International Conference on Multidisciplinary Perspectives in Advanced Computing and Technology (IMPACT 2026)
G. B. Pant University of Agriculture and Technology, Uttarakhand, India. Jan. 10-11, 2026
DMPedia Lecture Notes in Computer Science & Engineering
Volume 1, pp.507-523, 2026

AI and Machine Learning Applications in Smart Agriculture: A Comprehensive Review

Vicky Kumar

Department of Information Technology, Govind Ballabh Pant University of Agricultural and Technology, Pantnagar, Uttarakhand, India

vickykanojiya968@gmail.com, 62651gbpuat.ac.in

Abstract

Agriculture is undergoing a significant technological change through the use of artificial intelligence (AI) systems and machine learning (ML) techniques. The global demand for food security, sustainable practices, and environmental protection has created a need for intelligent systems that can produce the highest yields with the fewest resources. Machine learning has, over the last 10 years, become a major component of various agricultural fields, including yield forecasting, disease identification, irrigation management, and soil quality assessment. Various supervised and unsupervised learning methods, including convolutional neural networks (CNNs), long short-term memory (LSTMs), support vector machines (SVMs), and random forests, have been used by scientists, yielding prediction accuracies far beyond those of traditional statistical approaches. Deep learning, especially hybrid CNN–LSTM architectures, is very promising for detecting crop stress, nutrient deficiency, and yield variation, and can even be done in real time. The current review echoes the findings of 40 landmark research works that trace precision agriculture through machine learning from 2021 to 2025 and examine how it has contributed to the development of sustainable farming methods. Besides, it substantiates the existence of fixed obstacles, such as data heterogeneity, scalability, and ethical data governance; it also pinpoints new and forthcoming research on AI-IoT convergence and multi-sensor fusion frameworks aimed at generating resilient, resource-efficient agricultural ecosystems.

Keywords: *Machine Learning, Agriculture, LSTM, Artificial Intelligence. CNN, Deep learning, Remote Sensing.*

1. Introduction

Farming systems worldwide are under growing pressure to produce more food, even as environmental conditions continue to worsen. At the same time, climate change, declining soil quality, and increasing pressure on freshwater resources are making agricultural production more challenging. According to estimates by the United Nations, global food production will need to increase by nearly 70% to support a population expected to cross 9.6 billion by 2050. Under these conditions, traditional farming practices are no longer sufficient to meet future food demands in a sustainable way [1-4], [9], [13].

In this context, machine learning (ML), as a core component of artificial intelligence (AI), has emerged as a powerful tool for improving crop management and agricultural decision-making. Data collected in real time from IoT sensors, satellites, drones, and weather monitoring systems are increasingly combined with ML-based analytical models to support more informed farming practices. These systems help farmers make accurate decisions related to sowing, irrigation, and resource allocation based on actual field conditions rather than assumptions or past experience [3], [12], [15], [16]. As a result, agriculture is gradually shifting from intuition-based practices to data-driven approaches that align closely with big data technologies, leading to improved efficiency and enhanced sustainability across the different stages of agricultural production [1], [9].

1.1 Role of Machine Learning in Modern Agriculture

Machine learning is no longer the simple linear regression model it was at the beginning, when it was used for yield forecasting; now it employs complex deep-learning architectures that understand not only the spatial but also the temporal dynamics of agricultural ecosystems. [1], [6], [9], [18] Digitised farming applications' fundamental technologies include such algorithms as Random Forest (RF), Support Vector Machine (SVM), Artificial Neural Networks (ANNs), and Convolutional Neural Networks (CNNs). [1],[2],[6],[15]. These techniques help improve the accuracy of crop classification, yield measurement, and disease prediction, achieving much higher precision than traditional models. [1], [2], [6], [7].

Moreover, machine learning is helpful in precision irrigation, fertiliser optimisation, and integrated pest management, which leads to less waste and more productivity. [1], [2], [3], [4] As for the precision agriculture part, machine learning is one of the elements that work hand in hand with remote sensing and geographical information systems (GIS) to make sense of almost unlimited data sets - for example, those containing information about soil moisture, vegetation indices (NDVI, EVI), and crop phenotypic traits - by doing so, they not only facilitate but also make possible the implementation of farm management strategies that are both well-informed and localized. [1], [2], [3], [4]

1.2 Evolution of Research in Precision Agriculture

The study of ML techniques applicable to agriculture has been ongoing since the 1980s, and it was then that simple regression and rule-based systems were first documented. [3], [4], [12], [19]. Nevertheless, since 2018, the massive leap in computing power, along with the effects of sensor and big-data analytics, has been the main driving force behind a rapid transition to scenarios of increased utilisation of advanced models. [3], [12], [16]

1.3 Objectives of the Review

The purpose of this review is to consolidate the current state of research on the role of ML in transforming modern agriculture. Specifically, it seeks to

