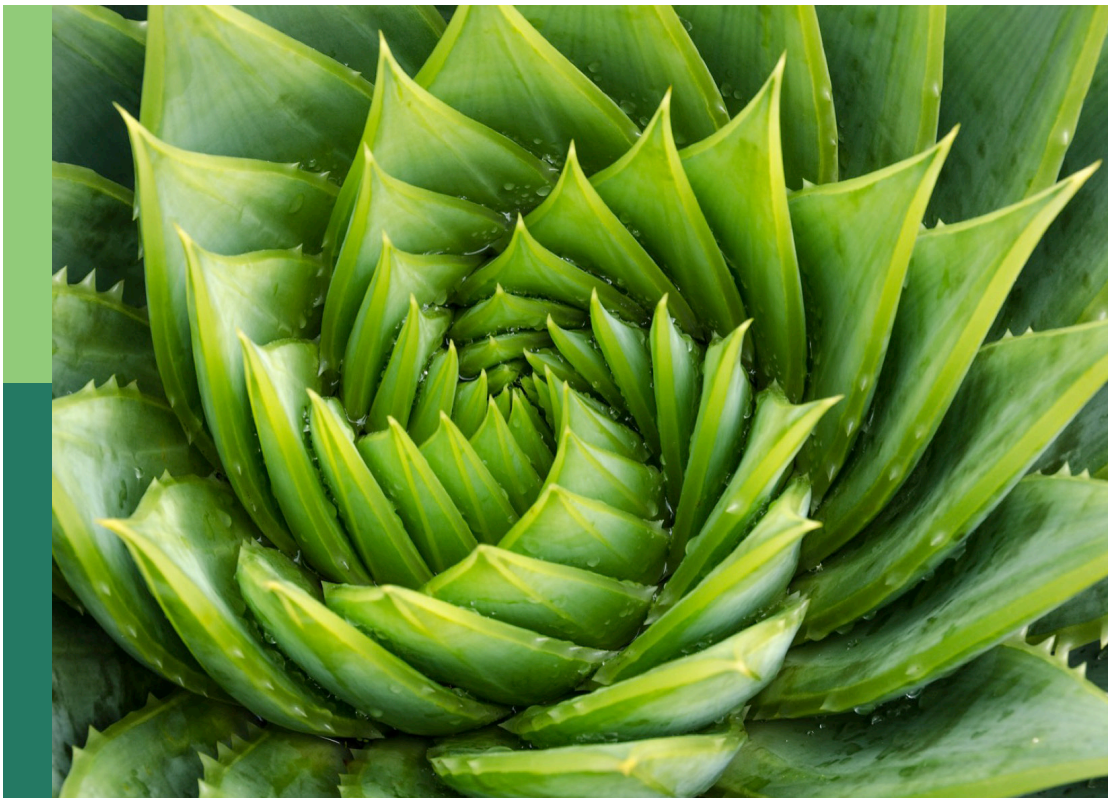
frontiers | Research Topics

Intelligent computing research with applications in biotechnology

Edited by
Yu Xue, Ferrante Neri and Ali Wagdy Mohamed

Published in
Frontiers in Plant Science



ISSN 1664-8714
ISBN 978-2-8325-4335-1
DOI 10.3389/978-2-8325-4335-1

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Intelligent computing research with applications in biotechnology

Topic editors

Yu Xue — Nanjing University of Information Science and Technology, China

Ferrante Neri — University of Surrey, United Kingdom

Ali Wagdy Mohamed — Cairo University, Egypt

Citation

Xue, Y., Neri, F., Mohamed, A. W., eds. (2024). *Intelligent computing research with applications in biotechnology*. Lausanne: Frontiers Media SA.

doi: 10.3389/978-2-8325-4335-1

Table of contents

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 18 October 2022
DOI 10.3389/fpls.2022.1035077

OPEN ACCESS

EDITED BY
Jian Su,
Nanjing University of Information Science and Technology, China

REVIEWED BY
Yao Xie,
Hunan University, China
Jinjin Rong,
Hanseo University, South Korea
Alex Liu,
Michigan State University, United States

*CORRESPONDENCE
Fei Zhou
tbwfvebx@hotmail.com

SPECIALTY SECTION
This article was submitted to Sustainable and Intelligent Phytoprotection, a section of the journal Frontiers in Plant Science

RECEIVED 02 September 2022
ACCEPTED 23 September 2022
PUBLISHED 18 October 2022

CITATION
Hu L, Zhou F and Fu X (2022) Texture synthesis of ecological plant protection image based on convolution neural network. *Front. Plant Sci.* 13:1035077. doi: 10.3389/fpls.2022.1035077

Texture synthesis of ecological plant protection image based on convolution neural network

Libing Hu[1], Fei Zhou[2*] and Xianjun Fu[3]

[1]Book Information Center, Zhejiang College of Security Technology, Wenzhou, China, [2]Business School, Wenzhou University, Wenzhou, China, [3]College of Artificial Intelligence, Zhejiang College of Security Technology, Wenzhou, China

Texture synthesis technology is an important realistic rendering technology. Texture synthesis technology also has a good application prospect in image rendering and other fields. Convolutional neural network is a very popular technology in recent years. Convolutional neural network model can learn the features in data and realize intelligent processing through the feature learning in data. Later, with the rapid improvement of convolutional neural network, texture synthesis technology based on neural network came into being. The purpose of this paper is to study the texture synthesis method of ecological plant protection image based on convolutional neural network. By studying the context and research implications, the definition of textures as well as texture synthesis methods, convolutional neural networks, and based on convolutional neural network. In the experiment, the experimental environment is established, and the subjective evaluation and objective evaluation of the image texture synthesis method experiment are investigated and studied by using swap algorithm. The experimental results show that the method used in this paper is superior to other methods.

KEYWORDS

convolutional neural network, ecological plant protection, image processing, texture synthesis method, realistic rendering technology

Introduction

Convolutional neural network is proposed according to the visual cognitive mechanism. Due to its local connection and weight sharing characteristics and the local translation invariance characteristics brought by the idea of spatial downsampling, it has made outstanding achievements in computer vision, speech recognition and other fields in recent years. The application of convolutional neural network to realize texture synthesis has also become one of the current research hotspots (Watanabe et al., 2018).

Images in nature or human real life contain various textures, and these textures have different characteristics. A part of these texture images is intercepted, whether it is a regular texture image or an irregular texture image, a new texture image can be formed.

As a significant branch of computer graphics, texture synthesis has always been the focus of researchers. The purpose of the Medeiros f a study was to determine whether the retinal nerve fiber layer thickness prediction obtained by the deep learning model applied to the fundus photos could detect progressive glaucoma changes over time. Design a retrospective cohort study of participants on color fundus photographs and spectral domain optical coherence tomography. Methods a deep learning convolutional neural network was trained to evaluate the fundus photos and predict the global retinal nerve fiber layer thickness measurement by tomography. The model was then tested on an independent eye sample with longitudinal follow-up by fundus photography and tomography. The ability to detect eyes with statistically significant tomographic change slopes was assessed by the subject operating characteristic curve. The repeatability of retinal nerve fiber layer thickness prediction was studied by measuring results obtained from multiple photos taken on the same day (Medeiros et al., 2021). Large scenes such as the building facade and other building structures of Labrie Larrivee F usually contain repetitive elements, such as the same window and brick patterns. A new method is proposed to improve the resolution and geometry of 3D meshes of large scenes with such repetitive elements. By using the structure from motion reconstruction and the ready-made depth sensor, the method captures small samples of the scene with high resolution and automatically extends the information to similar areas of the scene. The method uses RGB and SFM depth information as guidance, simple geometry as canvas, and powerful image-based texture synthesis method to expand the high-resolution mesh. The final result improves the standard SFM reconstruction with higher detail. Compared with full RGBD reconstruction, the method benefits from reduced manual labor and can be much cheaper than lidar based solutions (Labrie-Larrivee et al., 2019). The research of texture synthesis technology has made great progress, and a large number of new methods have appeared which lays a more solid theoretical foundation for digital image restoration.

In the second chapter, according to the background and research significance of the topic, the definition of texture and texture synthesis method, convolutional neural network, and the texture synthesis method of ecological plant protection image based on convolutional neural network are described. In the experiment of the third chapter, the experimental environment is established, and the subjective evaluation and objective evaluation of the image texture synthesis method experiment are investigated and studied in the fourth chapter by using the Swap algorithm. The experimental results show that the method used in this paper outperforms other methods.

Research on texture synthesis of ecological plant protection image based on convolution neural network

Topic background and research significance

The issue of image realism has always been the main course in computer graphics, and is also the focus of scientific researchers (Sardar and Tileubaeva, 2019; Larras et al., 2022). At first, people directly expressed the subtle structure of objective things through geometric models. However, the objective world is complex and changeable. This method has a large amount of calculation and modeling is also very difficult to meet the actual needs. In order to make up for the deficiency of geometric model expression, people combine the image processing technology with computer graphics technology, and use the image processing, analysis and synthesis technology to realize the realistic display of images, and have made a long-term improvement.

In recent years, texture synthesis technology, as an important image-based realistic rendering technology, has always received the attention of researchers and has high practical application value (Jassim and Harte, 2020; Shenson et al., 2021). Deep learning refers to interpreting data by simulating the thinking mode of human brain through different machine learning algorithms. It is a learning algorithm closest to human brain and has become one of the important branches of artificial intelligence research. Unlike traditional neural networks, deep learning networks are deeper in network level, larger in scale and higher in complexity. They can count large-scale data, learn fundamental characteristics from massive data, achieve breakthroughs in the field of artificial intelligence, and lead an innovative revolutionary upsurge. Convolutional neural network has the characteristics of weight sharing and pooling layer dimension reduction, which greatly shortens the training time of network model and improves the training time of model. Therefore, many remarkable research achievements have been made in image recognition and image processing.

Definition of texture and texture synthesis method

Texture is a concept often used in the field of computer graphics and photorealistic rendering. At present, it is generally defined as: the expression on the surface of any object is regarded as texture (Hsu and Yeh, 2018; Nikitina et al., 2018). In the field of computer graphics and image processing, people agree with the definition that texture refers to repeating the basic texture element - texel.

The basic constituent unit of texel is pixel, and texture refers to the type of image containing a special attribute, which is the

realization of a random process, which is local and stable (Williams et al., 2020; Preston et al., 2022). Texture image is different from ordinary image because of its locality and stability. Because of the complex changes of everything in the world, the classification of texture is also changeable. Texture can be obtained by image scanning, manual drawing, and computer image generation technology. According to the classification of texture synthesis, texture synthesis technology can be divided into three types: texture mapping method, process texture synthesis method and sample based texture synthesis method (Rsa et al., 2020; Borovik et al., 2022).

Convolutional neural network

Convolutional neural network is the most typical type of neural network, which is generally composed of input layer, hidden layer and output layer (Rafizah et al., 2018; Gribanov et al., 2021). The parameters of convolution layer include the number of input channels, the number of output channels, the size of convolution core, step size, filling, etc. among the many parameters of convolution layer, three are more important, which determine the function of convolution layer to a certain extent, namely, the number of convolution cores, the number of input channels and the number of output channels. Generally speaking, the convolution layer is composed of multiple convolution kernels. Any element in the convolution kernel is a numerical type weight coefficient that can be changed. In most cases, each convolution kernel corresponds to a numerical type weight coefficient, which is called the deviation amount. The coefficients of each convolution kernel in the convolution layer are weighted sum of multiple weight coefficients in the previous layer. In this sense, the convolution layer is a special feedforward neural network unit. Specifically, when the convolution kernel size is 1, the step size is 1, and no filling is performed, the convolution operation is equivalent to matrix multiplication, and the convolution layer is equivalent to a fully connected network. Some convolution networks use some special convolution structures, such as transposed convolution, extended convolution and separable convolution. These convolutions have their own characteristics and advantages, but also have some shortcomings. In most network models dealing with various tasks, there are often various convolution structures and basic units. Through mutual cooperation, a complex large-scale network that can complete the target task is constructed.

Texture synthesis method of ecological plant protection image based on convolutional neural network

The early texture synthesis technology based on neural network is mainly based on the theory of artificial neural network. A texture synthesis method based on continuous Hopfield network is proposed, which connects the neurons with the pixels of the image to synthesize different two tone and gray tone texture images; The theory of BP neural network is applied to the ecological plant protection texture image generation, and the application of artificial neural network in the ecological plant protection texture image synthesis is studied by using the analysis and feature extraction of the ecological plant protection texture. However, the effect of texture generation based on simple artificial neural network is not very ideal, and the extraction of image features is still based on artificial analysis, which makes the design cycle relatively long (Stepanova et al., 2018; Esquivel-Castro et al., 2019).

In recent years, with the rise and rapid improvement of deep learning, texture synthesis based on convolutional neural network has also achieved a lot of research results. Based on VGg network, the loss function is added to the statistical feature distribution of the image. These statistical distributions are calculated by the Clem matrix. This method has been extended by many researchers. Combined with Markov random field, Gatys et al., the method is extended to the fusion of ecological plant protection images, and the synthetic effect of ecological plant protection images is more stable. A confrontation generation network with 5-layer full convolution neural network as generator and discriminator is constructed. The synthetic texture ecological plant protection image is more realistic. Although the network training time is long, the test time is only 0.2S.

Investigation and research on texture synthesis of ecological plant protection image based on convolution neural network

Experimental environment

The experimental environment in this paper is a general experimental environment for deep learning, in which GPU processor of company a and CPU processor of company B are used. The training of convolutional neural network model requires that the neural network model be loaded into the GPU processor and then trained. At the same time, set the batch size during training. Batch size is the number of data samples selected during training the neural network model.

Research content

In the experiment, the subjective evaluation of the image texture synthesis method and the experiment will be investigated and analyzed, and finally the experimental results will be obtained.

Swap algorithm

The swap algorithm operates directly on the feature maps extracted by the convolutional neural network to find the similarity between the feature maps extracted by the loss network. The input of the loss network includes the output of the texture network, that is I_1 . The corresponding input model is M_1 set in the previous text, which can be assumed as $M_1(I_1)$. The value in $M_1(I_1)$ is composed of the characteristic diagram of each layer in M_1 neural network, namely:

$$M_1(I_1) = \left\{F_1^l, F_2^l, F_3^l, \ldots, F_n^l\right\}(n-1, 2, \ldots, n) \quad (1)$$

Among them, F_n^l represents the feature map of a layer in the M_1 model; at the same time, the I_1 corresponding to I_{gt} is used as an input and input to the neural network model red M_2 , that is, $M_2(I_{gt})$. The values in $M_2(I_{gt})$ consist of feature maps for each layer in the M_2 model, namely:

$$M_2(I_{gt}) = \left\{f_1^l, f_2^l, f_3^l, \ldots, f_n^l\right\}(n-1, 2, \ldots, n) \quad (2)$$

Among them, f_n^l represents the feature map of a certain layer in the model M_2 .

Finally, the swap algorithm designed in this paper is used to perform the matching calculation. Here, swap can be viewed as a network layer similar to the convolutional layer. The formula is expressed as follows:

$$\text{swap}_{texture} = swap(M_2(I_{gt}), M_1(I_1)) \quad (3)$$

Analysis and research on texture synthesis of ecological plant protection image based on convolution neural network

Subjective evaluation of image texture synthesis method experiment

The image (a) column is the original sample texture, and the image (b) column is the texture image generated by the method in this paper, the image (c) column is the texture image generated by the method of Gayts et al. By making a questionnaire, we select some students and teachers to participate in the evaluation of the generated texture image results, and we subjectively get the superiority of a certain method through their scores.

Since the field of study and study of each teacher and student is different, this method can be regarded as an auxiliary evaluation method and cannot play a decisive role. Let these images be evaluated by 2 randomly selected people. The evaluation tables of these 2 people are shown in Table 1 and Figure 1, and Table 2 and Figure 2:

Therefore, we can conclude that our method has certain advantages compared with other methods. However, the use of subjective evaluation methods to evaluate the generated results has certain limitations, because the subjective evaluation due to the existence of human interference factors, due to personal experience, personal preferences, and differences in the knowledge level of each person will make the same texture synthesis result. There is a large difference in the evaluation of the subjective evaluation method, which will make the subjective evaluation method have a certain error and will not be very accurate. Therefore, an objective evaluation method is also introduced in this paper to ensure the objectivity of the generation effect evaluation results.

Objective evaluation of image texture synthesis method experiment

In this section, the objective evaluation methods used are introduced, these methods are: Mean Squared Error (MSE), Peak Signal-Noise Ration (PSNR), Structural Similarity (Structural Similarity, SSIM). Similarly, in the objective evaluation section, use the objective evaluation criteria to compare the parameters of the image evaluation table as shown in Table 3 and Figure 3, and Table 4 and Figure 4, Table 5 and Figure 5:

Conclusions

The convolutional neural network model has made significant breakthroughs and achievements in many aspects, such as image synthesis, image classification, image segmentation and other fields. In this paper, the related things about texture synthesis algorithm are mainly introduced. Texture synthesis algorithms are divided into traditional and modern. There are some defects in the traditional texture synthesis method. For example, the texture

TABLE 1 Evaluation Table.

Picture	1	2	3	4
(a)	9.6	4.2	9.8	3.1
(b)	8.6	8.4	9.3	9.6
(c)	9.6	6.4	9.4	8.7

10.3389/fpls.2022.1035077

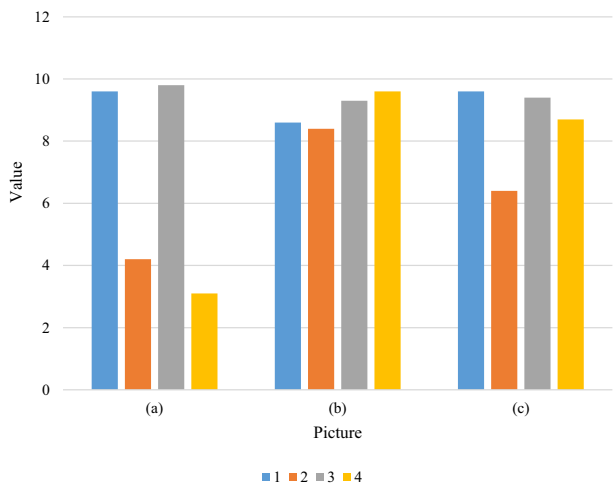


FIGURE 1
Evaluation data comparison is shown. The image (a) column is the original sample texture, the image (b) column is the texture image generated by this method, and the image (c) column is the text image generated by Gayts and other methods.

TABLE 2 Teacher evaluation form.

Picture	1	2	3	4
(a)	7.1	7.6	7.9	8.0
(b)	8.4	8.7	8.8	8.5
(c)	7.5	7.6	8.5	8.9

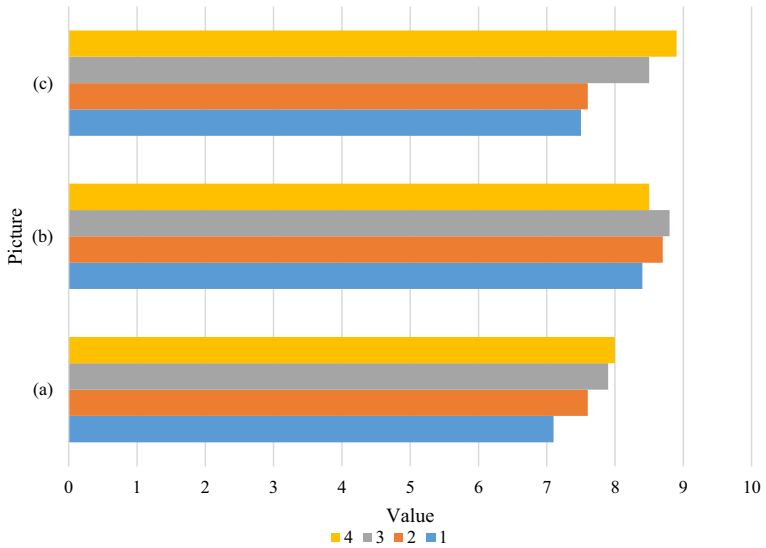


FIGURE 2
Subjective evaluation data for comparison. Using the subjective evaluation method, we obtained 2 evaluation sheets by randomly selecting 2 people. The 2 evaluation sheets have each person's evaluation and score for the texture results. Through calculation, we can get: The average score of the method in this paper is 8.2 points. The method in image **(c)** has a highest score of 9.5.Although some texture synthesis results of other methods may be superior to our results, our method outperforms other methods in many texture synthesis image results.

TABLE 3 Image evaluation table.

Evaluation criterion	The method of this paper	Other method
PSNR	16.87	18.51
SSIM	0.34	0.23
MSE	0.57	0.69

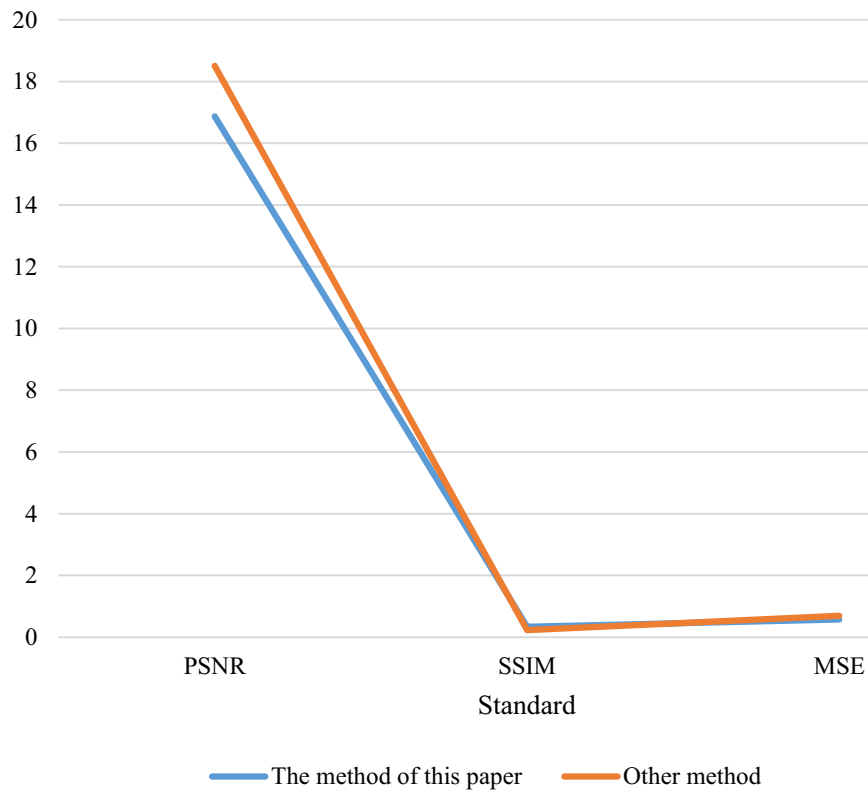


FIGURE 3
Image evaluation data in Figure.

TABLE 4 Objective evaluation table.

Evaluation criterion	The method of this paper	Other method
PSNR	10.54	9.45
SSIM	0.11	0.15
MSE	0.45	0.51

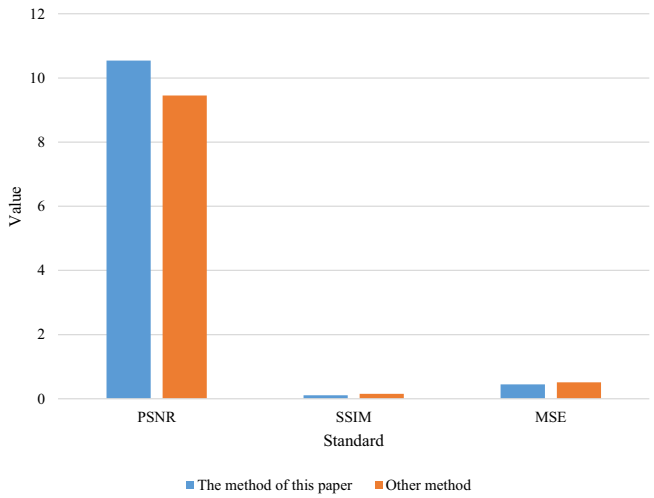


FIGURE 4
Objective evaluation of the data in Figure.

TABLE 5 Line C image evaluation table.

Evaluation criterion	The method of this paper	Other method
PSNR	6.54	8.65
SSIM	0.05	0.04
MSE	0.57	0.65

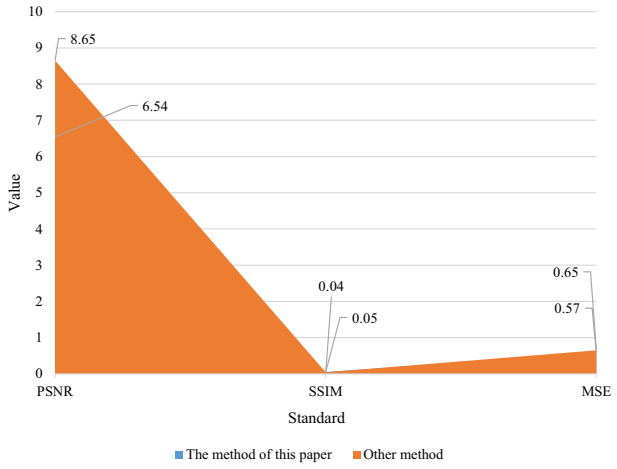


FIGURE 5
Comparison diagram of the data in row C. Through the evaluation of objective methods, it can be seen that the method in this paper can be close to other methods in some texture synthesis. The biggest innovation of our method is the synthesis at multiple scales, which surpasses other methods in this respect. The method in this paper is compared with other texture synthesis algorithms based on convolutional neural network, and the conclusions are drawn from the perspectives of subjective evaluation and objective evaluation. Through the evaluation of objective methods, it can be seen that the method in this paper can be close to other methods in some texture synthesis.

synthesized in texture mapping may have seams that change; the problem in the procedural texture synthesis method is that new parameters need to be adjusted when a new texture is not generated., Different textures correspond to different parameters, which is a troublesome way to generate textures; the textures existing in the texture synthesis method based on sample images are the sample images needed to synthesize textures. If the selection is not appropriate, the synthesized textures may be The real feeling will be lost.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material. Further inquiries can be directed to the corresponding author.

Author contributions

FZ: Conceptualization, Funding Acquisition, Resources, Supervision, Writing - Review and Editing. LH: Conceptualization, Methodology, Software, Investigation, Formal Analysis, Writing - Original Draft; XF: Visualization, Investigation; Resources, Supervision; Software, Validation. All authors contributed to the article and approved the submitted version.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Borovik, P., Oestreicher, V., Angelomé, P. C., Barja, B. C., and Jobbágy, M.. (2022). Room temperature synthesis of lanthanum phosphates with controlled nanotexture as host for Ln(III) through the epoxide route. *J. Sol-Gel Sci. Technol.* 102 (1), 279–287. doi: 10.1007/s10971-022-05744-w

Esquivel-Castro, T. A., Martínez-Luévanos, A., García-Cerda, L. A., Contreras-Esquivel, JC, Pérez, P. B., González Aguilera, E. N., et al. (2019). Effect of the drying on morphology and texture of aerogels and zirconia cryogels. *MRS Adv.* 4 (64), 1–9. doi: 10.1557/adv.2019.450

Gribanov, E. N., Gorshkov, A. I., Sinitsyn, E. A., Yu Khripunov, V., and Oskotskaya, E. R . (2021). On the synthesis and morphology and formation peculiarities of an alumosilicate film on a substrate. *J. Surf Invest. X-ray Synchrotron Neutron Tech* 15 (1), 16–23. doi: 10.1134/S1027451021010079

Hsu, P. K., and Yeh, J. (2018). PS02.070: Deep neural network to predict poor prognostic factors in patients with esophageal cancer. *Dis. Esophagus* 31 (13), 140–140. doi: 10.1093/dote/doy089.PS02.070

Jassim, W. A., and Harte, N. (2020). Estimation of *a priori* signal-to-noise ratio using neurograms for speech enhancement. *J. Acoust Soc. America* 147 (6), 3830–3848. doi: 10.1121/10.0001324

Labrie-Larrivee, F., Laurendeau, D., and Lalonde, J. F. (2019). Depth texture synthesis for high-resolution reconstruction of large scenes. *Mach. Vision Appl.* 30 (4), 795–806. doi: 10.1007/s00138-019-01030-y

Larras, F., Charles, S., Chaumot, A., Pelosi, C., LeGall, M., Mamy, L., et al. (2022). A critical review of effect modeling for ecological risk assessment of plant protection products. *Environ. Sci. Pollut. Res.* 29 (29), 43448–43500. doi: 10.1007/s11356-022-19111-3

Medeiros, F. A., Jammal, A. A., and Mariottoni, E. B. (2021). Detection of progressive glaucomatous optic nerve damage on fundus photographs with deep learning. *Ophthalmology* 128 (3), 383–392. doi: 10.1016/j.ophtha.2020.07.045

Nikitina, M. A., Pchelkina, V. A., and Kuznetsova, O. A. (2018). Technological solutions for intelligent data processing in the food industry[J]. *Proc. Voronezh State Univ. Eng. Technol.* 80 (2), 256–263. doi: 10.20914/2310-1202-2018-2-256-263

Preston, F. G., Meng, Y., Burgess, J., Ferdousi, M., Azmi, S., Petropoulos, I. N., et al. (2022). Artificial intelligence utilising corneal confocal microscopy for the diagnosis of peripheral neuropathy in diabetes mellitus and prediabetes. *Diabetologia* 65 (3), 457–466. doi: 10.1007/s00125-021-05617-x

Rafizah, Z., Omar, N. A., and Ibrahim, N. (2018). Morphology and chemical structure of sn(Oct)2 thin layer added binder *via* sol gel method. *Malaysian J. Anal Sci.* 22 (2), 311–317. doi: 10.17576/mjas-2018-2202-17

Suarez, A. R., Muñoz, F. F., Bonelli, P, Cukierman, A. L, and Larrondo, S. A.. (2020). Hierarchical, template-free self-assembly morphologies in CeO2 synthesized *via* urea-hydrothermal method - ScienceDirect. *Ceram Int.* 46 (8), 11776–11785. doi: 10.1016/j.ceramint.2020.01.212

Sardar, A. A., and Tileubaeva, Z. S. (2019). Influence of tillage methods and plant protection agents on the ecological parameters of soil cover and barley yield. *Agric. Machinery Technol.* 13 (3), 8–10. doi: 10.22314/2073-7599-2019-13-3-8-10

Shenson, J. A., Liu, G. S., Farrell, J., and Blevins, N. H. (2021). Multispectral imaging for automated tissue identification of normal human surgical Specimens:[J]. *Otolaryngol–Head Neck Surg.* 164 (2), 328–335. doi: 10.1177/0194599820941013

Stepanova, L. N., Belskaya, O. B., Vasilevich, A. V., Leont'eva, N. N., Baklanova, O. N., Likholobov, V. A., et al. (2018). Effect of the composition of initial components and the conditions of activation on the mechanochemical synthesis of magnesium–aluminum layered double hydroxides. *Kinet Catal* 59 (4), 521–531. doi: 10.1134/S0023158418040134

Watanabe, T., Oyama, T., and Fukumi, M. (2018). Estimation of tongue motion and vowels of silent speech based on EMG from suprahyoid muscles using CNN. *IEEJ Trans. Electron. Inf. Syst.* 138 (7), 828–837. doi: 10.1541/ieejeiss.138.828

Williams, B. M., Borroni, D., Liu, R., Zhao, Y, Zhang, J, Lim, J, et al. (2020). An artificial intelligence-based deep learning algorithm for the diagnosis of diabetic neuropathy using corneal confocal microscopy: a development and validation study. *Diabetologia* 63 (2), 419–430. doi: 10.1007/s00125-019-05023-4

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 26 January 2023
DOI 10.3389/fpls.2023.1122788

OPEN ACCESS

EDITED BY
Yu Xue,
Nanjing University of Information Science and Technology, China

REVIEWED BY
Ran Li,
Xinyang Normal University, China
Guangjun Liang,
Jiangsu Police Officer College, China
Jianfang Xin,
Harbin Institute of Technology, China

*CORRESPONDENCE
Linguo Li
llg-1212@163.com

SPECIALTY SECTION
This article was submitted to Sustainable and Intelligent Phytoprotection, a section of the journal Frontiers in Plant Science

RECEIVED 13 December 2022
ACCEPTED 02 January 2023
PUBLISHED 26 January 2023

CITATION
Li S, Li Z, Li Q, Zhang M and Li L (2023) Hybrid improved capuchin search algorithm for plant image thresholding. *Front. Plant Sci.* 14:1122788. doi: 10.3389/fpls.2023.1122788

Hybrid improved capuchin search algorithm for plant image thresholding

Shujing Li[1], Zhangfei Li[1], Qinghe Li[1], Mingyu Zhang[1] and Linguo Li[1,2]*

[1]School of Computer and Information Engineering, Fuyang Normal University, Fuyang, China, [2]School of Computer, Nanjing University of Posts and Telecommunications, Nanjing, China

With the development and wider application of meta-heuristic optimization algorithms, researchers increasingly apply them to threshold optimization of multi-level image segmentation. This paper explores the performance and effects of Capuchin Search Algorithm (CAPSA) in threshold optimization. To solve problems of uneven distribution in the initial population of Capuchin Search Algorithm, low levels of global search performance and premature falling into local optima, this paper proposes an improved Capuchin Search Algorithm (ICAPSA) through a multi-strategy approach. ICAPSA uses chaotic opposite-based learning strategy to initialize the positions of individual capuchins, and improve the quality of the initial population. In the iterative position updating process, Levy Flight disturbance strategy is introduced to balance the global optimization and local exploitation of the algorithm. Finally, taking Kapur as the objective function, this paper applies ICAPSA to multi-level thresholding in the plant images, and compares its segmentation effects with the original CAPSA, the Fuzzy Artificial Bee Colony algorithm (FABC), the Differential Coyote Optimization Algorithm (DCOA), the Modified Whale Optimization Algorithm (MWOA) and Improved Satin Bowerbird Optimization Algorithm (ISBO). Through comparison, it is found that ICAPSA demonstrates superior segmentation effect, both in the visual effects of image segmentation and in data comparison.

