PLANT SCIENCE TODAY
Vol 12(sp3): 01–15
https://doi.org/10.14719/pst.8757

eISSN 2348-1900

REVIEW ARTICLE

Artificial intelligence and agriculture: Transforming plant breeding for a sustainable future

Sanku Gowthami[1], V Nirubana[2], M Jayaramachandran[3], M Theradimani[4] & S Juliet Hepziba[5]*

[1]Centre for Plant Breeding and Genetics, Agricultural College and Research Institute, Tamil Nadu Agricultural University, Coimbatore 641 003, Tamil Nadu, India
[2]Department of Plant Breeding and Genetics, Agricultural College and Research Institute, Tamil Nadu Agricultural University, Madurai 625 104, Tamil Nadu, India
[3]Department of Plant Breeding and Genetics, Agricultural College and Research Institute, Tamil Nadu Agricultural University, Chettinad 630 102, Tamil Nadu, India
[4]V.O. Chidambaranar Agricultural College and Research Institute, Tamil Nadu Agricultural University, Killikulam 628 252, Tamil Nadu, India
[5]Department of Genetics and Plant Breeding, V.O. Chidambaranar Agricultural College and Research Institute, Tamil Nadu Agricultural University, Killikulam 628 252, Tamil Nadu, India

*Correspondence email - juliethepziba.s@tnau.ac.in

Received: 08 April 2025; Accepted: 11 June 2025; Available online: Version 1.0: 11 November 2025

Cite this article: Sanku G, Nirubana V, Jayaramachandran M, Theradimani M, Hepziba SJ. Artificial intelligence and agriculture: Transforming plant breeding for a sustainable future. Plant Science Today. 2025;12(sp3):01–15. https://doi.org/10.14719/pst.8757

Abstract

Plant Breeding is a reliable assurance of agriculture productivity and nutritional security, while information technologies offer efficient means of fostering advancements in plant variety development. Processing large amounts of multidimensional breeding data over generations is complex, even though breeding information technologies provide an accessible and scientifically sound approach. Therefore, decision support tools help breeders to extract relevant and valuable information by introducing the golden seed breeding cloud platform. This platform is a cutting-edge AI driven agriculture system designed to maximize seed breeding for increased sustainability, resilience and yield. This brought a paradigm shift in farming by developing AI-powered solutions. To accomplish data integration and feature identification for stress phenotyping, it is necessary to take advantage of machine learning algorithms to extract patterns and features from the massive repository of data. Currently, plant breeding is propelling a revolution driven by state-of-the-art amenities for crop phenotyping and genome sequencing. Advanced phenotyping and genotyping when coupled with machine learning and cognitive sciences, are improving the accuracy of identifying the underlying genetic causes of attributes. Another advancement i.e., Next-generation AI using big data envisages how it can deal with the challenges by interfacing with the multi-omics big data to accelerate plant breeding, particularly for climate-resilient agriculture. Successful implementation of the proposed model based on big data characteristics will facilitate the evolution of breeding from "art" to "science" and eventually to "intelligence" in the era of Artificial Intelligence.

Keywords: AI; crop improvement; machine learning; next-gen AI; plant breeding

Introduction

Agriculture and livestock husbandry have existed since the dawn of human civilization, when agricultural techniques were dependent on animal power. It persisted in the industrial revolution, when animal power was replaced by combustion engines, in tandem with synthetic fertilizers and pesticides (1). Following this, a variety of contemporary technologies were integrated, including sensors, geolocation systems, computer programs, etc., with the goal of determining the optimal amounts, timing and locations for the use of livestock, agricultural products and practices. Increasing agricultural yield remains a difficult task considering the world's growing population and the increasing effects of climate change-related weather patterns. Improvements in selection intensity, precision, additive genetic variation and generation time frame are necessary to achieve genetic gain, which is a measure of enhanced crop productivity over time, according to Breeder's equation. The applications of modern innovations coupled with plant breeding, resulting in high genetic gain is depicted in Fig. 1. Currently, agriculture production systems are facing disconcerting challenges worldwide, including climate change, declining water supply for irrigation, increasing production costs and an overall reduction in the farm workforce over the past several decades (2). Besides, a research gap exists in analyzing large volumes of phenotyping data and in extracting important and useful components. Managing vast quantities of multidimensional data with high accuracy has been a major challenge. Scientific and technological innovations are essential to secure production systems by utilizing state-of-the-art technologies and exploring ways to integrate them (3). This integration is still formidable today, with the incorporation of technologies like the Internet of Things, Big Data and Cloud Computing. Dynamic crop simulation models have been

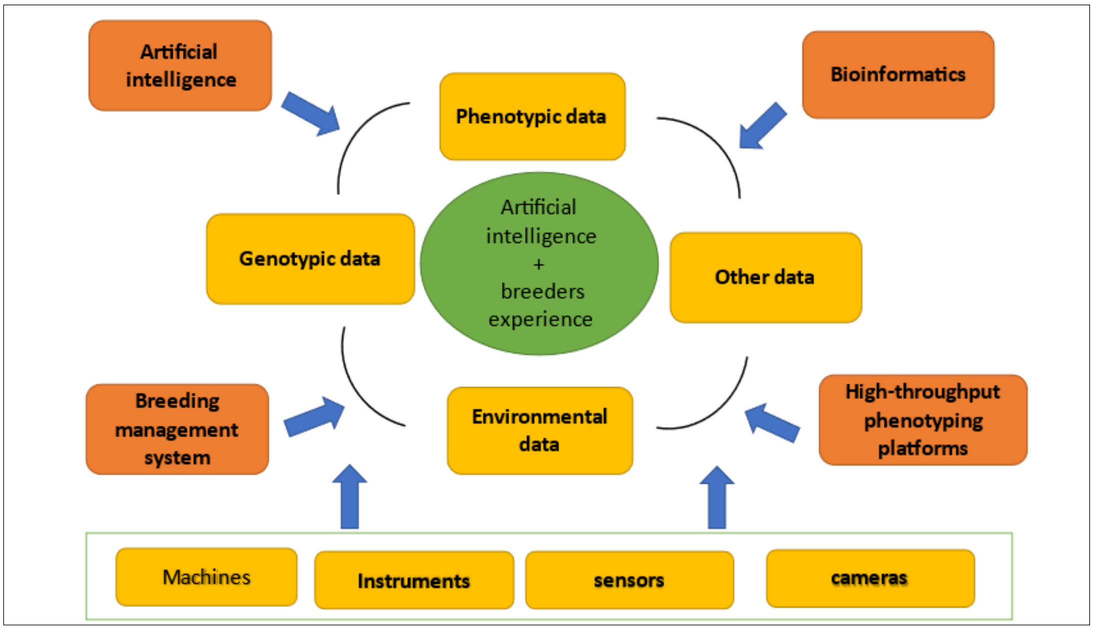

Fig. 1. Upcoming plant breeding in crop improvement.

practical tools for integrating distinctive components of agricultural systems, allowing us to explore how those components function within the system (4). The current review emphasizes the importance of Artificial Intelligence (AI) in agriculture, explaining how to overcome the existing gaps with higher precision.

AI is an approach that enables a computer, a robot, or any machine to think the way human thinks and resolve problems (5, 6). AI, according to Professor McCarthy, is "the science and engineering of developing intelligent devices, especially intelligent computer software". AI technology based on big data is deemed critical in the shift from Breeding 3.0 to Breeding 4.0 (7). In genomics and phenomics, big data has become a reality, but existing AI systems still have interpretability issues and cannot effectively produce information from overwhelming and chaotic data (8). Future AI technologies are designed to gradually tackle these difficulties while making good use of highly heterogeneous and complex data, bridging the genotype-phenotype gap and promoting breeding progress (9). AI is recently gaining significant attention within agriculture disciplines because of its potential to leverage big data. This is achieved through numerous tools such as machine learning, deep learning, image processing, artificial neural networks, Wireless Sensor Network (WSN) technology, wireless communication, robotics, Internet of Things (IoT), different genetic algorithms, fuzzy logic and computer vision.

Machine learning is a subset of AI that has undergone quick and substantial advancements (10). It solves problems that are too difficult for humans to solve by delegating decision -making to the algorithm (11-13). Machine learning, as defined by AI pioneer Arthur Samuel in 1959, is the "area of study that allows computers to learn without being explicitly programmed." Machine learning aims to create a prediction engine for a specific use case by building software for each sort of item that needs to be identified (14, 15).

Artificial Neural Networks (ANNs) are a type of deep learning that has three levels: input, output and several layers called "deep neural networks or hidden layers". a single node in input layer represents a specific feature and transfers the received data to the hidden layers (16); the received data is then processed in the hidden layers through an array of weighted connections deploying activation functions to produce an output (17, 18). The processed data is subsequently passed on to output layer to transform the activations into the desired format (16). ANN was inspired by the biological nervous system, particularly the brain, as an information-processing paradigm (19, 20). ANNs are models that emulate a network of biological neurons, which can quickly process a large amount of data and recognize patterns based on self-learning (21). The neural network is constructed from three types of layers, *viz.*, input, hidden and output layers (Fig. 2).

AI-powered deep learning techniques are increasingly used in remote sensing applications (22). These technologies produced significant gains in a variety of fields and piqued the interest of both academic and industrial communities. The most extensively used deep learning approach for image recognition is deep convolutional neural networks (CNNs). These networks require a significant amount of input, to construct hierarchical features and offer semantic information in the output. With the increasing availability of large volumes of aerial imagery from UAVs (Unmanned aerial vehicles) and satellites, CNNs can play a key role in analyzing this data and extracting useful information. The main objective of this study is to explain the applications of AI and its tools in crop breeding and improvement with its ability to analyze big data employing various algorithms and digital processing technologies.

Artificial Intelligence

AI is an intelligent entity created by humans capable of thinking rationally and humanely without being explicitly instructed. As defined by McCarthy, "Every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans and improve themselves".

