A Socio-Technical Synthesis of Artificial Intelligence in the Global Agrifood Transformation

Israt Jahan Saima College of Physical Science and Technology,Yangzhou University,Yangzhou 225009,China

ABSTRACT

The current agrifood industry is challenged by population growth and climate change, necessitating a shift towards sustainable data-driven production. Artificial intelligence (AI) is one of the most important solutions, though adoption is inconsistent in various agricultural fields. The present review discusses the recent developments in the field of agriculture that are driven by AI through the application of Preferred Reporting Items for Systematic Reviews and Meta-Analyses framework (PRISMA). The applications have been classified into four major themes, which include deep learning-based crop perception, UAV-assisted remote sensing, precision decision support systems, and autonomous robotics. A cross-study analysis on staple food crops such as Triticum aestivum (wheat), Oryza sativa (rice), Zea mays (maize) and diseases such as Magnaporthe oryzae (rice blast) and Phytophthora infestans (late blight) indicate that AI-based irrigation, disease identification, and yield prediction have accuracy levels above 85%.Some major issues to be solved include data standardization, lack of generalizability in models, and expensive initial investments for small farmers. There is an ongoing mismatch between highly technical agricultural production research and the goals of sustainability, with a focus on the manipulation, multimodal sensors, and digitalization of rural areas still understudied. The development of attention mechanisms and digital twins by current research reveals a problem of implementing such solutions at large scale. In terms of the UN SDGs, this review is a way forward towards aligning performance measures with sustainability.

ABSTRAK

Industri agrifood saat ini menghadapi tantangan akibat pertumbuhan populasi dan perubahan iklim, sehingga memerlukan pergeseran menuju produksi berbasis data yang berkelanjutan. Kecerdasan buatan (AI) merupakan salah satu solusi terpenting, meskipun adopsinya tidak konsisten di berbagai bidang pertanian. Tinjauan ini membahas perkembangan terkini di bidang pertanian yang didorong oleh AI melalui penerapan kerangka kerja Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA). Aplikasi-aplikasi tersebut telah diklasifikasikan ke dalam empat tema utama, yang meliputi persepsi tanaman berbasis deep learning, penginderaan jarak jauh dengan bantuan UAV, sistem pendukung keputusan presisi, dan robotika otonom. Analisis lintas studi pada tanaman pangan pokok seperti Triticum aestivum (gandum), Oryza sativa (padi), Zea mays (jagung) dan penyakit seperti Magnaporthe oryzae (penyakit blas padi) dan Phytophthora infestans (penyakit hawar daun)

CONTACT

Saimaisrat223@gmail.com

KEYWORDS

Artificial Intelligence, Precision Agriculture, Deep Learning, Agricultural Digital Twins, Agri-Robotics, Machine Learning, Crop Yield Prediction, Smart Farming, Socio-Technical Systems, Food Sustainability

menunjukkan bahwa irigasi berbasis AI, identifikasi penyakit, dan prediksi hasil panen memiliki tingkat akurasi di atas 85%. Beberapa masalah utama yang perlu dipecahkan meliputi standarisasi data, kurangnya kemampuan generalisasi pada model, dan investasi awal yang mahal bagi petani kecil. Masih terdapat ketidaksesuaian antara penelitian produksi pertanian yang sangat teknis dengan tujuan keberlanjutan, dengan fokus pada manipulasi, sensor multimodal, dan digitalisasi wilayah pedesaan yang masih kurang diteliti. Pengembangan mekanisme perhatian (attention mechanisms) dan kembaran digital (digital twins) oleh penelitian terkini mengungkapkan adanya kendala dalam mengimplementasikan solusi-solusi tersebut dalam skala besar. Dalam kaitannya dengan SDGs PBB, tinjauan ini merupakan langkah maju untuk menyelaraskan ukuran kinerja dengan keberlanjutan.

INTRODUCTION

The agrifood system is expected to be able to provide greater reliability in delivering nutrition while using smaller amounts of land, water, energy, chemicals, under increased climatic variability, soil degradation, labor shortage and market uncertainty. Thus, digital agriculture has evolved from focusing solely on guiding machines to establishing an integrated solution that combines sensing, forecasting, automation, and management on the farm. The key role in achieving this goal belongs to artificial intelligence (AI) which is capable of discovering patterns in heterogeneous data which cannot be handled effectively by standard rule-based analytics. Combined with agronomical knowledge, AI will allow for early disease detection, optimal application of inputs and quick adaptation to environmental changes (Basso & Antle, 2020; Liakos et al., 2018).

The existing literature on technology tends to be structured according to specific tasks. Computer vision detects plant diseases, weeds, fruit, nutrient deficiencies, and plant stresses; machine learning provides predictions for yield, irrigation needs, and management; unmanned aerial vehicles (UAVs) allow observation at the leaf level and extend it to the field level; and robots translate perception into physical action. Studies have shown rapid advances in all of these areas, especially after the emergence of cheap cameras, cloud computing power, open-source learning libraries, and large annotated databases of images (Jha et al., 2019; Kamilaris & Prenafeta-Boldú, 2018). Nonetheless, the information generated by this process is scattered. A method may provide outstanding classification rates but operate in an idealized lab setting, a genetically pure cultivar or inaccessible sensor technology to the target farmer.

Plant perception is a case in point. Convolutional neural network models developed using hand-curated repositories of leaf images have been found to recognize multiple plant diseases with accuracy approaching 100%. Mohanty et al. (2016), Ferentinos (2018), and Too et al. (2019) showed that transfer learning and deeper architectures allow for discrimination of many crop-disease pairs with very high classification rates. However, achieving high accuracy on random splits of a curated data set does not translate into reliability of diagnosis in practice when there are overlaps in symptoms, complex backgrounds, occluded leaves, varying illumination conditions, unbalanced occurrence of diseases and possibly co-existence of multiple stresses. Thus, the gap between achieving benchmarks and ensuring validity is a key issue, rather than a technical one.