KEYWORDS

capuchin search algorithm, chaotic mapping, opposite-based learning, levy flight, plant image thresholding

1 Introduction

Due to the advantages of fast convergence rate and high accuracy of meta- heuristic optimization algorithms, many researchers increasingly apply them to real-world problems, to improve the application effects of computer-aided design in the engineering field (Pare et al., 2020). Ma et al. (2011) combine Artificial Bee Colony algorithm (ABC) with image segmentation to improve the segmentation accuracy of Synthetic Aperture Radar (SAR) images. Hemasian-Etefagh and Safi-Esfahani, (2019) apply Whale Optimization Algorithm (WOA) to solve the scheduling problem of cloud computing, thus reducing the execution and

response time of scheduling tasks, and increasing the computing throughput in the context of a cloud environment. (Gholizadeh and Baghchevan, 2017) apply Firefly Algorithm (FA) in the seismic design of steel frames to locate the optimal goal more quickly. Specifically, in the field of plant image processing, with the change of environment and the development of information technology, many countries regard ecological construction and the development of smart agriculture as national strategies. In order to improve the benefits of ecological environment and agricultural development, scholars have deeply integrated traditional plant planting technology with the internet of things, 5G and artificial intelligence technology (Maray et al., 2022; Ruwona and Scherm, 2022). This development model reduces the labor of workers and improves production efficiency. However, in the current development process, the main problem encountered is that the frequent occurrence of diseases has led to reduction of output or even crop failure. In order to reduce the loss of plant quality and economy caused by diseases, targeted screening and diagnosis in advance are required during plant growth (Patel et al., 2023). The traditional way is manual identification, relying on experience, high cost and low accuracy. In recent years, the recognition based on computational image processing is more efficient and accurate, the processing steps include: image preprocessing, image segmentation, feature extraction and recognition, and the higher accuracy of image segmentation, the higher accuracy of recognition (Xue et al., 2021; Hasan et al., 2022). However, segmentation accuracy and efficiency directly affect the application of segmentation technology in plant applications. Making use of the advantages of meta-heuristic algorithms in multi-level image thresholding is thus considered an effective plant disease assisted treatment (Li et al., 2022).

As one of the main technical means of image segmentation, multi-level thresholding based on specific objective function is fused with a meta-heuristic algorithm, and has been thus used successfully (Merzban and Elbayoumi, 2019; Xue et al., 2022). Image segmentation divides an image into several regions according to characteristics such as texture, color, brightness, contrast, shape, and size (Pare et al., 2020; Rodríguez-Esparza et al., 2020; Xue et al., 2019). Specifically, thresholding segmentation is divided into bi-level (Bao et al., 2019) and multi-level thresholding (Ma and Yue, 2022). Bi-level is used mainly for image binarization, and its application field is limited. The multi-level can dynamically improve segmentation accuracy in response to actual needs by adjusting the number of thresholds, although such increase of thresholds leads to an explosive increase in computational complexity. Therefore, a meta-heuristic algorithm based on specific objective functions can effectively strike the balance between these problems (Ma and Yue, 2022). The common objective functions include Kapur entropy (Upadhyay and Chhabra, 2020), Minimum Cross Entropy (MCE) (Wang and Song, 2022), Tsallis (Lin et al., 2020).

Regarding the selection of objective functions, Sathya et al. (2021) compare and analyze the segmentation effect of Exchange Market Algorithm (EMA) using Kapur, Otsu and MCE. Through comparison of visual effects and quantitative data after standard image segmentation, they conclude that the Kapur-based method has faster processing speed and better segmentation effect. Li et al. (2016) taking Kapur as the objective function, prove the efficiency of the fuzzy ABC (FABC) in multi-level image thresholding through experimental comparison using BSD500 dataset. Kalyani et al. (2021) use EMA algorithm based on Kapur and MCE to segment three different images, comparing it with Krill Herd (KH), Teaching-Learning based Optimization (TLBO) and Cuckoo Search Algorithm (CSA). They find that the Kapur-based algorithm has faster convergence speed and segmentation accuracy, and thus has better practical application value. Rajinikanth et al. (2021) improve Moth Flame Optimization (MFO) algorithm and compare the image thresholding segmentation effects based on Kapur and Tsallis. Through comparison of multiple experimental results, they find this method achieves better segmentation effect than other methods of the same kind. Chen et al. (2022) taking Kapur as the objective function, integrate the multi-strategy driven Shuffled Frog-Leaping Algorithm with Horizontal and Vertical Crossover search (HVSFLA) to perform multi-level thresholding, achieving remarkable segmentation effect.

Compared with traditional methods (Kittler and Illingworth, 1986; Otsu, 1979), meta-heuristic optimization algorithms perform more efficiently in multi-level image thresholding. Zhao et al. (2021) use random spare strategy and logistic chaos enhancement strategy to optimize Ant Colony Optimization (ACO), effectively improving the convergence speed and accuracy of ACO and enhancing the ability of the algorithm to evade local optima. The experimental results show that the improved ACO algorithm is satisfactory. To optimize the multi-level thresholding of gray-scale images, Abdel-Basset et al. (2022) combine Linearly Convergence Increasing and Local Minima Avoidance Technique (LCMA) and Ranking-based Update Method (RUM) to improve the Wheel Optimizer Algorithm (WOA). The former moves the individual at the worst position in the population into the range of the best current scheme to avoid local optimization. The latter replaces the unfavorable solution with a better one. Similarly, based on Kapur, 13 test images of Berkeley segmentation dataset BSD are verified, demonstrating that the improved WOA is superior to other methods in both segmentation image quality and convergence speed. Ewees et al. (2020) propose an Improved Artificial Bee Colony Algorithm integrating the Sine Cosine Algorithm (ABCSCA). This algorithm uses ABC to narrow the search scope and optimize the threshold, while SCA can determine the global optimal threshold and obtain the optimal solution. To measure the effects of ABCSCA, Otsu and fuzzy entropy are used as the objective functions to segment 19 images. Compared with the original ABC, SCA algorithms and Hybrid Swarm Optimization (FASSO), the algorithm has more obvious advantages in image segmentation effect, and convergence speed. Anitha et al. (2021) improve Modified Whale Optimization Algorithm (MWOA) by using cosine function, and obtain better image segmentation quality and convergence speed than PSO, ABC and other algorithms using Otsu as the objective function. Li et al. (2021) use the differential evolution strategy to improve the population updating mechanism of Coyote Optimization Algorithm (DCOA), not only improving the convergence speed of the algorithm, but also its image segmentation accuracy, rendering it superior to standard COA, Gray Wolf Optimizer (GWO), modified Discrete Gray Wolf Optimizer (DGWO) and other methods. To solve the problem of identifying corn pests and diseases, Chen et al. (2021) improve Particle Swarm Optimizer (PSO) with an elite based advantage scheme to form an Enhanced Comprehensive Learning Particle Swarm Optimizer

(ECLPSO). Compared with the Comprehensive Learning Particle Swarm Optimizer (CLPSO) and Hybridizing Sine Cosine Algorithm with Differential Evolution (SCADE) algorithm in the corn leaf disease image in the public database of a plant village company, the results show this method to be superior to other comparison algorithms in locating the best threshold, and have higher convergence accuracy. Li et al. (2022) proposed a strategy based on chaos initialization and Cauchy mutation to improve Satin Bowerbird Optimization Algorithm (SBO), and verified its values in Kaggle plant image dataset. The comparison between the fuzzy Modified Discrete Grey Wolf Optimizer with aggregation strategy (FMGWO) and the fuzzy Coyote Optimization Algorithm (FCOA) proves that the improved ISBO has higher accuracy in the field of plant image segmentation.

Meta-heuristic optimization algorithms can effectively improve the computational efficiency in multi-level image thresholding. However, there is room for further improvement in population initialization and global search ability (Pare et al., 2020; Li et al., 2022). Therefore, this paper seeks to improve CAPSA, comparing its efficacy in plant image segmentation with other algorithms. CAPSA is a novel meta-heuristic optimization algorithm proposed by Braik et al. (2021) in 2021. This algorithm divides the population into two groups with distinct functions so as to strike the balance between global search ability and local exploitation ability. Compared with other similar algorithms, it has higher convergence speed and accuracy, but in later iterations, CAPSA is also prone to fall into local optima.

Therefore, this paper uses Tent chaotic iterative mapping and Opposite-based learning strategy to initialize the population, improve the quality of the initial population, make its distribution more uniform, thus eschewing the premature local optima of CAPSA. In the position updating strategy of the algorithm, Levy Flight strategy is integrated to balance the ability of global search and local exploitation of the algorithm to form an Improved Capuchin Search Algorithm (ICAPSA). To verify the effects of ICAPSA, this paper uses Kapur entropy as the objective function to segment plant images into multiple thresholds, and compares the experimental results with the results of the FABC (Li et al., 2016), MWOA (Anitha et al., 2021), DCOA (Li et al., 2021) et al.

The remainder of this paper is as follows: In the second section, the original CAPSA and its model construction is described. In the third section, the improved strategy of the algorithm is presented in detail. In the fourth section, to verify the practical effect of ICAPSA in plant image segmentation, six plant images are selected for visual and quantitative data comparative analysis. Finally, the fifth section concludes.

2 Description and model construction of Capuchin Search Algorithm (CAPSA)

2.1 Original CAPSA description

Capuchin search algorithm (CAPSA) is a new algorithm that simulates the foraging behavior of capuchin populations in Brazil and South America. Each population includes about 10 to 35 capuchins, and each population has an alpha (α) monkey commanding this group, called the leader, who is responsible for finding food sources for this group. The remaining capuchins are called followers. If the alpha monkey cannot obtain sufficient food sources in time, the group will be divided into smaller sub-groups to forage independently.

Capuchin monkeys use jumping, swinging and climbing in the process of foraging for food. Jumping allows capuchin monkeys to have a wider search range. Swinging and climbing are used to improve local search ability. Followers will update their positions according to the leader's position and their own positions, and finally improve the foraging rate and success rate.

2.2 Initialization of CAPSA

Like other similar algorithms, capuchin search algorithm is also a population-based search algorithm. It randomly initializes the population. Each individual of the population represents the candidate scheme of the target problem. The initialized individuals are divided into two categories: alpha monkeys and followers.

Assume that the capuchin monkey population has n individuals, and the search space is d-dimensional. The initial position can then be expressed by the following matrix:

$$x = \begin{bmatrix} x_1^1 & x_2^1 & \cdots & \cdots & x_d^1 \\ x_1^2 & x_2^2 & \cdots & \cdots & x_d^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_1^n & x_2^n & \cdots & \cdots & x_d^n \end{bmatrix} \quad (1)$$

Where x represents the positions of capuchin monkeys, n the number of capuchins, d the dimension of the problem, and x_d^i the position of the i -th capuchin monkey in d-dimensional space.

Initialize the position of each capuchin individual by (2):

$$x^i = ub_j + r \times (ub_j - lb_j) \quad (2)$$

Where ub_j and lb_j represent the upper and lower bounds of the capuchin monkey in the dimensional space respectively, and r is a random number uniformly generated inside [0,1].

2.3 Evolution of CAPSA

In Capuchin search algorithm, the updating of capuchin positions depends on their current position and best position along with the location of food F. F is the target of capuchin monkeys in d-dimensional search space. The position of the leader and those of its followers relative to where food F is, are updated as follow steps.

Jump on the tree: the leader (α monkey) can jump from tree to tree or from the current branch to other branches of the same tree, then α monkey's position updating formula is as follows.

$$x_j^i = F_j + \frac{P_{bf}(v_j^i)^2 \sin(2\theta)}{g}, i < {}^{n}/_{2}, 0.1 < \epsilon \leq 0.20 \quad (3)$$

Where x_j^i represents the position of α monkey and its followers in the j dimension, F_j the position of food in the j-th dimension, $\in$ the random number generated inside [0, 1], P_{bf} is the probability that the

tail provides balance in the capuchin jumping process, g the gravitational acceleration, g=9.81, θ is the capuchin jumping angle, τ represents the life cycle, systematically decreases in the whole iteration process, v_j^i represents the speed of the i-th capuchin in the j-th dimension.

The jumping angle of a capuchin monkey can be defined by (4):

$$\theta = \frac{3}{2}r \tag{4}$$

Where r is a random number generated uniformly in the range [0,1]. To balance the global search ability and local exploitation ability, CAPSA introduces the concept of life cycle τ as shown in (5):

$$\tau = \beta_0 e - \beta_1 \left(\frac{k}{K}\right)^{\beta_2} \tag{5}$$

Where k and K respectively represents the current iteration number and the maximum iteration number. The values of parameters β_0, β_1 and β_2 are respectively 2, 21 and 2. The exponential function has a great impact on updating the positions of Capuchin monkeys, the exploration and development of regions, and the quick locating of food sources.

The velocity of the i-th capuchin monkey in the j dimension is shown in (6):

$$v_j^i = \rho v_j^i + \tau a_1 (x_{best_j}^i - x_j^i) r_1 + \tau a_2 (F_j - x_j^i) r_2 \tag{6}$$

Where i=1, 2, 3, …,n, j stands for the dimension of the problem, v_j^i the current velocity of the i-th capuchin monkey in the j dimension, x_j^i the current position of the i-th capuchin monkey in the j-th dimension. $x_{best_j}^i$ the best velocity of the i-th capuchin monkey in the j-th dimension, F_j the position of food in the j-th dimension. a_1 and a_2 are two normal numbers, their values can be taken at 1 or 0, representing the impact of $x_{best_j}^i$ and F_j on the velocity of capuchins. r_1 and r_2 are random numbers generated evenly in the range [0, 1]. ρ is the inertia coefficient with a value of 0.7, indicating the impact of the current velocity on the motion.

Jumping on the ground: Capuchins can jump from one place to another on the ground, from one side of the riverbank to the other, or wander normally to search for food. In this case, the position updating formula of α monkey and its followers is as follows:

$$x_j^i = F_j + \frac{P_{ef} P_{bf} (v_j^i)^2 \sin(2\theta)}{g}, \quad i < {}^n/_2, 0.2 < \epsilon \le 0.30 \tag{7}$$

Where P_{ef} represents the elasticity probability of a capuchin monkey moving on the ground, θ is defined in (4).

On the other hand, when α monkey wanders normally, the position updating can be shown in (8):

$$x_j^i = x_j^i + v_j^i, \ i < {}^n/_2, 0.3 < \epsilon \le 0.50 \tag{8}$$

From these two jumping mechanisms, it can see that capuchin monkeys have two basic parameters in the process of approaching food. the probability P_{bf} that the tail provides balance in the process of jumping, and the elasticity probability P_{ef} of moving on the ground. These two coefficients balance their global search and local exploitation ability, and their values are taken at 0.7 and 9, respectively.

Swing: some α monkeys and their followers will swing their bodies over the branches with their tails and perform local exploitation to forage food. The positions of capuchin monkeys are updated as follows:

$$x_j^i = F_j + \tau P_{bf} \times \sin(2\theta), \ i < {}^n/_2, 0.5 < \epsilon \le 0.75 \tag{9}$$

Where θ is defined in (4);

Climbing: in the process of foraging, some α monkeys and their followers will climb up a tree or branches, and then climb down. This behavior is also local exploitation. Their positions are as follows:

$$x_j^i = F_j + \tau P_{bf} (v_j^i - v_{j-1}^i), \ i < {}^n/_2, \ 0.75 < \epsilon \le 1.0 \tag{10}$$

Where v_j^i is the current velocity of the i-th capuchin in the j-th dimension, v_{j-1}^i is the previous velocity of the i-th capuchin in the j-th dimension.

Random migration of capuchin monkeys: in foraging food, capuchin monkeys will randomly search in several new directions to effectively explore the forest to search for better food sources. Random migration is shown in (11):

$$x_j^i = \tau \times \left[lb_j + \epsilon \times (ub_j - lb_j)\right], \ i < {}^n/_2, \epsilon < Pr \tag{11}$$

Where Pr is a normal number with a value of 0.1, representing the probability of capuchin monkeys performing random search. ub_j and lb_j are the upper and lower bounds of the j-th dimensional search space, respectively. The random migration of capuchins not only enhances the global search ability, but also prevents CAPSA from falling into local optima.

To sum up, as is shown in (3) to (10), the capuchins will change their positions depending on the availability of food, their search target. This situation is particularly apparent when r > 0.1.

On the other hand, when r≤0.1, capuchin monkeys are more likely to randomly change their positions in the search domain so as to explore different areas for food. In this case, parameter τ can enhance the search space available for exploration.

The positions of the followers are updated according to the position of α monkey, as is shown by (12).

$$x_f = x_i + v_0 t + \frac{1}{2} a t^2 \tag{12}$$

Where x_f is the final displacement, x_1 is the initial displacement. t is the time, v_0 represents the initial velocity. a is the acceleration, its value is shown in (13):

$$a = \frac{\Delta v}{\Delta t} = \frac{v_f - v_0}{t_1 - t_0} \tag{13}$$

Where t_1 and t_0 represent the last time and the first time respectively. The parameter v_f represents the final velocity, its value is shown in (14):

$$v_f = \frac{\Delta x}{\Delta t} = \frac{x_f - x_0}{t_1 - t_0} \tag{14}$$

Substitute (14) into (13), and set the initial velocity v_0=0, then a can be shown by (15):

$$a = \frac{x_f - x_0}{(t_1 - t_0)^2} \tag{15}$$

In optimization, t_1 represents iteration, t_1-t_{i-1} represents the difference between successive iterations, equal to a value of 1. Based on (12) and (15), (16) represents the position updating formula of the followers:

$$x_j^i = \frac{1}{2}(x_j'^i + x_j^{i-1}),\ n/2 \le i \le n \tag{16}$$

Where x_j^i represents the current position of the followers in the j-th dimension. x_j^{i-1} represents the previous position of the followers in j-1-th dimension. $x_j'^i$ the current position of the leader in the j-th dimension.

3 The Hybrid Improved CAPSA

3.1 Tent chaotic mapping

The Capuchin search algorithm performs well in convergence speed, and also considers the balance between global search and local exploitation. However, similar to other meta-heuristic optimization algorithms, there exists a certain probability of "local optimal" in later iterations, which is attributable to the population random initialization strategy of meta-heuristic optimization algorithm. In constructing the model in section 2.2, it is described that the population initialization of CAPSA is completed randomly according to (2), and the distribution of individual positions of the population is inherently random. As chaotic mapping has the advantages of ergodicity and universality, many researchers use it to optimize the population initialization of meta-heuristic optimization algorithms. Commonly used chaotic map methods include Circle map (Ewees and Abd Elaziz, 2020), Gauss map (Elaziz et al., 2021), Logistic map (Prasad et al., 2021) and Tent map (Li et al., 2020), as shown in (Figures 1A-D). Circle chaotic map and Logistic chaotic map have the characteristics of small humps and peaks at both ends, which may cause group aggregation and are not conducive to global search. Compared with Gauss chaotic map, Tent chaotic map has the advantages of more uniform distribution and smaller peak value, which will not affect the convergence speed of the algorithm. Therefore, Tent chaotic map model is selected in this paper. Tent chaotic map model is as follows:

$$x_{i+1} = \begin{cases} x_i/\beta & x_i \in (0, \beta] \\ (1 - x_i)/(1 - \beta) & x_i \in (\beta, 1] \end{cases} \tag{17}$$

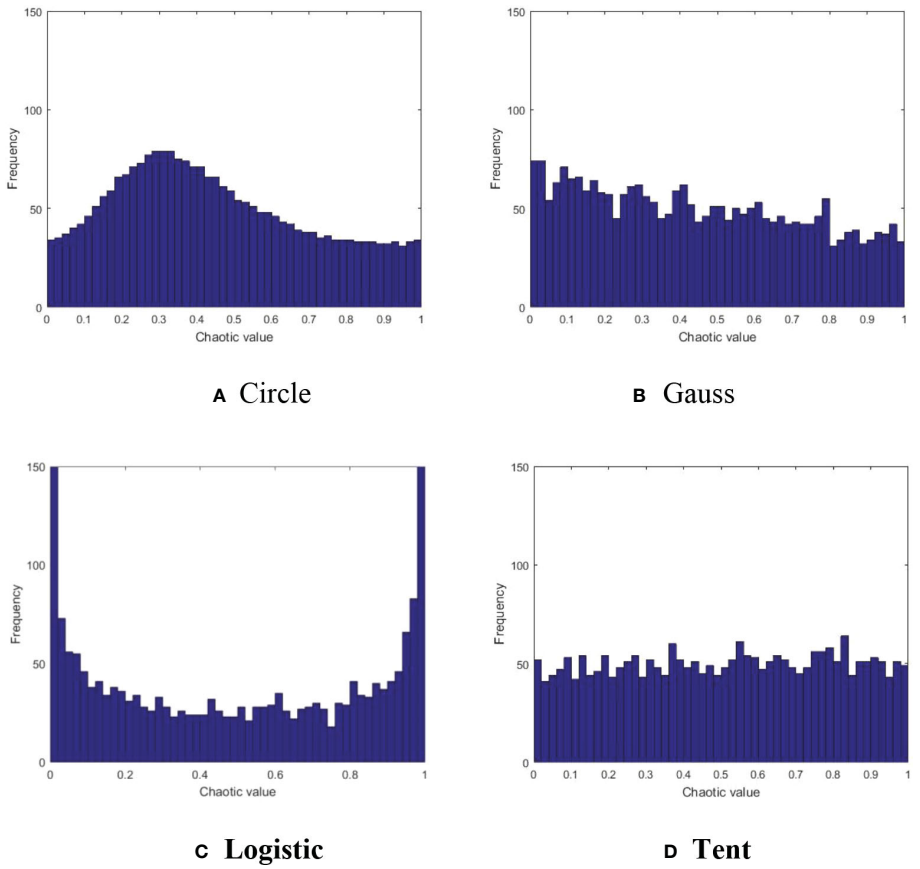


FIGURE 1
Comparison of four chaotic maps. **(A)** Circle **(B)** Gauss **(C)** Logistic **(D)** Tent.

3.2 Opposite-based learning (OBL)

Sihwail et al. (2020) use Opposite-based learning (OBL) to improve meta-heuristic algorithms, proving the ability of OBL to prevent the algorithm falling into local optima. Relying on chaotic reflection, and the initial population individuals, OBL compares the fitness values of the individuals before and after reflection and adds the higher fitness values to the initial population. It improves the quality and diversity of the population. The steps of OBL are as follows.

Assume that the initial population generated by Tent chaotic map is $x_{i,j}$,(i=1,2,3,…,N;j=1,2,3,…,d) , first, sequence the individuals in the current population, select the optimal value as the elite individual $x^e_{i,j} = (x^e_{i,1}, x^e_{i,2}, x^e_{i,3}, \ldots, x^e_{i,d})$, generate the chaotic elite inverse solution $\overline{x^e_{i,j}} = (\overline{x^e_{i,1}}, \overline{x^e_{i,2}}, \ldots, \overline{x^e_{i,d}})$ according to (17), and then deal with the points beyond the boundary according to (18).

$$\overline{x^e_{i,j}} = k \times (\alpha_j + \beta_j) - x^e_{i,j} \tag{18}$$

Where $k \in (0,1)$ is the inverse coefficient, α_j the minimum value of the feasible solution, β_j the maximum value of the feasible solution. Finally, the initial solution generated by chaotic map and the solution generated by OBL are sequenced, and the best first *N* individuals are selected as the initial of the population.

3.3 Position updating

From the model construction in section 2.2, it knows that when the leader α monkey finds the location of the best food, that is, the optimal solution of the problem. Other followers in the population will follow the leader to approach the location of the best food, indicating that the value will increasingly converge to the "optimal value". However, this optimal value is not guaranteed to be the optimal one of the whole search spaces. If the algorithm falls into local optimization, the optimization range will be unexpectedly narrowed, and optimization accuracy undermined. In order to reduce the probability of Capuchin algorithm falling into local optima in iteration, this paper introduces Levy Flight disturbance strategy (He et al., 2023) to disturb the population when updating the position of local exploitation, so that the global search and local exploitation can be balanced. As far as the application of ICAPSA is concerned, Levy Flight is integrated into Jump on the tree [formula (3)], Jumping on the ground [formula (7)-(8)], Swing [formula (9)], Climbing [formula (10)] and Random migration of capuchin monkeys [formula (11)], as described in the Pseudo code of section 3.4. In detail, Levy Flight runs through all stages of population position updating.

Levy Flight is a random walk strategy that conforms to Levy distribution. It is a strategy proposed by academics according to the foraging process of natural organisms. As it has the characteristics of long-distance and short-distance staggering motion and fully random direction, researchers often use it to optimize meta-heuristic algorithms, improve the global search range of the algorithm, and evade falling into the trap of local optima. Its mathematical model is shown in (19):

$$\text{levy}(\alpha) = 0.05 \times \frac{x}{|y|^{1/\alpha}} \tag{19}$$

Where *x* and *y* are two normally distributed variables subject to standard deviation σ_x and σ_y, the calculation formula is as follows:

$$x = \text{Normal}(0, \sigma x_2) \tag{20}$$

$$y = \text{Normal}(0, \sigma y_2) \tag{21}$$

$$\sigma_x = \left[\frac{\Gamma(1+\alpha)\sin(\frac{\pi\alpha}{2})}{\Gamma(\frac{1+\alpha}{2})\alpha 2^{\frac{(\alpha-1)}{2}}}\right]^{1/\alpha}, \sigma_y = 1, \alpha = 1.5 \tag{22}$$

3.4 Pseudo code of ICAPSA

After setting the initial parameter∈ and *P*, Tent Chaotic Mapping and OBL are used to optimize the positions of *N* capuchins. Tent Chaotic Mapping is used to ensure that the initial population has a higher randomness, and OBL strategy can improve the dispersion of the population, which to some extent reduces the risk of the algorithm falling into the local optimum. To balance the global search and local exploitation, Levy Flight is merged in all stages of ICAPSA's population position updating, as show in step 15, step 17, step 19, step 22, step 24 and step 27 of ICAPSA's pseudo code.

Pseudo code: ICAPSA

1: Initialization parameter∈ is a random number inside the range [0,1].
2: Initialization probability parameter *P*=0.5.
3: Initialize the positions of N capuchins with formulas (17) and (18).
4: Calculate the fitness value of each capuchin position.
5: Initialize the velocity of capuchin monkey.
6: Capuchins smaller than *n/2* are randomly selected as leaders and companions, and the remaining capuchins follow the leader.
7: while *t<maxitet*
8: Update the parameter life cycle according to formula (5).
9: For k=1: noP (noP is the number of Capuchins in the population)
10: if (*k<n/2*)
11: Use formula (6) to update the velocity of the leader.
12: if (∈ ≥ 0.1)
13: if (∈ ≥ *P*)
14: if (∈≤ 0.2)
15: Update the position of the leader jumping on the tree with formula (3) and (19).
16: else if (0.2<∈ ≥ 0.30)
17: Update the position of the leader who jumps the riverbank with formula (7) and (19).
18: else
19: Update the position of the leader wandering on the ground with formula (8) and (19).
20: end if
21: else if (0.5<∈ ≤ 0.75)
22: Update the position of the leader swinging between the branches with formula (9) and (19).
23: else if (0.75<∈ ≥ 1.0)
24: Update the position of the leader climbing the tree with formula (10) and (19).
25: end if
26: else
27: Update the position of the leader of the random migration with formula (11) and (19).
28: end if
29: else
30: Update the positions of the followers with formula (16).
31: end if
32: end for

(Continued)

Continued

Pseudo code: ICAPSA
33: Calculate the fitness value of each individual. 34: end while 35: Obtain the best solution

4 Analysis and comparison of experimental results

4.1 Parameter setting and discussion

To demonstrate the efficacy of the improved capuchin search algorithm in plant image thresholding, this paper selects six plant images (Figures 2A-F) in the Kaggle plant image dataset (https://www.kaggle.com/datasets/asheniranga/leaf-disease-dataset-combination, https://www.kaggle.com/datasets/vipoooool/new-plant-diseases-dataset) for verification. In addition to comparing the results with the original capuchin search algorithm, the results are also fully compared and analyzed with FABC (Li et al., 2016), MWOA (Anitha et al., 2021), DCOA (Li et al., 2021) et al. To reflect the best performance of ICAPSA, the parameter setting is discussed first. In this paper, the number of thresholds (NT) is set to 2, 3, 4, 5. the objective function is set as the Kapur which commonly used in thresholding image segmentation (Li et al., 2022), and the parameter selection feature similarity FSIM (Pare et al., 2020) and peak signal-to-noise ratio PSNR (Abualigah et al., 2022) are compared. The experiments performed in our work are run on Windows10-64bit, Intel processor and 16GB running memory and the programming software is Matlab 2016a. The initial experimental parameter values are shown in Table 1.

In Table 1, *noP* represents the number of capuchins, *NT* represents the number of thresholds, *maxiter* represents the maximum number of iterations. Next, we discuss the rationality optimization of parameter selection in Table 1 for plant image threshliding.

To reflect the impact of a single variable *noP* on the experimental results, the threshold number *NT* is set to 5 and the *maxiter* is set to 500. Then, by setting the number of individuals in different populations, the image threshold segmentation of the Leaf01 plant image is performed, and the PSNR value is recorded. The average value is calculated through multiple experiments as shown in Table 2. It can be observed that the best experimental results are achieved when the population individual number *noP* is set to 30.

On the premise that the number of individuals *noP* of Capuchin monkey population is set to 30, to determine the impact of the maximum number of iterations on the segmentation effect, Table 3 lists the maximum PSNR and FSIM values of Leaf01 image through numerous experiments. As can be seen, when the maximum number of iterations is set to 200, the segmentation effect is the best.

4.2 Plant image segmentation results with ICAPSA

To reflect the threshold segmentation efficacy of ICAPSA for different plant images, six plant images are selected, with Kapur as the objective function to obtain the optimal segmentation solution under different threshold numbers. Figures 2, 3 show the visual diagram of segmentation effect based on CAPSA and ICAPSA under different thresholds. Table 4 demonstrates the segmentation quality evaluation results (PSNR and FSIM values) of ICAPSA with different thresholds.

In Figures 2, 3, the segmentation effects are shown when *NT* was set to 2(Figure 2 from (A-i) to (F-i) and Figure 3 from (A-i) to (F-i)), 3 (Figure 2 from (A-ii) to (F-ii) and Figure 3 from (A-ii) to (F-ii)), 4 (Figure 2 from (A-iii) to (F-iii) and Figure 3 from (A-iii) to (F-iii)) and 5 (Figure 2 from (A-iv) to (F-iv) and Figure 3 from (A-iv) to

TABLE 1 Experimental related parameter setting.

Parameter	noP	NT	maxiter
Value	40	2,3,4,5	500

TABLE 2 Effect of experimental results by noP.

noP	10	20	30	40	50	60
PSNR	26.5038	26.7058	**26.7737**	26.6898	26.7129	26.5218
FSIM	0.8513	0.8579	**0.8624**	0.8577	0.8541	0.8508

TABLE 3 Effect of experimental results by maximum iteration number (maxiter).

maxiter	50	100	200	300	400	500
PSNR	25.1092	26.5741	**26.8305**	26.2416	26.6439	26.5458
FSIM	0.8329	0.8385	**0.8583**	0.8136	0.8583	0.8577

10.3389/fpls.2023.1122788

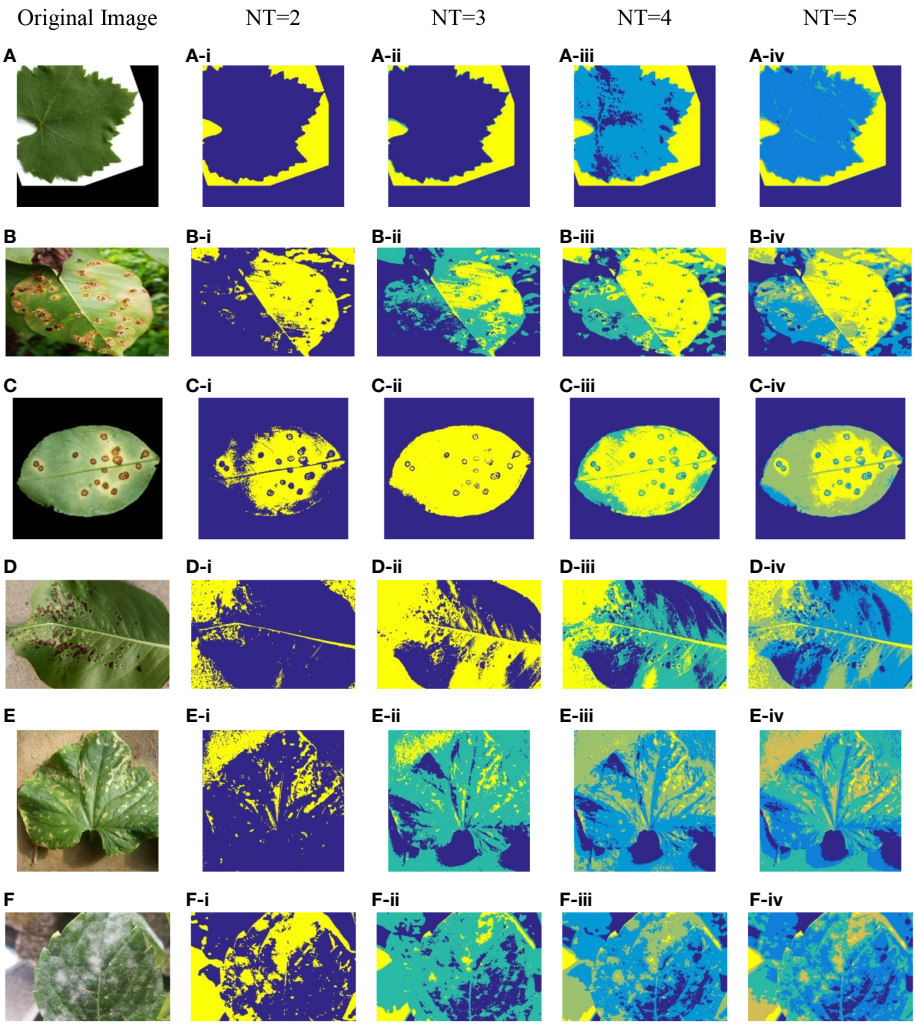


FIGURE 2
Plant image segmentation based on CAPSA. **(A)** Leaf01. **(B)** Apple Brown Spot. **(C)** Citrus Bacterial Canker. **(D)** Black Spot. **(E)** Bacterial Keratosis of Cucumber. **(F)** Powdery Mildew.

