

The pipelines of deep learning-based plant image processing

Kaiyue Hong¹, Yun Zhou²  and Han Han¹ 

Review

Cite this article: K. Hong et al. The pipelines of deep learning-based plant image processing. *Quantitative Plant Biology*, 6:e23, 1–14 <https://dx.doi.org/10.1017/qpb.2025.10018>

Received: 20 July 2024
Revised: 28 May 2025
Accepted: 18 June 2025

Keywords:

data science; deep learning; feature extraction; image recognition; machine learning.

Corresponding authors:

Han Han;
Email: hhan@njfu.edu.cn
Yun Zhou;
Email: zhouyun@purdue.edu

Associate Editor: Dr. Ross Sozzani

¹Co-Innovation Center for Sustainable Forestry in Southern China, College of Life Sciences, Nanjing Forestry University, Nanjing, China; ²Department of Botany and Plant Pathology, Center for Plant Biology, Purdue University, West Lafayette, IN, USA

Abstract

Recent advancements in data science and artificial intelligence have significantly transformed plant sciences, particularly through the integration of image recognition and deep learning technologies. These innovations have profoundly impacted various aspects of plant research, including species identification, disease detection, cellular signaling analysis, and growth monitoring. This review summarizes the latest computational tools and methodologies used in these areas. We emphasize the importance of data acquisition and preprocessing, discussing techniques such as high-resolution imaging and unmanned aerial vehicle (UAV) photography, along with image enhancement methods like cropping and scaling. Additionally, we review feature extraction techniques like colour histograms and texture analysis, which are essential for plant identification and health assessment. Finally, we discuss emerging trends, challenges, and future directions, offering insights into the applications of these technologies in advancing plant science research and practical implementations.

1. Introduction

In the digital age, large-scale plant image datasets are essential for advancing plant science, yet their efficient processing remains challenging; artificial intelligence (AI) and deep learning (DL) offer transformative solutions by enabling machines to simulate human intelligence in tasks like image recognition and decision-making (Williamson et al., 2023). In plant research, machine learning (ML), a subset of AI, enables automatic plant image analysis, allowing computers to learn and improve without explicit programming by identifying patterns in data to predict outcomes and make decisions through supervised, unsupervised, and reinforcement learning approaches (Silva et al., 2019). DL, a specialized branch of ML, uses multi-layer neural networks to process complex plant image data, automatically extract features, and perform tasks like classification and prediction, driving significant advancements in plant image analysis, especially in plant growth monitoring and disease detection (Saleem et al., 2019). Together, AI, ML, and DL propel innovation in plant science – ML enables learning from data, while DL leverages deep neural networks for advanced image analysis – driving transformative progress in plant research.

Processing and analyzing high-resolution plant images pose challenges for image processing algorithms due to plant diversity in colour, shape, and size, with additional complications from complex backgrounds and dense leaf structures affecting segmentation and feature extraction (Sachar & Kumar, 2021). To tackle these challenges, tailored methods in preprocessing, feature extraction, and data augmentation have been developed, showing strong effectiveness in plant image processing (Barbedo, 2016). For example, data augmentation methods like random rotation and flipping improve model adaptability to plant diversity by helping it learn more robust features (Cap et al., 2020). Furthermore, targeted approaches like colour normalization and background suppression improve feature recognition accuracy, reduce external interference, and highlight plants' distinct visual characteristics, optimizing workflows and enhancing plant image analysis accuracy and efficiency (Petrellis, 2019).

Challenges like data acquisition and the lack of high-quality annotated data hinder the widespread adoption of DL technologies in plant science. Image recognition, the first and

© The Author(s), 2025. Published by Cambridge University Press in association with John Innes Centre. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.


John Innes Centre
Unlocking Nature's Diversity

CAMBRIDGE
UNIVERSITY PRESS

crucial step in plant image processing, has been significantly advanced by the rapid development of DL, particularly Convolutional Neural Networks (CNNs) (Cai et al., 2023). CNNs are a type of feedforward neural network and a representative algorithm of DL, using convolutional calculations and possessing a deep structure (Kuo et al., 2019). The performance of CNNs in plant species recognition has been thoroughly evaluated on several large public wood image datasets, consistently demonstrating high accuracy. For example, CNN models achieved 97.3% accuracy on the Brazilian wood image database (Universidade Federal do Paraná, UFPR) and 96.4% on the Xylarium Digital Database (XDD), clearly outperforming traditional feature engineering methods. CNNs are both effective and generalizable for wood image recognition tasks (Hwang & Sugiyama, 2021).

Large language models (LLMs), such as ChatGPT, are advanced DL models that, when combined with domain-specific tools like the Agronomic Nucleotide Transformer (AgroNT) – a novel DNA-focused LLM – have demonstrated great potential in plant genetics and stress response studies (Mendoza-Revilla et al., 2024). By analyzing the genomes of 48 crop species and processing over 10 million cassava mutations, the LLM-based tools offer valuable insights into plant development, interactions, and traits, advancing gene expression profiling and opening new research possibilities (Agathokleous et al., 2024). Notably, LLMs have revealed new insights by uncovering non-obvious regulatory patterns in promoter regions, predicting the functional impacts of non-coding variants, and suggesting novel gene-stress associations that were previously unrecognized using traditional bioinformatics approaches (Mendoza-Revilla et al., 2024). For example, AgroNT has been shown to predict transcription factor binding affinities across diverse plant species with unprecedented accuracy, enabling the discovery of conserved stress-responsive elements in divergent genomes (Wang et al., 2025). Although still in its early stages, the application of language models in plant biology holds great potential to transform the field, despite currently lagging behind advancements in other domains.

This review evaluates key technologies in plant image processing, such as data acquisition, preprocessing, feature extraction,

and model training, examining their effectiveness, limitations, and potential to advance plant science research. It also compares various methodologies and ML models, highlighting their advantages, limitations, and challenges, providing a detailed framework to help researchers make informed decisions in plant image processing studies and applications.

2. Data acquisition and preprocessing

2.1. Data acquisition and plant feature extraction

Data acquisition and preprocessing are vital for ML in image processing, with high-resolution imaging, unmanned aerial vehicle (UAV) photography, and 3D scanning providing detailed morphological data for DL model foundation (Shahi et al., 2022) (Table 1). While high-resolution devices offer superior quality, regular cameras and smartphones provide greater accessibility and scalability, enabling large-scale data collection and enhancing dataset diversity and model robustness.

Feature extraction in plant image analysis integrates morphology, physiology, genetics, and ecology, starting with colour features (e.g., histograms, coherence vectors) and followed by morphological features (e.g., area, perimeter, shape descriptors) to identify plant traits (Mahajan et al., 2021). Texture features, capturing local variations in images, are crucial for species differentiation and disease detection, revealing surface structures like roughness and contrast (Mohan & Peeples, 2024). CNNs have proven effective in managing complex plant images, enhancing the classification and detection of diseases through robust feature extraction methods (Ahmad et al., 2022). Additionally, structural features, such as leaf morphology and spatial arrangements, are extracted using techniques like edge detection and shape description (Shoaib et al., 2023). Lastly, physiological features, including leaf count, size, and vein structure, provide valuable data on plant health and growth dynamics (Bühler et al., 2015). These features can be obtained manually or automatically through image processing, with recent studies favoring automated methods like segmentation and morphological analysis for high-throughput, objective phenotyping.

Table 1. Data acquisition techniques and their applications in plant sciences

Data acquisition techniques	Description	Type of camera	Type of platform	Type of applications	References
High-resolution imaging	Using high-resolution cameras to capture detailed images of plants.	Visual	Indoor	Plant growth dynamics, plant diseases diagnosis	Duncan et al. (2022)
UAV photography	Drone aerial photography provides detailed insights into plant community distribution and condition.	Visual	UAV	Monitoring crop growth in fields, assessing vegetation coverages, and observing environmental changes.	Wu et al. (2022)
3D scanning technology	Capturing plant spatial structure allows the creation of detailed 3D models of plant morphology.	Structured Light Scanner	Indoor	Plant morphology and growth	Nguyen et al. (2016)
Light detection and ranging (LiDAR)	Utilizing LiDAR to capture detailed spatial and structural data of plants.	LiDAR	Outdoor	Investigating plant morphology and analyzing their growth patterns.	Forero et al. (2022)
Spectral imaging technology	Capturing plant images in specific wavelengths to gather critical information about the health status of plants.	Multispectral	Indoor & Outdoor	Analyzing the photosynthetic efficiency, water content, and nutritional status of plants.	Zhang et al. (2022)
Public databases and resources	Collection of large-scale image datasets from various platforms that are accessible for research on plant species, health, and environmental impact.	Multi-platform		Expanding research datasets for species recognition and ecological studies.	Mano et al. (2009)

A key advantage of CNNs is their ability to learn hierarchical features from raw images, eliminating manual feature engineering and enhancing model adaptability and performance in diverse plant phenotyping tasks.

2.2. Preprocessing techniques

Data preprocessing in plant image analysis includes key steps like cropping, resizing, enhancing, augmenting, and annotating to optimize images for ML models. Cropping and resizing standardize dimensions, enhancing computational efficiency and reducing model complexity (Maraveas, 2024). Data augmentation modifies original images to generate new datasets, with techniques like contrast adjustment, denoising, and sharpening enhancing detail visibility and accuracy (Abebe et al., 2023), while augmentation strategies like rotation and flipping diversify the dataset to prevent overfitting and improve model generalization (Syarovy et al., 2024). Despite their benefits, preprocessing steps can cause information loss, requiring a balance between simplifying the model and preserving critical information. While most preprocessing is not labor-intensive, annotating and labeling training data remains highly labor-intensive, often becoming bottlenecks that hinder project progress. Accurate annotation, often referred to as 'ground truth', is essential for supervised learning, as it provides a reliable benchmark for model training and evaluation. In both research and practical applications – such as image recognition, natural language processing, and predictive analytics – the quality of labelled data directly influences model accuracy and reliability (Zhou et al., 2018a, 2018b).

To achieve optimal results, recommended sizes of datasets vary by task complexity. For binary classification, 1,000 to 2,000 images per class are typically sufficient (Singh et al., 2020). Multi-class classification requires 500 to 1,000 images per class, with higher requirements as the number of classes increases (Mühlenstädt & Frtunikj, 2024). More complex tasks, such as object detection, demand larger datasets, often up to 5,000 images per object (Cai et al., 2022). DL models like CNNs generally need 10,000 to 50,000 images, with larger models requiring 100,000+ images (Greydanus & Kobak, 2020). Data augmentation can multiply dataset size by 2–5 times (Shorten & Khoshgoftaar, 2019). Additionally, Transfer Learning, a machine-learning model, is effective for smaller datasets, requiring as few as 100 to 200 images per class for successful training (Zhu et al., 2021).