Unmanned aerial vehicles enable observation at a new scale and at a decision-relevant time-point. High-resolution RGB, multispectral, hyperspectral and thermal imaging allows capturing vigor, water stress, weeds and abnormalities caused by diseases at a spatial resolution that cannot be reached with satellite imagery. Repeated observations with UAVs enable capturing data at decision-relevant times thus making UAVs a useful tool for phenotyping and site-specific management (Maes & Steppe, 2019; Tsouros et al., 2019). However, to become useful, UAV data need to be calibrated, geo-referenced, flights carried out safely, tolerate adverse weather conditions, be interpreted appropriately and comply with local regulations. Decision Support System (DSS) is an application that links predictions to actions. Yield prediction involves information about weather, soil, genotype and management. Irrigation systems incorporate soil moisture, evapotranspiration, and crop development stage, while nutrient applications rely on spatio-temporal data instead of averages. Machine-learning algorithms outperform naive statistical approaches provided that there are enough representative data, yet the perceived superiority is prone to data leakage, validation setup and variability in seasonality and geographical location

(Chlingaryan et al., 2018; van Klompenburg et al., 2020). A model that shows good performance in a certain production area may become useless once cultivar changes, unusual climate occurs or management changes. DSS should be able to provide some information on uncertainties and preserve meaningful human involvement in decision-making process.

The most challenging application of agriculture artificial intelligence is autonomous robotics. Agricultural robots have to identify a target organism, localize themselves, make a motion plan, control implements and navigate through dust, vibration, changing lighting conditions and uneven surface. There have been advancements in the development of robots for weeding, selective spraying, crop scouting and harvesting, but manipulating fragile biological products still poses challenges. Commercial viability will depend not only on detection precision, but also on factors such as throughput, repairability, safety, energy consumption and compatibility with farming practices (Fountas et al., 2020; Shamshiri et al., 2018). Digital twins expand on this by combining real-time data from sensors with models of crops, fields or machinery, which may enable prediction and simulation prior to carrying out an intervention (Pylianidis et al., 2021; Verdouw et al., 2021).

A socio-technical perspective is required because agrifood technologies are situated within institutional, market, labor, infrastructural and local contexts. The issues of data ownership, interoperability, privacy, cybersecurity and platform dependency influence the users of an AI technology and its benefits. Smart farming technologies can shift decision-making power away from the farmer in favor of equipment manufacturers, software companies, and data platforms even if the technical goal is efficiency (Klerkx et al., 2019; Wolfert et al., 2017). As such, research on responsible innovation places greater importance on the anticipation of negative effects, inclusion of all relevant parties, reflexivity, and responsiveness within the technology design process and not at the implementation stage (Eastwood et al., 2019; Rose & Chilvers, 2018).

This problem of distribution is particularly acute for smallholder farmers. The cost of acquisition of high-end sensors or robotic technologies can be too high in the context of low connectivity, unreliable infrastructure, fragmented land holdings, inadequate record keeping, uncertain payoffs, and lack of reliable services. Subscription services can help with upfront costs but can lead to dependence and recurrent costs. On the other hand, collective farming, cooperatives, public extension systems, and result-based advisory services can distribute the fixed costs and keep the services local. Thus, the critical question is not whether AI can help perform a task under ideal conditions, but under which ownership, service, and governance structures that improvement can be achieved (Birner et al., 2021; FAO, 2022).

Sustainability, however, is more than just productivity. Several AI applications maximize yield, disease detection or treatment timing, without taking into account energy usage, impact on hardware, rebound effects, labor quality or whom exactly can benefit from resource savings. An approach may save pesticides but produce more electronic waste and cloud computing needs; increase yield, while excluding farmers who cannot go digital; or make work processes automated, yet provide no ways to ensure the transition to automation for labor in safety. Agriculture 4.0 therefore must be assessed on a variety of outcomes, such as resilience, environmental impact, economics, capability building and equity (Klerkx & Rose, 2020; Saiz-Rubio & Rovira-Más, 2020).

There are three main research questions that have informed this review. First, what are the consistencies in performance observed among the deep learning perception for crops, UAV remote sensing, precision decision making and autonomy of robots? Second, what are the reasons for the persisting gap between testing and implementation in terms of the underlying technology and institutions? Third, what is the implementation sequence that would ensure AI innovations are aligned with SDGs?

The review has made three contributions to knowledge. First, it brings together four disparate technologial literatures under one evidence base; second, it analyzes the performance metrics in terms of the wider framework of generalizability, cost, infrastructure, and governance; third, it offers a sequential path that sees rural connectivity, standards of data, field testing, farmer engagement, and viable business models as co-requisites to scale. Such an approach steers clear of two extremes: first, technological determinism that believes a better business model will solve agricultural problems, and technological pessimism that overlooks actual advances in sensing, prediction, and automation technologies.

Also, changes in the development of the field affect the unit of analysis. Initially, precision agriculture involved the study of georeferenced variation at the field level and the variable rate of application of inputs. Smart farming includes continuous connectivity, cloud platforms and data-driven coordination. In turn, Agriculture 5.0 pays more attention to human-centred automation, resilience, and ecological performance. Stages of development coexist, not replace each other. Thus, one may apply computer vision while recording manually or have high connectivity with low ability to cope with uncertainty. Maturity, accordingly, is determined through sensing, data handling, decision-making, and institutional capabilities, but not with the help of one specific advanced technology (Saiz-Rubio & Rovira-Más, 2020; Tzounis et al., 2017).

In this context, the review separates four types of value. The first type is the analytical value, which refers to the detection or prediction of a target. The second one is the decision value, which implies the capability of altering a timely decision. The third type is the operational value, which concerns the actual implementation of the decision with constraints in terms of cost, safety, and manpower. The fourth type is the public value, which is associated with resiliency, sustainability, and equal accessibility. Many studies prove the analytical value; some studies prove the decision value; and very few studies prove operational or public values over seasons.

The rest of the paper is laid out according to the IMRAD CAR format stipulated by the proceedings template. In the Methodology part, the purposeful review design in line with PRISMA, along with evidence gathering, is described. In Results and Discussions, the content is structured along the four technical pillars and then the socio-technical perspective is addressed.

