

Machine Learning for Plant Breeding/Biotechnology

Subjects: Agriculture, Dairy & Animal Science

Contributor: Mohsen Niazian, Gniewko Niedbala

Classical univariate and multivariate statistics are the most common methods used for data analysis in plant breeding and biotechnology studies. Evaluation of genetic diversity, classification of plant genotypes, analysis of yield components, yield stability analysis, assessment of biotic and abiotic stresses, prediction of parental combinations in hybrid breeding programs, and analysis of in vitro-based biotechnological experiments are mainly performed by classical statistical methods. Despite successful applications, these classical statistical methods have low efficiency in analyzing data obtained from plant studies, as the genotype, environment, and their interaction ($G \times E$) result in nondeterministic and nonlinear nature of plant characteristics. Large-scale data flow, including phenomics, metabolomics, genomics, and big data, must be analyzed for efficient interpretation of results affected by $G \times E$. Nonlinear nonparametric machine learning techniques are more efficient than classical statistical models in handling large amounts of complex and nondeterministic information with "multiple-independent variables versus multiple-dependent variables" nature. Neural networks, partial least square regression, random forest, and support vector machines are some of the most fascinating machine learning models that have been widely applied to analyze nonlinear and complex data in both classical plant breeding and in vitro-based biotechnological studies. High interpretive power of machine learning algorithms has made them popular in the analysis of plant complex multifactorial characteristics. The classification of different plant genotypes with morphological and molecular markers, modeling and predicting important quantitative characteristics of plants, the interpretation of complex and nonlinear relationships of plant characteristics, and predicting and optimizing of in vitro breeding methods are the examples of applications of machine learning in conventional plant breeding and in vitro-based biotechnological studies. Precision agriculture is possible through accurate measurement of plant characteristics using imaging techniques and then efficient analysis of reliable extracted data using machine learning algorithms. Perfect interpretation of high-throughput phenotyping data is applicable through coupled machine learning-image processing.

This entry shows how nonlinear machine learning algorithms can be used in different branches of classical plant breeding and in vitro-based methods. An idea is provided at the end of the entry that shows how coupled image processing-machine learning (especially deep CNN) could be used to identify the ploidy level of plants. It could be used in laboratories without flowcytometry equipment and/or in plant species without an established chromosome counting protocol.

Keywords: artificial neural networks ; big data ; high-throughput phenotyping

1. Introduction

Due to climate change (global warming), increasing food requirements and depletion of resources in consequence of increasing global population, it is necessary to use modern technologies in agriculture and food sciences ^[1]. Plant breeding is a dynamic branch of agricultural science. It started with simple selection of impressive plants with superior characteristics. Later, genetics and statistics were involved in classical plant breeding, mainly after the discoveries of Gregor Mendel and Sir Ronald Aylmer Fisher. Next, modern plant breeding emerged with the advancements in genetic and biotechnology approaches. Classical plant breeding methods mainly included assessment and classification of genetic diversity, yield components analysis (indirect selection of superior genotypes with impressive economic characteristics), yield stability analysis (genotype \times environment interaction), enhanced tolerance to biotic and abiotic stresses, and hybrid breeding programs. In vitro-based biotechnological breeding methods mainly included in vitro micropropagation, doubled haploid production, artificial polyploidy induction, and *Agrobacterium*-mediated gene transformation. In in vitro micropropagation studies, researchers want to investigate the effects of influential factors (inputs), such as combination of culture medium components, combination and concentrations of plant growth regulators (PGRs), and interactions of plant genotype \times culture medium \times PGRs \times explant type \times explant age \times elicitor additives \times type and concentration of carbohydrate source \times etc., on regeneration efficiency (outputs) of their desired plants. Classical statistical techniques have been employed to analyze and interpret the results of both classical and in vitro-based plant breeding studies. These analytical techniques are mainly based on variance and linear regression models to assess the

relationship of variables and predict the effect of independent variables on dependent variables. One regression model is required to assess the effect of a group of independent variables ($X_1, X_2, X_3, \dots, X_n$) on one dependent variable (Y), according to the multiple linear relationships [2]. However, nonlinear and nondeterministic properties are inextricably linked with plant biological systems [3]. Therefore, despite of successful applications, the classical linear regression-based models are unable to interpret highly nonlinear and complex relationships between dependent and independent variables. Most of these plant breeding approaches are “multiple-independent variables versus multiple-dependent variables.” Under these conditions, one regression model is required for each output [4]. Powerful data mining tools are employed in plant breeding studies to predict and explain complex data.