(F-iv)) respectively based on original CAPSA and ICAPSA. As can be observed, when *NT* was 2, the background and disease areas were better segmented and located. When *NT* was set to 3, 4 and 5, the segmentation effect was gradually refined, and the regional details were better defined. But in comparison with CAPSA and other algorithms, it is difficult for ICAPSA to directly show its advantages visually. Therefore, in Table 4, the values of PSNR and FSIM obtained in thresholding segmentation process are recorded too. As the number of thresholds gradually increased from 2 to 5 in Table 4, the distribution of thresholds in the range [0 255] was more balanced. In combination with the visual effects of Figure 3, the effectiveness of ICAPSA in plant image thresholding is illustrated from the data level. In addition, more detailed data comparison will be given in Section 4.3.

4.3 Comparison and analysis of similar algorithms

The visual and segmentation data analysis of ICAPSA alone cannot fully demonstrate its consistently reliable performance. To demonstrate the effects of ICAPSA more fully, we refer in this paper to the evaluation methods used in relevant comparative literature. We take PSNR value as the measurement standard, and compare it with the original CAPSA, FABC, DCOA, MWOA and ISBO. On the basis of keeping the original parameters of FABC, DCOA, MWOA and ISBO, the results are provided in Table 5. From the data comparison in Table 5, the effect of ICAPSA is visibly superior to the original CAPSA. The most obvious improvement occurred at Citrus Bacterial Canker image when the threshold number was 2, the value of PSNR increased by 6.0525, a proportional increase of 39.7%. In the worst case is at Powdery Mildew image, when the threshold was 2, the PSNR value also increased by 0.0142, with an increase of 0.07%. But overall, ICAPSA was proved about 10.6% higher than the original CAPSA. Compared with the experimental data of FABC, the effect of ICAPSA is also better than that of FABC. In the Citrus Bactrial Canker image, when the threshold is 2, the effect improvement is the largest, with an increase of about 91.3%. In general, the effect of ICAPSA is 12.5% higher than that of FABC on average. Compared with experimental results of DCOA, similarly, ICAPSA's segmentation effect is slightly higher, with an average increase of 11.7%.

10.3389/fpls.2023.1122788

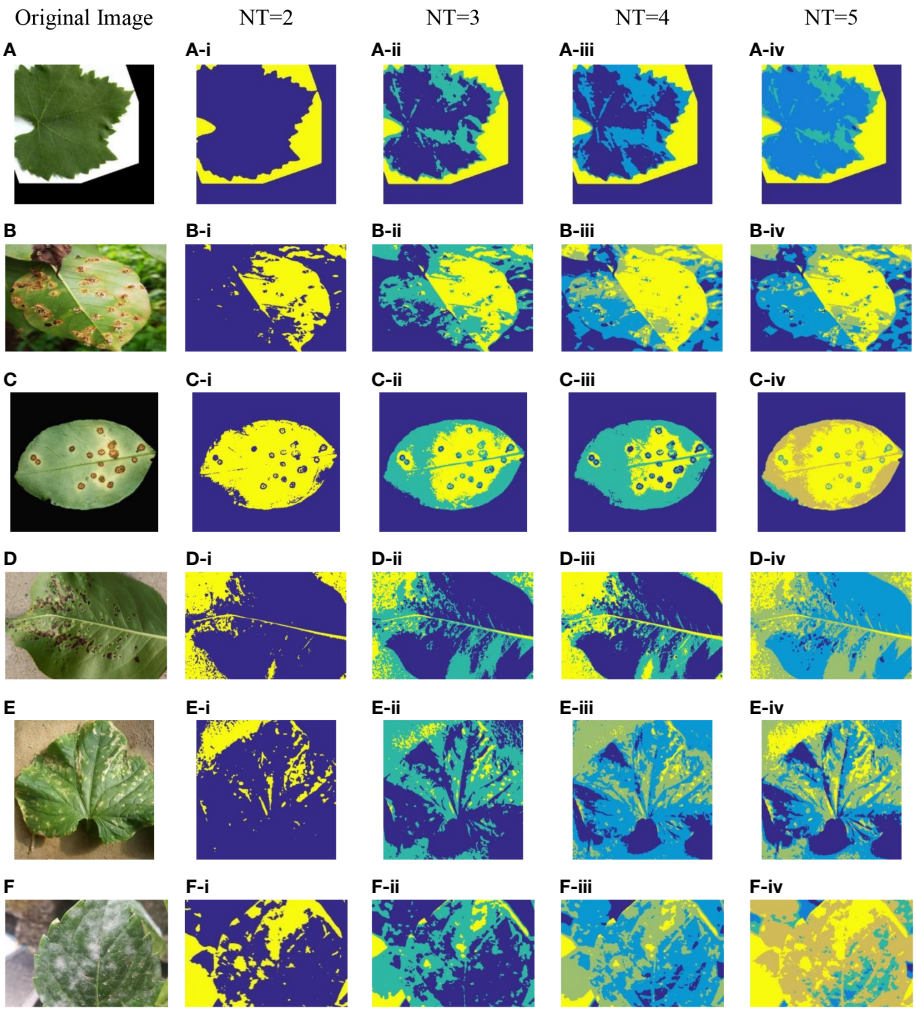


FIGURE 3
Plant image segmentation based on ICAPSA. **(A)** Leaf01. **(B)** Apple Brown Spot. **(C)** Citrus Bacterial Canker. **(D)** Black Spot. **(E)** Bacterial Keratosis of Cucumber. **(F)** Powdery Mildew.

Compared with the experimental data of MWOA, it is found that the segmentation effect of ICAPSA is generally slightly higher than that of MWOA, with an average improvement of 10%. Finally, from the comparison results with ISBO, except at Powdery Mildew image, when the threshold is 4, the effect of ICAPSA is lower than that of ISBO, and the PSNR difference is only 0.2647. From the overall analysis of the results, ICAPSA was about 7.5% higher than ISBO. From the analysis of the above results, it can be concluded that ICAPSA has better effects in plant image segmentation.

TABLE 4 Experimental results of ICAPSA with different threshold numbers.

Image	NT	Thresholds	PSNR	FSIM
Leaf01	2	141 213	16.2911	0.7229
	3	21 155 225	22.5087	0.7858
	4	51 119 190 224	25.7779	0.8015
	5	5 61 130 159 233	26.8305	0.8583
Apple Brown Spot	2	147 221	18.6521	0.6066
	3	63 137 207	21.2129	0.6904
	4	33 82 138 180	22.7506	0.7467
	5	27 116 156 192 252	23.9403	0.7953
Citrus Bacterial Canker	2	44 204	21.3084	0.7655
	3	60 168 232	24.4853	0.8379
	4	79 142 200 254	26.0432	0.8805
	5	13 36 58 122 168	27.1299	0.8943

(Continued)

10.3389/fpls.2023.1122788

TABLE 4 Continued

Image	NT	Thresholds	PSNR	FSIM
Black Spot	2	98 211	21.5449	0.6556
	3	97 162 245	23.7496	0.7590
	4	49 165 218 248	25.1650	0.8137
	5	17 67 134 202 243	26.2449	0.8436
Bacterial keratosis of cucumber	2	134 231	19.2370	0.6028
	3	87 167 214	21.8199	0.7081
	4	28 92 146 246	23.6029	0.7808
	5	36 98 140 204 255	24.6088	0.8402
Powdery Mildew	2	90 205	20.0592	0.5600
	3	94 159 220	21.8992	0.6061
	4	42 111 153 236	23.2750	0.6390
	5	18 58 78 104 200	24.9361	0.7018

TABLE 5 Thresholding results with different segmentation algorithms.

Image	NT	PSNR					
		ICAPSA	CAPSA	FABC	DCOA	MWOA	ISBO
Leaf01	2	**16.2911**	16.1719	16.2788	15.9541	16.1141	16.2897
	3	**22.5087**	16.5023	16.5314	16.5394	16.7720	16.6751
	4	**25.7779**	21.9052	16.6181	16.7702	24.6964	17.5584
	5	**26.8305**	23.7037	16.8697	17.1991	26.2811	26.1164
Apple Brown Spot	2	**18.6521**	17.6643	18.5825	16.2137	17.2934	18.0229
	3	**21.2129**	20.0332	20.1522	20.9419	20.7366	20.5594
	4	**22.7506**	21.1680	21.9783	21.6583	21.8136	21.9044
	5	**23.9403**	22.8429	23.6587	23.3634	23.2905	23.2198
Citrus Bacterial Canker	2	**21.3084**	15.2559	11.1408	20.8239	13.7563	20.0875
	3	**24.4853**	21.1620	24.0375	22.6063	23.0258	23.0702
	4	**26.0432**	23.8553	25.4951	24.8464	24.1417	25.2638
	5	**27.1299**	25.6433	26.5732	26.4201	25.5010	25.7739
Black Spot	2	**21.5449**	18.6999	20.9316	18.8026	15.6189	20.2820
	3	**23.7496**	20.0288	22.1855	21.1415	21.0457	21.3856
	4	**25.1650**	23.7635	23.7159	24.2592	23.4749	23.6836
	5	**26.2449**	24.4805	24.6354	24.4448	24.9746	24.3882
Bacterial keratosis of cucumber	2	**19.2370**	17.3167	19.1072	17.5003	18.9013	19.1536
	3	**21.8199**	20.2393	20.9140	20.8431	19.9737	20.0132
	4	**23.6029**	22.0439	22.7280	22.8016	22.6521	22.1477
	5	**24.6088**	23.6152	23.4007	23.7169	23.4316	23.0604
Powdery Mildew	2	**20.0592**	20.0450	19.8182	17.4870	19.7587	19.3961
	3	**21.8992**	20.9081	21.4543	20.7328	21.3120	21.2948
	4	23.2750	22.1054	23.1606	22.0906	21.6392	**23.5397**
	5	**24.9361**	23.4687	24.3110	23.7426	22.3853	24.6649

5 Conclusion

To improve the accuracy and effects of plant image segmentation, this paper combines and improves the traditional thresholding image segmentation by improving the CAPSA. It uses Tent chaotic map sequence and Opposite-based learning to improve the quality of the initial population and the ability of global optimization in ICAPSA. To avoid the problem of local optimization, Levy Flight disturbance strategy is introduced to make the algorithm mutate when updating the position, so as to balance the global optimization and local exploitation of ICAPSA. Finally, the Kapur entropy is used as the objective function to segment the plant images. The results are compared with CAPSA, FABC, DCOA, MWOA and ISBO. From these results, the improved CAPSA (ICAPSA) demonstrates superior segmentation effects in the field of plant image segmentation.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material. Further inquiries can be directed to the corresponding author.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Funding

This paper is supported by the National Youth Natural Science Foundation of China under Grant 61802208, the Natural Science Foundation of Anhui under Grant 1908085MF207, KJ2020A1215, KJ2020A1216 and KJ2021A1251, the Excellent Youth Talent Support Foundation of Anhui under Grant gxyqZD2019097 and gxyqZD2021142, the Postdoctoral Foundation of Jiangsu under Grant 2018K009B, the Foundation of Fuyang Normal University under Grant TDJC2021008 and the Quality Engineering Project of Anhui under Grant 2021jyxm1117, 2021kcszsfkc307 and 2019sjjd81.

Acknowledgments

Thanks for the support and help of the team when writing the paper. Thanks to the reviewers and experts of this magazine for their valuable opinions on the article revision.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Abdel-Basset, M., Mohamed, R., and Abouhawwash, M. A. (2022). A new fusion of whale optimizer algorithm with kapur's entropy for multi-threshold image segmentation: analysis and validations. *Artif. Intell. Rev.* 55, 6389–6459. doi: 10.1007/s10462-022-10157-w

Abualigah, L., Almotairi, K. H., and Elaziz, M. A. (2022). Multilevel thresholding image segmentation using meta-heuristic optimizationalgorithms: comparative analysis, open challenges and new trends. *Artif. Intell. Rev.* 9, 1–51. doi: 10.1007/s10489-022-04064-4

Anitha, J., Pandian, S. I. A., and Agnes, S. A. (2021). An efficient multilevel color image thresholding based on modified whale optimization algorithm. *Expert Syst. Appl.* 178, 1–12. doi: 10.1016/j.eswa.2021.115003

Bao, X., Jia, H., and Lang, C. (2019). A novel hybrid harris hawks optimization for color image multilevel thresholding segmentation. *IEEE Access.* 7, 76529–76546. doi: 10.1109/ACCESS.2019.2921545

Braik, M., Sheta, A., and Al-Hiary, H. (2021). A novel meta-heuristic search algorithm for solving optimization problems: capuchin search algorithm. *Neural Comput. Appl.* 33, 2515–2547. doi: 10.1007/s00521-020-05145-6

Chen, Y., Wang, M., Heidari, A., Shi, B., Hu, Z., Zhang, Q., et al. (2022). Multi-threshold image segmentation using a multi-strategy shuffled frog leaping algorithm. *Expert Syst. Appl.* 194, 1–12. doi: 10.1016/j.eswa.2022.116511

Chen, C., Wang, X., Heidari, A., Yu, H., and Chen, H. (2021). Multi-threshold image segmentation of maize diseases based on elite comprehensive particle swarm optimization and otsu. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.789911

Elaziz, M. A., Yousri, D., and Mirjalili, S. (2021). A hybrid Harris hawks-moth-flame optimization algorithm including fractional-order chaos maps and evolutionary population dynamics. *Adv. Eng. Software* 154, 1–13. doi: 10.1016/j.advengsoft.2021.102973

Ewees, A., and Abd Elaziz, M. (2020). Performance analysis of chaotic multi-verse harris hawks optimization: a case study on solving engineering problems. *Eng. Appl. Artif. Intel.* 88, 1–15. doi: 10.1016/j.engappai.2019.103370

Ewees, A. A., Elaziz, M. A., Al-Qaness, M. A., Khalil, H. A., and Kim, S. (2020). Improved artificial bee colony using sine-cosine algorithm for multi-level thresholding image segmentation. *IEEE Access.* 8, 26304–26315. doi: 10.1109/ACCESS.2020.2971249

Gholizadeh, S., and Baghchevan, A. (2017). Multi-objective seismic design optimization of steel frames by a chaotic meta-heuristic algorithm. *Eng. Comput.* 33, 1045–1060. doi: 10.1007/s00366-017-0515-0

Hasan, S., Jahan, S., and Islam, M. I. (2022). Disease detection of apple leaf with combination of color segmentation and modified DWT. *J. King Saud Univ-com* 34, 7212–7224. doi: 10.1016/j.jksuci.2022.07.004

He, Q., Liu, H., Ding, G., and Tu, L. (2023). A modified levy flight distribution for solving high-dimensional numerical optimization problems. *Math. Comput. Simulat.* 204, 376–400. doi: 10.1016/j.matcom.2022.08.017

Hemasian-Etefagh, F., and Safi-Esfahani, F. (2019). Dynamic scheduling applying new population grouping of whales meta-heuristic in cloud computing. *J. Supercomput.* 75, 6386–6450. doi: 10.1007/s11227-019-02832-7

Kalyani, R., Sathya, P. D., and Sakthivel, V. P. (2021). Medical image segmentation using exchange market algorithm. *Alex. Eng. J.* 60, 5039–5063. doi: 10.1016/j.aej.2021.04.054

Kittler, J., and Illingworth, J. (1986). Minimum error thresholding. *Pattern Recogn.* 19, 41–47. doi: 10.1016/0031-3203(86)90030-0

Li, Y., Han, M., and Guo, Q. (2020). Modified whale optimization algorithm based on tent chaotic mapping and its application in structural optimization. *KSCE J. Civ. Eng.* 24, 3703–3713. doi: 10.1007/s12205-020-0504-5

Lin, Q., Zhang, L., Wu, T., and Mean, T. (2020). Application of tsallis cross-entropy in image thresholding segmentation. *Sensor. Ma-ter.* 32, 2687–2696. doi: 10.18494/SAM.2020.2798

Li, L., Qian, S., Li, Z., and Li, S. (2022). Application of improved satin bowerbird optimizer in image segmentation. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.915811

Li, L., Sun, L., Jian, G., Chong, H., and Li, S. (2016). Fuzzy multilevel image thresholding based on modified quick artificial bee colony algorithm and local information aggregation. *Math. Probl. Eng.* 2016, 1–18. doi: 10.1155/2016/5985616

Li, L., Sun, L., Xue, Y., Li, S., Huang, X., and Mansour, R. F. (2021). Fuzzy multilevel image thresholding based on improved coyote optimization algorithm. *IEEE Access.* 9, 33595–33607. doi: 10.1109/ACCESS.2021.3060749

Ma, M., Liang, J., Guo, M., Fan, Y., and Yin, Y. (2011). SAR image segmentation based on artificial bee colony algorithm. *Appl. Soft Comput.* 11, 5205–5214. doi: 10.1016/j.asoc.2011.05.039

Maray, M., Albraikan, A. A., Alotaibi, S. S., Alabdan, R., Al Duhayyim, M., Al-Azzawi, W. K., et al. (2022). Artificial intelligence-enabled coconut tree disease detection and classification model for smart agriculture. *Comput. Electr. Eng.* 104, 1–15. doi: 10.1016/j.compeleceng.2022.108399

Ma, G., and Yue, X. (2022). An improved whale optimization algorithm based on multilevel threshold image segmentation using the otsu method. *Eng. Appl. Artif. Intel.* 113, 1–28. doi: 10.1016/j.engappai.2022.104960

Merzban, M. H., and Elbayoumi, M. (2019). Efficient solution of otsu multilevel image thresholding: A comparative study. *Expert Syst. Appl.* 166, 299–309. doi: 10.1016/j.eswa.2018.09.008

Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE Trans.Syst. Man Cybernet.* 9, 62–66. doi: 10.1109/TSMC.1979.4310076

Pare, S., Kumar, A., Singh, G. K., and Bajaj, V. (2020). Image segmentation using multilevel thresholding: a research review. *IJST-T Electr. Eng.* 44, 1–29. doi: 10.1007/s40998-019-00251-1

Patel, R., Mitra, B., Vinchurkar, M., Adami, A., Patkar, R., Giacomozzi, F., et al. (2023). Plant pathogenicity and associated/related detection systems a review. *Talanta* 251, 1–9. doi: 10.1016/j.talanta.2022.123808

Prasad, D., Mukherjee, A., and Mukherjee, V. (2021). Temperature dependent optimal power flow using chaotic whale optimization algorithm. *Expert Syst.* 38, 1–10. doi: 10.1111/exsy.12685

Rajinikanth, V., Kadry, S., Crespo, R. G., and Verdu, E. (2021). A study on RGB image multi-thresholding using Kapur/Tsallis entropy and moth-flame algorithm. *Int. J. Interact. Multi.* 7, 163–171. doi: 10.9781/ijimai.2021.11.008

Rodríguez-Esparza, E., Zanella-Calzada, L. A., Oliva, D., Heidari, A. A., Zalldivar, D., Pérez-Cisneros, M., et al. (2020). An efficient harris hawks-inspired image segmentation method. *Expert Syst. Appl.* 155, 1–12. doi: 10.1016/j.eswa.2020.113428

Ruwona, J., and Scherm, H. (2022). Sensing and imaging of plant disease through the lens of science mapping. *Trop. Plant Pathol.* 47, 74–84. doi: 10.1007/s40858-021-00478-6

Sathya, P. D., Kalyani, R., and Sakthivel, V. P. (2021). Color image segmentation using kapur, otsu and minimum cross entropy functions based on exchange market algorithm. *Expert Syst. Appl.* 172, 1–16. doi: 10.1016/j.eswa.2021.114636

Sihwail, R., Omar, K., Ariffin, K. A. Z., and Tubishat, M. (2020). Improved harris hawks optimization using elite opposition-based learning and novel search mechanism for feature selection. *IEEE Access.* 8, 121127–121145. doi: 10.1109/ACCESS.2020.3006473

Upadhyay, P., and Chhabra, J. K. (2020). Kapur's entropy based optimal multilevel image segmentation using crow search algorithm. *Appl. Soft Comput.* 97, 1–15. doi: 10.1016/j.asoc.2019.105522

Wang, Y., and Song, S. (2022). An adaptive firefly algorithm for multilevel image thresholding based on minimum cross entropy. *J. Supercomput.* 78, 11580–11600. doi: 10.1007/s11227-021-04281-7

Xue, Y., Tang, Y., Xu, X., Liang, J., and Neri, F. (2022). Multi-objective feature selection with missing data in classification. *IEEE Trans. Emerg. Top. Comput. Intell.* 6, 355–364. doi: 10.1109/TETCI.2021.3074147

Xue, Y., Xue, B., and Zhang, M. (2019). Self-adaptive particle swarm optimization for large-scale feature selection in classification. *ACM T. Knowl. Discov D.* 13, 1–27. doi: 10.1145/3340848

Xue, Y., Zhu, H., Liang, J., and Slowik, A. (2021). Adaptive crossover operator based multi-objective binary genetic algorithm for feature selection in classification. *Knowl-Based Syst.* 227, 1–9. doi: 10.1016/j.knosys.2021.107218

Zhao, D., Liu, L., Yu, F., Heidari, A. A., and Chen, H. (2021). Chaotic random spare ant colony optimization for multi-threshold image seg-mentation of 2D kapur entropy. *Knowl-Based Syst.* 216, 1–18. doi: 10.1016/j.knosys.2020.106510

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 14 February 2023
DOI 10.3389/fpls.2023.1146388

OPEN ACCESS

EDITED BY
Ferrante Neri,
University of Surrey, United Kingdom

REVIEWED BY
Hongxiu Wan,
Nanjing Institute of Geography and Limnology, CAS, China
Jun Geng,
Hefei University of Technology, China

*CORRESPONDENCE
Yongming Xu
xym30@nuist.edu.cn

SPECIALTY SECTION
This article was submitted to Sustainable and Intelligent Phytoprotection, a section of the journal Frontiers in Plant Science

RECEIVED 17 January 2023
ACCEPTED 31 January 2023
PUBLISHED 14 February 2023

CITATION
Lu X, Chen Y, Sun Y, Xu Y, Xin Y and Mo Y (2023) Spatial and temporal variations of net ecosystem productivity in Xinjiang Autonomous Region, China based on remote sensing.
Front. Plant Sci. 14:1146388.
doi: 10.3389/fpls.2023.1146388

Spatial and temporal variations of net ecosystem productivity in Xinjiang Autonomous Region, China based on remote sensing

Xiangjun Lu[1], Yang Chen[2,3], Yuyin Sun[1], Yongming Xu[2]*, Yan Xin[2] and Yaping Mo[2]

[1]The First Ecological and Environment Monitoring Station of Xinjiang Production and Construction Corps, Urumchi, China, [2]School of Remote Sensing and Geomatics Engineering, Nanjing University of Information Science & Technology, Nanjing, China, [3]Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, China

Net ecosystem productivity (NEP), which plays a key role in the carbon cycle, is an important indicator of the ecosystem's carbon budget. In this paper, the spatial and temporal variations of NEP over Xinjiang Autonomous Region, China from 2001 to 2020 were studied based on remote sensing and climate re-analysis data. The modified Carnegie Ames Stanford Approach (CASA) model was employed to estimate net primary productivity (NPP), and the soil heterotrophic respiration model was used to calculate soil heterotrophic respiration. Then NEP was obtained by calculating the difference between NPP and heterotrophic respiration. The annual mean NEP of the study area was high in the east and low in the west, high in the north and low in the south. The 20-year mean vegetation NEP of the study area is 128.54 $gC{\cdot}m^{-2}$, indicating that the study area is a carbon sink on the whole. From 2001 to 2020, the annual mean vegetation NEP ranged between 93.12 and 158.05 $gC{\cdot}m^{-2}$, and exhibited an increasing trend in general. 71.46% of the vegetation area showed increasing trends of NEP. NEP exhibited a positive relationship with precipitation and a negative relationship with air temperature, and the correlation with air temperature was more significant. The work reveals the spatio-temporal dynamics of NEP in Xinjiang Autonomous Region and can provide a valuable reference for assessing regional carbon sequestration capacity.

KEYWORDS

NEP, spatio-temporal variations, remote sensing, Xinjiang, CASA

1 Introduction

Since the industrial revolution, the concentration of carbon dioxide in the atmosphere has dramatically increased with the development of technology and the widespread use of fossil fuels. In 1705, the global atmospheric CO_2 concentration is 278 ppm, but in 2022, it has reached 417 ppm, an increase of about 50% (Met Office, 2022; NOAA, 2022). The increase in atmospheric CO_2 concentration has caused a series of climate and environmental problems,

including global warming, extreme weather events, sea level rising, and glacier retreat, which have important impacts on human survival and development (Zhao and Luo, 1998; Hoegh and Bruno, 2010; IPCC, 2014; Baker et al., 2018; Luo and Keenan, 2022; Xue and Qin, 2022). Carbon cycle research has attached strong interest from governments and scientists. Terrestrial ecosystem is an important part of the global carbon cycle and a key indicator of regional environment and human activities (Schimel et al., 2015; Friend et al., 2014). Net ecosystem productivity (NEP), which is the difference between net primary productivity (NPP) and soil heterotrophic respiration (Rh), can effectively indicate carbon absorption and emission capacity, and is generally used as an indicator of carbon source or carbon sink (Keenan et al., 2016; Li et al., 2020). A positive NEP means that the carbon absorbed by the ecosystem is more than the carbon emitted, which is a carbon sink. A negative NEP indicates that the carbon absorbed by the ecosystem is less than the carbon emitted, which is a carbon source. Accurate knowledge of NEP is critically important for assessing regional carbon sequestration capacity and developing strategies to stabilize the CO_2 emissions.

Traditional investigation methods of NPP include the biomass survey method, the direct harvest method, Eddy Covariance method, chlorophyll estimation method, and the raw material consumption measurement method. Most of these methods are time-consuming and can only provide NPP estimations at small scales. Recently, satellite remote sensing data are employed to map NPP at large scales, and several methods were developed for estimating NPP from remote sensing data, including light energy utilization models, physiological and ecological process models, and climate productivity models (Goetz et al., 1999; Parton et al., 1993; Veroustraete et al., 1994). Among these methods, the Carnegie-Ames-Stanford Approach (CASA) model, which is a simple light energy utilization model that uses remotely sensed normalized difference vegetation index (NDVI) and meteorological data as input variables, has been widely used. Many studies have been carried out on the spatial and temporal variations of NPP based on remote sensing datasets (Bradford et al., 2005; Chirici et al., 2007; Pachavo and Murwira, 2014; Ji et al., 2021; Xue et al., 2021; Mngadi et al., 2022). NPP can well quantify the carbon absorption by plants; while NEP can depict both the carbon absorption by plants and carbon release by soils. Compared with NPP, NEP can better depict regional carbon sequestration capacity. However, there are quite fewer studies investigating NEP from remote sensing at large scales. Yamaji et al. (2008) combined Moderate Resolution Imaging Spectroradiometer (MODIS) data and flux-based observations to obtain NEP maps of deciduous forests in Japan for the years 2002-2003 using a scaling-up technology. Nayak et al. (2015) analyzed the spatial and temporal variations of NEP over India during 1981–2006 using the CASA model and regional database and also investigated its relationship with climatic parameters. He et al. (2018) used the CASA model and a soil microbial respiration model to estimate NEP in the Yellow River Basin during 1982-2015 and analyzed its spatial response under diurnal asymmetric warming. Zhao et al. (2019) used an improved individual-based forest ecosystem carbon budget model and Advanced Very High Resolution Radiometer (AVHRR) remote sensing data to estimate the NEP of global forest ecosystems from 1982 to 2011, and then investigated the impacts of climate change on the NEP of different forest types. Guo et al. (2021) applied the CASA model and a heterotrophic respiration model to map NEP in the Hindu Kush Himalayan region and studied the temporal and spatial changes of NEP magnitude from 2001 to 2018. Zhang J. et al. (2021) combined the CASA and empirical models to map the NEP in Central Asia during 2001-2019 and evaluated the impact of drought on the carbon source and sink.

Xinjiang Autonomous Region is located in an arid and semi-arid climate zone in northwestern China. It is one of the most sensitive regions to climate change, which is characterized by relatively low ecosystem productivity and weak system regulation capacity. Quantitative investigating the ecosystem NEP in this region and the response to climate is of great significance. This paper aims to integrate satellite remote sensing data and meteorological reanalysis data to study the spatial and temporal variations of net ecosystem productivity in the region in the past 20 years, which is of great significance for the understanding of changes in carbon balances and expenditures in Xinjiang, and can provide a reference for the regulation of carbon cycle and ecological protection strategy in the region.

2 Study area and data

2.1 Study area

Xinjiang Uygur Autonomous Region is located in the northwest of China (Figure 1). It ranges from 73°20' to 96°25' E and 34°15' to 49° 10' N, with a total area of about 1.66 million km^2. The region has a complex and diverse topography, with the Altai, Tianshan, and Kunlun Mountains distributed from north to south. The Junggar and Tarim basins are located on the north and south sides of Tianshan Mountain, respectively. Xinjiang belongs to a temperate continental arid and semi-arid climate, characterized by long hours of sunshine, strong evaporation, hot summers, and cold winters. In general, precipitation is higher in the north than in the south, and temperature is higher in the south than in the north. The vegetation cover is low, and the vast majority of the area is covered by desert, followed by grassland, which accounts for 70.90% and 22.95% of the total area, respectively.

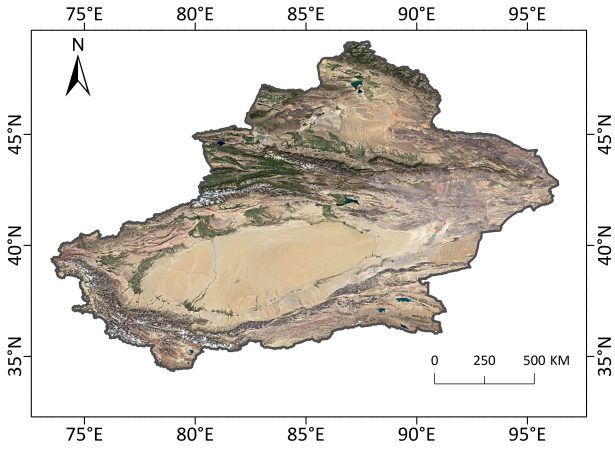


FIGURE 1
Study area.

2.2 Data

The remote sensing data used in this study include MODIS vegetation index product (MOD13Q1) and MODIS land cover product (MCD12Q1) from 2001 to 2020. MOD13Q1 is a 250m resolution 16-day composite MODIS vegetation index product that provides two primary vegetation indices: NDVI and enhanced vegetation index (EVI). The multi-temporal composite algorithm selects the best available pixel values during the 16-day period to produce the composite vegetation index values, with the criteria of low cloud, low view angle, and the highest NDVI/EVI value. In this study, the NDVI was extracted and then monthly composited. MCD12Q1 is a 500m resolution annual MODIS land cover product that provides land cover data under five different classification schemes. It is derived using supervised classification methods based on Terra and Aqua MODIS reflectance data and ancillary information. In this study, the land cover using the International Geosphere–Biosphere Programme (IGBP) scheme was used. In addition, the Shuttle Radar Topography Mission (SRTM) Digital Elevation Model (DEM) was also employed in this study to provide 90m resolution elevation.

The meteorological reanalysis data is the ERA5-Land meteorological data, which combines model data with observation data into a globally complete and consistent dataset using the laws of physics. It provides a series of meteorological elements with a spatial resolution of 0.1 degree. In this study, monthly mean temperature and monthly total precipitation from 2001 to 2020 were extracted.

The observed monthly sunlight hours and precipitation data were collected from 53 meteorological stations in Xinjiang. The sunlight hours were interpolated to gridded data with a spatial resolution of 1km based on the ordinary kriging method. Compared with the reanalysis precipitation and meteorological observed precipitation, the reanalysis precipitation is generally higher than observed precipitation. To reduce the error of reanalysis precipitation, observed precipitation from 53 meteorological stations in Xinjiang were employed to develop linear regressions between the reanalysis precipitation and meteorological observed precipitation month by month. Then the reanalysis precipitation were adjusted based on these equations.

All data were processed in Google Earth Engineering (GEE) platform, including mosaic, spatial clip, projection conversion, mask, monthly composite processing, spatial resampling, etc. Finally, all data were converted into Albers projection with a spatial resolution of 500m.

