RESEARCH ARTICLE

Rapid breeding framework of flower disease resistance based on deep learning and genome wide association study

Yuting Xiao[1, *], Shanmei Han[1], Kandratau Yauhen Valerievich[2], Belavusava Natalia Lenidovna[2]

[1]Luoyang Polytechnic, Luoyang, Henan, China. [2]Central Botanical Garden of the National Academy of Sciences of Belarus, Minsk, Belarus.

Received: January 14, 2026; **accepted:** May 19, 2026.

Flower breeding for disease resistance has long been constrained by lengthy field screening cycles, typically spanning 3 to 5 years, as well as by subjective manual phenotyping. To address these challenges, this study developed a rapid breeding framework integrating deep learning image recognition, genome wide association study (GWAS), and a fusion prediction model using phenotypic, image, and genotype data collected from three consecutive years of 2022 to 2024 under both greenhouse and open field conditions. A disease identification model based on convolutional neural networks was established, enabling automatic extraction of lesion features and improving the objectivity of phenotypic evaluation. The results showed that GWAS identified several significant genetic regions associated with resistance, providing insights into resistance mechanisms. The fusion model, which combined visual features from images and genomic information from GWAS significant loci, achieved superior prediction accuracy with F1 of 0.94, area under the curve (AUC) of 0.97, and demonstrated stable performance in cross year validation. These results indicated that multi source data fusion could significantly shorten the resistance screening period to within a single growing season and improve breeding efficiency. This framework offered a systematic, intelligent technical pathway for flower disease resistance breeding and supported early stage material screening and candidate gene mapping.

Keywords: deep learning; genome-wide association study; flower diseases; rapid breeding; disease resistance.

[*]**Corresponding author:** Yuting Xiao, Luoyang Polytechnic, Luoyang, Henan 471000, China. Email: ytjy6637@outlook.com.

Introduction

The flower industry plays an important role in ornamental gardening, urban and rural greening, and landscape engineering. Many common flowers are easily infected with fungal diseases such as powdery mildew, gray mold, and rust under the conditions of high temperature and high humidity or continuous rainfall. After infection, the photosynthetic efficiency of plants decreases, the deformity rate of flower organs increases, and the yield and ornamental quality decrease by 30 - 60% [1]. Traditional resistance breeding relies on field identification for many years, which takes 3 - 5 years to screen stable materials [2]. The occurrence of diseases is affected by climate fluctuation, and the phenotypic identification results are prone to deviation [3].

Recent advances in deep learning and high-throughput sequencing have provided new technical opportunities for efficient, quantitative, and automated resistance screening [4]. Deep learning has shown advantages in visual recognition tasks with

related technologies gradually extending to agricultural production and biological breeding research [5]. Aizenstein *et al*. suggested that deep learning had advantages in dealing with high-dimensional complex features, showing strong adaptability in multi-modal data modeling [1]. Estrada-Molina *et al*. systematically combed the deep learning model in the open learning environment and found that the convolutional network was stable in the image classification task [2]. AbuSalim *et al*. summarized the performance of deep learning in medical image diagnosis and indicated that convolution structure was sensitive to texture changes and helpful to identify the characteristics of subtle lesions [3]. Himeur *et al*. demonstrated the effect of transfer learning to improve the generalization ability of the model in the study of mask recognition in urban scenes [4]. Iqbal *et al*. emphasized the influence of data cleaning and labeling quality on model performance in cancer image classification [5]. In the agricultural field, deep learning has been used to identify growth stages, monitor diseases, and judge agronomic traits. Schneider *et al*. used YOLOv8 to classify the growth period of pepper, showing the robustness of the model in high-noise environment [6]. In terms of genetic research, Guintivano pointed out that cross-cohort GWAS integration coluld improve the stability of related signals [7]. Hellwege *et al*. emphasized the importance of population structure correction in multiancestor GWAS [8]. Poudel *et al*. confirmed the decisive role of genetic information in the selection of breeding materials in macadamia nut breeding research [9]. Liatukiene *et al*. focused on the synchronous monitoring of agronomic traits and disease development in alfalfa research [10], while Santanasto *et al*. showed that large-sample genetic element analysis could improve the analytical ability of traits [11]. Pachiyappan *et al*. reported that the controlled environment could improve the stability of phenotypic collection and provide a more reliable data base for deep learning model training [12].

