Review

Not peer-reviewed version

Breeding of Solanaceous Crops Using AI: Machine and Deep Learning Approaches

Maria Gerakari , Anastasios Katsileros , Konstantina Kleftogianni , Eleni Tani , Penelope J. Bebeli ,
Vasileios Papasotiropoulos *

Posted Date: 17 February 2025

doi: 10.20944/preprints202502.1208.v1

Keywords: artificial intelligence; machine learning; big data; deep learning; plant breeding; Solanaceae; tomato; potato; eggplant; pepper

Review

Breeding of Solanaceous Crops Using AI: Machine and Deep Learning Approaches

Maria Gerakari †, Anastasios Katsileros †, Konstantina Kleftogianni, Eleni Tani, Penelope J. Bebeli and Vasileios Papasotiropoulos *

Laboratory of Plant Breeding & Biometry, Department of Crop Science, Agricultural University of Athens, Iera Odos 75, 11855 Athens, Greece

* Correspondence: vpapasot@aua.gr; Tel.: +30-210-5294622

† These authors contributed equally to this work.

Abstract: Artificial intelligence (AI), including machine learning (ML) and deep learning (DL), has become an essential tool in modern agriculture, revolutionizing traditional practices and offering sustainable solutions to critical challenges, such as climate change, population growth, and resource scarcity. Through advanced algorithms and predictive models, ML and DL enhance precise genomic selection (GS), trait characterization, and the acceleration of crop breeding processes. These technologies facilitate the identification and optimization of key traits, including increased yield, improved quality, pest resistance, and tolerance to extreme climatic conditions. Additionally, ML-driven tools support gene-editing technologies, such as CRISPR-Cas9, contributing to the development of resilient and adaptable crops. By leveraging big data analytics and omic technologies, they provide valuable insights into linking genetic and phenotypic data, fostering the development of sustainable agricultural practices. This research explores the transformative potential of AI, particularly ML and DL, in Solanaceous crops by developing advanced breeding strategies to address challenges posed by climate change and rapid population growth. Furthermore, this study highlights the significant role of these technologies in creating novel crop varieties that are resilient to environmental stressors, while exhibiting superior agronomic and quality traits. AI and its applications, such as ML and DL, contribute to the genetic improvement of Solanaceous crops, strengthening agricultural resilience, ensuring food security, and promoting environmental sustainability

Keywords: artificial intelligence; machine learning; big data; deep learning; plant breeding; Solanaceae; tomato; potato; eggplant; pepper

---

1. Introduction

Artificial intelligence (AI) and specifically machine learning (ML) and its subset Deep Learning (DL) have become key components of modern technological advancement, with transformative applications across diverse fields. Agriculture, has increasingly embraced these technologies to address challenges such as food security, environmental sustainability, and resource optimization. AI and ML are extremely impactful in various sectors of agriculture, such as precision farming and field management [1], enabling farmers to make informed decisions about irrigation, fertilization, pest control, and harvesting [2]. AI and ML are playing a crucial role in optimizing field monitoring, through the analysis of soil quality and health, which is fundamental for agricultural productivity [3]. Moreover, AI-driven tools have revolutionized supply chain management in agriculture, enhancing inventory tracking, demand forecasting, and logistics [4]. By leveraging AI's capability to simulate human intelligence and ML's capacity to analyze and learn from vast datasets, agricultural systems are now better equipped to enhance efficiency and resilience [5].

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 February 2025 doi:10.20944/preprints202502.1208.v1

ML algorithms process massively produced datasets allowing farmers to address specific needs of their crops, while minimizing waste and resource inefficiency. These vast datasets are also called "Big Data" and they consist of 5Vs key attributes (Volume, Velocity, Variety, Veracity, and Value). The utilization of Big Data analysis in agriculture, through AI and ML, enables the processing of large-scale, diverse, and real-time datasets to make precise predictions, manage resources efficiently, and support data-driven decision-making, ultimately optimizing crop selection and improving overall productivity [6] (Figure 1).

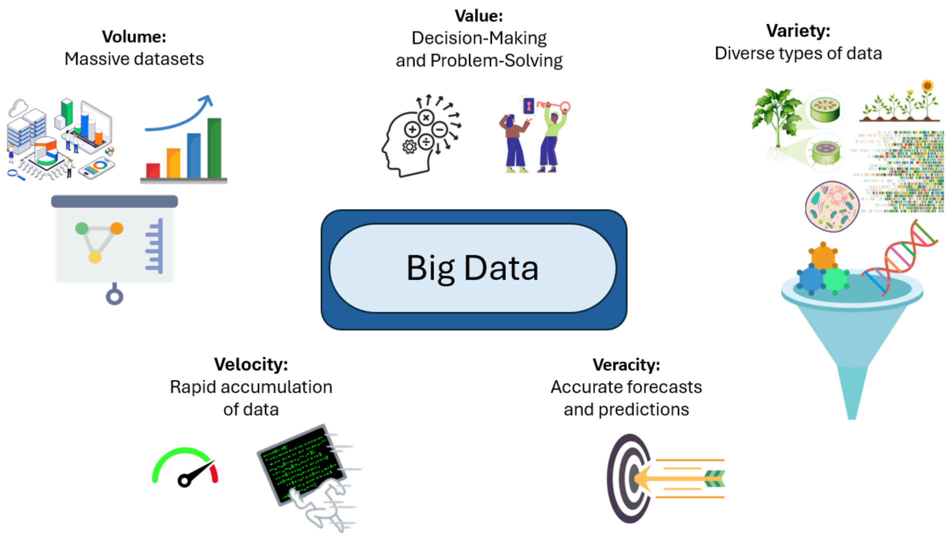

Figure 1. The five key attributes of Big Data (Volume, Velocity, Variety, Veracity, and Value). Each component describes a critical aspect of its nature and challenges.

A transformative use of AI and AI enabled learning methods like ML and DL in agriculture lies in the field of crop genetics and breeding. Crop breeding has been historically a time-intensive process, requiring years of research and experimentation to develop improved varieties. AI technologies like ML, DL, robotics, and computer vision can analyze vast amounts of data and identify patterns useful for selecting superior plant varieties with desired characteristics [7]. ML has dramatically accelerated this process by enabling the analysis of complex genetic, phenotypic, and environmental datasets encompassing crucial quantitative and qualitative characteristics [8]. ML also plays a role in addressing both biotic and abiotic plant stressors [9,10]. In this context ML algorithms predict disease outbreaks and identify resistant crop varieties through genetic screening and selection, helping breeders to develop more resilient crops [11]. These advancements are crucial for breeding resilient crops ensuring sustainability, particularly in the face of climate change.

ML models like random forests (RFs), support vector machines (SVMs), and gradient boosting algorithms (GBAs) allow breeders to identify plants with desirable traits—such as high yield, pest resistance, and drought tolerance—much faster and with greater accuracy than traditional methods [12]. RF is a commonly used supervised ML algorithm for both classification and regression. It combines multiple decision trees to improve accuracy and robustness and can effectively handle high-dimensional data, making it suitable for genomic data analysis [13]. Moreover, SVMs, which also belong to supervised ML models, are used for classification and regression (support vector regression) problems. They are powerful classifiers that work by finding the hyperplane that best separates different classes in the feature space. SVMs can be particularly useful in classifying plant varieties based on genetic markers, such as grapevine [14]. **GBAs** are a type of ML technique used for regression and classification tasks. They build an ensemble of weak learners, typically decision trees, in a sequential manner. Each new model is trained to correct the errors made by the previous models, aiming to minimize the overall prediction error using gradient descent [15].

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 February 2025 doi:10.20944/preprints202502.1208.v1

Deep Learning (DL), an advanced and specialized subset of Machine Learning (ML), has made a significant impact on the field of plant breeding by enabling the processing of high-dimensional and complex data. It is important to emphasize that DL is built upon the core principles of ML but distinguishes itself through its ability to automatically extract features from raw data and model intricate patterns, particularly when working with large datasets. For example, convolutional neural networks (CNNs), a class of DL models particularly well-suited for tasks involving spatial data, such as images or sequences, can analyze plant images to identify subtle phenotypic traits that might escape human observation, such as disease symptoms or growth anomalies [16].

The integration of ML techniques into plant phenotyping enables rapid and accurate analysis of multiple traits, which are influenced by environmental factors [17,18]. Recurrent neural networks (RNNs) designed to process sequential data, are used to study temporal patterns, such as changes in plant growth over time under varying environmental conditions [19]. Artificial neural networks (ANNs) which are computational models inspired by the structure and function of biological neural networks in animal and human brains, are used for prediction and identification of different species by analysing accurately and effectively complex morphological traits [20]. These tools provide breeders with detailed insights that inform the selection and development of superior crop varieties.

The emergence of big data and omic technologies has significantly enhanced the capabilities of ML in plant genetics and breeding. The power of big data analytics lies in its ability to integrate these extensive and varied datasets, creating a comprehensive framework for understanding plant biology. ML algorithms are particularly adept at linking omics data with phenotypic traits, enabling researchers to discover previously obscured relationships and patterns. With the exponential growth of genomic data and advances in AI algorithms, breeders can now leverage ML and DL techniques to revolutionise crop improvement [21].

In particular, genomic selection (GS) is a type of marker-assisted selection that uses dense molecular markers such as single nucleotide polymorphisms (SNPs) from the entire genome simultaneously in a linear regression model to predict the breeding potential of plants [22,23]. GS using AI enabled learning techniques, has been used to analyze big genomic data for superior genotypes identification and prediction of breeding values, optimizing the genotype selection process [24]. The key principle of GS is to build an accurate prediction model based on a training population consisting of individuals with both genotypic and phenotypic data [25]. ML models have proven invaluable in enhancing the efficiency of GS. By analyzing complex patterns within large datasets ML and DL algorithms forecast accurately phenotypic outcomes, enhancing the efficiency of plant breeding programs helping breeders make more strategic decisions [26].

ML and DL have also been implemented for prediction of genotypes-by-environment interactions, enhancing breeders' choices for well adaptive traits under diverse environments [27,28]. Moreover, reinforcement learning, one of the three basic ML paradigms, where the model is exposed to an environment and receives feedback in the form of rewards or penalties based on its actions [29], has been applied in simulated environments to optimize breeding strategies streamlining the GS process by efficiently allocating resources across different breeding generations [30].

