Journal of Agriculture, Food, Environment and Animal Sciences
Tarım, Gıda, Çevre ve Hayvancılık Bilimleri Dergisi
http://www.jafeas.com, ISSN: 2757-5659

J. Agric. Food, Environ. Anim. Sci.
6(2): 626-655, 2025

Predicting Crop Diseases with Machine Learning Methods and Computer Vision Techniques

Praveena PALANIAPPAN[1]*, Vijay Aadhithya CHANDRASEKARAN[2], Nishanthini KALAIMANI[3], Vetha Vardhini SUBRAMANIAN[4]

[1-4]Department of Biotechnology, Sethu Institute of Technology, Pulloor, Virudhunagar 626 115, Tamil Nadu, INDIA

[1]https://orcid.org/0009-0003-6857-0974, [2]https://orcid.org/0009-0009-0051-3985
[3]https://orcid.org/0009-0003-2799-4092, [4]https://orcid.org/0009-0003-9027-9710

*Corresponding Author: praveena@sethu.ac.in

Review Article

Article History:
Received: 17 September 2025
Accepted: 12 November 2025
Published online: 15 December 2025

***Keywords*:**
Machine Learning
Smart Farming
Computer Vision
Agriculture Crop
Deep Learning

ABSTRACT

Crop diseases and insect pests pose a serious threat to global food security, with annual losses estimated at 20–40% of production. Climate change further intensifies these challenges, driving pest damage and frequent disease outbreaks. Traditional methods of detection and diagnosis remain essential, but they are often limited in speed and accuracy. In this context, machine learning (ML) and deep learning (DL) offer promising solutions. Advanced models and algorithms such as Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks enable accurate identification and prediction of plant diseases. By detecting problems early, these systems support timely interventions, better crop yield, and reduced pesticide use. This review explores the role of ML in agriculture, emphasizing its potential to revolutionize plant disease management. State-of-the-art models currently used for disease identification include ResNet, DenseNet, Inception, GoogleNet, MobileNet, and LSTM. Their integration with traditional agricultural practices enhances crop management, optimizes resource utilization, and helps mitigate the impacts of climate change. However, challenges remain. These technologies face issues related to data quality, computational complexity, and the need for interdisciplinary collaboration between agriculture, computer science, and policy. Despite these limitations, ML-driven approaches are becoming increasingly significant for sustainable agriculture. By combining conventional practices with intelligent systems, the sector can adapt to rising global demands while minimizing environmental impact. As technology continues to advance, the future of agriculture will depend heavily on data-driven solutions and smart decision-making, ensuring food security for growing populations.

To Cite : Palaniappan P, Vijay Aadhithya C, Nishanthini K, Vetha Vardhini S., 2025. Predicting Crop Diseases with Machine Learning Methods and Computer Vision Techniques. Journal of Agriculture, Food, Environment and Animal Sciences, 6(2): 626-655.

INTRODUCTION

Infestations of pests and diseases of plants are the basic factors which lead to inadequate availability and hygienic conditions of food. They are seasonal and depend on pathogens, environmental factors, and crop variety. Traditional methods used for disease detection and localization take a lot of time and require many people to be involved in the process. The progress of sensor technologies and data processing opened new perspectives in the field of disease surveillance, including automatic monitoring and recognition systems. The data fusion techniques have been applied, with their machine learning-based data fusion exercised in agriculture for the optimization of profitability, sustainability, and protection of land resources (Ouhami et al., 2021). These infections and diseases in plants and crops may largely diminish the quality and quantity of crops, hence affecting economies. In India, agriculture forms the livelihood of 70% of the population, contributing towards 17% of the country's GDP. Crop pathologies are influenced by various factors such as environmental problems, unsuitable crops, climatic change, pests, and weeds. Crop growers, however, remain uninformed of how to adopt proper strategies of infection control. The traditional methods of physical observation are pretty useless in detecting a new infection. Computer vision, machine learning, and deep learning technologies can be leveraged and put into service to detect crop diseases accurately and at speed (Wani et al., 2021). The human population has been increasing very fast, and it is expected to achieve 9.7 billion by the year 2050 due to the enhanced medical care. With the increase in the human population comes the need for increased demand for agricultural products, which, in turn, requires growing output of crops by a minimum of double digits before the year 2050. Conventional practices in agricultural production are unsustainable. Output is threatened by diseases and insect pests. Climate change increases damage from and resistance to pests. Crop monitoring and forecasting should be automated, therefore reducing environmental damage and the costs incurred during production (Domingues et al., 2022). The ever-increasing population is the reason behind food and water shortages. These are threats to the environment, damage to agriculture production, and deformation of quality of life. Detection of the early signs of diseases can be done through the information gathering about plant health with the help of sensors and images. At last, it will help to overcome the deficiencies regarding food and environmental degradation, which will enhance the quality of life (Sinha and Shekhawat, 2020). Population pressure, together with improved food demands, has put the agriculture sector in jeopardy. Diseases can crop up due to bacteria, fungi, and viruses, reducing crop yields. To circumvent all these, plant disease detection approaches are in use. Machine learning and deep learning techniques are being used in disease identification, where deep learning has superior performance on computer vision. Comparative studies have illustrated the effectiveness of deep learning over machine learning in the detection of leaf diseases by captured images, hence evading massive crop losses. (Jackulin and Murugavalli, 2022). Computer vision has been turning into an essential methodology for phenomics in plants—a science to differentiate between genetic diversities and phenotypic characteristics of plant species. Developments in DNA sequencing technologies have given the ability to rapidly access genomic variations within population. Integration of high-throughput analytical platforms with plant phenotyping tools lends extensive power to explore genetic factors conditioning complex quantitative characters, like growth, environmental stress tolerance, disease resistance, and yield. High-throughput phenotyping enabled by computer vision and various sensors, combined with the power of algorithms, could improve crop yields able to support future scenarios associated with population demography and climate change. Machine learning offers data-driven predictions in image analysis. This improves throughput and accuracy (Mochida et al., 2018). AI has many subcategories, including machine learning and a wide range of algorithms. Deep learning also creates artificial neural networks. Researchers apply these models in leaf retrieval, image segmentation methods, and detailed analysis of big data. In detecting and recognizing plant diseases, deep learning models like the Chan-Vese algorithm and migration learning algorithms have proven effective. Convolutional neural networks (CNNs) are powerful techniques for recognizing patterns and handling large data sets. Thus, they are essential tools in agriculture. For example, one model has been developed that achieves an accuracy of 91 to 98%. This model can identify 13 diseases across five plant types using internet databases. Continued work in this area and the integration of newer technologies may open up fresh opportunities in agriculture (Das et al., 2023). In recent years, machine learning (ML) algorithms have been applied in agriculture. They help increase crop yield while lowering input costs through techniques that allow farmers to choose better crops, estimate yields, and predict diseases. While these methods don't solve all farming problems, they offer a powerful set of tools for various applications. The technology is reliable, does not cause damage, and produces consistent results in tasks like weed and disease detection as well as crop management. Although ML techniques can address some agricultural challenges, they cannot solve everything (Rehman et al., 2019). As shown in Figure 1, the workflow enables early identification of leaf-level symptoms

