CURRENT TRENDS IN AI TOOLS DEVELOPMENT IN PLANT BIOLOGY
CONFERENCE SUBMISSIONS

Anonymous authors
Paper under double-blind review

ABSTRACT

AI tools have revolutionized plant biology and agrobiology by enabling high-throughput analysis of complex biological data, from genomes to phenotypes, supporting precision breeding and sustainable agriculture. We highlight key applications in bioimage analysis using YOLO algorithms, crop plant genome analysis, and computational high-throughput phenotyping tools, discussing recent advancements. Classical genetic selection demands long time and labor-intensive opening way to novel computational modeling. Modern omics technologies, high-throughput genomics, high-throughput plant phenotyping, development of remote sensing devices on-farm provided vast amount of data available in databases. Such platforms and associated data generation have contributed to a booming AI industry in agriculture. Computer methods for plant genome data analysis oriented on stress resistance and crop yield rely on AI approaches including pattern recognition, neural networks, and LLMs. As an example of the development of AI tools for crop bioinformatics, we present a collaborative project between Russia and China "Smart Crop - Cognitive Platform for Reconstruction, Visualization and Analysis of Stress Response Networks based on ANDSystem and Multiomics in Rice and Wheat". The study of molecular genetic mechanisms of plant resistance to unfavorable biotic and abiotic factors (high or low temperature, drought, salinity, soil diseases, pathogens and pests) requires the study of the functioning of entire molecular genetic systems, including complex signaling, regulatory, transport and metabolic pathways. Finally, we discuss the challenges and current trends of AI applications in computational plant biology.

1 INTRODUCTION. PLANT GENOME ANALYSIS AND PROJECT 'SMART CROP'

Artificial intelligence (AI) is revolutionizing plant sciences by enabling accurate identification of plant species, early disease diagnosis, prediction of crop yield, and precision agriculture optimization (Gupta et al., 2024; Khan et al., 2025; Kassem, 2025). AI applications in plant research and agriculture have benefited large-scale industrial farming, with RD investment focused on commodity crops such as wheat, rice, and maize, and high-value horticulture crops such as soft fruits. The challenge is related to its importance for global food security. Extreme environmental impacts, including those caused by adverse climate changes, plant diseases and insect pests, lead to significant crop losses. AI-driven multi-omics integrates genomics, transcriptomics, and metabolomics for crop improvement, using deep learning to predict traits like yield and stress resistance (Javaid et al., 2023). Models such as convolutional neural networks analyze pan-genomes in crops like wheat and maize, identifying QTLs (Quantitative Trait Loci) for seed quality and accelerating breeding. Recent reviews highlight AI's role in plant genomics, where ML dissects regulatory networks and enhances genomic selection for climate-resilient varieties (Eftekhari et al., 2024; Khan et al., 2025).

Within the framework of the project, using artificial intelligence methods, a cognitive software and information platform "Smart Crop" was created, which provides a full cycle of knowledge engineering in the field of plant biology for the reconstruction and analysis of gene networks, focused on solving problems of studying the molecular genetic mechanisms of the genotype-phenotype relationship - medium for agriculturally valuable crops of rice and wheat (Cheng et al., 2024). The platform is focused on solving such substantive tasks as interpreting the results of omics data (establishing links between gene sets and biological processes, phenotypic traits, etc.); reconstruction of gene networks describing the relationship between molecular genetic objects and objects corresponding to the concepts of selection, phenomics and seed production, phytopathology, diagnostics, means of protection; identification of regulatory and signaling pathways of plant response to specific growth conditions, biotic and abiotic stresses (high or low temperature, drought, salinity, soil contamination with metals, response to fertilizers, the action of hormones, etc.); prediction of candidate genes for genotyping; search for markers for marker-based selection; candidate targets for the directed action of substances (including external factors) on plants to solve the problems of early/uniform emergence, better vegetative growth, efficient absorption of nutrients (Kleshchev et al., 2024; Antropova et al., 2024).

