Poultry Health and Disease Management: A Review. *J. Comput. Netw. Commun.* **2024**, *2024*, 8674250. [CrossRef]
63. Xiao, S.; Dhand, N.K.; Wang, Z.; Hu, K.; Thomson, P.C.; House, J.K.; Khatkar, M.S. Review of Applications of Deep Learning in Veterinary Diagnostics and Animal Health. *Front. Vet. Sci.* **2025**, *12*, 1511522. [CrossRef] [PubMed]
64. Piccoli, F.; Rossini, M.; Colombo, R.; Schettini, R.; Napoletano, P. A Deep Scalable Neural Architecture for Soil Properties Estimation from Spectral Information. *Comput. Geosci.* **2023**, *180*, 105433. [CrossRef]
65. Hosseini, F.S.; Razavi-Termeh, S.V.; Sadeghi-Niaraki, A.; Choi, S.M.; Jamshidi, M. Spatial Prediction of Physical and Chemical Properties of Soil Using Optical Satellite Imagery: A State-of-the-Art Hybridization of Deep Learning Algorithm. *Front. Environ. Sci.* **2023**, *11*, 1279712. [CrossRef]
66. Meier, M.; Souza, E.D.; Francelino, M.; Filho, E.F.; Schaefer, C. Digital Soil Mapping Using Machine Learning Algorithms in a Tropical Mountainous Area. *Rev. Bras. Ciência Solo* **2018**, *42*, e0170421. [CrossRef]
67. Huang, Y.; Srivastava, R.; Ngo, C.; Gao, J.; Wu, J.; Chiao, S. Data-Driven Soil Analysis and Evaluation for Smart Farming Using Machine Learning Approaches. *Agriculture* **2023**, *13*, 1777. [CrossRef]
68. Chandel, N.S.; Chakraborty, S.K.; Rajwade, Y.A.; Dubey, K.; Tiwari, M.K.; Jat, D. Identifying Crop Water Stress Using Deep Learning Models. *Neural Comput. Appl.* **2021**, *33*, 5353–5367. [CrossRef]
69. Islam, M.P.; Yamane, T. HortNet417v1—A Deep-Learning Architecture for the Automatic Detection of Pot-Cultivated Peach Plant Water Stress. *Sensors* **2021**, *21*, 7924. [CrossRef] [PubMed]
70. Sun, Y.; Zhang, D.; Miao, Z.; Yang, S.; Liu, Q.; Qu, Z. Season-Specific CNN and TVDI Approach for Soil Moisture and Irrigation Monitoring in the Hetao Irrigation District, China. *Agriculture* **2025**, *15*, 1946. [CrossRef]
71. Bu, L.; Chen, C.; Hu, G.; Sugirbay, A.; Sun, H.; Chen, J. Design and Evaluation of a Robotic Apple Harvester Using Optimized Picking Patterns. *Comput. Electron. Agric.* **2022**, *198*, 107092. [CrossRef]
72. P, A. AI-Enhanced Agricultural Robotics: From Precision Farming to Autonomous Harvesting. *Int. J. Sci. Eng. Technol.* **2025**, *13*, 1–6. [CrossRef]
73. Sapkota, R.; Ahmed, D.; Karkee, M. Comparing YOLOv8 and Mask R-CNN for Instance Segmentation in Complex Orchard Environments. *Artif. Intell. Agric.* **2024**, *13*, 84–99. [CrossRef]
74. Shi, X.; Wang, S.; Zhang, B.; Ding, X.; Qi, P.; Qu, H.; Li, N.; Wu, J.; Yang, H. Advances in Object Detection and Localization Techniques for Fruit Harvesting Robots. *Agronomy* **2025**, *15*, 145. [CrossRef]
75. Durrant, A.; Markovic, M.; Matthews, D.; May, D.; Enright, J.; Leontidis, G. The Role of Cross-Silo Federated Learning in Facilitating Data Sharing in the Agri-Food Sector. *Comput. Electron. Agric.* **2021**, *193*, 106648. [CrossRef]
76. Akbari, M.; Syed, A.; Member, I.J.S.H.; Erol-Kantarci, M. Constrained Federated Learning for AoI-Limited SFC in UAV-Aided MEC for Smart Agriculture. *IEEE Trans. Mach. Learn. Commun. Netw.* **2023**, *1*, 277–295. [CrossRef]