METHODS

A semi-systematic review methodology following the PRISMA framework was employed in this study. The aim was not to make any assertions of comprehensiveness regarding the use of databases, nor was there an attempt to compute an effect size; rather, the study aimed at employing clear criteria for inclusion, screening, data extraction and synthesis on a selected group of studies. The selected corpus of sources amounted to 34 peer-reviewed journals and two framework or policy documents. Most of the sources were from 2016 onwards, while two older sources were kept due to their relevance to the topic.

Eligibility was determined based on four criteria. The first criterion was that the publication should have dealt with AI, machine learning, deep learning, UAV analytics, agricultural robotics, digital twin technologies, or any similar data-intensive technologies for decision-making. The second criterion was that the research should relate to crop production or agrifood field management and not to any general computer science benchmark problem. The third criterion was that the research should include empirical results, synthesis of empirical research, or analysis of deployment/governance issues. Finally, the fourth criterion was that sufficient methodological information should be included in the full text so that the evidence could be understood.

The literature search has been performed using literature from the field of machine learning, deep learning, UAV remote sensing, decision support, robotics, and digital agriculture. Literature search through backward citation has been carried out for the purpose of identifying seminal empirical research whereas forward thematic chaining has also been employed for literature pertaining to the topics of digital twins, responsible innovation, and technology access. Concepts used in the literature search include artificial intelligence, machine learning, deep learning, precision agriculture, crop disease, yield prediction, irrigation, UAV, agricultural robot, digital twin, smart farming, sustainability, smallholder and data governance.

The extraction form documented the year of publication, area of operation, agricultural/biological application, sensor/data type, algorithmic strategy, validation procedure, performance reported, decision function and any deployment constraints. Reported accuracy, success rate and comparison with models were kept as originally published, without any efforts at conversion or analysis. The critical factor in validating models was whether it was done on curated images, independent data sets, multiple seasons or at separate locations.

The thematic synthesis had two phases. First, technical evidence was clustered into four themes: deep-learning crop recognition, UAV-based remote sensing, precision decision-making support, and autonomous robotic systems with digital-twin technology. Next, the evidence was classified by five socio-technical categories: data and interoperability, infrastructure and maintenance, human skills and trust, economic feasibility, and governance or distributional impacts. Both the highest possible performance and the preconditions of reproduction of value were compared by the synthesis.

A qualitative evaluation of the evidence quality was conducted in each empirical claim. More robust evidence had separate training and testing data, external testing sites or seasons, explicit description of the sample, correct metrics and a comparator as a state-of-the-art. Poor-quality evidence was based only on internal random splits, lacked information on the class distribution, provided just one positive result or assessed an algorithm in isolation from the workflow. Review papers were employed to define a domain and recurrent results. Primary studies took precedence for numerical performance in Table 1.

For consistency of interpretation, triangulation was conducted between technical and social science claims. The technical side would be a robotics paper on the state of the art in navigation and actuation, while the social science side would be the responsible innovation literature on labor, safety, and governance issues which dictate adoption. These findings were not averaged out. Instead, the differences were explored in terms of crop type, sensor type, validation method, farm size, and farm infrastructure.

However, there are limitations with this approach. The data set is purposeful in nature and heavily relies on English-based publications, which means that there is a possibility of under-representation of local language implementation reports as well as unpublished negative findings. The diversity of crop types, sensors, indicators and validation techniques used does not allow for meta-analysis to be performed. It is also clear that there is publication bias, as high benchmark scores are more prominent than failures.

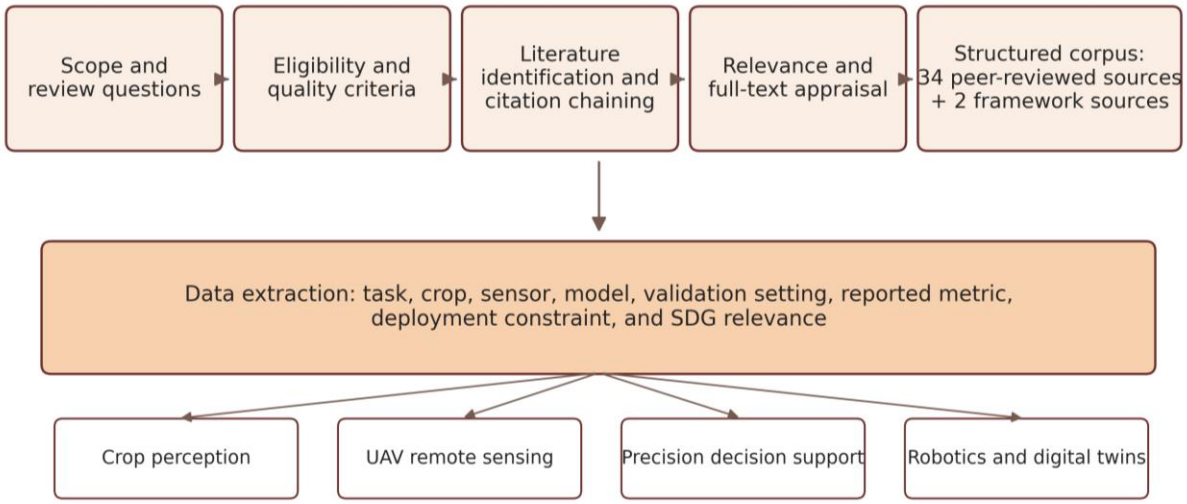


Figure 1. PRISMA-aligned evidence identification, extraction and thematic synthesis workflow.

RESULTS AND DISCUSSIONS

The findings from the literature reveal a lot of progress within algorithms but also an inherent trend: peak performance can only be achieved through well-controlled tasks, sensors, and environments. Figure 2 will capture the relationship among data sources, the four technical foundations, socio-technical governance and agrifood results. Table 1 contains sample evidence used during the synthesis.