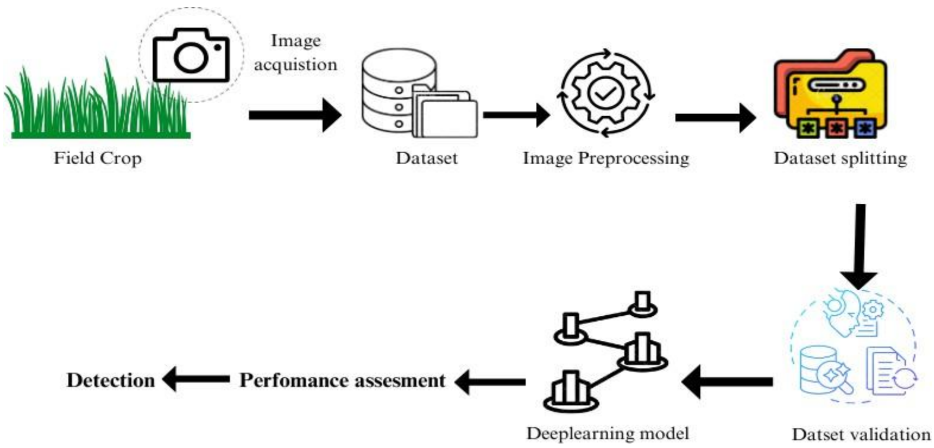

Figure 1. Efficient workflow for plant disease assessment

Machine learning is currently popular due to the growth of artificial intelligence, which allows machines to learn on their own. The main goals of ML techniques include feature extraction, pattern recognition, object detection, and classification. ML plays a key role in computer vision, impacting areas such as surveillance systems, optical character recognition, robotics, and suspect detection. In medical imaging, this technology improves image quality and identifies important features. This paper reviews the role of ML in computer vision and image processing, discussing tools, applications, datasets, and challenges (Khan et al., 2018). Detecting biotic stress early is vital for protecting crops. New non-invasive optical sensors and data analysis methods have greatly improved weed, plant disease, and insect pest detection. Machine learning methods like support vector machines and neural networks are used in precision agriculture for early identification of plant diseases and weeds based on spectral features (Behmann et al., 2014). Plant and crop diseases are affected by weather and environmental factors, with the disease triangle model being fundamental. A disease occurs when a pathogen encounters a host organism under suitable conditions. Over time, plant pathologists have expanded this triangle to include factors like humans, vectors, and time. Research shows that considering multiple climate change factors is crucial for understanding the impact of plant disease outbreaks. Combining climate, crop growth, and disease models can lead to forecasts for different scenarios (Fenu and Malloci, 2021). Some machine learning techniques used in disease detection include artificial neural networks, decision trees, support vector machines, and K-means. All these methods need to convert images into data that computers can interpret. Thus, basic knowledge of computer coding and programming is necessary in this field. These technologies may not work directly with images taken in the field (Khan et al., 2021). In agriculture, machine learning is equally important for predicting plant diseases accurately. This can minimize diseases by processing photos captured by cameras, leading to increased plant yield and reduced labor requirements (Shirahatti et al., 2018). The future of agricultural automation entails challenges as well as opportunities. Computer vision, using cameras and computers to identify and track targets, forms the backbone of small field farming. That said, it brings along some challenges that include expanding into new applications, building large data sets, and making sure of robust performance in complex environments. In the near future, with the aid of deep learning and other intelligent technologies, computer vision technology will be applied to the management of agricultural production, improving economic performance and solving current problems in agriculture (Tian et al., 2020). Challenges in the agricultural sector are many: a growing world population, decreasing arable land, and a lack of labor. Computer vision can help in standardizing crop data extraction, thereby reducing labor costs for better health and quality of plants. These systems will also be able to assist farmers in finding out the growth stages of crops so that timely decisions about harvests can be made for better yield. Traditionally, such monitoring of plants requires physical efforts, expertise, and focus. This would mean a call for the development of intelligent and automated vision-guided systems in farming in view of agricultural success (Ghazal et al., 2024). The literature review directs the creation of machine learning (ML)-based smart farming tools that provide farmers with data-driven decision-making support systems. In addition to maintaining and improving crop quality and yield, this helps lessen the need for pesticide application and the harm that comes with it. In the end, this helps to satisfy the world's food needs while causing less harm to the environment. This article actively promotes the advancement, growth, and success of smart farming initiatives, as ML-based methods are essential to their emergence and development (Moysiadis et al., 2021).

Machine Learning Techniques and Its Types:

ML applications on precision agriculture for yield prediction, disease detection, crop quality, and species recognition are far more. Automating harvesting, immature green citrus detection, grassland biomass estimates, and wheat, disease, and yield prediction have few practices. Most of the research is on disease detection and yield prediction, among which the receptivity of others dealing with better waste reduction in the identification of features and betterment of product prices. Capable of estimating the geographic origin plant species, along with time reduction in classification (Liakos et al., 2018). Agriculture is the backbone of developing countries. However, farmers face many problems related to climate change, soil degradation, nutrient deficiencies, and the need for trying different crop varieties. This paper discusses prediction techniques of crop types through algorithms like K-Nearest Neighbors, Decision Trees, and Random Forest with respect to different climatic conditions and different levels of nutrients present in the soil. Random Forest performed as the best predictor and outperformed KNN, decision trees, and entropy in terms of accuracy (Rao et al., 2022). Through historical meteorological and soil information, crop yield forecasting and prediction are achievable with the help of machine learning techniques. In pre-season agriculture yield forecasting, a neural network system was proposed which achieved an accuracy more than 75%. Naïve Bayes MapReduce Precision Agricultural Model was proposed for the region belts of India. Predictive algorithms in data mining were tested for soil classification and resulted in increased farming productivity without reliance on high quantities of fertilizers. Supervised machine learning techniques were applied to agricultural output predictions in production (Olalere et al., 2024). It is worthwhile to plan for agriculture, which could predict crop yield and monitor the quality of crops. In order to develop an estimate of crop yields and make relevant decisions about crop planting and growing seasons, machine learning is used. The measurement of NPK and pH values would tell the moisture content available in the soil. These values may predict the type of crop that decides the type of crop to be sown. This paper aims at bettering the accuracy and error rate of crop yield prediction using decision tree supervised machine learning algorithms for improved economic growth and sustainable farming practices (Paudel et al., 2022). Crop yield estimation is quite a complex task as it is dependent on the interrelation of the environmental factors, amongst which changes in weather, soil properties, or the farmer's choices and practices of crop rotation. Correct methods for forecasting crop yield must be developed after assessing the weather and soil conditions through crop and soil management experiments. The two most common approaches for predicting yield before harvest are crop growth models and data-driven models. Crop growth models aim to mathematically represent the interactions between plant physiology and the environment. On the other hand, data-driven models analyze crop yield data from several years to identify the key factors that affect yield variations. Both approaches have their advantages and disadvantages. However, this research requires a unified framework to model the nonlinear relationships among soil factors, weather, biomass, and crop yield (Bali and Singla, 2021). Machine learning shows great promise in agriculture due to the shortage of skilled farmers. It helps detect crops, prevent diseases, manage topsoil, and more. Machine learning models support quick and efficient decision-making, thus improving productivity and management in areas like soil classification, disease detection, species management, water management, yield prediction, crop quality, and weed detection. This overview includes various machine learning approaches proposed in the last five years, focusing on their advantages and disadvantages. It will also compare different algorithms used in modern agriculture (Veeragandham, 2020).

Analysis of Machine Learning Approaches

In this regard, in the present study, a crop recommendation system with an accuracy rate of 96% has been developed using Artificial Neural Networks and soil data from Karnataka, India. (Madhuri and Indiramma, 2021). It has presented a crop prediction methodology using feature selection methods in order to improve crop growth and productivity. Here, three selection procedures are used: Recursive Feature Elimination, Sequential Forward Feature Selection, and Boruta, which turned out to give accuracies such as 92.72%, 91.26%, and 91%, respectively. (Suruliandi et al., 2021)

It contributes a Naive Bayes-based crop recommendation algorithm with parameters of weather, selling price, and yield. Data cleaning methods combined with selection techniques and cross-validation, in this way, have returned an accuracy as high as 85.71% and robustness with low feature dimensionality, but at the cost of time consumption and high error occurrences (Palakshappa et al., 2024).

It developed a crop recommendation system to aid Indian farmers in maximizing crop production by identifying suitable crops based on climatic changes. Ensemble model learning procedures reduced training time due to various features. (Reddy et al., 2019).

This research uses temperature, pH, moisture, and NPK sensors in monitoring the parameters of the soil, combining an IoT system with machine learning to minimize soil degradation and foster crop health. The analysis of data was carried out using random forest and CNN models (Gosai et al., 2021). It presents a crop recommendation method that uses mining techniques and a harvest suggestion framework with the help of a support vector machine to improve efficiency. This approach will provide farmers with more information on soil supplements for better harvests, but concerns about efficiency arise (Pujar et al., 2024). A method for forecasting irrigation recommendations has been developed using data from 22 soil sensors over two years. Four major plots were analyzed, and various classification and regression methods were used with reasonable accuracy in regression (Goldstein et al., 2017). The authors expect better outcomes for Indian farmers by utilizing soil databases and expert advice sourced from a lab dataset. They implemented an ensemble model based on majority voting, incorporating Artificial Neural Networks and Support Vector Machines to recommend suitable crops based on soil characteristics (Jajur and Raju, n.d.). In their study, the authors employed equilibrium optimizer-based kernel extreme learning machines for crop recommendations and random forest machines for precise yield predictions. The simulations achieved a maximum accuracy of 97.91% (Gopi and Karthikeyan, 2023).

Traditional Disease Diagnosis: Foundations and Limitations;

Disease control has for many years relied heavily on agricultural management strategies that address cultural practices, such as crop rotation, soil solarization, and the cultivation of tolerant cultivars, either alone or in conjunction with other methods, such as the use of biological agents and low dosages of pesticides (Khalili et al., 2020). Conventional techniques, such as linear regression and linear discriminant analysis, depend on predetermined distributions and model assumptions, guaranteeing accuracy only for data that satisfies these criteria. Due to their requirement for linear class separation and representation by a unimodal Gaussian distribution, these techniques have a limited range of applications (Behmann et al., 2014). Traditional plant diseases and pests identification processes based on machine vision usually use traditional image processing algorithms or manually constructed characteristics and machine learning algorithms. These approaches raise the cost of implementation while reducing the complexity of traditional algorithm design. However, it is difficult to completely eradicate the influence of altered scenes on recognizing results via classical algorithms in complex natural environments. Small variations between lesion areas and backgrounds, low contrast, large scale variations, noise in lesion images, and disturbances from natural light conditions are some of the difficulties that must be addressed. Due to all of these factors, traditional classical methods are becoming almost useless and difficult to use in order to obtain better detection results (Liu and Wang, 2021).