- Examine the leading ML algorithms and architectures used for soil assessment, yield prediction, and irrigation management. [1], [3], [4], [6]
- Compare hybrid frameworks that integrate AI, IoT, and remote sensing for precision farming.
- Identify ongoing challenges related to scalability, model interpretability, and data privacy. [11]

- Outline future directions in deep transfer learning, federated learning, and multi-sensor fusion for developing robust, adaptive, and sustainable farming systems. [3], [11], [12], [16]

2. Machine Learning Techniques in Agriculture

2.1 Overview of Algorithms

Machine learning (ML) strategies in farm land can be broadly divided into three classes, viz. supervised, unsupervised and reinforcement learning. [1],[6],[17],[18] Out of these, the most frequent case is that of supervised learning because of its potency in predictive and classification operations, for instance, yield estimation, disease detection, and crop type identification. [1], [2], [6], [7].

- Supervised Learning: To perform image classification, yield forecasting, and pest detection, researchers have employed algorithms like Support Vector Machines (SVM), Random Forests (RF), and Artificial Neural Networks (ANNs). [1],[2],[6],[7]. In doing so, these models acquire the capacity to achieve a certain level of reliability in the prediction of results, even when the available datasets are limited or partly labelled, especially for nutrient mapping in soil and pest population analysis. [1], [2], [3], [4].
- Unsupervised Learning: K-Means Clustering and Principal Component Analysis (PCA) are just two of the many techniques that can be used to expose hidden relations in datasets of large size and complexity. [1],[6],[17],[18] Besides, they take part in soil categorization, climate pattern analysis, and genotype clustering, where the presence of labelled data is improbable. [3], [4], [8], [9].
- Reinforcement Learning (RL): RL is a new paradigm where the emphasis is placed upon systems that, through the use of feedback and reward mechanisms, learn the optimal strategies. [1], [6], [17], [18] In the case of agriculture, the aforementioned methods have contributed to irrigation management, pesticide spraying optimisation, and autonomous drone navigation. [3],[4],[5],[7] The incorporation of Q-Learning and Proximal Policy Optimisation (PPO) into the frameworks extends the benefit of resource efficiency over a longer period, demonstrating sustainability in precision farming. [1], [2], [18].Besides, deep learning—a major subarea of machine learning—comes up with models like Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM), and Transformers. [2], [6], [15], [17] The deep structures, in fact, were able to outperform the shallow ones in terms of accuracy in image-based disease diagnosis, remote sensing analytics, and phenotypic trait extraction. [2], [4], [5], [6].

2.2 Applications and Performance Metrics

Convolutional Neural Networks (CNNs) are the core of the image processing pipeline in the Ag-Tech sector, and one of their main applications is plant disease identification. [2],[6],[7],15]

Once trained, CNNs are capable of mastering the low-level to high-level features of input images and consequently, recognizing pathological patterns in a pixel-wise manner. [1],[6],[17],[18] For instance, CNNs trained on the PlantVillage dataset were able to classify the diseases on tomato and rice leaves with an accuracy of more than 95%. [1][6][17][18] The use of VGGNet, ResNet, and DenseNet-based architectures has become a common practice for leaf phenotyping as well as for crop classification from UAV imagery. [2],[5],[6],[7]

Random Forest (RF) and Gradient Boosted Trees (GBT) are machine learning algorithms that focus on maximizing the prediction accuracy of a certain parameter/value and, as a result, lead to excellent solutions for yield forecasting given a plethora of data such as the weather trends, temperature, and soil characteristics. [1],[3],[4],[6] A report cites that with RF models, the error rates can be dropped to almost 1/5 of the linear regression ones in the case of wheat yield estimation. [1],[6],[9],[18]

Long Short-Term Memory (LSTM) units are a recurrent network that is able to remember and utilize the information of the past to the current and future time steps, and this makes them promising candidates for the analysis of temporal series collected from a sensor network. [2],[3],[6],[12] Deploying LSTM models in the realm of the forecasting of maize and rice yields has brought about a 12–15% increment in accuracy compared to the classical ARIMA method. [2],[6],[15],[17]

Various statistical metrics such as Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Coefficient of Determination (R^2) are used to assess the quality of the model output. [1],[2],[18] The algorithm involved in the solution should be the deciding factor for the use of the performance metric. [1],[6],[17],[18]. For instance, image-based applications are best handled by CNNs; thus, evaluating their performance using metrics like accuracy and F1-score makes sense, whereas RF and GBT models that are numerically oriented and used for regression-based predictions entail the use of RMSE, MAE, and R^2 metrics. [1], [2], [18], [19].

2.3 Hybrid and Ensemble Frameworks

Hybrid learning systems that combine algorithms and data sources have been widely adopted to address the intricate nature of agricultural environments. [1], [6], [17], [18].