Despite these advances, a critical gap remains. Few studies have integrated deep-learning-based image phenotyping, GWAS, and a fusion prediction model into a single operational framework for rapid flower disease resistance breeding. The traditional breeding cycle is long (3 - 5 years), manual scoring errors are large, and the synergy between visual disease features and genetic variation has not been systematically exploited to shorten the screening period. To meet the needs of rapid breeding of flower disease resistance, this study focused on the digitalization of disease phenotypes and established a deep learning recognition model covering the characteristics of lesion distribution, color change, and tissue wilting degree. In addition, the study focused on mining of resistance genes by using millions of SNP data to carry out GWAS, determining significant loci and candidate regulatory intervals. The research constructed a fusion framework, which combined the image phenotypic features and GWAS features into the resistance prediction model to form a scoring system that could be directly used for material optimization. The proposed framework provided a systematic technical path for flower disease resistance breeding and laid a foundation for the development of intelligent breeding methods. By enabling the screening process to be completed within a single growing season, it significantly shortened the breeding cycle, improved the efficiency of resistance gene mining, and provided accurately matched resistant materials for different cultivation areas. The approach broke through the limitation of long traditional breeding cycle and large manual interpretation error, promoting the development of flower resistance breeding in an intelligent, repeatable, and extensible direction.

Materials and methods

Data collection

Phenotypic data came from the flower disease identification experiment conducted from 2022 to 2024 at the Institute of Ornamental Horticulture, Fujian Agriculture and Forestry University, Fuzhou, Fujian, China in the

greenhouse and the open field adjacent to the greenhouse (26°05'N, 119°14'E). Three types of flower diseases were identified including powdery mildew (*Erysiphe cichoracearum*), gray mold (*Botrytis cinerea*), and rust (*Puccinia* spp.). Disease indices recorded included lesion area ratio, leaf color change, and tissue moisture decline with scoring intervals set at 0 - 9. The data of plant images, phenotypes, and genotypes were collected and stored under unified naming rules. Images and phenotypes were matched by unique plant numbers, while genotypes were linked to the original plant barcodes, forming a traceable data structure. The overall collection process was consistent at the time nodes with continuous recording completed within 7 - 14 days after disease onset, ensuring that phenotypic fluctuations remained stable**.**

Sample selection

Samples were selected from 18 major ornamental flower varieties including *Rosa hybrida* "Carola", *R*. *hybrida* "Avalanche", *Chrysanthemum morifolium* "Jinba", *C*. *morifolium* "Shenma", *Lilium brownii* "Siberia", *L*. *brownii* "Sorbonne", *Dianthus caryophyllus* "Master", *D*. *caryophyllus* "Pink Kiss", *Gerbera jamesonii* "Rosalin", *G*. *jamesonii* "Sunshine", *Hydrangea macrophylla* "Endless Summer", *H*. *macrophylla* "Magical", *Paeonia suffruticosa* "Luoyang Hong", *P*. *suffruticosa* "Er Qiao", *Rhododendron simsii* "Hongshan", *R*. *simsii* "Baihua", *Camellia japonica* "Shiqi", and *C*. *japonica* "Naidong". A total of 1,530 individuals were included with healthy plants accounting for about 40% and plants with mild, moderate, and severe natural infection accounting for 22%, 25%, and 13%, respectively. The sample layout used a randomized block design. Image data were collected under natural light and uniform supplementary light conditions by using a Canon EOS 5D Mark IV camera (Canon Inc., Tokyo, Japan) with 100 mm macro lens and 3,264 × 2,448 pixels resolution, covering 3 to 5 growth stages with at least 6 leaf images and 2 flower images recorded per plant, totaling 9,800 images. Genotype data were obtained from the same batch of plants by flash freezing fresh leaf samples at -80°C followed by genomic DNA extraction and sequencing using Illumina NovaSeq 6000 platform (Illumina, Inc., San Diego, CA, USA). The number of single nucleotide polymorphisms (SNPs) per plant ranged from 3.0M to 3.5M.