Machine learning—the science of programming computers so they can learn from data—has been widely applied in both classical and in vitro-based plant breeding studies to interpret the flow of information about plants from the DNA sequence to the observed phenotypes. There are three ways to classify machine learning methods, including supervised and supervised models, linear and nonlinear algorithms, and shallow and deep learning models (Figure 1). Artificial neural networks (ANNs), deep neural networks (DNNs), convolutional neural networks (CNNs), random forest (RF), and support vector machines (SVMs) are examples of nonlinear nonparametric machine learning algorithms, applied for processing nonlinear data in plant studies [5]. These data-driven models are able to parse and interpret non-normal, nonlinear, and nondeterministic unpredictable data sets, through the full use of all spectral data and avoid irrelevant spectral bands and multicollinearity [6][7]. Among different learning algorithms, including supervised, unsupervised, reinforcement, sparse dictionary, and rule-based, supervised learning is more suitable and efficient for life science problems [8]. Supervised learning can be used for classification (predicting non-numeric answers) and regression (predicting numeric answers) [9]. Formless datasets such as data obtained by photo imaging or sequencing can be interpreted through machine learning algorithms [10]. Genome sequencing data can be used in machine learning models for the identification and classification of transposable elements [11]. By using machine learning algorithms, breeders are able to predict multiple outputs (multiple-dependent variables) through different combinations of multiple inputs in one model and reduce required analyses.

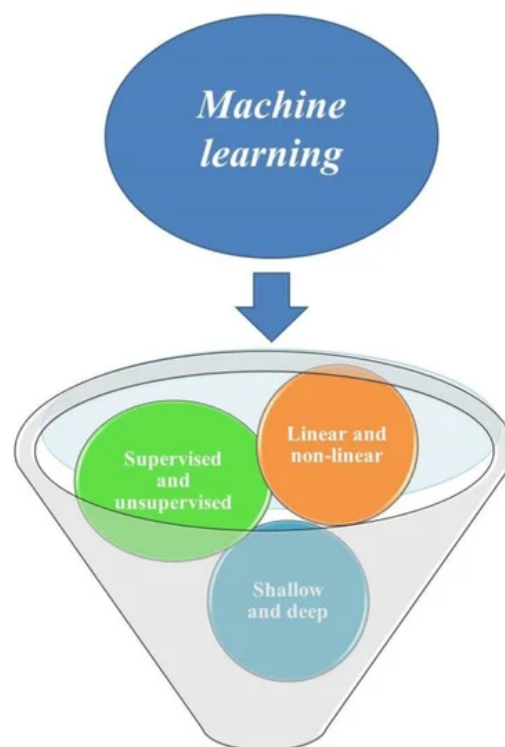


Figure 1. Different categories of machine learning algorithms.

Artificial neural networks, consist of an input, an output, and several hidden layers, are nonlinear nonparametric models which do not require a prior structure for data and detailed information about the physical processes to be modeled and to tolerate data loss [12][13]. Because of their more hidden layers, DNNs have greater predictive power than ANNs. Convolutional neural networks, as state-of-the-art deep learning architecture, are inspired by the natural visual perception mechanism of the living creatures and consist of convolutional, pooling, fully-connected layers, and an output layer [14]. CNNs are suitable for classification studies because of automatic feature extraction [9]. Image classification, object detection, object tracking, pose estimation, text detection and recognition, visual saliency detection, action recognition, scene labeling, speech, and natural language processing are some of the typical applications of CNNs [14]. Neural networks have low interpretability of the features (lack the interpretation capability), especially CNN in which the features