- CNN–LSTM hybrids integrate spatial and temporal learning to process multispectral imagery across multiple growing seasons, offering superior performance in yield prediction and stress detection.
- RF–SVM ensembles leverage Random Forest's feature selection abilities to SVM's classification capabilities. This makes the model more reliable, thus giving it the ability to work correctly even under noisy data conditions.
- By using pre-trained models (e.g., ImageNet, ResNet) in transfer learning, general features can be easily adapted to fit agricultural imagery, thus attaining high precision even in the case of limited datasets.

- Fuzzy logic-based ML systems are becoming more popular for making decisions in situations of uncertainty. Fuzzy–ANN hybrids are applied to irrigation scheduling and soil moisture classification, which are particularly useful in semi-arid areas where climatic parameters change unpredictably. [3], [4], [12], [19].

2.4. Optimization Techniques.

In Agriculture, machine-learning models need to be optimized to improve their efficiency and accuracy. The use of Genetic Algorithms (GA), Particle Swarm Optimisation (PSO), and Bayesian Optimisation are very significant for hyper-parameter tuning and feature selection. [1],[18]

One example is the use of a genetic algorithm to optimise an SVM model to achieve an accuracy of 89% in recognizing paddy fields under mixed cropping conditions. At the same time, particle swarm optimisation is used on a network that performs pest classification, and it is improved by 15%. Similarly, Bayesian optimisation is the current step for the integration phase in the AutoML framework, where automation is done for model selection and weighting in crop yield analysis. [1],[6],[9],[18]

2.5 Key Challenges in Algorithmic Implementation

As a matter of fact, significant technological advances have been achieved, but still, the number of problems that prevent large-scale ML use in agriculture is relatively high [1],[6],[17],[18]

Data Scarcity and Imbalance: Often, agricultural datasets are small, narrowly focused on specific regions, and/or unbalanced, which leads to the problem of overfitting and poor model generalisation. [1],[6],[17],[18]

Model Transferability: Changes in soil types, climate conditions, and crop species make the universal application of trained models a challenging task. [3], [4], [9], [13]

Computational Constraints: Processing high-resolution images requires heavy computational power, and that is the main reason why many rural farming communities are still deprived of it. [1], [6], [17], [18].

3. Precision Agriculture: Smart Irrigation and Soil Health

3.1 Principles of Precision Agriculture

Precision Agriculture (PA) refers to a transformation of the standard practice of uniform, heavily input farming to a data-led, site-specific management approach. It combines cutting-edge sensing technologies, IoT systems, and machine learning (ML) algorithms to non-stop monitor the main agricultural parameters like soil moisture, nutrient composition, weather patterns, and crop health. Farmers are enabled to implement precisely, through data, decisions that optimise the use of inputs while minimising waste. [3], [4], [12].

Precision agriculture through the use of AI and ML analytics enables on-the-spot detection of soil fertility and water needs variations in different locations, thus making the process more efficient and environmentally friendly. To illustrate, the studies on ML-based irrigation systems have shown that these technologies are capable of water saving by up to 27% of water use while at the same time keeping yields at a stable level in dry areas. [In addition, large-area application is supported by the combination of ground sensor data with remote sensing images that can even be done in a complex landscape. [3], [4],[5],[6]

3.2 Machine Learning in Smart Irrigation Systems

Traditional irrigation scheduling typically depends on static or empirical models that do not consider climatic and soil variations of specific areas. On the other hand, machine learning-based irrigation systems modify water needs on their own according to continuous environmental input, thus water use is not only efficient but also more effective. [3], [4], [12].

Some of the typical machine learning models employed for irrigation are:

- **Artificial Neural Networks (ANNs):** These networks are employed to estimate evapotranspiration on the basis of factors like temperature, humidity, and soil moisture content. ANN models have reached a coefficient of determination (R^2) of around 0.92, which is significantly higher than that of regression-based models. [1], [2], [3], [4]
- **Support Vector Machines (SVM):** Support Vector Machines are invaluable in designating irrigation zones as well as pinpointing anomalies within sensor networks, thereby ensuring reliability under uncertain field conditions. [1],[3],[4],[12]
- **Long Short-Term Memory (LSTM) Networks:** These networks are the perfect choice for time-series prediction, such as forecasting water consumption. The utilization of LSTM-based models has led to a reduction in the frequency of irrigation by 15–20% in controlled-environment agriculture research.
- **Random Forest (RF):** Deals with a variety of variables—rainfall, evapotranspiration, and crop growth stage—to facilitate multi-factor irrigation scheduling. [1], [4], [12].

Recent tiled publications comparing different methods have revealed that deep learning techniques are superior to conventional ones when it comes to water resource optimization. In addition, through the use of IoT-enabled sensors, ML can also identify anomalies such as pipe leaks or sensor malfunctions promptly, hence uninterruptedness and operational efficiency can be maintained. [3], [12], [16].