Data preprocessing and cleaning

Data preprocessing was carried out for phenotype, image, and genotype data to eliminate acquisition differences and noises, making input features expressed on a unified scale [13]. The image's brightness correction, contrast equalization, and edge enhancement were used to eliminate overexposed and out-of-focus images. All images were scaled to 256 × 256 pixels to adapt to the input structure of convolution network. Pixel-level normalization was performed with mean and standard deviation to reduce the offset caused by illumination difference as follows.

$$X' = \frac{X - \mu}{\sigma} \quad (1)$$

where *x* was the original pixel value. μ was the average value of pixels in the training set. σ was the standard deviation of pixels. The phenotypic data was checked by interval. The records inconsistent with the test time nodes were deleted, and the repeated observations of the same plant were averaged to reduce the subjectivity of manual scoring. Genotype data were screened by deletion rate and sequencing depth to generate high-quality SNP matrix, and the sample relationship matrix and allele frequency file were then constructed [14]. After preprocessing, the three types of data were summarized into the unified database after being numbered, and the conflict-free multi-source input could be directly called in the model construction stage. Data cleaning mainly focused on genotype quality control and systematically eliminated low-quality loci and abnormal samples. The SNP deletion rate was controlled within 10%, which meant that, when the number of missing points in the overall sample exceeded

the threshold, the index was calculated as below.

$$\mathrm{MissingRate}(s_i) = \frac{\mathrm{missing}(s_i)}{N} \quad (2)$$

where *N* was the total number of samples. $\mathrm{MissingRate}(s_i)$ was the missing number of loci s_i. To improve the stability of subsequent analysis, loci with allele frequency less than 0.05 did not participate in modeling to avoid weak information loci interfering with GWAS signals. Meanwhile, individuals whose genotype deletion rate was more than 15% or whose sequencing depth was less than 5× were taken as the elimination objects to ensure that there was no large gap in the input matrix [15]. The image records with extremely dark light, too miscellaneous background, or unrecognizable lesion areas were excluded. The abnormal values of phenotypes were tested by box chart, and the extreme observations that deviated from the distribution range were excluded. The cleaned data met the requirements of deep learning training and GWAS modeling in quantity, distribution, quality and provided a reliable database for the construction of a fusion resistance prediction framework.

Construction of deep learning disease identification model

The convolutional neural network (CNN) structure constructed in this study consisted of four convolutional modules and a fully connected layer. The image input size was fixed at 256 × 256 × 3. The first two convolutional layers had 32 and 64 channels with kernel size of 3 × 3 and stride of 1. The next two layers had 128 and 256 channels. Each convolutional module was followed by ReLU activation and a 2 × 2 max pooling layer. The fully connected layer output 128-dimensional features, which were then input into a Softmax layer to complete three-class disease identification. The network contained approximately 1.44 million trainable parameters. The CNN model was implemented by using PyTorch (version 1.13) (https://pytorch.org/) with the structural details as Conv1 of 256 × 256 × 32 and 896 parameters, Conv2 of 128 × 128 × 64 and 18,496 parameters, Conv3 of 64 × 64 × 128 and 73,856 parameters, Conv4 of 32 × 32 × 256 and 295,168 parameters, and FC layer of 128 and 1,048,576 parameters.

Genome-wide association study model

Mixed linear model (MLM) was employed at the genome level to handle background noise caused by population structure and kinship. The genome matrix consisted of 3.0M – 3.5M SNP loci. After filtering by minor allele frequency (MAF) with threshold of 0.05 and deletion rate with threshold of 10%, the allele frequency vector was constructed. The model took the disease phenotype score as the dependent variable and SNP genotype as the independent variable. The MLM was implemented using the GAPIT package (version 3.0) (https://zzlab.net/GAPIT/) in R (R Foundation for Statistical Computing, Vienna, Austria). The threshold for screening significant loci was set using suggestive significance as $P < 1 \times 10^{-5}$. The MLM was expressed as follows.

$$y = X\beta + Zu + \varepsilon \quad (3)$$

where y was the phenotypic vector. X was the fixed effect matrix. β was the fixed effect parameter. Z was the random effect matrix. u followed N(0, Kσ_g²). K was the kinship matrix. ε was the residual term.

Deep learning - GWAS fusion framework

The fusion framework was used to construct a resistance judgment system by using both phenotypic visual features and genetic locus information. The left channel input 128-dimensional image features extracted by CNN, while the right channel input the coding matrix of significant SNPs screened by GWAS. Features from both sides were concatenated in the middle layer and corrected by BatchNorm and Dropout (rate 0.5) and then sent to a fully connected layer to output the predicted resistance level. This structure related disease phenotype expression and functional site variation.