Explaining or forecasting phenotypes from underlying genotypes across multiple environmental variables has become a challenge in fundamental and practical research as

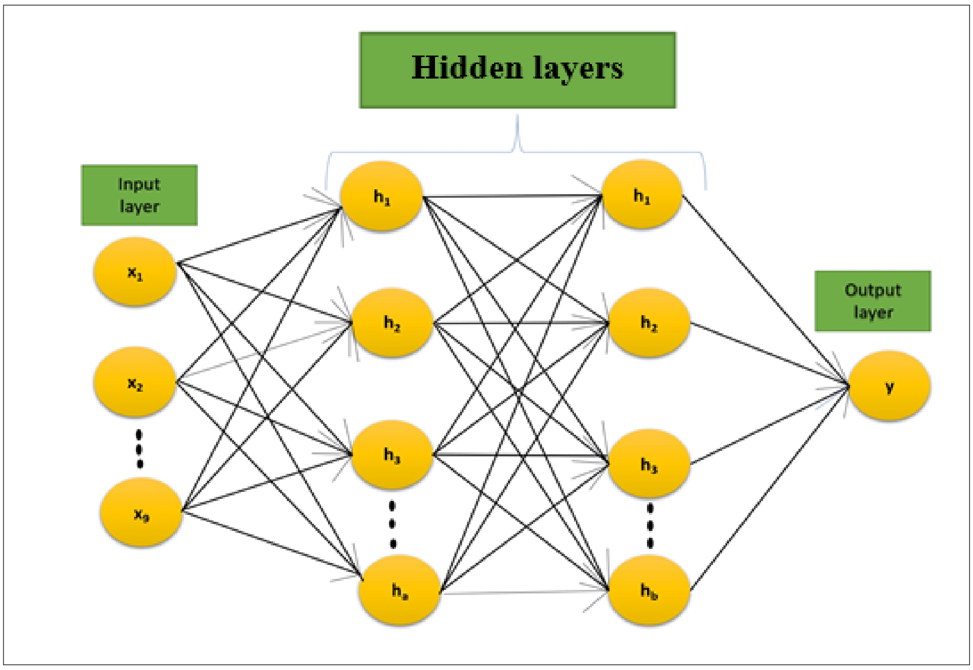

Fig. 2. Neural network architecture.

well as breeding applications since the dawn of the "big data" age in plant sciences (23). Genotype variation, on the other hand, causes differences in cell biochemistry, which affects plant development and organ formation. In the end, it integrates agriculturally significant qualities with the environment, such as pest, stress tolerance and yield. Analyzing environmental influences as well as phenotypic variation resulting from genotypic differences might help manage important plant physiology and development processes. It is crucial for current molecular plant breeding to determine quality and yield traits expressed by genotypes under specific environmental conditions. Examining multiple-level phenotypes or relating these phenotypes to genotypes, demands the assimilation and processing of ever-larger, diverse and noisy datasets (23). However, AI is a creative technique that simulates human intelligence and capability processes using technology, such as most computer systems, robots and digital equipment (24).

Plant breeding is being accelerated by interdisciplinary approaches. Phenomics, genomics, epigenomics, 3D genomics, transcriptomics, metabolomics, proteomics and ionomics have all benefited from technological developments in various disciplines as presented in Fig. 3. Furthermore, improvements in big data technologies and computer science have aided AI research significantly. By using AI integrated panomics, researchers can better understand how phenotypes are determined, allowing for more precise breeding (25). In the field of omics, AI is known to generate significant interest, with machine learning primed to flourish and deep learning expected to drive the AI boom.

Artificial intelligence in phenomics

To effectively run the machines for a particular task, AI methods were known to use large repositories of data. To strengthen early detection and improve decision-making, AI systems require vast volumes of data for training (26). In non-destructive phenomics, the data collection process entails combining data from instruments/sensors (such as digital cameras and spectrometers), which often operate using proprietary network topologies, into AI algorithms. AI is described in three primary facets of phenomics data: i) algorithms and strategies for processing sensory messages into phenotypic information; ii) formulating models to comprehend genotype-phenotype relationships alongside ecological factors and iii) administering databases for disseminating knowledge and resources. The important characteristics of AI, machine learning (ML), deep learning (DL) and computer vision have all been used in phenomics to a certain degree, as shown in the flow chart (Fig. 4).

Machine learning

ML stands for machine learning, which is a collection of computerized algorithms for modeling that can uncover associations in data and make autonomous judgments without having to implement explicit rules. The basic goal of machine learning is to find underlying structures, similarities and differences in data to appropriately describe or classify new events or scenarios. The capacity of ML technologies to extrapolate the patterns from given data is one of its most important capabilities. In both conventional and *in vitro* plant breeding studies, the field of machine learning which trains computers to learn from data has been widely employed to examine a stream of information about plants, from DNA sequence to observed phenotypes. Machine learning methods are divided into two categories: supervised and unsupervised models, which are further categorized into classification regression and clustering, respectively as illustrated in Fig. 5.

Nonlinear nonparametric machine learning techniques used in plant research include artificial neural networks (ANNs), deep neural networks (DNNs), convolutional neural networks (CNNs), random forest (RF), support vector machines (SVMs), multilayer perceptron (MLP) and radial basis function (RBF) (27). These data-driven models can parse and comprehend non -normal, nonlinear and nondeterministic data sets by using full spectrum data while avoiding unnecessary spectral bands and multi-collinearity (28). The pros and cons of these models for their applicability in classical and plant breeding studies are listed in Table 1. Supervised learning is rather suited and

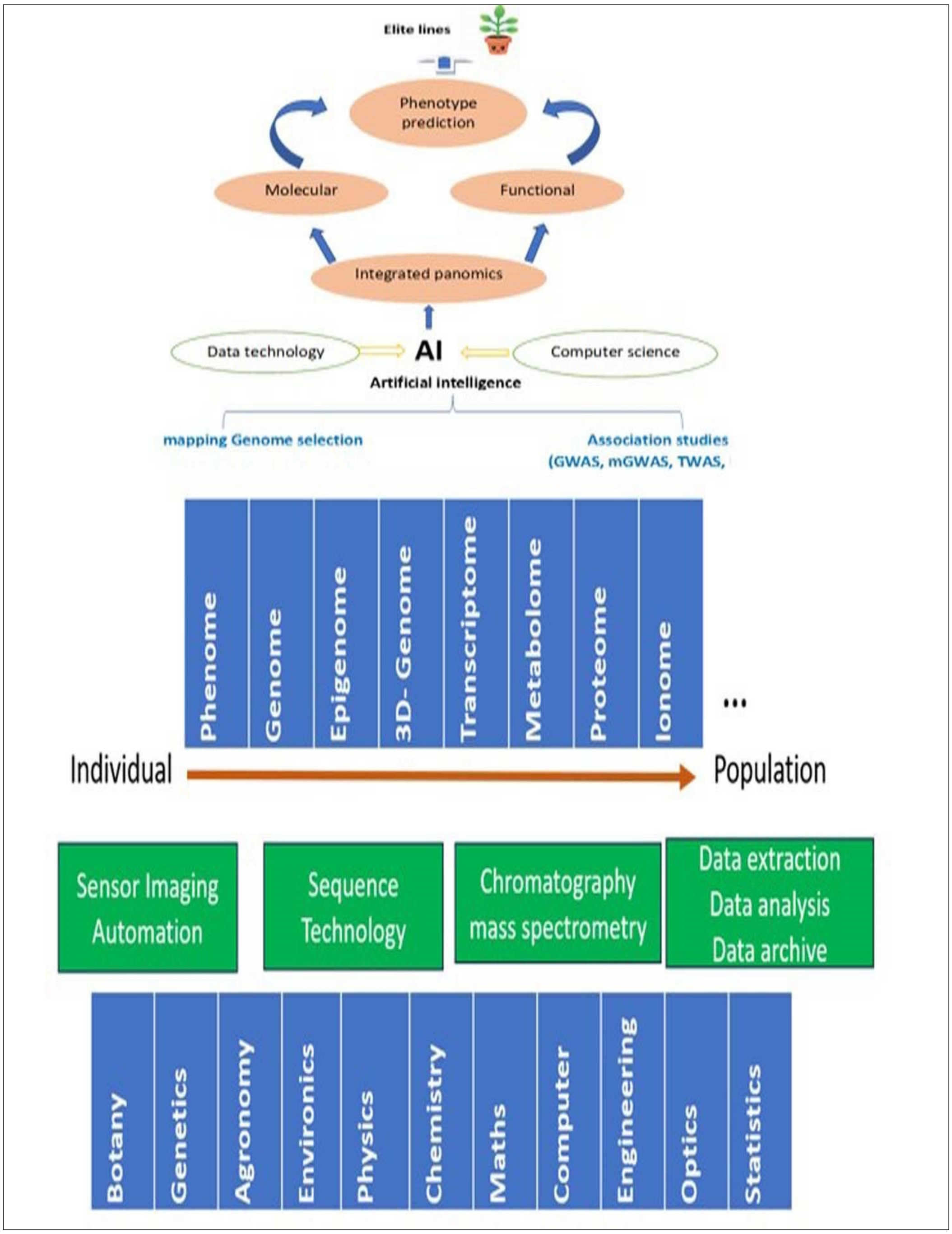

Fig. 3. Integrated panomics with artificial intelligence in plant breeding.

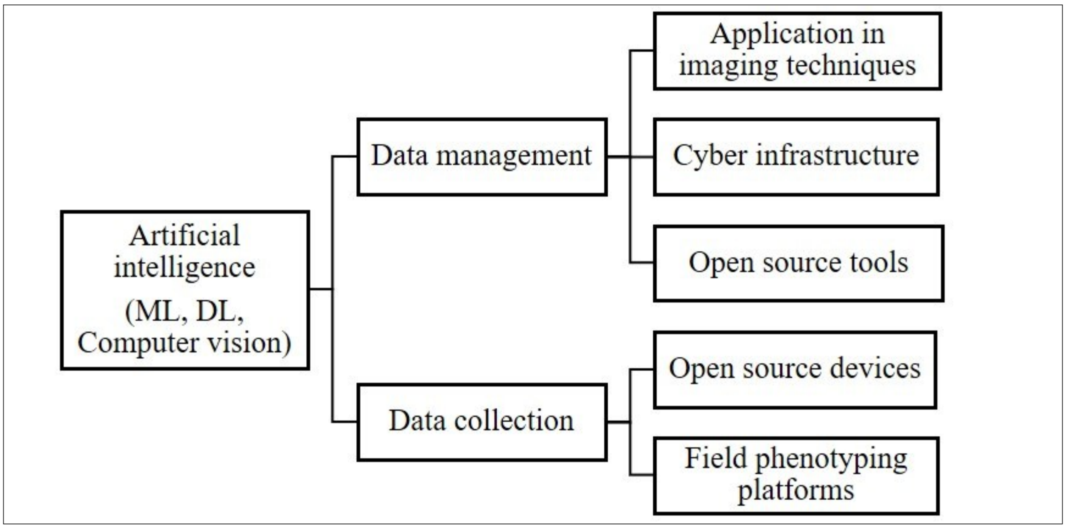

Fig. 4. Applications of artificial intelligence.