3 Methods

3.1 NPP estimation

The CASA (Carnegie–Ames–Stanford approach) model was employed for NPP estimation. It has the advantages of high accuracy and fewer input parameters and, and has been widely used for NPP estimation. Its main formula is:

$$NPP_{(x,t)} = APAR_{(x,t)} \cdot \epsilon_{(x,t)} \quad (1)$$

where $APAR_{(x,t)}$ is the absorbed photosynthetic active radiation of pixel x in month t ($MJ{\cdot}m^{-2}$); $\epsilon_{(x,t)}$ is light utilization efficiency of pixel x in month t ($gC{\cdot}MJ^{-1}$).

APAR is calculated by the following equation:

$$APAR_{(x,t)} = SOL_{(x,t)} \cdot FPAR_{(x,t)} \cdot 0.5 \quad (2)$$

Where $SOL_{(x,t)}$ is the total solar radiation of pixel x in month t ($MJ{\cdot}m^{2}$), FPAR is the fraction of absorbed photosynthetically active radiation.

The total solar radiation was calculated based on the empirical equation proposed by He and Xie (2010), which was developed for Western China:

$$Q = Q_0(a + b \cdot T_S/T_A) \quad (3)$$

Where Q_0 is the astronomical radiation, which is calculated based on geographical latitude, solar declination and other parameters; T_S is the observed sunlight hours; T_A is the theoretical sunlight hours which is calculated from latitude and solar declination; a and b are empirical coefficients (a=0.185; b=0.595).

FPAR is calculated by Equation 4~7:

$$FPAR_{(x,t)} = \alpha \mathrm{FPAR_{NDVI}} + (1-\alpha)\mathrm{FPAR_{SR}} \quad (4)$$

$$\mathrm{FPAR_{NDVI}} = \frac{\mathrm{NDVI}_{(x,t)} - \mathrm{NDVI}_{(i,min)}}{\mathrm{NDVI}_{(i,max)} - \mathrm{NDVI}_{(i,min)}} \times (\mathrm{FPRA}_{max} - \mathrm{FPRA_{min}}) + \mathrm{FPRA_{min}} \quad (5)$$

$$\mathrm{FPAR_{SR}} = \frac{\mathrm{SR}_{(x,t)} - \mathrm{SR}_{(i,min)}}{\mathrm{SR}_{(i,max)} - \mathrm{SR}_{(i,min)}} \times (\mathrm{FPRA}_{max} - \mathrm{FPRA_{min}}) + \mathrm{FPRA_{min}} \quad (6)$$

$$\mathrm{SR}_{(x,t)} = \frac{1 + \mathrm{NDVI}_{(x,t)}}{1 - \mathrm{NDVI}_{(x,t)}} \quad (7)$$

Where $NDVI_{(i,min)}$ and $NDVI_{(i,max)}$ are the minimum and maximum values of NDVI for vegetation type i, respectively; the values of $FPAR_{min}$ and $FPAR_{max}$ were 0.001 and 0.95, respectively; $SR_{(i,min)}$ and $SR_{(i,min)}$ are the 5% and 95% percentile of NDVI values for vegetation type i, respectively.

The light utilization energy ϵ is calculated by the following equation:

$$\epsilon_{(x,t)} = T_{\epsilon 1(x,t)} \cdot T_{\epsilon 2(x,t)} \cdot W_{\epsilon(x,t)} \cdot \epsilon_{max} \quad (8)$$

Where $T_{\epsilon 1}$ and $T_{\epsilon 2}$ represent the stressing effects of low and high temperatures on light energy utilization, respectively; $W_{\epsilon(x,t)}$ is the water stress effect coefficient; ϵ_{max} is the maximum light utilization efficiency under ideal conditions($gC{\cdot}MJ^{-1}$).

Table 1 gives the values of $NDVI_{(i,min)}$, $NDVI_{(i,max)}$, $NDVI_{(i,min)}$ and $NDVI_{(i,max)}$ and ϵ_{max} of different vegetation types (Zhu et al., 2007).

3.2 NEP calculation

Net Ecosystem Productivity (NEP) is defined as the difference between NPP and soil heterotrophic respiration:

$$NEP = NPP - R_H \quad (9)$$

where NEP is net ecosystem productivity of vegetation ($gC{\cdot}m^{-2}$); NPP is net primary productivity of vegetation ($gC{\cdot}m^{-2}$); RH is soil heterotrophic respiration ($gC{\cdot}m^{-2}$).

TABLE 1 The values of $NDVI_{(i,min)}$, $NDVI_{(i,max)}$, $NDVI_{(i,min)}$ and $NDVI_{(i,max)}$ and ϵ_{max} of different vegetation types.

Vegetation Type	$NDVI_{max}$	$NDVI_{min}$	SR_{max}	SR_{min}	ϵ_{max}
Evergreen needleleaved forest	0.647	0.023	4.67	1.05	0.389
Deciduous coniferous forest	0.738	0.023	6.63	1.05	0.485
Deciduous broadleaf forest	0.747	0.023	6.91	1.05	0.692
Mixed forest	0.702	0.023	5.84	1.05	0.475
Shrubland	0.636	0.023	4.49	1.05	0.429
Grassland	0.634	0.023	4.46	1.05	0.542
Cropland	0.634	0.023	4.46	1.05	0.542
Water	0.634	0.023	4.46	1.05	0.542
Impervious surface	0.634	0.023	4.46	1.05	0.542
Barren	0.634	0.023	4.46	1.05	0.542

Soil heterotrophic respiration is calculated by the empirical equation developed by Pei et al. (2003).

$$R_H = 3.069 \cdot (e^{0.0912 \cdot T} + \ln(0.3145 \cdot R + 1)) \quad (10)$$

Where RH is soil heterotrophic respiration ($gC \cdot m^{-2}$); T is air temperature (°C); R is precipitation (mm).

3.3 Spatio-temporal variations

The Theil-Sen Median (Sen) method was employed to determine the trend of NEP over the study area. This method is a robust nonparametric trend method, which does not require the data to follow a certain distribution (Fensholt et al., 2012; Zhang Z. et al., 2021). It has been widely used in the trend analysis of long-time data series.

The Sen's slope value (β) indicates the magnitude of the NEP trend. A positive β value suggests an upward trend and a negative β value suggests a downward trend. The calculation formula for β is as follows:

$$\beta = \text{Median}\left(\frac{NEP_j - NEP_i}{j - i}\right) \quad 2001 \le i < j \le 2020 \quad (11)$$

where β is Sen's slope value; NEP_i and NEP_j are NEP in year i and j respectively.

The Mann–Kendall test is used to assess the significance of NEP trends. The test statistics S value is calculated as:

$$S = \sum_{i-1}^{n-1} \sum_{j=1+1}^{n} (\text{sgn} NEP_j \_ NEP_i) \quad (12)$$

$$\text{sgn}(NEP_j \_ NEP_i) = \begin{cases} 1 & NEP_j \_ NEP_i > 0 \\ 0 & NEP_j \_ NEP_i = 0 \\ -1 & NEP_j \_ NEP_i < 0 \end{cases}$$

The test statistic Z value was obtained by standardizing S:

$$Z = \begin{cases} \frac{S-1}{\sqrt{\text{var}(S)}} & S > 0 \\ 0 & S = 0 \\ \frac{S+1}{\sqrt{\text{var}(S)}} & S < 0 \end{cases} \quad (13)$$

Where the function var is calculated as:

$$\text{var}(S) = \frac{(n(n-1)(2n+5) - \sum_{i=1}^{m} t_i(t_i-1)(2t_i+5))}{18} \quad (14)$$

where n is the number of sequence samples; m is the number of affiliated groups in the data sequence; t_i is the number of input values inside the affiliated group i.

To validate the significance of the trend, the absolute z-score value |Z| is compared with the critical value $Z_{1-\alpha/2}$ at a given significance level α. if |Z| is higher than $Z_{1-\alpha/2}$, the trend is considered significant. The $Z_{1-\alpha/2}$ values were obtained from the standard normal distribution table. For the significance level α of 0.05 and 0.01, the critical $Z_{1-\alpha/2}$ values are 1.96 and 2.58, respectively

Partial correlation is used to analyze the relationship between NEP and meteorological factors. Its formula is:

$$r_{ij \cdot k} = \frac{r_{ij} - r_{ik} r_{jk}}{\sqrt{(1 - r_{ik}^2)(1 - r_{jk}^2)}} \quad (15)$$

where $R_{ij,k}$ is the partial correlation coefficient between variable i and j after variable k is fixed. r_{ij}, r_{ih}, r_{jh} are correlation coefficients for the variables i and j, j and k, and k and i, respectively.

4 Results and discussions

4.1 Spatial distributions of NEP

Figure 2 shows the spatial distributions of the 20-year mean NEP of Xinjiang Autonomous Region. NEP in Xinjiang shows a spatial distribution pattern of high in the east and low in the west, high in the north and low in the south. During 2001-2020, the overall average vegetation NEP of the study area was 128.54 $gC \cdot m^{-2}$, suggesting an overall performance of carbon sink. However, the spatial differences were obvious. 61.56% of the vegetation area in the region had positive NEP, showing carbon sink effects. 38.44% of the vegetation area had negative NEP, showing a carbon source effect. The carbon source is mainly distributed in the central Junggar Basin, Altay Mountain, and southern Tacheng Basin.

4.2 Temporal variations of NEP

Figure 3 shows the temporal variations of annual mean vegetation NEP over Xinjiang from 2001 to 2020. Generally, the annual mean vegetation NEP ranged from 93.12 to 158.05 $gC \cdot m^{-2}$, exhibiting an increasing trend with a Sen's slope of 1.59 $gC \cdot m^{-2} \cdot a^{-1}$. During these 20 years, the annual mean vegetation NEP was the lowest in 2008 (93.12 $gC \cdot m^{-2}$) and the highest in 2013 (158.05 $gC \cdot m^{-2}$).

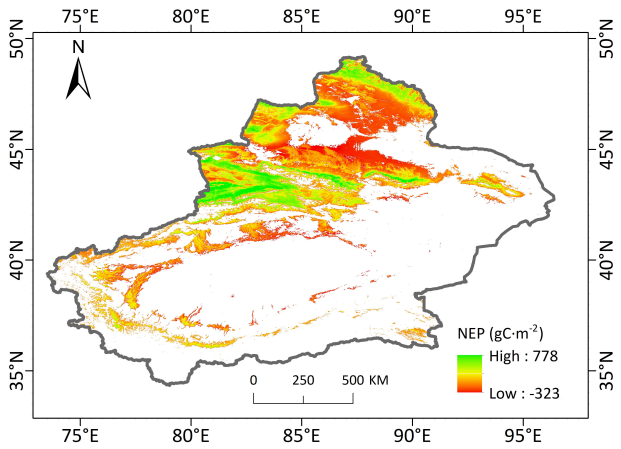


FIGURE 2
Spatial distribution of 20-year mean NEP in Xinjiang.

The monthly mean vegetation NPP in Xinjiang ranged from -4.62 to 39.81 gC·m^{-2}, showing obvious single-peaked characteristics (Figure 4). Among the 12 months, NEP showed positive values from April to October and achieved the highest value in July (39.81 gC·m^{-2}), indicating that during these months Xinjiang vegetation ecosystem acted as a carbon sink. From November to March, Xinjiang vegetation ecosystem showed a carbon source effect, and the lowest monthly NEP value was -4.62 gC·m^{-2} in February.

To better understand the spatial and temporal variations of NEP over the study area, the Sen's trend analysis and Mann–Kendall test were performed pixel by pixel. Figure 5 gives the spatial distributions of the trend of vegetation NEP and the significance of the trend over Xinjiang from 2001 to 2020. During this period, most of the vegetation area in the region showed increasing trends. The area with increasing NEP trend was ~327, 400 km^2, accounting for 71.46% of the total vegetation area. Among these areas, about 51.63% had the NEP trend lower than 0.15 gC·m^{-2}·a^{-1}, 21.65% had the NEP trend between 0.15 and 0.25 m^{-2}·a^{-1}, and 20.09% had the NEP trend between 0.25 and 0.5 gC·m^{-2}·a^{-1}. 6.64% had NEP growth rates greater than 0.5 gC·m^{-2}·a^{-1}, mainly in the oasis belt of the northern slopes of the Tianshan Mountain, the oasis belt around the Tarim Basin, the Altai Mountain, and the northern part of the Tacheng Basin. The area where NEP showed decreasing trends was ~130, 800 km^2, accounting for 28.54% of the total vegetation area. Most of these areas had NEP decline rates below 0.15 gC·m^{-2}·a^{-1}, accounting for about 79.61% of the vegetation area with decreasing NEP trends. These areas were mainly located in the Ili Basin, the northern part of the Junggar Basin, the Altai Mountain and the south part of the Tacheng Basin. 11.34% had NEP decline rates between 0.15 and 0.25 gC·m^{-2}·a^{-1}, 7.92% had NEP decline rates between 0.25 and 0.5 gC·m^{-2}·a^{-1}. Only 1.13% had NEP decline rates over 0.5 gC·m^{-2}·a^{-1}, which were sporadically distributed throughout the region. In terms of the distribution of significant levels of trends (Figure 5B), 22.33% of the vegetation area in Xinjiang showed highly significant increasing trends, mainly in the oasis belt in the middle of the northern slope of the Tianshan Mountain and the oasis belt around the Tarim Basin. There were also 11.01% of vegetation area showed significant increasing trends. Most of the vegetation area in the study area showed insignificant changes, accounting for 63.23% of the total vegetation area. Among them, 38.12% had insignificantly increasing trends and 25.11% had insignificantly decreasing trends. Only 2.16% and 1.27% of the vegetation area showed significant decreasing and highly significant decreasing trends, respectively. In general, the area with increasing NEP was obviously larger than that with decreasing NEP, and the NEP decreasing trends were mostly insignificant, suggesting good vegetation recovery in Xinjiang.

4.3 Response of NEP to meteorological factors

Figure 6 shows the 20-year average annual precipitation and temperature over Xinjiang. This regions exhibited strong spatial difference in precipitation. The Yili Basin region and the northern slopes of the Tianshan Mountain had the highest precipitation. This can be attributed to the fact that the Tianshan Mountain block the warm and humid Atlantic air currents, creating more precipitation on the windward slopes. And the open valley floor in the western part of the Yili Basin allows the entry of the humid Atlantic air currents. The Tarim Basin, Turpan Basin and Junggar Basin are surrounded by high mountains, making it difficult for water vapor to get in, resulting in

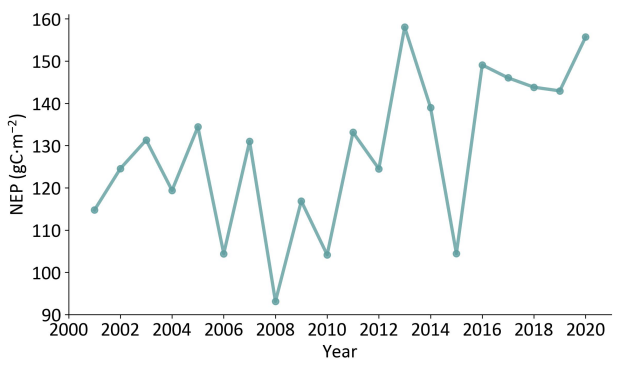


FIGURE 3
Temporal variations of annual mean vegetation NEP over Xinjiang from 2001 to 2020.

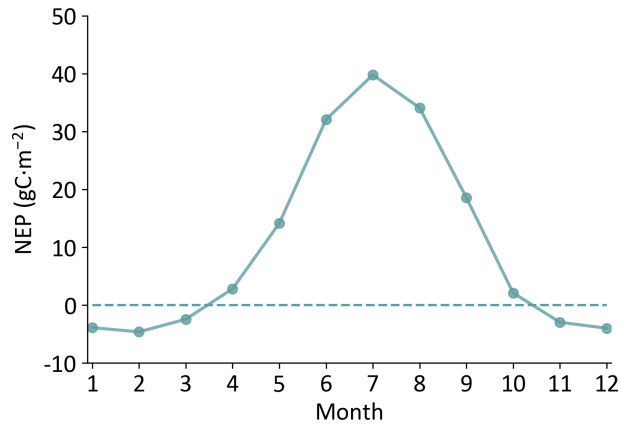


FIGURE 4
Temporal variations of monthly mean NEP over Xinjiang.

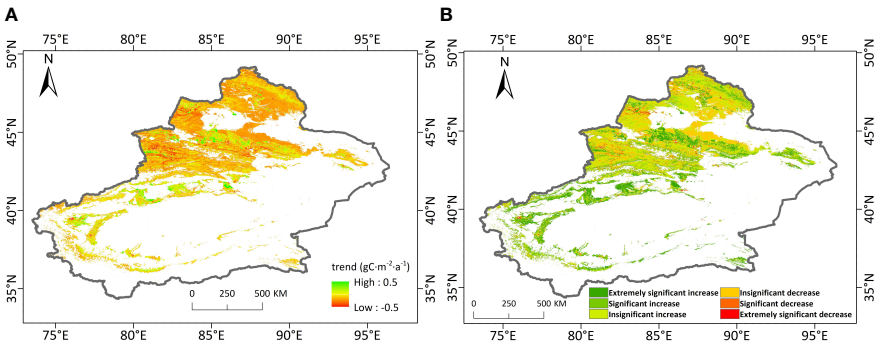


FIGURE 5
Trend of NEP **(A)** and the significance of the trend **(B)** over Xinjiang from 2001 to 2020.

low precipitation and the arid climate. The air temperature also varies widely among different regions, which is mainly influenced by the topography. The Altai Mountain, Tianshan Mountain, Kunlun Mountain and Ali Mountain have high altitude and low temperature, while the Junggar Basin, Ili Basin, Tarim Basin, Tacheng Basin and Turpan Basin have low altitude and high temperature. Compared with the spatial distribution of NEP, precipitation and temperature, the areas with high NEP overlapped with areas with high precipitation and low temperature to a high degree.

Figure 7 shows the temporal variations of annual precipitation and air temperature over Xinjiang from 2001 to 2020. The annual precipitation of vegetation region ranged between 225.99 mm in 2020 and 303.53 mm in 2016, and the 20-year average annual precipitation was 261.32 mm. During this period, the annual precipitation showed an overall fluctuating decreasing trend. The annual temperature of vegetation region ranged from 3.11°C (2003) to 5.10°C (2007), and the 20-year average annual temperature was 4.26°C. Different from precipitation, the annual temperature exhibited an overall fluctuating increasing trend.

The overall partial correlation coefficient between vegetation NEP and precipitation in Xinjiang was 0.145, which showed a positive correlation in general, indicating that precipitation mainly contributed to vegetation NEP in the study area. The partial correlation analysis between NEP and climate factors (precipitation and air temperature) were also carried out pixel by pixel. Figure 8 gives the spatial distribution of the significance of the partial correlation between annual NEP and the two climate factors. 52.57% of the vegetation NEP showed significant positive correlations with precipitation, and 15.66% of the vegetation NEP showed insignificant positive correlations with precipitation. The areas with negative correlation between vegetation NEP and precipitation are relatively small, of which 29.74% show significant negative correlations and 2.04% show insignificant negative correlations, mainly in the oasis zones around the Tarim Basin, the Tianshan Mountains in the north and south of the Ili Valley, the Altay Mountains and the southern side of the Junggar Basin. Precipitation in these regions showed decreasing trends over the past 20 years, while NEP showed increasing trends. This can be attributed to the implementation of ecological water replenishment, afforestation and other ecological protection projects in Xinjiang Autonomous Region, which resulted in abundant water resources and less influence by natural precipitation. Moreover, the decrease in precipitation reduced soil microbial respiration, therefore increased organic carbon consumption and therefore increased vegetation NEP. The partial correlation coefficient between vegetation NEP and annual air temperature in Xinjiang was -0.0368, suggesting that high temperature leads to low NEP. Most of the vegetation NEP showed significant negative correlation with temperature, accounting for 52.77% of the total vegetation area in the study area. The increase in temperature increased the transpiration of vegetation, intensified the evaporation rate of soil water, and reduced the NPP accumulated

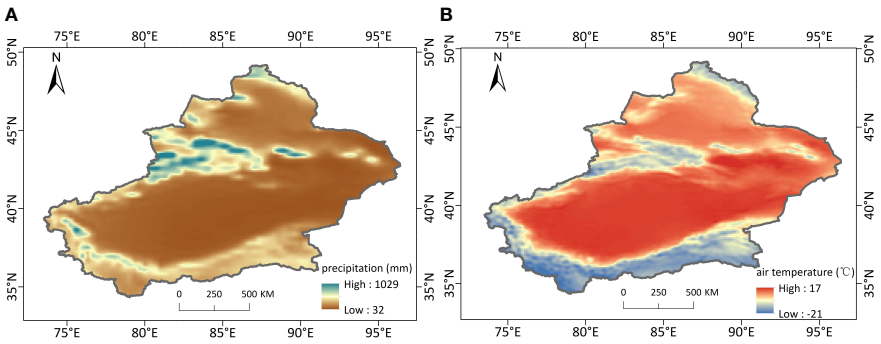


FIGURE 6
Spatial distributions of 20-year mean annual precipitation **(A)** and air temperature **(B)** over Xinjiang from 2001 to 2020.

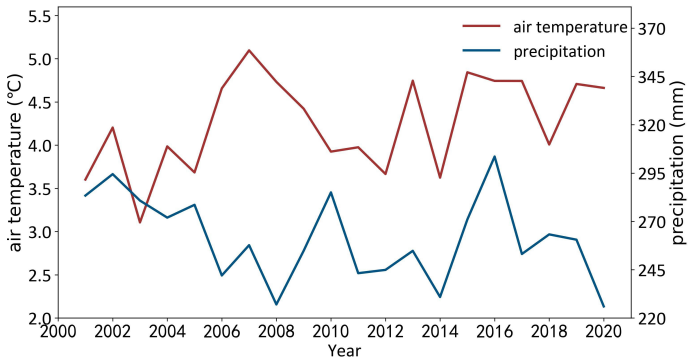


FIGURE 7
Temporal variations of annual precipitation **(A)** and air temperature **(B)** over Xinjiang from 2001 to 2020.

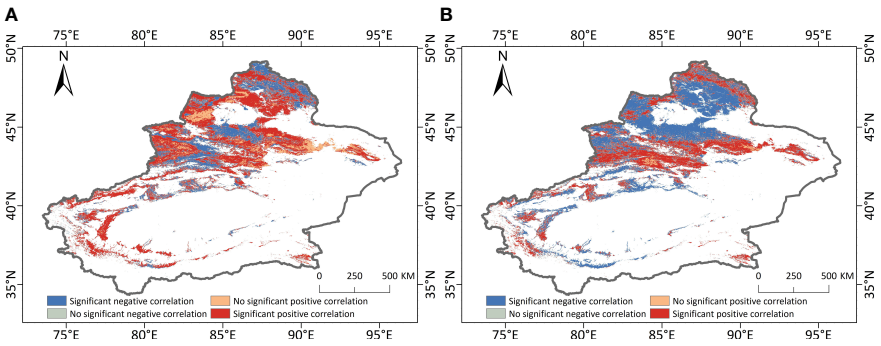


FIGURE 8
Spatial distributions of the significance of the partial correlation between annual NEP and precipitation **(A)** and air temperature **(B)** over Xinjiang from 2001 to 2020.

by vegetation, while the increase in temperature also increased the respiration of soil microorganisms, therefore the vegetation NEP value decreased. 40.84% of the vegetation NEP was significantly and positively correlated with temperature, mainly in the Tianshan Mountains and Altai Mountains. This can be attribute to the fact that higher temperature produces more alpine snow and ice melt water, which increases the amount of water in the region, promotes the growth and development of vegetation and enhances the vegetation's carbon sequestration capacity. Only 2.57% of the vegetation NEP showed insignificant positive correlation and 3.82% showed significant negative correlation with temperature. Generally, both precipitation and temperature had an influence on NEP in Xinjiang, and temperature had less influence than precipitation.

5 Discussions

Though a lot of studies have been carried out on mapping NPP and studying the spatial and temporal variations, the studies on the spatiotemporal variations of NEP are much fewer. Compared with NPP, NEP can better indicate the carbon absorption and emission capacity, which are more meaningful for carbon cycle research. This study is the first research that estimates NEP in Xinjiang, a typical ecologically fragile area in China, providing valuable knowledge for understanding the regional carbon sequestration capacity and for developing strategies to stabilize CO_2 emissions.

Previous studies generally use station-based observations to produce spatial continuous meteorological factors to estimate NEP. However, in Xinjiang there are quite a few meteorological stations and the land and atmospheric characteristics have obvious spatial differences. Spatial interpretation cannot well depict the spatial variations of meteorological factors. Based on this consideration, ERA5-land reanalysis data were employed in this study to derive gridded precipitation and air temperature, which can provide much more spatial details than interpolated meteorological observations in the study area. This study also has a limitation that there is no *in-situ* NPP observation data for validation. We compared our results with that of the studies of adjacent regions, such as Qinghai Province and Inner Mongolia Autonomous Region. The NEP values for typical vegetation types are similar. Moreover, the CASA model has been widely used in many regions and has proven its applicability. Therefore, the estimated NEP results are credible for the spatial and temporal variations.

Despite the relatively small vegetation coverage in Xinjiang, the terrestrial ecosystem still exhibits an overall carbon sink effect from 2001 to 2020. Moreover, the terrestrial carbon sink showed an

obvious increasing trend in the past 20 years. The partial correlation analysis between annual NEP, precipitation, and air temperature indicated that NEP was positively correlated with precipitation and negatively correlated with temperature. Precipitation exerted a stronger influence than temperature, indicating that precipitation is the dominant driving factor that influences the temporal trend of carbon sequestration capacity in this arid ecosystem.

6 Conclusions

By integrating satellite remote sensing data and reanalysis meteorological data, the spatio-temporal variations of NEP in Xinjiang Autonomous Region, China from 2001 to 2020 were studied based on the CASA model and the soil heterotrophic respiration model. The NEP in Xinjiang showed a spatial pattern that was high in the east and low in the west, high in the north and low in the south. The overall annual mean vegetation NEP of 128.54 $gC{\cdot}m^{-2}$. During the past 20 years, the annual mean NEP over Xinjiang increased with the Sen's slope of 1.59 $gC{\cdot}m^{-2}{\cdot}a^{-1}$, indicating that the carbon sink effect of the vegetation ecosystem in Xinjiang was enhanced. The trend of NEP also exhibited a large spatial heterogeneity. The area with increasing NEP trend accounted for 71.46% of the total vegetation area. Annual NEP was positively correlated with precipitation but negatively correlated with temperature. Compared with temperature, NEP was more sensitive to precipitation in this region. This study provides a reference for the sustainable development of the terrestrial ecosystem and the impact of climate change on the carbon cycle in Xinjiang, and also provides a template for NEP investigation in other regions.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

XL carried out the study and wrote the manuscript; YC and YS collected and processed the data, provided suggestions and edited the manuscript; YoX designed the work, helped organize the manuscript structure, and directed the study; YaX drew the figures; YM provided suggestions. All authors contributed to the article and approved the submitted version.

Funding

This research was funded by the Science and Technology Research Plan in Key areas of Xinjiang Production and Construction Corps (2022AB016), the National Natural Science Foundation of China (42271351, 42171101), and the Postgraduate Research and Practice Innovation Program of Jiangsu Province (KYCX22_1175).

Conflict of interest

Author XL was employed by the company The First Ecological and Environment Monitoring Station of Xinjiang Production and Construction Corps.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Baker, H. S., Millar, R. J., Karoly, D. J., Beyerle, U., Guillod, B. P., Mitchell, D., et al. (2018). Higher CO_2 concentrations increase extreme event risk in a 1.5°C world. *Nat. Clim. Change* 8, 604–608. doi: 10.1038/s41558-018-0190-1

Bradford, J. B., Hicke, J. A., and Lauenroth, W. K. (2005). The relative importance of light-use efficiency modifications from environmental conditions and cultivation for estimation of large-scale net primary productivity. *Remote Sens Environ.* 96, 246–255. doi: 10.1016/j.rse.2005.02.013

Chirici, G., Barbati, A., and Maselli, F. (2007). Modelling of Italian forest net primary productivity by the integration of remotely sensed and GIS data. *For. Ecol. Manage.* 246, 285–295. doi: 10.1016/j.foreco.2007.04.033

Fensholt, R., Langanke, T., Rasmussen, K., Reenberg, A., Prince, S., Tucker, C., et al. (2012). Greenness in semi-arid areas across the globe 1981–2007—An earth observing satellite based analysis of trends and drivers. *Remote Sens Environ.* 121, 144–158. doi: 10.1016/j.rse.2012.01.017

Friend, A. D., Lucht, W., Rademacher, T. T., Asner, G. P., Saatchi, S., Townsend, P., et al. (2014). Carbon residence time dominates uncertainty in terrestrial vegetation responses to future climate and atmospheric CO_2. *Proc. Natl. Acad. Sci. U.S.A.* 111, 3280–3285. doi: 10.1073/pnas.1222477110

Goetz, S. J., Prince, S. D., Goward, S. N., Thawley, M. M., and Small, J. (1999). Satellite remote sensing of primary production: An improved production efficiency modeling approach. *Ecol. Model.* 122, 239–255. doi: 10.1016/S0304-3800(99)00140-4

Guo, D., Song, X., Hu, R., Zhu, X., Jiang, Y., Cai, S., et al. (2021). Large-Scale analysis of the spatiotemporal changes of net ecosystem production in Hindu kush Himalayan region. *Remote Sens.* 13, 1180. doi: 10.3390/rs13061180

He, Q., and Xie, Y. (2010). Research on the climatological calculation method of solar radiation in. *China. J. Nat. Res.* 25, 308–319. doi: 10.11849/zrzyxb.2010.02.015

He, J., Zhang, P., Jing, W., and Yan, Y. (2018). Spatial responses of net ecosystem productivity of the yellow river basin under diurnal asymmetric warming. *Sustainability* 10, 3646. doi: 10.3390/su10103646

Hoegh, G. O., and Bruno, J. F. (2010). The impact of climate change on the world's marine ecosystems. *Science* 328, 1523–1528. doi: 10.1126/science.1189930

IPCC (2014). *Climate change 2014: Impacts, adaptation, and vulnerability* (Cambridge, UK: Cambridge University Press (IPCC Secretariat).

Ji, R., Tan, K., Wang, X., Pan, C., and Xin, L. (2021). Spatiotemporal monitoring of a grassland ecosystem and its net primary production using Google earth engine: A case study of inner Mongolia from 2000 to 2020. *Remote Sens.* 13, 4480. doi: 10.3390/rs13214480

Keenan, T. F., Prentice, I. C., Canadell, J., Williams, C. A., Wang, H., Raupach, M., et al. (2016). Recent pause in the growth rate of atmospheric CO_2 due to enhanced terrestrial carbon uptake. *Nat. Commun.* 7, 13428. doi: 10.1038/ncomms13428

Li, Z., Chen, Y., Zhang, Q., and Li, Y. (2020). Spatial patterns of vegetation carbon sinks and sources under water constraint in central Asia. *J. Hydrol.* 590, 125355. doi: 10.1016/j.jhydrol.2020.125355

Luo, X., and Keenan, T. F. (2022). Tropical extreme droughts drive long-term increase in atmospheric CO2 growth rate variability. *Nat. Commun.* 13, 1193. doi: 10.1038/s41467-022-28824-5

Met Office (2022) *Mauna loa carbon dioxide forecast for 2022*. Available at: https://www.metoffice.gov.uk/research/climate/seasonal-to-decadal/long-range/forecasts/co2-forecast.

Mngadi, M., Odindi, J., Mutanga, O., and Sibanda, M. (2022). Estimating aboveground net primary productivity of reforested trees in an urban landscape using biophysical variables and remotely sensed data. *Sci. Total Environ.* 802, 149958. doi: 10.1016/j.scitotenv.2021.149958

Nayak, R. K., Patel, N. R., and Dadhwal, V. K. (2015). Spatio-temporal variability of net ecosystem productivity over India and its relationship to climatic variables. *Environ. Earth Sci.* 74, 1743–1753. doi: 10.1007/s12665-015-4182-4

NOAA (2022) *Trends in atmospheric carbon dioxide*. Available at: https://gml.noaa.gov/ccgg/trends/.

Pachavo, G., and Murwira, A. (2014). Remote sensing net primary productivity (NPP) estimation with the aid of GIS modelled shortwave radiation (SWR) in a southern African savanna. *Int. J. Appl. Earth Obs. Geoinf.* 30, 217–226. doi: 10.1016/j.jag.2014.02.007

Parton, W. J., Scurlock, J. M. O., Ojima, D. S., Gilmanov, T. G., Scholes, R. J., Kinyamario, J. I., Schimel, D. S., et al. (1993). Observations and modeling of biomass and soil organic matter dynamics for the grassland biome worldwide. *Glob. Biogeochem. Cycles* 7, 785–809. doi: 10.1029/93GB02042

Pei, Z., Ouyang, H., Zhou, C., and Xu, X. (2003). CO_2 processes in an alpine grassland ecosystem on the Tibetan plateau. *J. Geogr. Sci.* 4, 47–55. doi: 10.1007/BF02837881

Schimel, D., Stephens, B. B., and Fisher, J. B. (2015). Effect of increasing CO_2 on the terrestrial carbon cycle. *Proc. Natl. Acad. Sci. U.S.A.* 112, 436–441. doi: 10.1073/pnas.1407302112

Veroustraete, F., Patyn, J., and Myneni, R. B. (1994). "Forcing of a simple ecosystem model with fAPAR and climatic data to estimate regional scale photosynthetic assimilation," in *Vegetation, modelling and climate change effects*. Eds. F. Veroustraete, R. Ceulemans, I. Impens and J. Van Rensbergen (The Hague, The Netherlands: SPB Academic Publishing), 151–177.