Deep-Learning-Based Crop Perception

Deep learning-based crop perception is the most advanced technology regarding benchmark size. Deep learning models learn hierarchical features based on pixels and are independent of hand-crafted descriptors for colors, textures, and shapes. Transfer learning also helps researchers to transfer pre-trained deep learning models, which were trained using large image databases to a smaller agricultural data set. The studies conducted by the Plant Village proved their importance because they indicated that a single model was capable of detecting many different categories of crops in healthy and infected states with high accuracy (Mohanty et al., 2016; Too et al., 2019).

The key limitation is domain shift. The curated leaves will be centrally aligned, properly lit, and photographed against plain backgrounds, but the field images would have soil, stems, adjacent leaves, shadows, water droplets, insect damage, and other disease symptoms. Splitting randomly selected near duplicate images collected in the same way into the training and testing datasets would further boost the performance. Hence, the operational testing should be done on external farms, previously unseen cultivars, different instruments, actual incidence of disease, and multiple seasons. Performance measures should cover sensitivity to important diseases, confidence calibration, cost of false alarm, and time to diagnose.

Table 1. Representative evidence and the difference between source-reported performance and deployment requirements.

Study	Domain/Data	AI Approach	Reported Result	Deployment Interpretation
Mohanty et al. (2016)	PlantVillage dataset: 54,306 images across 38 crop–disease classes	Transfer-learned convolutional neural network (CNN)	Achieved 99.35% accuracy on the held-out test set	Performance declined substantially on external and field images, indicating that background and image-acquisition biases affect real-world generalization.
Ferentinos (2018)	87,848 images representing 58 plant-condition classes	Deep convolutional neural networks	Best model achieved a success rate of 99.53%	High benchmark performance does not guarantee field readiness; validation across farms, cultivars, environmental conditions, and imaging devices is still required.
Too et al. (2019)	PlantVillage dataset with 38 crop–disease classes	Fine-tuned DenseNet-169	Achieved 99.75% test accuracy	Independent farm data and naturally occurring disease prevalence should be evaluated before making reliable diagnostic claims.
Khaki and Wang (2019)	Multi-environment maize-yield records	Deep neural network	Outperformed LASSO, regression-tree, and shallow neural-network baselines	Model performance depends heavily on representative genotype, weather, soil, and crop-management data.
van Klompenburg et al. (2020)	Systematic review of 50 crop-yield prediction studies	Random forest, artificial neural networks, CNNs, and related machine-learning methods	Weather, soil, and management variables were repeatedly identified as important predictors	Considerable variation in validation methods and reporting practices limited direct comparison among studies.
Tsouros et al. (2019)	Literature on unmanned aerial vehicles in precision agriculture	RGB, multispectral, and thermal-image analytics	Demonstrated applications in crop monitoring, disease detection, weed mapping, and irrigation assessment	Repeated operational use is constrained by sensor calibration, regulations, weather conditions, and integration with farm workflows.
Fountas et al. (2020)	Literature on agricultural field robotics	Computer vision, autonomous navigation, and robotic actuation	Reported substantial progress in weeding, crop scouting, and harvesting prototypes	Commercial readiness depends on reliability, safety, field capacity, maintenance requirements, and economic feasibility.
Pylianidis et al. (2021); Verdouw et al. (2021)	Literature on agricultural digital twins	Internet of Things-linked dynamic models	Identified strong potential for prediction, simulation, monitoring, and adaptive control	Operational digital twins remain limited by fragmented data systems, inconsistent standards, and poor model interoperability.

The way the task is defined is important as well. Classification of an image presupposes that the necessary leaf or organ has been selected already, while on-field use usually implies the need for object detection and segmentation in order to locate several symptoms within a noisy environment. The problem of severity is even harder, since it needs accurate measuring and clinical thresholds. The annotation process needs to separate disease from other factors such as nutrient deficiencies, insect damage and abiotic stress, and uncertain cases should be taken into account, not forced into a particular category.

Explanability can be used to assist validation but cannot replace it. Saliency maps will show whether the neural network pays attention to the lesion or some other background information, while counterfactuals will demonstrate the dependence on lighting or size. Edge computing can help speed up the process and cut out the dependency on connectivity, but the accuracy of a compressed model needs to be checked in relation to devices and image quality. The best option is a system of triaging rather than an autonomous one – it describes the case, its level of confidence and when the intervention from an expert is needed.

Attention and transformers can help in integrating long-range dependencies and multimodal inputs, but they still require a representative dataset. The realistic path is the combination of a diagnostic process: AI makes quick checks and, based on thresholds of uncertainty, sends problematic cases to agronomists, and the cases are sent back to training system as confirmed ones. Such a loop can transform model errors into supervised learning.

UAV-Enabled Remote Sensing

UAV-based remote sensing extends AI beyond object detection to spatial decision making. RGB sensors facilitate plant counts and lodging scores; multispectral sensors allow canopy vigor measurements and chlorophyll variations; thermal sensors may be able to detect water stress; and hyperspectral sensors can measure biochemical signatures. UAVs can offer better spatial resolution and more convenient timing than satellite data for experimental plots and heterogeneous fields (Maes & Steppe, 2019; Sishodia et al., 2020). Machine learning can produce weed maps, stress zones, biomass assessments or prescriptions from orthomosaics.

There are three common issues in translation. First, images taken at different times have different lighting conditions and hence cannot be compared directly because sensor temperatures, camera altitudes and calibrations are not equal. Second, decisions are only useful if a complete process from planning and stitching to agricultural analysis and field operations are available. And third, legal constraints, weather, power supply and operator expertise can prevent repeated monitoring. UAV-based research can thus be more impactful if done backwards from decisions, such as scouting areas, irrigation and spraying, rather than forwards from imagery (Tsouros et al., 2019).

Multimodal data fusion is attractive since no single type of sensor detects the reason behind the stress phenomenon. The presence of thermal abnormalities may imply water deficiency, but such abnormalities can also be caused by diseases, tree canopy structure, and atmospheric conditions. Multimodal data, including thermal data, multispectral data, weather parameters and soil measurements, can eliminate this ambiguity if the sensors' data is collected simultaneously and ground information is also available. Field sampling is important both for training and testing purposes; UAV data should serve to detect where to sample and not be self-explanatory.