2.3. Commonly used public dataset

The Plant Village dataset is a widely used public resource for DL-based plant disease diagnosis research (Mohameth et al., 2020). It serves as a valuable tool in agricultural and plant disease research, offering a comprehensive collection of labelled images essential for developing and testing ML models for plant health monitoring (Pandey et al., 2024). Its accessibility, diversity, and standardized format make it a benchmark for algorithm development in precision agriculture, contributing to early disease detection and yield management, and addressing global challenges like food security and sustainable farming (Majji & Kumaravelan, 2021). Promoting the use of such datasets can enhance collaboration among researchers, standardize methodologies, and support scalable solutions across various agricultural environments (Ahmad et al., 2021).

Similar to the plant village dataset, other plant image datasets include the plant doc dataset, which contains images from various

plant species for plant disease diagnosis (Singh et al., 2020). The crop disease dataset features images of diseases in multiple crops, making it suitable for training DL models, especially for crop disease classification (Yuan et al., 2022). The tomato leaf disease dataset focuses on disease images specific to tomato leaves, supporting research in tomato disease recognition and detection (Ahmad et al., 2020). These datasets are widely used in agriculture, particularly for plant disease detection, crop growth studies, and plant health management, driving the ongoing development of intelligent agricultural technologies.

3. Model development and training

3.1. The selection of ML model

Image classification, used to categorize input images into predefined groups, is commonly applied in plant identification and disease diagnosis. CNNs, with their strong hierarchical feature extraction abilities, excel in these tasks. Models like AlexNet and ResNet are frequently used to classify plant species and developmental stages (Zhu et al., 2018; Malagol et al., 2025). ResNet, by incorporating residual learning and skip connections, addresses gradient vanishing and degradation in deep networks. Its enhanced model has been applied in high-throughput quantification of grape leaf trichomes, supporting phenotypic analysis and disease resistance studies. CNN-based models generally achieve over 90% accuracy on public datasets, validating their 'High' performance in comparative evaluations (Yu et al., 2021; Yao et al., 2024).

Simpler models like K-nearest neighbors (K-NN) and support vector machines (SVMs) are ideal for smaller datasets with less complex features. Though computationally efficient and easy to implement, they are more sensitive to noise and tend to perform less effectively on complex image data (Ghosh et al., 2022). K-NN, for instance, classifies samples based on proximity in feature space and can be enhanced using surrogate loss training (Picek et al., 2022). SVMs utilize kernel functions like the Radial Basis Function (RBF) to handle non-linear data and prevent overfitting (Sharma et al., 2024). Both are typically rated as 'Medium' in performance due to their limitations in handling large-scale, high-dimensional data (Azlah et al., 2019).

For object detection tasks, which require both classification and localization, models like Faster R-CNN (FRCNN) and You Only Look Once version 5 (YOLOv5) offer high spatial accuracy. FRCNN uses a region proposal network (RPN) and shared convolutional layers to efficiently predict object categories and bounding boxes (Deepika & Arthi, 2022). YOLOv5 enables real-time detection and has been applied to UAV-based monitoring for early detection of pine wilt disease (Yu et al., 2021). In dense slash pine forests, improved FRCNN achieved 95.26% accuracy and an R^2 of 0.95 in crown detection, showcasing the utility of deep learning for woody plant monitoring (Cai et al., 2023).

Other advanced models include deep belief networks (DBNs), which use stacked restricted Boltzmann machines (RBMs) for unsupervised hierarchical learning and are fine-tuned via back-propagation (Lu et al., 2022). Recurrent neural networks (RNNs), particularly long short-term memory (LSTM) networks, are effective for modeling temporal dependencies, such as plant growth simulation using time-lapse imagery (Xing et al., 2023; Liu et al., 2024a, 2024b). Graph neural networks (GNNs) are increasingly used for modeling complex relationships in plant stress response and gene regulation, although they require significant training effort and parameter tuning (Chang et al., 2024).

Table 2. A comparison of common ML models in plant recognition and classification

ML models	Task	Advantages	Disadvantages	Applications	Performance score	References
CNN	Image classification	Highly accurate and suitable for complex image recognition.	Requires a large amount of data with high computational cost.	Recognizing and categorizing complex plant images	High	Yu et al. (2021)
SVM	Classification of high-dimensional data	Effectively handles high-dimensional data.	Sensitive to parameter selection and long training times.	Classifying small to medium-sized plant datasets.	Medium	Ghosh et al. (2022)
RFs	Multi-class classification	Suitable for multi-class classification tasks and demonstrate robust performance.	The model has high complexity, resulting in longer training times.	Handling complex feature plant classification problems.	High	Pandey and Vir (2024)
K-NN	Instance-based classification	Simple and easy to implement, suitable for small datasets.	Sensitive to noise; classification performance depends on distance selection.	Analyzing small-scale datasets or performing preliminary plant recognition.	Medium	Azlah et al. (2019)
DBNs	Deep feature learning	Offers powerful representation capabilities through deep feature learning.	Training is complex and requires adjustment of multiple hyperparameters.	Deep-level plant feature learning and classification	Medium	Shoaib et al. (2023)
Transfer Learning	Transfer knowledge from pre-trained models	Using pre-trained models reduces the amount of training data required.	Adequate pre-trained models related to the task are required.	Quickly applicable to new plant classification tasks	Medium to High	Shahoveisi et al. (2023)

In scenarios with limited labelled data or domain shifts, transfer learning is particularly valuable. By leveraging pre-trained models, it enables medium to high performance in plant classification and disease recognition tasks (Wu et al., 2022). Advanced architectures like GoogLeNet, with its multi-scale inception module, further enhance classification accuracy. For instance, a GoogLeNet model achieved *F*-scores of 0.9988 and 0.9982 in classifying broadleaf and coniferous tree species, respectively, after 100 training epochs (Minowa et al., 2022).

In summary, each model class exhibits distinct strengths. CNNs specialize in image classification; SVMs and K-NN are optimal for simpler datasets; FRCNN and YOLOv5 excel in object detection; DBNs and RNNs support hierarchical and temporal modeling; GNNs tackle high-dimensional interactions; and transfer learning enhances adaptability across domains. The qualitative performance ratings (High/Medium) presented in Table 2 synthesize evaluation metrics like accuracy, precision, recall, and F1-score across representative studies, enabling researchers to select appropriate models based on task complexity and dataset characteristics.

3.2. The integration of language models

In recent years, with the widespread application of LLMs such as BERT and the GPT series in natural language processing and cross-modal learning, plant science research has also begun exploring the integration of LLMs into areas such as gene function prediction, literature-based knowledge mining, and bioinformatic inference. For instance, LLMs can automatically extract potential functional annotation information from a large body of plant gene literature, aiding in the construction of plant gene regulatory networks. Moreover, due to their powerful contextual understanding capabilities, LLMs demonstrate enhanced accuracy and generalizability in predicting gene expression patterns across species (Zhang et al., 2024). The application of LLMs in plant biology is beginning to

transform the field by driving advancements in chemical mapping, genetic research, and disease diagnostics (Eftekhari et al., 2024). For instance, by analyzing data from over 2,500 publications, researchers have revealed the phylogenetic distribution of plant compounds and enabled the creation of systematic chemical maps with improved automation and accuracy (Busta et al., 2024). LLMs and protein language models (PLMs) also enhance the analysis of nucleic acid and protein sequences, advancing genetic improvements and supporting sustainable agricultural systems (Liu et al., 2024a, 2024b). In disease diagnostics, models like contrastive language image pre-training (CLIP) utilize high-quality images and textual annotations to improve classification accuracy for plant diseases, achieving significant precision gains on datasets such as plant village and field plant (Eftekhari et al., 2024).

Similarly, a CNN-based system combining InceptionV3 with GPT-3.5 Turbo achieved 99.85% training accuracy and 88.75% validation accuracy in detecting tomato diseases, providing practical treatment recommendations (Madaki et al., 2024). The feature fusion contrastive language-image pre-training (FF-CLIP) model further enhances this approach by integrating visual and textual data to identify complex disease textures, achieving a 33.38% improvement in Top-1 accuracy for unseen data in zero-shot plant disease identification (Liaw et al., 2025). These advancements highlight the transformative potential of language models in advancing plant biology research and driving sustainable agricultural innovation.

In summary, two notable new insights brought by LLMs in plant science include: (1) the ability to identify and summarize ‘potential regulatory information within non-coding sequences’, which is often overlooked by traditional models; and (2) the promotion of holistic modeling of plant trait complexity through multimodal integration – such as combining sequence, image, and textual data – offering new avenues for complex trait prediction and breeding design.