Early detection and prevention of plant diseases are an important part of agricultural technology, as plant pathologies lead to vast ecological and agricultural losses. The traditional manual way of visual observation is inefficient and time-consuming; it also increases overhead costs. In the past years, due to improvement in computer vision in precision agriculture technology, there is an increase in efficiency in disease detection and crop production. Automatic plant disease detection using machine learning algorithms became an effective way to perform precision agriculture. Traditional ML-based techniques, such as K-means clustering and support vector machines, are less accurate and slow for real-time disease detection due to lengthy image preprocessing steps and sequential extraction of features. Models using convolutional neural networks have become popular due to higher accuracy in object detection (Roy and Bhaduri, 2021). Traditional methods of information and knowledge management in agriculture are labour-intensive, time-consuming, and prone to errors. For the farming sector to become 'smarter,' there is a requirement for technological advancement in remote sensing, digital applications, sensors, advanced imaging systems, cloud data storage, and intelligent data analysis with decision support systems. Smart agriculture shall make use of the latest technologies in IoT, Machine Learning, Cloud computing, and Blockchain in enhancing food productivity and mitigating emerging challenges in the sector (Dhanya et al., 2022). One of the techniques used in capturing crops and plant leaves at ground level for disease detection is crop ground imaging. Researchers are however faced with challenges in capturing leaves in field conditions due to complex backgrounds, shadows, and unstable luminosity. Plant spectral characteristics are impacted by stress, which makes it possible to identify diseased and uninfected leaves and categorize the severity of the disease. The methods used today are based on machine learning algorithms because of their ability to make predictions. For example, support vector machine models are used to detect plant diseases. Conventional machine learning algorithms have been used to identify diseases on plant leaves and to detect drought stress in barley early. K-nearest neighbor and the decision tree-based classifier C5.0 have been used in classifying grey mold infection severities on tomato leaves with hyperspectral images. Using a broad spectrum model, both normal and infectious leaves can be distinguished correctly with 92.86 percent and 85.71 percent precision, respectively (Ouhami et al., 2021). Increased insect metabolic rates brought on by climate change are increasing food consumption and causing pest damage and development. The use of chemical treatments and pesticides to protect crops is increasing, which harms the environment and has detrimental effects on human health. But it also raises the possibility that pests will become resistant to pesticides. The naked-eye observation method used in traditional plant disease detection is time-consuming and impractical for large farms. Crop monitoring and forecasting must be done automatically to prevent overuse of chemicals and pesticides, which lowers production costs and environmental harm (Domingues et al., 2022). Detection and classification of rice diseases are important because traditional methods resulting from visual assessment, pathogen-induced symptoms, and laboratory identification are subjective and lead to false diagnoses. The visual classification is also difficult because one plant may be affected by many diseases at the same time, and multiple laboratories are laborious and time-consuming in identification. High-level professionals are often unavailable, especially in small farms. It has, hence, led to increasing interest in applying machine learning methods for automatic detection and classification of crop diseases using digital images. This way, there will be improved efficiency and accuracy in the detection and classification of diseases (Arinichev et al., 2021). Computer vision and machine learning are disrupting the traditional agriculture crop production cycle, thereby increasing productivity and efficiency. These innovative ICT solutions enable gathering and representation of spatial and spatiotemporal information, while learning allows for planning, reasoning, and inference. Their junction of computer vision and machine learning automatizes the tasks of a crop cycle, improving fidelity and efficiency. Advancements related to computer vision, remote sensing platforms, parallel computing, open source machine learning libraries, large datasets, reproducibility in research, crop health monitoring, pest control, automatic harvesting, and automated farm management - these all contributed to innovations in precision agriculture (Raval et al., 2021). Traditional computer vision detection of plant diseases involves complicated and time-consuming feature engineering. Traditional methods often focus on a limited number of classes within one crop, such as tomato powdery mildew, apple scab, citrus huanglongbing, and tomato yellow leaf curl virus. Machine learning has been hugely applied in plant phenotyping, but it suffers from the inability to revisit features and datasets because of changes in the problem or dataset (Mohanty et al., 2016).

Machine Learning Techniques For Crop Production And Pest Reduction :

An increasing human population requires agricultural products, so there is a need to double the yield of crops by 2050. Traditional practices are unsustainable because they reduce biodiversity and causes emissions of greenhouse gases. Plant diseases and insects present major challenges as their damages are valued at US $220 and US $70 billion respectively. The risk is increased by climate change and pesticide resistance. These are now being promoted through automated methods, such as ML, for higher-yield crops and detection of diseases and water and soil management. There are several commercial tools, like Plantix and CropDiagnosis, available for identifying plant diseases by farmers (Domingues et al., 2022). Machine learning as well as deep learning methods are used to identify plant diseases. The prediction has been significantly impacted by RF, SVM, DT, and NB. To enhance disease prediction in cotton, rice, and citrus plants, deep learning models like CNN, LSTM, DCNN, and DBN are suggested. However, current methods lack predictive capabilities. After that, a hybrid machine learning/deep learning approach will be created to enhance prediction and classification performance. Modernization and intelligence in the agricultural sector will result from the integration of internet technologies like mobile terminal processors and agricultural IoT into the grain storage warehouse. This will enable continuous tracking and pest identification (Chithambarathanu and Jeyakumar, 2023). In agriculture, machine learning has proven to be a very useful tool, particularly for crop and soil monitoring, fruit grading, plant disease detection, and insect pest identification. The latest advancements have concentrated on computer vision-based quality inspection for the detection and classification of insects in grain storage conditions. The best results in pest classification and detection have been obtained by advanced machine learning models. Nevertheless, some insects in the wild provide distinct locations and comparable features, which makes identification and classification challenging. On top of the SVM and ANN algorithms, different parameter sets were used to create deep learning models with improved performance. This study is set out to classify and detect corn, soybean, and wheat pests at the early stages using machine learning techniques and algorithms for insect pest detection. The insects used different shape features in their classification, with image-processing techniques applied to segment the foreground insect and locate its position in the image (Kasinathan et al., 2021). Research has demonstrated that machine learning techniques and pattern recognition in various crop sciences have sparked interest in creating strategies for biotic stress identification and early detection. Benefiting from the identification of crop variability or heterogeneity brought on by biotic stressors like diseases or weeds is the main objective of applying machine learning techniques to precision crop protection. Weeds and plant diseases were identified and mapped using optical remotely and near-sensed data. Compared to conventional statistics-based discrimination models, machine-learning analytics can speed up and improve the accuracy of data analysis procedures (Behmann et al., 2014). Machine learning is making a difference in agriculture, from optimizing fertilization, irrigation, and pest control to ensuring the most efficient practices. Those machine learning models build detailed maps of crop growth, nutrient levels, and moisture from large amounts of data coming from satellite images, drone shots, and soil sensors. These maps can be used in regulating farming practices for maximum output and minimum wastage. It can also predict crop demand with respect to the market and environmental situations, thus showing the best times and places of planting. It can further gauge the quality of harvested crops by estimating the degree of ripeness and quality. Notwithstanding the challenges in lack of groundwork on data, high costs of sensors, and specialized expertise, the potential benefits of machine learning in agriculture are patent. The more farms that adopt precision agriculture, the more research has to be done to develop the full potential of the technology (Elbasi et al., 2023). The two biggest issues facing agriculture in the coming decades will be addressing climate change, which typically disrupts farming practices, making crop seasons unpredictable due to heat and water scarcity, and producing food for the expanding human and animal populations while maintaining food security by utilizing fewer agrochemicals and enforcing stringent controls at every point of the agricultural supply chain. It should advocate for transdisciplinary research and interdisciplinary collaborations while putting forward innovative solutions based on AI, ML algorithms, and other technological advancements. One of the key areas of the Ag5.0 revolution, which calls for new methods and approaches to significantly cut down on the use of agrochemicals to control diseases, weeds, and plagues, will be precision crop protection (Mesías-Ruiz et al., 2023). Smart agriculture and precision farming are crucial in the agriculture industry for reducing resource usage, water usage, and chemicals. These technologies require data collection and analysis, particularly web technologies, which handle vast volumes of data. Planning the use of soil and water, evaluating crop health, and reducing pollutants like pesticides and herbicides are all aspects of smart farming. Performance is dependent on computer vision technology and historical data. Classification is a common application of machine learning techniques in precision agriculture, including Support Vector Machine, Decision Tree, K-NN, Random Forest, Genetic algorithm, and Fuzzy logic. SVM is not appropriate for large or noisy data sets, but it is appropriate for both classification and regression tasks (Shaikh et al., 2022). Machine learning is self-learning, so it is strategy-oriented. The technology opposes challenges in the development of high-performance predictive models, and hence one has to choose the proper algorithm and critical features. There are a good number of attributes in large datasets, and, hence, in machine learning, dimensionality reduction is very important. There are two techniques to reduce features: feature selection and feature extraction. While Feature selection selects a subset of the original features containing relevant information; Feature extraction entails a transformation of the input space to a lower-dimensional subspace. Current research directions focus on the application of the feature selection algorithms by crop yield prediction models, including its assessment of the contribution, investigation of the best way against the features without selections (Pham et al., 2022). Automatic pest identification has become a significant research topic, with computer vision, machine learning, and deep learning technologies being used to detect plant diseases. However, comparisons between different techniques are often limited, with many works focusing on a specific technological approach. Advances in computer vision and object recognition have led to the Large Scale Visual Recognition Challenge (ILSVRC) as a benchmark for visualization-related problems. Traditional image classification approaches rely on features detection algorithms, which can be challenging to update. Deep learning techniques can solve this issue by automatically extracting features. Recent developments in machine learning and deep learning have significantly improved object recognition and detection accuracy, with machine learning making informed decisions based on input data, and deep learning making intelligent decisions using layers of structures (Gutierrez et al., 2019). Technological advancements have transformed agriculture, especially in response to climate change. Smart farming, a sustainable approach, uses technologies like sensors, IoT, artificial intelligence, and robotics to meet food needs while minimizing negative effects. Data mining and IoT-based agriculture systems help farmers adopt precautionary measures and customize crop management practices. Researchers are developing new algorithms for detailed information, including fertilizer application suggestions, disease prediction, soil moisture detection, and irrigation scheduling. These techniques are used in yield estimation, irrigation, fertilizer management, and insect, pest, and disease monitoring in crop production (Ali et al., 2023). Smart farming has combined the basis of IoT and IT to collect and organize data from different sources in order to predict and organize farm activities. This may include sensing of temperature, moisture, foliage, intensity, and direction of sunlight. Such data is fundamental in climatic disease forecasting, meteorological information, and crop protection. Data mining techniques are inseparable parts of big data patterns analysis, which can be used in identifying pests, detecting diseases, predicting yields, and planning fertilizer and pesticide uses. It can also be used in crop management and farm organization. Direct integration of ICT in agriculture domains affects agricultural efficiency and provides better results. The use of data mining techniques in agriculture domains is inseparable for bringing enhancements in agricultural efficiency (Shaikh, Mir, et al., 2022).