2 CROP PLANT GENOME ANALYSIS

Extreme environmental effects, diseases and insect pests result in significant crop loss. Therefore, more researchers turn their attention to the study of crop resistance. Earlier present, research on single stress-responsive genes in crops has been performed. To date, multiple studies have systematically revealed a part of molecular regulatory network in rice and wheat (Li et al., 2025). Recently, Artificial intelligence (AI) has opened new possibilities for the systematic analysis of molecular networks. Text-mining is a field in AI that aims to extract information from collections of text documents based on machine learning and natural language processing (NLP) techniques. Text mining is considered a useful tool for integrative biological research involving genes, proteins and phenotypes. In this project we plan to apply ANDSystem - bioinformatics tool that builds molecular networks by text-mining and data-mining of PubMed indexed publications and multiple omics by Russian partners (Ivanisenko, et al., 2020; Demenkov, et al., 2011; Ivanisenko, et al., 2022), for reconstruction and analysis of molecular networks under specific growth conditions, biotic and abiotic stresses in rice and wheat.

Therefore, the “Smart Crop” cognitive software and information platform was created, providing a full cycle of knowledge engineering about the genotype-phenotype-environment relationships in the field of agriculturally valuable rice and wheat crop research using AI and text-mining methods. The platform included three modules: (1) a module for automated extraction of knowledge from the texts of scientific publications, patents, databases; (2) a knowledge base containing extracted information in the form of a knowledge graph; (3) a module designed for the reconstruction and analysis of gene networks that describe the molecular mechanisms, regulatory and signaling pathways of a plant’s response to specific growth conditions, biotic and abiotic stresses (Demenkov et al., 2025).

Non-coding RNA (ncRNA) study is an important research field in life sciences. ncRNAs are involved in various biological processes and molecular regulations. Based on AI applications on plant ncRNAs and their regulatory network construction, we explored high-throughput sequencing datasets such as Ribo-seq, RNA-seq, sRNA-seq and other genomic/epigenomic data, to develop and improve molecular networks under stress response in rice and wheat (Shen et al., 2024; Zhou et al., 2025).

3 BIOIMAGE ANALYSIS WITH YOLO (YOU ONLY LOOK ONCE) MODELS

YOLO (You Only Look Once) models, known for real-time object detection, have been adapted for plant bioimaging to identify diseases, pests, and structures with high speed and accuracy (Alhwaiti et al., 2025). YOLOv3 and YOLOv4 excel in localizing diseased areas on leaves, achieving superior mAP, precision, and recall compared to traditional methods, aiding early intervention in crops like fruits. Recent 2025 studies demonstrate YOLO’s integration with multispectral imaging for non-invasive trait extraction, enhancing biodiversity monitoring and precision agriculture.

New approaches for detecting plant diseases, based on image processing, Machine Learning (ML), Deep Learning (DL) techniques, and hyperspectral imaging have been developed as a result of technological advancements (Kaur et al., 2025). For instance, the applications included multi-class SVM for detecting three distinct classes of apple diseases, a feed-forward back propagation neural network for detecting mildew from grape leaves, machine learning techniques for the prediction

Table 1: Recent works in AI – driven plant bioinformatics (2024-2026)

Area	**Key contribution**	**Publication**
Bioimage	Real-time detection	Alhwaiti et al. 2025
Genome	Deep networks for omics	Syeda, 2025
Inter-disciplinary	GPT model for plants	Zhang et al., 2025
Phenotyping	Synthetic data for trait algorithms	Roggiolani et al., 2025
Genome/ Phenotype	Multi-omics for breeding values	Kassem, 2025
Phenotype	Phenotyping in field using drones	Lu et al., 2025
Inter-disciplinary	Plant disease diagnosis	Sheikh et al., 2024
Phenotyping	Yield prediction	Pugh et al., 2024

of crop yield. DL-based approaches for detecting plant diseases have attracted growing interest in recent years due to their capacity to acquire complex representations from vast amounts of data.