Precision Decision-Support Systems

Decision support systems based on precision agriculture are temporal integration of observations. Models of yield prediction include genotype, environment and management factors, models of irrigation predict crop water needs and nutrient models derive spatial responses from data on soil, canopy and weather conditions. Review studies confirm common use of random forests, artificial neural networks, support-vector machines and recently deep learning methods (Benos et al., 2021; van Klompenburg et al., 2020). Khaki and Wang (2019) have shown that deep neural networks can learn nonlinear interactions between genotypes and environments and are better than baseline algorithms in yield prediction in maize.

What is needed from the farmer's point of view is not necessarily the model with the lowest retrospective prediction error. Recommendations should be timely, interpretable, robust to missing data and linked to farm management practices. The estimates of the yield may be useless because it comes after input decisions, while the recommendations about irrigation have to consider the price of water, pumping capacity and risk aversion. Thus, decision support systems have to show the uncertainties, data provenance and allow scenario comparisons. Integration with farm management information systems and adoption of standards is a bottleneck (Tzounis et al., 2017; Zhai et al.,2020).

Prediction needs to be disentangled from causation recommendation as well. While the model might be able to link a high application of nitrogen to high yields as fertile farms have been using more fertilizers in the past, it doesn't

mean that by applying fertilizers to any farm, you would boost its productivity. Prescriptive models need field testing, either causal understanding or experiments to gauge the effect of interventions. Economic thresholds must allow converting uncertainty into decision-making: when to water, when to scout, when to spray or wait.

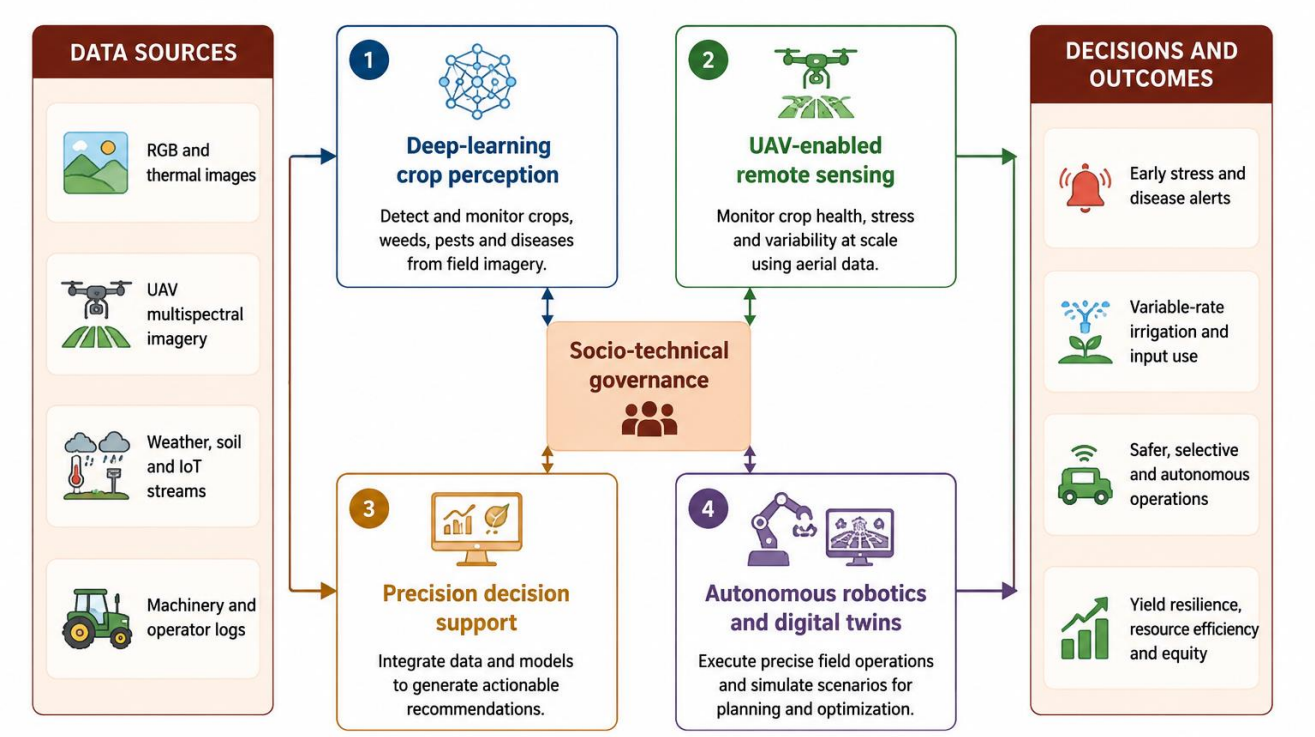


Figure 2. Four-pillar socio-technical architecture for AI-enabled agrifood decisions and outcomes.

Autonomous Robotics and Agricultural Digital Twins

Autonomous robots turn perception and prediction into actions. Robotic weeders can weed mechanically, electrically or using micro-spraying of herbicides; robotic scouts can measure repeatedly in high resolution; and robotic harvesters integrate perception of fruit location, estimation of fruit pose, motion planning and action on the fruit. These systems can potentially minimize human exposure and allow selective treatment, but field conditions present an especially challenging reliability problem, since the interaction between dust, mud, vibrations, irregular terrain, occlusions and biological variation is hard to simulate in a lab setting (Fountas et al., 2020; Shamshiri et al., 2018).

Gentle manipulation is poorly researched compared to fruit detection. It is possible for a detector to estimate fruit location correctly, but the manipulator to bruise the fruit, miss a partially occluded stem or act too slowly to keep up with commercial harvesting pace. Thus, development of these technologies requires design of sensing, soft materials, control and work process in parallel. Also, safety certifications, teleoperation, maintenance and liability issues are a must where humans, animals and vehicles operate alongside robots.