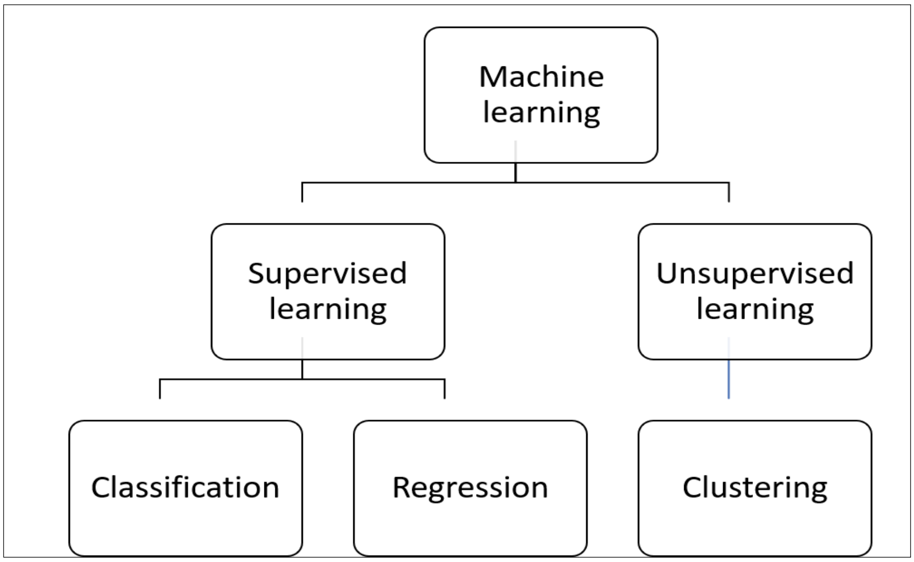

Fig. 5. Classification of machine learning.

Table 1. List of ML algorithms in classical and plant breeding studies.

Algorithm	Pros	Cons
Artificial neural networks (ANNs)	Effective learning abilities	Inability to interpret In a small number of training data, generalization error and local minima can occur using only a finite count of hidden neurons
Convolutional neural networks (CNNs)	Streamlined feature retrieval	Inadequate interpretation abilities Massive quantities of data are required for training, as well as substantial skill and experience in choosing appropriate hyperparameters
Multilayer Perceptron (MLP)	Manages large data and small data with equal accuracy Quick prediction after training	The proper functioning of the model depends on the quality of the training.
Random Forest (RF)	Interference tolerance Managing a large number of features forgoing overfitting	The architecture was indeed shallow
Support vector machines (SVMs)	Employs a large number of hidden units. A quadratic optimization task is required.	The architecture was indeed shallow
Radial Basis Function (RBF)	Faster requires fewer interactions Easy to parameterize	Require more time for classification

reliable than unsupervised, reinforcement, sparse dictionary or rule-based learning (29). Classification (predicting non-numeric responses) and regression can both benefit from supervised learning (predicting numeric answers) (30).

Machine learning algorithms may analyze abstract datasets such as data received by optical imaging or sequencing (31). Transposable elements can be spotted and characterized through machine learning approaches with genomic sequencing data (32). Breeders can predict many outputs (multiple-dependent variables) from multiple juxtapositions of numerous inputs in one model using machine learning techniques, thereby reducing the number of required analyses (Fig. 6). Each machine learning algorithm is known to have a specific artificial neural network to process the input through a series of hidden layers, using activation functions within the layers to provide the required output.

Artificial neural networks (ANNs) are non-linear nonparametric models that is capable of handling incomplete data and do not necessitate prior data structures or detailed comprehension of the physical processes to be simulated (33). Deep neural networks (DNNs) offer more predictive potential than ANNs due to their numerous hidden layers. The innate visual perception systems of living things serve as the inspiration for convolutional neural networks, a forefront deep learning architecture featuring a combination of multilayer, amalgamation, fully interconnected layers and an output layer (34). Because of their automatic feature extraction, CNNs are ideal for categorization investigations. CNNs are commonly used for image classification, object identification, object tracking, posture estimation, text character recognition, feature extraction detection, activity recognition, scene tagging, voice and natural language processing. The accuracy of features in neural networks is low, especially in CNNs, where the extracted virtues are concealed. SVMs are a more advanced machine learning technique for grouping, categorization and predictive analysis of data sets. They use a supervised learning algorithm to detect both linear & nonlinear associations in data. In comparison to the multilayer perceptron of ANN, SVM employs a greater sum of obscured units and performs better in the specification of the learning confront, followed by the algebraic optimization technique. Using a multitude of decision trees to categorize data, random forest regression is a tree-based machine learning paradigm for regression tasks. It entails choosing the number of trees, the frequency of random features and the training stop criterion. RF is better suited for spectral data analysis and overfitting can be avoided by combining multiple independent risk factors (35, 36). Deep learning CNN can be utilized more efficiently in semantic segmentation approaches such as automated phenotyping and plant disease diagnosis over superficial learning models of

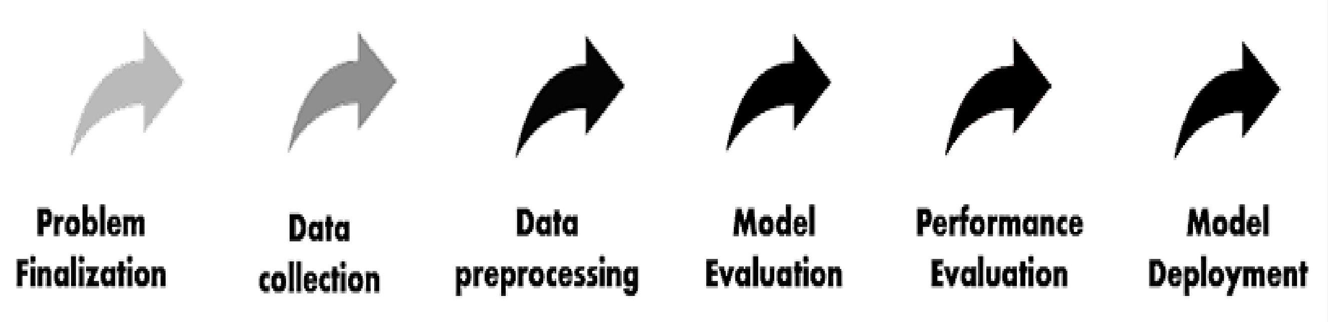

Fig. 6. Steps involved in the employment of machine learning model.

SVMs and RF. The challenge of requiring significant manually crafted features can be solved by using image augmentation and a small explicitly summarized factual dataset to fine-tune a synthetically bootstrapped CNN (37). Deep convolutional neural networks have been generally employed in biotic & abiotic stress diagnosis and classification by combining image feature extraction and classification within a single conduit (38). Based on a data set comprising 39 distinct classes of plant leaves and background images, a nine-layer deep CNN model was programmed to figure out plant leaf diseases. A classification accuracy of 96.46 % was documented, surpassing that of more conventional machine learning techniques like SVM, decision trees, logistic regression and K-nearest neighbour (KNN). In addition to this, CNNs can also be used in remote sensing for object detection and similarity analysis. It has been reported that deep CNN-based segmentation with data extracted from visual assessment of UAV-based high-resolution RGB imagery can achieve maximum accuracy (84 %) for fine-grained mapping of vegetation species and communities (39).

A thorough review of the uses of these nonlinear machine learning approaches in *in vitro* plant breeding research is explained in following sections of this article as illustrated in Fig. 7.

Genetic diversity assessment

Classical multivariate analyses, such as cluster analysis, discriminant function analysis and principal component analysis (PCA), are often utilized to characterize & group genotypes in various plant species using morphological, biochemical, physiological and molecular markers (40-43). Deep learning algorithms for object detection can be used to proficiently assess genetic diversity and categorize plant genotypes. CNN is found to be a feasible substitute to the conventional classification methods such as a K-nearest neighbour, probabilistic neural network, support vector machine, genetic algorithm and PCA, which are often labor-intensive and require lengthy feature extraction (44,45). The genetic diversity of 90 soybean accessions (*Glycine max* (L.) Merr.) was revealed using high-throughput stomatal density analysis (46). Deep CNN was used to distinguish between morphologically similar species in *Cinnamomum osmophloeum Kanehira* (Lauraceae), with CNN classifiers outperforming SVM classifiers (96.7 % *vs.* 74.6 % (47). To evaluate genetic diversity and categorize ten plant populations, a comparative analysis of ANN was performed with Fisher's classical multivariate statistical technique and Anderson's discriminant functions (48). They unearthed that ANN-classified populations with high and low differentiation performed beyond the traditional methods, with fewer incorrectly classified individuals.

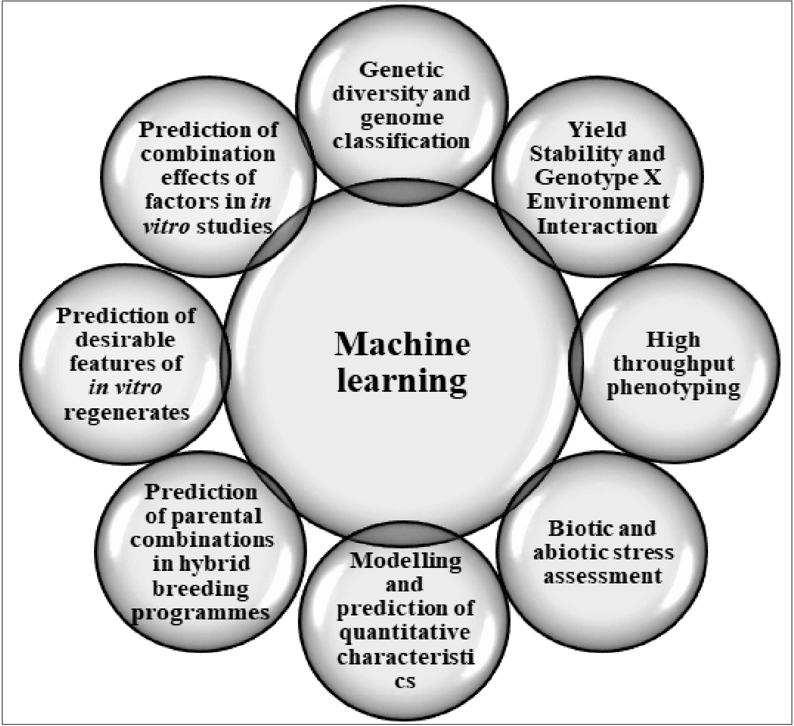

Fig. 7. Application of machine learning in crop breeding.

Yield Stability and Genotype X Environment Interaction

Principal component analysis (PCA), cluster analysis, additive main effect and multiplicative interaction (AMMI) as well as GGE biplot are multivariate procedures that allow for the exploration of the GEI's multifaceted attributes by illustrating a genotype's reaction in an E-dimensional space. Multivariate stability analyses outperform univariate approaches in terms of power and precision. A prime constraint of these methods is a lack of available software (49). In a recent study, both linear and nonlinear regression models were used to estimate the influence of climate variables on growth and yield-related characteristics. Linear regression equations typically offer lower interpretability than that of nonlinear equations (50). The resultant effect of the environmental variables on the cotton growth metrics at numerous locations had not been determined, nor were the optimal equations for the cotton growth indices identified in single research. This was challenging due to a relatively complex nonlinear relationship or a linear relationship between the indicators of cotton growth and the variables. In addition, one regression model was constructed by the authors for all possible circumstances. When several explanatory variables & predictor variables re involved, ANN can reduce the number of assessments required and contribute to increased accuracy (51). Comprehensive plant-environment interactions can be comprehended through the amalgamation of AI with deep phenotyping (7).