YOLO-LeafNet approach is proposed for detecting diseases from leaf images of four distinct species, namely, grape, bell pepper, corn, and potato. About 9 thousand leaf images have been acquired for this work from five different publicly available datasets on Kaggle. All the acquired images were pre-processed by applying four different image pre-processing operations. The number of images in the training dataset was tripled for better model performance by applying five different augmentation operations.

4 PLANT PHENOTYPE ESTIMATION TOOLS

Computer vision tools automate phenotyping, quantifying growth, morphology, and physiology from images and 3D models (Kassem, 2025). Recent generative AI creates synthetic 3D leaf point clouds for trait estimation, addressing data scarcity and improving yield predictions (Plant Phenomics, 2025). Platforms like those using CNNs enable field-scale analysis via drones, modeling genotype-phenotype links for nowcasting and forecasting. Tools such as scKAN and AI phenomics pipelines integrate multi-omics for precise seed trait mapping

The study by Roggiolani and colleagues (Roggiolani et al., 2025) introduces a generative model capable of producing lifelike 3D leaf point clouds with known geometric traits, accelerating crop improvement and optimize yield predictions through data-driven modeling.

The research team trained a 3D convolutional neural network to learn how to generate realistic leaf structures from skeletonized representations of real leaves. Using datasets from sugar beet, maize, and tomato plants, they extracted the “skeleton” of each leaf—the petiole and main and lateral axes that define its shape—and then expanded these skeletons into dense point clouds using a Gaussian mixture model. The neural network, designed as a 3D U-Net architecture, predicts per-point offsets to reconstruct the complete leaf shape while maintaining its structural traits.

AI in plant breeding faces significant hurdles like data scarcity and model interpretability but holds promise for predictive breeding and climate resilience. Recent reviews emphasize addressing these to unlock faster genetic gains (Plant Phenomics, 2025).

5 FIELD SURVEILLANCE USING DRONES

Digital phenotyping using drones (Unmanned Aerial Vehicles - UAVs) enables high-throughput, non-destructive trait assessment in breeding plots, accelerating selection for yield, stress tolerance, and maturity (Eftekhari et al., 2024). These applications integrate multispectral/RGB imaging with AI for scalable field data collection. Drones quantify key traits like plant height, biomass, stand count, canopy cover, and relative maturity with 90 percent accuracy, replacing manual labor in crops such as maize, soybean, and dry beans. RGB and multispectral sensors capture vegetation indices (e.g., NDVI) at multiple growth stages, strongly correlated with agronomic performance, such as yield and height.

UAVs detect pests, plant diseases and abiotic stresses via hyperspectral imaging, enabling early selection of resistant genotypes in breeding nurseries. High-resolution imagery identifies symptoms at the plot level, supporting genomic selection for traits such as drought tolerance through digital twins.

For example, in soybean breeding, drone HTP refines genotype selection via vegetation indices, optimizing research cycles and reducing costs. Cost-effective solutions like VITO's allow in-house deployment for large-scale trials, enhancing precision in variety development. Drones bridge genomics-phenomics gaps, simulating analysis of genotype-environment interactions for predictive breeding.

6 KEY CHALLENGES OF AI IN PLANT BIOLOGY

Data management is a primary barrier, with issues in standardization, quality, and availability of high-throughput phenomics and multi-omics datasets, especially for orphan crops (Williamson et al., 2023). High computational demands, biases in training data, and lack of field validation hinder model reliability and scalability. Interpretability remains elusive, as ”black-box” AI struggles with explaining predictions for regulatory approval, while ethical concerns like inequality in access and GMO risks persist.

Note current challenges in AI applications in agrobiology:

Data scarcity/quality

Federated learning, synthetic data generation

Computational cost

Field-to-lab

Hybrid AI-process-based models will integrate mechanistic simulations with ML for better extrapolation to new environments and traits. Advances in high-throughput phenotyping via drones, robotics, and digital twins promise real-time, scalable trait assessment. AI-driven gene editing (e.g., CRISPR optimization) and predictive simulations could shorten breeding cycles to months, enabling proactive designs for stress tolerance.

ACKNOWLEDGMENTS

The work was supported by Russian Science Foundation and NSFC.