Training resistance prediction model

The resistance prediction model was trained by using deep learning and genome characteristics and maintained consistent logic in sample division, loss function setting, and optimization strategy. The samples were divided into training set, verification set, and test set according to the ratio of 7.0:1.5:1.5 to ensure the consistent distribution of different years and disease grades in the three types of datasets and avoid the model learning from being biased towards a certain environment or a certain disease performance [18]. The loss function adopted cross entropy, which was used to deal with the deviation between the prediction probability and the real label in the classification task and was expressed as below.

$$L = -\sum_{i=1}^{N} y_i \ln(\hat{y}_i) \quad (4)$$

where y_i was the real label. $\hat{y}_i$ was the model prediction probability. Adam algorithm was adopted in the optimization stage. The training used a batch size of 32, the learning rate of 1×10^{-4}, and early stopping strategy. The training process included four stages of feature input, loss calculation, gradient update, and performance evaluation of verification set. The whole training path formed a repeatable resistance evaluation system, which kept the prediction model stable in different environmental data.

Data division and training strategy

The distribution of samples was 6,860 images for training, 1,470 images for validation, 1,470 images for test, while 1,064 phenotypes for training, 228 for validation, 228 for test. A total of 1,530 SNP profiles were used across all splits without sample leakage, i.e., the same plant's image and genotype were kept in the same dataset split.

Model performance evaluation index

The accuracy rate, recall rate, and F1 value were employed as judging indicators. The area under the curve (AUC) was used as a reference for robustness in distribution intervals. The performance of two-class or multi-class models in different categories was expressed, and a comprehensive evaluation formula was introduced. In the classification task, the definition of accuracy was to predict the proportion of correct samples to all samples and was expressed as follows.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (5)$$

where *TP*, *TN*, *FP*, and *FN* corresponded to the number of correctly predicted positive samples, correctly predicted negative samples, negative samples incorrectly predicted as positive, and positive samples incorrectly predicted as negative, respectively. In the task of disease classification, the F1 value could more comprehensively express the performance of the model under the condition of unbalanced categories and was defined as follows.

$$\text{F1} = \frac{2PR}{P + R} \quad (6)$$

where *P* was the accuracy rate. *R* was the recall rate. To ensure the stability of the index, the model was calculated five times on the test set, and the average and standard deviation were recorded to reduce the random fluctuation caused by a single evaluation. All indicators were implemented in a unified evaluation environment to avoid deviations caused by different data distribution or calculation methods [20]. The model evaluation results were used to support the validity test of the fusion framework and served as an important reference standard for ablation experiments and generalization experiments.

Cross-validation and generalization ability analysis

A 50% cross-validation scheme was adopted, and the complete sample was randomly divided into

five equal subsets, four of which were used as the training set and the remaining one as the validation set in each round of training. The process was repeated for five rounds, and each subset went through the verification stage at least once. In addition to routine cross-validation, the generalization ability was studied and tested. Three annual samples in 2022, 2023, and 2024 were taken as independent test sets, and the performance of the model was quantified under unknown environmental conditions. The generalization ability was evaluated based on the accuracy of the test set and the change range of F1 value. If the value was within the range of 0.05 between the training set and the verification set, the model had strong generalization. The cross-environmental stability of the model was an important prerequisite for whether the resistance prediction system could be used in real breeding scenarios.

Ablation experiment

Three types of ablation schemes were designed in this research, which included image feature ablation, genetic feature ablation, and fusion structure ablation. The image feature ablation completely shielded CNN feature channels and kept GWAS salient points as input to test the performance decline of the model after losing visual phenotypic information [21]. The genetic feature ablation removed GWAS locus from the input vector, while CNN image features were used in the training to determine the enhancement degree of genetic information on model prediction. The fusion structure ablation replaced dual-channel splicing with single-channel feature stacking to test the advantages of standard fusion method in resistance prediction. Each type of ablation experiment was performed under the condition of unified data division and training parameters. The comparison results were not disturbed by training differences [22]. Accuracy and F1 were the main factors to determine the performance degradation. If the index dropped more than 0.05 after a module was removed, it indicated that the removed module made a significant contribution to the model performance.

Results

Disease image recognition

The recognition results showed high stability with the accuracy and recall between different diseases being kept above 0.85. Pink disease and rust disease had clear boundaries and stable texture expression, and the recognition results were concentrated in the range of 0.92 - 0.96. Gray mold and leaf spot were affected by light and tissue corrosion. The image characteristics changed greatly, making the F1 value slightly lower. Healthy leaves had the highest recognition performance because of uniform color distribution and consistent texture features. Overall, the model had a high generalization ability, and it still maintained a stable image feature extraction ability under complex background and different lighting conditions, which provided a reliable phenotypic feature input for the subsequent fusion model (Figure 1).