extracted are hidden. More advanced machine learning technique of SVMs, which uses a supervised learning algorithm to find both linear and nonlinear relationships in data, can be used for clustering, classification, and regression analysis of data sets. In comparison with multilayer perceptron (MLP) of ANN, SVM uses a large number of hidden units and has better performance in the formulation of the learning problem, subsequently quadratic optimization task [15]. Random forest regression is a regression tree-based machine learning that uses multiple decision trees to classify data and needs setting the number of trees, the number of random features, and the stop criteria for training. RF is more suitable for spectral data analysis and overfitting can be controlled through combining different independent predictors [16][17]. In semantic segmentation methods, such as automated phenotyping and plant disease detection, deep learning CNN can be more effective than shallow learning models of SVMs and RF and problem of required large manually crafted features can be solved by using image augmentation and small manually annotated empirical dataset for fine-tuning a synthetically bootstrapped CNN [18]. Through the integrating image feature extraction with classification in a single pipeline, deep convolutional neural networks have been considered as mainstream in biotic and abiotic stress diagnosis and classification [19]. A nine-layer deep CNN model was trained for identification of plant leaf diseases using data set with 39 different classes of plant leaves and background images and 96.46% classification accuracy was reported, which is greater than traditional machine learning approaches of SVM, decision tree, logistic regression, and K-NN [20]. CNNs are also applicable in remote sensing for object detection and pattern recognition. High accuracy (84%) for fine-grained mapping of vegetation species and communities using deep CNN-based segmentation, trained by data directly derived from visual interpretation of unmanned aerial vehicles (UAV)-based high-resolution Red-Green-Blue (RGB) imagery, has been reported [21].

A lot of training data is required in ANN for the optimization of sigmoid functions belonging to the hidden layer's neurons, as overfitting and local minima may happen by small number of training data. Therefore, the optimization process cannot be properly carried using back-propagation algorithms, when the number of training samples is small [8]. Through the short review on studies that used SVM and ANN techniques for identifying disease in plants, it was concluded that the ANN-based methods are better than SVM-based methods, as few samples and features are used in SVM-based methods to identify the disease-affected plants [22]. Conversely, in modeling in vitro culture of *Chrysanthemum* (*Dendranthema* × *grandiflorum*), better performance accuracy of SVR ($R^2 > 0.92$) than MLP ($R^2 > 0.82$) has been reported [15]. Applying different algorithm and comparing their performance is an appropriate solution to find the best algorithm in a particular data set. In tea plant (*Camellia sinensis* L.), partial least squares discriminative analysis (PLS-DA) and least squares-support vector machines (LS-SVM) were used for the classification of different nitrogen nutrition status under field condition and better performance with correct classification of LS-SVM than PLS-DA was reported [23].

Different application areas for nonlinear machine learning technologies in classical and in vitro-based plant breeding studies are shown in Figure 2. The following sections of the article provide a comprehensive review of the applications of these nonlinear machine learning techniques in classical and in vitro-based plant breeding studies.

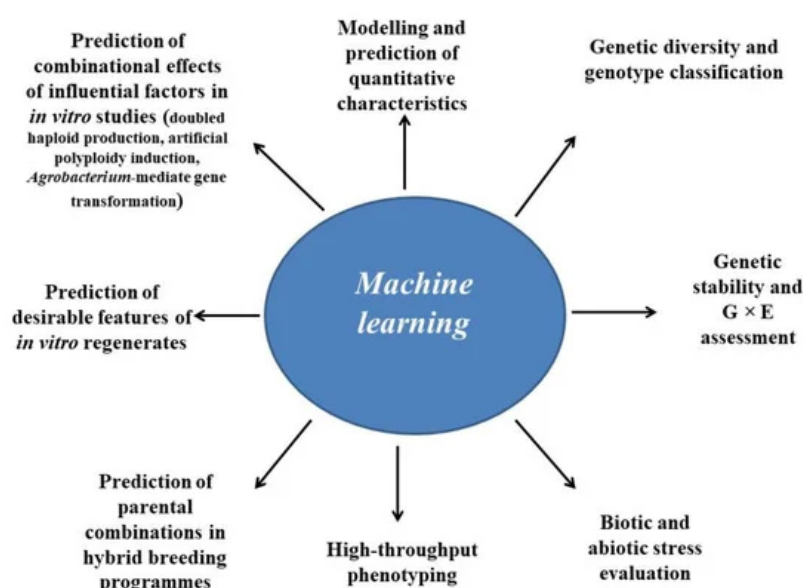


Figure 2. Potential applications of machine learning techniques in classical and modern plant breeding.