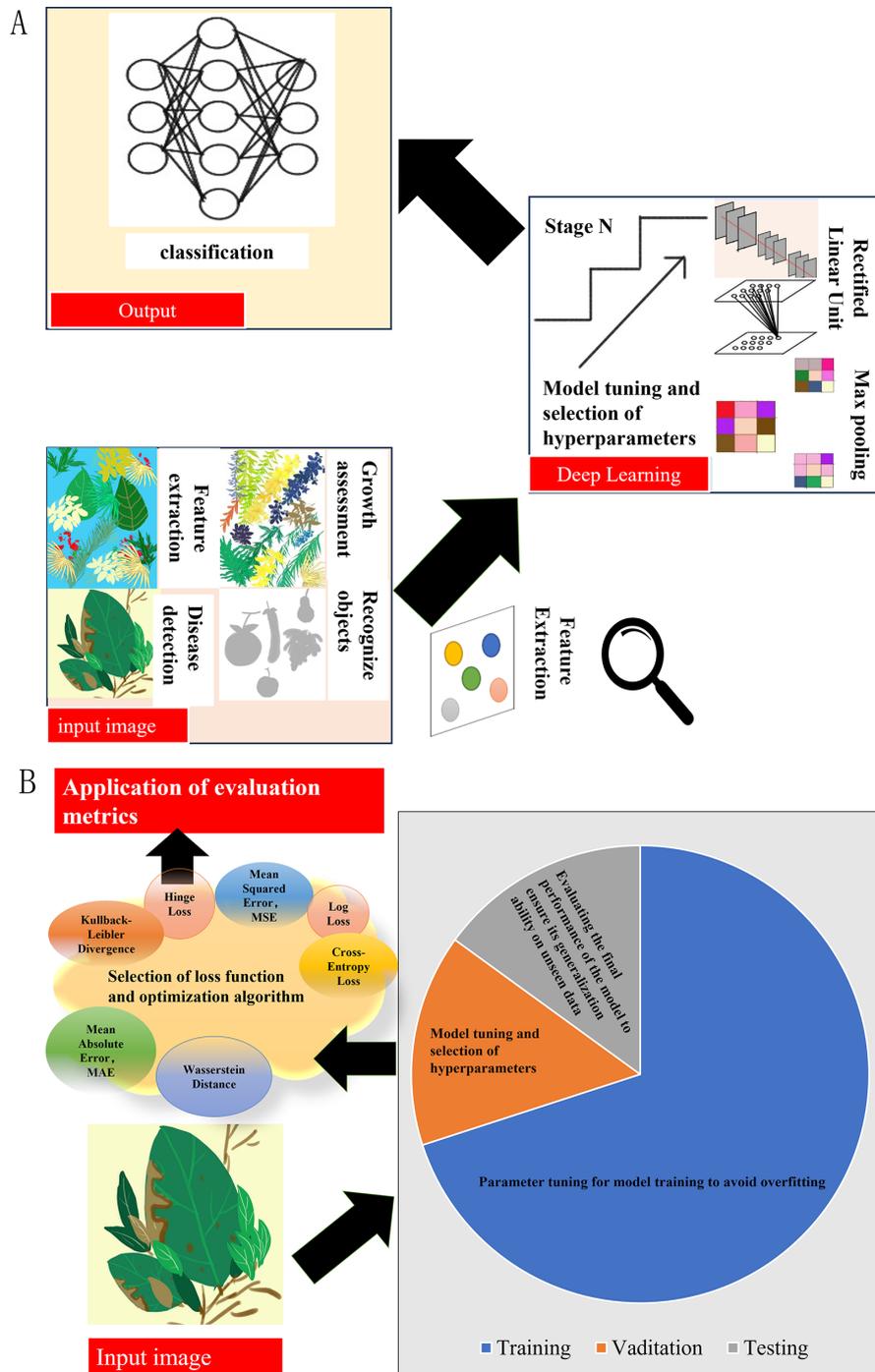


Figure 1. CNN and data training process flowchart. A. DL-based image processing flowchart; B. Data training process flowchart.

3.3. Model training and evaluation methods

Effective ML model training relies on three key components: data partitioning, loss functions, and optimization strategies. Properly splitting the data (commonly 70:15:15 for training, validation, and testing) ensures good generalization (Figure 1). The training set fits model parameters, the validation set guides hyperparameter tuning, and the test set evaluates final performance, enhancing model robustness (Ghazi et al., 2017). For instance, Mohanty et al. used the PlantVillage dataset (54,306 images), applied an 80:10:10 split and cross-entropy loss with SGD optimization, achieving over 99% accuracy and demonstrating the efficiency of deep learning

in plant disease identification (Mohanty et al., 2016). Similarly, Ferentinos used a publicly available plant image dataset with 87,848 leaf images, splitting it into 80% for training and 20% for testing, achieving 99.53% accuracy on the test set. This study highlights the effectiveness of data partitioning and CNNs in plant species classification and disease detection (Ferentinos, 2018).

The loss function quantifies the difference between predicted and true values, minimized during training. For regression tasks, common loss functions include mean squared error (MSE), root mean squared error (RMSE), and mean absolute error (MAE), with MSE penalizing larger errors, RMSE providing interpretable

results, and MAE being more robust to outliers (Picek et al., 2022). For classification tasks, binary classification problems often use binary cross-entropy to measure the accuracy of predictions involving ‘yes/no’ or ‘true/false’ decisions (Bai et al., 2023). Custom loss functions can also be defined to suit specific project needs. For instance, Gillespie et al. developed a deep learning model called ‘Deepbiosphere’ and designed a sampling bias-aware binary cross-entropy loss function, which significantly improved the model’s performance in monitoring changes in rare plant species (Gillespie et al., 2024).

Optimization algorithms are equally critical in DL. Common optimizers include stochastic gradient descent (SGD), adaptive moment estimation (ADAM) (Saleem et al., 2020), and RMSprop (Kanna et al., 2023). SGD updates parameters using randomly selected samples per iteration, ADAM combines momentum with adaptive learning rates, and RMSprop adjusts the learning rate of each parameter using moving averages, effectively reducing gradient oscillation (Mokhtari et al., 2023). For example, Sun et al. employed the SGD optimizer with momentum techniques in plant disease recognition tasks, which improved the convergence speed and stability of the model, thereby enhancing its accuracy (Sun et al., 2021). In a similar vein, Kavitha et al. trained six ImageNet-pretrained CNN models on an RMP dataset for rural medicinal plant classification and reported that MobileNet, optimized with SGD, achieved the best classification performance, highlighting its effectiveness in medicinal plant recognition (Kavitha et al., 2023). In another study, Labhsetwar et al. compared different optimizers for plant disease classification and found that the Adam optimizer achieved the highest validation accuracy of 98%, demonstrating its strong performance in this context. (Labhsetwar et al., 2021). Complementing these findings, Praharsha et al. evaluated multiple optimizers in CNNs and found that RMSprop, with a learning rate of 0.001 and L2 regularization of 0.0001, achieved the highest validation accuracy of 89.09%, outperforming Adam and SGD, and proving especially effective for plant pest classification tasks (Praharsha et al., 2024).

Evaluation metrics like accuracy, recall, F1 score, and area under the curve (AUC) are essential for assessing model performance, with accuracy being effective for balanced datasets but potentially misleading for imbalanced data. (Naidu et al., 2023). Recall emphasizes the model’s ability to identify all relevant positive cases, essential for tasks like plant species identification, while the F1 score, the harmonic mean of precision and recall, offers a balanced evaluation, particularly for imbalanced datasets (Fourure et al., 2021). AUC evaluates model classification ability across different thresholds and is particularly useful in imbalanced classification tasks (Vakili et al., 2020). For example, Tariku et al. developed an automated plant species classification system using UAV-captured Red, Green, Blue (RGB) images and transfer learning, achieving 0.99 accuracy, precision, recall, and 0.995 F1 score, highlighting the effectiveness of these metrics in real-world tasks and the importance of recall and F1 score for handling diverse and imbalanced datasets (Tariku et al., 2023). In another study, Sa et al. introduced WeedMap, a large-scale semantic weed mapping framework using UAV-captured multispectral imagery and deep neural networks, achieving AUCs of 0.839, 0.863, and 0.782 for background, crop, and weed classes, respectively, highlighting the role of AUC in evaluating model performance across multiple categories in real-world agricultural applications (Sa et al., 2018).

Evaluation should be conducted after training and before deployment. Re-evaluation is also necessary when there are changes in data distribution or environmental conditions (Reich

and Barai, 1999). For example, when encountering new plant species or ecological conditions, retraining or fine-tuning may be needed to maintain strong performance on novel inputs (Soares et al., 2017).

4. The applications of ML in plant research

4.1. Biotic and abiotic stress management

DL models are instrumental in analysing plant leaf images to detect diseases and pests, a vital component of plant protection (Shoab et al., 2023). For instance, a 2019 study developed a back propagation neural network (BPNN) – a multilayer feedforward model trained via backpropagation and optimized through one-way ANOVA – that effectively identified rice diseases like blast and blight, highlighting its strength in pattern recognition and classification (Chaudhari & Malathi, 2023). Additionally, hyperspectral remote sensing combined with ML enables rapid detection of plant viruses like *Solanum tuberosum* virus Y, enhancing early disease identification (Polder et al., 2019). The YOLOv5s algorithm processes RGBdrone images for real-time pine wilt disease detection, suitable for large-scale monitoring (Du et al., 2024). Generative adversarial networks (GANs), consisting of a generator and a discriminator that improve through adversarial learning, have been applied in plant science for tasks such as data augmentation, plant disease detection, and growth simulation (Gandhi et al., 2018).

ML and DL advance abiotic stress management through sensors and drones for early detection, precise predictions, and improved plant resilience (Patil et al., 2024). These technologies also optimize plant stress responses, with further advancements expected in agricultural applications (Sharma et al., 2024). Hyperspectral imaging aids early disease detection related to abiotic stresses (Lowe et al., 2017). The ‘ASmiR’ framework predicts plant miRNAs under abiotic stresses, supporting stress-resistant crop breeding (Pradhan et al., 2023). GNN, a DL model tailored for graph-structured data, learns node representations via relationships and adjacency, and has been effectively used to predict miRNA associations with abiotic stresses by capturing complex structural patterns (Chang et al., 2024). Interested readers are encouraged to refer to the cited review, which highlights the role of bioinformatics and AI in managing abiotic stresses for food security, aiding stress gene analysis, and improving crop resilience to drought and salinity (Chang et al., 2024).

4.2. Plant species identification and classification

DL aids large-scale growth monitoring, identifying plant growth patterns, health, and predicting yields. An intelligent greenhouse management system uses ML and mobile networks for automated phenotypic monitoring (Rahman et al., 2024). Shapley Additive Explanations (SHAP), which quantifies each feature’s contribution to model predictions, is commonly used to interpret and evaluate the performance of yield prediction models (Sun et al., 2019). Lightweight SegNet (LW-SegNet) is a CNN architecture tailored for image segmentation tasks, designed to reduce network parameters and computational demands, ensuring efficient and accurate results, particularly in resource-limited environments. Lightweight networks like LW-SegNet and Lightweight U-Net (LW-Unet) enable efficient segmentation of rice varieties in plant research (Zhang et al., 2023a, 2023b, 2023c). A hybrid model combining RF regression and radiative transfer simulation estimates wheat leaf area index (LAI) using UAV multispectral

imaging (Sahoo et al., 2023). Self-supervised Learning (SSL) trains models using patterns in unlabelled data without manual labeling, accelerating the training process; while it speeds up plant breeding with unlabelled datasets, supervised pre-training still generally outperforms SSL, particularly in tasks like leaf counting (Ogidi et al., 2023).

CNNs enable fast and accurate plant species identification. In sustainable agriculture, a CNN-based DL model was developed to classify weeds, optimizing herbicide use for eco-friendly control (Corceiro et al., 2023). CNNs (VGG-16, GoogleNet, ResNet-50, ResNet-101) were employed to identify 23 wild grape species, showcasing the effectiveness of DL in leaf recognition and crop variety identification (Pan et al., 2024).