Significant SNPs of GWAS

A total of 2.9M SNP was entered into MLM calculation. The result demonstrated that significant loci were concentrated in areas related to immune regulation, pathogen identification, and signal transduction such as WRKY, LRR-RLP, and reactive oxygen species scavenging gene cluster. The range of related signals on Chr3 and Chr7 was narrow, indicating that their effects were concentrated and might correspond to core resistance genes. The association region of Chr11 segment was large, which reflected that disease resistance might be regulated by multiple genes. The range of phenotypic variation explained by significant loci was between 7% and 12%, which was consistent with the phenotypic performance of images and provided genetic information with clear biological significance for the construction of resistance scoring model (Figure 2).

Performance comparison of fusion models

The results showed that the fusion model was superior to the single-source model in all indicators, and the prediction results were more stable. CNN model could capture the texture

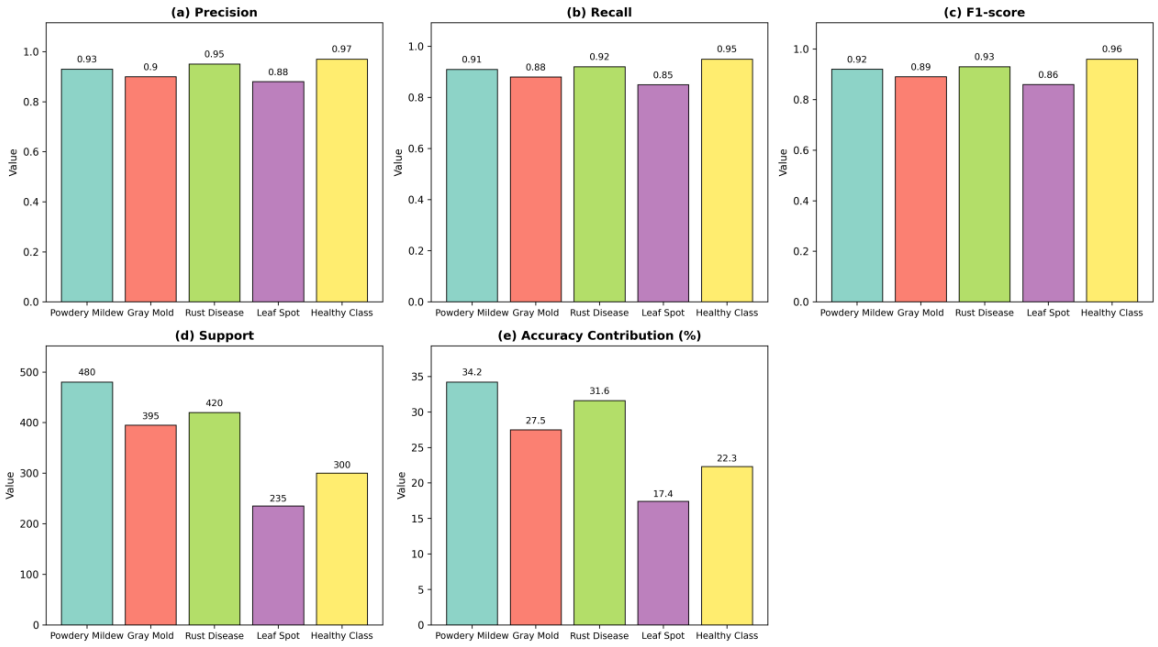

Figure 1. CNN disease image recognition performance.