In parallel, ML extends to gene editing technologies like CRISPR-Cas9, which have revolutionized modern biology by allowing precise modifications of plant genomes [31]. ML models support this process by analyzing genomic datasets to identify target regions for editing, enabling the introduction of traits such as salinity tolerance, improved nutrient profiles, and pest resistance [32,33]. By combining CRISPR-Cas9 with ML, researchers can develop crops that are better adapted to changing environmental conditions, ensuring long-term agricultural sustainability.

Members of the Solanaceae family are among the world's most important agricultural species. These crops are not only economically significant but also essential for global diets and nutrition. Several crops of this family e.g tomato (*Solanum lycopersicum*), potato (*Solanum tuberosum*), eggplant (*Solanum melongena*), tobacco (*Nicotiana tabacum*) and pepper (*Capsicum annuum*), have been subjected to intensive breeding for improved agricultural traits that lead to higher yields, resistance to biotic and abiotic stresses, longer shelf-life, as well as better taste and superior nutritional quality [34].

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 February 2025 doi:10.20944/preprints202502.1208.v1

However, long-term breeding objectives could be negated by alterations to the economic or physical environment [35].

AI stands as a transformative force, revolutionising Solanaceae crop breeding efforts with its array of groundbreaking technologies. The integration of ML applications in Solanaceaous crop breeding, highlights their transformative potential in accelerating trait prediction, disease detection and stress resilience, enabling breeders to efficiently analyze complex datasets, uncover genetic insights, and develop high-quality, resilient cultivars tailored to meet the challenges of modern agriculture in the perspective of climate change challenges (Figure 2). ML applications in Solanaceae research are being used to analyze large datasets, either focusing on specific traits individually or combining various types of data to predict cross-breeding and estimate different plants' performance under diverse conditions, analyzing numerous data and genotype-phenotype interactions [36,37].

The purpose of this study is to investigate the transformative applications of AI and specifically ML and DL in breeding Solanaceous crops, focusing on major representatives of this family, such as tomatoes, potatoes, peppers, and eggplants and to highlight the beneficial effects of AI methods to enhance important agricultural traits, and to promote innovations supporting global food security and environmental sustainability.

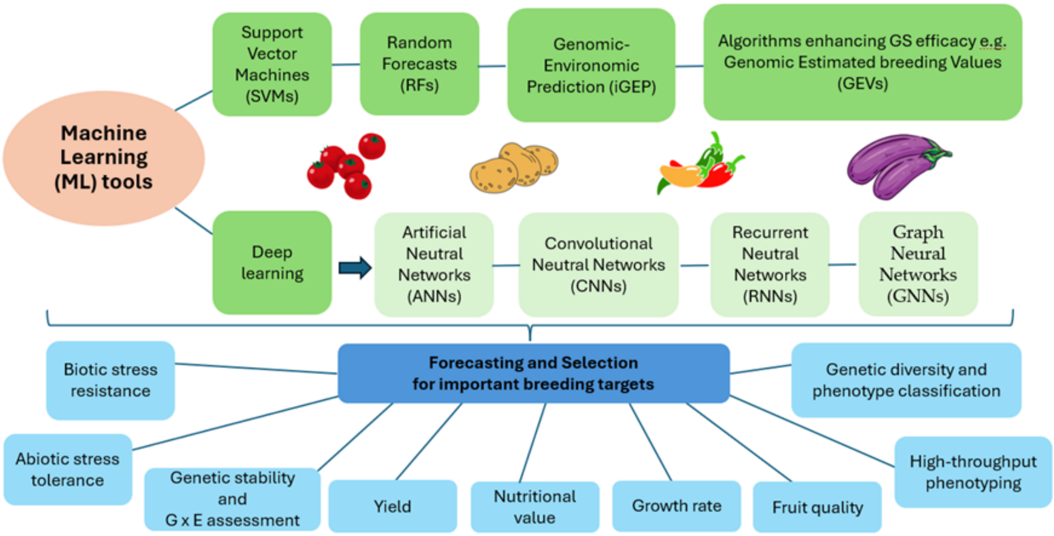

Figure 2. ML applications in Solanaceous crops to enhance breeding strategies for important selection targets.

2. Applications of Machine Learning in Solanaceaous Crop Breeding

2.1. Tomato

2.1.1. ML Applications for Productivity Monitoring and Yield Prediction

Advancements in AI, Augmented Reality (AR), and remote sensing technologies are revolutionizing tomato yield prediction, enabling more precise, data-driven decision-making for breeders and farmers. An innovative development in tomato yield estimation is the ARIA (Augmented Reality and Artificial Intelligence) mobile app, designed to detect, count, and classify tomatoes by capturing images through markerless AR technology. This app gives insights into the quality and maturity of preharvest tomatoes and provides reliable yield estimations in breeding programs [38]. Darra et al. [39], introduced satellite imagery ML model "*Sentinel-2 imaginary*", which provides satellite-based data with high-resolution spatial and temporal analysis, capturing key crop characteristics [39]. A collaborative study conducted in 2020, applied autonomous and remote-control techniques to remotely operating greenhouse models for cherry tomatoes production. Researchers employed a dynamic virtual greenhouse model that integrates KASPRO (KAS PROcesmodel, which is Dutch for "Greenhouse Process Model" (KAS = greenhouse, PROcesmodel =

process model)) (for climate modeling) and INTKAM (INTelligent Knowledge-based Adaptive Model) (for crop modeling) to enhance yield, quality, and net profit while reducing resource usage and costs [40].

Moreover, researchers utilized ANNs to predict tomato yields in greenhouse production. By conducting a sensitivity analysis they determined which input variables significantly impact prediction accuracy and showed that combining ANNs with sensitivity analysis can effectively enhance decision-making in greenhouse cultivation [41]. Furthermore, ANNs and multiple linear regression (MLR) models have been utilized for identification of the most sensitive traits affecting fruit yield in segregating generations of tomato during breeding programs [42].

2.1.2. ML Applications for Quality Traits and Seed Selection

Multiple prediction models have been developed in tomato to estimate critical quality traits, important to be incorporated in genotypes during breeding processes, such as aspartate content, fruit weight, firmness, ripeness, elasticity, soluble solids, pH, acidity, sugar, and carotene levels [43–47]. In addition, AI algorithms have been used to analyze fruit color in order to classify ripeness stages, achieving high accuracy in the process automatically [48]. Vazquez et al. [49], demonstrated a ML-based system which automatically classifies tomato fruits according to their shape, improving the efficiency of phenotypic characterization, highlighting its significance for tomato breeding and genetics. The researchers used ML algorithms to train a model for recognition and classification of different tomato fruit shapes. They also used image analysis techniques to extract shape-related information and they created a database containing all these features which served as the training and testing base for the model. The applications of this computational tool, could enhance the knowledge regarding the relationship between fruit shape and related genes, thus facilitating breeding programs, cultivar description and varietal registration and at the same time increase the classification accuracy for consumer-preferred fruit shape characteristics [49].

Yeon et al. [50], utilized two tomato germplasm collections (TGC1 with 162 accessions and TGC2 with 191 accessions) employing a 51K Axiom™ SNP array, and they estimated the GEVs of five quality characteristics (fruit weight, fruit width, fruit height, pericarp thickness, and total soluble solids (TSS) content). The researchers implemented parametric models (RR-BLUP, Bayes A, Bayesian LASSO) and non-parametric models (RKHS, SVM, RF) evaluating prediction accuracy across various cross-validation methods of GS models and marker sets. They concluded that the selection of the appropriate GS model can unravel preferable fruit traits in tomato breeding programs, potentially accelerating the development of elite cultivars [50].

AI genomic prediction (GP) models, which enable precise selection based on traits such as fruit morphology, color, and yield have also been used to forecast key visual and size traits in tomatoes, proving that GS has effectively accelerated the breeding process [51]. GP models have also been applied for morphometric and colormetric traits in tomato fruits, intergrating gemomic data with ML tools, like genomic best linear unbiased prediction (GBLUP), Bayesian models, and ML approaches such as RFs and SVMs, to identify key genetic markers associated with fruit shape, size, and color enchancing dission making and tomato breeding prosseses [36]. An innovative hybrid tomato AI breeding program has also been proposed to accelerate time to market by combining Seed-X's advanced computer vision and AI technology with TomaTech's. The researchers reported that this combination significantly increased the likelihood of desirable market characteristics and enhanced prediction capabilities in the breeding process. Furthermore, under the perspective of breeding for fruit quality, Khan and Adem in 2023, utilized the AI model "Connoisseur" which leverages the consumers' sensory scores for various tomato varieties, coupled with their chemical composition for identifying either these compounds are positively or negatively related to flavor [52].

ML algorithms like SVMs, RFs and Neutral Networks have also been implemented to classify and discriminate tomato seeds cultivars. Images of tomato seeds from various cultivars were captured using a consistent imaging setup and seed morphology-related traits were extracted. The ML models, after they tested for their accuracy and precision, applied to distinguish tomato seeds

 doi:10.20944/preprints202502.1208.v1

which belong to different cultivars based on their physical and morphological traits. The researchers demonstrate that the findings of the present study highlight the potential of integrating ML with seed imaging to modernize agricultural practices and facilitates the identification and tracking of promising cultivars in breeding experiments, emphasizing that this approach can be extended to other crops apart from tomato [53].

2.1.3. ML Applications for Breeding Against Environmental Stressors

ANNs have been employed to monitor tomato yield under environmental stresses like salinity, moisture deficiency, and diseases. Recent studies have integrated ML models, such as RFs and decision trees (DT), to develop predictive frameworks for managing irrigation and improving tomato production. These models have effectively captured the interactions between environmental and plant variables, supporting their application in crop management and breeding programs under limited water availability. Similarly, Zhang et al. in their study in 2023 [54], highlighted the potential of AI in breeding stress-tolerant tomato cultivars under various abiotic stressors including drought, salinity, cold and heat stress. They emphasized the use of machine-learning algorithms to identify genetic markers, DL models to analyze multi-omic and environmental data, and AI-driven tools for phenotype prediction. Additionally, advanced decision-support systems were proposed to assist breeders in selecting and optimizing genotypes for stress-prone environments. These approaches demonstrate AI's transformative role in precision breeding and crop management [54]. Moreover a study by Chowdhury et.al, focused on identifying drought-responsive genes in tomato utilizing various ML tools to analyse gene expression data. Researchers applied SVMs and RFs along with statistical analysis tools like Principal Component Analysis (PCA), to unravel critical drought-related genes and pathways, enhancing our understanding of drought tolerance and providing targets for genetic improvement in tomatoes [55]. Additionally, the applications of ML in plant metabolomics, to elucidate the biochemical effects of a non-microbial biostimulant on tomato plants under salt stress conditions, has also been studied by Chele et al. [56]. The research focused on how ML techniques can analyze complex metabolomic data to understand and enhance salt stress tolerance in tomatoes, highlighting the potential of integrating ML and metabolomics to improve crop resilience to environmental stressors.