3.3 Soil Health Monitoring and Nutrient Analysis

Healthy soil is the basis of nature-friendly farming. To a large extent, ML algorithms have taken over the job of monitoring soil through multispectral imaging, geospatial analysis, and chemical sensing. [3],[4],[5],[6] Major application areas are

- Soil Moisture Estimation: Regression and ANN models take the data from the ground-based sensors as well as satellite radar (e.g., Sentinel-1) to find out surface and root-zone moisture levels. [3], [4], [5], [6].
- Soil Organic Matter (SOM) Prediction: Random Forest (RF) and Gradient Boosted Decision Trees (GBDT) models utilize hyperspectral reflectance data to estimate SOM with 8–12% improvement in accuracy over the conventional method of sampling. [1], [3], [4], [5].
- Nutrient Availability Mapping: By fusing ML with soil sample data, satellite images, and weather parameters, nutrient distribution maps are generated that serve as a guide for the variable-rate fertigation techniques. These models are most effective in optimising the nitrogen application process, cutting down on fertiliser that is not needed and at the same time reducing the emission of greenhouse gases. [3], [4], [12].

Also, present-day soil analysis designs implement data fusion pipelines that combine satellite vegetation indices (e.g., NDVI, EVI) with locally collected chemical data. This process allows for the exact location of macronutrient shortages (for example, nitrogen, potassium), which are very difficult to detect through manual sampling; hence, both yield and ecosystem health are improved. [1], [6], [9], [18].

3.4 Integration of IoT and Machine Learning

The Internet of Things (IoT) is a powerful tool in precision farming systems, enabling non-stop recording and sending of data from field sensors to AI cloud platforms. By incorporating ML algorithms, IoT communication systems are capable of [3], [12], [16].

- Irrigation schedules are being automated by microcontroller-based valve systems.
- Fertilizer mixture being changed on the fly depending on fresh soil and plant data.
- Besides that, irrigation needs of the future can be anticipated through evapotranspiration predictions along with soil retention models. [3], [4], [12].

3.5 Environmental and Economic Implications

As a result of machine learning, the soil and irrigation system has become more environmentally friendly and cheaper. These systems that optimize irrigation schedules not only lessen the groundwater depletion but also prevent nutrient runoff that, therefore, doesn't pollute the water bodies surrounding the area from where the nutrients come. With variable-rate irrigation, waste is hardly possible, and thus the farmer can expect to get back their money invested, usually within 2–3 years for medium-scale operations. [3], [4], [12].

Moreover, intelligent systems can do a lot for the environment, as they can enhance climate resilience. This is done by them recognizing, for example, drought or nutrient stress even before it happens, as the system sends out an alert. The combination with microclimate forecasting models further helps in dealing with sudden extreme weather, like dry days or rain at odd times, and thus becomes a food producer that is able to sustain the demand. [3], [4], [12].

3.6 Challenges and Future Directions

The innovative irrigation solutions designed to measure soil health and give the corresponding alert system are still struggling to resolve the existing technical and infrastructural problems, despite their success

- Data Heterogeneity: Differences in sensor calibration and data resolution make precise data fusion more difficult. [1], [3], [12], [16].
- Hardware Costs: The high cost of the setup discourages the deployment of IoT networks, most of all in developing regions. [3], [12], [16].
- Connectivity Barriers: In many rural areas, the internet connection is so weak or unstable that real-time data transmission is hardly possible. [15], 16].
- Model Adaptability: Algorithms designed to serve one location may perform poorly in other locations with different agro-ecological zones. [3], [4], [12].

Eventually, research will focus on edge computing, which allows data to be processed locally, thus communication with the server can be done at times other than when data is waiting for transmission, and on transfer learning, which can offer a solution to the problem of models trained in one region and then inefficiently applied to another. Besides that, Blockchain technology, when fully integrated, could be instrumental in enhancing both data traceability and security, thus providing a great need to the farming ecosystem in such a way that decision-making would become more foolproof and verifiable and done in a digital manner. [10]

4. Crop Yield Prediction and Remote Sensing

4.1 Importance of Crop Yield Prediction

Very important that crop yield prediction is done accurately in order to provide food for the world's population in a safe way, to manage resources in an efficient manner and to plan agricultural activities in a sustainable way. Traditional methods of data analysis, such as linear and multiple regressions, are not powerful enough to capture complex, non-linear relationships between natural, biological, and climatic factors influencing crop productivity. [1], [5], [6], [9].

The shortcomings of machine learning (ML) methods are getting resolved by ML methods that take into account data of different origins and types, e.g. of climate variables, soil characteristics, and remote sensing imagery. They are able to model dependencies of any complexity. Yield prediction is a tool process which allows different stakeholders- farmers, policymakers, and supply chain managers - to achieve the most efficient use of inputs, risk minimization and, in the case of weather calamities, pests or diseases, to hedge against possible yield losses. [1], [6], [9], [18].