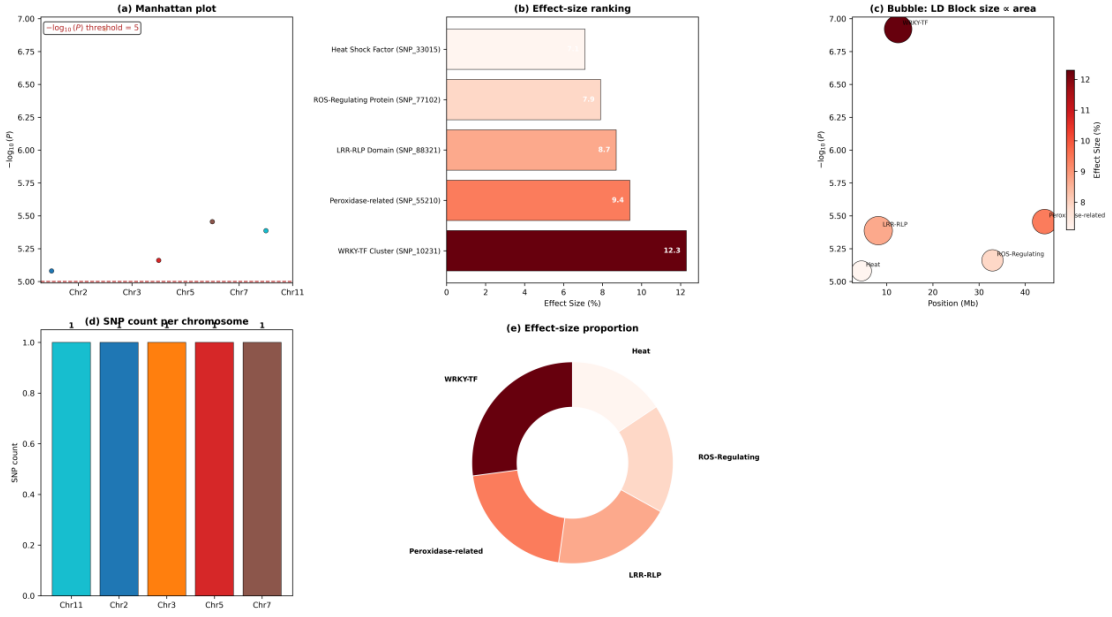

Figure 2. Significant SNP information map related to disease resistance.

information of disease phenotype and lacked the genetic explanation. The contribution of GWAS model to resistance scores was more obviously affected by sample size, and the overall performance was low. After two-channel fusion, the model could use both visual features and significant locus information with F1 value rising to 0.94 and AUC reaching 0.97, while the fluctuation range was lower, indicating that the fusion method had advantages in the task of resistance prediction (Figure 3).

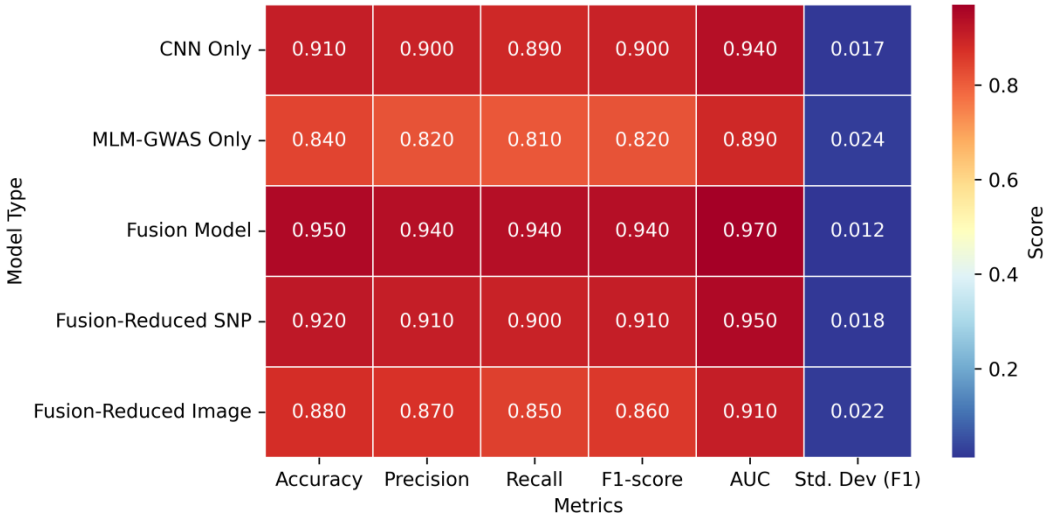

Figure 3. Thermal diagram of performance comparison of various models.

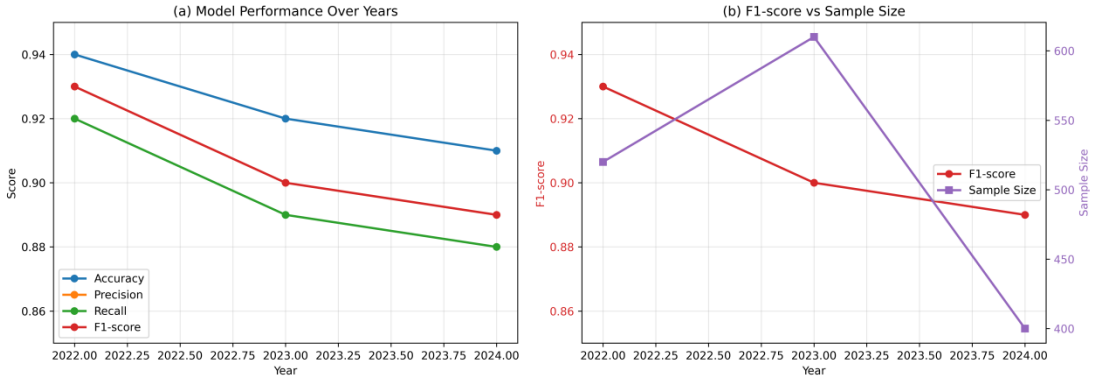

Figure 4. Comparison of model generalization results for data from different years.