Biotic and abiotic stress assessment

Plants are known to experience stress from both biotic and abiotic factors. To evaluate plant genotypes' resilience and adaptation to various stresses and to pinpoint superior genetic variants, a variety of practises have been deployed. Combining phenomic, metabolomic and genomic data is an effective approach for evaluating plant responses to abiotic and biotic stresses (51). Traditional multivariate statistical methods are inefficient for dealing with such large datasets. Linear regression is the most frequently used approach designed to detect nutritional deficiencies using the RGB image technique. However, linear regression cannot explain features extracted from digital images that have a curvilinear relationship with nutrient content (52). By analysing all phenomic and omics (metabolomics and genomics) data, machine learning techniques and digital images could be used as a model to predict genotypes' responses to stressful conditions as well as to identify those conferring resistance to stress and non-stress environments. Machine learning can be used for high-throughput stress phenotyping interpreting large data imaging and remote sensing data (53).

Convolutional Networks have indeed been frequently used to detect and quantify biotic stress in plants (54, 55). Deep CNN was used to detect and recognize maize leaf diseases such as northern corn leaf blight (*Exserohilum*), common rust (*Puccinia sorghi*) and grey leaf spot (*Cercospora*), with an accuracy of 93.35 % (56) and have been applied for banana disease and pest detection (57). A semantic segmentation model based on CNN was devised on cucumber (*Cucumis sativus*) to sequence the powdery mildew disease on leaf images at the pixel level., The CNN model achieved a pixel accuracy of 96.08 %), outperforming other segmentation methods such as K-means, random forest and GBDT. DNNs were used to identify mildew disease (*Sclerospora graminicola*) in pearl millet and the developed model achieved an accuracy of 95.00 % (58).

Machine learning applications in in-vitro based plant biotechnology

Through the process of forecasting the optimal blend of these supplements in conjunction with other decisive elements like the genotype of the plant, the ambient conditions of the donor plants, physical treatments (inductive stresses) applied to the cultured gametophytic cells, the initial gametophytic cells' developmental stage and the constituents of the medium in which they are cultured, ANN models may resolve the issue of resistant species/genotypes by increasing the effectiveness of *in vitro* doubled haploid development. The ANN anticipated the callus induction percentage in androgenesis (anther culture) of tomato (*Lycopersicon esculentum* L.) more accurate than the MLR model under the influence of plant genotype, 2,4-D, kinetin and gum arabic concentration (59).

Ploidy level determination through image processing

The identification of plant ploidy levels can be accomplished using machine learning algorithms. Recently, a deep learning-based feature detection engine was designed to quantify stomatal density and analyze variations in stomatal CNN was trained using *R1-nj* anthocyanin color marker data derived from 1230 haploid and 1770 diploid maize seed images to distinguish between haploid and diploid seeds. The ploidy of quaking aspen (*Populus tremuloides* Michx.) has also been determined using remote sensing technique (61).

Another method for determining plant ploidy levels involves analyzing visual data from the cellular pattern of the epidermal layer. Compared to doubled haploids, haploids have narrower, more densely packed epidermal and mesophyll cells (more cells per same unit area). As a result, haploids and their diploid counterparts have different amounts of DNA content per unit leaf area (62). As a result, epidermal cell patterning (size, shape and number) could be used as ploidy level classification and segmentation criteria because of its uniqueness in each ploidy group. In chromosome engineering studies across various vegetation, an image-driven model has been developed to estimate ploidy levels. These kodels use imaging modalities for precise image retrieval of leaf punch samples (the cellular pattern, including cell size and number) and subsequent modelling of captured images using deep learning approaches, particularly CNN (Fig. 8). This is a more precise, quick and cost-effective approach to ploidy differentiation that might also be applied in other fields.

Machine learning in genomics

Evaluating phenotypes obtained at multiple phases or increasing phenotypic-genotype correspondence necessitates the handling and embedment of massive, noisy and diverse data sets. Machine learning (ML), a collection of computer algorithms to track down the data patterns, is becoming increasingly crucial in these endeavours (63-65). ML has driven a wave of recent advances in a variety of scientific and technical disciplines and is now poised to apply the same principles to plant science (66). ML facilitates breakthroughs in plant research and plant breeding by focusing on applications in biochemical and macroscopic studies, as well as by correlating genotype data and phenotypic traits.

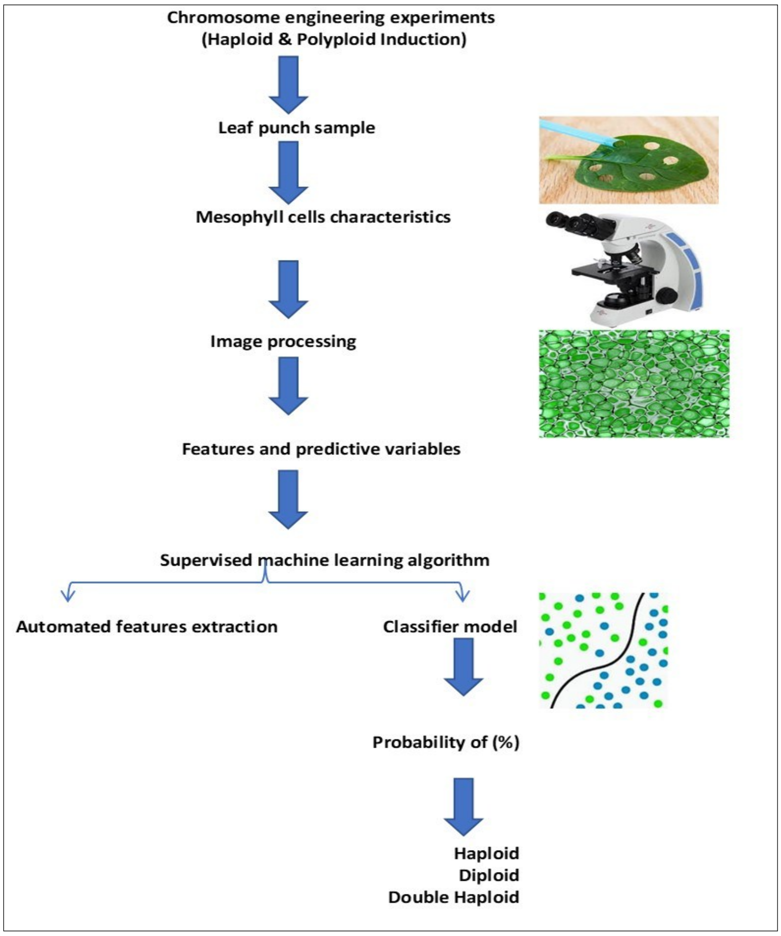

Fig. 8. Coupled image processing-supervised machine learning.

Machine learning at biochemical level

In the context of biochemical science, omics technology describes the data sources such as genomics, phenomics, metabolomics, epigenomics, methylation, proteomics and transcriptomics. These are often supplemented by microscopy techniques to examine biomolecule abundance and its localization. Recent advanced technologies, which are predominantly centred around DNA sequencing in high-throughput approach, have considerably improved the scale at which molecules and interactions can be examined. Large data sets limited mechanistic expertise and the necessity for data integration to decipher the use of machine learning to interpret outcomes. Also, there are recent advances, particularly scenarios in which machine learning allows us to examine the significance of active constituents in the formation of plant phenotypes.

One area where machine learning has proven effective is in discriminating between different types of genomic regions. In maize, for instance, DNA sequence areas have been classified as active genes vs. (inactive) pseudogenes based on DNA methylation (68). ML has also been used to predict genomic regions where crossovers or exchange of genetic material from the paternal and maternal genomes are more inclined to happen (69). Machine learning has also begun to be used in population genetics, but largely in humans thus far.

Single-cell RNA sequencing is an interesting new field in which ML plays a critical role [70, 71], enabling researchers to study the formation and response of diverse tissues at the cellular level. With thousands of cells and tens of thousands of genes worth of information, the resulting data sets are huge and complicated to analyze by unsupervised machine learning. This implies that, unlike most of the previous investigations, there is no unique "label" that can be anticipated. Instead, patterns are discovered, which aid in the data organization and interpretation. Clustering and manifold learning approaches try to uncover data structure in the same way as PCA.

For instance, transcription factors were revealed to be strong candidates for causative genes in the gene-for-trait hierarchical study (72). Another piece of evidence is the discovery that many DNA structure traits, some of which were indicative of crossover incidence across all plant species and others specific to individual species, were predictive of crossover occurrence (69). This form of model interpretation is increasingly being deployed to provide verifiable estimates, such as targeting key genomic areas, potential genes or protein residues for further examination in the lab (Fig. 9).

Machine learning at macroscopic phenotypes

Sensors capable of promptly recognizing the key parameters, such as weight, temperature, water intake, light, humidity and gaseous proportions, are available for monitoring environmental factors and some plant features (73). The processing of data required for these sensors is minimal, consisting mostly of background subtraction. However, because these sensors only measure at a region of space, they are unable to determine morphological and geometrical parameters, which are critical for plant phenotyping. The remainder of machine learning at the macroscopic level also focuses on image-based plant phenotyping because imaging sensors are commonly employed to obtain such data.