In a study conducted by Bupi et al. [57] an integrated ML framework was developed to assess the severity of Tomato Yellow Leaf Curl Virus (TYLCV) infections. By leveraging advanced algorithms to identify patterns and predict infection severity, the framework incorporates tools for data preprocessing, feature selection, and model optimization. The researchers suggest that this ML framework can effectively and timely contribute to managing TYLCV in agricultural systems. Additionally, they highlight its potential applications in precision agriculture and its role in developing resistant cultivars for utilization in breeding programs. Moreover, Oliveira Dias et al. [58], employed RFs models to predict late blight severity in tomato plants caused by *Phytophthora infestans*. They utilized multispectral images captured by unmanned aerial vehicles (UAVs) to calculate vegetation indices, which served as input features for the models, integrating remote sensing applications. Two methodologies were tested: the first utilized images from the final day of evaluation, while the second incorporated images from four different evaluation days. The researchers emphasize the effectiveness and potential of integrating remote sensing and ML for phenotyping in real-world conditions. They highlight its applications in breeding programs aimed at developing late blight-resistant tomato varieties, as well as its utility for informed decision-making in crop management and accelerating the selection of superior cultivars. Gadade and Kirange [59], explored the use of classical ML techniques like SVMs, for identifying tomato leaf disease progress across different developmental plant stages. Although this study focuses on disease identification, it offers tools and methods that are highly beneficial for breeding programs by streamlining the disease evaluation process, providing accurate solutions for disease resistance screening and supporting the developments of tomato cultivars with enhanced disease resistance. Furthermore, Tan et. al. [60], tried to compare classical ML techniques (SVMs, k-Nearest Neighbors (k-NN), and RFs) and DL

 doi:10.20944/preprints202502.1208.v1

methods CNNs and pre-trained models) for classifying tomato leaf diseases. The results indicated that DL models outperformed classical ML in terms of classification accuracy and robustness and eliminated the need for manual feature extraction by learning features directly from the data. This study can contribute in tomato breeding, by highlighting that DL methods can provide a tool to enhance disease resistance screening and improve breeding efficiency [60].

Finally, Johansen et al., conducted field and unmanned aerial vehicles (UAVs) phenotyping experiment of 199 accessions of a wild tomato species (*Solanum pimpinellifolium*) for biomass and yield prediction under normal and saline conditions on UAV imagery based on shape characters. The researchers demonstrated the feasibility of biomass and yield predictions (two indicators of salt tolerance) up to eight weeks prior to harvesting [61], facilitating the identification of salt-tolerant accessions. Their tolerance traits could then be introgressed into commercial cultivars.

2.1.4. ML Applications for Multiple Trait Combinations and GP

Researchers have applied AI based learning tools focusing on breeding for multiple traits simultaneously. ML algorithms have been utilized to interpret genomic data to improve the efficiency of predictions for complex quantitative traits, such as yield and disease resistance, further advancing breeding efficiency by merging genomics with AI-driven computational tools [36,62]. Yamamoto et al. [63], discussed a simulation-based breeding strategy that integrates whole-genome prediction to improve complex characteristics like yield and flavor by leveraging genomic information. The main aspects studied were the integration of breeding stimulation and GP, the construction of phenotype prediction models, and the simulation for yield improvement and flavor-related traits. Recent research introduces the term Integrated Genomic-Environomic Prediction (iGEP), extending the traditional GP approach. It combines multi-omics data, big data technologies, and AI, particularly using machine and DL to increase the precision and credibility of phenotypic prediction. Although these studies' primary emphasis is on crops like rice, wheat, and maize the researchers highlight that the same methodologies and techniques can be applied to tomatoes successfully [64,65]. ANNs and MLR models have also been employed to investigate tomato callus formation, anther culture and the factors that influence these phenomena. Analysing the effects of multiple parameters such as plant genotype, concentration of plant growth regulators, cold temperature duration and flower length through ANNs and MLRs researchers can investigate the callus induction percentage and the numbers of regenerated calli [42]. Rearchers have also reviewed the potential of GP in tomato breeding, emphasizing that its implementation requires the optimization of varius factors, including field trial management, agronomic practices, seed production, phenotyping, and sequencing. Furthermore they highlight that a careful evaluation of parameters such as inbreeding levels, marker metrics, and the number of individuals to assess is essential. The researchers conclude that integrating GP into breeding programs like the single seed descent scheme and backcrossing can reduce the number of generations and streamline the selection process in tomato breeding. Additionally they emphasize that genotyping platforms can facilitate the identification of desirable and undesirable genotypes, thereby enhancing introgression of favourable traits [51,66].

2.2. Eggplant

2.2.1. ML Applications for Selecting Superior Plants Based on Yield Prediction

Various studies have explored AI-based methods to enhance yield estimation and agricultural management in eggplants. ML models using spectral vegetation indices derived from remote sensing data have significantly improved eggplant yield predictions. SVIs are mathematical combinations of reflectance values from different wavelengths captured by remote sensing devices. They are widely used to quantify vegetation characteristics such as health, biomass, chlorophyll content and water status [67]. Taşan et al. [68], demonstrated the accuracy of five distinct machine learning models—ANN, kNN, SVR, RF, and AB—were examined for their capacity to forecast eggplant yield at field scale, with varying input combinations, highlighting data-driven approaches to optimize precision

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 February 2025 doi:10.20944/preprints202502.1208.v1

agriculture [68]. Additionally, a study by Islam et al. [69], applied predictive algorithms, including regression and boosting techniques, for precise eggplant yield prediction of 130 locally collected eggplant genotypes. The study's overall findings demonstrated that combining vegetation index and crop data can significantly improve eggplant production modeling using ANN-based remote sensing, even though the data collected over three growing seasons is insufficient to make definitive judgments. These two studies underscore AI's utility in aiding breeders with selection of superior genotypes.

2.2.2. ML Applications for Growth Parameters and Seed Quality

Other studies have leveraged AI for crop quality enhancement. García-Fortea et al. [70], introduced MicroScan, a DL-based tool to identify the ideal stage for pollen induction in androgenesis. As a result, a more efficient method for producing doubled haploid lines provides a valuable tool for research in plant genetics and breeding, facilitating the production of doubled haploid lines [70]. Additionally, Sun et al. [71], used multispectral imaging and ML to classify eggplant seeds with greater accuracy, benefiting seed quality assessment through improved classification models. Furthermore, research conducted by Nomura et al. in 2023 [72], focused on developing a hybrid AI model for canopy photosynthesis rate estimation in eggplants, combining different data-driven techniques. The model combines ML methods and traditional modeling techniques to create an accurate and trustworthy system for predicting canopy photosynthesis rates under various environmental conditions and their impact on fruit quality, while researchers demonstrate that it can be applied effectively for greenhouse management optimizations in eggplants [72].

2.2.3. ML Applications for Breeding Against Environmental Stressors

Kaniyassery et al. [73], developed an AI-based disease detection system for eggplant, focusing on leaf spot and fruit rot diseases. The research addressed two primary aspects: the impact of meteorological variables on disease incidence and the AI-based classification of diseases using techniques such as image recognition and pattern analysis. The study utilized the YOLOv8 (You Only Look Once v8) model, a state-of-the-art DL algorithm for object detection, to accurately identify and classify disease symptoms from images. The researchers concluded that combining weather-based disease modeling with AI-driven classification offers a comprehensive approach to managing plant diseases, enhancing productivity and decision-making processes in eggplant breeding programs [73]. In another recent study, Lajom et al. [74], employed a SVM model integrated with near-infrared spectroscopy (NIRS) to detect eggplant fruit and shoot borer (EFSB) (*Leucinodes orbonalis*) infestations accurately at early stages. The results demonstrated a high degree of accuracy in identifying EFSB, marking a significant advancement in the integration of modern technology to agricultural pest management. This approach provides a valuable tool for eggplant farmers and breeders aiding in the selection of resistant genotypes and improving pest control strategies.. Additionally, Zhang et. al, [54] detected *Verticillium wilt* in eggplant leaves, combined VGG16, which is a convolutional neural network (CNN) architecture enhanced with a triplet attention mechanism. That trained VGG16-triplet attention model achieved a precision of 86.73% on the test set, demonstrating its effectiveness in detecting the disease and contributing to eggplant breeding efforts by addressing disease management and resistance traits in breeding programs [54].

2.2.4. ML Applications for Breeding Multiple Traits

Cemek et al. [75], addressed water management challenges by applying AI techniques to predict crop evapotranspiration (ET) for eggplants. Models like ANNs and SVMs provided reliable ET estimates based on environmental and crop data, supporting efficient water usage in irrigation. ANNs have also been utilized in studies aiming to model the relationship between integrated nutrient management practices and eggplant yield and quality. The study of Thingujam et al. [76],

 doi:10.20944/preprints202502.1208.v1

incorporates numerous nutrient management strategies combining organic and inorganic nutrient sources aiming to assess their impact on the growth, yield, and quality of eggplants. Overall, the ANNs models application can effectively be utilized as an optimal nutrient management guide to making decisions for better eggplant fruit quality and yields, while using nutrient availability effectively and sustainably.

DL techniques have also been proposed to address challenges in horticultural crops including eggplants. More specifically, AlexNet, a pioneering deep convolutional neural network (CNN) architecture, particularly based on image recognition and computer vision tasks, and VGG-16, which is a convolutional neural network (CNN) architecture, widely used in computer vision tasks, are introduced by a review article from Yang et al. 2021. Researchers propose AlexNet and VGG-16 for five eggplant diseases classification using smartphones revealing promising results.

The aforementioned advancements underscore AI's potential in improving agricultural practices, particularly in yield prediction, water management, and crop quality breeding for eggplants.