4.2 Role of Remote Sensing in Agricultural Prediction

Remote sensing (RS) technologies are one of the revolutions in agricultural monitoring. These technologies are holding multispectral and hyperspectral imaging from satellites, UAVs, and drones that provide high-resolution and field-scale data. Also, these data are compatible with ML algorithms, which, when used together with these data, have made it possible to carry out a very detailed examination of crop phenology, soil variability, and vegetation health. [3], [4]

The continuous imagery provided by RS platforms such as Sentinel-2, Landsat-8, and MODIS is used by ML models to obtain vegetation indices like the Normalised Difference Vegetation Index (NDVI), Enhanced Vegetation Index (EVI), and Leaf Area Index (LAI). [5], [6], [18].

New developments like CNN-based segmentation frameworks and CNN–LSTM hybrids can not only consider the spatial but also the temporal aspects of crop growth, thus they are a step ahead of the traditional models. For example, CNN models trained on global maize datasets achieved R^2 scores exceeding 0.9; thus, they were able to predict maize yield more accurately than traditional growth models. [1], [2], [6], [9].

4.3 Machine Learning Models for Yield Prediction

Crop yield estimation has radically changed over the years. Initially, it involved basic regression methods; however, it has now incorporated advanced deep learning models that can handle heterogeneous datasets.

- Random Forest (RF) and Gradient Boosting Models (GBM): These tree-based models effectively analyze environmental features like soil moisture, precipitation, and temperature. [Their ability to determine feature importance helps identify dominant yield factors for specific crops and regions. [1], [6], [9], [18].
- Support Vector Machines (SVM): Used for classifying fields into distinct yield zones based on spectral and reflectance characteristics. [1], [4], [5], [6].
- Convolutional Neural Networks (CNNs): Applied to aerial and satellite imagery for estimating canopy density, biomass, and yield potential. When combined with NDVI and reflectance data, CNNs provide highly accurate, spatially resolved predictions. [4], [5], [6], [18].
- Recurrent Neural Networks (RNNs) Long Short-Term Memory (LSTM) models: These structures are designed to capture temporal dependencies and help yield prediction by incorporating daily and seasonal variations of the climate data. [1], [2], [6] ,[9].

In addition to that, data fusion approaches that integrate satellite images, ground sensor data, and meteorological data strengthen the reliability of the prediction. For instance, an ensemble CNN–LSTM model that was created for wheat yield forecasting showed an 18% improvement in seasonal prediction accuracy as compared to single-model architectures. [1], [2], [6], [9].

4.4 Crop Disease Detection and Early Warning Systems

Diseases in crops have been one of the major challenges to yield stability. Annually, the losses caused by diseases in crops worldwide exceed $200 billion. Machine Learning (ML) and deep learning (DL) models have increasingly become the main instruments in automating disease detection and forecasting processes. [2], [6], [7], [15].

It has been proven that advanced CNN architectures, for example, VGG-19, Inception-V3, DenseNet, trained by transfer learning, have surpassed 95% accuracy in detection for different crops like rice, banana, potato and maize. Simple networks like MobileNetV2 on-device inference through smartphones and UAVs enable instant disease diagnosis in the field. [15], [16]

The main points of the current state of research are

• AI-based early warning systems analysing color, texture, and chlorophyll changes to detect diseases that have not yet shown visible symptoms. [2], [7], [15], [16].

• Hybrid AI–IoT surveillance frameworks, combining drone imagery, soil and weather sensors, and prediction models for rapid outbreak tracking. [1], [3], [4], [5].

• Explainable AI (XAI) methods that provide more transparency in diagnosis, thus giving more chances to agronomists to model decisions for better management. [8]

State-of-the-art AI systems are capable of recognizing early-stage bacterial and fungal infections with 92–94% accuracy, thus enabling the implementation of proactive pest control and integrated disease management strategies. [2], [7], 15], [16].

4.5 Emerging Trends and Technologies

The latest breakthroughs in ML-powered yield prediction and remote sensing unveil numerous forward-looking research topics [1], [4], [5], [6].

•Transformer Architectures: Significantly enhance temporal attention for multi-season yield prediction and can scale faster and learn more efficiently than recurrent models. [1] ,[2], [6], [9].

•Edge Computing: Localises the processing of ML inferences on drones, sensors, or field devices, thus cutting down on the time delay and reducing the need for a cloud-based infrastructure. [15], [16].

Blockchain-enabled Traceability: Facilitates open and verified data sharing throughout the agricultural supply chain; thus, by reporting yields, it is helping to maintain accountability. [10]

When put together, these innovations signify the birth of an advanced technology-driven, sustainable agricultural system based on AI, which is capable of self-learning, self-adapting, and real-time yield improvements. [1], [6], [9], [15].