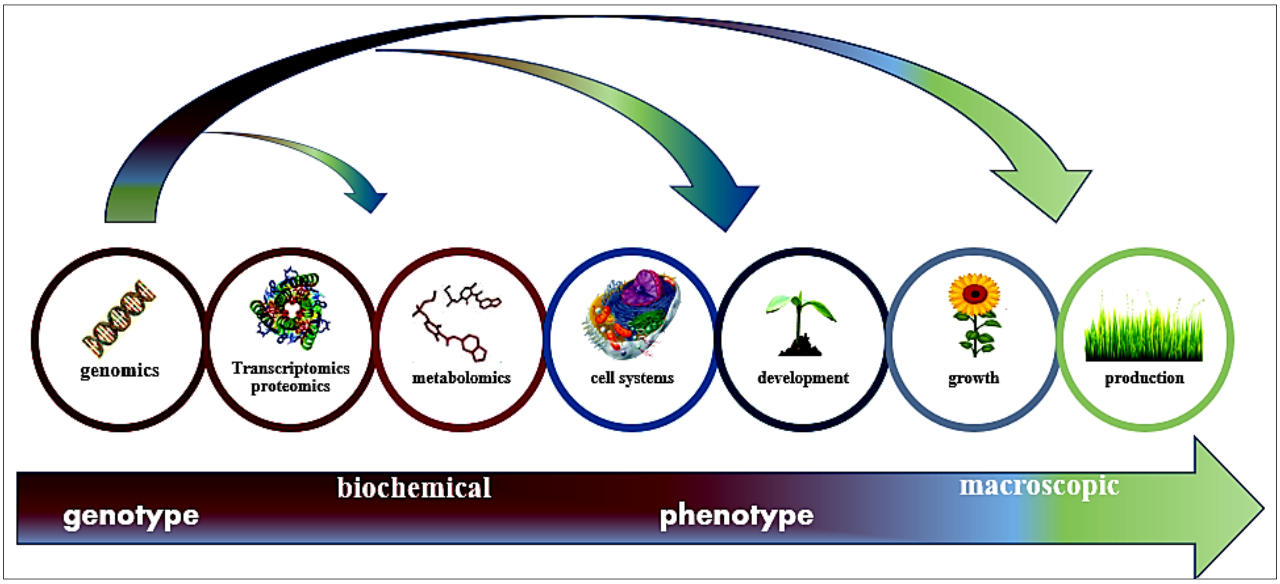

Fig. 9. Biochemical levels of organization depicting the association of phenotypic variation with variation in genotype spectrum at the macroscopic level at different scales (Source: 67).

Microscopy allows for a detailed study of plant organ development (74), while an entire plant growth can be phenotyped in growth chambers (75) or in greenhouses and open fields using robotic devices (76), with drones further assisting the monitoring of plots or complete fields (77). Colour, multispectral and hyperspectral imaging, long-wave infrared (thermal), chlorophyll fluorescence and tomographic techniques including - magnetic resonance imaging (MRI), positron emission tomography [PET], computerized tomography [CT]) are all examples of imaging sensors that measure different parts of the electromagnetic spectrum (78).

DL has also been used successfully in plant phenotyping. To explore plant growth, CNN has been employed to recognize and categorize spikes & spikelets in photos of wheat (79). CNN was applied to segment distinct plant components in both 2D and 3D photos of tomato seedlings (2). To address the challenges arising from limited labelled data, synthetic data have been effectively utilized to train the circuits for leaf segmentation despite the lack of labelled data (80). Auto-encoders including recurrent neural networks (RNNs) have been employed to analyze plant development (81). RNNs capture temporal trends via the use of feedback connections Patial Least Squares (PLS) announced superior performance than the Support Vector Machines (SVM) classifier in identifying rice varieties based on seed images.

Machine learning in genomic prediction

One of the most important goals in plant research and breeding is to explain complex features like yield as a function of genotypic, phenotypic and environmental data. The steps involved are

a. Identify quantitative trait loci (QTLs) or genetic areas linked to a specific trait through QTL mapping in experimental populations or genome-wide association studies in diverse genetic panels.

b. Determine the genetic framework by calculating the impact of individual loci and the percentage of variance explained by all loci combined for the related traits. Analysing genetic correlations between characteristics are also valuable since they quantify the degree of overlap between genetic signals.

c. Forecast anticipated trait values for new genotypes with only marker data, a process known as genomic prediction (GP).

High throughput stress phenotyping (HTSP)

With increasing advancements in the light of machine learning, it is obvious that ML-enabled HTSP will assist scientists from multiple disciplines in addition to extension workers and farmers by providing accurate, precise and rapid screening of various stress factors. This will also enable the introduction of novel selection techniques for complex quantitative features like biotic and abiotic stressors, as well as the streamlining of the gene investigative process. The use of ML-enabled HTSP will also advance our understanding of pathogen–plant relationships (83) as well as plant-stress responses.

UAVs are potential platforms for plant phenotyping because of their capability to acquire high spatial and temporal resolution data (84, 85). Traditional phenotyping is based on manual data collection, which significantly limits the available data, thereby reducing the efficiency of linking phenotypic information with genomic data. High-throughput image-based phenotyping approaches for field crops provide a non-destructive, non-invasive alternative to accelerate genotype-phenotype association studies, leading to reduced breeding cycles for the acquisition of desired traits (84). Advances in automated and high-throughput imaging technologies have resulted in the generation of high-resolution images and sensor data of plants (86, 87). However, extracting patterns and features from this large dataset requires the use of machine learning (ML) tools to enable data assimilation and feature identification for stress phenotyping (84). The applications of machine learning enabled HTSP are presented in Table 2.

The four decision cycle phases where various ML approaches can be applied in plant stress phenotyping and breeding operations are i) Identification; ii) Classification; iii) Quantification and iv) Prediction.

Identification

The detection of a specific stress among various potential stressors in the field is the basis for identification procedures. SVM, neural networks (NNs), kernel techniques and instance-

Table 2. Machine learning enabled high throughput stress phenotyping in crop plants

ML Approach	ML Algorithm	Sensor	Stress type	Trait phenotyped	Plant species
Identification	SVM variant	Scanned images	Nutrient deficiency	NPK stress	Rice
Identification	Gaussian mixture model	RGB images	Disease	Wheat streak mosaic virus	Wheat
Identification	SVM	Hyperspectral	Drought	Water stress	Tomato
Identification & Classification	Linear Discriminant analysis and k- means	RGB images	Pollution	Ozone	Clover
Classification	SVM and Bayesian classifier	RGB images	Disease	Salmonella bacteria	Arabidopsis
Quantification	SVM	RGB images and spectral reflectance	Insect	Leaf miner	Tomato
Quantification	Pre-processing *via* segmentation	RGB images	Disease	Yellow vein virus	Chilli Pepper
Prediction	SVM, GRNN	Manual severity rating	Disease	Rice blast	Rice

based approaches have all been utilized in the past to identify various stressors. The SVM approach has been successfully utilized to identify stress in plants across a range of settings.

Classification

Instead of recognizing single stress among multiple potential stressors, a classifier is utilized to categorize stress into labelled classes based on perceived stress and signatures. A standard workflow in classification algorithms involves a pre-processing phase that includes segmentation followed by the extraction of features that provides input into a classifier. Drought stress, for example, could be divided into three categories: no stress, moderate stress and severe stress.

Quantification

Quantification approaches are a type of categorization that quantifies each class based on the degree of stress. Disease severity can be used to characterize various plant illnesses in the case of plant diseases (88). Severity of rust in wheat, for example, may be measured on a scale of 0–100 %. A pre-processing stage to segregate the forefront from the backdrop, edge detection and contrast enhancement is commonly recommended before using quantification techniques.

Prediction

The ability to detect plant stress before it becomes visible to the naked eye has significant implications for stress management that are both timely and cost-effective. There have been a few reports of utilizing machine learning to anticipate stress which widened and explored the scope of machine learning approaches. This holds significant implications for precision agriculture and prescription farming.

Next-Generation Artificial Intelligence

Advancements in genomics and field phenomics investigate new methodologies and challenges for integrating multi-omics big data using next-generation (Next-Gen) AI, providing a viable route for plant breeding.

As stated by the 'no free lunch' theorem which demonstrates that no single optimizer is universally effective for all problems; several models or algorithms must be utilized to discover the optimum suit for each specific circumstance (89, 90). Similarly, the speed and reliability of AI algorithms vary greatly, especially when dealing with large amounts of data. Summit, the world's most powerful supercomputer, equipped with world's largest collection of graphics processing units (GPUs) (27,000), has recently been introduced, paving the path for the convergence of AI and scientific breakthroughs. As a result, Next-Gen AI is expected to use many models and scales to compile an ensemble of optimum trainees for each subtask while harnessing ever-increasing processing capacity (7).

In breeding programmes, Next-Gen AI is designed not only to forecast breeding value for distinctive properties across habitats and timeframes, but also to learn and improve repeatedly. This requires sophisticated and efficient data mining that accurately underpins the study of biological systems and environment, along with human involvement at both the input and output ends.

Furthermore, making multi-omics data findable, accessible, identifiable, and reusable (FAIR) encourages data reuse and discovery through proper data management (91, 92). Using international standards, such as the breeding API (BrAPI), facilitates data integration and interoperability across several datasets (93). As shown in Fig. 10, analyzing multi-omics information and producing intelligent features in plant breeding innovation was possible using the 'input–system–output' big data analytics approach. Data capture, analysis and interpretation are the three main determinants of the process, leading to insights into the underlying biology and improved breeding decisions. High-resolution measurements of genome, environment and phenome sequencing, sensing and other high-throughput technologies are continually expanding the amount of genotypic, environmental and phenotypic data available for large breeding populations in multi-environment experimental studies (94). The combination of this data with next-generation AI approach can help researchers in identifying better alleles for critical genes and breed superior genotypes and ideotypes to suit climate-resilient agriculture.

Challenges, ethical concerns and future prospects

Despite the benefits of AI, its' practical applicability remains limited across locations. AI use in agriculture is constrained by the lack of easy-to-implement technologies. Farmers often lack the necessary time and digital skills to study AI possibilities unaided. To effectively integrate AI into agriculture, new solutions must relate to current infrastructure and processes (95). Agriculture cannot fully rely on AI as it cannot function outside of its programming. Farmers in rural regions sometimes lack technical skills and awareness of available technology. Nevertheless, the distribution of mechanization may be unequal, with certain regions having limited access to