Model generalization

The three-year test results showed that the model maintained stable accuracy in the year with obvious environmental fluctuation. The fluctuation range of F1 value was controlled in the range of 0.89 - 0.93. In 2023, due to the high rainfall, the fluctuation of leaf water content increased, which led to the blurred boundary between image texture characteristics and disease areas, and the prediction performance was slightly lower, which was still in an acceptable range. In 2024, the temperature changed greatly, resulting in the difference of disease development cycle. The recall rate of the model decreased slightly (Figure 4). The overall performance showed that the fusion model could maintain accuracy under the condition of strong environmental interference, which showed that it had application value in real breeding environment and adapted to the phenotypic variation of cross-year materials.

Discussion

The research results showed high stability in phenotypic digitization, genetic information analysis, and resistance prediction, reflecting the practical value of this method system in breeding processes. The deep learning model extracted fine lesion textures from complex backgrounds and multi-illumination conditions, reducing

manual interpretation errors. In traditional breeding, manual grading deviations often reach 1 – 2 classes, which affects early-stage material selection. The model compressed the dispersion of disease scores, making evaluation more consistent. The GWAS-detected loci clustered around several gene clusters including WRKY, LRR-RLP, ROS scavenging, providing explanatory evidence for the genetic basis of flower resistance. After fusing genetic features, prediction accuracy improved significantly, indicating strong coupling between phenotype and genotype. In terms of breeding efficiency, the fusion model enabled material screening to be completed within a single growing season, saving time and cost. The research framework was embeddable in different breeding stages. In early material screening, the deep learning model could quickly process tens of thousands of images. In mid-term resistance identification, GWAS results helped locate candidate gene regions, making subsequent validation more efficient. In later-stage material optimization, the fusion model provided resistance scores based on image features and significant loci, supporting rapid selection. For commercial applications, horticultural enterprises could use this model for disease monitoring and early warning.

The proposed model faced challenges from multiple environmental variables in actual deployment. Field illumination, background complexity, and plant density affected image texture information. Disease occurrence rhythms differed across regions, and lesion spread was influenced by temperature and humidity. Genotype data sources in field trials were more complex, affecting GWAS signal detection stability. Additionally, equipment configuration and image acquisition standards were not completely consistent across different institutions, requiring additional adaptation for cross-institution deployment. Future work will focus on model interpretability, feature enhancement, and cross-environment migration ability. The deep learning module needs improvement in lesion area localization, and providing heat maps would enhance its interpretive value for breeders. More environmental variables and time-series phenotypes should be added to construct multi-dimensional gene-environment interaction models. Transfer learning structures need to be designed to maintain model stability across different ecological regions. Graph neural networks may be introduced to more fully represent genome structure information. Overall, follow-up research should aim to improve generalization ability, enhance biological interpretation, and achieve automated breeding processes.

Focusing on the demand for rapid breeding of flower disease resistance, this study constructed a comprehensive framework consisting of deep learning phenotype recognition, genome-wide association study, and a fusion prediction model. The system formed a complete technical chain at three levels of phenotypic digitization, genetic signal analysis, and resistance prediction. The deep learning model stably identified disease characteristics under different lighting conditions and complex backgrounds. GWAS detected significant loci reflecting the genetic basis of resistance traits in immune regulation and pathogen recognition pathways. The fusion model integrated image features with genetic variation information, significantly improving prediction accuracy and showing strong generalization ability across years and environments. The results demonstrated that multi-source data fusion could complete material screening in a single season, making the resistance breeding process more efficient. The system demonstrated application potential in early material screening, candidate gene interval location, and final optimization, serving as an auxiliary decision-making tool for flower breeding units.

References

1. Aizenstein H, Moore RC, Vahia I, Ciarleglio A. 2024. Deep learning and geriatric mental health. Am J Geriatr Psychiatry. 32(3):270-279.