77. Kumar, P.; Gupta, G.P.; Tripathi, R. PEFL: Deep Privacy-Encoding-Based Federated Learning Framework for Smart Agriculture. *IEEE Micro* **2022**, *42*, 33–40. [CrossRef]

78. Yu, C.; Shen, S.; Zhang, K.; Zhao, H.; Shi, Y. Energy-Aware Device Scheduling for Joint Federated Learning in Edge-Assisted Internet of Agriculture Things. In *2022 IEEE Wireless Communications and Networking Conference (WCNC)*; IEEE: Piscataway, NJ, USA, 2022; pp. 1140–1145. [CrossRef]
79. Mehta, S.; Kukreja, V.; Vats, S. Empowering Farmers with AI: Federated Learning of CNNs for Wheat Disease Multi-Classification. In *2023 4th International Conference for Emerging Technology (INCET)*; IEEE: Piscataway, NJ, USA, 2023; pp. 1–6. [CrossRef]
80. Idoje, G.; Dagiuklas, T.; Iqbal, M. Federated Learning: Crop Classification in a Smart Farm Decentralised Network. *Smart Agric. Technol.* **2023**, *5*, 100277. [CrossRef]
81. Mehta, S.; Kukreja, V.; Vats, S. Propelling Agritech Forward: Federated Learning and CNN for Sweet Potato Disease Diagnosis. In *2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT)*; IEEE: Piscataway, NJ, USA, 2023; pp. 1–6. [CrossRef]
82. Shukla, A.N.; Joshi, K.; Yadav, A.P.S.; Kukreja, V.; Mehta, S. Precision Phytopathology in Agriculture: A Federated Learning CNN Framework for Banana Leaf Disease Classification. In *2024 11th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)*; IEEE: Piscataway, NJ, USA, 2024; pp. 1–7. [CrossRef]
83. Bansal, A.; Vats, S.; Prasad, C.; Kukreja, V.; Mehta, S. Bridging the Strawberry Cultivation Gap: Federated Learning CNN for Disease Detection. In *2023 12th International Conference on System Modeling & Advancement in Research Trends (SMART)*; IEEE: Piscataway, NJ, USA, 2023; pp. 433–439. [CrossRef]
84. Ha, J.G.; Moon, H.; Kwak, J.T.; Hassan, S.I.; Dang, M.; Lee, O.N.; Park, H.Y. Deep Convolutional Neural Network for Classifying Fusarium Wilt of Radish from Unmanned Aerial Vehicles. *J. Appl. Remote Sens.* **2017**, *11*, 1. [CrossRef]
85. Lu, Y.; Yi, S.; Zeng, N.; Liu, Y.; Zhang, Y. Identification of Rice Diseases Using Deep Convolutional Neural Networks. *Neurocomputing* **2017**, *267*, 378–384. [CrossRef]
86. Dosovitskiy, A.; Beyer, L.; Kolesnikov, A.; Weissenborn, D.; Zhai, X.; Unterthiner, T.; Dehghani, M.; Minderer, M.; Heigold, G.; Gelly, S.; et al. An Image Is Worth 16x16 Words: Transformers for Image Recognition at Scale. In Proceedings of the International Conference on Learning Representations (ICLR), Virtual, 3–7 May 2021.
87. Zeng, W.; Li, M. Crop Leaf Disease Recognition Based on Self-Attention Convolutional Neural Network. *Comput. Electron. Agric.* **2020**, *172*, 105341. [CrossRef]
88. Unal, Y. Integrating CBAM and Squeeze-and-Excitation Networks for Accurate Grapevine Leaf Disease Diagnosis. *Food Sci. Nutr.* **2025**, *13*, e70377. [CrossRef] [PubMed]
89. Jocher, G.; Chaurasia, A.; Qiu, J. *Ultralytics YOLOv8. Software*, Version 8.0.0; Ultralytics, Inc.: Frederick, MD, USA, 2023.