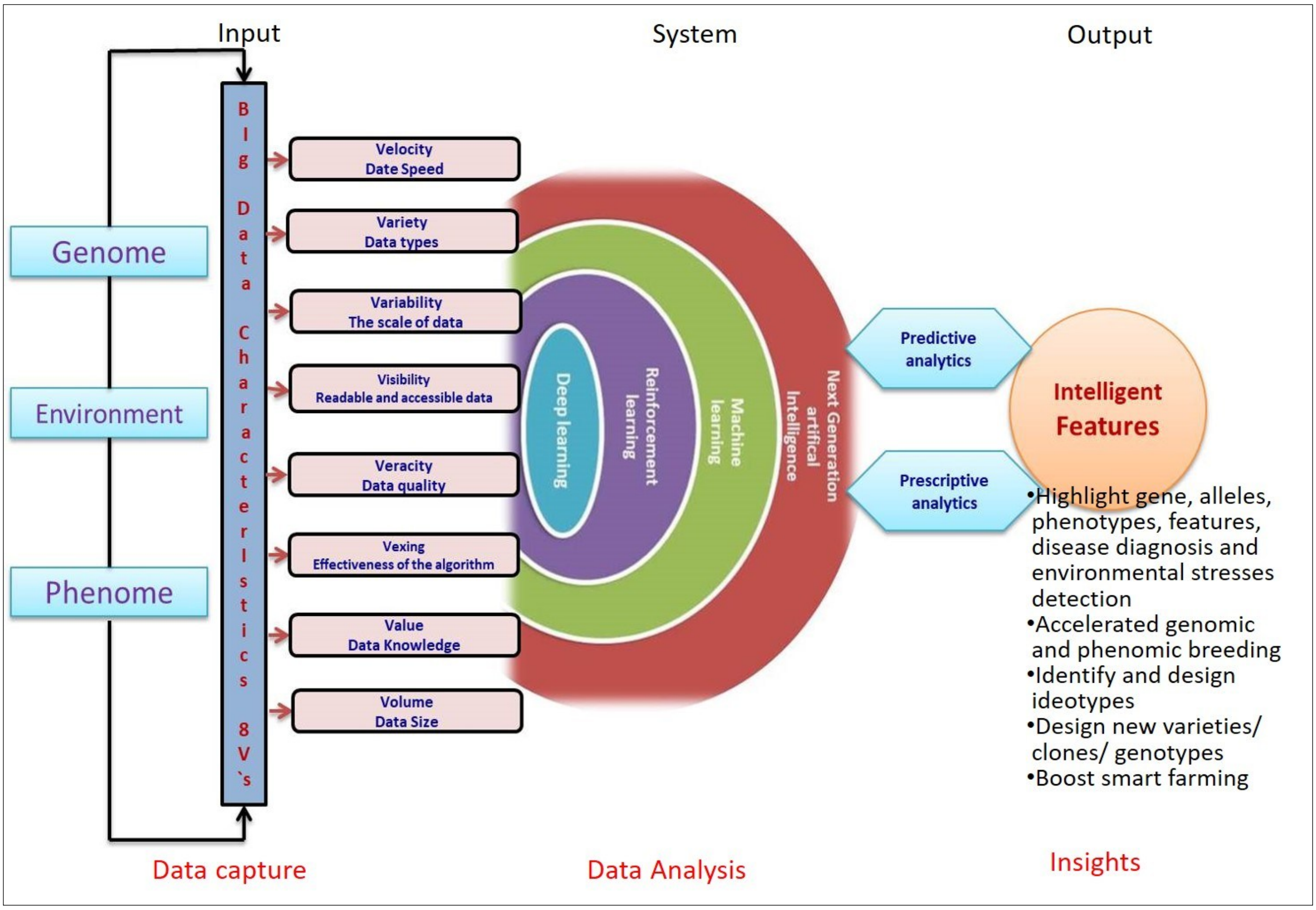

Fig. 10. Combination of next generation artificial intelligence with big data in plant breeding.

resources and circumstances that are unchangeable regardless advancements in science and technology. Since most AI systems rely on the internet, for instance, their use may be limited in rural or isolated locations without online services and expertise in managing AI operations. For AI systems to teach machines and produce precise predictions, a lot of data is needed. Spatial data is simple, but obtaining temporal data is more difficult, particularly crop- specific data, which can only be acquired once a year. The development of a strong machine learning model takes a long time as data infrastructure requires time to evolve. In addition to the above, data privacy is the major concern where cloud servers have been subjected to tampering in the most recent times. Given the existing laws and regulations, the problems of data inaccuracy relating to biased decisions or incorrect recommendations brought about by flawed data analysis could endanger farmers' own crops. These matters pertain to both civil and criminal law (96). One of the most contentious issues in AI is data ownership where farmers biggest fear is that their data my fall into wrong hands and be used against him. Hence, there is a necessity to protect the farmers' data by ensuring intellectual property rights through AI based technology providers and national legislations (97). "Digital division" is another economic issue in adopting AI in agriculture contributing to dissolution of small farms and the disproportionate growth of large ones. This concern may lead to farm inequalities which could be addressed through adopting policies that provide equitable access to AI based technologies for small-scale farmers (98).

AI technologies are expected to provide creative and accurate answers to the main agricultural problems that farmers face globally in the future. AI could assist significantly in areas including weather forecasting, pest management and agricultural labor support through further technological advancements. Besides, AI also may help farmers transform into agricultural pioneers in the future by utilizing data to maximize plant yields for specific rows. Robots could be designed to efficiently perform a variety of tasks in agricultural environments, harvest crops faster and more thoroughly than humans. These robots can be programmed to check for weeds and crop quality during harvesting and packing crops. These robots may also be capable of overcoming the challenges that agricultural labor faces.

Conclusion

Before the rise of technological advancements, plant breeding faced several limitations *viz.,* manual data collection, low-throughput phenotyping and real time monitoring. This hindered the precision breeding necessitating the need of AI driven technologies. AI technologies rely heavily on vast amounts of data, presenting opportunities to enhance customization, data mining and execution accuracy in agricultural applications. The integration of next-generation AI, including advanced deep learning, generative AI and self-learning algorithms**,** can drive scientific breakthroughs by refining predictive models and optimizing real-time decision-making. The application of deep learning in phenotyping data analysis is particularly promising, as it has already demonstrated success in complex material analysis and pattern recognition. In plant research, particularly in plant breeding, machine learning (ML) plays a pivotal role in deciphering large-scale phenotyping datasets and establishing precise genotype-phenotype correlations at biochemical and macroscopic levels. With high-throughput stress phenotyping, becoming increasingly sophisticated, the seamless integration of AI-powered analytics into data collection, processing and interpretation workflows is essential. Next-generation AI models, capable of self-improvement and adaptive learning, may further enhance efficiency, accuracy and automation in agricultural research. Moreover, AI-driven innovations in predictive analytics, climate adaptation strategies and precision breeding hold immense potential for addressing the pressing challenge of global food security. As environmental conditions shift rapidly, leveraging AI to develop resilient crops, optimize resource use and mitigate risks within agricultural ecosystems will be crucial. By integrating AI at every stage of the agricultural pipeline, from genetic research to field applications, we can build a more sustainable, adaptive and high-yield food production system to support the growing global population.

Acknowledgements

We acknowledge the Agricultural College and Research Institute, Tamil Nadu Agricultural University, Madurai for providing technical support in the manuscript preparation.

Authors' contributions

SG and SJH drafted the manuscript. SG, VN and MJ collected the literature and MT guided in preparation of the article. All authors reviewed the contents and approved the final version of the manuscript.

Compliance with ethical standards

Conflict of interest: The authors declare that there is no conflict to interest related to this paper

Ethical issues: None

References

1. Elbasi E, Mostafa N, AlArnaout Z, Zreikat AI, Cina E, Varghese G, Shdefat A, Topcu AE, et al. Artificial intelligence technology in the agricultural sector: a systematic literature review. IEEE Access. 2022;11:171–202. https://doi.org/10.1109/ACCESS.2022.3232485
2. Lybbert TJ, Sumner DA. Agricultural technologies for climate change in developing countries: policy options for innovation and technology diffusion. Food Policy. 2012;37(1):114–23. https://doi.org/10.1016/j.foodpol.2011.11.001
3. Shi X, An X, Zhao Q, Liu H, Xia L, Sun X, Guo Y. State-of-the-art Internet of Things in protected agriculture. Sensors (Basel). 2019;19(8):1833. https://doi.org/10.3390/s19081833
4. Walters JP, Archer DW, Sassenrath GF, Hendrickson JR, Hanson JD, Halloran JM, Vadas P, et al. Exploring agricultural production systems and their fundamental components with system dynamics modelling. Ecol Model. 2016;333:51–65. https://doi.org/10.1016/j.ecolmodel.2016.04.015
5. Shah G, Shah A, Shah M. Panacea of challenges in real-world application of big data analytics in healthcare sector. J Data Inf Manag. 2019;1:107–16. https://doi.org/10.1007/s42488-019-00010-1
6. Sukhadia A, Upadhyay K, Gundeti M, Shah S, Shah M. Optimization of smart traffic governance system using artificial intelligence. Aug Human Res. 2020;5(1):13. https://doi.org/10.1007/s41133-020-00035-x
7. Wallace JG, Rodgers-Melnick E, Buckler ES. On the road to breeding 4.0: unraveling the good, the bad, and the boring of crop quantitative genomics. Annu Rev Genet. 2018;52(1):421–44. https://doi.org/10.1146/annurev-genet-120116-024846
8. Harfouche AL, Jacobson DA, Kainer D, Romero JC, Harfouche AH, Mugnozza GS, Moshelion M, et al. Accelerating climate-resilient plant breeding by applying next-generation artificial intelligence. Trends Biotechnol. 2019;37(11):1217–35. https://doi.org/10.1016/j.tibtech.2019.05.007
9. Zhao X, Fan W, Chen B, Tang R. Automated machine learning for recommendations: fundamentals and advances. In: Proceedings of the 31st ACM Web Conference (WWW '22). New York: ACM; 2022. p. 398.
10. Kakkad V, Patel M, Shah M. Biometric authentication and image encryption for image security in cloud framework. Multiscale Multidiscip Model Exp Des. 2019;2(4):233–48. https://doi.org/10.1007/s41939-019-00049-y
11. Patel D, Shah Y, Thakkar N, Shah K, Shah M. Implementation of artificial intelligence techniques for cancer detection. Aug Human Res. 2020;5:1–10. https://doi.org/10.1007/s41133-019-0024-3
12. Patel H, Prajapati D, Mahida D, Shah M. Transforming petroleum downstream sector through big data: a holistic review. J Pet Explor Prod Technol. 2020;10:2601–11. https://doi.org/10.1007/s13202-020-00889-2
13. Panchiwala S, Shah M. A comprehensive study on critical security issues and challenges of the IoT world. J Data Inf Manag. 2020;2:257–78. https://doi.org/10.1007/s42488-020-00030-2
14. Gavhale KR, Gawande U. An overview of the research on plant leaves disease detection using image processing techniques. IOSR J Comput Eng. 2014;16(1):10–6.
15. Jani K, Chaudhuri M, Patel H, Shah M. Machine learning in films: an approach towards automation in film censoring. J Data Inf Manag. 2020;2:55–64. https://doi.org/10.1007/s42488-019-00016-9
16. Zakaria M, Mabrouka AS, Sarhan S. Artificial neural network: a brief overview. Int J Comput Sci Eng. 2014;4(2):7–12.
17. Xiao D, Li B, Mao Y. A multiple hidden layers extreme learning machine method and its application. Math Probl Eng. 2017;2017(1):4670187. https://doi.org/10.1155/2017/4670187
18. Hesamifard E, Takabi H, Ghasemi M. Cryptodl: deep neural networks over encrypted data. arXiv. 2017;1711.05189. https://doi.org/10.48550/arXiv.1711.05189
19. Sladojevic S, Arsenovic M, Anderla A, Culibrk D, Stefanovic D. Deep neural networks based recognition of plant diseases by leaf image classification. Comput Intell Neurosci. 2016;2016(1):3289801. https://doi.org/10.1155/2016/3289801
20. Pandya R, Nadiadwala S, Shah R, Shah M. Buildout of methodology for meticulous diagnosis of K-complex in EEG for aiding the detection of Alzheimer's by artificial intelligence. Aug Human Res. 2020;5:1–8. https://doi.org/10.1007/s41133-019-0021-6
21. Abiodun OI, Jantan A, Omolara AE, Dada KV, Mohamed NA, Arshad H. State-of-the-art in artificial neural network applications: a survey. Heliyon. 2018;4(11):e00938. https://doi.org/10.1016/j.heliyon.2018.e00938
22. Nabwire S, Suh HK, Kim MS, Baek I, Cho BK. Application of artificial intelligence in phenomics. Sensors (Basel). 2021;21(13):4363. https://doi.org/10.3390/s21134363
23. William JR. Some definitions of artificial intelligence. 2022. Available from: https://cse.buffalo.edu/~Rapaport/definitions.of.ai.html
24. van Dijk AD, Kootstra G, Kruijer W, de Ridder D. Machine learning in plant science and plant breeding. iScience. 2021;24(1):101890. https://doi.org/10.1016/j.isci.2020.101890