The main stages of image-based plant pathology are detailed in Figure 2.

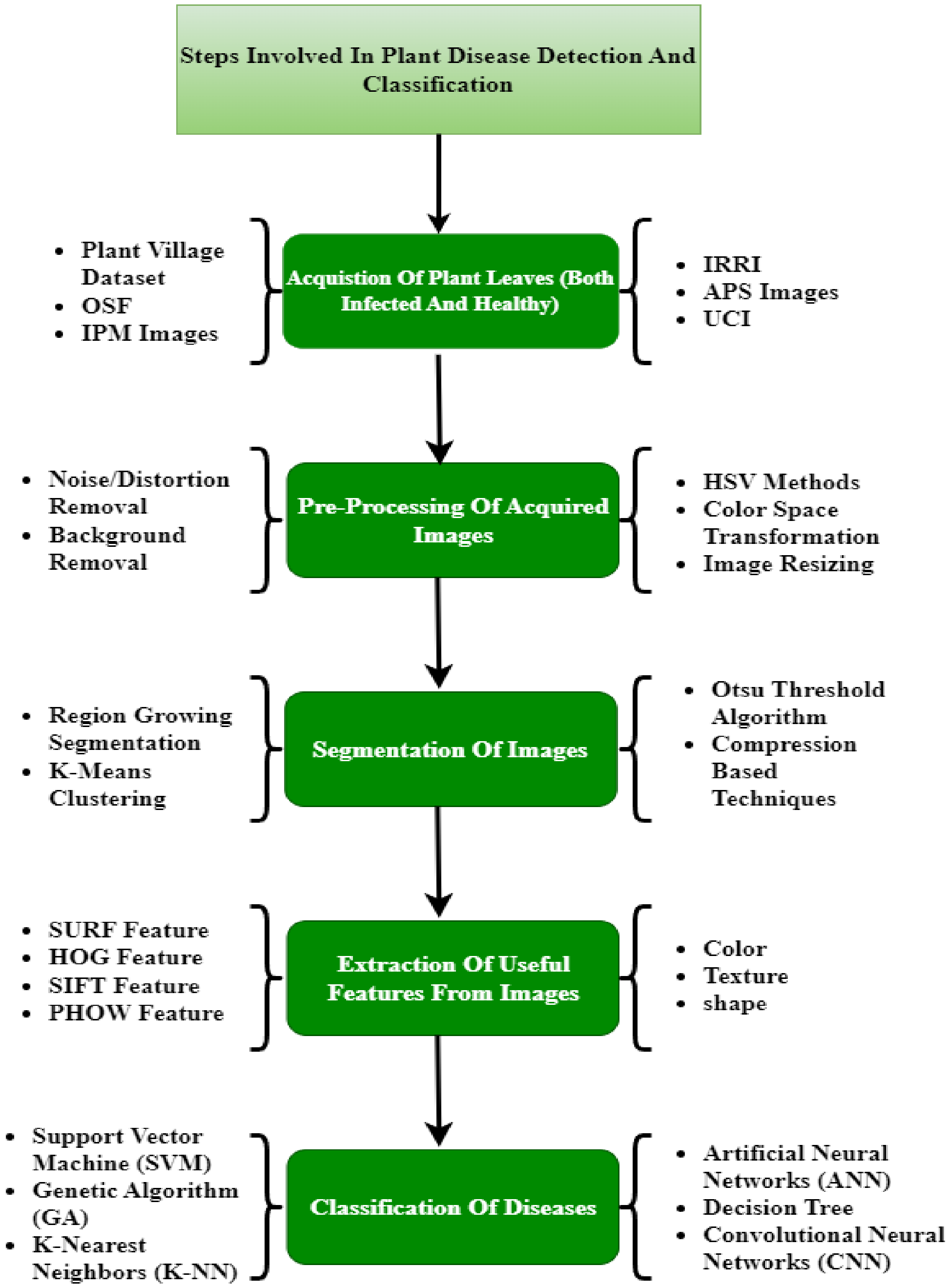

Figure 2. Steps involved in plant pathology using image analysis

Hyperspectral Imaging Processing and Machine Learning Techniques for Crop Disease Analysis and Identification: An Advanced Research Methodology

Hyperspectral images, originally used for military purposes, are now being used in precision agriculture for crop classification and disease detection. By combining them with machine learning techniques, these images can monitor crops like cereals, oilseeds, vegetables, and fruits, enabling early diagnosis of crop conditions like maturity index and nutrient status (García-Vera et al., 2024).

Hyperspectral ML approaches are grouped into major categories in Figure 3.

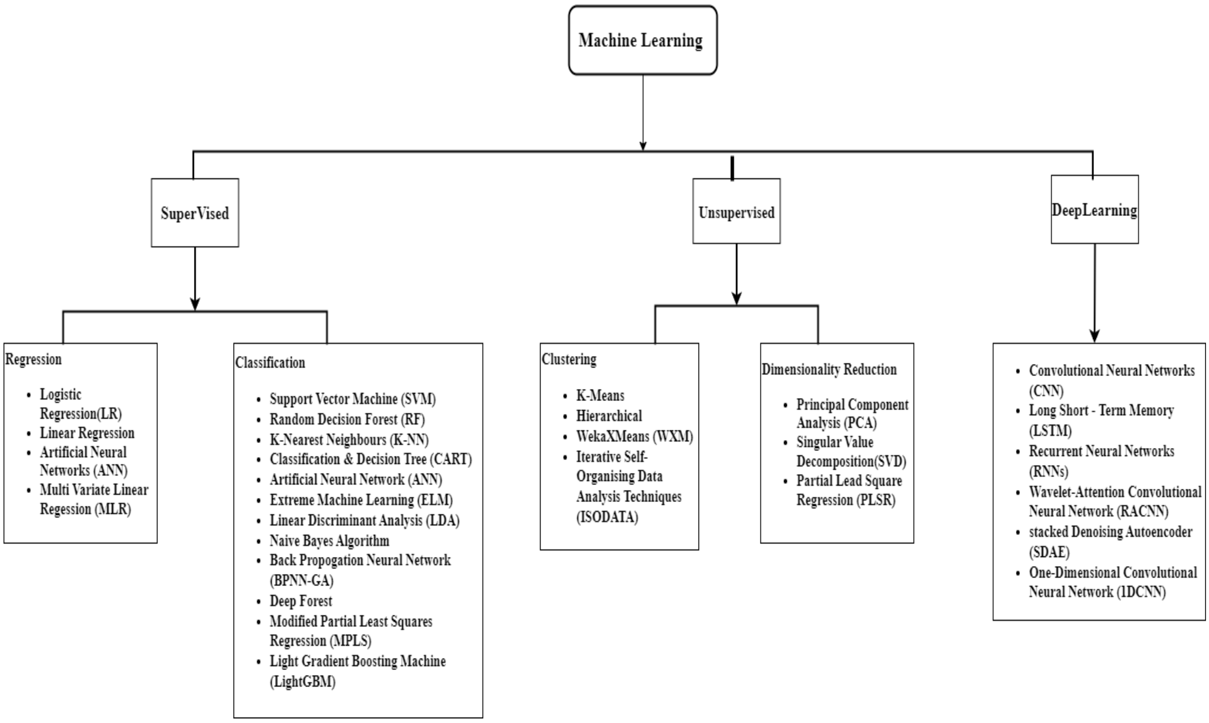

Figure 3. Classification of Machine Learning Approaches Hyperspectral Imaging Technology

Machine learning has already been combined with hyperspectral imaging quite successfully in case studies like viticulture. Grapevines are very prone to diseases like powdery and downy mildew; therefore, early detection and classification of these diseases are possible using this method (Eller, 2022).