5. Disease Detection and Pest Control

5.1 Introduction

Diseases in crops and pest have been the leading causes of global yield losses. [1]. their share in the deficiencies of total food production per year is almost 30–40%. Traditional ways of recognizing diseases, which are based on manual scouting, microscopy, and farmer observation, are frequently labor-intensive, time-consuming, and error-prone. Machine Learning (ML) and Deep Learning (DL) Convolutional Neural Networks (CNNs) and hybrid architectures have radically changed the field of plant pathology by allowing automated, image-based disease diagnosis with great precision. By focusing on the spectral, morphological, and textural attributes of diseased plant parts, ML solutions are now proficient in determining early-stage infections that are even beyond human vision. Moreover, these systems acquire information from drones, IoT-enabled sensors, and mobile imaging devices, producing early signals of disease that not only lower yield attrition but also help in saving pesticides. [1], [2], [3], [6].

5.2 Deep Learning Architectures for Disease Detection

Recently, it has been shown that deep learning architectures significantly outperform traditional ML classifiers, i.e., Support Vector Machines (SVM) and Decision Trees—in the aspects of scalability, adaptability, and predictive accuracy. Main deep learning architectures are [2], [6], [15], [17].

- VGG-16 and VGG-19: Extremely efficient in hierarchical visual feature extraction, resulting in 99.7% accuracy for recognition of nine tomato leaf diseases based on the Plant Village dataset. [2],,[7]

- Inception-V3 and DenseNet-121: Mainly used for differentiating the color and texture of different crop datasets, achieving 98%+ accuracy of banana leaf disease classification. [7], [15]

- MobileNetV2 and EfficientNetB0: Compact and specially designed for edge devices, thus, real-time disease detection through smartphones and drones is possible. Field experiments conducted by Bilal et al. (2025) showed F1-scores of 95.6% with the inference being done for less than 10 milliseconds per image. [1], [2], [18].

- Swin Transformers: State-of-the-art attention models that achieve 97.7% accuracy in locating early-stage plant diseases in products like potato and fig. [2], [6], [15], [17].

The best trade-off between accuracy, model compression, and live detection abilities has been achieved through fusion architectures that integrate MobileNetV2 and EfficientNetB0. These hybrid structures have been effectively implemented for in-field testing on Android smartphones and Raspberry Pi devices. [2], [7], [15], [16].

5.3 Machine Learning Models for Pest Prediction

Machine learning is the leading technology that supports not only disease identification, but also pest species classification and population forecasting. To that end, scientists have trained the Random Forest (RF), Gradient Boosted Trees (GBT), and SVM algorithms to identify pest types by wing venation, body morphology, and damage patterns on leaves. [1], [2], [7], [15].

Feature extraction through CNNs may be greatly improved if the CCMT (crops: cashew, maize, cassava, and tomato) dataset is used for real-time pest identification, where precision and recall rates both exceed 95%. [1], [2], [7], [15].

Hybrid CNN–SVM frameworks have especially excelled in separating visually highly similar pest species like fruit flies, aphids, and caterpillars. [1], [2], [6], [7]. Besides that, predictive pest population models that bring in environmental factors like temperature, moisture, and pest life-cycle data can forecast pest outbreaks several workdays ahead, thus making timely pesticide application and biological control interventions possible. [1], [2], [3], [4].

5.4 Data Sources and Datasets

Providing good quality, free access datasets has significantly contributed to research in AI-driven agricultural science. Some of the significant examples are

- Banana-LSD Dataset: More than 2,500 field pictures representing seven different categories of banana leaf diseases, which are commonly utilized in the development of CNN-based hybrid models.

- Custard Apple Disease Dataset: Labelled dataset consisting of 8,000 images; classifiers such as SVM and KNN that were trained on this dataset reached 99.5% accuracy.

- Potato Leaf Image Dataset: Around 3,000 annotated samples of fungal diseases, bacterial diseases, and viral infections that were used for a DenseNet-based classification with 97% accuracy.

- Fig Leafworm Dataset: The dataset enables deep learning models to distinguish between the two classes of infested and non-infested leaves with 86%+ accuracy. [2], [6], [15], [17].

Well-designed datasets like these offer the fundamental reference points necessary for the assessment and validation of AI frameworks in worldwide agriculture.

5.5 Real-Time Implementation and Edge Computing

In order to facilitate the introduction of new technologies in agriculture directly outside the lab, AI-powered systems are getting more and more fine-tuned for edge computing, where the operations are carried out right on low-power units without the need for a continuous cloud connection. TensorFlow Lite and PyTorch Mobile are the technologies that allow on-the-fly calculations coming from environmental data stored in microcontrollers, smartphones, and UAVs.

For instance, MobileNetV2–EfficientNetB0 fusion models achieved sub-10 millisecond detection latency on handheld instruments and embedded systems like Raspberry Pi boards. [15][16] These technologies, when combined with drone-mounted high-resolution cameras, allow automated pesticide spraying to be enabled as fast as pest detection, thereby greatly saving time and also cutting down the chemical waste drastically. [2], [5], [7], [15].

5.7 Challenges and Future Prospects

In the draft of the paper, the authors described up to the present time the deployment of AI-based disease and pest management systems with significant achievements; however, they also mentioned the following challenges that need to be solved [2], [7], [15], [16].