2. Applications of Machine Learning in In Vitro-Based Plant Biotechnology

Biotechnology-based breeding methods (BBBMs) complement classical breeding methods in rapid plant improvement. In vitro regeneration, as the main core of many in-vitro-based breeding methods, has numerous plant breeding applications. In situ and ex situ conservation and micropropagation (proliferation) are direct applications of in vitro regeneration [24]. In endangered rare plant species, like medicinal plants, in vitro culture is an effective strategy for mass propagation, germplasm conservation, and production of bioactive compounds [25]. Several factors determine the fate of cultured cells in in vitro regeneration of plants. These are the plant genotype, plant growth regulators (PGRs), culture medium components, explant type, explant age, enhancer additives-elicitors, etc. [24]. These factors can be divided into three main categories: initial triggers of regeneration (environmental signal inputs and physical stimuli), epigenetic and transcriptional cellular responses to the initial triggers, and molecules that manage the formation and development of the new stem cell niche [26]. The combination and interactions between these factors lead to multifactorial nature of the in vitro plant regeneration process. Basal culture medium components, plant genotype, PGRs, explant type, and explant age are all multilevel factors with different applicable combinations. The inclusion of other factors results in a very complex situation for interpretation. Plant cells and tissues have nondeterministic and nonlinear developmental patterns in a stressful in vitro environment [27]. The analysis of variance of factorial experiments and simple means comparison analysis with classical methods such as LSD, Tukey's HSD, and Duncan's test, are the main statistical methods used to interpret the effects of interaction between effective factors in most in vitro regeneration studies [25][28][29].

Murashige and Skoog (MS), modified MS (MMS), Gamborg's B5 medium Woody Plant Medium (WPM), and Driver and Kuniyuki Woody Plant Medium (DKW) are the most commonly used basal culture media in in vitro regeneration studies. Basal medium manipulation is a promoting strategy that has been applied to increase the output of in vitro studies [30]. However, due to the large number of micro- and macroelements in the culture medium, it is difficult to manipulate their concentrations. In this situation, prediction of the effect of culture media components on the target characteristics of in vitro regenerants is the right solution. Artificial neural networks have been applied in these experiments to predict the best culture media components for efficient propagation of different plant species [31][32][33].

Different combinations of auxin and cytokinin PGRs can determine the developmental fate of cultured cells and tissues toward organogenesis and/or somatic embryogenesis. The cytokinin/auxin ratio is also very important in in vitro studies [34]. Niazi et al. [29] found that 2,4-dichlorophenoxyacetic acid (2,4-D) combined with kinetin resulted in indirect somatic embryogenesis of cultured hypocotyl segments of ajowan medicinal plants, whereas a combination of 3-methoxy-(6-benzylamino-9-tetrahydropyran-2-yl) purine and naphthalene acetic acid led cultivated explants toward an indirect shoot regeneration pathway. Arab et al. [35] combined artificial neural networks and genetic algorithms to predict and optimize the effect of cytokinin–auxin plant hormone (BAP, KIN, TDZ, IBA, and NAA) combinations and concentrations on the number of microshoots per explant, the length of microshoots, developed callus weight, and the quality index of plantlets in in vitro proliferation of Garnem (G × N15) rootstock. The ANN model predicted the number and length of microshoots with high accuracy. The highest values of the variable sensitivity ratio for the proliferation rate were related to the BAP (19.3), KIN (9.64), and IBA (2.63) inputs. An MLP-ANN was developed to predict the physical properties of embryogenic callus and the number of somatic embryos in in vitro regeneration of ajowan under the effect of different combinations of the explant age, concentrations of 2,4-D, kinetin, and sucrose inputs [36]. The ANN model predicted the physical properties of embryogenic callus (area, perimeter, Feret diameter, roundness, and true density) and the number of somatic embryos better than the multiple linear regressions. Fifteen-day-old hypocotyl explants × 1.5 mg/L 2,4-D × 0.5 mg/L Kin × 2.5% (w/v) sucrose was the best combination of inputs with the highest measured and predicted number of somatic embryos [36].