Xue, Y., and Qin, J. (2022). Partial connection based on channel attention for differentiable neural architecture search. *IEEE Trans. Industr. Inform*, 1–10. doi: 10.1109/TII.2022.3184700

Xue, Y., Wang, K., and J. and Slowik, A. (2021). A self-adaptive mutation neural architecture search algorithm based on blocks. *IEEE Comput. Intell. Mag.* 16, 67–78. doi: 10.1109/MCI.2021.3084435

Yamaji, T., Sakai, T., Endo, T., Baruah, P. J., Akiyama, T., Saigusa, N., et al. (2008). Scaling-up technique for net ecosystem productivity of deciduous broadleaved forests in Japan using MODIS data. *Ecol. Res.* 23, 765–775. doi: 10.1007/s11284-007-0438-0

Zhang, J., Hao, X., Hao, H., Fan, X., and Li, Y. (2021). Climate change decreased net ecosystem productivity in the arid region of central Asia. *Remote Sens.* 13, 4449. doi: 10.3390/rs13214449

Zhang, Z. Q., Liu, H., Zuo, Q. T., Yu, J. T., and Li, Y. (2021). Spatiotemporal change of fractional vegetation cover in the yellow river basin during 2000–2019. *Resour. Sci.* 4, 849–858. doi: 10.3390/rs13214449

Zhao, S., and Luo, T. (1998). Approaches to the regional scale bio-productivity of terrestrial ecosystems. *Resour. Sci.* 20, 23–34.

Zhao, J., Ma, J., and Zhu, Y. (2019). Evaluating impacts of climate change on net ecosystem productivity (NEP) of global different forest types based on an individual tree-based model FORCCHN and remote sensing. *Glob. Planet Change*, 103010. doi: 10.1016/j.gloplacha.2019.103010

Zhu, W., Pan, Y., and Zhang, J. (2007). Estimation of net primary productivity of Chinese terrestrial vegetation based on remote sensing. *J. Plant Ecol.* 31, 413–424. doi: 10.17521/cjpe.2007.0050

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 14 February 2023
DOI 10.3389/fpls.2023.1048016

OPEN ACCESS

EDITED BY
Ferrante Neri,
University of Surrey, United Kingdom

REVIEWED BY
Nisha Pillai,
Mississippi State University, United States
Bin Yang,
Aalborg University, Denmark

*CORRESPONDENCE
Peisen Yuan
peiseny@njau.edu.cn
Zhaoyu Zhai
2021004@njau.edu.cn

SPECIALTY SECTION
This article was submitted to Sustainable and Intelligent Phytoprotection, a section of the journal Frontiers in Plant Science

RECEIVED 19 September 2022
ACCEPTED 13 January 2023
PUBLISHED 14 February 2023

CITATION
Yuan P, Xu S, Zhai Z and Xu H (2023) Research of intelligent reasoning system of *Arabidopsis thaliana* phenotype based on automated multi-task machine learning. *Front. Plant Sci.* 14:1048016. doi: 10.3389/fpls.2023.1048016

Research of intelligent reasoning system of *Arabidopsis thaliana* phenotype based on automated multi-task machine learning

Peisen Yuan*, Shuning Xu, Zhaoyu Zhai* and Huanliang Xu

College of Artificial Intelligence, Nanjing Agricultural University, Nanjing, China

Traditional machine learning in plant phenotyping research requires the assistance of professional data scientists and domain experts to adjust the structure and hyperparameters tuning of neural network models with much human intervention, making the model training and deployment ineffective. In this paper, the automated machine learning method is researched to construct a multi-task learning model for *Arabidopsis thaliana* genotype classification, leaf number, and leaf area regression tasks. The experimental results show that the genotype classification task's accuracy and recall achieved 98.78%, precision reached 98.83%, and classification F_1 value reached 98.79%, as well as the R^2 of leaf number regression task and leaf area regression task reached 0.9925 and 0.9997 respectively. The experimental results demonstrated that the multi-task automated machine learning model can combine the benefits of multi-task learning and automated machine learning, which achieved more bias information from related tasks and improved the overall classification and prediction effect. Additionally, the model can be created automatically and has a high degree of generalization for better phenotype reasoning. In addition, the trained model and system can be deployed on cloud platforms for convenient application.

KEYWORDS

plant phenotype reasoning, multi-task learning, automated machine learning, *Arabidopsis thaliana*, cloud deployment

1 Introduction

Plant phenotypes are recognizable morphological, physiological, and biochemical characteristics and attributes that result in part or entirely from the interaction of genes with the environment (Dobrescu et al., 2020; Cheng et al., 2021; Saric et al., 2022), which is widely used in ecological protection (Carvalho et al., 2021), plant breeding (van Dijk et al., 2021) and so on. Currently, machine learning has rapidly evolved and is now widely applied in science in general and in plant genotyping and phenotyping (Ubbens and Stavness, 2017; van Dijk et al., 2021). Different phenotypic qualities of plants are connected with one another, and this gives us the hints to leverage the benefits of multi-task learning to enhance the

effectiveness of individual learning activity. In order to improve the classification and prediction performance of numerous related tasks, multi-task learning for the same plant enables better reasoning of the link between various phenotypic variables as well as training with less data and annotated information.

The construction of multi-task learning models requires a lot of manual time to tune the hyperparameters of the model for making the model have a high performance (Zhang et al., 2019; Vandenhende et al., 2022).

Therefore, there are limitations in human thinking to consider the model structure and parameters in all possible cases.

At the present time, machine learning has become an essential part of daily applications (van Dijk et al., 2021), however, building well-performing machine learning models still requires the help of data scientists and domain experts. To solve this problem, Automated Machine Learning (AutoML for short) (He et al., 2021) was proposed and researched. AutoML automates the process of constructing network structure, adjusting network structure, adjusting hyperparameters, and model evaluation (Truong et al., 2019; Xue et al., 2021) through its own set of algorithms, which turns the original structure adjustment and parameter tuning into structured and orderly adjustment through the well-designed algorithms, which lowers the threshold of machine learning and shortens the whole modeling process. Using automated machine learning methods enables deep learning techniques to be applied to more fields in a simpler way to build better network models for machine learning tasks with high accuracy. AutoML brings a way for researchers without AI knowledge and the help of machine learning experts to build their AI system (Zöller and Huber, 2021).

Based on the above pros and cons, we propose the AutoML to build a multi-task learning model for *Arabidopsis thaliana* phenotype reasoning, which can combine the advantages of both approaches, and take into account both the correlation between multiple phenotypic variables and the diverse model structures and parameter pairings. For multi-task learning, the use of an automated machine learning method to construct models provides a viable new approach for subsequent research on other multitasks. And for AutoML, applying the knowledge of multitask learning allows for better finding the network suitable for each task by taking into account the correlation between tasks when searching for neural network architectures.

Currently, Zhou et al. (2021) introduced a deep learning-based maize image analysis software that can automatically solve a variety of image-based maize phenotyping tasks, such as internal length, stem diameter, and leaf count, for high-throughput plant phenotyping. Similarly, P. Hüther et al. (Hüther et al., 2020). analyze the phenotype of *Arabidopsis thaliana* using transfer learning by centering our pipeline around the well-established deep-learning model DeepLabV3+ for batch automated plant leaf state analysis, and no automated generation of the model was implemented.

This paper focuses on automated machine learning methods for multi-task learning models. Taking *Arabidopsis thaliana* as an example, three tasks concerning the processing and analysis of plant phenotypic characteristics were finally realized: 1)infer the genotype of *Arabidopsis thaliana*; 2) predict the total number of leaves in *Arabidopsis thaliana*; and 3)predict the leaf area of *Arabidopsis thaliana*. For the above analysis tasks of *Arabidopsis thaliana* dataset, AutoML based multi-task models are researched and constructed for the three tasks mentioned above for training, and the different metrics of each model are compared to produce the best classification and regression results.

The main contribution of this paper is the use of an automated machine learning approach that automatically adjusts hyperparameters as well as model structure as a way to construct a multi-task learning model for *Arabidopsis thaliana* phenotype reasoning tasks. And the experiment results show that it has better classification and regression results compared to previous state-of-the-art models.

The rest of this paper is organized as follows. Section 2 introduces the relevant principles and workflow of multi-task AutoML. The details of the multi-task AutoML model used in this paper are explained in Section 3. Experiment and comparison of the proposed method for *Arabidopsis thaliana* phenotype multi-task reasoning are presented in Section 4. System implementation and deployment are presented in Section 5. Finally, Section 6 draws conclusions and provides an in-depth analysis and an outlook on future work.

2 Related works

Multi-Task Learning (MTL for short) has been proposed with the intention of leveraging the useful information contained in multiple related tasks to help improve the generalization performance of all the tasks (Zhang and Yang, 2022). While the phenotypic traits of plants are correlated to some extent, using multi-task learning, the network can be trained with less data and less labeled information to achieve better classification and prediction results for multiple related tasks.

Among the two basic frameworks for multi-task learning, soft parameter sharing does not make any assumptions about task relevance, but the number of required parameters is large. In contrast, hard parameter sharing is mostly applied to networks with strong task relevance. For the study of *Arabidopsis thaliana* phenotypes, there are strong correlations among phenotypes, and thus the hard parameter sharing framework is mostly used to build relevant models. For example, the first application of multi-task learning to plant phenotypes was proposed by Pound et al. (2017) with the ability to both detect and count the number of wheat ears and to classify the presence of wheat awns, and Dobrescu et al. (2020) present a hard parameter sharing framework of multi-task learning for plant phenotyping to infer leaf count, projected leaf area, and genotype classification.

With the development of multi-task learning, the simple hard parameter sharing model can no longer satisfy the needs of *Arabidopsis thaliana* phenotype multi-task reasoning applications, and people start to add different strategies to the hard parameter sharing framework to obtain higher performance. Jiang et al. (2021) incorporates a migration learning approach in a hard parameter sharing framework to achieve simultaneous recognition of leaf diseases in rice and wheat. Keceli et al. (2022) combined CNN features and transfer features to construct a multi-input multi-task learning model to improve the efficiency of plant type and disease detection.

Automated Machine Learning is the process of automating the end-to-end process of applying machine learning to real-world problems, enabling models to automatically learn appropriate parameters and models without human intervention.

Currently, the open source AutoML frameworks include Auto-sklearn (Feurer et al., 2015), TPOT (Olson et al., 2016), Auto-Keras (Jin et al., 2019), H2O (LeDell and Poirier, 2020), etc. Auto-sklearn (Feurer et al., 2015) and H2O (LeDell and Poirier, 2020) are mainly oriented to traditional machine learning for automatic modeling. TPOT (Olson et al., 2016) mainly applies genetic algorithms for feature and model selection. Auto-Keras (Jin et al., 2019) is a Keras-based AutoML system that can achieve the powerful function of neural architecture search with only a few lines of code and is easy to get started and use.

Nowadays, More and more advanced methods are being applied to AutoML to improve the performance of the models. Wong et al. (2018) proposed Transfer Neural AutoML, which reduces the computational cost of neural AutoML by migration learning. Xue et al. (2019) proposed a migratable AutoML method that uses previously trained models to speed up the search process for new tasks and datasets, accelerating the overall search time for multiple datasets with negligible accuracy loss. Ferreira et al. (2021) conducted a comparative study of hundreds of computational experiments based on a total of three scenarios: general-purpose machine learning, deep learning, and XGBoost, with GML achieving the best prediction results and the GML AutoML tool obtaining the most competitive results, while confirming the potential of the general-purpose AutoML tool to fully automate machine learning algorithm selection and tuning. Zöller et al. (2022) proposed an XAutoML for interpreting arbitrary AutoML optimization processes and ML pipelines constructed by AutoML. And the framework we use is optimized for AutoML mainly in the Neural Architecture Search part.

Neural Architecture Search (NAS) (Elsken et al., 2019) aims at automatically designing well-performing neural network architectures for specific target tasks, which requires huge computational resources. Ying et al. (2019) introduced NAS-Bench-101 to ameliorate these problems. And Dong et al (Dong and Yang, 2020). proposed NAS-Bench-201 with a fixed search space, which provides a unified benchmark for almost all the latest NAS algorithms and is an extension of NAS-Bench-101. To overcome the efficiency challenges of simple weight sharing in NAS, Shen et al. (2022) introduce DASH, a differentiable NAS algorithm that achieves better asymptotic complexity and up to 10 times faster search time in practice. Luo et al. (2020) proposed SemiNAS, a semi-supervised NAS approach that uses a trained accuracy predictor to predict the accuracy of a large number of architectures, reducing computational cost and achieving higher accuracy at the same computational cost with the same accuracy guarantee, e.g., it achieves 94.02% test accuracy on NASBench-101, using the same number of architectures outperformed all baselines. Xue et al. (Xue and Qin, 2022). proposed a partial channel connection based on channel attention for differentiable neural architecture search. Auto-Keras (Jin et al., 2019) uses an efficient neural architecture search method with network morphism, combined with Bayesian optimization, which makes the search space exploration more efficient and has better performance for the current optimal baseline model.

In the field of plant phenotype research, Koh et al. (2021) investigated the application of AutoML in image-based plant phenotyping, comparing the performance of the open source AutoML framework Auto-Keras with migration learning using a convolutional neural network architecture. In the classification task, migration learning with Xception (Chollet, 2017) and DenseNet-201 (Huang et al., 2017) achieved the best classification accuracy of 93.20%, while Auto-Keras achieves 92.40% accuracy. With similar accuracy, Auto-Keras speeds up the model's inference time by a factor of 40 and has great potential for enhancing plant phenotyping capabilities applicable to crop breeding and precision agriculture.

In summary, we proposed an intelligent reasoning system for *Arabidopsis thaliana* phenotype based on automated Multi-task machine learning with Auto-Keras (Jin et al., 2019), which can take the both advantages of AutoML and multi-task learning.

3 Intelligent reasoning of *Arabidopsis thaliana* phenotype based on multi-task automated machine learning

3.1 Problem statement

In AutoML, model generation and evaluation are done by neural network architecture search. As the backbone of deep AutoML, Neural Architecture Search(NAS for short) (Xue and Qin, 2022) can define and optimize the neural network architectures and tune hyperparameters automatically, which enables people with little expertise and knowledge to perform machine learning tasks for obtaining highly accurate, and even discover unproposed network architectures for some specific tasks.

The basic process of NAS is shown in Figure 1. First, a specific structure a is selected from the predefined search space A according to the search strategy, and the specific structure is evaluated by the performance evaluation module, which returns the performance estimate to the search strategy and guides the next structure selection, and so on until a model a∗ satisfying the predefined performance requirements is produced as the output of the problem.

The optimization strategy used by NAS to obtain the optimal network configuration is shown in Equation 1.

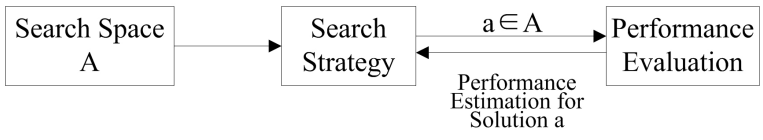


FIGURE 1
Basic processing procedure of neural architecture search for auto machine learning.

$$a* = \underset{a \in A}{\arg\max}\, O\left(\Lambda(a, d_t), d_v\right) = \underset{a \in A}{\arg\max}\, f(a) \quad (1)$$

where O represents the target function for training network structure parameters, d_t represents training data, and d_v represents validation data.

Among them, the calculation formula of $\Lambda(a, d)$ is shown in Equation 2.

$$\Lambda(a, d) = \underset{m_{a,\theta} \in M_a}{\arg\min}\, L(m_{a,\theta}, d_t) + R(\theta) \quad (2)$$

where, M is the model space, L represents the loss parameter used for the training network, θ represents the network parameter, and R represents the loss function part used for regularization.

3.2 Workflow of *Arabidopsis thaliana* phenotype analysis based on automated multitask machine learning

AutoML can be divided into two types: traditional AutoML and deep AutoML. Traditional AutoML combines the three steps of feature engineering, model selection, and optimization algorithm selection into a single pipeline for automatic learning. Deep AutoML uses neural architecture search (Elsken et al., 2019) to optimize the three problems and thus learn the optimal network structure automatically. The deep AutoML for neural network modeling in deep learning compose of four processes: data preparation, feature engineering, model generation and evaluation, and the workflow of *Arabidopsis thaliana* phenotype with auto multi-task reasoning is shown in Figure 2.

3.2.1 Data preparation

The preparation of *Arabidopsis thaliana* phenotype data mainly includes data collection and data cleaning.

1. Data collection consists mainly of data collection, data tagging, and data improvement (Roh et al., 2019), which tunes the completed raw data into the storage systems.

Data collection includes the following two steps. The main purpose is to convert the Ara2013-Canon dataset (Minervini et al., 2016) to Visual Object Class (VOC) format to obtain direct information about leaf area.

In the first step, the RGB segmentation annotation image data set in CVPPP is converted into JSON files in Common Objects in Context(COCO) format. A separate black and white image of each leaf is generated from the color annotated leaf images provided in the dataset, and Figure 3 gives an example of the completed conversion of a particular image, and a JavaScript Object Notation (JSON) file containing all the image information is generated.

In the second step, a tag file in XML format unique to each image is generated from the JSON file. In this tag file, the location of the image, the genotype, and the bounding box where each leaf is located are included, so all information such as the genotype, number of leaves, and leaf area of the image can be directly obtained through this tag file.

2. Data cleaning is mainly to remove irrelevant data and duplicate data from the original data set, smooth out noisy data, filter out data irrelevant to the mining topic, and deal with missing values, abnormal values, etc (Brownlee, 2020). Data cleaning in this paper focuses on comparing the number of leaves in each image obtained according to the data collection step with the number of leaves given in the original dataset and eliminating the parts with different numbers of leaves.

3.2.2 Feature engineering

Feature engineering extracts features from the processed data in the data preparation phase and transforms them into formats that are suitable for the machine learning model (Zheng and Casari, 2018). It mainly includes three parts: feature selection, feature extraction, and feature construction. Feature selection reduces feature redundancy by selecting important features, feature extraction reduces the dimensionality of features by applying a specific mapping function, and feature construction extends the original feature space. In addition, it also includes feature improvement, feature dimensionality reduction, and other contents. Feature engineering maximizes the extraction of features for use in subsequent NAS processes.

3.3 Model architecture search

Neural architecture search is a sophisticated and systematic work, which is mainly based on the key components of NAS: search space, search strategy, and evaluation strategy (Ren et al., 2021).

Bayesian optimization(BO) (Shahriari et al., 2015) is an effective way for hyperparameter optimization, and has recently emerged as a

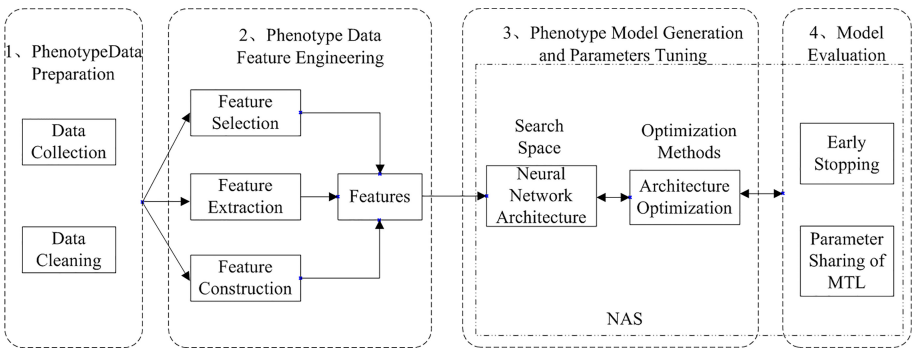


FIGURE 2
Workflow of *Arabidopsis thaliana* phenotype analysis based on automated multi-task machine learning (He et al., 2021).

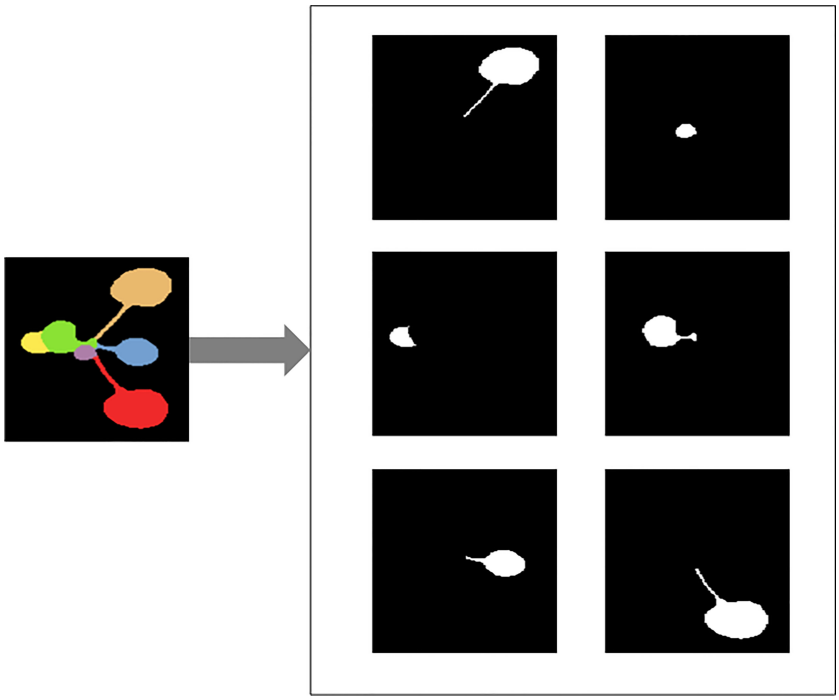


FIGURE 3
Illustration of converting color image of *Arabidopsis thaliana* to black and white images for each leaf.

very promising strategy for NAS. Bayesian optimization puts the optimization issue into a probabilistic framework by representing the agent function as a probability distribution, and then updating this distribution with new information. The acquisition function is used to evaluate the probability of obtaining a better result at a particular point in the exploitation space based on a known prior. The key to this is the balance between exploration and exploitation.

Auto-Keras (Jin et al., 2019) is guided by a Bayesian optimization algorithm to explore the search space by changing the neural structure. The range of fluctuations of the true target function values (i.e., mean and variance) is first estimated based on the function values of the already searched points, usually implemented with Gaussian process regression. Afterward, the acquisition function can be constructed from the mean and variance, i.e., an estimate of the probability that each point is the extreme point of the function, reflecting the degree to which each point is worth searching, and the extreme point of this function is the next search point. Finally, the newly sampled data is added to the set of observations and then recursive execution is performed until convergence or exhaustion of budgetary resources.

For the search algorithm, Auto-Keras uses A^* algorithm (Yao et al., 2010) for searching and simulated annealing, inspired by various heuristic search algorithms that explore tree-like search spaces and optimization methods that explore and exploit tradeoffs.

Whenever NAS generates a new neural network, it is first evaluated for performance. If the network is trained until convergence and then its performance is evaluated, it will consume a lot of time and computational resources. So the early stopping, low fidelity, surrogate, and parameter sharing skills are selected to speed up the evaluation. Auto-Keras uses network morphisms for the purpose of network weight parameter sharing. The sub-models share weights with each other, so there is no need to re-train the sub-models each time. It also uses the early stop method to stop the computation of configuration models that are expected to perform poorly on the validation set.

3.4 Execution process

In this paper, we construct amdeep AutoML model for *Arabidopsis thaliana* phenotype reasoning based on Auto-Keras (Jin et al., 2019). The execution process of automatic parameter tuning and network structure selection for generating the optimal model is shown in Figure 4.

There are 5 steps for optimizing multi-task phenotype reasoning:

(1) Put the network module into the generator as a seed to initialize the model. Three models including CNN, ResNet, and DenseNet can be selected. The number of initial network models constructed during initialization can be set by itself when calling the API;

(2) After the initialization, all the generated models will be put into the training queue, and the models from the queue head are taken out in turn for training;

(3) When training, the model is evaluated and put into the search queue, because the best model will eventually be trained again, so full training is not required at this time, and thus the neural network architecture search can be performed while training. The performance of the model is used as feedback to the Best Model Searcher to update the Gaussian Process;

(4) The model is removed from the search queue, the Bayesian Optimizer in the Searcher would generate a new architecture

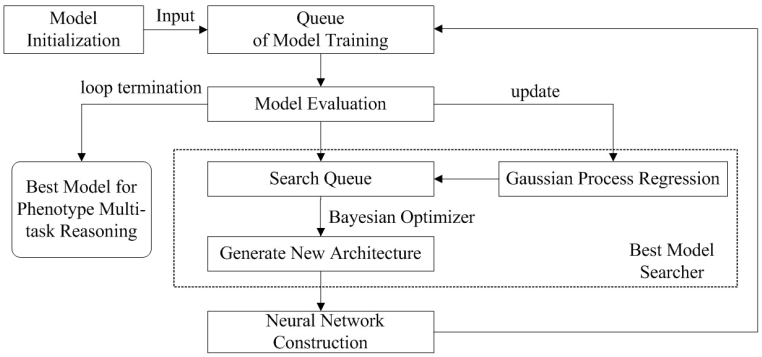


FIGURE 4
The parameters tuning and network structure selection process for *Arabidopsis thaliana* multi- task phenotype reasoning based on Auto-Keras.

and the annealing algorithm is used to determine whether to perform network morphism. The following types of morphism are provided in Auto-Keras: depth, width, and connection between layers. Morphism is random, e.g. the choice of which morph is random, or when choosing to increase the width, the choice of which layer to widen is also random. If a new network architecture is generated by network morphism and that network architecture is not in the existing model, a Gaussian Process Regression is used to predict the better network structure, and the best-performing network is recorded and added to the model search queue and training queue;

(5) The search and training process in steps 3 and 4 is repeated continuously according to the predetermined number of network models to obtain the model with the best results.

After setting the number of training models, the number of iteration rounds, and the system resources available for training by Auto-Keras, the program will automatically adjust the model structure and each parameter according to the process shown in Figure 4, and using the visualization component, the final model structure after training can be obtained.

4 Experiments and analysis

4.1 Experimental environment

The experimental platform is Windows 10 with Intel(R) Xeon(R) Gold 5218 CPU, 32G RAM, GeForce RTX 2080 Ti GPU, and 11G video memory. Models are implemented with Tensorflow 2.0.0+, autokeras 1.0. 18, Python 3.7+, and CUDA 10.0+.

4.2 Dataset details

This experiment uses the Ara2013-Canon dataset (Minervini et al., 2016), a publicly available dataset obtained from the CVPPP leaf segmentation and counting challenge, for training. First, color segmentation of annotated images in CVPPP is used to generate JSON files in COCO format from the original data set. Then, the COCO format data set is converted to a VOC format data set, and the genotype, leaf count, and leaf area information of *Arabidopsis thaliana* are obtained directly from the XML label file.

There are five *Arabidopsis thaliana* genotypes in this dataset: col-0, ein2, ctr, adh1, and pgm. Each *Arabidopsis thaliana* image has label information of genotype, leaf number, leaf position box images, as shown in Table 1, with a total of 165 *Arabidopsis thaliana* phenotype images.

4.3 Evaluation metric

4.3.1 Classification evaluation metrics

In classification tasks, Accuracy A is a frequently used evaluation metric, and it measures the proportion of correctly predicted samples to all samples. The formula is displayed in Equation 3. The prediction effect of the model is better in terms of accuracy the closer its value is near 1.

$$A = \frac{TP + TN}{TP + TN + FP + FN} \tag{3}$$

TABLE 1 Sample image of the Ara2013-Canon dataset.

Genotype	Col-0	ein2	pgm	ctr	adh 1
Original Image					
Leaf Segmentation Image					

where *TP* represents the positive sample predicted by the model as a positive class, *TN* represents the negative sample predicted by the model as a negative class, *FP* represents the negative sample predicted by the model as a positive class, and *FN* represents the positive sample predicted by the model as a negative class.

4.3.2 Regression evaluation metrics

The reliability of the change in the dependent variable in the regression task is indicated by the coefficient of determination, R^2, which is defined as Equation 4. R^2 is a numerical feature used to illustrate the link between a random variable and many other random variables. The coefficient of determination might have a maximum value of 1. The regression line fits the predicted value better and becomes closer to the true value as the value gets closer to 1.

$$R^2 = \frac{\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \quad (4)$$

The extreme errors caused by the squaring amplify the mean squared error(MSE), which is determined as the mean of the squared difference between the anticipated and actual observed values. Predicted values that deviate more from the genuine value are penalized more harshly than those that vary less. The prediction effect is more closely related to the true value the lower the MSE value, which is defined in Equation 5.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \quad (5)$$

In the formula given above, n stands for the quantity of samples, y_i for the sample's true value, $\hat{y}_i$ for its predicted value, and $\bar{y}$ for the average of the true values of all the samples.

4.4 Experimental results

4.4.1 Auto network generation

A multi-task model for automated machine learning was built using Auto-Keras (Jin et al., 2019), setting the maximum number of trials to 10 and the number of iteration rounds to 300, and training the model with an input image of 28 × 28 × 3.

Auto-Keras applies the Early Stop model evaluation criteria and does not fully train all the searched models. Thus only the best-performing models in the evaluation process will be fully trained. After training 10 models using the neural network architecture to search and automatically tune the models as well as the parameters, the model with the best predictions from the 10 trials was selected for the 11th full training. The final structure of the automated machine learning model generated by Auto-Keras' model visualization tool for accomplishing multi-task learning is shown in Figure 5.

As can be seen in Figure 5, the multi-task learning model generated with the automated machine learning approach is a hard parameter sharing model, and the optimal network structure and parameters are automatically selected based on the neural network architecture search algorithm with flexible addition of network layers such as random deactivation. The total number of parameters that have been trained for the model is 118,247, and the specific parameters are shown in Table 2.

4.4.2 Training loss

The last complete model training process of the loss value curve is shown in Figure 6, the loss decreases quickly with epoch, and when the epoch is lower than 15. Figure 6 shows that our model performs well during training without oscillation and can eventually reach convergence.

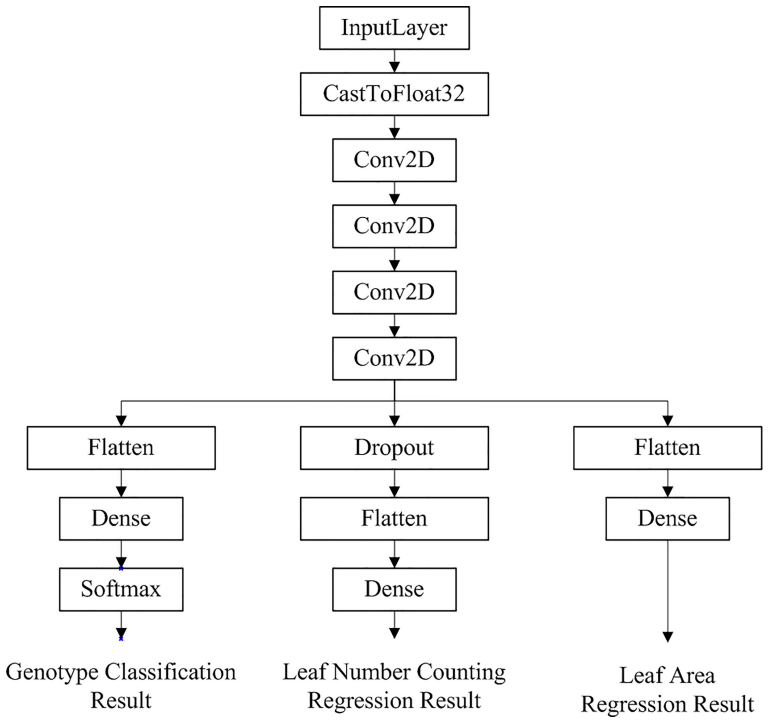


FIGURE 5
Network structure of the optimal multi-task learning model generated by Auto-Keras for *Arabidopsis thaliana* phenotype reasoning tasks.

TABLE 2 Network parameters of the trained and selected model shown in Figure 5 generated by AutoML for *Arabidopsis thaliana* phenotype reasoning.

No.	Layer	Output Shape	Parameters	Connected to
1	input_1	(None,28,28,3)	0	–
2	cast_to_float32	(None,28,28,3)	0	input_1
3	conv2d	(None,26,26,32)	896	cast_to_float32
4	conv2d_1	(None,24,24,32)	9248	con2d
5	conv2d_2	(None,22,22,32)	9248	con2d_1
6	conv2d_3	(None,20,20,32)	9248	con2d_2
7	flatten	(None,12800)	0	con2d_3
8	dropout	(None,20,20,32)	0	con2d_3
9	dense	(None,5)	64005	flatten
10	flatten_1	(None,12800)	0	dropout
11	flatten_2	(None,12800)	0	cov2d_3
12	classification_head_1	(None,5)	0	dense
13	regression_head_1	(None,1)	12801	flatten_1
14	regression_head_2	(None,1)	12801	flatten_2

4.4.3 Results reasoning and comparison

For the genotype classification task, the confusion matrix of the final model obtained by the aforementioned automated machine learning approach is shown in Figure 7. As shown in Figure 7, only two *Arabidopsis thaliana* samples belonging to Col0 type were mistakenly classified as other genotypes, which indicates the great accuracy of our model in the task of genotype classification.

Results comparison with Dobrescu et al. (2020) of the in terms of classification accuracy metrics is shown in Figure 8. From Figure 8, it can be seen that the model trained using the auto multi-task outperforms the model of Dobrescu et al. (2020) in the genotype classification task, with an improvement in classification accuracy of 7.68%.