4.3. Plant growth simulation

ML explores complex molecular and cellular mechanisms in plant growth and development, such as stem cell homeostasis in arabidopsis shoot apical meristems (SAMs) (Hohm et al., 2010), leaf development (Richardson et al., 2021), and sepal giant cell development (Roeder, 2021), and the simulation of weed growth in crop fields over decades (Zhang et al., 2023a, 2023b, 2023c).

Various algorithms have been applied to study plant development. For example, image processing and ML using SVM and RF were used to analyze Cannabis sativa callus morphology, with SVM showing higher accuracy, while genetic algorithms optimized PGR concentrations to validate the model (Hesami and Jones, 2021). A microfluidic chip was created to simulate pollen tube growth, and the ‘Physical microenvironment Assay (SPA)’ method was established to study mechanical signal transmission during pollen tube penetration of pistil tissues (Zhou et al., 2023).

Many groups are utilizing the latest computational technologies and algorithms to develop tools that enhance the efficiency and precision of plant biology research (Muller & Martre, 2019). The virtual plant tissue (VPTissue) software simulates plant developmental processes, facilitating the integration of functional modules and cross-model coupling to efficiently simulate cellular-level plant growth (De Vos et al., 2017). ADAM-Plant software uses stochastic techniques to simulate breeding plans for self- and cross-pollinated crops, tracking genetic changes across scenarios and supporting diverse population structures, genomic models, and selection strategies for optimized breeding design (Liu et al., 2019). The L-Py framework, a Python-based L-system simulation tool, simplifies plant architecture simulation and analysis, with dynamic features that enhance programming flexibility, making plant growth model development more convenient (Boudon et al., 2012). Many groups are also developing advanced computational tools to accurately simulate plant morphological changes at various stages of growth (Boudon et al., 2015). A 3D maize canopy model was created using a t-distribution for the initial model, treating the maize whorl – leaves, stem segments, and buds – as an agent to precisely simulate the canopy’s spatial dynamics and structure (Wu et al., 2024).

Significant advancements from 2012 to 2023 have enhanced our understanding of plant biology. In 2012, researchers used live imaging combined with computational analysis to monitor cellular and tissue dynamics in *A. thaliana* (Cunha et al., 2012). The introduction of the Cellzilla platform in 2013 enabled simulation of plant tissue growth at the cellular level (Shapiro et al., 2013). A pivotal study in 2014 focused on the WOX5-IAA17 feedback loop, which

is essential for maintaining the auxin gradient in *A. thaliana* (Tian et al., 2014). By 2016, research explored plant signaling pathways and mechanical models to analyze sepal growth and morphology (Hervieux et al., 2016). In 2018, studies delineated the expression pattern of the *CLV3* gene in SAMs (Zhou et al., 2018a, 2018b), followed by 2019 research on leaf development and chloroplast ultrastructure (Kierzkowski et al., 2019), and the *TCX2* gene’s role in maintaining stem cell identity (Clark et al., 2019). Research in 2020 focused on epidermis-specific transcription factors affecting stem cell niches (Han et al., 2020), and 2021 introduced new modeling techniques for root tip growth and stem cell division (Marconi et al., 2021). In 2022, 3D bioprinting was used to study cellular dynamics in both *A. thaliana* and *Glycine max* (Van den Broeck et al., 2022). The latest studies in 2023 provided new insights into weed evolution and applied advanced DL techniques for plant cell analysis (Feng et al., 2023). These milestones demonstrate the integration of computational tools and empirical datasets in plant science, enabling innovative methods and applications that propel the field forward.

4.4. Plant cell segmentation

Accurate cell segmentation is crucial for understanding plant cell morphology, developmental processes, and tissue organization. Recent advancements in DL and computer vision have led to the development of various specialized tools for segmenting plant cell structures from complex microscopy data. This section provides an overview of key tools, highlighting their core methodologies, applications, and advantages in plant research.

PlantSeg is a neural network-based tool designed for high-resolution plant cell segmentation. It starts with image preprocessing, including scaling and normalization, followed by U-Net-based boundary prediction to identify cell boundaries (Wei et al., 2024). The boundary map is transformed into a region adjacency graph (RAG), where nodes represent image regions and edges represent boundary predictions. Graph segmentation algorithms, such as Multicut or GASP, partition the graph into individual cells, and post-processing ensures the segmentations align with the original resolution and corrects over-segmentation (Wolny et al., 2020). By streamlining these processes, PlantSeg supports high-throughput analysis of plant cell dynamics, particularly for confocal and light sheet microscopy data (Vijayan et al., 2024). This tool not only improves segmentation efficiency but also handles large-scale datasets, providing robust support for long-term monitoring of plant cell behavior.

Complementing PlantSeg, the Soybean-MVS dataset leverages multi-view stereo (MVS) technology to provide a 3D imaging resource, capturing the full growth cycle of soybeans and enabling precise 3D segmentation of plant organs (Sun et al., 2023). This dataset plays a significant role in plant growth and developmental research, offering fine-grained data support for dynamic analysis of long-term growth processes.

Other tools focus on generalizability and adaptability across various plant species and imaging modalities. Cellpose utilizes a convolutional neural network (CNN), which performs well in segmenting different cell types and shapes, especially in dynamic plant structures and large-scale image analysis, improving scalability (Stringer et al., 2021). This feature enables Cellpose to maintain high accuracy and robustness under diverse experimental conditions when processing plant cells.

DeepCell uses CNNs for plant cell segmentation and supports cell tracking and morphology analysis, offering robust tools

Table 3. Challenges and future trends of ML and image recognition technologies in plant science

	Category	Content	References
Challenges	Image quality and diversity	Imaging variations, such as lighting, angle, and background, can significantly affect recognition accuracy.	Shoib et al. (2023)
	Inter-class similarity	The similarity in appearance among different plant species can present challenges to accurate classification.	Jeyapoornima et al. (2023)
	Intra-class variation	Individuals of the same plant species can show significant variations in appearance, including growth stages and seasonal changes.	Harris (1913)
	Insufficient data	Limited image data for rare or endangered plants makes training effective ML models challenging.	Cong and Zhou (2023)
	Complexity of background	Plant images often have complex backgrounds, such as soil and other plants, which can interfere with recognition and feature extraction.	Wang et al. (2008)
	Real-time processing requirements	Real-time plant identification on mobile devices faces challenges due to high demands on processing speed and resource consumption.	Padhiary et al. (2023)
	The multi-label classification problem	An individual plant image may require multiple labels, like species and disease type, increasing classification complexity.	Anh et al. (2022)
	Adaptability and scalability	Identification systems need good adaptability and scalability to accommodate new species discoveries and updated classification standards.	Rao et al. (2022)
	Future trends	Further applications of DL	DL, especially CNNs, is effective for image recognition and has potential for optimizing plant sample identification and categorization.
The application of weakly supervised learning and unsupervised learning		Due to high annotation costs, weakly supervised and unsupervised learning methods using unlabelled data are expected to be more widely adopted as cost-effective solutions.	Adke et al. (2022)
The development of fine-grained image recognition		Fine-grained image recognition targets distinguishing highly similar species, with future research addressing high inter-class similarity.	Šulc and Matas (2017)
The exploration of cross-domain learning techniques		Transferring advanced technologies from other domains to plant image recognition may help overcome specific challenges.	Chulif et al. (2023)
Increasing interpretability and transparency		As ML models gain prominence in plant science, their interpretability and transparency are key to understanding decision-making processes.	Paudel et al. (2023)
The usage of mobile devices and edge computing		Mobile devices and edge computing for image capture and initial processing will enable real-time plant recognition and data collection in the field.	Khan et al. (2023)
The application of multimodal learning		The system's accuracy and robustness can be improved by integrating image data with genetic and ecological information through multimodal learning.	Zhou et al. (2021)
The integration of cloud computing and big data technologies		Cloud computing and big data will increasingly manage large plant image datasets, offering improved computational resources and storage.	Singh (2018)
Sustainability and environmental monitoring applications		ML and image recognition will support plant conservation, species monitoring, and environmental assessments, aiding sustainable development goals.	Wongchai et al. (2022)

keyword network analysis of DL in plant research publications, while Table 3 outlines the challenges and future trends in ML and image recognition technologies within plant science.

In summary, we examined the usage of DL-based image recognition in plant science, covering plant feature extraction, classification, disease detection, and growth analysis. We highlight the importance of data acquisition and preprocessing methods like high-resolution imaging, drone photography, and 3D scanning, as well as techniques for improving data quality. Various feature extraction methods – such as colour histograms, shape descriptors, and texture features – are reviewed for plant identification. The development of ML models, especially CNNs, is also discussed, alongside current challenges and future prospects. Despite progress, challenges remain. Future research should aim to apply methods across diverse plant systems, refine data acquisition, and enhance algorithm efficiency. Advancements will likely improve model generalization and interpretability, with interdisciplinary collaboration in plant biology, mathematics, and computer science being crucial to addressing upcoming challenges.

Open peer review. To view the open peer review materials for this article, please visit <http://doi.org/10.1017/qpb.2025.10018>.

Acknowledgements

We acknowledge Manjun Shang for critical reading our manuscript. We apologized to colleagues whose work was not included or described in this review due to limited space.

Competing interest. None.

Data availability statement. No data and coding involved in this manuscript.

Author contributions. KYH, HH and YZ conceived the study. KYH wrote the manuscript. HH and YZ revised the manuscript.

Funding statement. This work was supported by National Natural Science Foundation of China (No. 32202496 to H. H.), Nanjing Forestry University (start-up funding to H.H.), and Jiangsu Sheng Tepin Jiaoshou Program (to H.H.).