90. Wang, D.; Shang, S.; Wang, D.; He, X.; Feng, K.; Zhu, H. Plant Disease Detection and Classification Method Based on the Optimized Lightweight YOLOv5 Model. *Agriculture* **2022**, *12*, 931. [CrossRef]
91. Safonova, A.; Stiller, S.; Yordanov, M.; Ryo, M. Self-supervised learning outperforms supervised learning for crop classification by annotating only 5% of images. *Precis. Agric.* **2025**, *26*, 129. [CrossRef]
92. David, E.; Madec, S.; Sadeghi-Tehran, P.; Aasen, H.; Zheng, B.; Liu, S.; Kirchgessner, N.; Ishikawa, G.; Nagasawa, K.; Badhon, M.A.; et al. Global Wheat Head Detection (GWHD) Dataset: A Large and Diverse Dataset of High-Resolution RGB-Labelled Images to Develop and Benchmark Wheat Head Detection Methods. *Plant Phenomics* **2020**, *2020*, 3521852. [CrossRef] [PubMed]
93. Gautam, K.R.; Zhang, G.; Landwehr, N.; Adolphs, J. Machine Learning for Improvement of Thermal Conditions inside a Hybrid Ventilated Animal Building. *Comput. Electron. Agric.* **2021**, *187*, 106259. [CrossRef]
94. Ryo, M. Explainable Artificial Intelligence and Interpretable Machine Learning for Agricultural Data Analysis. *Artif. Intell. Agric.* **2022**, *6*, 257–265. [CrossRef]
95. Selvaraju, R.R.; Cogswell, M.; Das, A.; Vedantam, R.; Parikh, D.; Batra, D. Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. *Int. J. Comput. Vis.* **2020**, *128*, 336–359. [CrossRef]
96. Konecný, J.; McMahan, H.B.; Yu, F.X.; Richtárik, P.; Suresh, A.; Bacon, D. Federated Learning: Strategies for Improving Communication Efficiency. In Proceedings of the NIPS Workshop on Private Multi-Party Machine Learning, Barcelona, Spain, 9 December 2016.
97. Adkisson, M.; Kimmell, J.C.; Gupta, M.; Abdelsalam, M. Autoencoder-Based Anomaly Detection in Smart Farming Ecosystem. In Proceedings of the *2021 IEEE International Conference on Big Data (Big Data)*, Orlando, FL, USA, 15–18 December 2021; IEEE: Piscataway, NJ, USA, 2021. [CrossRef]
98. Mujkic, E.; Philipsen, M.P.; Moeslund, T.B.; Christiansen, M.P.; Ravn, O. Anomaly Detection for Agricultural Vehicles Using Autoencoders. *Sensors* **2022**, *22*, 3608. [CrossRef] [PubMed]
99. Muhammad, A.; Salman, Z.; Lee, K.; Han, D. Harnessing the Power of Diffusion Models for Plant Disease Image Augmentation. *Front. Plant Sci.* **2023**, *14*, 1280496. [CrossRef] [PubMed]
100. Li, Y.; Guo, J.; Qiu, H.; Chen, F.; Zhang, J. Denoising Diffusion Probabilistic Models and Transfer Learning for Citrus Disease Diagnosis. *Front. Plant Sci.* **2023**, *14*, 1267810. [CrossRef] [PubMed]
101. Yang, S.; Liu, Z.; Mayer, W. ShizishanGPT: An Agricultural Large Language Model Integrating Tools and Resources. In Proceedings of the Web Information Systems Engineering (WISE 2024), Doha, Qatar, 2–5 December 2024.

102. Sawant, S.; Nair, R.; Hariharan, S. Empowering Farmers with Artificial Intelligence: A Retrieval-Augmented Generation Based Large Language Model Advisory Framework. *J. Agric. Eng.* **2026**, *57*, 1908. [CrossRef]
103. Menezes, G.L.; Mazon, G.; Ferreira, R.E.P.; Cabrera, V.E.; Dorea, J.R.R. Artificial Intelligence for Livestock: A Narrative Review of the Applications of Computer Vision Systems and Large Language Models for Animal Farming. *Anim. Front.* **2024**, *14*, 42–53. [CrossRef] [PubMed]
104. Upadhyay, N.; Bhargava, A. Artificial Intelligence in Agriculture: Applications, Approaches, and Adversities Across Pre-Harvesting, Harvesting, and Post-Harvesting Phases. *Iran. J. Comput. Sci.* **2025**, *8*, 749–772. [CrossRef]
105. Tomotaki-Dawoud, K.; Chaliganti, R.; Chauhan, S.C.; Wolffram, P.; Türkes, O.; Bosse, S. Fine-Tuning the Prithvi Foundation Model for Crop-Type and Health Mapping in Smallholder Farms. *Procedia Comput. Sci.* **2025**, *270*, 5055–5064. [CrossRef]

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.