Figure 9 compares the two models according to R^2 on the task of leaf number regression. It can be seen that the multi-task learning model built by the AutoML not only makes more accurate prediction of the leaf number, but also improves its R^2 value by 4.25% over the previous model.

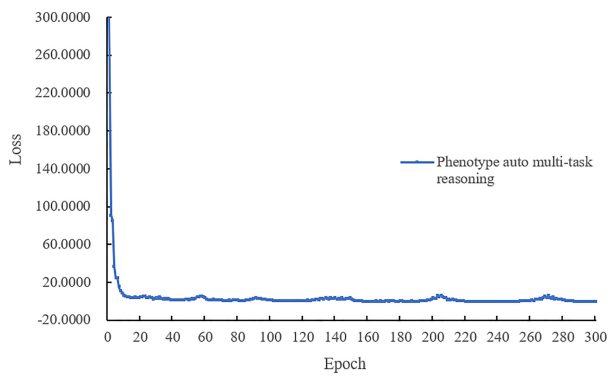


FIGURE 6
Total loss curve of the optimal model during training.

In the leaf area regression task, the comparison of the two models in terms of MSE value is presented in Figure 10. The multi-task learning model developed using AutoML had the smallest MSE error value for leaf area prediction and also reduced by 1.02% compared to the previous model.

Combining the experimental results of the two tasks, it can be seen that automated machine learning can automatically adjust the structure and hyperparameters of the model to obtain a higher model classification accuracy and a lower prediction error without a lot of human intervention. It not only makes up for the shortcomings of manually constructed models, but also improves classification accuracy and prediction accuracy compared with the previous model.

Although it takes longer to train the model than other methods, the stochastic nature of model construction makes it difficult for the best model obtained to be reproduced by others. Automated machine learning facilitates the model building process and can be used by a wide range of people, making it easy for almost anyone to build a model suitable for their task, and allowing experts and scholars to focus their research on more important goals rather than spending a lot of time in tweaking the model.

4.5 Complexity discussion

For the model training of the *Arabidopsis thaliana* phenotype reasoning system, the time complexity depends primarily on NAS as $O(nt + t_{best})$, where n is the number of network architectures to be searched, and t is the average training time for all networks. Auto-Keras reduces t by generating a new network structure on top of the original network through network morphism, which allows the new network to perform better with fewer iterations. And t_{best} is the time for one complete training of the final selected optimal model. When the model training is completed, our system simply feeds the image to be analyzed into the already trained model to obtain the information

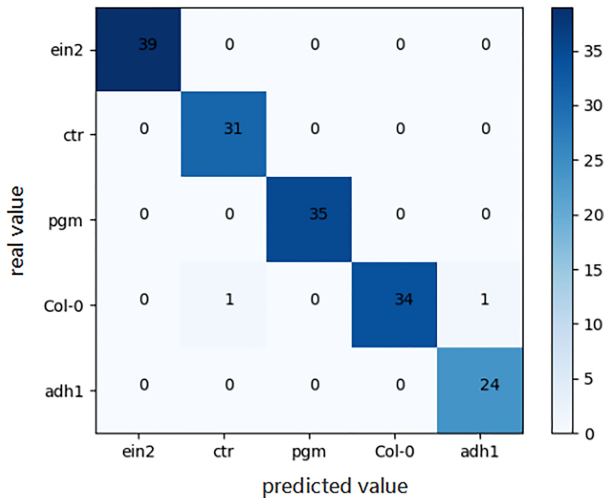


FIGURE 7
Confusion matrix of the optimal model on genotype classification task.

on genotype, leaf number, and leaf area with time complexity of $O(1)$. Thus, the time complexity of the system depends on the efficiency of the NAS and the size of the search space.

Therefore, we can draw the conclusion that the strength of using AutoML to construct multi-task learning for *Arabidopsis thaliana* phenotype reasoning not only considering the correlation between tasks, but also achieves a joint improvement of multiple objectives of tasks through parameter sharing and joint training. Furthermore, it also takes advantage of the points of automatic machine learning to select the best models and adjust hyperparameters tuning, and finally obtain better performance.

5 System implementation and deployment

5.1 System workflow

The system includes two components of the user and the server, such as AWS of Amazon, or Alibaba Cloud, which can be deployed on the cloud. The server deploys the trained AutoML multi-task model

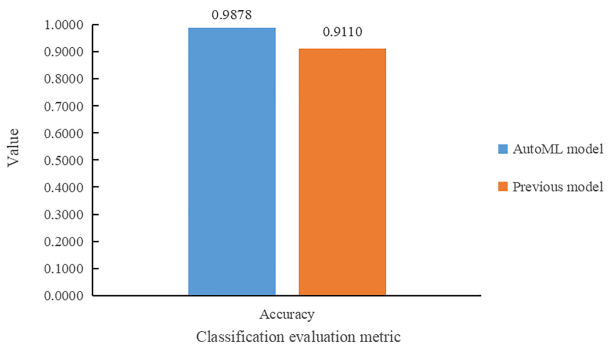


FIGURE 8
Comparative histogram of the two models in terms of classification accuracy.

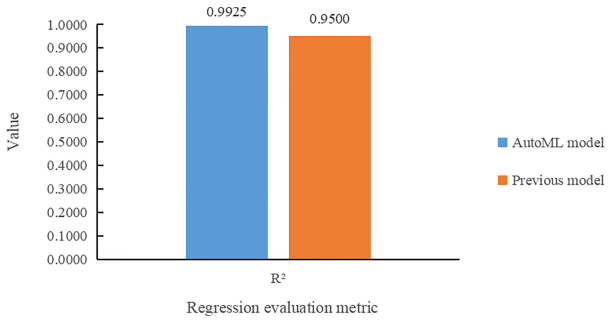


FIGURE 9
Comparative histogram of R^2 for the two models in the leaf number regression task.

and leverages GPU or CPU for inference while the user primarily handles the action of picking the recognized images for the user. Figure 11 depicts the unique workflow when using the system to analyze plant photos.

As shown in Figure 11, There are 5 steps in all: 1) A recognition request is submitted after the user chooses the image to be recognized using the system at their end; 2) The front-end system sends the back-end server the data it has been asked for; 3) Based on the requested input, the server reads in photos and feeds them into a model that has been trained using the AutoML approach; 4) Obtain the expected data following the conclusion of the model processing and deliver the analysis findings; and 5) The user interface shows the image processing outcomes so that users may see the data graphically.

5.2 Main functions

Based on its server's URL address, the server will generate a hyperlink address for users to access. The home page of the online *Arabidopsis thaliana* phenotypic reasoning system based on AutoML can be accessed by this address, as shown in Figure 12.

After choosing the image that needs to be processed and analyzed, click "click here to upload your file" in the main interface. The image's file name will then be presented in the main interface, as shown in Figure 13.

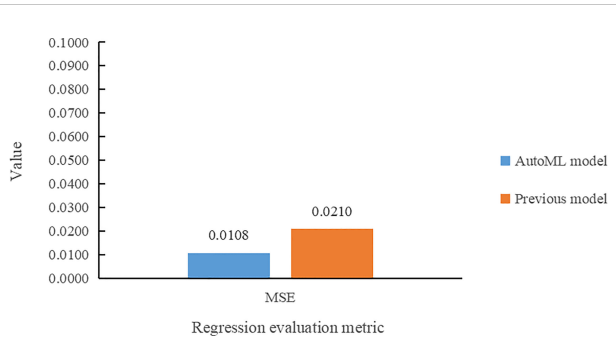


FIGURE 10
Comparative histogram of MSE for the two models in the leaf area regression task.

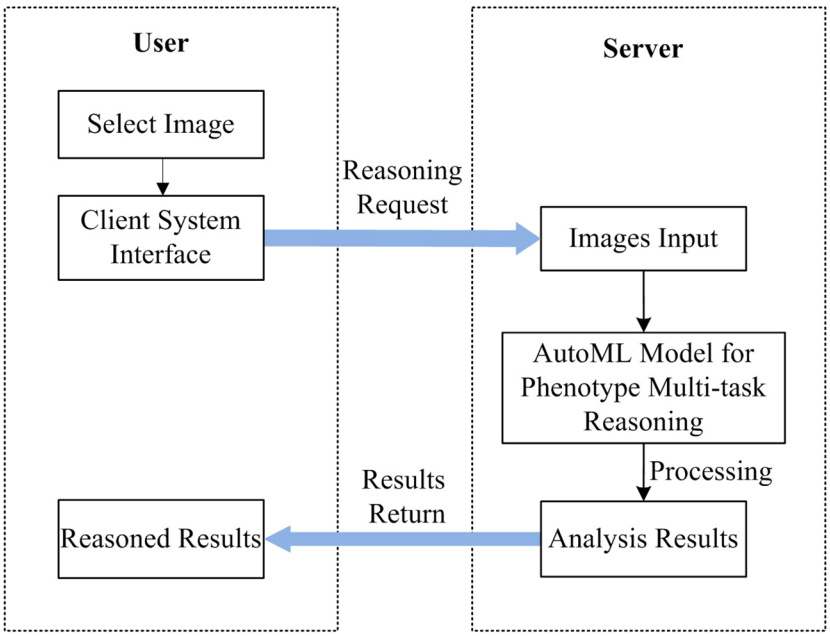


FIGURE 11
Workflow of *Arabidopsis thaliana* phenotype reasoning system based on AutoML.

The user will create the request data and send it to the back-end server by clicking the "Identify" button. To perform inference and retrieve the *Arabidopsis thaliana* results of genotype classification, leaf number regression, and leaf area regression, the server will read in the images and input them into the trained model based on the request *Arabidopsis thaliana* data. The outcomes of the three jobs will be returned to the user client by the model on the server, and these will be illustrated on the user client as illustrated in Figure 14.

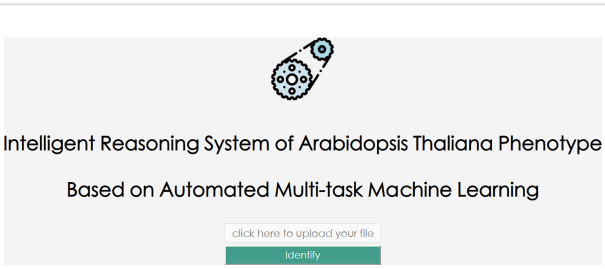


FIGURE 12
User's client of *Arabidopsis thaliana* phenotype reasoning system.

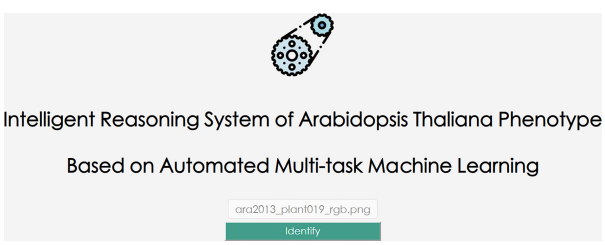


FIGURE 13
Upload *Arabidopsis thaliana* phenotype image for multi-task reasoning.

6 Conclusions and future work

In this paper, we propose AutoML based multi-task intelligent reasoning system for the *Arabidopsis thaliana* phenotype. Our method can select the best model and perform parameter tuning automatically for multi-task learning for *Arabidopsis thaliana* phenotype analysis. The optimal genotypic classification, leaf number, and leaf area prediction results of the present *Arabidopsis thaliana* data set were obtained by training the multi-task learning model with AutoML. The conclusions are summarized as the following.

(1) The multi-task learning model trained by AutoML of Auto-Keras achieved 98.78% accuracy in *Arabidopsis thaliana* genotype classification task, and 7.68% higher than Dobrescu's model. In the leaf counting regression task, the value of R^2 is 0.9925, and 4.25% higher than the previous model. In leaf area regression task, the MSE value is 0.0108,

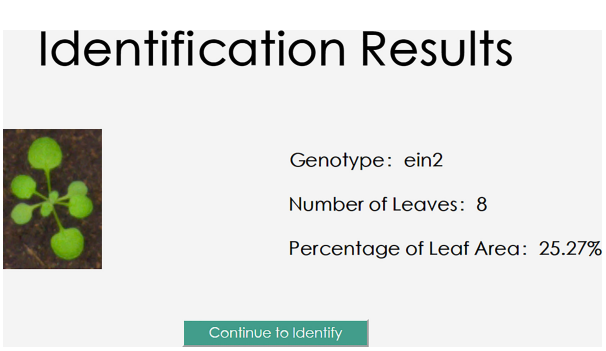


FIGURE 14
Illustration of *Arabidopsis thaliana* phenotype multi-task reasoning results.

which is 1.02% lower than Dobrescu's work (Dobrescu et al., 2020).

(2) Our method can train and adjust model structure and parameter tuning automatically for plant phenotype multi-task reasoning, and improve the classification and regression ability of models automatically without human intervention.

The dataset used in this paper is relatively small and can be expanded in subsequent studies. In future research, various different AutoML frameworks can be used to build the model and compare which method can obtain better overall performance with this dataset. More phenotypic classification or regression tasks can be added to the multitask learning model, and a system dedicated to analyzing plant phenotypes can be built, which can be extended to other plant phenotypic studies.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

PY: Conceptualization, Methodology, Reviewing, Supervision. SX: Data analysis, Writing Original draft,Software. ZZ: Reviewing and Editing. HX: Reviewing, Supervision. All authors contributed to the article and approved the submitted version

Funding

This work is supported by the Jiangsu Agriculture Science and Technology Innovation Fund (JASTIF) (SCX(21)3059), Shanghai Big Data Management System Engineering Research Centre Open Fund (HYSY21022).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Brownlee, J. (2020). "Data preparation for machine learning: Data cleaning, feature selection, and data transforms in Python," (Shenzhen, China: Machine Learning Mastery IEEE).

Carvalho, C. S., Forester, B. R., Mitre, S. K., Alves, R., and Jaff, R. (2021). Combining genotype, phenotype, and environmental data to delineate site–adjusted provenance strategies for ecological restoration. *Mol. Ecol. Resour.* 21, 44–58. doi: 10.1111/1755-0998.13191

Cheng, C.–Y., Li, Y., Varala, K., Bubert, J., Huang, J., Kim, G. J., et al. (2021). Evolutionarily informed machine learning enhances the power of predictive gene–to–phenotype relationships. *Nat. Commun.* 12, 1–15. doi: 10.1038/s41467-021-25893-w

Chollet, F. (2017). "Xception: Deep learning with depthwise separable convolutions," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).* 1251–1258.

Dobrescu, A., Giuffrida, M. V., and Tsaftaris, S. A. (2020). Doing more with less: A multitask deep learning approach in plant phenotyping. *Front. Plant Sci.* 11, 141. doi: 10.3389/fpls.2020.00141

Dong, X., and Yang, Y. (2020). Nas–bench–201: Extending the scope of reproducible neural architecture search. *arXiv*. doi: 10.48550/arXiv.2001.00326

Elsken, T., Metzen, J. H., and Hutter, F. (2019). Neural architecture search: A survey. *J. Mach. Learn. Res.* 20, 1997–2017. doi: 10.5555/3322706.3361996

Ferreira, L., Pilastri, A., Martins, C. M., Pires, P. M., and Cortez, P. (2021). "A comparison of automl tools for machine learning, deep learning and xgboost," in *A Comparison of AutoML Tools for Machine Learning, Deep Learning and XGBoost in 2021 International Joint Conference on Neural Networks (IJCNN)* (Shenzhen, China: IEEE), 1–8.

Feurer, M., Klein, A., Eggensperger, K., Springenberg, J., Blum, M., and Hutter, F. (2015). Efficient and Robust Automated Machine Learning. *Adv. Neural Inf. Process. Syst.* 28, 2755-2763.

He, X., Zhao, K., and Chu, X. (2021). Automl: A survey of the state–of–the–art. *Knowledge– Based Syst.* 212, 106622. doi: 10.1016/j.knosys.2020.106622

Hüther, P., Schandry, N., Jandrasits, K., Bezrukov, I., and Becker, C. (2020). Aradeepopsis, an automated workflow for top–view plant phenomics using semantic segmentation of leaf states. *Plant Cell* 32, 3674–3688. doi: 10.1105/tpc.20.00318

Huang, G., Liu, Z., van der Maaten, L., and Weinberger, K. Q. (2017). "Densely connected convolutional networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).* 4700–4708.

Jiang, Z., Dong, Z., Jiang, W., and Yang, Y. (2021). Recognition of rice leaf diseases and wheat leaf diseases based on multi–task deep transfer learning. *Comput. Electron. Agric.* 186, 106184. doi: 10.1016/j.compag.2021.106184

Jin, H., Song, Q., and Hu, X. (2019). "Auto-Keras: An Efficient Neural Architecture Search System," in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (Anchorage, AK, USA: Association for Computing Machinery). 1946–1956.

Keceli, A. S., Kaya, A., Catal, C., and Tekinerdogan, B. (2022). Deep learning–based multi–task prediction system for plant disease and species detection. *Ecol. Inf.* 69, 101679. doi: 10.1016/j.ecoinf.2022.101679

Koh, J. C., Spangenberg, G., and Kant, S. (2021). Automated machine learning for high–throughput image–based plant phenotyping. *Remote Sens.* 13, 858. doi: 10.3390/rs13050858

LeDell, E., and Poirier, S. (2020). H2o automl: Scalable automatic machine learning. *Proc. AutoML Workshop at ICML* 2020c, 1–16.

Luo, R., Tan, X., Wang, R., Qin, T., Chen, E., and Liu, T.–Y. (2020). Semi–supervised neural architecture search. *Adv. Neural Inf. Process. Syst.* 33, 10547–10557.

Minervini, M., Fischbach, A., Scharr, H., and Tsaftaris, S. A. (2016). Finely–grained annotated datasets for image–based plant phenotyping. *Pattern recognition Lett.* 81, 80–89. doi: 10.1016/j.patrec.2015.10.013

Olson, R. S., Bartley, N., Urbanowicz, R. J., and Moore, J. H. (2016). "Evaluation of a tree– based pipeline optimization tool for automating data science," in *Proceedings of the genetic and evolutionary computation conference 2016* (Denver, Colorado, USA: Association for Computing Machinery). 485–492.

Pound, M. P., Atkinson, J. A., Wells, D. M., Pridmore, T. P., and French, A. P. (2017). "Deep learning for multi–task plant phenotyping," in *Proceedings of the IEEE International Conference on Computer Vision (ICCV) Workshops*. 2055–2063.

Ren, P., Xiao, Y., Chang, X., Huang, P.–Y., Li, Z., Chen, X., et al. (2021). A compre–hensive survey of neural architecture search: Challenges and solutions. *ACM Computing Surveys (CSUR)* 54, 1–34. doi: 10.1145/3447582

Roh, Y., Heo, G., and Whang, S. E. (2019). A survey on data collection for machine learning: A big data–ai integration perspective. *IEEE Trans. Knowledge Data Eng.* 33, 1328–1347. doi: 10.1109/TKDE.2019.2946162

Saric, R., Nguyen, V. D., Burge, T., Berkowitz, O., Trtilek, M., Whelan, J., et al. (2022). Applications of hyperspectral imaging in plant phenotyping. *Trends Plant Sci.* 27, 301-315. doi: 10.1016/j.tplants.2021.12.003

Shahriari, B., Swersky, K., Wang, Z., Adams, R. P., and De Freitas, N. (2015). Taking the human out of the loop: A review of bayesian optimization. *Proc. IEEE* 104, 148–175. doi: 10.1109/JPROC.2015.2494218

Shen, J., Khodak, M., and Talwalkar, A. (2022). Efficient architecture search for diverse tasks. *arXiv*. doi: 10.48550/arXiv.2204.07554

Truong, A., Walters, A., Goodsitt, J., Hines, K., Bruss, C. B., and Farivar, R. (2019). "Towards automated machine learning: Evaluation and comparison of automl approaches and tools," in *2019 IEEE 31st International Conference on Tools with Artificial Intelligence (ICTAI)* (Portland, OR: IEEE), 1471–1479.

Ubbens, J. R., and Stavness, I. (2017). Deep plant phenomics: A deep learning platform for complex plant phenotyping tasks. *Front. Plant Sci.* 8, 1190. doi: 10.3389/fpls.2017.01190

Vandenhende, S., Georgoulis, S., Van Gansbeke, W., Proesmans, M., Dai, D., and Van Gool, L. (2022). Multi–task learning for dense prediction tasks: A survey. *IEEE Trans. Pattern Anal. Mach. Intell.* 44, 3614–3633. doi: 10.1109/TPAMI.2021.3054719

van Dijk, A. D. J., Kootstra, G., Kruijer, W., and de Ridder, D. (2021). Machine learning in plant science and plant breeding. *Iscience* 24, 101890. doi: 10.1016/j.isci.2020.101890

Wong, C., Houlsby, N., Lu, Y., and Gesmundo, A. (2018). Transfer learning with neural automl. *Adv. Neural Inf. Process. Syst.* 31, 8366-8375. doi: 10.48550/arXiv.1903.08362

Xue, Y., and Qin, J. (2022). Partial connection based on channel attention for differentiable neural architecture search. *IEEE Trans. Ind. Inf.*, 1–10. doi: 10.1109/TII.2022.3184700

Xue, Y., Wang, Y., Liang, J., and Slowik, A. (2021). A self–adaptive mutation neural architecture search algorithm based on blocks. *IEEE Comput. Intell. Maga– zine* 16, 67–78. doi: 10.1109/MCI.2021.3084435

Xue, C., Yan, J., Yan, R., Chu, S. M., Hu, Y., and Lin, Y. (2019). "Transferable automl by model sharing over grouped datasets," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 9002–9011.

Yao, J., Lin, C., Xie, X., Wang, A. J., and Hung, C.–C. (2010). "Path planning for virtual human motion using improved a* star algorithm," in *Path Planning for Virtual Human Motion Using Improved A* Star Algorithm in 2010 Seventh International Conference on Information Technology: New Generations* (Las Vegas, NV, USA: IEEE), 1154–1158.

Ying, C., Klein, A., Christiansen, E., Real, E., Murphy, K., and Hutter, F. (2019). "Nas–bench–101: Towards reproducible neural architecture search," in *NAS-Bench-101: Towards Reproducible Neural Architecture Search in Proceedings of the 36th International Conference on Machine Learning* (PMLR). 7105–7114 .

Zhang, Y., and Yang, Q. (2022). A survey on multi–task learning. *IEEE Trans. Knowledge Data Eng.* 34, 5586–5609. doi: 10.1109/TKDE.2021.3070203

Zhang, J., Zhang, J., Ghosh, S., Li, D., Zhu, J., Zhang, H., et al. (2019). Regularize, ex–pand and compress: Multi–task based lifelong learning *via* nonexpansive automl. *arXiv*.

Zheng, A., and Casari, A. (2018). *Feature engineering for machine learning: Principles and techniques for data scientists* (Sebastopol, CA: O'Reilly Media, Inc).

Zhou, S., Chai, X., Yang, Z., Wang, H., Yang, C., and Sun, T. (2021). Maize–ias: Maize-IAS: A maize image analysis software using deep learning for high–throughput plant phenotyping. *Plant Methods* 17, 1–17. doi: 10.1186/s13007-021-00747-0

Zöller, M.–A., Titov, W., Schlegel, T., and Huber, M. F. (2022). Xautoml: A visual analytics tool for establishing trust in automated machine learning. *arXiv*. doi: 10.48550/arXiv.2202.11954

Zöller, M.–A., and Huber, M. F. (2021). Benchmark and survey of automated machine learning frameworks. *J. Artif. Intell. Res.* 70, 409–472. doi: 10.1613/jair.1.11854

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 15 February 2023
DOI 10.3389/fpls.2023.1123696

OPEN ACCESS

EDITED BY
Ferrante Neri,
University of Surrey, United Kingdom

REVIEWED BY
Huiling Chen,
Wenzhou University, China
Huaming Chen,
The University of Sydney, Australia

*CORRESPONDENCE
Zhiguo Chen
chenzhiguo@nuist.edu.cn

SPECIALTY SECTION
This article was submitted to Sustainable and Intelligent Phytoprotection, a section of the journal Frontiers in Plant Science

RECEIVED 14 December 2022
ACCEPTED 27 January 2023
PUBLISHED 15 February 2023

CITATION
Chen Z, Xing S and Ren X (2023) Efficient Windows malware identification and classification scheme for plant protection information systems. *Front. Plant Sci.* 14:1123696. doi: 10.3389/fpls.2023.1123696

Efficient Windows malware identification and classification scheme for plant protection information systems

Zhiguo Chen[1,2]*, Shuangshuang Xing[1,2] and Xuanyu Ren[1,2]

[1]Engineering Research Center of Digital Forensics, Ministry of Education, Nanjing University of Information Science and Technology, Nanjing, China, [2]School of Computer and Software, Nanjing University of Information Science and Technology, Nanjing, China

Due to developments in science and technology, the field of plant protection and the information industry have become increasingly integrated, which has resulted in the creation of plant protection information systems. Plant protection information systems have modernized how pest levels are monitored and improved overall control capabilities. They also provide data to support crop pest monitoring and early warnings and promote the sustainable development of plant protection networks, visualization, and digitization. However, cybercriminals use technologies such as code reuse and automation to generate malware variants, resulting in continuous attacks on plant protection information terminals. Therefore, effective identification of rapidly growing malware and its variants has become critical. Recent studies have shown that malware and its variants can be effectively identified and classified using convolutional neural networks (CNNs) to analyze the similarity between malware binary images. However, the malware images generated by such schemes have the problem of image size imbalance, which affects the accuracy of malware classification. In order to solve the above problems, this paper proposes a malware identification and classification scheme based on bicubic interpolation to improve the security of a plant protection information terminal system. We used the bicubic interpolation algorithm to reconstruct the generated malware images to solve the problem of image size imbalance. We used the Cycle-GAN model for data augmentation to balance the number of samples among malware families and build an efficient malware classification model based on CNNs to improve the malware identification and classification performance of the system. Experimental results show that the system can significantly improve malware classification efficiency. The accuracy of RGB and gray images generated by the Microsoft Malware Classification Challenge Dataset (BIG2015) can reach 99.76% and 99.62%, respectively.

KEYWORDS

protection information system, terminal protection, malware classification, image enhancement, data augmentation, deep learning

1 Introduction

Due to increasing levels of industrialization and urbanization, dozens of major diseases and pests found on 2 billion hectares of land around the world all year round (Sun et al., 2019). The management of these diseases and pests requires a significant amount of manual input for agricultural plant protection operations, resulting in a sharp rise in labor costs (Yongliang et al., 2019; Brown et al., 2022). Therefore, intelligent plant protection information systems such as rice canopy pest monitoring systems (Li et al., 2022), field pest monitoring and forecasting systems (Liu et al., 2022), meteorological monitoring systems, and crop disease real-time monitoring and early warning systems have been widely used. The visualization and digitization of pest information improve the efficiency of pest forecasting and reduces the amount of work for plant protection staff at the grassroots level. Users can view data in real-time and manage equipment remotely through a cloud platform or mobile application to realize information management, so as to complete wireless transmission, transportation control, and information data sharing among information collection stations at all levels. However, agricultural-related data storage terminals face increasingly complex agricultural and network security situations, are threatened by various malicious software, and bear security risks such as data leakage, data theft, data loss, and data trafficking. Therefore, security systems must respond quickly to malware using new attack techniques, protect terminals from attacks, maintain the security and integrity of plant protection data, and protect the interests of farmers and the benefits of agricultural production. This paper aims to find an effective method to accurately classify malware and its variants into their families, so as to improve the malware identification and classification efficiency and enhance the comprehensive security protection capabilities of terminal systems in the construction of plant protection informatization. Many companies and scholars have proposed various malware classification techniques, which are mainly divided into two categories: signature-based classification and anomaly-based classification (Gandotra et al., 2014). Most commercial antivirus products use a signature-based approach to determine whether the software is malicious by scanning and matching signatures of known malware. This approach can quickly identify existing malware in a malware library with a low error rate but cannot identify unknown malware. Due to developments in computing power and artificial intelligence, the anomaly-based method has attracted much attention. Researchers have proposed many malware classification schemes based on this technique, which effectively overcome the limitations of signature-based methods. Malware classification schemes based on the anomaly method mainly extract features through static and dynamic analysis and selects a classification algorithm to build a model.

Dynamic analysis is the observation of the real behavior of a program at runtime which is achieved by monitoring the program's execution in a sandbox or a virtual machine (Galal et al., 2016; Jamalpur et al., 2018). During monitoring, actions performed by programs (such as library usage, API calls (Salehi et al., 2017), network traffic, etc.) are recorded as reports. Researchers analyze characteristics in the reports to effectively categorize malware. Dynamic analysis methods attempt to discover all the actual operations of a program based on its behavior. Therefore, unknown and variant malware samples can be identified to improve the efficiency of malware classification (Ghiasi et al., 2015). However, dynamic analysis has certain limitations, such as possible infection of terminal systems, lack of suitability for real-time classification, and compromised monitoring due to evasion techniques.

Static analysis is a method of identifying and classifying executable programs without running them. It scrutinizes the "genes" of a file, rather than the current behavior which can be changed or delayed to an unexpected time in order to evade the dynamic analysis (Nissim et al., 2014). Static analysis has been proposed that mostly used by anti-malware products for automatic malware analysis. This technique allows the study of different features to build a classification system that effectively distinguishes the families to which malware belongs, such as opcode instructions (Lu, 2019), binary (Lad and Adamuthe, 2020), API (D'Angelo et al., 2020), PE header information (Rezaei et al., 2021), etc. This method can classify unknown malware and its variants, and is easier to implement than dynamic analysis. However, static malware analysis suffers from low accuracy and a high false positive rate. To overcome these shortcomings, most existing systems combine a large number of different types of features (Kim et al., 2018). Using a large number of features will cause time consumption and memory overhead, and is not suitable for real-time classification.

In recent research, static analysis methods combining malware visualization and deep learning (Liao et al., 2021; Chen et al., 2023) effectively alleviated the pressure of feature engineering technology in processing a large number of features, reduce time overhead, and make up for the shortcomings of traditional static classification methods (Yuan et al., 2020), which has achieved success in the field of malware classification. These methods visualize the binary sequence of malware as gray, RGB, or other types of images as the input of the models and use deep learning algorithms to build effective classification models that are most conducive to distinguishing the families to which malware belongs (Pinhero et al., 2021). However, in the process of visualizing the binary sequence of malware into image representation, most researchers use the zero-filling method which generates images with redundant and irrelevant features (Tekerek and Yapici, 2022). This affects the accuracy of malware classification. The existing malware benchmark datasets have the problem of unbalanced malware family data. The researchers proposed using GAN network (Park et al., 2020; Wang et al., 2022) to expand the data of small class samples to improve the efficiency of malware identification and classification. In response to the above problems, this paper proposes a malware classification system based on bicubic interpolation, Cycle-GAN, and CNNs. The accuracy of the test on the BIG2015 dataset provided by Microsoft can reach 99.76%.

The main contributions of the paper are as follows:

(i)We performed image enhancement on images converted from byte files using bicubic interpolation to preserve the integrity of malware data, addressing malware image size imbalances and image conversion distortions.

(ii)We used Cycle-GAN to perform data augmentation on gray and RGB images transformed from the BIG2015 dataset, solving the data imbalance among malware families.

(iii)We used the optimized DenseNet model to build a system to improve the efficiency of malware classification and the security capabilities of plant protection information terminal systems.

This paper is organized as follows. Section 2 provides an overview of the related work. Section 3 presents the materials and methodology. Section 4 describes the proposed system based on deep learning. Section 5 presents the experimental results and analysis. Section 6 summarizes this paper and future work.

2 Related work

2.1 Malware identification based on deep learning

Deep learning techniques such as CNN and recurrent neural networks (RNNs) have been widely used in the field of malware identification. Kumar et al (Kumar, 2021) proposed a malware classification system based on a fine-tuned convolutional neural network (MCFT-CNN). Without prior knowledge of feature engineering, binary code analysis, reverse engineering, detection, and avoidance, the system can effectively identify unknown malware samples. The classification accuracy of MalImg and BIG2015 datasets reached 99.18% and 98.63%, and the prediction time was 5.14ms and 5.15ms, respectively. The experimental results demonstrated the high efficiency of the system in identifying unknown malware, and the results on different datasets verify the universality of the system. Vasan et al. (2020) proposed an image-based malware classification system that uses a CNN architecture (IMCEC) to identify packed and unpacked malware. The experimental results show that the classification accuracy of packaged and unpackaged malware on the MalImg dataset reaches 98% and 99%, respectively. Vasan et al. (2020) proposed a malware classification system based on deep learning. The proposed fine-tuned convolutional neural network architecture (IMCFN) can effectively detect hidden code, obfuscated malware, and its variants. Experimental results show that the classification accuracy of MalImg and IoT-android datasets can reach 98.82% and 97.35%, respectively. Wang et al. (2021) proposed a gray image-based malware detection and classification system consisting of a deep efficient attention module (DEAM) and a DenseNet module. A detection accuracy of 99.3% was achieved on a dataset constructed from 1,087 benign samples collected by the authors and 1,087 malware samples randomly selected from the MalImg and BIG2015 datasets. The classification accuracy of 98.5% and 97.3% on the MalImg and BIG2015 datasets also verifies that the system can significantly improve the efficiency of malware classification. Gilbert et al (Gibert et al., 2018) proposed a deep learning system based on entropy flow to classify malware. The system used the entropy signal of wavelet transform to describe the change of entropy energy and achieved the purpose of classification by mining the similarity between the malware's entropy streams. Experimental results show that the classification accuracy of the BIG2015 dataset reached 98.28%. Gao et al. (2020) proposed a cloud-based semi-supervised transfer learning (SSTL) framework consisting of detection, prediction, and transfer components. Experimental results on the BIG2015 dataset show that semi-supervised transfer learning can improve the accuracy of detecting components from 94.72% to 96.9%. Hemalatha et al. (2021) proposed an efficient malware classification system based on deep learning methods. The system uses a high-weight class-balanced loss function in the final classification layer of the DenseNet model, which achieves remarkable results in malware classification by addressing the data imbalance problem. The classification accuracy of the system on the MalImg, BIG2015, MaleVis, and Malicia datasets reached 98.23%, 98.46%, 98.21%, and 89.48%, respectively.