2.3. Potato

2.3.1. ML Applications for Productivity Monitoring and Yield Prediction

ML has emerged as a pivotal tool in agriculture; specifically in potato (*Solanum tuberosum* L.) research, by providing robust solutions to complex challenges in yield forecasting, quality monitoring, disease detection, and overall crop management. ML has facilitated significant advancements for yield prediction, which could have potential implications in plant breding, as evidenced by numerous studies that integrate satellite imagery, climatic data, and agronomic parameters. Salvador et al. [78], employed a combination of meteorological data, field observations, and satellite imagery with five ML algorithms—RF, support vector machine linear (svmL), support vector machine polynomial (svmP), support vector machine radial (svmR), and general linear model (GLM)—across six time frames to assess yield prediction models in Mexico. The SVM-polynomial model, when trained with the first five months of data post-sowing, was the most effective for predicting yield during the summer cycle, while the RF model performed best in the winter cycle with only three months of data [78]. The proposed methodology can predict potato yield prior to harvest, making it highly valuable for developing food security strategies

Similarly, Gómez et al. [79], in Spain developed predictive models using Sentinel satellite imagery to support precision agriculture. By testing nine ML algorithms in their initial study—ranging from generalized linear models (GLM) to k-nearest neighbors (KNN) and model-averaged neural networks (avNNet)—they were able to identify the models best suited for potato yield forecasting. In a subsequent study, Gómez et al. [79] focused on SVM-radial and RF algorithms and introduced the Potato Productivity Index (PPI), a novel metric for yield prediction. Their findings validated the effectiveness of the PPI index, underscoring the potential of ML and remote sensing data to refine yield estimations in regional potato production [79,80]. Additionally, Kurek et al. [81], conducted research in Poland, utilizing agronomic, climatic, soil, and satellite data across five growing seasons on 114 commercial potato fields. By applying ML techniques such as linear regression, ridge, Lasso, Elastic Net, XGBoost, RF, multilayer perceptron (MLP), stochastic gradient descent (SGD), and support vector regression (SVR), they developed three predictive models: non-satellite, satellite, and hybrid, the latter achieving the lowest mean absolute percentage error (MAPE) [81]. El-Kenawy et al. [82] assessed several predictive models—such as K-nearest neighbors (KNN), gradient boosting, XGBoost, multilayer perceptron (MLP), graph neural networks (GNNs), gated recurrent units (GRUs), and long short-term memory networks (LSTMs)—using metrics like mean squared error (MSE), root mean squared error (RMSE), and mean absolute error (MAE) to predict potato yield. Their results indicate that, GNNs and LSTMs offer superior accuracy and effectively capture complex spatial and temporal patterns [82].

Li et al. [83], combined cultivar-specific data with UAV (unmanned aerial vehicle) remote sensing to improve yield predictions in Minnesota. Using RF regression and SVR, they found that early-season UAV spectral data—particularly at the tuber initiation stage in late June—correlated strongly with marketable yield. Their results revealed that combining high-resolution UAV imagery with cultivar data significantly outperformed yield prediction models that lacked cultivar-specific information, highlighting the potential of early detection for yield optimization [83].

Coulibali et al. [84], studied gradients in the elemental composition of a potato leaf tissue (i.e. its ionome) that are linked to crop potential and therefore have applications in plant breeding. Because the ionome is a function of genetics and environmental conditions, practitioners aim at fine-tuning fertilization to obtain an optimal ionome based on the needs of potato cultivars. Their objective was to assess the validity of cultivar grouping and predict potato tuber yields using foliar ionomes. Their dataset comprised 3382 observations in Québec (Canada) from 1970 to 2017. The first mature leaves from top were sampled at the beginning of flowering for total N, P, K, Ca, and Mg analysis. They also used the preprocessed ionomes to assess their effects on tuber yield classes (high- and low-yields) on a cultivar basis using k-nearest neighbors, RF and SVMs classification algorithms. Their ML models returned an average accuracy of 70%, a fair diagnostic potential to detect in-season nutrient imbalance of potato cultivars [84].

Yu et al. [85] on the other hand highlighted the importance of accurately estimating potato Leaf Area Index (LAI) for optimizing yield prediction and management practices. Using UAV-based remote sensing, their study combined data from RGB images, LiDAR, and hyperspectral imaging (HSI). Four ML models—SVR, Random Forest Regression (RFR), Histogram-based Gradient Boosting Regression Tree (HGBR), and Partial Least-Squares Regression (PLSR)— analyzed features from these data sources, with HSI showing the highest predictive accuracy due to its rich spectral information. Combining all features across sensors achieved the highest R^2 (0.782), with RF Regression excelling in feature integration. This approach not only advances LAI estimation but also has potential applications in breeding programs and precision agriculture [85].

2.3.2. ML Applications for Variey Identification and Potato Tuber Quality

Rahman et al. [86]. explored the use of DL models for potato breed recognition, employing five state-of-the-art convolutional neural network models, namely: VGG16, ResNet50, MobileNet, Inception-v3, and another custom CNN model. These models were trained on images of various potato breeds to differentiate them based on visual traits such as size, shape, color, texture, and skin pattern. Performance evaluation revealed that the customized CNN model achieved the highest accuracy at 94.84%, demonstrating its superiority for this task [86]. Similarly, Azizi et al. [87], proposed a method for identifying and differentiating 10 potato varieties by integrating machine vision and ANNs. Non-linear ANNs achieved a perfect classification accuracy of 100%. The findings underscore the efficacy of combining machine vision with neural networks for precise potato variety identification [87].

2.3.3. ML Applications for Breeding Against Environmental Stressors

Potato crops are highly susceptible to fungal diseases like early blight (*Alternaria solani*) and late blight (*Phytophthora infestans*), leading to significant yield losses. ML, through image analysis [88], monitoring of stress factors and optimizing nutrient management [88], has improved disease detection. A plethora of ML tools like SVMs, RFs, ANN and CNNs, have been implemented in various studies for efficient detection of plant diseases [89–96] enhancing genotype selection in breeding programs. ML applications in disease detection also extend to viral infections, underscoring the critical role of ML in advancing virus detection and supporting healthier crop management practices [43,98,99].

Sugiura et al. [98], developed an image classification method to detect virus-infected plants in potato seed production fields in Japan, aiming to improve the roguing process during selection programs. In this study, RGB images were captured using an unmanned aerial vehicle (UAV) from 5 to 10 meters above the ground. A convolutional neural network (CNN) achieved 96% accuracy in training and 84% in validation, demonstrating the potential of UAV-based image classification for effective virus detection in potato fields. This method is particularly important for plant breeding, as it enables the efficient identification of virus-infected plants, ensuring the production of virus-free seed tubers and contributing to the overall health and productivity of potato crops [98].

ML has also been applied to monitor stress factors and optimize nutrient management in potato crops. Gold et al. [89], analyzed physiological responses in potato cultivars with varied resistance to late blight by examining their spectral reflectance following exposure to *Phytophthora infestans*. Using ML algorithms, including RF and partial least squares discriminant analysis (PLS-DA), they showed that specific genotypic traits significantly influence disease response, providing insights into the complex host-pathogen interactions and helping identify cultivars with natural disease resistance. These findings highlight the potential of ML to improve understanding of crop resilience and facilitate the selection of stress-resistant varieties [89].

Boguszewska-Mańkowska et al. [100], investigated drought tolerance variability among 50 potato cultivars by analyzing morphological traits under different water regimes over 11 consecutive years. The study focused on tuber yield, plant tolerance indices, and Climatic Water Balance to assess stability in drought conditions. To enhance the classification of drought tolerance groups, several ML algorithms, including Quadratic Discriminant Analysis, RF, Extra Trees, AdaBoost, and extreme gradient boosting, were evaluated. Extreme gradient boosting emerged as the most effective classifier, achieving an accuracy of 96.7% [100].

Lapajne et al. [101], explored the use of hyperspectral imaging and attention-based DL models to detect drought stress in potato plants. Their study involved two potato cultivars exposed to water-deficient conditions and used dual-sensor hyperspectral imaging (Visible and Near-Infrared/VNIR and Short-Wave Infrared/ SWIR) to identify critical wavelengths related to drought stress.

These applications can be utilized in plant breeding by improving detection of the effects of environmental stressors in breeding programs, allowing for more efficient selection of plants with desirable traits, thus enhancing breeding outcomes and crop sustainability.

2.4. Pepper

2.4.1. ML Applications for Yield Prediction and Favourable Agronomic Traits

Lozada et al. [102], implemented ridge regression and DL-based models to estimate genomic breeding values for yield and agronomic traits in 204 Capsicum genotypes evaluated across multi-environment trials in New Mexico, USA. Their study aimed to assess the accuracy of GP for traits related to yield, morphology, and phenology, examine the impact of marker subsets on prediction accuracy, and evaluate selection responses for various strategies. Using six models they highlighted the promise of genome-wide selection for chile pepper breeding. The study underscored the importance of large training datasets to enhance the accuracy of DL models [102].

2.4.2. ML Applications for Variety Identification, Chemical Clasification, Seed Selection and Fruit Quality

Sabanci et al. [103], explored the use of computer vision and AI to classify pepper seeds from different cultivars, which is crucial for breeding programs. In this study, images of seeds from green, orange, red, and yellow pepper cultivars were captured using a flatbed scanner. The following approaches were proposed for classification: the first involved training CNN models (ResNet18 and ResNet50), achieving accuracies of 98.05% and 97.07%, respectively. The second approach involved fusing features from pre-trained CNN models and applying feature selection before classifying with a SVM. The CNN-SVM-Cubic model achieved up to 99.02% accuracy offerings high precision and efficiency in plant breeding [103]. Moreover, Kurtulmuş et al. [104], developed a classification method to discriminate pepper seed varieties using neural networks and computer vision. The multilayer perceptron model with 30 neurons in the hidden layer, trained using resilient back propagation,

 doi:10.20944/preprints202502.1208.v1

achieved an accuracy of 84.94% in classifying eight pepper seed varieties [104]. Additionally, Tu et al. [105], focused on improving the selection of high-quality pepper seeds by automating the recognition of seed features. The study identified several physical traits, such as color, size, and weight, as key indicators of seed vigor. The best predictive model, based on a multilayer perceptron (MLP) neural network using 15 physical traits, achieved a high stability rate of 99.4%. The model significantly improved germination rates and selection efficiency, reaching up to 79.4% germination and 90% selection rate. This automated approach shows potential for reducing costs and labor in seed selection, making it an effective tool for quality control in pepper breeding programs [105].