25. Patel UK, Anwar A, Saleem S, Malik P, Rasul B, Patel K, Yao R, et al. Artificial intelligence as an emerging technology in the current care of neurological disorders. J Neurol. 2021;268:1623–42. https://doi.org/10.1007/s00415-019-09518-3
26. Yanting SH, Clarke R, Tingting ZH, Clifton D; inventors; Oxford University Innovation Ltd, assignee. Generating neural network models, classifying physiological data, and classifying patients into clinical classifications. US Patent Application US 18/280,751. 2024 Sep 12.
27. Frey LJ. Artificial intelligence and integrated genotype-phenotype identification. Genes (Basel). 2018;10(1):18. https://doi.org/10.3390/genes10010018
28. Zheng H, Li W, Jiang J, Liu Y, Cheng T, Tian Y, Zhu Y, Cao W, Zhang Y, Yao X. A comparative assessment of different modeling algorithms for estimating leaf nitrogen content in winter wheat using multispectral images from an unmanned aerial vehicle. Remote Sens (Basel). 2018;10(12):2026. https://doi.org/10.3390/rs10122026
29. Salehi M, Farhadi S, Moieni A, Safaie N, Ahmadi H. Mathematical modeling of growth and paclitaxel biosynthesis in *Corylus avellana* cell culture responding to fungal elicitors using multilayer perceptron-genetic algorithm. Front Plant Sci. 2020;11:1148. https://doi.org/10.3389/fpls.2020.01148
30. Asefpour Vakilian K. Machine learning improves our knowledge about miRNA functions towards plant abiotic stresses. Sci Rep. 2020;10(1):3041. https://doi.org/10.1038/s41598-020-59981-6
31. Wang MW, Goodman JM, Allen TE. Machine learning in predictive toxicology: recent applications and future directions for classification models. Chem Res Toxicol. 2020;34(2):217–39. https://doi.org/10.1021/acs.chemrestox.0c00316
32. Hu H, Scheben A, Edwards D. Advances in integrating genomics and bioinformatics in the plant breeding pipeline. Agriculture (Basel). 2018;8(6):75. https://doi.org/10.3390/agriculture8060075
33. Orozco-Arias S, Isaza G, Guyot R. Retrotransposons in plant genomes: structure, identification, and classification through bioinformatics and machine learning. Int J Mol Sci. 2019;20(15):3837. https://doi.org/10.3390/ijms20153837
34. Azevedo AM, Andrade VC, Pedrosa CE, Oliveira CM, Dornas MF, Cruz CD, Valadares NR. Application of artificial neural networks in indirect selection: a case study on the breeding of lettuce. Bragantia. 2015;74(4):387–93. https://doi.org/10.1590/1678-4499.0088
35. Gu J, Wang Z, Kuen J, Ma L, Shahroudy A, Shuai B, et al. Recent advances in convolutional neural networks. Pattern Recogn. 2018;77:354–77. https://doi.org/10.1016/j.patcog.2017.10.013
36. Ali AM, Darvishzadeh R, Skidmore A, Gara TW, Heurich M. Machine learning methods' performance in radiative transfer model inversion to retrieve plant traits from Sentinel-2 data of a mixed mountain forest. Int J Digit Earth. 2021;14(1):106–20. https://doi.org/10.1080/17538947.2020.1794064
37. Gold KM, Townsend PA, Herrmann I, Gevens AJ. Investigating potato late blight physiological differences across potato cultivars with spectroscopy and machine learning. Plant Sci. 2020;295:110316. https://doi.org/10.1016/j.plantsci.2019.110316
38. Barth R, IJsselmuiden J, Hemming J, Van Henten EJ. Synthetic bootstrapping of convolutional neural networks for semantic plant part segmentation. Comput Electron Agric. 2019;161:291–304. https://doi.org/10.1016/j.compag.2017.11.040
39. An J, Li W, Li M, Cui S, Yue H. Identification and classification of maize drought stress using deep convolutional neural network. Symmetry (Basel). 2019;11(2):256. https://doi.org/10.3390/sym11020256
40. Kattenborn T, Leitloff J, Schiefer F, Hinz S. Review on convolutional neural networks (CNN) in vegetation remote sensing. ISPRS J Photogramm Remote Sens. 2021;173:24–49. https://doi.org/10.1016/j.isprsjprs.2020.12.010
41. Srivastava A, Gupta S, Shanker K, Gupta N, Gupta AK, Lal RK. Genetic diversity in Indian poppy (*Papaver somniferum* L.) germplasm using multivariate and SCoT marker analyses. Ind Crops Prod. 2020;144:112050. https://doi.org/10.1016/j.indcrop.2019.112050
42. Boonsrangsom T. Genetic diversity of 'Wan Chak Motluk' (*Curcuma comosa* Roxb.) in Thailand using morphological characteristics and RAPD markers. S Afr J Bot. 2020;130:224–30. https://doi.org/10.1016/j.sajb.2020.01.005
43. Raza A, Saher MS, Farwa A, Ahmad KR. Genetic diversity analysis of *Brassica* species using PCR-based SSR markers. Gesunde Pflanzen. 2019;71(1):1–7. https://doi.org/10.1007/s10343-018-0435-y
44. Xu S, Shen J, Wei Y, Li Y, He Y, Hu H, Feng X. Automatic plant phenotyping analysis of Melon (*Cucumis melo* L.) germplasm resources using deep learning methods and computer vision. Plant Methods. 2024;20:166. https://doi.org/10.1186/s13007-024-01293-1
45. Pandolfi C, Mugnai S, Azzarello E, Bergamasco S, Masi E, Mancuso S. Artificial neural networks as a tool for plant identification: a case study on Vietnamese tea accessions. Euphytica. 2009;166:411–21. https://doi.org/10.1007/s10681-008-9828-9
46. Saini G, Khamparia A, Luhach AK. Classification of plants using convolutional neural network. In: Proceedings of the First International Conference on Sustainable Technologies for Computational Intelligence: ICTSCI 2019. Singapore: Springer; 2020. p 551–61. https://doi.org/10.1007/978-981-15-0029-9_44
47. Sakoda K, Watanabe T, Sukemura S, Kobayashi S, Nagasaki Y, Tanaka Y, Shiraiwa T. Genetic diversity in stomatal density among soybeans elucidated using high-throughput technique based on an algorithm for object detection. Sci Rep. 2019;9(1):7610. https://doi.org/10.1038/s41598-019-44127-0
48. Yang HW, Hsu HC, Yang CK, Tsai MJ, Kuo YF. Differentiating between morphologically similar species in genus *Cinnamomum* (Lauraceae) using deep convolutional neural networks. Comput Electron Agric. 2019;162:739–48. https://doi.org/10.1016/j.compag.2019.05.003
49. Sant'Anna IC, Tomaz RS, Silva GN, Nascimento M, Bhering LL, Cruz CD. Superiority of artificial neural networks for a genetic classification procedure. Genet Mol Res. 2015;14(4):9898–9906. https://doi.org/10.4238/2015.august.19.24
50. Niedbała G, Wróbel B, Piekutowska M, Zielewicz W, Paszkiewicz-Jasińska A, Wojciechowski T, Niazian M. Application of artificial neural networks sensitivity analysis for the pre-identification of highly significant factors influencing the yield and digestibility of grassland sward in the climatic conditions of central Poland. Agronomy. 2022;12(5):1133. https://doi.org/10.3390/agronomy12051133
51. Li N, Lin H, Wang T, Li Y, Liu Y, Chen X, et al. Impact of climate change on cotton growth and yields in Xinjiang, China. Field Crops Res. 2020;247:107590. https://doi.org/10.1016/j.fcr.2019.107590
52. Niazian M, Sadat-Noori SA, Abdipour M, Tohidfar M, Mortazavian SM. Image processing and artificial neural network-based models to measure and predict physical properties of embryogenic callus and number of somatic embryos in ajowan (*Trachyspermum ammi* (L.) Sprague). *In Vitro* Cell Dev Biol-Plant. 2018;54:54–68. https://doi.org/10.1007/s11627-017-9877-7
53. Singh A, Ganapathysubramanian B, Singh AK, Sarkar S. Machine learning for high-throughput stress phenotyping in plants. Trends Plant Sci. 2016;21(2):110–24. https://doi.org/10.1016/j.tplants.2015.10.015
54. Picon A, Alvarez-Gila A, Seitz M, Ortiz-Barredo A, Echazarra J, Johannes A. Deep convolutional neural networks for mobile capture device-based crop disease classification in the wild. Comput Electron Agric. 2019;161:280–90. https://doi.org/10.1016/j.compag.2018.04.002
55. Pacal I, Kunduracioglu I, Alma MH, Deveci M, Kadry S, Nedoma J,

et al. A systematic review of deep learning techniques for plant diseases. Artif Intell Rev. 2024;57(11):304. https://doi.org/10.1007/s10462-024-10944-7

56. Sibiya M, Sumbwanyambe M. A computational procedure for the recognition and classification of maize leaf diseases out of healthy leaves using convolutional neural networks. Agriengineering. 2019;1(1):119–31. https://doi.org/10.3390/agriengineering1010009

57. Selvaraj MG, Vergara A, Ruiz H, Safari N, Elayabalan S, Ocimati W, et al. AI-powered banana diseases and pest detection. Plant Methods. 2019;15:1–1. https://doi.org/10.1186/s13007-019-0475-z

58. Coulibaly S, Kamsu-Foguem B, Kamissoko D, Traore D. Deep neural networks with transfer learning in millet crop images. Comput Ind. 2019;108:115–20. https://doi.org/10.1016/j.compind.2019.02.003