Deep learning technology can achieve more flexible malware feature representation, abstract all kinds of information contained in malware images layer by layer, and help to develop automatic and general models for identifying and classifying malware. Therefore, this paper uses the DenseNet deep learning model to build a malware classification system to effectively identify and classify malware and its variants.

2.2 Malware identification based on visualization technology

In the field of malware identification, researchers use visualization technology to visualize malware samples as image representations and identify malware by analyzing the visual similarity between images. Jian et al (Gao et al., 2020). proposed a deep neural network-based malware classification system (SERLA). The system utilizes image visualization and data augmentation techniques to convert the BIG2015 dataset into three-channel RGB images as input to the SERLA system. The experimental results show that the classification accuracy of the SERLA system on the BIG2015 dataset is 98.13%. Gibert et al. (2019) proposed a malware classification system based on gray images and deep learning methods. The system can capture similar characteristics between malware variants and precisely classify them into families. Experimental results show that applying the CNN model to the BIG2015 dataset achieves a classification accuracy of 97.5% and an average classification time of 0.001s. Ni et al. (2018) proposed an efficient malware classification system based on the CNN model and SimHash. The authors converted the disassembly malware code from the BIG2015 dataset into SimHash-based gray images, extracted pixel features through the CNN model, and effectively identified the family of malware. The experimental results show that the classification accuracy of the system on the BIG2015 dataset can reach 99.26%. Kalash et al. (2018) proposed a gray image-based malware classification system. The authors converted malware binary files into gray images and efficiently classified them through a CNN model. The experimental results show that the classification accuracy of the system on the Malimg and BIG2015 datasets reaches 98.52% and 99.97%, respectively. Jang et al. (2020) proposed a fastText-based local feature visualization method. This method extracts local features such as opcodes and API function names from malware and selects important local features in each malware family for embedding and visualizing through a word frequency-inverse document frequency algorithm. The experimental results show that the classification accuracy of this method on the BIG2015 dataset is about 99.65%.

The malware classification method based on visualization technology does not require disassembly and a time-consuming feature extraction process and can capture the difference between malware and its variants, so as to effectively classify malware. Therefore, malware visualization methods are beneficial to improve

classification efficiency while reducing system complexity. Moreover, the visualization method can be applied to large-scale malware classification tasks without employing feature engineering techniques. This paper leverages visualization techniques to convert the BIG2015 dataset into gray and RGB image representations for the efficient classification of malware families.

2.3 Malware identification based on GAN networks

Generative adversarial networks (GANs), which consist of generative networks and discriminative networks, can be used for image-to-image translation and to generate high-quality images. In the field of malware classification, researchers use GANs to augment the data of classes with a small number of samples, so as to solve the problem of unbalanced malware datasets and improve classification efficiency. Tekerek et al (Tekerek and Yapici, 2022). proposed a malware classification system composed of cycle-consistent generative adversarial networks (Cycle-GAN) and DenseNet121 models. The byte files of the BIG2015 dataset were converted into gray and RGB images by B2IMG, the Cycle-GAN model was used to expand the data of the small sample family, and the DenseNet121 model was used to effectively classify the malware. The experimental results show that a classification accuracy of 99.73% is achieved on RGB images converted from the BIG2015 dataset. Rigaki et al (Rigaki and Garcia, 2018) proposed a method of generating network traffic with GANs to simulate other types of traffic. The authors modified the source code of the malware by receiving parameters from the GAN to modify the behavior of its command-and-control (C2) channel, thereby simulating Facebook chat network traffic. Experimental results show that GAN provides effective sample data for malware classification while successfully modifying malware traffic. Won et al (Won et al., 2022) proposed a generative adversarial network-based malware simulation framework (PlausMal-GAN) to augment malware image data. Experimental results show that the framework is beneficial for identifying and predicting zero-day malware-like images. Gao et al. (2022) proposed an efficient classification framework (MaliCage) for packaged malware. Experimental results show that the MaliCage framework composed of a packer detector, malware classifier, and packer GAN can classify packed malware with an accuracy of 91.66%. Singh et al. (2019) proposed a GAN-based malware image generation model (MIGAN). Experimental results show that MIGAN can improve the performance of classifiers by performing data augmentation on malware images generated from binary files, intrusion detection, and log files.

The creation of data labels for the benchmark dataset requires manual marking and is time consuming, however, GANs can learn features from real data and generate similar data without data labels. GANs can be used to generate network traffic and simulate malware data to expand the dataset, thus effectively improving the identification and classification performance. Therefore, this paper uses the Cycle-GAN model to expand the image data to balance the number of samples.

3 Materials and methodology

3.1 Dataset

The Microsoft Malware Classification Challenge Dataset (BIG2015) (Ronen et al., 2018) is a benchmark dataset in the field of malware classification. The dataset contains more than 20,000 assembly and bytecode files composed of 9 different malware families: Ramnit, Lollipop, Kelihos_ver3, Vundo, Simda, Tracur, Kelihos_ver1, Obfuscator.ACY, and Gatak. The specific data distribution is shown in Table 1. Each byte file contains a hexadecimal representation of the file's binary content, excluding headers. Each ASM file contains various metadata information extracted from the binary file, such as logs of function calls, strings, etc. This paper uses the byte files in this dataset for system verification and analysis.

3.2 Bicubic interpolation

Bicubic, Lanczos, and other bicubic interpolation algorithms have been successfully applied to data enhancement, digital splicing of multiple scenes, and information extraction (Rifman, 1973; Bernstein,

TABLE 1 Malware families in the training dataset.

No.	Family	Number of samples	Type
1	Ramnit	1541	Worm
2	Lollipop	2478	Adware
3	Kelihos_ver3	2942	Backdoor
4	Vundo	475	Trojan
5	Simda	42	Backdoor
6	Tracur	751	TrojanDownloader
7	Kelihos_ver1	398	Backdoor
8	Obfuscator.ACY	1,228	Any kind of obfuscated malware
9	Gatak	1,013	Backdoor
	Total	10,868	

1976). Bicubic interpolation preserves the details of the original image as much as possible by interpolating or increasing the number/density of pixels in the image. In this algorithm, the value of the function f at the point (x, y) is obtained by calculating the weighted average of the nearest 16 sample points in the rectangular grid. The interpolation function in each direction is calculated using the formulas of Eq. (1) and Eq. (2).

$$W(x) = \begin{cases} (\alpha+2)|x|^3-(\alpha+3|x|^2+1 & for \quad |x| \le 1 \\ \alpha|x|^3-5\alpha|x|^2+8\alpha|x|-4\alpha & for \quad 1<|x|<2 \\ 0 otherwise & \end{cases} \quad (1)$$

Where x is the distance between the pixel point (x, y) and the last 16 sample points, α is usually 1 or 0.5.

For the interpolated pixel point (x, y) (x, y can be floating numbers), select a point near 4 × 4 and use Eq (2) to calculate the weighted sum.

$$f(x,y) = \sum_{i=0}^{3}\sum_{j=0}^{3} f(x_i,y_j)W(x-x_i)W(y-y_i) \quad (2)$$

As shown in Figure 1, suppose the size of the source image A is $m \times n$, and the size of the scaled target image B is $M \times N$. According to the ratio, the corresponding coordinates of $B(X, Y)$ on A can be obtained from $A(x, y) = A(X \times (m/M), Y \times (n/N))$. of the target image. Point P is the coordinate at (X, Y) corresponding to the target image B on the source image A. Assume that the coordinates of P are $P(x + u, y+v)$, where x, y represent the integer part, and u, v represent the fractional part. As shown in Figure 1, the position of the nearest 16 pixels is represented by α(i, j=1, 2, 3, 4). According to Eq. (1), the influence factor W on the pixel value of point P is used to obtain the pixel value of the corresponding point of the target image, so as to achieve the purpose of image scaling.

In the field of malware classification, bicubic interpolation can effectively balance image size and correct distorted images (Keys, 1981). Dai et al. (2018) proposed a gray image-based malware classification system. The authors used the bicubic interpolation algorithm to equalize the size of the gray image converted from the memory dump file and used the image features extracted from the gradient histogram as input for malware classification. Experimental results show that the system achieves a classification accuracy of 95.2% on the Open Malware Benchmark dataset. Cui et al. (2018) proposed a deep learning-based malware classification system. The author used a bicubic interpolation algorithm to equalize the size of gray images converted from malicious code and used a CNN model to classify malware images. Experimental results show that the accuracy of the system on the MalImg dataset can reach 94.5%.

In this paper, the bicubic interpolation algorithm is used to enhance the gray and RGB images generated by the BIG2015 dataset to overcome the problems of pixel distortion and image size imbalance in the image conversion process. Experimental results show that image enhancement is beneficial to remove redundant and irrelevant features and improve the accuracy of malware classification.

4 Proposed system

The plant protection information system is also threatened by malware when it is monitoring and defending against pests and diseases. In order to ensure the safe and stable operation of the plant protection information system, we proposed a static identification and classification system architecture of malware, as shown in Figure 2. The classification system utilizes bicubic interpolation, Cycle-GAN, and DenseNet121 to improve the efficiency of malware classification. The system mainly includes three parts: (1) Image generation and image enhancement, (2) Data augmentation, and (3) Classification model. Image Generation and Image Enhancement: The hexadecimal features of the byte files in the BIG2015 dataset are converted to decimal features between 0 and 255, and the malware is visualized as gray and RGB images. We use the bicubic interpolation algorithm to enhance the gray and RGB images to solve the problems of image distortion and size imbalance. Data augmentation: We use the Cycle-GAN model to perform data augmentation on a small number of samples in the BIG2015 dataset to address data imbalances among malware families. Classification model: We build an efficient malware identification and classification system using deep learning algorithms (DenseNet121).

4.1 Image generation and image enhancement

This paper uses the byte-to-image method (B2IMG) proposed by Tekerek et al (Tekerek and Yapici, 2022) to convert the byte files of

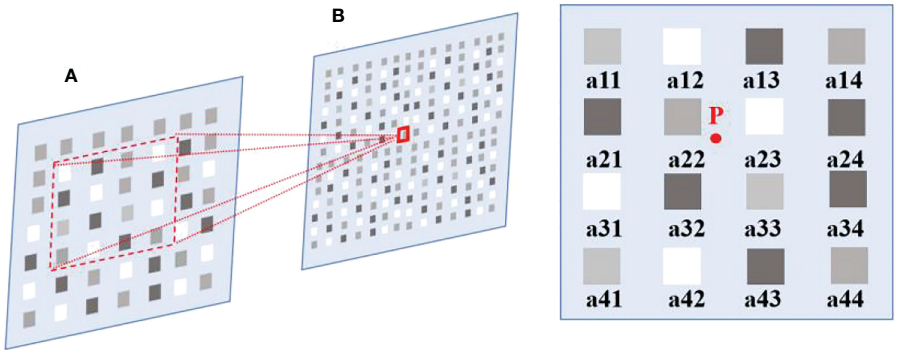


FIGURE 1
Position of the last 16 pixels of point P.

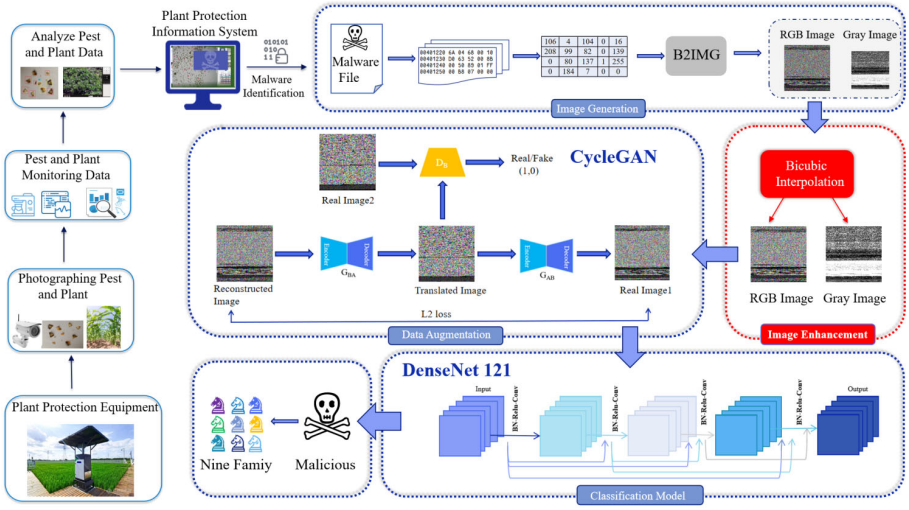


FIGURE 2
Architecture of the proposed system.

the BIG2015 dataset into gray and RGB images, as shown in Algorithm 1. Firstly, the algorithm detects and removes meaningless line numbers, characters, and numbers such as "??" and "00". Secondly, the remaining hexadecimal number is converted to a decimal value between 0 and 255 and is loaded into the pixel array. The aspect ratio of the image is obtained by dividing the total number of decimal array elements by the number of channels in the image and taking the square root of it. Finally, the decimal pixel array elements between 0 and 255 are loaded into the 2-dimensional gray image and the 3-dimensional all-0 value matrix of the RGB image to obtain the image of each malware.

We used the bicubic interpolation algorithm to perform image enhancement on the RGB and gray images generated by the B2IMG method, as shown in Figures 3A, B. All image sizes are unified to 224 × 224 images as input to the DenseNet121 model.

Step 1: While (Read Line with (filename))
Step 2: Line split in pixel array according to the spaces
Step 3: Foreach (item in pixel array)
Step 4: IF (item ==?? OR item<=00)
Step 5: Clear item
Step 6: ELSE
Step 7: Convert item hexadecimal to decimal
Step 8: Load the converted item in pixel array
Step 9: End While
Step 10: image size = Ceil ($\sqrt{\frac{pixel\ array\ length}{color\ channel}}$)
Step 11: Create a matrix with the size of (image size X image size X color channel)
Step 12: Load 0 values in matrix
Step 13: Reshape pixel array with (image size X image size X color channel)
Step 14: Load pixel array in matrix
Step 15: Convert matrix to image

ALGORITHM 1
Algorithm of B2IMG.

4.2 Data augmentation

GANs usually require paired data, but paired data for malware images is hard to obtain in practical applications. Cycle-GAN

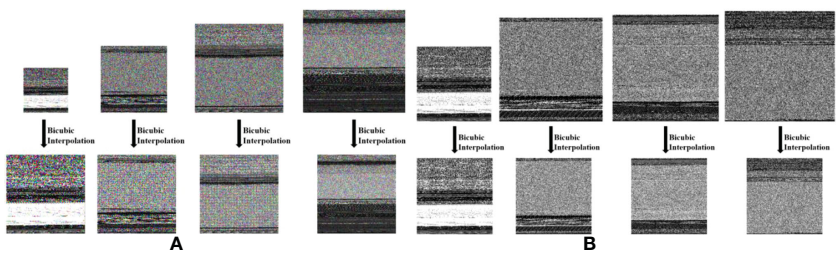


FIGURE 3
Images generated using the bicubic interpolation technique: **(A)** RGB image, **(B)** Gray image.

generates image data without pairing data (Zhu et al., 2017), which greatly reduces the difficulty of malware image augmentation. Therefore, in order to solve the problem of unbalanced malware family samples, this paper uses the Cycle-GAN model to learn the features between different images of the same malware family, so as to augment the data of small sample family malware. Figure 4 shows the malware data augmentation process, which includes 2 generative models and 2 discriminative models.

The specific process of generating the reconstructed malware image A from the real malware image A is as follows (the real malware image B generated the reconstructed malware image B is the same).

Firstly, train the generative model G_{AB}(G_{AB}:$A \rightarrow B$ and the discriminative model D_B, obtain the adversarial loss function L_{GAN} minimized by Eq. (3) and Eq. (4) to obtain the optimal model, and convert the real malware image A into the simulated malware image B.

$$\mathcal{L}_{GAN}(G_{AB}, D_B, A, B) = \mathbb{E}_{a \sim P_{data}(a)}[(D_B(G_{AB}(a)) - 1)^2] \tag{3}$$

$$\mathcal{L}_{GAN}(G_{AB}, D_B) = \frac{\mathbb{E}_{a \sim P_{data}(a)}[(D_B(G_{AB}(a)))^2] + \mathbb{E}_{b \sim P_{data}(b)}[(D_B(b) - 1)^2]}{2} \tag{4}$$

denotes the collection of malware images belonging to category A and $\{b_j\}_{j=1}^{M}$ denotes the collection of malware images belonging to category B. $a \sim P_{data}(a)$ denotes the data distribution of malware images of category A and $b \sim P_{data}(b)$ denotes the malware images of category B data distribution. Secondly, the simulated malware image B is reconstructed into malware image A. By minimizing the cyclic consistency loss function $\mathcal{L}_{cyc}$ in Eq. (5) and identifying loss function $\mathcal{L}_{idt}$ in Eq. (6), the parameters of the generated model were adjusted to ensure the similarity between the reconstructed malware image A and the real malware image A.

$$\begin{aligned}\mathcal{L}_{cyc}(G_{AB}, G_{BA}, \lambda_A, \lambda_B) &= \lambda_A \cdot \mathbb{E}_{a \sim P_{data}(a)}[||G_{BA}(G_{AB}(a)) - a||_1] \\ &+ \lambda_B \cdot \mathbb{E}_{b \sim P_{data}(b)}[||G_{AB}(G_{BA}(b)) - b||_1]\end{aligned} \tag{5}$$

$$\begin{aligned}\mathcal{L}_{idt}(G_{AB}, G_{BA}, \lambda_A, \lambda_B, \lambda_{idt}) &= \mathbb{E}_{a \sim P_{data}(a)}[||G_{BA}(G_{AB}(a)) - a||_1] \cdot \lambda_A \cdot \lambda_{idt} \\ &+ \mathbb{E}_{b \sim P_{data}(b)}[||G_{AB}(G_{BA}(b)) - b||_1] \cdot \lambda_B \cdot \lambda_{idt}\end{aligned} \tag{6}$$

The generative model can be expressed as a mapping function G_{AB}: $A \rightarrow B$, G_{BA}: $B \rightarrow A$. The discriminative model is expressed as D_A, D_B. Λ_A and Λ_B represent the cycle consistency loss weights of A and B images, respectively. Λ_{idt} denotes the identity loss weight of the reconstructed image A and the real image A.

Finally, the optimal performance of the Cycle-GAN network is obtained by minimizing the functions G_{AB}*, G_{BA}*in Eq. (8).

$$\begin{aligned}\mathcal{L}(G_{AB}, G_{BA}, D_A, D_B, \lambda_A, \lambda_B, \lambda_{idt}) &= \mathcal{L}_{GAN}(G_{AB}, D_B, A, B) + \mathcal{L}_{GAN}(G_{BA}, D_A, B, A) \\ &+ \mathcal{L}_{cyc}(G_{AB}, G_{BA}, \lambda_A, \lambda_B) + \mathcal{L}_{idt}(G_{AB}, G_{BA}, \lambda_A, \lambda_B, \lambda_{idt})\end{aligned} \tag{7}$$

$$G_{AB}*, G_{BA}* = arg\, D_A, D_B \max_{G_{AB}, G_{BA}} min\mathcal{L}(G_{AB}, G_{BA}, D_A, D_B\lambda_A, \lambda_B, \lambda_{idt}) \tag{8}$$

The function $\mathcal{L}_{cyc}(G_{AB}, G_{BA}, D_A, D_B\lambda_A, \lambda_B, \lambda_{idt})$ in Eq. (7) represents the sum of loss functions. As shown in Table 2, we used the Cycle-GAN model to augment 300, 738, 400, 100, and 400 samples for the 5 families of Vundo, Simda, Tracur, Kelihos_ver1, and Obfuscator.ACY, respectively.

4.3 Classification model

With improvements in computing power and the scale of the explosion of malware data, traditional machine-learning algorithms are no longer sufficient to identify and classify malware families effectively. The image-based deep learning method does not require specialized domain knowledge and manual parameter adjustment, and can learn independently through the model to improve classification efficiency. As shown in Figure 5; Huang et al. (2017) proposed a dense convolutional network (DenseNet) consisting of three dense blocks in CVPR in 2017. In each dense block module, the

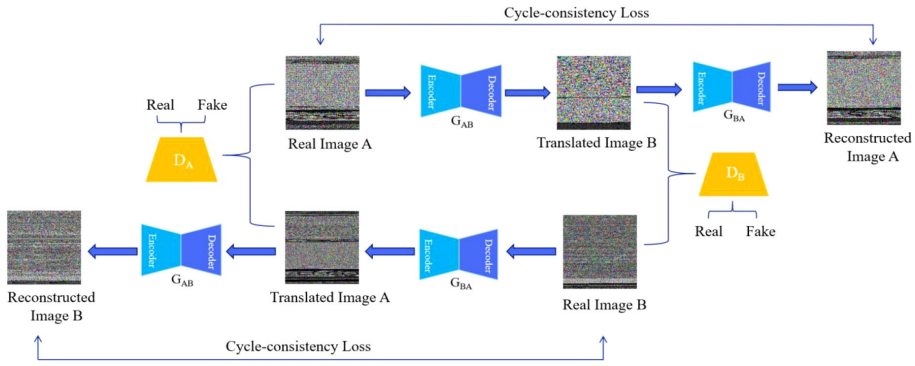


FIGURE 4
Data augmentation methods.

TABLE 2 The number of trainings, augmented trainings, and test data for BIG2015.

No.	Family	Train Data	Augmented Train Data	Test Data
1	Ramnit	1,079	0	462
2	Lollipop	1,735	0	743
3	Kelihos_ver3	2,060	0	882
4	Vundo	333	300	142
5	Simda	30	738	12
6	Tracur	526	400	225
7	Kelihos_ver1	279	100	119
8	Obfuscator.ACY	860	400	368
9	Gatak	710	0	303

output features of all previous layers are used as the input of subsequent layers. The reuse of features can reduce network parameters and reduce model complexity. Compared with other networks, the DenseNet optimization problem is less difficult and can be extended to hundreds of layers. The DenseNet structure integrates identity mapping, deep supervision and attributes of different depths, which can alleviate the problem of gradient disappearance and enhance feature transfer and usage efficiency. Therefore, this paper uses the classic DenseNet (DenseNet121) to build a malware identification and classification model based on malware images. In order to prevent overfitting, we used dropout to simplify the network structure and improved the model's generalization ability. After the FC layer of DenseNet121, we added an FC layer of size 512 to prevent overfitting and reduce redundant parameters.

5 Experimental results and analysis

5.1 Experimental setup

The BIG2015 dataset was split into two as 80% training and 20% testing. Test data was not used during the training phase. All experiments used 10-fold cross validation to prevent overfitting. According to the principle of the 10-fold CV model, 90% of the data at each fold training phase was used for training, and the remaining 10% was used in the validation phase. The final results were obtained with test data never present in the training phase. The experiment mainly uses Pytorch and the programming language is Python 3.8. Stochastic Gradient Descent (SGD) was used for optimization in this experiment, the value of learning rate was 0.03 and the value of momentum was 0.9.

This paper uses metrics such as precision, recall, accuracy and F1-score to evaluate the effectiveness of the proposed system. These metrics are widely used in the field of classification and can objectively measure the performance of malware classification systems.

Accuracy is the most commonly used measure of evaluation, and is defined as the number of samples that correctly predict the malware's family divided by the total number of samples. Specificity represents the proportion of the sum of predicted and actual sample number not in this malware family to the sum of actual sample number not in this malware family. Precision represents the proportion of the number of samples that are correctly predicted to belong to the actual malware family to the number of samples that are predicted to belong to that malware family. Recall represents the ratio of the number of samples that are correctly predicted to belong to the family of malware to the number of families that the samples actually belong to. F1-score is a comprehensive evaluation index for measuring precision and recall

$$Accuracy = \frac{TP + TN}{TP + FN + TN + FP} \tag{9}$$

$$Specificity = \frac{TN}{TN + FP} \tag{10}$$

$$Precision = \frac{TP}{TP + FP} \tag{11}$$

$$Recall = \frac{TP}{TP + FN} \tag{12}$$

$$F1 = 2\frac{Precision \times Recall}{Precision + Recall} \tag{13}$$

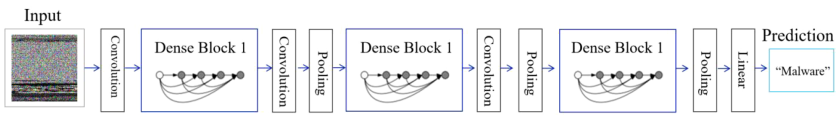


FIGURE 5
Network structure of DenseNet.

Where *TP* is true positive, *TN* is true negative, *FP* is false positive, and *FN* is false negative.

5.2 Experiment with data augmentation

In order to verify the effectiveness of the system, this paper conducts experiments on gray images and RGB images based on the original BIG2015 dataset and the augmented dataset.

5.2.1 Experimental results on gray images

We used the method described in Section 4.1 to convert the malware into a gray image representation and reconstructed the generated malware gray images using the bicubic interpolation algorithm. The Cycle-GAN model was used for data augmentation to build a malware identification and classification system based on DenseNet121.

Table 3 shows that implementing the data augmentation method on gray images can improve the AUC, specificity, precision, recall, F1-score, and classification accuracy of most malware families. The accuracy of Vundo and Obfuscator.ACY families before data augmentation were 99.66% and 99.08%. After data augmentation the accuracy increased by 0.06% and 0.28%, reaching 99.72% and 99.36%. It is worth noting that after data augmentation, the AUC and recall of the Simda family both reached 100% from 95.81% and 91.67%, respectively, indicating that all Simda family malware were accurately classified. The specificity and precision of the Tracur family were increased from 99.67% and 95.56% to 99.70% and 96.99% after data augmentation. Although the accuracy of the Kelihos_ver1 family decreased by 0.16% after data augmentation compared with that before data augmentation, the misclassification of a small data in a small sample family has little impact on the overall performance of the system. Therefore, data augmentation based on gray images can effectively improve the performance of malware classification systems.

5.2.2 Experimental results on RGB images

In order to verify the generalization of the proposed system to feature images that have different textures and the effectiveness of the system, we used the method described in Section 4.1 to convert malware into RGB image representations to construct a malware identification and classification system.

As shown in Table 4, the use of data augmentation on RGB images can significantly improve the evaluation indicators such as AUC, specificity, precision, recall, F1-score, and classification accuracy of 8 malware families such as Ramnit, Kelihos_ver3, and Vundo. After data augmentation, the classification accuracy of Vundo, Simda, Tracur and Obfuscator.ACY families increased from 99.85%, 99.94%, 99.48%, and 99.29% to 99.91%, 99.97%, 99.60%, and 99.36%, which was an increase of 0.06%, 0.03%, 0.12%, and 0.07%. Although the classification accuracy of the Ramnit, Kelihos_ver1, and Gatak families did improve, the F1-score of the Ramnit family as a measure of precision and recall increased by 0.03%, reaching 98.18%. The Kelihos_ver1 and Gatak families still maintain high classification accuracy of 99.97% and 99.88%. Experimental results show that the augmentation of RGB image data can solve the problem of sample imbalance among malware families, which is beneficial to improve the performance of malware identification and classification systems.

The results in Tables 3, 4 verify that our proposed malware identification and classification system based on bicubic interpolation, Cycle-GAN, and DenseNet121 model can effectively identify and classify malware into their corresponding families. Meanwhile, it can be seen that the classification performance of the system based on the RGB image representation of malware is better than that of the gray image representation. The classification accuracy of Lollipop, Kelihos_ver3, and Gatak families without data augmentation is 99.69%, 99.94%, and 99.88% for RGB images. These values are higher than the 99.42%, 99.85% and 99.63% accuracy for gray images. After data augmentation, the classification performance of the system based on RGB images was greatly improved in almost all families compared to gray images. Notably, the classification system based on RGB images achieved an accuracy of over 99.3% on each malware family. In particular, the classification accuracy on the Kelihos_ver3 family reached 100%. The experimental results show that the RGB image representation of malware has richer texture patterns and more feature information than the gray image

TABLE 3 Experimental results of gray images converted from BIG2015 dataset.

		Ramnit	Lollipop	Kelihos_ver3	Vundo	Simda	Tracur	Kelihos_ver1	Obfuscator.ACY	Gatak
Without Augmentation	Accuracy	0.9954	0.9942	0.9985	0.9966	0.9994	0.9948	0.9991	0.9908	0.9963
	AUC	0.9901	0.9924	0.9982	0.9695	0.9581	0.9826	0.9915	0.9770	0.9905
	Specificity	0.9975	0.9956	0.9987	0.9994	0.9997	0.9967	0.9997	0.9948	0.9976
	Precision	0.9849	0.9852	0.9966	0.9859	0.9167	0.9556	0.9916	0.9592	0.9769
	Recall	0.9827	0.9892	0.9977	0.9396	0.9167	0.9685	0.9833	0.9592	0.9834
	F1-Score	0.9838	0.9872	0.9972	0.9622	0.9167	0.9620	0.9875	0.9593	0.9801
With Augmentation	Accuracy	0.9942	0.9926	0.9985	0.9972	0.9997	0.9942	0.9975	0.9936	0.9982
	AUC	0.9810	0.9905	0.9986	0.9759	1.0000	0.9764	0.9753	0.9939	0.9960
	Specificity	0.9996	0.9944	0.9983	0.9994	0.9997	0.9970	0.9994	0.9935	0.9987
	Precision	0.9978	0.9912	0.9955	0.9859	0.9367	0.9699	0.9832	0.9384	0.9868
	Recall	0.9624	0.9865	0.9989	0.9524	1.0000	0.9558	0.9512	0.9943	0.9934
	F1-Score	0.9798	0.9838	0.9972	0.9689	0.9565	0.9579	0.9669	0.9708	0.9901

TABLE 4 Experimental results of RGB images converted from BIG2015 dataset.

		Ramnit	Lollipop	Kelihos_ver3	Vundo	Simda	Tracur	Kelihos_ver1	Obfuscator.ACY	Gatak
Without Augmentation	Accuracy	0.9948	0.9969	0.9994	0.9985	0.9994	0.9948	0.9997	0.9929	0.9988
	AUC	0.9924	0.9961	0.9989	0.9830	0.9582	0.9730	0.9998	0.9841	0.9964
	Specificity	0.9957	0.9976	1.0000	1.0000	0.9997	0.9980	0.9997	0.9955	0.9993
	Precision	0.9740	0.9919	1.0000	1.0000	0.9167	0.9733	0.9916	0.9647	0.9934
	Recall	0.9890	0.9946	0.9977	0.9660	0.9167	0.9481	1.0000	0.9726	0.9934
	F1-Score	0.9815	0.9933	0.9989	0.9827	0.9167	0.9605	0.9958	0.9686	0.9934
With Augmentation	Accuracy	0.9948	0.9963	1.0000	0.9991	0.9997	0.9960	0.9997	0.9936	0.9988
	AUC	0.9855	0.9952	1.0000	0.9897	0.9615	0.9780	0.9998	0.9927	0.9978
	Specificity	0.9986	0.9972	1.0000	1.0000	1.0000	0.9990	0.9997	0.9938	0.9990
	Precision	0.9913	0.9906	1.0000	1.0000	1.0000	0.9867	0.9916	0.9511	0.9901
	Recall	0.9724	0.9933	1.0000	0.9793	0.9231	0.9569	1.0000	0.9915	0.9967
	F1-Score	0.9818	0.9919	1.0000	0.9896	0.9600	0.9716	0.9958	0.9709	0.9934

representation and is more conducive to the construction of classification systems.

5.3 Comparison and discussion

To demonstrate the effectiveness of our proposed system, Table 5 shows a comparison between the results of this paper and existing relevant studies based on the BIG2015 dataset. The studies used CNN, DenseNet, EfficientNetB1 and EfficientNetB7 models, as well as frameworks such as SERLA, RNN+ SSTL and CNN+ Cycle-GAN.

We converted byte files into image representations and adopted an improved CNN model (DenseNet121) to build the system. Compared with Gilbert et al (Gibert et al., 2018; Gibert et al., 2019). and Hemalatha et al. (2021) using the CNN model, our classification accuracy has increased by 1.44%, 2.22%, and 1.26%, reaching 99.72%.

We used the Cycle-GAN model to balance the number of samples in the malware dataset to build an identification and classification system. Compared to the classification system composed of DEAM and DenseNet (Wang et al., 2021), the accuracy, precision, and F1-Score are improved by 2.46%, 3.6%, and 2.99%, respectively. In terms of accuracy, precision, and F1-Score, our system improved by 1.45%, 0.22%, and 0.09% compared with the SERLA model (Gao et al., 2020), which was also constructed based on RGB images generated by the BIG2015 dataset, reaching 99.76%, 98.9%, and 98.39%. Compared to the above hybrid models composed of multiple classification modules,

TABLE 5 Comparison of the proposed system to systems in the literature using the BIG2015 dataset.