Utilization and structure of robots impact robot economics. While a single-function device can do its job extraordinarily well but sit around unused the rest of the year, a versatile design will spread costs of navigation, energy, and maintenance among functions such as reconnaissance, weeding, and transport. Robot fleets can help with redundancy and less soil compaction, but they require more coordination and higher cybersecurity needs. Comparative studies should present hectares per hour, success rate of interventions, downtime, energy consumption, supplies used, supervision and maintenance time, so that the total cost could be compared with human labor and standard machinery.

Automation also transforms work as opposed to eliminating it. It can minimize tedious or dangerous exposure, but at the same time, it can create responsibilities such as monitoring, labeling data, or maintaining equipment, or shift control to the outsourced service providers. The transition plan has to address who gets skilled and who is taking new risks and whether unemployed workers have access to more rewarding jobs. This employment effect has to be included into the performance evaluation, as untrustworthy automation always transfers its workload to people.

Digital Twins Offer a Coordination Layer Option. Digital twins are not just simulations, but rather involve a continuous interaction between a physical object and its digital counterpart via information flows. In the case of

agricultural applications, a twin could include a crop, a greenhouse, a field, a machinery or logistics system, combining sensor inputs with mechanistic and learning models. The digital twins could help with what-if scenario development, predictive maintenance and adaptive control (Pylianidis et al., 2021; Verdouw et al., 2021). However, most agricultural twins are partial due to incomplete data, different scale of models and the high cost of synchronization in real time.

From Benchmark Performance to Deployment Readiness

In all four pillars, 85% accuracy, frequently cited as a rule of thumb, is not necessarily a measure of readiness across all tasks. An error means different things depending on the task at hand: the error of detecting a non-existent disease leads to excessive spraying; missing obstacles results in injury; the bias in yield estimation leads to distorted credit and insurance claims; and a tiny error in irrigation becomes a large error in water usage. It is, thus, necessary to assess performance of an algorithm in the context of its real-world agronomic and social impacts.

The constraints of data quality and standardization stand out as the most powerful common constraints. There is a lot of agricultural data collected in different seasons using numerous devices, and there are different labels and units of measure, different coordinate systems and data ownership rules that cause problems for the integration process. The big companies can afford the costs of the process of integration, but small farmers cannot. The open standards, interfaces, portable records, ontologies and consent will help to avoid lock-in and allow innovations. At the same time, it should not mean that anyone can extract the farm data freely.

Generalizability of AI models is likewise a systems problem. Increased complexity can help fit benchmarks better while adding to the computational cost, energy requirements and opaqueness. Attention architectures and multimodal learning are promising as they can incorporate image, meteorology, soil and management features but require sampling and validation to be meaningful. In regions lacking adequate data, an interpretable model with local calibration may provide higher societal benefit than a powerful model trained elsewhere.

An evidence ladder with four steps would help make claims comparable. Step 1 consists of internal benchmarking with clean separation between the training and validation datasets. Step 2 involves external validation on unseen farms, growing seasons, crop varieties or devices. Step 3 is the prospective field trial where decisions based on the AI results are compared with standard practice. Step 4 involves operational evaluation in terms of reliability, cost, maintenance, safety, environment and distribution of benefits. Papers must declare the level they reached. This would avoid misleading reporting of Level 1 precision as being Level 4 ready.

A Socio-Technical Implementation Roadmap

The results call for a phased approach to implementation in lieu of a technology-driven procurement process. The first phase focuses on creating the basic infrastructure required for rural digitization—stable electricity, decent connectivity, maintainable sensors, calibration of devices and local technical assistance. These are no trivial concerns. Cloud computing cannot offer timely advice regarding irrigation if network connectivity is patchy and faulty sensors could skew every subsequent prediction silently.

The second phase is about setting up interoperable and accountable data agreements. This involves defining common variables, units and metadata, data testing for quality and documenting sensor provenance. Access, portability, correction and deletion rights need to be determined at the outset as cybersecurity needs to be taken into account due to the risk of operational vulnerability arising from connected machines and irrigation systems.

Stage 3 tests the entire decision process through field trials. These need to be cross-seasonal, cross-geographical, cross-varietal and cross-instrument type in order to compare the artificial intelligence-enabled decision process with the existing one. Reports need to take into account any cases of failure, uncertainties, labor hours, maintenance requirements, cost and the distribution of benefits. The human-centric approach is key: farmers, agronomists, operators and extension workers need to contribute to deciding on alert criteria, presentation form and alternatives for the cases when the algorithm is unreliable.

Stage 4 builds up the necessary skills, finance and service eco-systems. Purchasing the equipment upfront is just one way to go. Robotic systems that could be rented out by hour, hectare or some other measure of area, UAV fleets organized within cooperatives and advisory packages all allow for translating the initial cost into the operational one. However, agreements need to guarantee data portability and transparency.

The roles of the institutions need to be made clear through all the stages. The public institutions can invest in connectivity, openness, extension capacity, and independent evaluation; universities can create regional benchmarking stations and build interdisciplinary teams; the cooperatives can build up demand and negotiate data rights; and the vendors can deliver service-level agreements, interoperable interfaces, and update policies. The procurement process needs to value results and portability more than proprietary capabilities. Long-term programs need to include calibration, replacement, cybersecurity, and user support, as the pilots will often fail when the initial grants run out.

Only stage 5 scales once public-value measures are defined. The productivity benefits should be counted in combination with water, nutrient, and pesticide utilization, energy considerations, greenhouse gas effects, worker safety, labor quality, cost, and accessibility by farm size and by gender. This helps avoid confusion of adoption and success. Figure 3 provides an overview of the process: infrastructure leads to governance, which allows reliable validation, which leads to sustainable services.