Apart from culture medium components and PGRs combination, ANN has been applied to model the sterilization step of in vitro regeneration. Hesami et al. [37] applied an MLP-ANN along with a genetic algorithm to model and optimize the contamination frequency and explant viability under the influence of seven input variables, i.e., HgCl₂, Ca(ClO)₂, nanosilver, H₂O₂, NaOCl, AgNO₃, and immersion times, in an in vitro culture of chrysanthemum. The lowest contamination frequency (0%) and the highest explant viability (99.98%) resulted from 1.62% NaOCl at 13.96 min immersion time. The sensitivity analysis of the ANN showed that the immersion time was the most important variable affecting the contamination frequency and explant viability [37]. ANNs are also used to simulate in vitro growth of plant tissue cultures, distinguish embryos from nonembryos, predict the formation of plantlets from embryos, estimate the biomass of cell cultures, simulate the distribution of temperature in a culture vessel, identify and estimate the in vitro induced shoot length, and cluster in vitro regenerated plantlets [27].

Other in vitro-based breeding methods, such as artificial polyploidy induction, doubled haploid production, plant gene transformation, and genome editing methods also have multifactorial nature and require multivariate statistical methods to interpret the results. Different chemical enhancers can be used in in vitro doubled haploid production methods (induced

parthenogenesis and androgenesis) to improve the haploid induction efficiency, e.g., PGRs, osmoprotectants, cellular antioxidants, reactive oxygen species scavengers, polyamins, stress hormones, chlormequat chloride, compatible solutes, DNA demethylating agents, histone deacetylase inhibitors, cell wall remodeling agents, ethylene inhibitors, and other applicable additives. They enhance tolerance to inductive stresses and improve the final efficiency of doubled haploid production [38]. ANN models may improve the efficiency of in vitro doubled haploid production and solve the problem of recalcitrant species/genotypes by predicting the best combination(s) of these additives in interaction with other influencing factors, such as the plant genotype, the surrounding environment of donor plants, physical treatments (inductive stresses) of cultured gametophytic cells, the developmental stage of initial gametophytic cells, and culture medium components. The ANN predicted the callus induction percentage in androgenesis (anther culture) of tomato (*Lycopersicon esculentum* L.) under the influence of plant genotype, the concentrations of 2,4-D and kinetin PGRs, and the concentration of gum Arabic better than the MLR model [39].

Plants' vigor and performance are commonly enhanced by mitotic-induced polyploidy. It consists in in vivo and in vitro application of mitotic spindle poisons [40]. In vitro-induced polyploidy is a multifactorial procedure. The efficiency of in vitro-induced polyploidy may be affected not only by in vitro regeneration parameters (basal culture medium components, combination of PGRs, additives, etc.) but also by the plant genotype, the developmental stage of initial explants as well as the type, dosage, and duration (exposure time) of the application of the antimitotic agent. Due to the genotype dependency, different genotypes of plant species exhibit different responses to concentrations of the antimitotic agent applied [41]. This results in significant interaction of the plant genotype and antimitotic agent in artificial polyploidy induction. Although there have been no reports on the application of ANN to model and predict the results of in vitro-induced artificial polyploidy, it might increase the efficiency by predicting and finding the best combination and interaction of all influential factors.

Agrobacterium-mediated gene transformation is a well-known method of plant gene transformation and genetic engineering. However, various parameters must be optimized for an efficient gene delivery, including the *Agrobacterium* strain cell density, the time of inoculation, the type and concentration of antibiotics to kill *Agrobacterium*, the type and concentration of selectable antibiotics, and the concentration of acetosyringone [42]. These influencing factors along with in vitro regeneration factors result in a multi-variable nature of *Agrobacterium*-mediated gene transformation [24]. It is obvious that machine learning algorithms could be used to predict and optimize *Agrobacterium*-mediated gene transformation, especially in important *Agrobacterium*-recalcitrant plant species.