References

- Abebe, A. M., Kim, Y., Kim, J., Kim, S. L., & Baek, J. (2023). Image-based high-throughput phenotyping in horticultural crops. *Plants*, *12*(10), 2061. <https://doi.org/10.3390/plants12102061>
- Adke, S., Li, C., Rasheed, K. M., & Maier, F. W. (2022). Supervised and weakly supervised deep learning for segmentation and counting of cotton bolls using proximal imagery. *Sensors*, *22*(10), 3688. <https://doi.org/10.3390/s22103688>
- Agathokleous, E., Rillig, M. C., Peñuelas, J., & Yu, Z. (2024). One hundred important questions facing plant science derived using a large language model. *Trends in Plant Science*, *29*(2), 210–218. <https://doi.org/10.1016/j.tplants.2023.06.008>
- Ahmad, I., Hamid, M., Yousaf, S., Shah, S. T., & Ahmad, M. O. (2020). Optimizing pretrained convolutional neural networks for tomato leaf disease detection. *Complexity*, *2020*(1), 8812019. <https://doi.org/10.1109/ACCESS.2021.3119655>
- Ahmad, M., Abdullah, M., Moon, H., & Han, D. (2021). Plant disease detection in imbalanced datasets using efficient convolutional neural networks with stepwise transfer learning. *IEEE Access*, *9*, 140565–140580. <https://doi.org/10.1109/ACCESS.2021.3119655>
- Ahmad, M. U., Ashiq, S., Badshah, G., Khan, A. H., & Hussain, M. (2022). Feature extraction of plant leaf using deep learning. *Complexity*, *2022*(1), 6976112. <https://doi.org/10.1155/2022/6976112>
- Anh, P. T. Q., Thuyet, D. Q., & Kobayashi, Y. (2022). Image classification of root-trimmed garlic using multi-label and multi-class classification with deep convolutional neural network. *Postharvest Biology and Technology*, *190*, 111956. <https://doi.org/10.1016/j.postharvbio.2022.111956>
- Azlah, M. A. F., Chua, L. S., Rahmad, F. R., Abdullah, F. I., & Wan Alwi, S. R. (2019). Review on techniques for plant leaf classification and recognition. *Computers*, *8*(4), 77. <https://doi.org/10.3390/computers8040077>
- Bai, X., Gu, S., Liu, P., Yang, A., Cai, Z., Wang, J., & Yao, J. (2023). Rpnnet: Rice plant counting after tillering stage based on plant attention and multiple supervision network. *The Crop Journal*, *11*(5), 1586–1594. <https://doi.org/10.1016/j.cj.2023.04.005>
- Barbedo, J. G. A. (2016). A review on the main challenges in automatic plant disease identification based on visible range images. *Biosystems Engineering*, *144*, 52–60. <https://doi.org/10.1016/j.biosystemseng.2016.01.017>
- Boudon, F., Pradal, C., Cokelaer, T., Prusinkiewicz, P., & Godin, C. (2012). L-Py: An L-system simulation framework for modeling plant architecture development based on a dynamic language. *Frontiers in Plant Science*, *3*, 76. <https://doi.org/10.3389/fpls.2012.00076>
- Boudon, F., Chopard, J., Ali, O., Gilles, B., Hamant, O., Boudaoud, A., et al. (2015). A computational framework for 3D mechanical modeling of plant morphogenesis with cellular resolution. *PLoS Computational Biology*, *11*(1), e1003950. <https://doi.org/10.1371/journal.pcbi.1003950>
- Bühler, J., Rishmawi, L., Pflugfelder, D., Huber, G., Scharr, H., Hülskamp, M., ... Jahnke, S. (2015). Phenovein – A tool for leaf vein segmentation and analysis. *Plant Physiology*, *169*(4), 2359–2370. <https://doi.org/10.1104/pp.15.00974>
- Busta, L., Hall, D., Johnson, B., Schaut, M., Hanson, C. M., Gupta, A., ... Maeda, A. (2024). Mapping of specialized metabolite terms onto a plant phylogeny using text mining and large language models. *The Plant Journal*, *120*(1), 406–419. <https://doi.org/10.1111/tpj.16906>
- Cai, L., Zhang, Z., Zhu, Y., Zhang, L., Li, M., & Xue, X. (2022). Bigdetection: A large-scale benchmark for improved object detector pre-training. In *Paper presented at the Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. <https://doi.org/10.48550/arXiv.2203.13249>
- Cai, C., Xu, H., Chen, S., Yang, L., Weng, Y., Huang, S., ... Lou, X. (2023). Tree recognition and crown width extraction based on novel faster-RCNN in a dense loblolly pine environment. *Forests*, *14*(5), 863. <https://doi.org/10.3390/f14050863>
- Cap, Q. H., Uga, H., Kagiwada, S., & Iyatomi, H. (2020). Leafgan: An effective data augmentation method for practical plant disease diagnosis. *IEEE Transactions on Automation Science and Engineering*, *19*(2), 1258–1267. <https://doi.org/10.1109/TASE.2020.3041499>
- Chang, L., Jin, X., Rao, Y., & Zhang, X. (2024). Predicting abiotic stress-responsive miRNA in plants based on multi-source features fusion and graph neural network. *Plant Methods*, *20*(1), 33. <https://doi.org/10.1186/s13007-024-01158-7>
- Chaudhari, D. J., & Malathi, K. (2023). Detection and prediction of rice leaf disease using a hybrid CNN-SVM model. *Optical Memory and Neural Networks*, *32*(1), 39–57. <https://doi.org/10.3103/S1060992X2301006X>
- Chen, Y., Huang, Y., Zhang, Z., Wang, Z., Liu, B., Liu, C., ... Wan, F. (2023). Plant image recognition with deep learning: A review. *Computers and Electronics in Agriculture*, *212*, 108072. <https://doi.org/10.1016/j.compag.2023.108072>
- Chulif, S., Lee, S. H., Chang, Y. L., & Chai, K. C. (2023). A machine learning approach for cross-domain plant identification using herbarium specimens. *Neural Computing and Applications*, *35*(8), 5963–5985. <https://doi.org/10.1007/s00521-022-07951-6>
- Clark, N. M., Buckner, E., Fisher, A. P., Nelson, E. C., Nguyen, T. T., Simmons, A. R., et al. (2019). Stem-cell-ubiquitous genes spatiotemporally coordinate division through regulation of stem-cell-specific gene networks. *Nature Communications*, *10*(1), 5574. <https://doi.org/10.1038/s41467-019-13132-2>
- Cong, S., & Zhou, Y. (2023). A review of convolutional neural network architectures and their optimizations. *Artificial Intelligence Review*, *56*(3), 1905–1969. <https://doi.org/10.1007/s10462-022-10213-5>
- Corceiro, A., Alibabaei, K., Assunção, E., Gaspar, P. D., & Pereira, N. (2023). Methods for detecting and classifying weeds, diseases and fruits using AI to improve the sustainability of agricultural crops: A review. *PRO*, *11*(4), 1263. <https://doi.org/10.3390/pr11041263>
- Cunha, A., Tarr, P. T., Roeder, A. H., Altinok, A., Mjolsness, E., & Meyerowitz, E. M. (2012). Computational analysis of live cell images of the *Arabidopsis thaliana* plant. In *Methods in cell biology* (Vol. 110, pp. 285–323): Elsevier. <https://doi.org/10.1016/B978-0-12-388403-9.00012-6>
- De Vos, D., Dzhurakhalov, A., Stijven, S., Klosiewicz, P., Beemster, G. T., & Broeckhove, J. (2017). Virtual plant tissue: Building blocks for next-generation plant growth simulation. *Frontiers in Plant Science*, *8*, 686. <https://doi.org/10.3389/fpls.2017.00686>
- Deepika, P., & Arthi, B. (2022). Prediction of plant pest detection using improved mask FRCNN in cloud environment. *Measurement: Sensors*, *24*, 100549. <https://doi.org/10.1016/j.measen.2022.100549>
- Du, Z., Wu, S., Wen, Q., Zheng, X., Lin, S., & Wu, D. (2024). Pine wilt disease detection algorithm based on improved YOLOv5. *Frontiers in Plant Science*, *15*, 1302361. <https://doi.org/10.3389/fpls.2024.1302361>
- Duncan, K. E., Czymmek, K. J., Jiang, N., Thies, A. C., & Topp, C. N. (2022). X-ray microscopy enables multiscale high-resolution 3D imaging of plant cells, tissues, and organs. *Plant Physiology*, *188*(2), 831–845. <https://doi.org/10.1093/plphys/kiab405>
- Eftekhari, P., Serra, G., Schoeffmann, K., & Portelli, D. B. (2024). Using natural language processing to enhance visual models for plant leaf diseases classification. <https://netlibrary.aau.at/obvuklhs/download/pdf/10268703>
- Feng, X., Yu, Z., Fang, H., Jiang, H., Yang, G., Chen, L., ... Hu, G. (2023). Plantorganelle hunter is an effective deep-learning-based method for plant organelle phenotyping in electron microscopy. *Nature Plants*, *9*(10), 1760–1775. <https://doi.org/10.1038/s41477-023-01527-5>
- Ferentinos, K. P. (2018). Deep learning models for plant disease detection and diagnosis. *Computers and Electronics in Agriculture*, *145*, 311–318. <https://doi.org/10.1016/j.compag.2018.01.009>
- Forero, M. G., Murcia, H. F., Mendez, D., & Betancourt-Lozano, J. (2022). LiDAR platform for acquisition of 3D plant phenotyping database. *Plants*, *11*(17), 2199. <https://doi.org/10.3390/plants11172199>
- Fourure, D., Javaid, M. U., Posocco, N., & Tihon, S. (2021). *Anomaly detection: How to artificially increase your f1-score with a biased evaluation protocol*. Paper presented at the Joint European conference on machine learning and knowledge discovery in databases. https://doi.org/10.1007/978-3-030-86514-6_1
- Gandhi, R., Nimbalkar, S., Yelamanchili, N., & Ponskhe, S. (2018). *Plant disease detection using CNNs and GANs as an augmentative approach*. Paper presented at the +2018 IEEE International Conference on Innovative Research and Development (ICIRD). <https://doi.org/10.1109/ICIRD.2018.8376321>
- Ghazi, M. M., Yanikoglu, B., & Aptoula, E. (2017). Plant identification using deep neural networks via optimization of transfer learning parameters. *Neurocomputing*, *235*, 228–235. <https://doi.org/10.1016/j.neucom.2017.01.018>