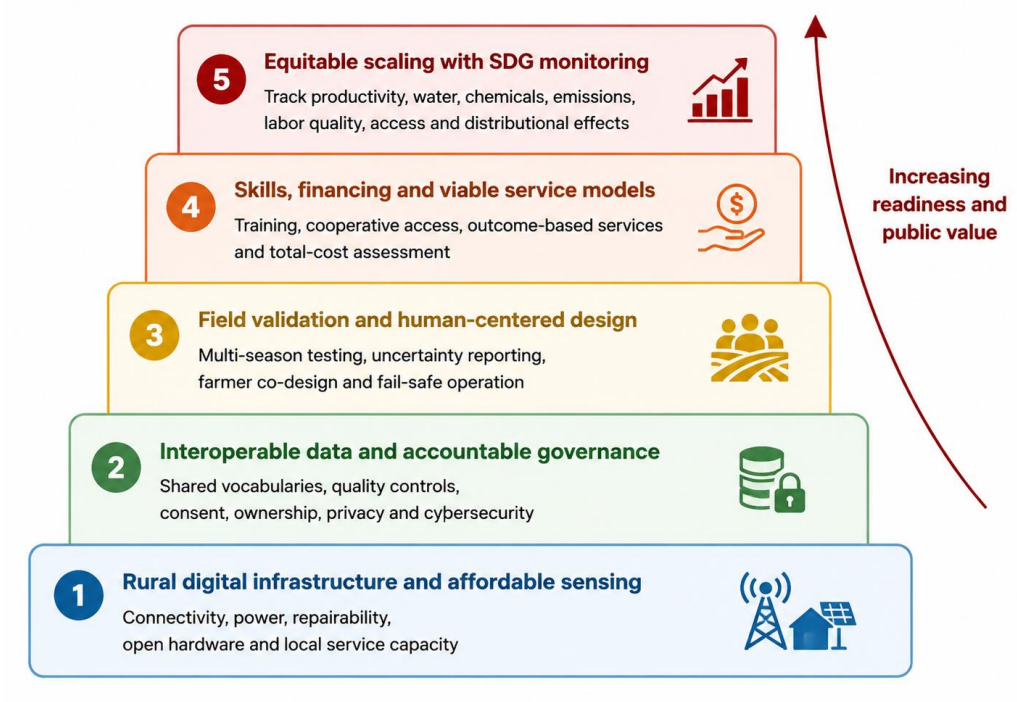


Figure 3. Five-stage roadmap for scalable, equitable and accountable agricultural AI implementation.

Alignment With the Sustainable Development Goals

The plan corresponds with SDG 2 (Zero Hunger) directly because of resilient productivity and minimized losses of crops; SDG 6 because of more efficient water management; SDG 8 through enhanced safety and capacity building; SDG 9 through rural infrastructure for innovation; SDG 12 through optimized resource management and the management of lifecycles of equipment; SDG 13 by adapting to climate change and managing emissions; and SDG 10 explicitly because the impact on access will be monitored (United Nations, 2015). This relationship is contingent on a number of factors. AI helps to meet an SDG as long as the operational improvement contributes to meeting the objective without shifting costs or risks disproportionately elsewhere.

An initial dashboard should thus comprise technical and socio-economic performance metrics. Technical metrics include external accuracy, calibration, uptime, latency, and intervention effectiveness. Agronomic metrics include yield stability, loss prevention, water efficiency, and input savings. Economic metrics include total cost of ownership, payback, availability, and sensitivity to farm size. Social metrics include user trust, training effort, workload, safety, data control, and access by underprivileged farmers. Environmental metrics include energy consumption, chemical leakage, emission and end-of-life management of sensors, batteries, and electronics.

Trade-offs and rebound effects need to be tracked. Increased irrigation efficiency may decrease water usage per kilogram but lead to a larger area of irrigated land; cheaper disease testing could result in more spraying due to incorrect calibration of alerts; and intensive monitoring at a high resolution could increase energy and e-waste impacts. Life cycle assessment and counterfactual comparison will thus have to be performed to evaluate whether an intervention decreases the aggregate pressure on the environment. Similarly, distributional analysis will have to evaluate whether benefits flow to small and distant farms or merely consolidate benefits of the digitized farms.

The heterogeneity and the selected purposive corpus limits this review. It cannot give a pooled accuracy estimate, and the source reported results could have some biases due to publication bias. The research is less focused on livestock, fisheries, food processing and logistics compared to crop production. There should be a replication of the screening process in different bibliographic databases, use of local language data, registration of field experiments, reporting of negative findings, and development of benchmark datasets.

CONCLUSIONS

At present, artificial intelligence enables high-performing crop perception, spatial sensing, prediction and automation. While best controlled experiments achieve well over 85% accuracy – most notably in plant disease classification – these figures cannot be used to prove the readiness for field deployment. In the entire literature, the bottlenecks are representative data, external validation, sensor calibration, interoperability, connectivity, cost, maintenance, user capabilities and governance.

The four technical pillars complement each other. Crop perception detects biological states; UAVs map them; decision support systems process information into recommendations; robots and digital twins execute and orchestrate actions. Their benefits are realized only in the presence of people, organizations and infrastructure within the system. The outlined five-step plan starts from building rural digital infrastructure and data governance, moves on to field validation and realistic business models, and scales up with multidimensional SDG tracking.

ACKNOWLEDGEMENTS

The authors acknowledge the researchers whose empirical and conceptual work formed the evidence base for this synthesis. No external funding was declared for this draft.

REFERENCES

Ayaz, M., Ammad-Uddin, M., Sharif, Z., Mansour, A., & Aggoune, E.-H. M. (2019). Internet-of-Things (IoT)-based smart agriculture: Toward making the fields talk. IEEE Access, 7, 129551-129583.

Basso, B., & Antle, J. (2020). Digital agriculture to design sustainable agricultural systems. Nature Sustainability, 3, 254-256.

Benos, L., Tagarakis, A. C., Dolias, G., Berruto, R., Kateris, D., & Bochtis, D. (2021). Machine learning in agriculture: A comprehensive updated review. Sensors, 21(11), 3758.

Birner, R., Daum, T., & Pray, C. (2021). Who drives the digital revolution in agriculture? A review of supply-side trends, players and challenges. Applied Economic Perspectives and Policy, 43(4), 1260-1285.