Ramírez-Meraz et al. [106], applied 1H NMR-based metabolomics combined with ML, specifically RF, to study the metabolic fingerprinting of ten experimental races of *Capsicum annuum* cv. Jalapeño. Their analysis classified and evaluated these races based on differential metabolite profiles, commercial traits, and multivariate data analysis. The study revealed variations in carbohydrate, amino acid, nucleotide, and organic acid contents among the races. RF identified length, width, weight, and yield as key variables for accurately distinguishing between the races, highlighting critical traits for commercial and breeding applications.

2.4.3. ML Applications for Breeding Against Environmental Stressors

AI can facilitate and improve selection efficiency for resilient genotypes against environmental stressors. Dissanayake et al. [107], developed an effective method for detecting diseases and nutrient deficiencies in bell peppers, focusing on the rapid spread of powdery mildew and magnesium deficiency. The study integrated CNNs to enhance detection accuracy, achieving a 93% success rate in distinguishing the health status of bell pepper leaves, with 97% accuracy in identifying magnesium deficiency and powdery mildew. The approach also demonstrated 98% accuracy in assessing the progression of powdery mildew and 96% in magnesium deficiency [107]. Haque et al. [108], highlighted the importance of detecting pepper diseases quickly and accurately to prevent significant losses in pepper production. The study utilized several pre-trained DL models, including VGG-19, Xception, NasNet Mobile, MobileNet-V2, ResNet-152-V2, and Inception-ResNet-V2, to extract deep features from pepper plant images for disease identification. The customized CNN models achieved high accuracy, with VGG-19 and ResNet-152-V2 reaching an impressive 96.26% accuracy. Additionally, Xception outperformed Inception-ResNet-V2, MobileNet-V2, and NasNet-Mobile, achieving a 93.46% accuracy. These results suggest that DL models can be effectively used for early disease detection in pepper crops, helping farmers minimize losses by enabling rapid identification and treatment of diseases and for breeding programs to ensure disease resistance in pepper cultivars [108].

Fumia et al. [109], conducted a comparative study of genomic and phenomic selection methodologies to identify heat-tolerant genotypes within a core collection of 300 *Capsicum annuum* accessions, representing 84.1% of the species' diversity. Initially, anomaly analysis via k-means clustering was utilized to identify individuals exhibiting anomalous behavior under heat stress compared to optimal conditions, based on phenotypic data. This analysis informed the training of a RF, ML model capable of classifying heat-tolerant genotypes with near-perfect accuracy using only data from trials under optimal conditions. Subsequently, a genomic-based predictive analysis was performed, leveraging genomic data to predict component traits and generate a weighted rank-sum selection index (WRSSI) to identify heat-tolerant lines. Finally, the selected lines were compared across three selection methodologies: (1) breeder's intuition, (2) phenomics-based anomaly analysis, and (3) genomics-based predictive modeling and selection index. The study concluded that integrating classical and multispectral phenotyping techniques enhances selection efficiency and outcomes [109].

Moreover, Islam et al. [69], developed a method for classifying early-stage stress symptoms in pepper seedlings using image processing and a SVM. The study investigated the effects of different environmental factors (temperature, light intensity, and day-night cycles) on stress symptoms. Using RGB camera images, the researchers extracted 18 color features, nine texture features, and one

 doi:10.20944/preprints202502.1208.v1

morphological feature. The SVM model, validated with cross-validation, achieved an accuracy of 85%. This system provides a way for real-time stress monitoring, enabling growers to optimize environmental conditions for improved seedling growth. This approach could also assist in accelerating the identification of stress-resistant traits, aiding in the development of improved cultivars with enhanced resilience to environmental stresses.

Ataş et al. [110], explored the use of hyperspectral imaging for detecting aflatoxin contamination in chili peppers, offering a rapid, non-destructive alternative to traditional chemical methods. The study utilized both UV and Halogen excitations, extracting features from individual spectral bands and their differences. ML classifiers, including multi-layer perceptrons (MLPs) and linear discriminant analysis (LDA), were applied, achieving robust classification performance with fewer spectral bands. This method could be useful in breeding programs for selecting aflatoxin-resistant cultivars, enhancing food safety and quality.

3. Conclusions

The history and evolution of ML provides a robust framework for understanding its contemporary applications in plant breeding. As methodologies continue to evolve, AI ML and DL applications could become a cornerstone for addressing critical matters in agriculture and food security in the future. ML and DL techniques can revolutionize plant breeding by speeding up decision-making and improving precision through the creation of advanced predictive models that quickly respond to economic and environmental challenges. The combination of the innovative AI tools, ML models and Big Data Science with traditional breeding methods can optimize Solanaceae crop breeding, and enhance efficiency by accelerating the pace and precision of breeding efforts creating new advanced varieties with superior agronomic traits [42].

The capacity of ML algorithms to uncover hidden data relationships makes them essential partners in developing sophisticated breeding strategies that integrate multiple parameters, supporting the creation of crops that are not only highly productive but also resilient and sustainable across diverse agricultural environments.Therefore AI advancements aim to help breeders to utilize the new technologies and their transformative impact on agriculture, to develop cultivars well-adapted to various cultivation systems. It also aims to enhance their effort to make decisions rapidly and precisely, as the new trend of cultivation cropping systems demands [11,46,65].

The present study underscores the transformative potential of AI and more specifically ML and DL, in driving advancements in the genetic improvement of tomatoes, potatoes, peppers, and eggplants. By harnessing advanced algorithms, big data analytics, multiomics and gene-editing technologies, these innovations not only accelerate breeding cycles but also enable precise selection of traits such as yield, pest resistance, and climate adaptability, creating thus opportunities for personalized breeding strategies on Solanaceous crops, tailored to the specific needs of different regions and agricultural conditions. The adoption of these emerging technologies enables plant breeders to develop resilient and high-yielding Solanaceous crops capable of addressing all critical challenges such as food security, climate change, and resource scarcity ensuring that innovations effectively address the diverse agricultural needs worldwide.

The future of ML and DL in plant genetic improvement appears exceptionally promising as new technologies continue to evolve. Multimodal AI systems capable of analyzing and integrating diverse data types — such as genetic information, plant images, and environmental parameters — are set to revolutionize crop management by providing breeders with a more holistic approach to decision-making. Automated ML (AutoML) will further facilitate the use of ML by researchers without prior experience in data analysis, allowing for faster development of models that can be integrated into genetic breeding programs. Additionally, the convergence of quantum computing and ML holds the potential to significantly enhance data processing speeds and analytical capabilities, unlocking unprecedented opportunities for innovation in plant breeding.

Author Contributions: Conceptualization, V.P., E.T. and P.J.B; writing—original draft preparation, M.G., A.K., and K.K.; writing—review and editing, V.P., E.T., and P.J.B.; visualization, M.G. and A.K.; supervision, V.P., E.T., and P.J.B.; project administration, V.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Not Applicable

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

References

1. Monteiro. A.; Santos, S.;, Gonçalves, P. Precision Agriculture for Crop and Livestock Farming-Brief Review. Animals (Basel). 2021 Aug 9;11(8):2345. doi: 10.3390/ani11082345. PMID: 34438802; PMCID: PMC8388655.
2. Talaviya, T.; Shah, D.; Patel, N.; Yagnik, H.; Shah, M. Implementation of artificial intelligence in agriculture for optimisation of irrigation and application of pesticides and herbicides, Artificial Intelligence in Agriculture, Volume 4, 2020, Pages 58-73, doi.org/10.1016/j.aiia.2020.04.002.
3. Dengiz, O.; Alaboz, P.; Saygın, F.; Adem, k.; Yüksek, E. Evaluation of soil quality of cultivated lands with classification and regression-based machine learning algorithms optimization under humid environmental condition, Advances in Space Research, 2024, 74, 11, 5514-5529, doi.org/10.1016/j.asr.2024.08.048.
4. Khandelwal, C.;, Singhal, M.; Gaurav, G.: Dangayach, G.S.; Meena, M.L. Agriculture Supply Chain Management: A Review (2010–2020), Materials Today: Proceedings, Volume 47, Part 11, 2021, Pages 3144-3153, doi.org/10.1016/j.matpr.2021.06.193
5. Mogili, U.M.R.;. Deepak, B.B.V.L. Review on application of drone systems in precision agriculture International Conference on Robotics and Smart Manufacturing. Procedia Computer Science 133, 2018, pp. 502-509
6. Qiang, J.; Zhang, X.. Data analytics for crop management: A big data view. *Journal of Big Data*, 2022, 9(1), Article 52. https://doi.org/10.1186/s40537-022-00668-2
7. Balasubramanian, V.; Guo, W.; Chandra, A.; Desai, S.V. Computer Vision with Deep Learning for Plant Phenotyping in Agriculture: A Survey. Advanced Computing and Communications, 2020, 10.34048/ACC.2020.1.F1.
8. Fernandes, I.K.; Vieira, C.C.; Dias, K.O.G. Using machine learning to combine genetic and environmental data for maize grain yield predictions across multi-environment trials. Theor. Appl. Genet., 2024, 137, 189. https://doi.org/10.1007/s00122-024-04687-w
9. Gou, C.; Zafar, S.; Hasnain, Z.; Aslam, N.; Iqbal, N.; Abbas, S.; Li, H.; Li, J.; Chen, B.; Ragauskas, A.; Abbas, M. Machine and Deep Learning: Artificial Intelligence Application in Biotic and Abiotic Stress Management in Plants. Frontiers in Bioscience-Landmark, 2024, 29. 20. 10.31083/j.fbl2901020.
10. Moustaka, J.; Moustakas, M. Early-Stage Detection of Biotic and Abiotic Stress on Plants by Chlorophyll Fluorescence Imaging Analysis. Biosensors, 2023, 13, 796. https://doi.org/10.3390/bios1308079
11. Yoosefzadeh, Najafabadi M;, Hesami, M,; Eskandari, M. Machine Learning-Assisted Approaches in Modernized Plant Breeding Programs. Genes (Basel), 2023 Mar 23;14(4):777. doi: 10.3390/genes14040777. PMID: 37107535; PMCID: PMC10137951.
12. van Dijk, A.D.J.; Kootstra, G.; Kruijer, W.; de Ridder, D. Machine learning in plant science and plant breeding, iScience, Volume 24, Issue 1, 2021, 101890, https://doi.org/10.1016/j.isci.2020.101890.