- Limited Annotated Data: The availability of labelled datasets, which are of high quality, is still limited for rare or new diseases. [2], [7], [15], [16].

- Visual Variability: Differences in natural variations within species and between species make it difficult for models to generalise across different regions and types of crops. [2], [7], [15], [16].

- Environmental Noise: The condition under which the photos of the outdoor subjects are taken, for example, the change of the light source and the clutter of the background, will have a negative influence on the accuracy of the detection.

- Integration Gaps: Many current models are only at the conceptual stage and have not been completely integrated with practical farm management systems or easy-to-use mobile tools for end-users. [15], [16].

The next generation of work would focus more on technology like Explainable AI (XAI) that would enhance AI systems' interpretability and Federated Learning, whereby different partners can improve the models locally without the need for data exchanges. Besides, the adoption of the acoustic sensing and thermal imaging technology is another way of early pest-detection in the sub-surface pest activity.

Eventually, AI–IoT pest prediction networks are going to make self-surveillance possible, as well as enable targeted management interventions—thus contributing to the reduction of environmental pollution and upholding the principle of sustainable pest control at the same time.

Conclusions

An extensive literature review clearly demonstrates the pivotal and revolutionary role of the. In the agricultural sector. The application of AI technologies across areas such as soil health assessment, disease detection, yield forecasting, and genomic crop improvement demonstrates that these technologies are powerful not only in optimising resource utilisation and achieving higher productivity, but also in ensuring the long-term sustainability of the ecosystem. The use of machine learning algorithms such as CNNs, LSTMs, Random Forests, and Gradient Boosted Models, along with normalising remote sensing and IoT data, allows us to obtain a very accurate picture of plant health and other environmental variables. Thus, the interaction between these technologies leads to less waste, higher profits, and greater empowerment of farmers through real-time, data-driven decision-support tools. Moreover, AI-assisted breeding and high-throughput phenotyping methods are not only making the traditional plant genetic improvement pipeline more efficient but also helping uncover new genes, improve stress tolerance, and exponentially increase the potential of climate-smart agriculture. In summary, AI represents a significant shift from reactive to predictive and prescriptive approaches in agriculture, where decisions are not made by human norms but by scientific insight. These issues require strategies that enable different parties, such as governments, researchers, and technology-producing companies, to collaborate and create open and inclusive frameworks that ensure sustainable digital agriculture. Artificial Intelligence is no longer a futuristic concept; it is an active enabler of precision, productivity, and sustainability in agriculture. As digital tools become more accessible and affordable, AI-driven farming systems will empower even small-scale farmers to make informed, data-backed decisions. The fusion of AI, IoT, and remote sensing technologies will significantly alter the methods of planting, observing, and enhancing the cropsAI, along with research, policy support, and ethical management, can be a significant force for the emergence of a resilient, fair, and innovative agricultural ecosystem that will be able to feed the rising global population and at the same time, conserve nature for the future generations.

References

[1] Chlingaryan, S. Sukkarieh and B. Whelan, "Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review," *Computers and electronics in agriculture,* vol. 151, p. 61–69, 2018.

[2] A. K. Srivastav and P. Das, "Sustainable Soil Health and Fertility," in *Biotechnology and IoT in Agriculture and Food Production: Green Innovation*, Springer, 2025, p. 95–104.

[3] D. G. Pai, M. Balachandra and R. Kamath, "Explainable AI in agriculture: review of applications, methodologies, and future directions," *Engineering Research Express,* 2025.

[4] U. Nagavelli, T. Sreeja, V. Srilakshmi and M. V. Kumar, "Precision Agriculture with Remote Sensing: Integrating Deep Learning for Crop Monitoring," in *Sustainable Farming through Machine Learning*, CRC Press, 2024, p. 174–187.

[5] S. Ahmad, F. Siddiqui, D. Y. Perwej, H. Rizvi and D. N. Akhtar, "Modelling Crop Yield with Deep CNN-LSTM for Spatiotemporal Data Analysis," *International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN,* p. 2456–3307, 2025.

[6] Meruga, Vishal Bharadwaj. "LLM-Guided Multi-Agent Collaboration for Complex Task Automation." In 2025 5th International Conference on Emerging Research in Electronics, Computer Science and Technology (ICERECT), pp. 1-7. IEEE, 2025.

[7] J. Dhakshayani, B. Surendiran and J. Jyothsna, "Artificial intelligence in precision agriculture: A systematic review on tools, techniques, and applications," *Predictive Analytics in Smart Agriculture,* p. 37–57, 2023.

[8] M. I. Hossen, M. Awrangjeb, S. Pan and A. A. Mamun, "Transfer learning in agriculture: a review," *Artificial Intelligence Review,* vol. 58, p. 97, 2025.

[9] J. Zhao, D. Liu and R. Huang, "A review of climate-smart agriculture: Recent advancements, challenges, and future directions," *Sustainability,* vol. 15, p. 3404, 2023.