3. Conclusions

Most classical statistical methods use only simple statistics and few influential factors to assess the biological features of plants. For example, Y_p and Y_s are the only indices used to identify drought-tolerant plant genotypes in yield-based drought tolerance assessment methods. However, there are other influential factors, such as cellular, physiological, and phytochemical pathways, which are involved in plants' responses to environmental stress. The tolerance of different plant species to biotic and abiotic stresses, as complex biological processes, can be efficiently enhanced through large-scale analysis of phenomic, metabolomic, and genomic data. Machine learning models are capable of processing large amounts of data (imaging and remote-sensing data) for high-throughput stress phenotyping. The analysis of different omics and phenomic data may result in more precise interpretation of GEI and yield stability. Plants' qualitative and quantitative characteristics can be predicted more precisely by analysis of climate data (temperature, humidity, sunshine, precipitation, etc.), soil factors, agricultural operations data (harvest date, information on diseases, crop status, ground temperature, etc.), topographic, and meteorological data. Big data analysis enables more efficient classification of plants' phenotypes and genotypes. Machine learning techniques are able to manage large amounts of data in various areas of plant breeding, which can lead to more accurate results and better interoperation than classical statistical methods. Artificial neural networks can be used for pattern recognition, nonlinear regression, and classification purposes in plant tissue culture studies because they can handle binary, continuous, categorical, and fuzzy datasets. The present review can give an overview of applications of machine learning to plant breeders. It would be helpful to adopt the correct method of data analysis in future studies, which in turn can increase the output of studies.

References

1. Ali, S.; Shafique, O.; Mahmood, T.; Hanif, M.A.; Ahmed, I.; Khan, B.A. A Review about Perspectives of Nanotechnology in Agriculture. *Pakistan J. Agric. Res.* 2018, 31, 116–121.
2. Niazian, M.; Sadat-Noori, S.A.; Abdipour, M. Modeling the seed yield of Ajowan (*Trachyspermum ammi* L.) using artificial neural network and multiple linear regression models. *Ind. Crops Prod.* 2018, 117, 224–234.