Hyperspectral Plant Disease Diagnosis

It is being used in monitoring and prediction related to agri-food production through hyperspectral images for the sensing of biological microbial contaminants and diseases on the agricultural crops. These images realized in an integrated form with specialized sensors grant a plethora of information related to plant traits, soil nutrient levels, and state of disease for early diagnosis and early intervention in food supply production. (Xia and Wu, 2022) (Cheshkova, 2022).

Machine Learning

The identification of the severity of Verticillium wilt disease in olive crops was carried out using the algorithms: Support Vector Machine (SVM) and Linear Decomposition Algorithms (LDA). The SVM-based algorithm yielded an accuracy of 79.2% while the algorithm of LDA gave 59% accuracy (Calderón et al., 2015). Further, in the models for the estimation of chlorophyll in potatoes, different models were compared, which were analyzed for different growth stages, and finally an SVM approaches-based model was proposed, achieving a result with greater precision in the estimation of chlorophyll content (Applied Machine Learning,‖ 2023).

Supervised Machine Learning

This is a commonly used machine learning technique in which an algorithm predicts based on labeled examples. It can be used for a variety of analysis or forecasting tasks, such as crop analysis. Supervised machine learning models use labeled datasets, such as images or sensor data, to identify patterns related to crop traits, diseases, or quality factors (Panigrahi et al., 2023).

K-Nearest Neighbor (K-NN)

This instance-based algorithm is a method of machine learning that generates a prediction by finding k-nearest neighbors based on the similarity of data. This algorithm basically retains training examples only in memory and makes a prediction, explicitly not learning any model while training as other algorithms (Shuaibu et al., 2018).

Classification and Decision Tree (CART)

Decision trees have also proved to be very good machine learning toward classification tasks in agriculture, in particular, in hyperspectral imaging. They can do crop type classification, detect stress/disease, or monitor crop health by partitioning the data space based on the values of their features, making them particularly fitting for high-dimensional data from hyperspectral images (Baranowski et al., 2013).

Logistic Regression (LR)

A supervised learning algorithm for classification problems that uses the probability approach is called logistic regression. This algorithm is used in predictive analysis to identify healthy apples from those with bitter pit disease and predict thresholds. Unlike linear regression, which produces continuous values, logistic regression evaluates the relationship between the independent and dependent variables to produce discrete results (Jarolmasjed et al., 2018).

Linear Regression

Hyperspectral imaging uses spectral bands to capture detailed data on plant chemical composition. Linear regression aims to establish a linear relationship between independent variables like crop yield and dependent variables like chlorophyll content or disease presence/absence (Wang et al., 2021).

Multivariate Linear Regression (MLR)

The statistical method for expressing the relationship between multiple dependent variables and a single or multiple independent variables is called multivariate linear regression. It subsequently means that multivariate linear regression is an extension of simple linear regression, involving with only one dependent variable. Based on spectral data from hyperspectral sensors via MLR, hyperspectral imaging in crops can be used to predict agronomic variables pertaining to variables like biomass, chlorophyll content, moisture levels, and yield (Hastie et al., 2001).

Deep Forest

A novel machine learning model called Deep Forest has produced intriguing outcomes in image classification, which translates to remote sensing uses like hyperspectral imaging in agriculture. With a hierarchical structure for effective feature learning and representation, this method is especially suited to ensemble learning, much like decision trees. Crop types can be categorized by deep forests based on their spectral signature (Huang et al., 2023).

Backpropagation Neural Networks (BPNNs)

That said, the potential of BPNNs with GAs creates very powerful computational tools in hyperspectral imaging analysis in agriculture. They take advantage of the learning of complex patterns that neural networks do and make optimization on model parameters of greater value in precision agriculture applications. However, such gains in BPNN algorithms must be weighed against their increased computational cost, their hefty data requirements, and limitations on the interpretability of results (Golhani et al., 2018).

Linear Discriminant Analysis (LDA)

One of the popular statistical and machine learning techniques is supervised classification, which makes use of linear discriminant analysis in order to identify a linear combination of features and, hence, give an effective separation between multiple classes or groups in a dataset. This has been applied widely in agricultural products and weed discrimination using hyperspectral images (Calderón et al., 2015; Zhang et al., 2021).

Naïve Bayes Algorithm

Naïve Bayes is a potent machine learning classifier that describes outcome probabilities using conditional probability. Hyperspectral imaging systems with both visible and near-infrared sensors are used to classify agricultural crops and identify apple lesions. Utilizing remote sensors in the visible and near-infrared spectrum, it is also used to estimate nitrogen characteristics and comprehend rice crop development. This method is employed to comprehend how nitrogen and chlorophyll interact in crops (Wang et al., 2021b; Wang et al., 2021).

Modified Partial Least Squares Regression (MPLSR)

MPLSR is a statistical method for building predictive models that is primarily used with high-dimensional data sets, such as hyperspectral images. While keeping pertinent information for crop analysis, this method lowers the dimensionality of these data. In order to reduce the dimensionality of these data, MPLSR finds latent variables that capture the greatest variance in predictors and response variables (Cheng et al., 2006).

Light Gradient Boosting Machine (LightGBM)

One of the more valuable tools in solving problems is trees, due to their efficiency and scalability. They are especially useful in ranking, classification, and regression tasks, making them a valued solution to analyze hyperspectral image crops (Ke et al., 2017).

Unsupervised Machine Learning

Plant biology research is entering the big data era thanks to high-throughput omics technologies. Big data analysis relies heavily on machine learning (ML), which necessitates a large number of labeled samples. Getting enough labeled data requires the use of semi-supervised learning (SSL) and unsupervised learning (UL) paradigms. In this review, ML techniques, UL and SSL algorithms, recent developments, and their importance in plant systems biology and phenotyping research are introduced (Yan and Wang, 2022). Unsupervised machine learning is one technique that finds hidden patterns and structures of input data without labels or predefined responses; as a result, it will be helpful in situations where labels are not available or to reveal hidden information. It's an attractive alternative to traditional feature extraction methods, and significant progress has been reported for remote sensing image classification (Avola et al., 2023). Applications in the field of detectors of plant diseases are increasingly using machine learning. In this paper, a convolutional autoencoder-based unsupervised feature learning algorithm has been proposed for disease detection. This method does away with the necessities of handcrafted features, avoids labeling data, and makes the output from the auto-encoder act as input to SVM-based classifiers, hence much better than traditional auto-encoders with more hidden layers. This approach offers fast, inexpensive solutions for the diagnostics of plant diseases by tacking the challenge of high intra-variability and inter-variability in data (Pardede et al., 2018). In this respect, unsupervised machine learning methods have been utilized in precision agriculture to effectively process aerial image data. The technique was tested in rice fields, a plant nursery, and row crops for the estimation of accurate plant phenotypes without the requirement for supervised classification approaches. The method proves robust against ground elevational variations, thus efficient, needing only vehicle overflight, which, therefore, means its applications are quite promising in spectral plant pathosystems (Davis et al., 2020).

K-Means Clustering

It applies the k-means algorithm in arranging data into groups with high and low similarities in groups to eventually enhance the classification accuracy of PLDA models, such as categorized images of wheat crop ears (Almoujahed et al., 2022).

Hierarchical Clustering

Hierarchical clustering is a technique used in data mining to group data points into a dendrogram, and it is normally done using measures of similarity or distance. In relation to crop monitoring using hyperspectral images, this will help in the revelation of patterns and relationships, hence improving on agricultural outcomes through better crop health monitoring with early detection (Baker and Hubert, 1975).

WEKAXMeans (WXM)

WEKAXMeans is a Weka machine learning software-based X-means clustering algorithm for automatic determination of the optimum number of clusters based on the data, useful in the hyperspectral imaging of crops for classification or segmentation of crop regions based on spectral signatures (Aneece and Thenkabail, 2021).

Iterative Self-Organizing Data Analysis Technique (ISODATA)

ISODATA is a non-parametric, unsupervised methodology for the analysis of hyperspectral imaging data at different wavelengths, hence attaining a detailed view of the chemical composition of plants without any prior information on crops or health status (Memarsadeghi et al., 2007).

Dimensionality Reduction

Dimensionality reduction is indeed the most important technique in data analysis and machine learning because it reduces the feature set, conserving the essential information while avoiding computational complexity and fighting with overfitting (Hidalgo et al., 2019).

Principal Component Analysis (PCA)

PCA is a multivariate analysis technique that minimizes data dimensionality while preserving variability. It involves creating uncorrelated variables from potentially correlated ones, a method used for disease detection and classification in cereals like rice (Dong et al., 2014).