- Ghosh, S., Singh, A., Jhanjhi, N., Masud, M., & Aljahdali, S. (2022). SVM and KNN based CNN architectures for plant classification. *Computers, Materials & Continua*, *71*(3). <https://doi.org/10.32604/cmc.2022.023414>
- Gillespie, L. E., Ruffley, M., & Exposito-Alonso, M. (2024). Deep learning models map rapid plant species changes from citizen science and remote sensing data. *Proceedings of the National Academy of Sciences*, *121*(37), e2318296121. <https://doi.org/10.1073/pnas.2318296121>
- Greenwald, N. F., Miller, G., Moen, E., Kong, A., Kagel, A., Dougherty, T., et al. (2022). Whole-cell segmentation of tissue images with human-level performance using large-scale data annotation and deep learning. *Nature Biotechnology*, *40*(4), 555–565. <https://doi.org/10.1038/s41587-021-01094-0>
- Greydanus, S. J., & Kobak, D. (2020). *Scaling down deep learning with MNIST-1D*. Paper presented at the Forty-first International Conference on Machine Learning. <https://doi.org/10.48550/arXiv.2011.14439>
- Han, H., Yan, A., Li, L., Zhu, Y., Feng, B., Liu, X., & Zhou, Y. (2020). A signal cascade originated from epidermis defines apical-basal patterning of Arabidopsis shoot apical meristems. *Nature Communications*, *11*(1), 1214. <https://doi.org/10.1038/s41467-020-14989-4>
- Harris, J. A. (1913). On the calculation of intra-class and inter-class coefficients of correlation from class moments when the number of possible combinations is large. *Biometrika*, *9*(3/4), 446–472. <https://doi.org/10.1093/biomet/9.3-4.446>
- Hervieux, N., Dumond, M., Sapala, A., Routier-Kierzkowska, A.-L., Kierzkowski, D., Roeder, A. H., et al. (2016). A mechanical feedback restricts sepal growth and shape in Arabidopsis. *Current Biology*, *26*(8), 1019–1028. <https://doi.org/10.1016/j.cub.2016.03.004>
- Hesami, M., & Jones, A. M. P. (2021). Modeling and optimizing callus growth and development in Cannabis sativa using random forest and support vector machine in combination with a genetic algorithm. *Applied Microbiology and Biotechnology*, *105*(12), 5201–5212. <https://doi.org/10.1007/s00253-021-11375-y>
- Hohm, T., Zitzler, E., & Simon, R. (2010). A dynamic model for stem cell homeostasis and patterning in Arabidopsis meristems. *PLoS One*, *5*(2), e9189. <https://doi.org/10.1371/journal.pone.0009189>
- Hwang, S. W., & Sugiyama, J. (2021). Computer vision-based wood identification and its expansion and contribution potentials in wood science: A review. *Plant Methods*, *17*(1), 47. <https://doi.org/10.1186/s13007-021-00746-1>
- Jeyapoornima, B., Suganthi, P., Shankar, K., Chowdary, V. V., Sai, V. R. U., & Venkat, B. N. (2023). An Automated System Based on Informational Value of Images for the Purpose of Classifying Plant Diseases. Paper presented at the 2023 International Conference on Artificial Intelligence and Knowledge Discovery in Concurrent Engineering (ICECONF). <https://doi.org/10.1109/ICECONF57129.2023.10083618>
- Kanna, G. P., Kumar, S. J., Kumar, Y., Changela, A., Woźniak, M., Shafi, J., & Ijaz, M. F. (2023). Advanced deep learning techniques for early disease prediction in cauliflower plants. *Scientific Reports*, *13*(1), 18475. <https://doi.org/10.1038/s41598-023-45403-w>
- Kavitha, S., Kumar, T. S., Naresh, E., Kalmani, V. H., Bamane, K. D., & Pareek, P. K. (2023). Medicinal plant identification in real-time using deep learning model. *SN Computer Science*, *5*(1), 73. <https://doi.org/10.1007/s42979-023-02398-5>
- Kerstens, M., Strauss, S., Smith, R., & Willemsen, V. (2020). From stained plant tissues to quantitative cell segmentation analysis with MorphoGraphX. *Plant Embryogenesis: Methods and Protocols*, 63–83. https://doi.org/10.1007/978-1-0716-0342-0_6
- Khan, A. T., Jensen, S. M., Khan, A. R., & Li, S. (2023). Plant disease detection model for edge computing devices. *Frontiers in Plant Science*, *14*, 1308528. <https://doi.org/10.3389/fpls.2023.1308528>
- Kierzkowski, D., Runions, A., Vuolo, F., Strauss, S., Lymbouridou, R., Routier-Kierzkowska, A.-L., et al. (2019). A growth-based framework for leaf shape development and diversity. *Cell*, *177*(6), 1405, e1417–1418. <https://doi.org/10.1016/j.cell.2019.05.011>
- Kuo, C.-C. J., Zhang, M., Li, S., Duan, J., & Chen, Y. (2019). Interpretable convolutional neural networks via feedforward design. *Journal of Visual Communication and Image Representation*, *60*, 346–359. <https://doi.org/10.1016/j.jvcir.2019.03.010>
- Labhsetwar, S. R., Haridas, S., Panmand, R., Deshpande, R., Kolte, P. A., & Pati, S. (2021, January). Performance analysis of optimizers for plant disease classification with convolutional neural networks. In *2021 4th biennial international conference on nascent technologies in engineering (ICNTE)* (pp. 1–6). IEEE. <https://doi.org/10.1109/ICNTE51185.2021.9487698>
- Liaw, J. Z., Chai, A. Y. H., Lee, S. H., Bonnet, P., & Joly, A. (2025). *Can language improve visual features for distinguishing unseen plant diseases?* Paper presented at the International Conference on Pattern Recognition. https://doi.org/10.1007/978-3-031-78113-1_20
- Liu, H., Tessema, B. B., Jensen, J., Cericola, F., Andersen, J. R., & Sørensen, A. C. (2019). ADAM-plant: A software for stochastic simulations of plant breeding from molecular to phenotypic level and from simple selection to complex speed breeding programs. *Frontiers in Plant Science*, *9*, 425945. <https://doi.org/10.3389/fpls.2018.01926>
- Liu, G., Chen, L., Wu, Y., Han, Y., Bao, Y., & Zhang, T. (2024a). PDLLMs: A group of tailored DNA large language models for analyzing plant genomes. *Molecular Plant*. <https://doi.org/10.1016/j.molp.2024.12.006>
- Liu, X., Guo, J., Zheng, X., Yao, Z., Li, Y., & Li, Y. (2024b). Intelligent plant growth monitoring system based on LSTM network. *IEEE Sensors Journal*. <https://doi.org/10.1109/JSEN.2024.3376818>
- Lowe, A., Harrison, N., & French, A. P. (2017). Hyperspectral image analysis techniques for the detection and classification of the early onset of plant disease and stress. *Plant Methods*, *13*(1), 80. <https://doi.org/10.1186/s13007-017-0233-z>
- Lu, Y., Du, J., Liu, P., Zhang, Y., & Hao, Z. (2022). Image classification and recognition of rice diseases: A hybrid DBN and particle swarm optimization algorithm. *Frontiers in Bioengineering and Biotechnology*, *10*, 855667. <https://doi.org/10.3389/fbioe.2022.855667>
- Madaki, A. Y., Muhammad-Bello, B., & Kusharki, M. (2024). AI-driven detection and treatment of tomato plant diseases using convolutional neural networks and OpenAI language models. *Proceedings of the 5th International Electronic Conference on Applied Sciences, Computing and Artificial Intelligence*, 4–6 December 2024, MDPI; Basel, Switzerland. <https://sciforum.net/manuscripts/20907/slides.pdf>
- Mahajan, S., Raina, A., Gao, X.-Z., & Kant Pandit, A. (2021). Plant recognition using morphological feature extraction and transfer learning over SVM and AdaBoost. *Symmetry*, *13*(2), 356. <https://doi.org/10.3390/sym13020356>
- Majji, V. A., & Kumaravelan, G. (2021). Detection and classification of plant leaf disease using convolutional neural network on plant village dataset. *International Journal of Innovative Research in Applied Sciences and Engineering (IJIRASE)*, *4*(11), 931–935. <https://doi.org/10.29027/IJIRASE.v4.i11.2021.931-935>
- Malagol, N., Rao, T., Werner, A., Töpfer, R., & Hausmann, L. (2025). A high-throughput ResNet CNN approach for automated grapevine leaf hair quantification. *Scientific Reports*, *15*(1), 1590. <https://doi.org/10.1038/s41598-025-85336-0>
- Mano, S., Miwa, T., Nishikawa, S.-i., Mimura, T., & Nishimura, M. (2009). Seeing is believing: on the use of image databases for visually exploring plant organelle dynamics. *Plant and cell physiology*, *50*(12), 2000–2014. <https://doi.org/10.1093/pcp/pcp128>
- Maraveas, C. (2024). Image analysis artificial intelligence Technologies for Plant Phenotyping: Current state of the art. *AgriEngineering*, *6*(3), 3375–3407. <https://doi.org/10.3390/agriengineering6030193>
- Marconi, M., Gallemi, M., Benkova, E., & Wabnik, K. (2021). A coupled mechano-biochemical model for cell polarity guided anisotropic root growth. *eLife*, *10*, e72132. <https://doi.org/10.7554/elife.72132>
- Mendoza-Revilla, J., Trop, E., Gonzalez, L., Roller, M., Dalla-Torre, H., de Almeida, B. P., et al. (2024). A foundational large language model for edible plant genomes. *Communications Biology*, *7*(1), 835. <https://doi.org/10.1038/s42003-024-06465-2>
- Minowa, Y., Kubota, Y., & Nakatsukasa, S. (2022). Verification of a deep learning-based tree species identification model using images of broadleaf and coniferous tree leaves. *Forests*, *13*(6), 943. <https://doi.org/10.3390/f13060943>
- Mohameth, F., Bingcai, C., & Sada, K. A. (2020). Plant disease detection with deep learning and feature extraction using plant village. *Journal of Computer and Communications*, *8*(6), 10–22. <https://doi.org/10.4236/jcc.2020.86002>