3. Hesami, M.; Naderi, R.; Tohidfar, M.; Yoosefzadeh-Najafabadi, M. Application of Adaptive Neuro-Fuzzy Inference System-Non-dominated Sorting Genetic Algorithm-II (ANFIS-NSGAI) for Modeling and Optimizing Somatic Embryogenesis of Chrysanthemum. *Front. Plant Sci.* 2019, 10, doi:10.3389/fpls.2019.00869.
4. Chegini, G.R.; Khazaei, J.; Ghobadian, B.; Goudarzi, A.M. Prediction of process and product parameters in an orange juice spray dryer using artificial neural networks. *J. Food Eng.* 2008, 84, 534–543.
5. 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.* 2018, 10, 2026.
6. Hesami, M.; Naderi, R.; Yoosefzadeh-Najafabadi, M.; Rahmati, M. Data-driven modeling in plant tissue culture. *J. Appl. Environ. Biol. Sci.* 2017, 7, 37–44.
7. 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, doi: 10.3389/fpls.2020.01148.
8. Asefpour Vakilian, K. Machine learning improves our knowledge about miRNA functions towards plant abiotic stresses. *Sci. Rep.* 2020, 10, 3041.
9. Wang, H.; Cimen, E.; Singh, N.; Buckler, E. Deep learning for plant genomics and crop improvement. *Curr. Opin. Plant Biol.* 2020, 54, 34–41.
10. Hu, H.; Scheben, A.; Edwards, D. Advances in Integrating Genomics and Bioinformatics in the Plant Breeding Pipeline. *Agriculture* 2018, 8, 75.
11. 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, 3837.
12. Alvarez, R. Predicting average regional yield and production of wheat in the Argentine Pampas by an artificial neural network approach. *Eur. J. Agron.* 2009, 30, 70–77.
13. Azevedo, A.M.; Andrade Júnior, V.C.D.; Pedrosa, C.E.; Oliveira, C.M.D.; Dornas, M.F.S.; Cruz, C.D.; Valadares, N.R. Application of artificial neural networks in indirect selection: A case study on the breeding of lettuce. *Bragantia* 2015, 74, 387–393.
14. Gu, J.; Wang, Z.; Kuen, J.; Ma, L.; Shahroudy, A.; Shuai, B.; Liu, T.; Wang, X.; Wang, G.; Cai, J.; et al. Recent advances in convolutional neural networks. *Pattern Recognit.* 2018, 77, 354–377.
15. Hesami, M.; Naderi, R.; Tohidfar, M.; Yoosefzadeh-Najafabadi, M. Development of support vector machine-based model and comparative analysis with artificial neural network for modeling the plant tissue culture procedures: Effect of plant growth regulators on somatic embryogenesis of chrysanthemum, as a case study. *Plant Methods* 2020, 16, 112.
16. Ali, A.M.; Darvishzadeh, R.; Skidmore, A.; Gara, T.W.; 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* 2020, 1–15, doi:10.1080/17538947.2020.1794064.
17. 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.
18. Barth, R.; IJsselmuiden, J.; Hemming, J.; Van Henten, E.J. Synthetic bootstrapping of convolutional neural networks for semantic plant part segmentation. *Comput. Electron. Agric.* 2019, 161, 291–304.
19. An, J.; Li, W.; Li, M.; Cui, S.; Yue, H. Identification and Classification of Maize Drought Stress Using Deep Convolutional Neural Network. *Symmetry* 2019, 11, 256.
20. Geetharamani, G.; Arun Pandian, J. Identification of plant leaf diseases using a nine-layer deep convolutional neural network. *Comput. Electr. Eng.* 2019, 76, 323–338.
21. Kattenborn, T.; Eichel, J.; Wiser, S.; Burrows, L.; Fassnacht, F.E.; Schmidtlein, S. Convolutional Neural Networks accurately predict cover fractions of plant species and communities in Unmanned Aerial Vehicle imagery. *Remote Sens. Ecol. Conserv.* 2020, doi:10.1002/rse2.146.
22. Iniyar, S.; Jebakumar, R.; Mangalraj, P.; Mohit, M.; Nanda, A. Plant Disease Identification and Detection Using Support Vector Machines and Artificial Neural Networks. In *Artificial Intelligence and Evolutionary Computations in Engineering Systems; Advances in Intelligent Systems and Computing*; Dash, S., Lakshmi, C., Das, S., Panigrahi, B., Eds.; Springer: Singapore, 2020; pp. 15–27.
23. Wang, Y.; Li, T.; Jin, G.; Wei, Y.; Li, L.; Kalkhajeh, Y.K.; Ning, J.; Zhang, Z. Qualitative and quantitative diagnosis of nitrogen nutrition of tea plants under field condition using hyperspectral imaging coupled with chemometrics. *J. Sci. Food Agric.* 2020, 100, 161–167.