[10] D. Bhumichai, C. Smiliotopoulos, R. Benton, G. Kambourakis and D. Damopoulos, "The convergence of artificial intelligence and blockchain: The state of play and the road ahead," *Information,* vol. 15, p. 268, 2024.

[11] R. Dembani, I. Karvelas, N. A. Akbar, S. Rizou, D. Tegolo and S. Fountas, "Agricultural data privacy and federated learning: A review of challenges and opportunities," *Computers and Electronics in Agriculture,* vol. 232, p. 110048, 2025.

[12] S. Mansoor, S. Iqbal, S. M. Popescu, S. L. Kim, Y. S. Chung and J.-H. Baek, "Integration of smart sensors and IOT in precision agriculture: trends, challenges and future prospectives," *Frontiers in Plant Science,* vol. 16, p. 1587869, 2025.

[13] S. Ahmed, S. N. K. Marwat, G. B. Brahim, W. U. Khan, S. Khan, A. Al-Fuqaha and S. Koziel, "IoT based intelligent pest management system for precision agriculture," *Scientific Reports,* vol. 14, p. 31917, 2024.

[14] M. A. Mushtaq, H. G. M.-D. Ahmed and Y. Zeng, "Applications of artificial intelligence in wheat breeding for sustainable food security," *Sustainability,* vol. 16, p. 5688, 2024.

[15] L. Ale, A. Sheta, L. Li, Y. Wang and N. Zhang, "Deep learning based plant disease detection for smart agriculture," in *2019 IEEE Globecom Workshops (GC Wkshps)*, 2019.

[16] A. A. Ahmed and G. H. Reddy, "A mobile-based system for detecting plant leaf diseases using deep learning," *AgriEngineering,* vol. 3, p. 478–493, 2021.

[17] S. Fei, M. A. Hassan, Y. Xiao, X. Su, Z. Chen, Q. Cheng, F. Duan, R. Chen and Y. Ma, "UAV-based multi-sensor data fusion and machine learning algorithm for yield prediction in wheat," *Precision agriculture,* vol. 24, p. 187–212, 2023.

[18] H. Gunal, M. Budak, E. Gunal, M. Kilic, A. R. M. Al Tawaha, A. R. Al-Tawaha, A. Karnwal, N. Nesterova, A. Singh, V. D. Rajput and others, "Soil fertility in arid lands: strategies for sustainable management and fertilization," in *Sustainable Agriculture under Drought Stress*, Elsevier, 2025, p. 61–69.

Comparisons Table

S.N	Authors & Year	Focused Area	Data Type	Techniques Used	Key Contributions	Limitations / Gaps
1	Chlingaryan et al., 2018	Crop yield prediction & nitrogen estimation	Soil data, remote sensing, weather data	RF, SVM, ANN, Regression	Comprehensive review of ML methods for yield and nitrogen estimation; shows ML outperforming traditional methods	Limited discussion on deep learning and real-time systems
2	Srivastav & Das, 2025	Soil health & fertility management	Soil samples, IoT sensor data	ML models with IoT integration	Highlights sustainable soil management using AI and IoT	Focused more on soil, less on crop-level prediction
3	Pai et al., 2025	Explainable AI in agriculture	Multi-source agricultural data	XAI models (SHAP, LIME)	Improves transparency and trust in AI-based agricultural decisions	Limited real-world deployment examples
4	Nagavelli et al., 2024	Crop monitoring using remote sensing	Satellite & UAV imagery	CNN, Deep Learning	Demonstrates effectiveness of DL for crop health monitoring	High computational cost; requires large datasets
5	Ahmad et al., 2025	Crop yield prediction	Spatiotemporal satellite data	CNN–LSTM hybrid	Captures spatial and temporal dependencies; improves yield prediction accuracy	Model complexity and training time are high
6	Shahi et al., 2023	Crop disease detection	UAV imagery	CNN, Transfer Learning	High accuracy disease detection using UAV data	Sensitive to image quality and lighting conditions
7	Dhakshayani et al., 2023	Precision agriculture (review)	Multi-source data	ML, DL, IoT frameworks	Systematic review of AI tools and applications in agriculture	Does not provide experimental validation
8	Hossen et al., 2025	Transfer learning in agriculture	Agricultural image datasets	Pre-trained CNN models	Shows reduced training time and higher accuracy	Domain adaptation remains challenging
9	Zhao et al., 2023	Climate-smart agri-	Climate, soil, crop data	ML-based decision	Addresses sustainability and	Limited focus on real-time

S.N	Authors & Year	Focused Area	Data Type	Techniques Used	Key Contributions	Limitations / Gaps
		culture		systems	climate resilience	implementation
10	Bhumichai et al., 2024	AI + Blockchain integration	Agricultural supply chain data	AI models + Blockchain	Enhances data security, traceability, and trust	High infrastructure cost and scalability issues