- Mohan, A., & Peebles, J. (2024). *Lacunarity pooling layers for plant image classification using texture analysis*. Paper presented at the proceedings of the IEEE/CVF conference on computer vision and pattern recognition. <https://doi.org/10.48550/arXiv.2404.16268>
- Mohanty, S. P., Hughes, D. P., & Salathé, M. (2016). Using deep learning for image-based plant disease detection. *Frontiers in Plant Science*, 7, 1419. <https://doi.org/10.3389/fpls.2016.01419>
- Mokhtari, A., Ahmadnia, F., Nahavandi, M., & Rasoulzadeh, R. (2023). A comparative analysis of the Adam and RMSprop optimizers on a convolutional neural network model for predicting common diseases in strawberries. <https://www.researchgate.net/publication/375372415>.
- Mühlenstädt, T., & Frtunikj, J. (2024). How much data do you need? Part 2: Predicting DL class specific training dataset sizes. *arXiv preprint arXiv:2403.06311*. <https://doi.org/10.48550/arXiv.2403.06311>
- Muller, B., & Martre, P. (2019). Plant and crop simulation models: Powerful tools to link physiology, genetics, and phenomics. *Journal of Experimental Botany* 70, 2339–2344. <https://doi.org/10.1093/jxb/erz175>
- Naidu, G., Zuva, T., & Sibanda, E. M. (2023). *A review of evaluation metrics in machine learning algorithms*. Paper presented at the computer science online conference. https://doi.org/10.1007/978-3-031-35314-7_2
- Nguyen, C. V., Fripp, J., Lovell, D. R., Furbank, R., Kuffner, P., Daily, H., & Sirault, X. (2016). 3D scanning system for automatic high-resolution plant phenotyping. Paper presented at the 2016 international conference on digital image computing: techniques and applications (DICTA). <https://doi.org/10.48550/arXiv.1702.08112>
- Ogidi, F. C., Eramian, M. G., & Stavness, I. (2023). Benchmarking self-supervised contrastive learning methods for image-based plant phenotyping. *Plant Phenomics*, 5, 0037. <https://doi.org/10.34133/plantphenomics.0037>
- Padhiary, M., Rani, N., Saha, D., Barbhuiya, J. A., & Sethi, L. (2023). Efficient precision agriculture with python-based raspberry Pi image processing for real-time plant target identification. *IJRAR*, 10(3). https://www.researchgate.net/publication/373825210_Efficient_Precision_Agriculture_with_Pythonbased_Raspberry_Pi_Image_Processing_for_Real-time_Plant_Target_Identification
- Pan, B., Liu, C., Su, B., Ju, Y., Fan, X., Zhang, Y., et al. (2024). Research on species identification of wild grape leaves based on deep learning. *Scientia Horticulturae*, 327, 112821. <https://doi.org/10.1016/j.scienta.2023.112821>
- Pandey, A., & Vir, R. (2024). An Improved Approach to Classify Plant Disease using CNN and Random Forest. Paper presented at the 2024 IEEE 3rd World Conference on Applied Intelligence and Computing (AIC). <https://doi.org/10.1109/ICAIA57370.2023.10169830>
- Pandey, V., Tripathi, U., Singh, V. K., Gaur, Y. S., & Gupta, D. (2024). Survey of accuracy prediction on the plant village dataset using different ML techniques. *EAI Endorsed Transactions on Internet of Things*, 10. <https://doi.org/10.4108/eetiot.4578>
- Patil, S., Kolhar, S., & Jagtap, J. (2024). Deep learning in plant stress phenomics studies – A review. *International Journal of Computing and Digital Systems*, 16(1), 305–316. <https://doi.org/10.12785/ijcds/1571029493>
- Paudel, D., de Wit, A., Boogaard, H., Marcos, D., Osinga, S., & Athanasiadis, I. N. (2023). Interpretability of deep learning models for crop yield forecasting. *Computers and Electronics in Agriculture*, 206, 107663. <https://doi.org/10.1016/j.compag.2023.107663>
- Petrellis, N. (2019). *Plant disease diagnosis with colour normalization*. Paper presented at the 2019 8th international conference on modern circuits and systems technologies (MOCAST). <https://doi.org/10.1109/MOCAST.2019.8741614>
- Picek, L., Šulc, M., Patel, Y., & Matas, J. (2022). Plant recognition by AI: Deep neural nets, transformers, and kNN in deep embeddings. *Frontiers in Plant Science*, 13, 787527. <https://doi.org/10.3389/fpls.2022.787527>
- Polder, G., Blok, P. M., De Villiers, H. A., Van der Wolf, J. M., & Kamp, J. (2019). Potato virus Y detection in seed potatoes using deep learning on hyperspectral images. *Frontiers in Plant Science*, 10, 209. <https://doi.org/10.3389/fpls.2019.00209>
- Pradhan, U. K., Meher, P. K., Naha, S., Rao, A. R., Kumar, U., Pal, S., & Gupta, A. (2023). ASmiR: A machine learning framework for prediction of abiotic stress-specific miRNAs in plants. *Functional & Integrative Genomics*, 23(2), 92. <https://doi.org/10.1007/s10142-023-01014-2>
- Praharsha, C. H., Poulouse, A., & Badgajar, C. (2024). Comprehensive investigation of machine learning and deep learning networks for identifying multispecies tomato insect images. *Sensors*, 24(23), 7858. <https://doi.org/10.3390/s24237858>
- Rahman, H., Shah, U. M., Riaz, S. M., Kifayat, K., Moqurrah, S. A., & Yoo, J. (2024). Digital twin framework for smart greenhouse management using next-gen mobile networks and machine learning. *Future Generation Computer Systems*. <https://doi.org/10.1016/j.future.2024.03.023>
- Rao, D. S., Ch, R. B., Kiran, V. S., Rajasekhar, N., Srinivas, K., Akshay, P. S., ... Bharadwaj, B. L. (2022). Plant disease classification using deep bilinear CNN. *Intell. Autom. Soft Comput*, 31(1), 161–176. <https://doi.org/10.32604/iasc.2022.017706>
- Reich, Y., & Barai, S. (1999). Evaluating machine learning models for engineering problems. *Artificial Intelligence in Engineering*, 13(3), 257–272. [https://doi.org/10.1016/S0954-1810\(98\)00021-1](https://doi.org/10.1016/S0954-1810(98)00021-1)
- Richardson, A., Cheng, J., Johnston, R., Kennaway, R., Conlon, B., Rebocho, A., et al. (2021). Evolution of the grass leaf by primordium extension and petiole-lamina remodeling. *Science*, 374(6573), 1377–1381. <https://doi.org/10.1126/science.abf9407>
- Robinson, H. F., & Vink, J. N. (2024). Rapid and robust monitoring of Phytophthora infectivity through detached leaf assays with automated image analysis. In *Phytophthora: Methods and protocols* (pp. 93–104): Springer. https://doi.org/10.1007/978-1-0716-4330-3_7
- Roeder, A. H. (2021). Arabidopsis sepals: A model system for the emergent process of morphogenesis. *Quantitative Plant Biology*, 2, e14. <https://doi.org/10.1017/qpb.2021.12>
- Sa, I., Popović, M., Khanna, R., Chen, Z., Lottes, P., Liebisch, F., ... Siegwart, R. (2018). WeedMap: A large-scale semantic weed mapping framework using aerial multispectral imaging and deep neural network for precision farming. *Remote Sensing*, 10(9), 1423. <https://doi.org/10.3390/rs10091423>
- Sachar, S., & Kumar, A. (2021). Survey of feature extraction and classification techniques to identify plant through leaves. *Expert Systems with Applications*, 167, 114181. <https://doi.org/10.1016/j.eswa.2020.114181>
- Sahoo, R. N., Gakhar, S., Rejith, R. G., Verrelst, J., Ranjan, R., Kondraju, T., ... Kumar, S. (2023). Optimizing the retrieval of wheat crop traits from UAV-borne hyperspectral image with radiative transfer modelling using Gaussian process regression. *Remote Sensing*, 15(23), 5496. <https://doi.org/10.3390/rs15235496>
- Saleem, M. H., Potgieter, J., & Arif, K. M. (2019). Plant disease detection and classification by deep learning. *Plants*, 8(11), 468. <https://doi.org/10.3390/plants8110468>
- Saleem, M. H., Potgieter, J., & Arif, K. M. (2020). Plant disease classification: A comparative evaluation of convolutional neural networks and deep learning optimizers. *Plants*, 9(10), 1319. <https://doi.org/10.3390/plants9101319>
- Shahi, T. B., Xu, C. Y., Neupane, A., & Guo, W. (2022). Machine learning methods for precision agriculture with UAV imagery: A review. *Electronic Research Archive*, 30(12), 4277–4317. <https://doi.org/10.3934/era.2022218>
- Shapiro, B. E., Meyerowitz, E., & Mjolsness, E. (2013). Using cellzilla for plant growth simulations at the cellular level. *Frontiers in Plant Science*, 4, 57520. <https://doi.org/10.3389/fpls.2013.00408>
- Shahoveisi, F., Taheri Gorji, H., Shahabi, S., Hosseinirad, S., Markell, S., & Vasefi, F. (2023). Application of image processing and transfer learning for the detection of rust disease. *Scientific Reports*, 13(1), 5133. <https://doi.org/10.1038/s41598-023-31942-9>
- Sharma, G., Kumar, A., Gour, N., Saini, A. K., Upadhyay, A., & Kumar, A. (2024). Cognitive framework and learning paradigms of plant leaf classification using artificial neural network and support vector machine. *Journal of Experimental & Theoretical Artificial Intelligence*, 36(4), 585–610. <https://doi.org/10.1080/0952813X.2022.2096698>
- Shoab, M., Shah, B., Ei-Sappagh, S., Ali, A., Alenezi, F., Hussain, T., & Ali, F. (2023). An advanced deep learning models-based plant disease detection: A review of recent research. *Frontiers in Plant Science*, 14, 1158933. <https://www.x-mol.com/paperRedirect/1638206368800894976>.
- Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. *Journal of Big Data*, 6(1), 1–48. <https://doi.org/10.1186/s40537-019-0197-0>