59. Niazian M. Application of genetics and biotechnology for improving medicinal plants. Planta. 2019;249:953–73. https://doi.org/10.1007/s00425-019-03099-1

60. Altuntaş Y, Cömert Z, Kocamaz AF. Identification of haploid and diploid maize seeds using convolutional neural networks and a transfer learning approach. Comput Electron Agric. 2019;163:104874. https://doi.org/10.1016/j.compag.2019.104874

61. Blonder B, Graae BJ, Greer B, Haagsma M, Helsen K, Kapás RE, et al. Remote sensing of ploidy level in quaking aspen (*Populus tremuloides* Michx.). J Ecol. 2020;108(1):175–88. https://doi.org/10.1111/1365-2745.13296

62. Santeramo D, Howell J, Ji Y, Yu W, Liu W, Kelliher T. DNA content equivalence in haploid and diploid maize leaves. Planta. 2020;251(1):30. https://doi.org/10.1007/s00425-019-03320-1

63. Crossa J, Pérez-Rodríguez P, Cuevas J, Montesinos-López O, Jarquín D, De Los Campos G, et al. Genomic selection in plant breeding: methods, models, and perspectives. Trends Plant Sci. 2017;22(11):961–75. https://doi.org/10.1016/j.tplants.2017.08.011

64. Libbrecht MW, Noble WS. Machine learning applications in genetics and genomics. Nat Rev Genet. 2015;16(6):321–32. https://doi.org/10.1038/nrg3920

65. Perez-Sanz F, Navarro PJ, Egea-Cortines M. Plant phenomics: An overview of image acquisition technologies and image data analysis algorithms. GigaScience. 2017;6(11):gix092. https://doi.org/10.1093/gigascience/gix092

66. Ma C, Zhang HH, Wang X. Machine learning for big data analytics in plants. Trends Plant Sci. 2014;19(12):798–808. https://doi.org/10.1016/j.tplants.2014.08.004

67. van Dijk AD, Kootstra G, Kruijer W, de Ridder D. Machine learning in plant science and plant breeding. iScience. 2021;24(1):101890. https://doi.org/10.1016/j.isci.2020.101890

68. Sartor RC, Noshay J, Springer NM, Briggs SP. Identification of the expressome by machine learning on omics data. Proc Natl Acad Sci U S A. 2019;116(36):18119–25. https://doi.org/10.1073/pnas.1813645116

69. Demirci S, Peters SA, de Ridder D, van Dijk AD. DNA sequence and shape are predictive for meiotic crossovers throughout the plant kingdom. Plant J. 2018;95(4):686–99. https://doi.org/10.1111/tpj.13979

70. Denyer T, Ma X, Klesen S, Scacchi E, Nieselt K, Timmermans MC. Spatiotemporal developmental trajectories in the *Arabidopsis* root revealed using high-throughput single-cell RNA sequencing. Dev Cell. 2019;48(6):840–52. https://doi.org/10.1016/j.devcel.2019.02.022

71. Jean-Baptiste K, McFaline-Figueroa JL, Alexandre CM, Dorrity MW, Saunders L, Bubb KL, et al. Dynamics of gene expression in single root cells of *Arabidopsis thaliana*. Plant Cell. 2019;31(5):993–1011. https://doi.org/10.1105/tpc.18.00785

72. Lin F, Fan J, Rhee SY. QTG-Finder: a machine-learning based algorithm to prioritize causal genes of quantitative trait loci in *Arabidopsis* and rice. G3 (Bethesda). 2019;9(10):3129–38. https://doi.org/10.1534/g3.119.400319

73. Fahlgren N, Gehan MA, Baxter I. Lights, camera, action: high-throughput plant phenotyping is ready for a close-up. Curr Opin Plant Biol. 2015;24:93–9. https://doi.org/10.1016/j.pbi.2015.02.006

74. Dhondt S, Wuyts N, Inzé D. Cell to whole-plant phenotyping: the best is yet to come. Trends Plant Sci. 2013;18(8):428–39. https://doi.org/10.1016/j.tplants.2013.04.008

75. van Es SW, van der Auweraert EB, Silveira SR, Angenent GC, van Dijk AD, Immink RG. Comprehensive phenotyping reveals interactions and functions of *Arabidopsis thaliana* TCP genes in yield determination. Plant J. 2019;99(2):316–28. https://doi.org/10.1111/tpj.14326

76. Virlet N, Sabermanesh K, Sadeghi-Tehran P, Hawkesford MJ. Field Scanalyzer: An automated robotic field phenotyping platform for detailed crop monitoring. Funct Plant Biol. 2016;44(1):143–53. https://doi.org/10.1071/FP16163

77. Araus JL, Kefauver SC, Zaman-Allah M, Olsen MS, Cairns JE. Translating high-throughput phenotyping into genetic gain. Trends Plant Sci. 2018;23(5):451–66. https://doi.org/10.1016/j.tplants.2018.02.001

78. Li L, Zhang Q, Huang D. A review of imaging techniques for plant phenotyping. Sensors (Basel). 2014;14(11):20078–111. https://doi.org/10.3390/s141120078

79. Pound MP, Atkinson JA, Wells DM, Pridmore TP, French AP. Deep learning for multi-task plant phenotyping. Proc IEEE Int Conf Comput Vis Workshops. 2017:2055–63. https://doi.org/10.1101/204552

80. Ward D, Moghadam P, Hudson N. Deep leaf segmentation using synthetic data. arXiv. 2018. https://doi.org/10.48550/arXiv.1807.10931

81. Taghavi Namin S, Esmaeilzadeh M, Najafi M, Brown TB, Borevitz JO. Deep phenotyping: deep learning for temporal phenotype/genotype classification. Plant Methods. 2018;14:1–4. https://doi.org/10.1186/s13007-018-0333-4

82. Chandrika G, Hepziba J, Pillai MA, Pushpam KA, Vijayalakshmi R. Vision-based rice genotype identification using machine learning techniques. Biol Forum Int J. 2023;15(8A):285–91.

83. Kuska M, Wahabzada M, Leucker M, Dehne HW, Kersting K, Oerke EC, et al. Hyperspectral phenotyping on the microscopic scale: towards automated characterization of plant-pathogen interactions. Plant Methods. 2015;11:1–5. https://doi.org/10.1186/s13007-015-0073-7

84. Zhang Z, Zhu L. A review on unmanned aerial vehicle remote sensing: Platforms, sensors, data processing methods, and applications. Drones. 2023;7(6):398. https://doi.org/10.3390/drones7060398

85. Sharma V, Honkavaara E, Hayden M, Kant S. UAV remote sensing phenotyping of wheat collection for response to water stress and yield prediction using machine learning. Plant Stress. 2024;12:100464. https://doi.org/10.1016/j.stress.2024.100464

86. Das Choudhury S, Guadagno CR, Bashyam S, Mazis A, Ewers BE, Samal A, et al. Stress phenotyping analysis leveraging autofluorescence image sequences with machine learning. Front Plant Sci. 2024;15:1353110. https://doi.org/10.3389/fpls.2024.1353110

87. Arya S, Sahoo RN, Sehgal VK, Bandyopadhyay K, Rejith RG, Chinnusamy V, et al. High-throughput chlorophyll fluorescence image-based phenotyping for water deficit stress tolerance in wheat. Plant Physiol Rep. 2024;1–6. https://doi.org/10.1007/s40502-024-00783-7

88. Bock CH, Poole GH, Parker PE, Gottwald TR. Plant disease severity estimated visually, by digital photography and image analysis, and by hyperspectral imaging. Crit Rev Plant Sci. 2010;29(2):59–

107. https://doi.org/10.1080/07352681003617285

89. Sterkenburg TF, Grünwald PD. The no-free-lunch theorems of supervised learning. Synthese. 2021;199(3):9979–10015. https://doi.org/10.1007/s11229-021-03233-1

90. Sychrovský D, Šustr M, Davoodi E, Bowling M, Lanctot M, Schmid M. Learning not to regret. Proc AAAI Conf Artif Intell. 2024;38(14):15202–10. https://doi.org/10.1609/aaai.v38i14.29443

91. Niehues A, de Visser C, Hagenbeek FA, Kulkarni P, Pool R, Karu N, et al. A multi-omics data analysis workflow packaged as a FAIR Digital Object. GigaScience. 2024;13:giad115. https://doi.org/10.1093/gigascience/giad115

92. Doniparthi G, Mühlhaus T, Deßloch S. Integrating FAIR Experimental Metadata for Multi-omics Data Analysis. Datenbank-Spektrum. 2024;1–9. https://doi.org/10.1007/s13222-024-00473-6

93. Selby P, Abbeloos R, Backlund JE, Basterrechea Salido M, Bauchet G, Benites-Alfaro OE, et al. BrAPI-an application programming interface for plant breeding applications. Bioinformatics. 2019;35(20):4147–55. https://doi.org/10.1093/bioinformatics/btz190

94. Sheikh M, Iqra F, Ambreen H, Pravin KA, Ikra M, Chung YS. Integrating artificial intelligence and high-throughput phenotyping for crop improvement. J Integr Agric. 2024;23(6):1787–802. https://doi.org/10.1016/j.jia.2023.10.019

95. Oliveira RC, Silva RD. Artificial intelligence in agriculture: benefits, challenges, and trends. Appl Sci. 2023;13(13):7405. https://doi.org/10.3390/app13137405

96. Sparrow R, Howard M, Degeling C. Managing the risks of artificial intelligence in agriculture. NJAS Impact Agr Life Sci. 2021;93(1):172–96.https://doi.org/10.1080/27685241.2021.2008777

97. Ryan M. The social and ethical impacts of artificial intelligence in agriculture: mapping the agricultural AI literature. AI Soc. 2023;38(6):2473–85. https://doi.org/10.1007/s00146-021-01377-9

98. Uddin M, Chowdhury A, Kabir MA. Legal and ethical aspects of deploying artificial intelligence in climate-smart agriculture. AI Soc. 2024;39(1):221–34. https://doi.org/10.1007/s00146-022-01421-2

Additional information

Peer review: Publisher thanks Sectional Editor and the other anonymous reviewers for their contribution to the peer review of this work.

Reprints & permissions information is available at https://horizonepublishing.com/journals/index.php/PST/open_access_policy

Publisher's Note: Horizon e-Publishing Group remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Indexing: Plant Science Today, published by Horizon e-Publishing Group, is covered by Scopus, Web of Science, BIOSIS Previews, Clarivate Analytics, NAAS, UGC Care, etc
See https://horizonepublishing.com/journals/index.php/PST/indexing_abstracting

Publisher information: Plant Science Today is published by HORIZON e-Publishing Group with support from Empirion Publishers Private Limited, Thiruvananthapuram, India.