13. Montesinos López, O.A.; Montesinos López, A.; Crossa, J. Random Forest for Genomic Prediction. In: Multivariate Statistical Machine Learning Methods for Genomic Prediction, 2020, Springer, Cham. https://doi.org/10.1007/978-3-030-89010-0_15
14. Gutiérrez, S.; Tardaguila, J.; Fernández-Novales, J.; Diago, M.P. Support Vector Machine and Artificial Neural Network Models for the Classification of Grapevine Varieties Using a Portable NIR Spectrophotometer. PLOS ONE, 2015, 10(11): e0143197. https://doi.org/10.1371/journal.pone.0143197
15. Yan, J., Xu, Y., Cheng, Q. et al. LightGBM: accelerated genomically designed crop breeding through ensemble learning. Genome Biol, 2021, 22, 271. https://doi.org/10.1186/s13059-021-02492-y
16. Karahan, T.; Nabiyev, V. Plant identification with convolutional neural networks and transfer learning. Pamukkale University Journal of Engineering Sciences, 2021, 27. 638-645. 10.5505/pajes.2020.84042.
17. Singh, A.; Ganapathysubramanian, B.; Singh, A.K.; &Sarkar, S. Machine learning for high-throughput stress phenotyping in plants. Trends in plant science, 2016, 21(2), 110-124.
18. Gill, T.; Gill, S. K.;, Saini, D. K.; Chopra, Y.; de Koff, J. P.; Sandhu, K.S. A comprehensive review of high throughput phenotyping and machine learning for plant stress phenotyping. Phenomics, 2020, 2(3), 156-183.
19. Mohmed, G.; Heynes, X.; Naser, A. et al. Modelling daily plant growth response to environmental conditions in Chinese solar greenhouse using Bayesian neural network. Sci. Rep., 2023 13, 4379. https://doi.org/10.1038/s41598-023-30846-y
20. Ferreira, M. G., Azevedo, A. M., Siman, L. I., da Silva, G. H., Carneiro, C. D. S., Alves, F. M., Nick, C. Automation in accession classification of Brazilian Capsicum germplasm through artificial neural networks. Scientia Agricola, 2017, 74(3), 203-207
21. Esposito, S.; Ruggieri, V.; Tripodi, P. Editorial: Machine Learning for Big Data Analysis: Applications in Plant Breeding and Genomics. Front. Genet., 2020, 13:916462. doi: 10.3389/fgene.2022.916462
22. Nguyen, T.T.; Huang, J.; Wu, Q.; Nguyen, T.; Li, M. Genome-wide association data classification and SNPs selection using two-stage quality-based Random Forests. BMC Genomics, 2015; 16 Suppl 2:S5. doi: 10.1186/1471-2164-16-S2-S5. Epub 2015 Jan 21. PMID: 25708662; PMCID: PMC4331719.
23. Montesinos-López, O.A.; Montesinos-López, A.; Crossa, J.; Gianola, D.; Hernández-Suárez, C.M.; Martín-Vallejo, J. Multi-trait, Multi-environment Deep Learning Modeling for Genomic-Enabled Prediction of Plant Traits. G3 (Bethesda). 2018 Dec 10;8(12):3829-3840. doi: 10.1534/g3.118.200728
24. Crossa, J.; Pérez-Rodríguez, P.; Cuevas, J.; Montesinos-López, O.; Jarquín, D.; De Los Campos, G.; Varshney, R. K. (2017). Genomic selection in plant breeding: methods, models, and perspectives. Trends in plant science, 22(11), 961-975.
25. Montesinos-López, A.; Montesinos-López, O.A.; Gianola, D.; Crossa, J.; Hernández-Suárez, C.M. Multi-environment Genomic Prediction of Plant Traits Using Deep Learners With Dense Architecture, *G3 Genes|Genomes|Genetics*, Volume 8, Issue 12, 1 December 2018, Pages 3813–3828, https://doi.org/10.1534/g3.118.200740
26. Xu, Y., Zhang, X., Li, H., Zheng, H., Zhang, J., Olsen, M. S., Qian, Q. Smart breeding driven by big data, artificial intelligence, and integrated genomic-enviromic prediction. Molecular Plant, 2022, 15(11), 1664-1695.
27. Heslot, N., Yang, H. P., Sorrells, M. E., & Jannink, J. L. Genomic selection in plant breeding: a comparison of models. Crop science, 2012, 52(1), 146-160.
28. Birhanu, C., Mekbib, F., Lule, D., Bekeko, Z., Girma, G., Tirfessa, A., ... & Mengiste, T. Genotype by environment interactions and stability for grain yield and other agronomic traits in selected sorghum genotypes in Ethiopia. Agrosystems, Geosciences & Environment, 2024, 7(3), e20544.

29. Lee, D.; Seo, H.; Jung, M.W. Neural basis of reinforcement learning and decision making. Annu Rev Neurosci. 2012; 35:287-308. doi: 10.1146/annurev-neuro-062111-150512. Epub 2012 Mar 29. PMID: 22462543; PMCID: PMC3490621.
30. Moeinizade, S.; Hu, G.; Wang, L. A reinforcement Learning approach to resource allocation in genomic selection, Intelligent Systems with Applications, 14, 2022, 200076, https://doi.org/10.1016/j.iswa.2022.200076.
31. Amoah, P.; Oumarou Mahamane, A.R, ;Byiringiro M.H.; Mahula N.J.; Manneh, N, Oluwasegun, YR.; Assfaw, A.T.; Mukiti, H.M.; Garba, A.D.; Chiemeke, F.K.; Bernard, Ojuederie O.; Olasanmi, B. Genome editing in Sub-Saharan Africa: a game-changing strategy for climate change mitigation and sustainable agriculture. GM Crops Food. 2024 Dec 31;15(1):279-302. doi: 10.1080/21645698.2024.2411767. Epub 2024 Oct 31. PMID: 39481911; PMCID: PMC11533803.
32. Angon, P.B.; Mondal, S.; Akter, S.; Sakil, Md.A.; Jalil, Md.A Roles of CRISPR to mitigate drought and salinity stresses on plants, Plant Stress, 2023, 8, 100169,https://doi.org/10.1016/j.stress.2023.100169.
33. Manoj, K.; Ranjan, P.M.; Manish K.P.; Kumar, S.P.; Abhishek, B.; Baozhu, G.; Rajeev K.V. Application of CRISPR/Cas9-mediated gene editing for abiotic stress management in crop plants. Frontiers in Plant Science, 14, 2023, https://www.frontiersin.org/journals/plant-science/articles/10.3389/fpls.2023.1157678
34. Bai, Y.; Lindhout, P. (2007) Domestication and Breeding of Tomatoes: What Have We Gained and What Can We Gain in the Future. Annals of Botany, 100, 1085-1094. https://doi.org/10.1093/aob/mcm150
35. Gasparini, K.,; Figueiredo, Y. G.; Araújo, W. L.; Peres, L. E.; Zsögön, A. (2024). De novo domestication in the Solanaceae: advances and challenges. *Current Opinion in Biotechnology*, *89*, 103177.
36. Tong, H.; Nankar, A. N.; Liu, J.; Todorova, V.; Ganeva, D.; Grozeva, S.; Nikoloski, Z. (2022). Genomic prediction of morphometric and colorimetric traits in Solanaceous fruits. Horticulture research, 9, uhac072.
37. Palakshappa, A.; Kuthkunja, A.; Aditya, G.; Anvitha, V. Analysis of Leaf Disease Detection in the Solanaceae family plants using Machine Learning Algorithms. 2024, 1-6. 10.1109/IITCEE59897.2024.10467645.
38. Balaji Prabhu B.V. ARIA: Augmented Reality and Artificial Intelligence enable mobile application for Yield and grade prediction of tomato crops. Procedia Computer Science 235, 2024, 2693–2702
39. Darra, N.; Espejo-Garcia, B.; Kasimati, A.,; Kriezi, O.; Psomiadis, E.; Fountas, S. Can satellites predict yield? Ensemble machine learning and statistical analysis of Sentinel-2 imagery for processing tomato yield prediction. Sensors, 2023, 23(5), 2586.
40. Hemming, S.; Zwart, F.d.; Elings, A.; Petropoulou, A.; Righini, I. Cherry Tomato Production in Intelligent Greenhouses—Sensors and AI for Control of Climate, Irrigation, Crop Yield, and Quality. Sensors 2020, 20, 6430. https://doi.org/10.3390/s20226430
41. Belouz, K.;Nourani, A.; Zereg, S.; Bencheikh, A. Prediction of greenhouse tomato yield using artificial neural networks combined with sensitivity analysis. Scientia Horticulturae, 2022, 293, 110666.
42. Acharya, B.; Kumar, P.P.; Hazra, S.; Dutta, S.; Saha, S.; Roy, S.; Maji, A.; Chakraborty, I.; Chattopadhyay, A.; Hazra, P. Genetic control of important yield attributing characters predicted through machine learning in segregating generations of interspecific crosses of tomato (Solanum lycopersicum L.). Acta Physiologiae Plantarum, 2024, 46(8):78. doi:10.1007/s11738-024-03702-3.
43. Polder, G.; Blok, P.M.; De Villiers, H.A.C.; Van Der Wolf, J.M.; Kamp, J. Potato Virus Y Detection in Seed Potatoes Using Deep Learning on Hyperspectral Images. Front. Plant Sci., 2019, 10, 209. https://doi.org/10.3389/fpls.2019.00209