- Silva, J. C. F., Teixeira, R. M., Silva, F. F., Brommonschenkel, S. H., & Fontes, E. P. (2019). Machine learning approaches and their current application in plant molecular biology: A systematic review. *Plant Science*, **284**, 37–47. <https://doi.org/10.1016/j.plantsci.2019.03.020>
- Singh, D., Jain, N., Jain, P., Kayal, P., Kumawat, S., & Batra, N. (2020). PlantDoc: A dataset for visual plant disease detection. In *Proceedings of the 7th ACM IKDD CoDS and 25th COMAD* (pp. 249–253). <https://doi.org/10.1145/3371158.3371196>
- Singh, K. K. (2018). An artificial intelligence and cloud based collaborative platform for plant disease identification, tracking and forecasting for farmers. Paper presented at the 2018 IEEE international conference on cloud computing in emerging markets (CCEM). <https://doi.org/10.1109/CCEM.2018.00016>
- Soares, R., Ferreira, P., & Lopes, L. (2017). Can plant-pollinator network metrics indicate environmental quality? *Ecological Indicators*, **78**, 361–370. <https://doi.org/10.1016/j.ecolind.2017.03.037>
- Stringer, C., Wang, T., Michaelos, M., & Pachitariu, M. (2021). Cellpose: A generalist algorithm for cellular segmentation. *Nature Methods*, **18**(1), 100–106. <https://doi.org/10.1038/s41592-020-01018-x>
- Šulc, M., & Matas, J. (2017). Fine-grained recognition of plants from images. *Plant Methods*, **13**, 1–14. <https://doi.org/10.1186/s13007-017-0265-4>
- Sun, J., Di, L., Sun, Z., Shen, Y., & Lai, Z. (2019). County-level soybean yield prediction using deep CNN-LSTM model. *Sensors*, **19**(20), 4363. <https://doi.org/10.3390/s19204363>
- Sun, Y., Liu, Y., Zhou, H., & Hu, H. (2021). Plant diseases identification through a discount momentum optimizer in deep learning. *Applied Sciences*, **11**(20), 9468. <https://doi.org/10.3390/app11209468>
- Sun, Y., Zhang, Z., Sun, K., Li, S., Yu, J., Miao, L., et al. (2023). Soybean-MVS: Annotated three-dimensional model dataset of whole growth period soybeans for 3D plant organ segmentation. *Agriculture*, **13**(7), 1321. <https://doi.org/10.3390/agriculture13071321>
- Syarovy, M., Pradiko, I., Farrasati, R., Rasyid, S., Mardiana, C., Pane, R. et al. (2024). Pre-processing techniques using a machine learning approach to improve model accuracy in estimating oil palm leaf chlorophyll from portable chlorophyll meter measurement. Paper presented at the IOP conference series: Earth and environmental science. <https://doi.org/10.1088/1755-1315/1308/1/012054>
- Tariku, G., Ghiglieno, I., Gilioli, G., Gentilin, F., Armiraglio, S., & Serina, I. (2023). Automated identification and classification of plant species in heterogeneous plant areas using unmanned aerial vehicle-collected RGB images and transfer learning. *Drones*, **7**(10), 599. <https://doi.org/10.3390/drones7100599>
- Tian, H., Wabnik, K., Niu, T., Li, H., Yu, Q., Pollmann, S., et al. (2014). WOX5-IAA17 feedback circuit-mediated cellular auxin response is crucial for the patterning of root stem cell niches in Arabidopsis. *Molecular Plant*, **7**(2), 277–289. <https://doi.org/10.1093/mp/sst118>
- Vakili, M., Ghamsari, M., & Rezaei, M. (2020). Performance analysis and comparison of machine and deep learning algorithms for IoT data classification. *arXiv preprint arXiv:2001.09636*. <https://doi.org/10.48550/arXiv.1909.02061>
- Van den Broeck, L., Schwartz, M. F., Krishnamoorthy, S., Tahir, M. A., Spurney, R. J., Madison, I., et al. (2022). Establishing a reproducible approach to study cellular functions of plant cells with 3D bioprinting. *Science Advances*, **8**(41), eabp9906. <https://doi.org/10.1126/sciadv.abp9906>
- Vijayan, A., Mody, T. A., Yu, Q., Wolny, A., Cerrone, L., Strauss, S., ... Kreshuk, A. (2024). A deep learning-based toolkit for 3D nuclei segmentation and quantitative analysis in cellular and tissue context. *Development*, **151**(14). <https://doi.org/10.1242/dev.202800>
- Wang, X.-F., Huang, D.-S., Du, J.-X., Xu, H., & Heutte, L. (2008). Classification of plant leaf images with complicated background. *Applied mathematics and computation*, **205**(2), 916–926. <https://doi.org/10.1016/j.amc.2008.05.108>
- Wang, Z., Yuan, H., Yan, J., & Liu, J. (2025). Identification, characterization, and design of plant genome sequences using deep learning. *The Plant Journal*, **121**(1), e17190. <https://doi.org/10.1111/tpj.17190>
- Wei, T., Chen, Z., Yu, X., Chapman, S., Melloy, P., & Huang, Z. (2024). Plantseg: A large-scale in-the-wild dataset for plant disease segmentation. *arXiv preprint arXiv:2409.04038*. <https://doi.org/10.48550/arXiv.2409.04038>
- Williamson, H. F., Brettschneider, J., Caccamo, M., Davey, R. P., Goble, C., Kersey, P. J., et al. (2023). Data management challenges for artificial intelligence in plant and agricultural research. *F1000Research*, **10**, 324. <https://doi.org/10.12688/f1000research.52204.2>
- Wolny, A., Cerrone, L., Vijayan, A., Tofanelli, R., Barro, A. V., Louveaux, M., et al. (2020). Accurate and versatile 3D segmentation of plant tissues at cellular resolution. *eLife*, **9**, e57613. <https://doi.org/10.7554/eLife.57613>
- Wongchai, A., Rao Jenjeti, D., Priyadarsini, A. L., Deb, N., Bhardwaj, A., & Tomar, P. (2022). Farm monitoring and disease prediction by classification based on deep learning architectures in sustainable agriculture. *Ecological modelling*, **474**, 110167. <https://doi.org/10.1016/j.ecolmodel.2022.110167>
- Wu, Z., Jiang, F., & Cao, R. (2022). Research on recognition method of leaf diseases of woody fruit plants based on transfer learning. *Scientific Reports*, **12**(1), 15385. <https://doi.org/10.1038/s41598-022-18337-y>
- Wu, Y., Wen, W., Gu, S., Huang, G., Wang, C., Lu, X., et al. (2024). Three-dimensional Modeling of maize canopies based on computational intelligence. *Plant Phenomics*, **6**, 0160. <https://doi.org/10.34133/plantphenomics.0160>
- Xing, D., Wang, Y., Sun, P., Huang, H., & Lin, E. (2023). A CNN-LSTM-att hybrid model for classification and evaluation of growth status under drought and heat stress in chinese fir (*Cunninghamia lanceolata*). *Plant Methods*, **19**(1), 66. <https://doi.org/10.1186/s13007-023-01044-8>
- Xiong, J., Yu, D., Liu, S., Shu, L., Wang, X., & Liu, Z. (2021). A review of plant phenotypic image recognition technology based on deep learning. *Electronics*, **10**(1), 81. <https://doi.org/10.3390/electronics10010081>
- Yao, J., Tran, S. N., Garg, S., & Sawyer, S. (2024). Deep learning for plant identification and disease classification from leaf images: Multi-prediction approaches. *ACM Computing Surveys*, **56**(6), 1–37. <https://doi.org/10.48550/arXiv.2310.16273>
- Yu, R., Luo, Y., Zhou, Q., Zhang, X., Wu, D., & Ren, L. (2021). Early detection of pine wilt disease using deep learning algorithms and UAV-based multispectral imagery. *Forest Ecology and Management*, **497**, 119493. <https://doi.org/10.1016/j.foreco.2021.119493>
- Yuan, Y., Chen, L., Ren, Y., Wang, S., & Li, Y. (2022). Impact of dataset on the study of crop disease image recognition. *International Journal of Agricultural and Biological Engineering*, **15**(5), 181–186. <https://doi.org/10.25165/j.ijabe.20221505.7005>
- Zhang, J., Zhang, D., Cai, Z., Wang, L., Wang, J., Sun, L., ... Zhao, J. (2022). Spectral technology and multispectral imaging for estimating the photosynthetic pigments and SPAD of the Chinese cabbage based on machine learning. *Computers and Electronics in Agriculture*, **195**, 106814. <https://doi.org/10.1016/j.compag.2022.106814>
- Zhang, C., Liu, Y., & Tie, N. (2023a). Forest land resource information acquisition with sentinel-2 image utilizing support vector machine, K-nearest neighbor, random forest, decision trees and multi-layer perceptron. *Forests*, **14**(2), 254. <https://doi.org/10.3390/f14020254>
- Zhang, N., Zhou, X., Kang, M., Hu, B.-G., Heuvelink, E., & Marcelis, L. F. (2023b). Machine learning versus crop growth models: An ally, not a rival. *AoB Plants*, **15**(2), plac061. <https://doi.org/10.1093/aobpla/plac061>
- Zhang, P., Sun, X., Zhang, D., Yang, Y., & Wang, Z. (2023c). Lightweight deep learning models for high-precision rice seedling segmentation from UAV-based multispectral images. *Plant Phenomics*, **5**, 0123. <https://doi.org/10.34133/plantphenomics.0123>
- Zhang, X., Ibrahim, Z., Khaskheli, M. B., Raza, H., Zhou, F., & Shamsi, I. H. (2024). Integrative approaches to abiotic stress management in crops: Combining bioinformatics educational tools and artificial intelligence applications. *Sustainability*, **16**(17), 7651. <https://doi.org/10.3390/su16177651>
- Zhou, J., Li, J., Wang, C., Wu, H., Zhao, C., & Teng, G. (2021). Crop disease identification and interpretation method based on multimodal deep learning. *Computers and Electronics in Agriculture*, **189**, 106408. <https://doi.org/10.1016/j.compag.2021.106408>
- Zhou, N., Siegel, Z. D., Zarecor, S., Lee, N., Campbell, D. A., Andorf, C. M., et al. (2018a). Crowdsourcing image analysis for plant phenomics to generate ground truth data for machine learning. *PLoS Computational Biology*, **14**(7), e1006337. <https://doi.org/10.1371/journal.pcbi.1006337>
- Zhou, Y., Yan, A., Han, H., Li, T., Geng, Y., Liu, X., & Meyerowitz, E. M. (2018b). HAIRY MERISTEM with WUSCHEL confines

- CLAVATA3 expression to the outer apical meristem layers. *Science*, **361**(6401), 502–506. <https://doi.org/10.1126/science.aar8638>
- Zhou, X., Han, W., Dai, J., Liu, S., Gao, S., Guo, Y., ... Zhu, X.** (2023). SPA, a stigma-style-transmitting tract physical microenvironment assay for investigating mechano-signaling in pollen tubes. *Proceedings of the National Academy of Sciences*, **120**(49), e2314325120. <https://doi.org/10.1073/pnas.2314325120>
- Zhu, L., Li, Z., Li, C., Wu, J., & Yue, J.** (2018). High performance vegetable classification from images based on alexnet deep learning model. *International Journal of Agricultural and Biological Engineering*, **11**(4), 217–223. <https://doi.org/10.25165/IJABE.V11I4.2690>
- Zhu, W., Braun, B., Chiang, L. H., & Romagnoli, J. A.** (2021). Investigation of transfer learning for image classification and impact on training sample size. *Chemometrics and Intelligent Laboratory Systems*, **211**, 104269. <https://doi.org/10.1016/j.chemolab.2021.104269>