24. Niazian, M. Application of genetics and biotechnology for improving medicinal plants. *Planta* 2019, 249, 953–973.
25. Ayuso, M.; García-Pérez, P.; Ramil-Rego, P.; Gallego, P.P.; Barreal, M.E. In vitro culture of the endangered plant *Eryngium viviparum* as dual strategy for its ex situ conservation and source of bioactive compounds. *Plant Cell Tissue Organ Cult.* 2019, 138, 427–435.
26. Sugimoto, K.; Temman, H.; Kadokura, S.; Matsunaga, S. To regenerate or not to regenerate: Factors that drive plant regeneration. *Curr. Opin. Plant Biol.* 2019, 47, 138–150.
27. Prasad, V.S.S.; Gupta, S.D. Applications and potentials of artificial neural networks in plant tissue culture. In *Plant Tissue Culture Engineering*; Gupta, S.D., Ibaraki, Y., Eds.; Focus on Biotechnology; Springer: Dordrecht, The Netherlands, 2008; Volume 6, ISBN 978-1-4020-3594-4.
28. Niazian, M.; Noori, S.A.S.; Galuszka, P.; Tohidfar, M.; Mortazavian, S.M.M. Genetic stability of regenerated plants via indirect somatic embryogenesis and indirect shoot regeneration of *Carum copticum* L. *Ind. Crops Prod.* 2017, 97, 330–337.
29. Orłowska, R.; Bednarek, P.T. Precise evaluation of tissue culture-induced variation during optimisation of in vitro regeneration regime in barley. *Plant Mol. Biol.* 2020, 103, 33–50.
30. Phillips, G.C.; Garda, M. Plant tissue culture media and practices: An overview. *Vitr. Cell. Dev. Biol. Plant* 2019, 55, 242–257.
31. Alanagh, E.N.; Garoosi, G.; Haddad, R.; Maleki, S.; Landín, M.; Gallego, P.P. Design of tissue culture media for efficient *Prunus* rootstock micropropagation using artificial intelligence models. *Plant Cell Tissue Organ Cult.* 2014, 117, 349–359.
32. Arab, M.M.; Yadollahi, A.; Shojaeiyan, A.; Ahmadi, H. Artificial Neural Network Genetic Algorithm as Powerful Tool to Predict and Optimize In vitro Proliferation Mineral Medium for G × N15 Rootstock. *Front. Plant Sci.* 2016, 7, doi:10.3389/fpls.2016.01526.
33. Arab, M.M.; Yadollahi, A.; Eftekhari, M.; Ahmadi, H.; Akbari, M.; Khorami, S.S. Modeling and Optimizing a New Culture Medium for In Vitro Rooting of G×N15 *Prunus* Rootstock using Artificial Neural Network-Genetic Algorithm. *Sci. Rep.* 2018, 8, 9977.
34. Hesami, M.; Naderi, R.; Tohidfar, M. Modeling and Optimizing Medium Composition for Shoot Regeneration of *Chrysanthemum* via Radial Basis Function-Non-dominated Sorting Genetic Algorithm-II (RBF-NSGAI). *Sci. Rep.* 2019, 9, 18237.
35. Arab, M.M.; Yadollahi, A.; Ahmadi, H.; Eftekhari, M.; Maleki, M. Mathematical Modeling and Optimizing of in Vitro Hormonal Combination for G × N15 Vegetative Rootstock Proliferation Using Artificial Neural Network-Genetic Algorithm (ANN-GA). *Front. Plant Sci.* 2017, 8, doi:10.3389/fpls.2017.01853.
36. Niazian, M.; Sadat-Noori, S.A.; Abdipour, M.; Tohidfar, M.; Mortazavian, S.M.M. 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). *Vitr. Cell. Dev. Biol. Plant* 2018, 54, 54–68.
37. Hesami, M.; Naderi, R.; Tohidfar, M. Modeling and Optimizing in vitro Sterilization of *Chrysanthemum* via Multilayer Perceptron-Non-dominated Sorting Genetic Algorithm-II (MLP-NSGAI). *Front. Plant Sci.* 2019, 10, doi:10.3389/fpls.2019.00282.
38. Niazian, M.; Shariatpanahi, M.E. In vitro-based doubled haploid production: Recent improvements. *Euphytica* 2020, 216, 69.
39. Niazian, M.; Shariatpanahi, M.E.; Abdipour, M.; Oroojloo, M. Modeling callus induction and regeneration in an anther culture of tomato (*Lycopersicon esculentum* L.) using image processing and artificial neural network method. *Protoplasma* 2019, 256, 1317–1332.
40. Sadat Noori, S.A.; Norouzi, M.; Karimzadeh, G.; Shirkool, K.; Niazian, M. Effect of colchicine-induced polyploidy on morphological characteristics and essential oil composition of ajowan (*Trachyspermum ammi* L.). *Plant Cell Tissue Organ Cult.* 2017, 130, 543–551.
41. Castillo, A.M.; Cistué, L.; Vallés, M.P.; Soriano, M. Chromosome Doubling in Monocots. In *Advances in Haploid Production in Higher Plants*; Springer: Dordrecht, The Netherlands, 2009; pp. 329–338.
42. Niazian, M.; Sadat Noori, S.A.; Galuszka, P.; Mortazavian, S.M.M. Tissue culture-based *Agrobacterium*-mediated and in planta transformation methods. *Czech J. Genet. Plant Breed.* 2017, 53, 133–143.