REFERENCES

1. Alhwaiti Y, Khan M, Asim, M. et al. Leveraging YOLO deep learning models to enhance plant disease identification. Sci Rep 2025; 15, 7969. doi: 10.1038/s41598-025-92143-0

2. Antropova EA, Volyanskaya AR, Adamovskaya AV, Demenkov PS, Yatsyk IV, Ivanisenko TV, Orlov YL, Haoyu Ch, Chen M, Ivanisenko VA. Computational identification of promising genetic markers associated with molecular mechanisms of reduced rice resistance to Rhizoctonia solani under excess nitrogen fertilization using gene network reconstruction and analysis methods. Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding 2024;28(8):960-973. doi 10.18699/vjgb-24-103

3. Cheng M, Zhu Y, Yu H, Shao L, Zhang Y, Li L, Tu H, Xie L, Chao H, Zhang P, Xin S, Feng C, Ivanisenko V, Orlov Y, Chen D, Wong A, Yang YE, Chen M. Non-coding RNA notations, regulations and interactive resources. Functional and Integrative Genomics 2024;24(6):217. doi: 10.1007/s10142-024-01494-w

4. Demenkov PS, Ivanisenko TV, Kolchanov NA, Ivanisenko VA. ANDVisio: a new tool for graphic visualization and analysis of literature mined associative gene networks in the ANDSystem. In Silico Biol. 2011-2012;11(3-4):149-61. doi: 10.3233/ISB-2012-0449

5. Demenkov PS, Ivanisenko TV, Kleshchev MA, Antropova EA, Yatsyk IV, Volyanskaya AR, Adamovskaya AV, Maltseva AV, Venzel AS, Chao H, Chen M, Ivanisenko VA. SmartCrop: knowl-

edge base of molecular genetic mechanisms of rice and wheat adaptation to stress factors. Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov J Genet Breed. 2025;29(8):1221-1234. doi 10.18699/vjgb-25-129

6. Eftekhari M, Ma C, Orlov Y.L. Editorial: Applications of artificial intelligence, machine learning, and deep learning in plant breeding. Front. Plant Sci. 2024; 15:1420938. doi: 10.3389/fpls.2024.1420938

7. Gupta DK, Pagani A, Zamboni P, Singh AK. AI-powered revolution in plant sciences: advancements, applications, and challenges for sustainable agriculture and food security. Explor Foods Foodomics. 2024;2:443–59. doi: 10.37349/eff.2024.00045

8. Ivanisenko TV, Demenkov PS, Kolchanov NA, Ivanisenko VA. The New Version of the ANDDigest Tool with Improved AI-Based Short Names Recognition. Int J Mol Sci. 2022; 23(23):14934. doi: 10.3390/ijms232314934

9. Ivanisenko TV, Saik OV, Demenkov PS, Ivanisenko NV, Savostianov AN, Ivanisenko VA. ANDDigest: a new web-based module of ANDSystem for the search of knowledge in the scientific literature. BMC Bioinformatics. 2020;21(Suppl 11):228. doi: 10.1186/s12859-020-03557-8

10. Javaid M, Haleem A, Khan IH, Suman R. Understanding the potential applications of Artificial Intelligence in Agriculture Sector. Adv Agron. 2023;2:15–30. doi: 10.1016/j.aac.2022.10.001

11. Kassem MA. Harnessing Artificial Intelligence and Machine Learning for Identifying Quantitative Trait Loci (QTL) Associated with Seed Quality Traits in Crops. Plants (Basel). 2025 Jun 5;14(11):1727. doi: 10.3390/plants14111727. 12. Kaur R, Mittal U, Wadhawan A. et al. YOLO-LeafNet: a robust deep learning framework for multispecies plant disease detection with data augmentation. Sci Rep 15, 28513 (2025). doi: 10.1038/s41598-025-14021-z 13. Khan I, Khare BK. Integrating AI in plant science: A review of applications and future prospects. Plant Gene 2025; 44, 100542. doi: 10.1016/j.plgene.2025.100542