44. Schmilovitch, Z.; Ignat, T.; Alchanatis, V.; Gatker, J.; Ostrovsky, V.; Felföldi, J.Hyperspectral imaging of intact bell peppers, Biosystems Engineering, Volume 117, 2014, Pages 83-93, https://doi.org/10.1016/j.biosystemseng.2013.07.003.
45. Mollazade, K.; Omid, M.; Akhlaghian Tab, F.; Rezaei kalaj, Y.; Mohtasebi, S. Data Mining-Based Wavelength Selection for Monitoring Quality of Tomato Fruit by Backscattering and Multispectral Imaging. International Journal of Food Properties, 2015, 18. 880-896. 10.1080/10942912.2013.835822.
46. Zhu, W.; Li, W.; Zhang, H.; Li, L. Big data and artificial intelligence-aided crop breeding: Progress and prospects. Journal of Integrative Plant Biology, 2024, 10.1111/jipb.13791.
47. Duangjit, J.; Causse, M.; Sauvage, C. Efficiency of genomic selection for tomato fruit quality. Molecular Breeding, 2016, 36. 10.1007/s11032-016-0453-3.
48. El-Bendary, N.; El-hariri, E.; Hassanien, A.E.; Badr, A. Using machine learning techniques for evaluating tomato ripeness. Expert Systems with Applications, 2014, 42. 10.1016/j.eswa.2014.09.057.
49. Vazquez, D.V.; Spetale, F.E.; Nankar, A.N.; Grozeva, S.; Rodríguez, G.R. Machine Learning-Based Tomato Fruit Shape Classification System. Plants, 2024, 13(17), 2357. https://doi.org/10.3390/plants13172357
50. Yeon, J.; Nguyen, T.T.P.; Kim, M. et al. Prediction accuracy of genomic estimated breeding values for fruit traits in cultivated tomato (Solanum lycopersicum L.). BMC Plant Biol., 2024, 24, 222. https://doi.org/10.1186/s12870-024-04934-8
51. Cappetta, E.; Andolfo, G.; Di Matteo, A.; Barone, A.; Frusciante, L.; Ercolano, M.R. Accelerating tomato breeding by exploiting genomic selection approaches. Plants 9, 2020, 1236. doi: 10.3390/plants9091236.
52. Khan, T.; Adem I. An AI taste 'connoisseur' could be the future of crop breeding. Agriculture Dive, 2023, p. 1–3. https://www.agriculturedive.com/news/an-ai-taste-connoisseur-could-be-thefuture-of-crop-breeding/700086/
53. Ropelewska, E.; Piecko, J. Discrimination of tomato seeds belonging to different cultivars using machine learning. European Food Research and Technology, 2022, 248(3), 685-705.
54. Zhang, C.; Jiang, S.; Tian, Y.; Dong, X.; Xiao, J.; Lu, Y.; Xia, Z. Smart breeding driven by advances in sequencing technology. Modern Agriculture, 2023, 1(1), 43-56
55. Chowdhury, R.H.; Eti, F.S.; Ahmed, R.; Gupta, S.D.; Jhan, P.K.; Islam, T.; Bhuiyan, M.A.R.; Rubel, M.H.; Khayer, A. Drought-responsive genes in tomato: meta-analysis of gene expression using machine learning. Sci Rep., 2023, Nov 8;13(1):19374. doi: 10.1038/s41598-023-45942-2. PMID: 37938584; PMCID: PMC10632472.
56. Chele, K.; Tinte, M.; Piater, L.; Dubery, I.; Tugizimana, F. Soil Salinity, a Serious Environmental Issue and Plant Responses: A Metabolomics Perspective. Metabolites, 2021, 11. 724. 10.3390/metabo11110724.
57. Bupi, N.; Sangaraju, V.K.; Phan LT, Lal A, Vo TTB, Ho PT, Qureshi MA, Tabassum M, Lee S, Manavalan B. An Effective Integrated Machine Learning Framework for Identifying Severity of Tomato Yellow Leaf Curl Virus and Their Experimental Validation. Research (Wash D C). 2023, 6, 0016. doi: 10.34133/research.0016. Epub, 2023, Jan 10. PMID: 36930763; PMCID: PMC10013792.
58. Dias, F.; Valente, D.; Oliveira, C.; Dariva, F.; Copati, M. Remote sensing and machine learning techniques for high throughput phenotyping of late blight-resistant tomato plants in open field trials. International Journal of Remote Sensing, 2023, 44. 1900-1921. 10.1080/01431161.2023.2192878.
59. Gadade, H.; Kirange, D. Machine Learning Based Identification of Tomato Leaf Diseases at Various Stages of Development, 2021, 814-819. 10.1109/ICCMC51019.2021.9418263.
60. Tan, L.; Lu, J.; Jiang, H. Tomato Leaf Diseases Classification Based on Leaf Images: A Comparison between Classical Machine Learning and Deep Learning Methods. AgriEngineering, 2021, 3. 542-558. 10.3390/agriengineering3030035.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 February 2025 doi:10.20944/preprints202502.1208.v1

61. Johansen, K.; Morton, M.J.L.; Malbeteau, Y.; Aragon, B.; Al-Mashharawi, S.; Ziliani, M.G.; et al. Predicting biomass and yield in a tomato phenotyping experiment using UAV imagery and random forest. Front. Artif. Intell., 2020, 3. doi: 10.3389/frai.2020.00028
62. Bose, S.; Banerjee, S.; Kumar, S.; Saha, A.; Nandy, D.; Hazra, S. Review of applications of artificial intelligence (AI) methods in crop research. Journal of Applied Genetics, 2024, 65(2), 225-240.
63. Yamamoto, E.; Matsunaga, H.; Onogi, A.; Kajiya-Kanegae, H.; Minamikawa, M.; Suzuki, A.; Fukuoka, H. A simulation-based breeding design that uses whole-genome prediction in tomato. Scientific reports, 2016, 6(1), 19454.
64. Xu, Y.; Zhang, X.; Li, H.; Zheng, H.; Zhang, J.; Olsen, MS.; Varshney, RK.; Prasanna, B.M.; Qian, Q. Smart breeding driven by big data, artificial intelligence, and integrated genomic-enviromic prediction. Mol Plant, 2022, Nov 7;15(11):1664-1695. doi: 10.1016/j.molp.2022.09.001. Epub 2022 Sep 7. PMID: 36081348.
65. Ansari, R.; Manna, A.; Hazra, S.; Bose, S.; Chatterjee, A.; Sen, P. Breeding 4.0 vis-à-vis application of artificial intelligence (AI) in crop improvement: an overview. New Zealand Journal of Crop and Horticultural Science, 2024, 1-43.
66. Martina, M.; De Rosa, V.; Magon, G.; Acquadro, A.; Barchi, L.; Barcaccia, G.; De Paoli, E.; Vannozzi, A.; Portis, E. Revitalizing agriculture: next-generation genotyping and-omics technologies enabling molecular prediction of resilient traits in the Solanaceae family. Front. Plant Sci., 2024 , 15:1278760. doi: 10.3389/fpls.2024.1278760
67. Bannari, A.; Morin, D.; Bonn, F.; Huete, A.R. A review of vegetation indices. Remote Sensing Reviews, 1995, 13(1–2), 95–120. https://doi.org/10.1080/02757259509532298
68. Taşan, S.; Cemek, B.; Taşan, M.; Cantürk, A. Estimation of eggplant yield with machine learning methods using spectral vegetation indices. Computers and Electronics in Agriculture, 2022, 202, 107367. https://doi.org/10.1016/j.compag.2022.107367
69. Islam, S.; Samsuzzaman, Reza, M.N.; Lee, K.-H.; Ahmed, S.; Cho, Y.J.; Noh, D.H.; Chung, S.-O. Image Processing and Support Vector Machine (SVM) for Classifying Environmental Stress Symptoms of Pepper Seedlings Grown in a Plant Factory. Agronomy, 2024, 14, 2043. https://doi.org/10.3390/agronomy14092043
70. García-Fortea, E.; García-Pérez, A.; Gimeno-Páez, E.;, Sánchez-Gimeno. A.; Vilanova, S.; Prohens, J.; Pastor-Calle, D. A Deep Learning-Based System (Microscan) for the Identification of Pollen Development Stages and Its Application to Obtaining Doubled Haploid Lines in Eggplant. Biology (Basel), 2020, Sep 5;9(9):272. doi: 10.3390/biology9090272. PMID: 32899465; PMCID: PMC7564724.
71. Sun, C.; Feng, L.; Zhang, Z.; Ma, Y.; Crosby, T.; Naber, M.; Wang, Y. Prediction of End-Of-Season Tuber Yield and Tuber Set in Potatoes Using In-Season UAV-Based Hyperspectral Imagery and Machine Learning. Sensors, 2020, 20, 5293. https://doi.org/10.3390/s20185293
72. Nomura, K.; Kaneko, T.; Iwao, T.; Kitayama, M.; Goto, Y.; Kitano, M. Hybrid AI model for estimating the canopy photosynthesis of eggplants. Photosynthesis Research, 2023, 155(1), 77-92.
73. Kaniyassery, A.; Goyal, A.; Thorat, S.A.; Rao, M.R.; Chandrashekar, H.K.; Murali, T.S.; Muthusamy, A. Association of meteorological variables with leaf spot and fruit rot disease incidence in eggplant and YOLOv8-based disease classification, Ecological Informatics, 2024, 83, 102809, 1574-9541, https://doi.org/10.1016/j.ecoinf.2024.102809.
74. Lajom, M.P.; Remigio, J.P.; Arboleda, E.; Sacala, R.J.R. Design and Development of Eggplant Fruit and Shoot Borer (Leucinodes Orbonalis) Detector Using Near-Infrared Spectroscopy. Journal of Engineering and Sustainable Development, 2024, 28(4), 439-454. https://doi.org/10.31272/jeasd.28.4.3
75. Cemek, B.; Tasan, S.; Canturk, A. Machine learning techniques in estimation of eggplant crop evapotranspiration. Appl. Water Sci., 2023, 13, 136. https://doi.org/10.1007/s13201-023-01942-1