Singular Value Decomposition (SVD)

SVD is a potent tool for processing and analyzing hyperspectral images in agriculture. It improves crop health monitoring, type classification, and property prediction, resulting in more accurate and efficient management that eventually raises crop yield and quality (Lu et al., 2020).

Partial Least Squares Regression (PLSR)

One statistical technique used to model the connections between predictor and response variables is called PLSR. When determining micronutrients like calcium, magnesium, molecular iron, and zinc in cereals like wheat, it is highly helpful. Among the techniques used in dimensionality reduction and multivariate regression analysis, particularly in high-dimensional and multicollinearity datasets, is PLSR. It also incorporates elements of PCA (Geladi and Kowalski, 1986; Kandpal et al., 2015).

Deep Learning

Deep learning has made a revolution not only in data analysis but also in pattern recognition through various fields of computer vision, satellite image analysis, and precision agriculture by means of stacking multiple layers for learning hierarchical object features. Convolutional neural networks are the most used tool in the context of different computer vision tasks to capture higher-order semantic features from intermediate layers. Models like VGG, DenseNet, ResNet, and GoogleNet among others have given outstanding performance on topics pertaining to health informatics, remote sensing, drones, and natural language processing. CNNs are variants of neural networks both in terms of width and depth. (Shahi et al., 2023). Smartphone-assisted disease diagnosis is now possible thanks to the advent of deep learning and the widespread use of smartphones. Using 54,306 photos of 14 crop species and 26 diseases, they trained a deep convolutional neural network with a 99.35% accuracy rate on a publicly accessible dataset. This demonstrates that such a strategy can be applied globally (Mohanty et al., 2016).

Climate change and a lack of crop immunity increase crop diseases, resulting in crop destruction, decreased cultivation, and loss of money for the farmers. The identification and treatment of these diseases are quite problematic. Leaves' texture and visual similarities greatly help in the identification of diseases. Computer vision combined with deep learning can be of assistance in solving the problem (Kulkarni, 2018).

Convolutional Neural Network (CNN)

CNNs are powerful tools for the analysis of hyperspectral imagery in crops, allowing the handling of nonlinear relationships and retaining spatial information while conducting an automatic feature extraction. However, their computational cost, data requirements, and "black box" nature mean careful consideration is needed. CNNs are capable of capturing spatially variable issues in crops; this is very important, as it can help identify problems occurring in cereals like rice (Toda and Okura, 2019; Lu et al., 2021).

Recurrent Neural Networks (RNNs)

Time-series analysis and data point sequences, such as hyperspectral imaging in agriculture, can benefit from their use. RNNs enhance speech recognition and natural language processing applications by identifying spatial and temporal patterns in crop data recorded across multiple spectral bands (Gnanasaravanan et al., 2021).

Long Short-Term Memory (LSTM)

Particularly for time series analysis and sequential data, RNNs and LSTMs have shown themselves to be reasonably effective. They can be used to identify stressors like diseases, pests, and nutrient deficiencies in order to monitor crop health using hyperspectral imaging (Xiao et al., 2018). In this instance, the LSTM will pick up on the temporal trends of health indicators, providing early warnings along with useful information. Given that it takes into account variations in the spectral signature throughout the growing season, this will also aid in accurately predicting crop yield (Kim et al., 2017).

Stacked Denoising Autoencoder (SDAE)

SDAEs are deep learning techniques for feature extraction in the context of hyperspectral data analysis, particularly crop health monitoring and classification, where multiple layers of denoising auto-encoders are stacked on top, learning increasingly abstract representations of input data (Zilvan et al., 2019).

Residual Attention Convolutional Neural Networks (RACNNs)

In agriculture, convolutional operations are being used more and more for tasks like yield prediction, weed identification, and crop disease detection. Convolutional layers, wavelet analysis, and attention mechanisms are all part of this neural network architecture, which is intended to extract features from data, particularly in signal processing and image analysis (Weng et al., 2021).

Wavelet Attention Convolutional Neural Networks (WACNNs)

WACNNs are models that integrate wavelet transforms with an attention mechanism, leading to advanced feature extraction and improved model performance, useful in many agronomical tasks such as crop disease detection, soil moisture estimation, and plant classification (Bhattacharya and Pandey, 2024).

One-Dimensional Convolutional Neural Network (1D CNN)

One-dimensional Convolutional Neural Networks are a variant of neural network architecture tailored to process one-dimensional sequential data, be that time series or text data, as opposed to traditional CNNs, which are used with two-dimensional data (Pourdarbani et al., 2021). As shown in Table 1, deep learning architectures generally outperform traditional ML algorithms in multi-class plant disease detection.

Table 1. Comparative analysis of machine learning and deep learning techniques (Fahmi et al., 2018; Patrício and Rieder, 2018; Pallagani et al., 2019; Saleem et al., 2019; Rehman et al., 2019; Lee et al., 2020; Sujatha et al., 2021; Sambasivam and Opiyo, 2021)

Functions	**Techniques used**
Citrus leaf-disease detected	ML(SGD,RF,SVM)DL(Inception-V3,VGC-16,VGG- 19)
Isolated crops(maize,cassava,tomato,apple,wheat, citrus,potato)	DL(Convolutional neural networks)
57 diff seeds of weed species.	ML(supervised and unsupervised learning)
Palm oil plantation	GLCM(Gray level cooccurrence matrix method
Plants-image recognition (leaf)	Res Nets-Residual Neural Networks
Crop-disease Prediction (sugarcane)	PyTorch model
Grains:maize,rice,wheat,joyabean and barley-Disease detection and grain quality	AI(Deep belief networks DBN)
Plant village-tomato plant	DL architectures & Visualization AlexNet and Googlenet CNN
Cassava	Convolutional Neural networks

DISCUSSION

This review paper offers insight into the potential of machine learning and computer vision technologies in crop disease prediction, with emphasis on timely detection and accurate classification of infected plants. Such would offer a chance for the accurate prescription of treatments, thereby reducing associated costs by up to 90% in some infections. Thus, the efficiency of machine and deep learning algorithms for crop diseases diagnosis using image datasets is brought out.

Multispectral and hyperspectral imaging provide all information related to plant health and hence allow for accurate detection of the diseases. Also, thermal infrared imaging has great potential for bringing forth plant stress and disease symptoms when integrated with other imaging techniques. This paper reviews the potential of fusion models combining information of different origins to improve the accuracy and robustness of the predictions. For ease of dissemination and integration of these technologies into agricultural practices, automated tools have been developed. In the future, predictive models will not only predict diseases but also give recommendations for certain interventions. Machine learning and computer vision in agriculture can increase crop yields, reduce pesticide usage, enhance food security, and promote sustainable agriculture.

Future research directions would be aimed at enhancing data quality, improving model interpretability, improving scalability issues, and integrating these technologies with drones, sensors, and IoT devices. This review serves as a landmark platform for continued research and innovation in crop disease prediction using machine learning and computer vision that could foster a more sustainable and resilient agricultural future.

CONCLUSION

This comprehensive review shows the great strides that have been made in computer vision- based plant phenotyping and, from here, really can revolutionize agricultural practices. Applying machine learning algorithms on large image data sets with pre-trained models, recognition, taxonomic classification, and quality assessment of plants could be considerably automated. The results show just how far computer vision has made it possible to derive valuable data from plant images, hence opening up possibilities for precision agriculture. Pretrained model repositories have become integral resources that accelerate model development and increase the performance of models developed with these techniques. Deep learning techniques, especially convolutional neural networks, have been found to be quite efficient for most plant phenotyping tasks. Although deep learning techniques have made significant improvements in this area, there are still some bottlenecks related to data quality, model interpretability, and computational efficiency that must be overcome before these technologies find more widespread application in agriculture. In the future, it would be the integration of computer vision with other agri-tech tools, including drones and sensors, that would provide a holistic approach to precision agriculture. Developing species-specific or condition-specific models would improve performance, while ethical considerations are equally important as this technology will become prevalent. It is in this light that this review spots the huge potential of computer-vision-based plant phenotyping to increase agricultural productivity, sustainability, and efficiency, calling on researchers to keep pushing the boundaries and to solve at least some existing challenges for a more resilient agricultural future.

Conflict of Interest

The authors declare no conflict of interest

Authors contribution

Equal contributions was made by the authors

REFERENCE

Ali A, Hussain T, Tantashutikun N, Hussain N, Cocetta G., 2023. Application of smart techniques, internet of things and data mining for resource use efficient and sustainable crop production. Agriculture, 13(2): 397.

Almoujahed MB, Rangarajan AK, Whetton RL, Vincke D, Eylenbosch D, Vermeulen P, Mouazen AM., 2022. Detection of fusarium head blight in wheat under field conditions using a hyperspectral camera and machine learning. Computers and Electronics in Agriculture, 203: 107456.

Aneece I, Thenkabail PS., 2021. Classifying crop types using two generations of hyperspectral sensors (Hyperion and DESIS) with machine learning on the cloud. Remote Sensing, 13(22): 4704.

Applied Machine Learning., 2023. In MDPI eBooks.