14. Kleshchev MA, Maltseva AV, Antropova EA, Demenkov PS, Ivanisenko TV, Orlov YL, Chao H, Chen M, Kolchanov NA, Ivanisenko VA. Reconstruction and computational analysis of the microRNA regulation gene network in wheat drought response mechanisms. Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding. 2024;28(8):904-917. doi: 10.18699/vjgb-24-98

15. Li F, Chen J, Luo W, Liu J, Chen G, Shuai B, Hou Z, Gan Z, Zhao H, Zhan P, Bi C, Wang Z, Hu H, Wang S. DeepPGDB: A novel paradigm for AI-guided interactive plant genomic database. Plant Commun. 2025;6(12):101494. doi: 10.1016/j.xplc.2025.101494

16. Lu J, Qi Q, Zheng G, Eitel JUH, Zhang Q, Zhang J, Chen S, Zhang F, Fang W, Guan Z, Chen F. High-Throughput Field Phenotyping Using Unmanned Aerial Vehicles (UAVs) for Rapid Estimation of Photosynthetic Traits. Plant Phenomics. 2025;7(2):100045. doi: 10.1016/j.plaphe.2025.100045

17. Pugh NA, Young A, Ojha M, Emendack Y, Sanchez J, Xin Z, Puppala N. Yield prediction in a peanut breeding program using remote sensing data and machine learning algorithms. Front Plant Sci. 2024;15:1339864. doi: 10.3389/fpls.2024.1339864

18. Roggiolani G, Bailey B.N, Behley J, Stachniss C. Generation of labeled leaf point clouds for plants trait estimation. Plant Phenomics 2025; 7(3):100071. doi: 10.1016/j.plaphe.2025.100071

19. Sheikh M, Iqra F, Ambreen H, Pravin K.A, Ikra M, Chung Y.S. Integrating artificial intelligence and high-throughput phenotyping for crop improvement. Journal of Integrative Agriculture 2024; 23(6):1787-1802. doi: 10.1016/j.jia.2023.10.019

20. Shen Y, Liu L, Liu E, Li S, Orlov Y, Ivanisenko V, Chen M. AthRiboNC: an Arabidopsis database for ncRNAs with coding potential revealed from ribosome profiling. Database (Oxford). 2024;2024:baae123. doi: 10.1093/database/baae123.

21. Syeda A. Harnessing multi-omics and genome-editing technologies for climate-resilient agriculture: bridging AI-driven insights with sustainable crop improvement. Plant Mol Biol. 2025;115(6):120. doi: 10.1007/s11103-025-01650-1

22. Williamson HF, Brettschneider J, Caccamo M, Davey RP, Goble C, Kersey PJ, May S, Morris RJ, Ostler R, Pridmore T, Rawlings C, Studholme D, Tsaftaris SA, Leonelli S. Data management

challenges for artificial intelligence in plant and agricultural research. F1000Res. 2021;10:324. doi: 10.12688/f1000research.52204.2.

23. Wu H, Han R, Zhao L, Liu M, Chen H, Li W, Li L. AutoGP: An intelligent breeding platform for enhancing maize genomic selection. Plant Commun. 2025; 6(4):101240. doi: 10.1016/j.xplc.2025.101240

24. Zhang R, Wang Y, Yang W, Wen J, Liu W, Zhi S, Li G, Chai N, Huang J, Xie Y, Xie X, Chen L, Gu M, Liu YG, Zhu Q. PlantGPT: An Arabidopsis-Based Intelligent Agent that Answers Questions about Plant Functional Genomics. Adv Sci (Weinh). 2025; 12(30):e03926. doi: 10.1002/advs.202503926

25. Zhou X, Fan H-Y, Feng X-Y, Ruan Z, Yuan J, Han Q, He Z, You Y, Chao H, Chen M, Shao Z-Q, Xue J-Y, Chen D. PGCP: A comprehensive database of plant genomes for comparative phylogenomics. Plant Biotechnol. J 2025; 23: 2928-2930. doi: 10.1111/pbi.70110