76. Prusty, Meenakhi.; Panda, Narayan.; Dash, A.; Mishra, N. Nutrient Management Modules for Eggplant (Solanum melongena L.): Yield, Quality, Economics, Nutrient Uptake and Post-Harvest Soil Properties. Journal of the Indian Society of Coastal Agricultural Research, 2022, 40. 46-53. 10.54894/JISCAR.40.1.2022.112565.
77. Yang, B.; Xu, Y. Applications of deep-learning approaches in horticultural research: a review. *Horticulture Research*, 2021, *8*.
78. Salvador, P.; Gómez, D.; Sanz, J.; Casanova, J.L. Estimation of Potato Yield Using Satellite Data at a Municipal Level: A Machine Learning Approach. ISPRS Int. J. Geo-Inf., 2020, 9, 343. https://doi.org/10.3390/ijgi9060343
79. Gómez, D.; Salvador, P.; Sanz, J.; Casanova, J.L. 2019. Potato Yield Prediction Using Machine Learning Techniques and Sentinel 2 Data. Remote Sens, 2019, 11, 1745. https://doi.org/10.3390/rs11151745
80. Gómez, D.; Salvador, P.; Sanz, J.; Casanova, J.L. New spectral indicator Potato Productivity Index based on Sentinel-2 data to improve potato yield prediction: a machine learning approach. Int. J. Remote Sens., 2021, 42, 3426–3444. https://doi.org/10.1080/01431161.2020.1871102
81. Kurek, J.; Niedbała, G.; Wojciechowski, T.; Świderski, B.; Antoniuk, I.; Piekutowska, M.; Kruk, M.; Bobran, K. Prediction of Potato (Solanum tuberosum L.) Yield Based on Machine Learning Methods. Agriculture, 2023, 13, 2259. https://doi.org/10.3390/agriculture13122259
82. El-Kenawy, E.-S.M.; Alhussan, A.A.; Khodadadi, N.; Mirjalili, S.; Eid, M.M. Predicting Potato Crop Yield with Machine Learning and Deep Learning for Sustainable Agriculture, 2024, Potato Res. https://doi.org/10.1007/s11540-024-09753-w
83. Li, D.; Miao, Y.; Gupta, S.K.; Rosen, C.J.; Yuan, F.; Wang, C.; Wang, L.; Huang, Y. Improving Potato Yield Prediction by Combining Cultivar Information and UAV Remote Sensing Data Using Machine Learning. Remote Sens. 2021, 13, 3322. https://doi.org/10.3390/rs13163322
84. Coulibali, Z.; Cambouris, A.N.; Parent, S.-É. Cultivar-specific nutritional status of potato (Solanum tuberosum L.) crops. PLOS ONE, 2020, 15, e0230458. https://doi.org/10.1371/journal.pone.0230458
85. Yu, T.; Zhou, J.; Fan, J.; Wang, Y.; Zhang, Z. Potato Leaf Area Index Estimation Using Multi-Sensor Unmanned Aerial Vehicle (UAV) Imagery and Machine Learning. Remote Sens., 2023, 15, 4108. https://doi.org/10.3390/rs15164108
86. Rahman, Md.A.; Khan, A.A.; Hasan, Md.M.; Rahman, Md.S.; Habib, Md.T. Deep Learning Modeling for Potato Breed Recognition. IEEE Trans. AgriFood Electron., 2024, 2, 419–427. https://doi.org/10.1109/TAFE.2024.3406544
87. Azizi, A.; Abbaspour-Gilandeh, Y.; Nooshyar, M.; Afkari-Sayah, A. Identifying Potato Varieties Using Machine Vision and Artificial Neural Networks. International Journal of Food Properties, 2015, 19(3), 618–635. https://doi.org/10.1080/10942912.2015.1038834
88. Gao, J.; Westergaard, J.C.; Sundmark, E.H.R.; Bagge, M.; Liljeroth, E.; Alexandersson, E. Automatic late blight lesion recognition and severity quantification based on field imagery of diverse potato genotypes by deep learning. Knowl.-Based Syst., 2021, 214, 106723. https://doi.org/10.1016/j.knosys.2020.106723
89. Gold, K.M.; Townsend, P.A.; Herrmann, I.; Gevens, A.J. 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
90. Aparajita, Sharma, R.; Singh, A.; Dutta, M.K.; Riha, K.; Kriz, P.Image processing based automated identification of late blight disease from leaf images of potato crops. 40th Int. Conf. Telecommun. Signal Process, 2017, TSP 758–762. https://doi.org/10.1109/TSP.2017.8076090

 doi:10.20944/preprints202502.1208.v1

91. Iqbal, Md.A.; Talukder, K.H. Detection of Potato Disease Using Image Segmentation and Machine Learning, in: 2020 International Conference on Wireless Communications Signal Processing and Networking (WiSPNET). Presented at the 2020 International Conference on Wireless Communications Signal Processing and Networking (WiSPNET), IEEE, Chennai, India, 2020, pp. 43–47. https://doi.org/10.1109/WiSPNET48689.2020.9198563
92. Sholihati, R.A.; Sulistijono, I.A.; Risnumawan, A.; Kusumawati, E. Potato Leaf Disease Classification Using Deep Learning Approach, in: 2020 International Electronics Symposium (IES). Presented at the 2020 International Electronics Symposium (IES), IEEE, Surabaya, Indonesia, 2020, pp. 392–397. https://doi.org/10.1109/IES50839.2020.9231784
93. Tarik, M.I.; Akter, S.; Mamun, A.A.; Sattar, A. Potato Disease Detection Using Machine Learning, in: 2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV). Presented at the 2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV), IEEE, Tirunelveli, India, 2021, pp. 800–803. https://doi.org/10.1109/ICICV50876.2021.9388606
94. Rashid, J.; Khan, I.; Ali, G.; Almotiri, S.H.; AlGhamdi, M.A.; Masood, K. Multi-Level Deep Learning Model for Potato Leaf Disease Recognition. Electronics, 2021, 10, 2064. https://doi.org/10.3390/electronics10172064
95. Pasalkar, J.; Gorde, G.; More, C.; Memane, S.; Gaikwad, V. Potato Leaf Disease Detection using Machine Learning. Curr. Agric. Res. J., 2023, 11, 949–954. https://dx.doi.org/10.12944/CARJ.11.3.23
96. Shin, M.-Y.; Gonzalez Viejo, C.; Tongson, E.; Wiechel, T.; W.J. Taylor, P.; Fuentes, S. Early detection of Verticillium wilt of potatoes using near-infrared spectroscopy and machine learning modeling. Comput. Electron. Agric., 2023, 204. https://doi.org/10.1016/j.compag.2022.107567
97. Bhere, H.; Jariwala, V.; Sharma, A.; Nemade, V. Potato Plant Leaf Disease Classification Using Deep CNN, in: Potato Plant Leaf Disease Classification Using Deep CNN. Presented at the 5th International Conference, ESDA 2022, Springer, Singapore, pp. 367–378. https://doi.org/10.1007/978-981-97-0154-4_29
98. Sugiura, R.; Tsuda, S.; Tsuji, H.; Murakami, N. Virus-Infected Plant Detection in Potato Seed Production Field by UAV Imagery</i>, in: 2018 Detroit, Michigan July 29 - August 1, 2018. Presented at the 2018 Detroit, Michigan July 29 - August 1, 2018, American Society of Agricultural and Biological Engineers. https://doi.org/10.13031/aim.201800594
99. Griffel, L.M.; Delparte, D.; Edwards, J. Using Support Vector Machines classification to differentiate spectral signatures of potato plants infected with Potato Virus Y. Comput. Electron. Agric., 2018, 153, 318–324. https://doi.org/10.1016/j.compag.2018.08.027
100. Boguszewska-Mańkowska, D.; Ruszczak, B.; Zarzyńska, K. Classification of Potato Varieties Drought Stress Tolerance Using Supervised Learning. Appl. Sci., 2022, 12, 1939. https://doi.org/10.3390/app12041939
101. Lapajne, J.; Vojnović, A.; Vončina, A.; Žibrat, U. Enhancing Water-Deficient Potato Plant Identification: Assessing Realistic Performance of Attention-Based Deep Neural Networks and Hyperspectral Imaging for Agricultural Applications. Plants, 2024, 13, 1918. https://doi.org/10.3390/plants13141918
102. Lozada, D.N.; Sandhu, K.S.; Bhatta, M. Ridge regression and deep learning models for genome-wide selection of complex traits in New Mexican Chile peppers. BMC Genomic Data, 2023, 24, 80. https://doi.org/10.1186/s12863-023-01179-6
103. Sabanci, K.; Fatih Aslan, M.; Ropelewska, E.; Fahri Unlersen, M. A convolutional neural network-based comparative study for pepper seed classification: Analysis of selected deep features with support vector machine. J. Food Process Eng., 2021, 45. https://doi.org/10.1111/jfpe.13955
104. Kurtulmuş, F.; Kavdir, I.; Alibaş, İ. Classification of pepper seeds using machine vision based on neural network. Int. J. Agric. Biol. Eng., 2016, 9, 51－62. https://doi.org/10.3965/j.ijabe.20160901.1790

105. Tu, K.; Li, L.; Yang, L.; Wang, J. Sun, Q. Selection for high quality pepper seeds by machine vision and classifiers. Journal of Integrative Agriculture, 2016, 17. 1999-2006. 10.1016/S2095-3119(18)62031-3.
106. Ramírez-Meraz, M.; Méndez-Aguilar, R.; Hidalgo-Martínez, D.; Villa-Ruano, N.; Zepeda-Vallejo, L.G.; Vallejo-Contreras, F.; Hernández-Guerrero, C.J.; Becerra-Martínez, E. Experimental races of Capsicum annuum cv. jalapeño: Chemical characterization and classification by 1H NMR/machine learning. Food Res. Int., 2020, 138, 109763. https://doi.org/10.1016/j.foodres.2020.109763
107. Dissanayake, A.; Rajapaksha, I.; Gunarathna, R.; Jayasinghe, S.; De Silva, H.; Hettiarachchi, S. Detection of Diseases and Nutrition in Bell Pepper. Presented at the 5th International Conference on Advancements in Computing (ICAC), IEEE, Colombo, Sri Lanka, 2023, pp. 286–291. https://doi.org/10.1109/ICAC60630.2023.10417573
108. Haque, I.; Islam, Md.A.; Roy, K.; Rahaman, Md.M.; Shohan, A.A.; Islam, Md.S. Classifying Pepper Disease based on Transfer Learning: A Deep Learning Approach. Presented at the 2022 International Conference on Applied Artificial Intelligence and Computing (ICAAIC), Salem, India, 2022, pp. 620–629. https://doi.org/10.1109/ICAAIC53929.2022.9793178
109. Fumia, N.; Kantar, M.; Lin, Y.; Schafleitner, R.; Lefebvre, V.; Paran, I.; Börner, A.; Diez, M.J.; Prohens, J.; Bovy, A.; Boyaci, F.; Pasev, G.; Tripodi, P.; Barchi, L.; Giuliano, G.; Barchenger, D.W. Exploration of high-throughput data for heat tolerance selection in Capsicum annuum. Plant Phenome J., 2023, 6, e20071. https://doi.org/10.1002/ppj2.20071
110. Ataş, M.; Yardimci, Y.; Temizel, A. A new approach to aflatoxin detection in chili pepper by machine vision. Comput. Electron. Agric., 2012, 87, 129–141. https://doi.org/10.1016/j.compag.2012.06.001
111. Fang, J. Breeding 5.0: AI-Driven Revolution in Designed Plant Breeding. Molecular Plant Breeding, 2024, 15.