2. Estrada-Molina O, Mena J, López-Padrón A. 2024. The use of deep learning in open learning: a systematic review (2019 to 2023). Int Rev Res Open Dis. 25(3):370-393.
3. AbuSalim S, Zakaria N, Islam MR, Kumar G, Mokhtar N, Abdulkadir SJ. 2022. Analysis of deep learning techniques for dental informatics: A systematic literature review. Healthcare. 10(10):1892.
4. Himeur Y, Al-Maadeed S, Varlamis I, Al-Maadeed N, Abualsaud K, Mohamed A. 2023. Face mask detection in smart cities using deep and transfer learning: Lessons learned from the COVID-19 pandemic. Systems. 11(2):107.
5. Iqbal MS, Ahmad W, Alizadehsani R, Hussain S, Rehman R. 2022. Breast cancer dataset, classification and detection using deep learning. Healthcare. 10(12):2395.
6. Schneider F, Swiatek J, Jelali M. 2024. Detection of growth stages of chili plants in a hydroponic grower using machine vision and YOLOv8 deep learning algorithms. Sustainability. 16(15):6420.
7. Guintivano J. 2023. Meta-analyses of genome-wide association studies for postpartum depression. Am J Psychiatry. 180(12):884-895.
8. Hellwege JN, Kim J, Williams A, Noh H, Jones SH, Ruiz-Narváez EA, *et al*. 2023. Cross-ancestry genome-wide association meta-analyses of uterine fibroids. J Womens Health. 32(11):A14.
9. Poudel PD, Cowan M, Shaw L, De Faveri J, Topp B, Alam M. 2023. Macadamia breeding for reduced plant vigor: Progress and prospects for profitable and sustainable orchard systems. Sustainability. 15(19):14506.
10. Liatukiene A, Norkeviciene E, Danyte V, Liatukas Z. 2023. Diversity of agro-biological traits and development of diseases in alfalfa cultivars during contrasting vegetation seasons. Sustainability. 15(2):1445.
11. Santanasto A, Kiel D, Karasik D, Cvejkus R, Biggs M, Lunetta K, *et al*. 2022. Genome-wide association meta-analysis of physical performance in older adults. Innov Aging. 6(Suppl 1):214-215.
12. Pachiyappan P, Kumar P, Reddy KV, Kumar KNR, Konduru S, Paramesh V, *et al*. 2022. Protected cultivation of horticultural crops as a livelihood opportunity in western India: An economic assessment. Sustainability. 14(12):7430.
13. Gawde U, Hegde P, Desai K, Barai RS, Shekhar BR, Das DK, *et al*. 2022. Multimorbidity landscape of schizophrenia: Insights from meta-analysis of genome-wide association studies. Schizophrenia Res. 243:214-216.
14. Johnson EC, Demontis S, Thorgeirsson TE. 2020. A large-scale genome-wide association study meta-analysis of cannabis use disorder. Lancet Psychiatry. 7(11):1032-1045.
15. Ma LQ, Kong FQ, Sun K, Wang T, Guo T. 2021. From classical radiation to modern radiation: past, present, and future of radiation mutation breeding. Front Public Health. 9:768071.
16. Zhang MJ, Dong R, Wang XX. 2021. Plants with health risks undermine residents' perceived health status, evaluations and expectations of residential greenery. Landsc Urban Plan. 216:104236.
17. Havers L, Cardno A, Freeman D, Ronald A. 2021. Associations between the subdomains of negative symptoms in the general population and genome-wide polygenic scores for major depressive disorder and schizophrenia. Behav Genet. 51(6):709.
18. Melese YY, Kolech SA. 2021. Coffee (*Coffea arabica* L.): Methods, objectives, and future strategies of breeding in Ethiopia—Review. Sustainability. 13(19):10814.
19. Riehl C. 2021. Evolutionary origins of cooperative and communal breeding: Lessons from the crotophagine cuckoos. Ethology. 127(10):827-836.
20. Lenaerts B, de Mey Y, Demont M. 2022. Revisiting multi-stage models for upstream technology adoption: Evidence from rapid generation advance in rice breeding. J Agric Econ. 73(1):277-300.
21. Ceballos H, Hershey C, Iglesias C, Zhang XF. 2021. Fifty years of a public cassava breeding program: Evolution of breeding objectives, methods, and decision-making processes. Theor Appl Genet. 134(8):2335-2353.
22. Kumar M, Rawat S, Nagar B, Kumar A, Pala NA, Bhat JA, *et al*. 2021. Implementation of the use of ethnomedicinal plants for curing diseases in the Indian Himalayas and its role in sustainability of livelihoods and socioeconomic development. Int J Environ Res Public Health. 18(4):1509.