Arinichev IV, Polyanskikh SV, Volkova GV, Arinicheva IV., 2021. Rice fungal diseases recognition using modern computer vision techniques. International Journal of Fuzzy Logic and Intelligent Systems, 21(1): 1–11.

Avola G, Matese A, Riggi E., 2023. An overview of the special issue on "Precision Agriculture using Hyperspectral Images." Remote Sensing, 15(7): 1917.

Baker FB, Hubert LJ., 1975. Measuring the power of hierarchical cluster analysis. Journal of the American Statistical Association, 70(349): 31–38.

Bali N, Singla A., 2021. Emerging trends in machine learning to predict crop yield and study its influential factors: a survey. Archives of Computational Methods in Engineering, 29(1): 95–112.

Baranowski P, Mazurek W, Pastuszka-Woźniak J., 2013. Supervised classification of bruised apples with respect to the time after bruising on the basis of hyperspectral imaging data. Postharvest Biology and Technology, 86: 249–258.

Behmann J, Mahlein A, Rumpf T, Römer C, Plümer L., 2014. A review of advanced machine learning methods for the detection of biotic stress in precision crop protection. Precision Agriculture, 16(3): 239–260.

Bhattacharya S, Pandey M., 2024. PCFRIMDS: Smart next-generation approach for precision crop and fertilizer recommendations using integrated multimodal data fusion for sustainable agriculture. IEEE Transactions on Consumer Electronics, 1.

Calderón R, Navas-Cortés J, Zarco-Tejada P., 2015. Early detection and quantification of Verticillium wilt in olive using hyperspectral and thermal imagery over large areas. Remote Sensing, 7(5): 5584–5610.

Cheng B, Wu X, Altair Engineering Company, Renmin University of China., 2006. An modified PLSR method in prediction. Journal of Data Science, 4: 257–274.

Cheshkova AF., 2022. A review of hyperspectral image analysis techniques for plant disease detection and identification. Vavilov Journal of Genetics and Breeding, 26(2): 202–213.

Chithambarathanu M, Jeyakumar MK., 2023. Survey on crop pest detection using deep learning and machine learning approaches. Multimedia Tools and Applications, 82(27): 42277–42310.

Das S, Pattanayak S, Behera P, Kardasz P., 2023. Application of machine learning: a recent advancement in plant diseases detection. Journal of Plant Protection Research.

Davis RL II, Greene JK, Dou F, Jo Y, Chappell TM., 2020. A practical application of unsupervised machine learning for analyzing plant image data collected using unmanned aircraft systems. Agronomy, 10(5): 633.

Dhanya V, Subeesh A, Kushwaha N, Vishwakarma DK, Kumar TN, Ritika G, Singh A., 2022. Deep learning based computer vision approaches for smart agricultural applications. Artificial Intelligence in Agriculture, 6: 211–229.

Domingues T, Brandão T, Ferreira JC., 2022. Machine learning for detection and prediction of crop diseases and pests: a comprehensive survey. Agriculture, 12(9): 1350.

Dong C, Ye Y, Zhang J, Zhu H, Liu F., 2014. Detection of thrips defect on Green-Peel Citrus using hyperspectral imaging technology combining PCA and B-Spline lighting correction method. Journal of Integrative Agriculture, 13(10): 2229–2235.

Elbasi E, Zaki C, Topcu AE, Abdelbaki W, Zreikat AI, Cina E, Shdefat A, Saker L., 2023. Crop prediction model using machine learning algorithms. Applied Sciences, 13(16): 9288.

Eller JA., 2022. Using hyperspectral remote sensing techniques to identify Vitis vinifera powdery mildew Napa Valley, California.

Fahmi F, Trianda D, Andayani U, Siregar B., 2018. Image processing analysis of geospatial UAV orthophotos for palm oil plantation monitoring. Journal of Physics Conference Series, 978: 012064.

Fenu G, Malloci FM., 2021. Forecasting plant and crop disease: an explorative study on current algorithms. Big Data and Cognitive Computing, 5(1): 2.

Paudel S, Nakarmi R, Giri P, Karki SB., 2022. Prediction of crop yield based-on soil moisture using machine learning algorithms. In 2022 2nd international conference on technological advancements in computational sciences (ICTACS). pp. 491-495. IEEE.

García-Vera YE, Polochè-Arango A, Mendivelso-Fajardo CA, Gutiérrez-Bernal FJ., 2024. Hyperspectral image analysis and machine learning techniques for crop disease detection and identification: a review. Sustainability, 16(14): 6064.

Geladi P, Kowalski BR., 1986. Partial least-squares regression: a tutorial. Analytica Chimica Acta, 185: 1–17.

Ghazal S, Munir A, Qureshi WS., 2024. Computer vision in smart agriculture and precision farming: techniques and applications. Artificial Intelligence in Agriculture.

Goldstein A, Fink L, Meitin A, Bohadana S, Lutenberg O, Ravid G., 2017. Applying machine learning on sensor data for irrigation recommendations: revealing the agronomist’s tacit knowledge. Precision Agriculture, 19(3): 421–444.

Golhani K, Balasundram SK, Vadamalai G, Pradhan B., 2018. A review of neural networks in plant disease detection using hyperspectral data. Information Processing in Agriculture, 5(3): 354–371.

Gopi PSS, Karthikeyan M., 2023. Multimodal machine learning based crop recommendation and yield prediction model. Intelligent Automation & Soft Computing, 36(1): 313–326.

Gosai D, Raval C, Nayak R, Jayswal H, Patel A., 2021. Crop recommendation system using machine learning. International Journal of Scientific Research in Computer Science Engineering and Information Technology: 558–569.

Gutierrez A, Ansuategi A, Susperregi L, Tubío C, Rankić I, Lenža L., 2019. A benchmarking of learning strategies for pest detection and identification on tomato plants for autonomous scouting robots using internal databases. Journal of Sensors, 2019: 1–15.

Hastie T, Friedman J, Tibshirani R., 2001. The elements of statistical learning. Springer Series in Statistics.

Hidalgo DR, Cortés BB, Bravo EC., 2019. Una herramienta para el análisis de índices espectrales para la detección remota de vegetación y cultivos utilizando imágenes hiperespectrales. Entre Ciencia E Ingeniería, 13(26): 51–58.

Huang H, Fei X, Hu X, Tian J, Ju J, Luo H, Huang D., 2023. Analysis of the spectral and textural features of hyperspectral images for the nondestructive prediction of amylopectin and amylose contents of sorghum. Journal of Food Composition and Analysis, 115: 105018.

Jackulin C, Murugavalli S., 2022. A comprehensive review on detection of plant disease using machine learning and deep learning approaches. Measurement Sensors, 24: 100441.

Jajur SM, Raju G., n.d. Crop recommendation using machine learning techniques. International Journal of Innovative Technology and Exploring Engineering, 9(2S): 658–661.

Jarolmasjed S, Khot L, Sankaran S., 2018. Hyperspectral imaging and spectrometry-derived spectral features for bitter pit detection in storage apples. Sensors, 18(5): 1561.

Kandpal LM, Lee S, Kim MS, Bae H, Cho B., 2015. Short wave infrared (SWIR) hyperspectral imaging technique for examination of aflatoxin B1 (AFB1) on corn kernels. Food Control, 51: 171–176.

Kasinathan T, Singaraju D, Uyyala SR., 2021. Insect classification and detection in field crops using modern machine learning techniques. Information Processing in Agriculture, 8(3): 446–457.

Ke G, Meng Q, Finley T, Wang T, Chen W, Ma W, Ye Q, Liu T., 2017. LightGBM: a highly efficient gradient boosting decision tree. NeurIPS Proceedings.

Khalili E, Kouchaki S, Ramazi S, Ghanati F., 2020. Machine learning techniques for soybean charcoal rot disease prediction. Frontiers in Plant Science, 11.

Khan A, Laghari A, Awan S., 2018. Machine learning in computer vision: A review. ICST Transactions on Scalable Information Systems, 169418.

Khan RU, Khan K, Albattah W, Qamar AM., 2021. Image-based detection of plant diseases: from classical machine learning to deep learning journey. Wireless Communications and Mobile Computing, 2021: 1–13.

Kim Y, Roh J, Kim H., 2017. Early forecasting of rice blast disease using long short-term memory recurrent neural networks. Sustainability, 10(2): 34.

Kulkarni O., 2018. Crop disease detection using deep learning. IEEE.

Lee SH, Goëau H, Bonnet P, Joly A., 2020. New perspectives on plant disease characterization based on deep learning. Computers and Electronics in Agriculture, 170: 105220.

Liakos K, Busato P, Moshou D, Pearson S, Bochtis D., 2018. Machine learning in agriculture: a review. Sensors, 18(8): 2674.

Liu J, Wang X., 2021. Plant diseases and pests detection based on deep learning: a review. Plant Methods, 17(1).

Lu B, Dao P, Liu J, He Y, Shang J., 2020. Recent advances of hyperspectral imaging technology and applications in agriculture. Remote Sensing, 12(16): 2659.

Lu J, Tan L, Jiang H., 2021. Review on convolutional neural network (CNN) applied to plant leaf disease classification. Agriculture, 11(8): 707.