Chlingaryan, A., Sukkarieh, S., & Whelan, B. (2018). Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review. Computers and Electronics in Agriculture, 151, 61-69.

Coble, K. H., Mishra, A. K., Ferrell, S., & Griffin, T. (2018). Big data in agriculture: A challenge for the future. Applied Economic Perspectives and Policy, 40(1), 79-96.

Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., et al. (2021). An image is worth 16 x 16 words: Transformers for image recognition at scale. Proceedings of the International Conference on Learning Representations.

Eastwood, C., Klerkx, L., Ayre, M., & Dela Rue, B. (2019). Managing socio-ethical challenges in the development of smart farming: From a fragmented to a comprehensive approach for responsible research and innovation. Journal of Agricultural and Environmental Ethics, 32, 741-768.

Food and Agriculture Organization of the United Nations. (2022). The state of food and agriculture 2022: Leveraging automation in agriculture for transforming agrifood systems. FAO.

Ferentinos, K. P. (2018). Deep learning models for plant disease detection and diagnosis. Computers and Electronics in Agriculture, 145, 311-318.

Fountas, S., Mylonas, N., Malounas, I., Rodias, E., Santos, C. H., & Pekkeriet, E. (2020). Agricultural robotics for field operations. Sensors, 20(9), 2672.

Gebbers, R., & Adamchuk, V. I. (2010). Precision agriculture and food security. Science, 327(5967), 828-831.

Jha, K., Doshi, A., Patel, P., & Shah, M. (2019). A comprehensive review on automation in agriculture using artificial intelligence. Artificial Intelligence in Agriculture, 2, 1-12.

Kamilaris, A., & Prenafeta-Boldú, F. X. (2018). Deep learning in agriculture: A survey. Computers and Electronics in Agriculture, 147, 70-90.

Khaki, S., & Wang, L. (2019). Crop yield prediction using deep neural networks. Frontiers in Plant Science, 10, 621.
Klerkx, L., Jakku, E., & Labarthe, P. (2019). A review of social science on digital agriculture, smart farming and agriculture 4.0: New contributions and a future research agenda. NJAS - Wageningen Journal of Life Sciences, 90-91, 100315.
Klerkx, L., & Rose, D. (2020). Dealing with the game-changing technologies of Agriculture 4.0: How do we manage diversity and responsibility in food system transition pathways? Global Food Security, 24, 100347.
Koirala, A., Walsh, K. B., Wang, Z., & McCarthy, C. (2019). Deep learning - Method overview and review of use for fruit detection and yield estimation. Computers and Electronics in Agriculture, 162, 219-234.
Liakos, K. G., Busato, P., Moshou, D., Pearson, S., & Bochtis, D. (2018). Machine learning in agriculture: A review. Sensors, 18(8), 2674.
Maes, W. H., & Steppe, K. (2019). Perspectives for remote sensing with unmanned aerial vehicles in precision agriculture. Trends in Plant Science, 24(2), 152-164.
Mohanty, S. P., Hughes, D. P., & Salathé, M. (2016). Using deep learning for image-based plant disease detection. Frontiers in Plant Science, 7, 1419.
Pylianidis, C., Osinga, S., & Athanasiadis, I. N. (2021). Introducing digital twins to agriculture. Computers and Electronics in Agriculture, 184, 105942.
Rose, D. C., & Chilvers, J. (2018). Agriculture 4.0: Broadening responsible innovation in an era of smart farming. Frontiers in Sustainable Food Systems, 2, 87.
Saiz-Rubio, V., & Rovira-Más, F. (2020). From smart farming towards agriculture 5.0: A review on crop data management. Agronomy, 10(2), 207.
Shamshiri, R. R., Weltzien, C., Hameed, I. A., Yule, I. J., Grift, T. E., Balasundram, S. K., Pitonakova, L., Ahmad, D., & Chowdhary, G. (2018). Research and development in agricultural robotics: A perspective of digital farming. International Journal of Agricultural and Biological Engineering, 11(4), 1-14.
Singh, A., Ganapathysubramanian, B., Singh, A. K., & Sarkar, S. (2018). Machine learning for high-throughput stress phenotyping in plants. Trends in Plant Science, 23(10), 883-898.
Sishodia, R. P., Ray, R. L., & Singh, S. K. (2020). Applications of remote sensing in precision agriculture: A review. Remote Sensing, 12(19), 3136.
Talaviya, T., Shah, D., Patel, N., Yagnik, H., & Shah, M. (2020). Implementation of artificial intelligence in agriculture for optimisation of irrigation and application of pesticides and herbicides. Artificial Intelligence in Agriculture, 4, 58-73.
Too, E. C., Yujian, L., Njuki, S., & Yingchun, L. (2019). A comparative study of fine-tuning deep learning models for plant disease identification. Computers and Electronics in Agriculture, 161, 272-279.
Tsouros, D. C., Bibi, S., & Sarigiannidis, P. G. (2019). A review on UAV-based applications for precision agriculture. Information, 10(11), 349.
Tzounis, A., Katsoulas, N., Bartzanas, T., & Kittas, C. (2017). Internet of Things in agriculture, recent advances and future challenges. Biosystems Engineering, 164, 31-48.
United Nations. (2015). Transforming our world: The 2030 agenda for sustainable development. United Nations.
van Klompenburg, T., Kassahun, A., & Catal, C. (2020). Crop yield prediction using machine learning: A systematic literature review. Computers and Electronics in Agriculture, 177, 105709.
Verdouw, C., Tekinerdogan, B., Beulens, A., & Wolfert, S. (2021). Digital twins in smart farming. Agricultural Systems, 189, 103046.
Wolfert, S., Ge, L., Verdouw, C., & Bogaardt, M.-J. (2017). Big data in smart farming - A review. Agricultural Systems, 153, 69-80.
Zhai, Z., Martínez, J. F., Beltran, V., & Martínez, N. L. (2020). Decision support systems for agriculture 4.0: Survey and challenges. Computers and Electronics in Agriculture, 170, 105256.