Madhuri J, Indiramma M., 2021. Artificial neural networks based integrated crop recommendation system using soil and climatic parameters. Indian Journal of Science and Technology, 14(19): 1587–1597.

Moysiadis, V., Sarigiannidis, P., Vitsas, V., & Khelifi, A. (2021). Smart farming in Europe. Computer Science Review, *39*: 100345.

Memarsadeghi N, Mount DM, Netanyahu NS, Moigne JL., 2007. A fast implementation of the ISODATA clustering algorithm. International Journal of Computational Geometry & Applications, 17(01): 71–103.

Mesías-Ruiz GA, Pérez-Ortiz M, Dorado J, De Castro AI, Peña JM., 2023. Boosting precision crop protection towards agriculture 5.0 via machine learning and emerging technologies: a contextual review. Frontiers in Plant Science, 14.

Mochida K, Koda S, Inoue K, Hirayama T, Tanaka S, Nishii R, Melgani F., 2018. Computer vision-based phenotyping for improvement of plant productivity: a machine learning perspective. GigaScience, 8(1).

Mohanty SP, Hughes DP, Salathé M., 2016. Using deep learning for image-based plant disease detection. Frontiers in Plant Science, 7: 215232

Olalere M, Aimufua GI, Abdullahi MU, Egga BH., 2024. Ensemble-based predictive model for crop recommendation. Dutse Journal of Pure and Applied Sciences, 10(2c): 391–409.

Ouhami M, Hafiane A, Es-Saady Y, Hajji ME, Canals R., 2021. Computer vision, IoT and data fusion for crop disease detection using machine learning: a survey and ongoing research. Remote Sensing, 13(13): 2486.

Palakshappa A, Nanjappa SK, Mahadevappa P, Sinchana S., 2024. Smart irrigation with crop recommendation using machine learning approach. Bulletin of Electrical Engineering and Informatics, 13(3): 1952–1960.

Pallagani V, Khandelwal V, Chandra B, Udutalapally V, Das D, Mohanty SP., 2019. DCrop: A deep-learning based framework for accurate prediction of diseases of crops in smart agriculture. In *2019* IEEE international symposium on smart electronic systems (iSES)(formerly inis) (pp. 29-33). IEEE.

Panigrahi B, Kathala KCR, Sujatha M., 2023. A machine learning-based comparative approach to predict the crop yield using supervised learning with regression models. Procedia Computer Science, 218: 2684–2693.

Pardede HF, Suryawati E, Sustika R, Zilvan V., 2018. Unsupervised convolutional autoencoder-based feature learning for automatic detection of plant diseases. In 2018 international conference on computer, control, informatics and its applications (IC3INA) (pp. 158-162). IEEE.

Patrício DI, Rieder R., 2018. Computer vision and artificial intelligence in precision agriculture for grain crops: a systematic review. Computers and Electronics in Agriculture, 153: 69–81.

Pham HT, Awange J, Kuhn M., 2022. Evaluation of three feature dimension reduction techniques for machine learning-based crop yield prediction models. Sensors, 22(17): 6609.

Pourdarbani R, Sabzi S, Rohban MH, Hernández-Hernández JL, Gallardo-Bernal I, Herrera-Miranda I, García-Mateos G., 2021. One-dimensional convolutional neural networks for hyperspectral analysis of nitrogen in plant leaves. Applied Sciences, 11(24): 11853.

Pujar P, Kumar A, Kumar V., 2024. A survey on plant leaf disease identification and classification by various machine-learning techniques. IAES International Journal of Artificial Intelligence, 13(2): 1187.

Rao MS, Singh A, Reddy NS, Acharya DU., 2022. Crop prediction using machine learning. Journal of Physics Conference Series, 2161(1): 012033.

Raval MS, Chaudhary S, Adinarayana J., 2021. Computer vision and machine learning in agriculture. In: Studies in Big Data, pp. 97–126.

Reddy DA, Dadore B, Watekar A., 2019. Crop recommendation system to maximize crop yield in Ramtek region using machine learning. International Journal of Scientific Research in Science and Technology, 485–489.

Rehman TU, Mahmud MS, Chang YK, Jin J, Shin J., 2019. Current and future applications of statistical machine learning algorithms for agricultural machine vision systems. Computers and Electronics in Agriculture, 156: 585–605.

Roy AM, Bhaduri J., 2021. A deep learning enabled multi-class plant disease detection model based on computer vision. AI, 2(3): 413–428.

Gnanasaravanan S, Tharani B, Sahu M., 2021. Long short-term memory recurrent neural networks for plant disease identification. International Journal of Computer Communication and Informatics, 3(2): 51–62.

Saleem N, Potgieter N, Arif NM., 2019. Plant disease detection and classification by deep learning. Plants, 8(11): 468.

Sambasivam G, Opiyo GD., 2021. A predictive machine learning application in agriculture: cassava disease detection and classification with imbalanced dataset using convolutional neural networks. Egyptian Informatics Journal, 22(1): 27–34.

Shahi TB, Xu C, Neupane A, Guo W., 2023. Recent advances in crop disease detection using UAV and deep learning techniques. Remote Sensing, 15(9): 2450.

Shaikh TA, Mir WA, Rasool T, Sofi S., 2022. Machine learning for smart agriculture and precision farming: towards making the fields talk. Archives of Computational Methods in Engineering, 29(7): 4557–4597.

Shaikh TA, Rasool T, Lone FR., 2022. Towards leveraging the role of machine learning and artificial intelligence in precision agriculture and smart farming. Computers and Electronics in Agriculture, 198: 107119.

Shirahatti J, Patil R, Akulwar P., 2018. A survey paper on plant disease identification using machine learning approach. IEEE.

Shuaibu M, Lee WS, Schueller J, Gader P, Hong YK, Kim S., 2018. Unsupervised hyperspectral band selection for apple Marssonina blotch detection. Computers and Electronics in Agriculture, 148: 45–53.

Sinha A, Shekhawat RS., 2020. Review of image processing approaches for detecting plant diseases. IET Image Processing, 14(8): 1427–1439.

Sujatha R, Chatterjee JM, Jhanjhi N, Brohi SN., 2021. Performance of deep learning vs machine learning in plant leaf disease detection. Microprocessors and Microsystems, 80: 103615.

Suruliandi A, Mariammal G, Raja S., 2021. Crop prediction based on soil and environmental characteristics using feature selection techniques. Mathematical and Computer Modelling of Dynamical Systems, 27(1): 117–140.

Tian H, Wang T, Liu Y, Qiao X, Li Y., 2020. Computer vision technology in agricultural automation — a review. Information Processing in Agriculture, 7(1): 1–19.

Toda Y, Okura F., 2019. How convolutional neural networks diagnose plant disease. Plant Phenomics, 2019. DOI: 10.34133/2019/923713

Veeragandham S, H S., 2020. A review on the role of machine learning in agriculture. Scalable Computing Practice and Experience, 21(4): 583–589.

Wang L, Chen S, Li D, Wang C, Jiang H, Zheng Q, Peng Z., 2021. Estimation of paddy rice nitrogen content and accumulation both at leaf and plant levels from UAV hyperspectral imagery. Remote Sensing, 13(15): 2956.

Wang Z, Fan S, Wu J, Zhang C, Xu F, Yang X, Li J., 2021. Application of long-wave near infrared hyperspectral imaging for determination of moisture content of single maize seed. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 254: 119666.

Wani JA, Sharma S, Muzamil M, Ahmed S, Sharma S, Singh S., 2021. Machine learning and deep learning based computational techniques in automatic agricultural diseases detection: methodologies, applications, and challenges. Archives of Computational Methods in Engineering, 29(1): 641–677.

Weng S, Han K, Chu Z, Zhu G, Liu C, Zhu Z, Zhang Z, Zheng L, Huang L., 2021. Reflectance images of effective wavelengths from hyperspectral imaging for identification of Fusarium head blight-infected wheat kernels combined with a residual attention convolution neural network. Computers and Electronics in Agriculture, 190: 106483.

Xia L, Wu X., 2022. A review of hyperspectral remote sensing of crops. E3S Web of Conferences, 338: 01029.

Xiao Q, Li W, Chen P, Wang B., 2018. Prediction of crop pests and diseases in cotton by long short term memory network. In: Lecture Notes in Computer Science, pp. 11–16.

Yan J, Wang X., 2022. Unsupervised and semi-supervised learning: the next frontier in machine learning for plant systems biology. The Plant Journal, 111(6): 1527–1538.

Zhang Y, Xia C, Zhang X, Cheng X, Feng G, Wang Y, Gao Q., 2021. Estimating the maize biomass by crop height and narrowband vegetation indices derived from UAV-based hyperspectral images. Ecological Indicators, 129: 107985.

Zilvan V, Ramdan A, Suryawati E, Kusumo RBS, Krisnandi D, Pardede HF., 2019. Denoising convolutional variational autoencoders-based feature learning for automatic detection of plant diseases. In 2019 3rd International Conference on Informatics and Computational Sciences (ICICoS) (pp. 1-6). IEEE.