Authors	Year	Models	Dataset	AUC	Precision	Recall	F1-Score	Accuracy
Gibert et al. (2018)	2018	CNN	Bytes	–	–	–	96.36%	98.28%
Gibert et al. (2019)	2020	CNN	Bytes (Grayscale)	–	94.00%	–	–	97.5%
Hemalatha et al. (2021)	2021	CNN	Bytes (Grayscale)	–	98.58%	97.84%	98.21%	98.46%
Wang et al. (2021)	2021	CNN+ DEAM	Bytes (Grayscale)	–	95.3%	95.4%	95.4%	1
Gao et al. (2020)	2020	RNN+ SSTL	Bytes+ ASM	–	96.92%	96.9%	96.81%	96.9%
Jian et al (Gao et al., 2020)	2021	SERLA	Bytes+ ASM	–	98.68%	97.93%	98.3%	98.31%
Acharya et al. (2021)	2021	EfficientNetB1	Bytes (Grayscale)	–	96.00%	97.00%	97.00%	98.57%
Pratama et al (Pratama and Sidabutar, 2022)	2022	EfficientNetB7	Bytes (Grayscale)	98.01%	97.96%	97.93%	97.93%	99.56%
Pratama et al (Pratama and Sidabutar, 2022)	2022	EfficientNetB7	Bytes (RGB)	98.30%	98.36%	98.35%	98.34%	99.63%
Tekerek et al (Tekerek and Yapici, 2022)	2022	CNN+ Cycle-GAN	Bytes (Grayscale)	98.13%	97.53%	96.50%	96.93%	99.58%
Tekerek et al (Tekerek and Yapici, 2022)	2022	CNN+ Cycle-GAN	Bytes (RGB)	98.51%	98.52%	97.16%	97.76%	99.73%
Proposed System	–	**CNN**	**Bytes (Grayscale)**	**98.33%**	**97.25%**	**96.89%**	**97.07%**	**99.61%**
Proposed System	–	**CNN**	**Bytes (RGB)**	**98.69%**	**97.84%**	**97.53%**	**97.68%**	**99.72%**
Proposed System	–	**CNN+ Cycle-GAN**	**Bytes (Grayscale)**	**98.75%**	**97.62%**	**97.72%**	**97.47%**	**99.62%**
Proposed System	–	**CNN+ Cycle-GAN**	**Bytes (RGB)**	**98.89%**	**98.90%**	**97.92%**	**98.39%**	**99.76%**

Bold text highlights authors' contributions and experimental results.

we only used a CNN model to identify and classify the families to which malware belongs, which can effectively reduce model complexity, time, and memory consumption.

We used the bicubic interpolation algorithm to enhance the malware images generated by the BIG2015 dataset to solve the problem of image size imbalance and effectively improve the performance of the malware identification and classification system. The classification accuracy of gray images and RGB images is 0.04% and 0.03% higher than that of the same image representation in (Tekerek and Yapici, 2022), reaching 99.62% and 99.76%. Compared with the EfficientNet-B model proposed by Acharya et al. (2021), the classification accuracy on gray images increased by 1.05%, reaching 99.62%. We compared our model to the B2IMG-based EfficientNetB7 model in (Pratama and Sidabutar, 2022)and achieve improved accuracy of 0.06% and 0.13% for gray and RGB images, reaching 99.62% and 99.76% respectively.

We combined image enhancement and data augmentation techniques to preserve more malware classification information while maintaining image data integrity, generating high-quality malware images for small sample families that balance malware data distribution. The 99.76% accuracy, 98.9% precision, 97.92% recall, 98.39% F1-score, and 98.89% AUC on the RGB images generated by the BIG2015 dataset prove that our proposed system can effectively identify and classify malware.

6 Conclusions and future work

With the exponential growth of the number of malware and its variants, the threat to plant protection information systems that store massive amounts of agricultural data is increasing. As a result, it is critical to effectively identify and classify malware. Existing malware classification schemes based on malware visualization and deep learning mainly identify and classify malware variants by analyzing the similarity of malware binary images. However, the images generated by such schemes have the problem of unbalanced image size and contain irrelevant and redundant features, which affects the accuracy of malware classification. In addition, the unbalanced data affects the classification performance of the system. Therefore, we proposed a malware identification and classification scheme based on DenseNet121 and Cycle-GAN models. The scheme used bicubic interpolation technology to enhance malware images, which solved the problem of image distortion and size imbalance caused by removing redundant and irrelevant features. Using the Cycle-GAN model for data augmentation solved the problem of unbalanced samples of malware families and effectively improves the efficiency of malware classification. The experimental results show that the AUC, precision, recall, F1-score, and accuracy of the proposed system on gray images are 98.75%, 97.62%, 97.72%, 97.47%, and 99.62%. The system can achieve 98.89%, 98.90%, 97.92%, 98.39%, and 99.76% on RGB images. Therefore, the system deployed on the plant protection information terminal can effectively prevent malware attacks, maintain the safety and integrity of plant protection data, and protect farmers' interests and agricultural production efficiency. The BIG2015 dataset does not contain header information and cannot generate a complete image of malware. In future research, we will further collect complete malware samples for visual analysis and research, consider the problems of system complexity, cost, delay and throughput brought by system operation, balance the accuracy and time consumption, and further improve the efficiency of malware classification model.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here:https://www.kaggle.com/competitions/microsoft-malware-prediction/data.

Author contributions

ZC, SX, and XR Efficient Windows Malware Identification and Classification Scheme for Plant Protection Information Systems With the exponential growth of the number of malware and its variants, the threat to plant protection information systems that store massive amounts of agricultural data is increasing. It is critical to effectively identify and classify malware. Existing malware classification schemes based on malware visualization and deep learning mainly identify and classify malware variants by analyzing the similarity of malware binary images. However, the images generated by such schemes have the problem of unbalanced image size and contain irrelevant and redundant features, which affects the accuracy of malware classification. In addition, the unbalanced data affects the classification performance of the system. Therefore, we proposed a malware identification and classification scheme based on DenseNet121 and Cycle-GAN models. The scheme used bicubic interpolation technology to enhance malware images, which solved the problem of image distortion and size imbalance caused by removing redundant and irrelevant features. Using the Cycle-GAN model for data augmentation solved the problem of unbalanced samples of malware families. The system deployed on the plant protection information terminal can prevent malware attacks, maintain the safety and integrity of plant protection data. All authors contributed to the article and approved the submitted version.

Funding

This work is supported by the National Natural Science Foundation of China (Grant No. 62102190).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Acharya, V., Ravi, V., and Mohammad, N. (2021). "EfficientNet-based convolutional neural networks for malware classification," in *2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, Vol. 1-6.

Bernstein, R. (1976). Digital image processing of earth observation sensor data. *IBM J. Res. Dev.* 20 (1), 40–57. doi: 10.1147/rd.201.0040

Brown, J., Qiao, Y., Clark, C., Lomax, S., Rafifique, K., and Sukkarieh, S. (2022). Automated aerial animal detection when spatial resolution conditions are varied. *Comput. Electron. Agric.* 193, 106689. doi: 10.1016/j.compag.2022.106689

Chen, Y., Gan, H., Chen, H., Zeng, Y., Xu, L., Heidari, A. A., et al. (2023). Accurate iris segmentation and recognition using an end-to-end unified framework based on MADNet and DSANet. *Neurocomputing* 517, 264–278. doi: 10.1016/j.neucom.2022.10.064

Cui, Z., Xue, F., Cai, X., Cao, Y., Wang, G. G., and Chen, J. (2018). Detection of malicious code variants based on deep learning. *IEEE Trans. Ind. Informatics.* 14 (7), 3187–3196. doi: 10.1109/TII.2018.2822680

Dai, Y., Li, H., Qian, Y., and Lu, X. (2018). A malware classification method based on memory dump grayscale image. *Digital Invest.* 27, 30–37. doi: 10.1016/j.diin.2018.09.006

D'Angelo, G., Ficco, M., and Palmieri, F. (2020). Malware detection in mobile environments based on autoencoders and API-images. *J. Parallel Distributed Computing.* 137, 26–33. doi: 10.1016/j.jpdc.2019.11.001

Galal, H. S., Mahdy, Y. B., and Atiea, M. A. (2016). Behavior-based features model for malware detection. *J. Comput. Virol. Hacking Techniques* 12 (2), 59–67. doi: 10.1007/s11416-015-0244-0

Gandotra,, Ekta,, Divya, B., and Sanjeev, S. (2014). Malware analysis and classification: A survey. *J. Inf. Secur.* 5 (2), 56–64. doi: 10.4236/jis.2014.52006

Gao, X., Hu, C., Shan, C., and Han, W. (2022). MaliCage: A packed malware family classification framework based on DNN and GAN. *J. Inf. Secur. Applications.* 68, 103267. doi: 10.1016/j.jisa.2022.103267

Gao, X., Hu, C., Shan, C., Liu, B., Niu, Z., and Xie, H. (2020). Malware classification for the cloud *via* semi-supervised transfer learning. *J. Inf. Secur. Applications.* 55, 102661. doi: 10.1016/j.jisa.2020.102661

Ghiasi, M., Sami, A., and Salehi, Z. (2015). Dynamic VSA: a framework for malware detection based on register contents. *Eng. Appl. Artif. Intell.* 44, 111–122. doi: 10.1016/j.engappai.2015.05.008

Gibert, D., Mateu, C., Planes, J., and Vicens, R. (2018). "Classification of malware by using structural entropy on convolutional neural networks," in *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 32.

Gibert, D., Mateu, C., Planes, J., and Vicens, R. (2019). Using convolutional neural networks for classification of malware represented as images. *J. Comput. Virol. Hacking Techniques.* 15, 15–28. doi: 10.1007/s11416-018-0323-0

Hemalatha, J., Roseline, S. A., Geetha, S., Kadry, S., and Damaševičius, R. (2021). An efficient densenet-based deep learning model for malware detection. *Entropy* 23 (3), 344. doi: 10.3390/e23030344

Huang, G., Liu, Z., van der Maaten, L., and Weinberger, K. Q. (2017). "Densely connected convolutional networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 4700–4708.

Jamalpur, S., Navya, Y. S., Raja, P., Tagore, G., and Rao, G. R. K. (2018). "Dynamic malware analysis using cuckoo sandbox," in *2018 Second international conference on inventive communication and computational technologies (ICICCT)*. 1056–1060.

Jang, S., Li, S., and Sung, Y. (2020). Fasttext-based local feature visualization algorithm for merged image-based malware classification framework for cyber security and cyber defense. *Mathematics* 8 (3), 460. doi: 10.3390/math8030460

Kalash, M., Rochan, M., Mohammed, N., Bruce, N. D., Wang, Y., and Iqbal, F. (2018). "Malware classification with deep convolutional neural networks," in *2018 9th IFIP international conference on new technologies, mobility and security (NTMS)*. 1–5.

Keys, R. (1981). Cubic convolution interpolation for digital image processing. *IEEE Trans. acoustics speech Signal processing.* 29 (6), 1153–1160. doi: 10.1109/TASSP.1981.1163711

Kim, T., Kang, B., Rho, M., Sezer, S., and Im, E. G. (2018). A multimodal deep learning method for android malware detection using various features. *IEEE Trans. Inf. Forensics Security.* 14 (3), 773–788. doi: 10.1109/TIFS.2018.2866319

Kumar, S. (2021). MCFT-CNN: Malware classification with fine-tune convolution neural networks using traditional and transfer learning in Internet of things. *Future Generation Comput. Syst.* 125, 334–351. doi: 10.1016/j.future.2021.06.029

Lad, S. S., and Adamuthe, A. C. (2020). Malware classification with improved convolutional neural network model. *Int. J. Comput. Netw. Inf. Secur.* 12, 30–43. doi: 10.48550/arXiv.1906.04593

Liao, Z., Lan, P., Fan, X., Kelly, B., Innes, A., and Liao, Z. (2021). SIRVD-DL: A COVID-19 deep learning prediction model based on time-dependent SIRVD. *Comput. Biol. Med.* 138, 104868. doi: 10.1016/j.compbiomed.2021.104868

Li, S., Feng, Z., Yang, B., Li, H., Liao, F., Gao, Y., et al. (2022). An intelligent monitoring system of diseases and pests on rice canopy. *Front. Plant Science.* 13, 972286. doi: 10.3389/fpls.2022.972286

Liu, C., Zhai, Z., Zhang, R., Bai, J., and Zhang, M. (2022). Field pest monitoring and forecasting system for pest control. *Front. Plant Science.* 13, 990965. doi: 10.3389/fpls.2022.990965

Lu, R. (2019). Malware detection with lstm using opcode language. *arXiv preprint*, 1906.04593. doi: 10.48550/arXiv.1906.04593

Ni, S., Qian, Q., and Zhang, R. (2018). Malware identification using visualization images and deep learning. *Comput. Security.* 77, 871–885. doi: 10.1016/j.cose.2018.04.005

Nissim, N., Moskovitch, R., Rokach, L., and Elovici, Y. (2014). Novel active learning methods for enhanced PC malware detection in windows OS. *Expert Syst. Appl.* 41 (13), 5843–5857. doi: 10.1016/j.eswa.2014.02.053

Park, M., Tran, D. Q., Jung, D., and Park, S. (2020). Wildfire-detection method using DenseNet and CycleGAN data augmentation-based remote camera imagery. *Remote Sens.* 12 (22), 3715. doi: 10.3390/rs12223715

Pinhero, A., Anupama, M. L., Vinod, P., Visaggio, C. A., Aneesh, N., Abhijith, S., et al. (2021). Malware detection employed by visualization and deep neural network. *Comput. Secur.* 105, 102247. doi: 10.1016/j.cose.2021.102247

Pratama, H. Y., and Sidabutar, J. (2022). "Malware classification and visualization using EfficientNet and B2IMG algorithm," in *2022 International Conference on Advanced Computer Science and Information Systems (ICACSIS)*. 75–80.

Rezaei, T., Manavi, F., and Hamzeh, A. (2021). A PE header-based method for malware detection using clustering and deep embedding techniques. *J. Inf. Secur. Applications.* 60, 102876. doi: 10.1016/j.jisa.2021.102876

Rifman, S. S. (1973). "Digital rectification of ERTS multispectral imagery," in *NASA. Goddard space flight center symp. on significant results obtained from the ERTS-1*, vol. 1. , 1973.

Rigaki, M., and Garcia, S. (2018). "Bringing a gan to a knife-fight: Adapting malware communication to avoid detection," in *Proceedings of the 2018 IEEE security and privacy workshops (SPW)*. (San Francisco, CA, USA: IEEE), 70–75.

Ronen, R., Radu, M., Feuerstein, C., Yom-Tov, E., and Ahmadi, M. (2018). Microsoft Malware classification challenge. *arXiv preprint*, 1802.10135. doi: 10.48550/arXiv.1802.10135

Salehi, Z., Sami, A., and Ghiasi, M. (2017). MAAR: Robust features to detect malicious activity based on API calls, their arguments and return values. *Eng. Appl. Artif. Intell.* 59, 93–102. doi: 10.1016/j.engappai.2016.12.016

Singh, A., Dutta, D., and Saha, A. (2019). "MIGAN: malware image synthesis using GANs," in *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 33. 10033–10034.

Sun, F., Wang, X., and Zhang, R. (2019). "A new optimization method application to agricultural plant protection UAV scheduling," in *Proceedings of the 2019 6th International Conference on Information Science and Control Engineering (ICISCE)*. (Shanghai, China: IEEE), 80–84.

Tekerek, A., and Yapici, M. M. (2022). A novel malware classification and augmentation model based on convolutional neural network. *Comput. Security.* 112, 102515. doi: 10.1016/j.cose.2021.102515

Vasan, D., Alazab, M., Wassan, S., Naeem, H., Safaei, B., and Zheng, Q. (2020). IMCFN: Image-based malware classification using fine-tuned convolutional neural network architecture. *Comput. Networks.* 171, 107138. doi: 10.1016/j.comnet.2020.107138

Vasan, D., Alazab, M., Wassan, S., Safaei, B., and Zheng, Q. (2020). Image-based malware classification using ensemble of CNN architectures (IMCEC). *Comput. Secur.* 92, 101748. doi: 10.1016/j.cose.2020.101748

Wang, H., Lu, X., and Deng, F. (2022). "Improving CycleGAN for image-to-Image style transfer by DenseNet," in *2022 7th International Conference on Computer and Communication Systems (ICCCS)*. 326–330.

Wang, C., Zhao, Z., Wang, F., and Li, Q. (2021). A novel malware detection and family classification scheme for IoT based on DEAM and DenseNet. *Secur. Communication Networks.* 2021, 6658842. doi: 10.1155/2021/6658842

Won, D. O., Jang, Y. N., and Lee, S. W. (2022). PlausMal-GAN: Plausible malware training based on generative adversarial networks for analogous zero-day malware detection. *IEEE Trans. Emerging Topics Computing* 1. doi: 10.1109/TETC.2022.3170544

Yongliang, Q., Truman, M., and Sukkarieh, S. (2019). Cattle segmentation and contour extraction based on mask r-CNN for precision livestock farming. *Comput. Electron. Agric.* 165, 104958. doi: 10.1016/j.compag.2019.104958

Yuan, B., Wang, J., Liu, D., Guo, W., Wu, P., and Bao, X. (2020). Byte-level malware classification based on markov images and deep learning. *Comput. Secur.* 92, 101740. doi: 10.1016/j.cose.2020.101740

Zhu, J. Y., Park, T., Isola, P., and Efros, A. A. (2017). "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proceedings of the IEEE international conference on computer vision*. 2223–2232.

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 06 March 2023
DOI 10.3389/fpls.2023.1139094

OPEN ACCESS

EDITED BY
Yu Xue,
Nanjing University of Information Science and Technology, China

REVIEWED BY
Gangyan Xu,
Hong Kong Polytechnic University, Hong Kong SAR, China
Qinqin Fan,
Shanghai Maritime University, China

*CORRESPONDENCE
Yu-Jun Zheng
yujun.zheng@computer.org

SPECIALTY SECTION
This article was submitted to Sustainable and Intelligent Phytoprotection, a section of the journal Frontiers in Plant Science

RECEIVED 06 January 2023
ACCEPTED 03 February 2023
PUBLISHED 06 March 2023

CITATION
Liu L-C, Lv K-C and Zheng Y-J (2023) Crop cultivation planning with fuzzy estimation using water wave optimization. *Front. Plant Sci.* 14:1139094. doi: 10.3389/fpls.2023.1139094

Crop cultivation planning with fuzzy estimation using water wave optimization

Li-Chang Liu, Kang-Cong Lv and Yu-Jun Zheng*

School of Information Science and Technology, Hangzhou Normal University, Hangzhou, Zhejiang, China

In a complex agricultural region, determine the appropriate crop for each plot of land to maximize the expected total profit is the key problem in cultivation management. However, many factors such as cost, yield, and selling price are typically uncertain, which causes an exact programming method impractical. In this paper, we present a problem of crop cultivation planning, where the uncertain factors are estimated as fuzzy parameters. We adapt an efficient evolutionary algorithm, water wave optimization (WWO), to solve this problem, where each solution is evaluated based on three metrics including the expected, optimistic and pessimistic values, the combination of which enables the algorithm to search credible solutions under uncertain conditions. Test results on a set of agricultural regions in East China showed that the solutions of our fuzzy optimization approach obtained significantly higher profits than those of non-fuzzy optimization methods based on only the expected values.

KEYWORDS

crop cultivation planning, optimization, fuzzy parameters, evolutionary algorithms, water wave optimization (WWO).

1 Introduction

Many agricultural areas have complex and diverse topographic features (Rabia et al., 2022). In an area of one or several square kilometers, soil properties often change greatly, and different soil properties are suitable for different crops (Fu et al., 2023). The planning of crop cultivation in such a complex agricultural area needs to determine the appropriate crop for each plot of land to maximize the expected total profit, which is an important but difficult problem from an agricultural management point of view (Thilakarathne et al., 2023). The problem has to consider many factors including not only the topography and soil properties, but also the investment budget and cost of cultivation, expected yield of each plot, and selling price of each crop. However, the factors such as cost, yield, and selling price are typically uncertain and hard to estimate exactly. How to appropriately characterize these factors becomes a challenging task in the problem formulation.

In this paper, we present a problem of crop cultivation planning that aims to maximize the expected total profit under the constraint of investment budget and potential loss, where the uncertain factors are characterized as fuzzy parameters. Therefore, the problem

is formulated as a fuzzy optimization problem, which is much more complex than its crisp counterpart. Classical exact optimization approaches typically use methods such as expected functions and centroids to transform fuzzy values into crisp values, which inevitably lose important information contained in fuzzy parameters (Ekel et al., 1998; Liu, 2002; Zheng and Ling, 2013; Luhandjula, 2015).

To solve the problem more credibly by fully utilizing the information contained in fuzzy parameters, we propose an approach that evaluates the objective function with three metrics based on the expected, optimistic and pessimistic value models developed by Liu (2007). Based on the comprehensive fitness evaluation approach, we adapt an efficient evolutionary algorithm, water wave optimization (WWO) (Zheng, 2015) for the problem, which evolves the solutions to simultaneously improve the fitness in terms of the three related but different metrics. We conduct computational experiments on a variety of test instances constructed on agricultural regions in East China, and the results validate that the solutions obtained by the proposed WWO algorithm with fuzzy optimization obtain significantly higher profits than those of popular non-fuzzy evolutionary algorithms based on only the expected values. The main contributions of this paper can be summarized as follows:

- We present a crop cultivation planning problem that uses fuzzy parameters to characterize uncertain factors.
- We propose an adapted WWO algorithm to efficiently solve the fuzzy optimization problem.
- We validate the proposed method on a variety of test instances.

The remainder of the paper is organized as follows. Section 2 reviews the related work, Section 3 presents the crop cultivation planning problem, Section 4 describes the adapted WWO algorithm for the problem, Section 5 presents the test results, and Section 6 concludes with a discussion

2 Related work

Optimization models and algorithms have been widely used in agricultural planning. Zuo et al. (1991) studied a production planning problem for a large seed corn production company in North America in order to minimize the total cost by allocating the production of corn hybrids to different geographical areas; they developed a series of mathematical programming models and proposed a linear programming package and a mixed-integer programming package combined by a designed heuristic program to solve the problem. Sarker et al. (1997) presented a linear programming model for crop planning in Bangladesh that aims to maximize the overall contribution having satisfied the food demand, land availability, and capital constraints. Detlefsen and Jensen (2004) presented a decision support system, which calculates for each variety of winter wheat the expected net revenue as the expected gross revenue minus the expected costs for treatment of diseases and application of additional fertilization; the decision process was represented as a simple stochastic optimization model. Janová (2012) developed specific validation and verification procedures for the crop planning optimization models in agriculture when the randomness of harvests is considered and complex crop rotation restrictions must hold; the procedures were applied to stochastic programming model constructed as a decision support tool for crop plan optimization in South Moravian farm. López-Mata et al. (2016) developed a direct-solution algorithm capable of determining the crop ′ planning (area and volume of water per crop) that maximizes the profitability of an irrigation farm based on the data including the total cultivable area of the farm, the amount of available irrigation water, and the "gross margin vs. irrigation depth" functions of the considered crops. Esteso et al. (2022) presented a centralized multi-objective mathematical programming model to support the sustainable crop planning definition for a region that jointly optimize three objectives including supply chain profits maximization, waste minimization, and unfairness among farmers minimization; the multi-objective model was solved by applying the weighted sum method.

Except the simplest linear programming model, integer, mixed-integer, and multi-objective programming models are all NP-hard, for which exact optimization algorithms are applicable to only small- or medium size problem instances. Many recent efforts have been devoted to evolutionary algorithms for find near optimal or acceptable solutions to complex crop planning problems. Sarker and Ray (2009) formulated a crop-planning problem as a bi-objective optimization model that maximizes the total gross margin while minimizing the total working capital required; they solved two versions of the problem using multi objective evolutionary algorithms. Adeyemo et al. (2010) considered a multi-objective crop planning problem with three objectives including total net benefit maximization, agricultural output maximization, and total irrigation water minimization; they transformed the model into a single-objective one by taking the latter two as constraints, and then solved the single-objective optimization problem using differential evolution (DE) (Storn and Price, 1997). Márquez et al. (2011) modeled a water-saving crop planning problem as a multi-objective optimization problem that not only maximizes the economic benefits but also minimizes the water used; they solved the problem using two multi-objective evolutionary algorithms to search for Pareto-optimal solutions representing a trade-off between the two objectives. The water-saving crop planning problem considered by Wang et al. (2012) used four objective functions including maximum total net output, total grain yield, ecological benefits, and water productivity; they employed a multiple objective chaos particle swarm optimization (PSO) algorithm to solve the problem. Chetty and Adewumi (2014) compared a genetic algorithm (GA) and several swarm intelligence metaheuristics including cuckoo search, firefly algorithm, and glowworm swarm optimization, in solving an NP-hard annual crop planning problem. Zheng et al. (2013) studied a multiobjective oil crop fertilization problem, which takes into consideration not only crop yield and quality but also energy consumption and environmental effects; the authors proposed a hybrid multiobjective fireworks optimization algorithm that evolves

a set of solutions to the Pareto optimal front, using the concept of Pareto dominance for individual evaluation and selection. Fereidoon and Koch (2018) used a complex coupled simulation-optimization tool combining constrained PSO and LINGO-sub-optimization to solve crop planning in the Karkheh River Basin, Iran, under the impacts of climate change. Lin et al. (2021) presented a mathematical programming model for annual crop planning that allocates a land area for growing dryland and wetland crops to maximize the total profit and minimize the total irrigation water used for multiple cropping, and they proposed a simplified swarm optimization that improves PSO with four probabilities to determine the operations of updating solutions to effectively solve the problem.

In practice, agricultural systems are related to various uncertainty factors from the environment and market. However, only a few studies formulate these uncertainties into crop planning problems. Niu et al. (2016) developed an interactive two-stage fuzzy stochastic programming method for supporting crop planning and water resource allocation, where uncertainties are expressed as probability distributions and fuzzy-boundary intervals; the method enables decision makers to identify a trade-off between higher objective values and feasibility of constraints, and was applied to a real case of Hetao irrigation district in China. Alemany et al. (2021) developed a set of mathematical programming models to plan the planting and harvest of fresh tomatoes under a sustainable point of view for multi-farmer supply chains under uncertainty in different decision-making scenarios; for each distributed scenario, the individual solution per farmer as regards the planting and harvesting decisions per crop were integrated to obtain the overall supply to satisfy the markets demand. To the best of our knowledge, there are few studies conducted on evolutionary algorithms for solving large-size crop planning problems (typically of tens to hundreds of plots of lands and types of crops) with uncertain factors.

3 Problem description

3.1 Basic problem formulation

In the considered problem, we have an agricultural region that is divided into a set of m plots of lands, denoted by $\{P_1,P_2,\ldots,P_m\}$. The area of each plot P_i is a_i hectares; as the topographic conditions and soil properties inside a plot are homogeneous, each plot is allowed to cultivated with only one type of crop in the planning horizon (i.e., a particular season).

There are n types of candidate crops, denoted by $\{C_1,C_2,\ldots,C_n\}$. If plot P_i is cultivated with crop C_j, the basic investment is $\tilde{u}_{ij}$ (including investment for seeds, pesticides, fertilizers, irrigation, cultivation machines, etc.) per hectare, the expected yield is g_{ij} kg per hectare, and the cost for harvesting the crop is $_{ij}$ per kg. The expected selling price of crop C_j (after harvesting) is p_j per kg ($1\le i\le m;1\le j\le n$). The superscript ˜ indicates that due to uncertain conditions, the corresponding variable is difficult to determined exactly, and therefore is estimated as a fuzzy number.

The problem is to determine for each plot P_i the type of crop to be cultivated. Therefore, the decision variables can be expressed by an m-dimensional integer vector $\boldsymbol{x}=\{x_1,x_2,\ldots,x_m\}$, where x_i denotes the type of crop in P_i, i.e., P_i is cultivated with crop C_{x_i} ($1\le x_i\le n$).

Given a solution vector $\boldsymbol{x}=\{x_1,x_2,\ldots,x_m\}$, the expected overall revenue of the cultivation decision can be calculated as:

$$f(x)=\sum_{i=1}^{m}a_i(\tilde{g}_{i,x_i}(\tilde{p}_{x_i}-\tilde{v}_{i,x_i})-\tilde{u}_{i,x_i}) \tag{1}$$

The total budget is B, and the maximum loss that can be tolerated by the blueinvestor is L. Therefore, the budget constraint and loss constraint can be described as follows:

$$\sum_{i=1}^{m}a_i(\tilde{u}_{i,x_i}+\tilde{g}_{i,x_i}\tilde{v}_{i,x_i})\le B \tag{2}$$

$$\sum_{i=1}^{m}l(i,x_i)\le L \tag{3}$$

where $l(i,x_i)$ denotes the loss in plot P_i cultivated with crop C_{x_i}, which is calculated as:

$$l(i,x_i)=\begin{cases}\max\left(a_i\widetilde{u}_{i,x_i}-a_i\widetilde{g}_{i,x_i}(\widetilde{p}_{x_i}-\widetilde{v}_{i,x_i}),0\right), & \widetilde{p}_{x_i}\ge\widetilde{v}_{i,x_i}\\ a_i\widetilde{u}_{i,x_i}-a_i\widetilde{g}_{i,x_i}(\widetilde{v}_{i,x_i}-\widetilde{p}_{x_i}) & \widetilde{p}_{x_i}<\widetilde{v}_{i,x_i}\end{cases} \tag{4}$$

3.2 Evaluation of fuzzy parameters

If all input parameters are crisp values, the above formulation (1)–(4) can be regarded as an exact integer programming model. However, at the beginning of the planning horizon, some important parameters are difficult to estimated accurately. In this work, we express the investment $\tilde{u}_{ij}$, yield rate g_{ij}, and harvest cost $_{ij}$ as interval fuzzy numbers $[\underline{u}_{ij},\bar{u}_{ij}]$, $[\underline{g}_{ij},\bar{g}_{ij}]$, and $[\underline{linev}_{ij},\bar{v}_{ij}]$, respectively, where an underline denotes a lower limit and an overline denotes an upper limit ($1\le i\le m$, $1\le j\le n$); we express the expected selling price p_j as a Gaussian fuzzy number $N(\mu_j,\sigma_j)$, where μ_j is the mean value and σ_j is the deviation ($1\le j\le n$). Nevertheless, other types of fuzzy numbers (e.g., triangular and trapezoidal fuzzy numbers) are also allowable in the formulation of our fuzzy optimization problem. The fuzzy values can be estimated from historical data based on regression, fuzzy logic, and other machine learning methods that are capable of modeling uncertainty (Zheng et al., 2017a; Zheng et al., 2017b; Hernández and López, 2020; Gavahi et al., 2021).

As the budget constraint is a hard constraint, we use upper limits of investments and costs to transform the fuzzy constraint (2) as:

$$\sum_{i=1}^{m}a_i(\bar{u}_{i,x_i}+\bar{g}_{i,x_i}\bar{v}_{i,x_i})\le B \tag{5}$$

For the loss constraint, we evaluate the selling price of crop C_j as $p_j-3\sigma_j$ (the probability that the selling price is even smaller is less than 0.27% and is therefore negligible); moreover, if this selling price is larger than the harvest cost, we consider the lower limit of yield; otherwise, we consider the upper limit of yield; consequently,

equation (4) is transformed as:

$$l'(i,x_i)=\begin{cases}\max(a_i\overline{u}_{i,x_i}-a_i\underline{g}_{i,x_i}(\mu_{x_i}-2\sigma_{x_i}-\overline{v}_{i,x_i}),0), & \mu_{x_i}-2\sigma_{x_i}\geq\overline{v}_{i,x_i}\\ a_i\overline{u}_{i,x_i}-a_i\overline{g}_{i,x_i}(\overline{v}_{i,x_i}-\mu_{x_i}+2\sigma_{x_i}), & \mu_{x_i}-2\sigma_{x_i}<\overline{v}_{i,x_i}\end{cases} \tag{6}$$

Normally, the objective function (1) can be evaluated by using expected values μ_{x_i}, $u_{i,x_i}=(\underline{u}_{i,x_i}+\overline{u}_{i,x_i})/2$, $g_{i,x_i}=(\underline{g}_{i,x_i}+\overline{g}_{i,x_i})/2$, and $v_{i,x_i}=(\underline{v}_{i,x_i}+\overline{v}_{i,x_i})/2$ for fuzzy parameters:

$$E(x)=\sum_{i=1}^{m}a_i(g_{i,x_i}(\mu_{x_i}-v_{i,x_i})-u_{i,x_i}) \tag{7}$$

However, in practice, the expected value could deviate largely from the actual value. Therefore, we employ a credibility model (Liu, 2007) that calculates a credibility value $C_r(\xi)$ for a fuzzy variable ξ with membership function μ over the base set B of real numbers as follows:

$$C_r\{\xi\in B\}=(\sup_{\xi\in B}\mu(x)+1-\sup_{\xi\in B^C}\mu(x))/2 \tag{8}$$

Given a confidence $\theta\in(0,1]$, the θ-optimistic value and θ-pessimistic value of ξ are respectively defined as follows:

$$O(\xi,\theta)=\sup\{r\,|\,C_r\{\xi\geq r\}\geq\theta\} \tag{9}$$

$$P(\xi,\theta)=\inf\{r\,|\,C_r\{\xi\geq r\}\leq\theta\} \tag{10}$$

Based on the model, we also respectively evaluate an optimistic objective value and a pessimistic objective value as follows:

$$O(x,\theta)=\sum_{i=1}^{m}a_i(O(g_{i,x_i},\theta)(O(p_{x_i},\theta)-O(v_{i,x_i},\theta))-O(u_{i,x_i},\theta)) \tag{11}$$

$$P(x,\theta)=\sum_{i=1}^{m}a_i(P(g_{i,x_i},\theta)(P(p_{x_i},\theta)-P(v_{i,x_i},\theta))-P(u_{i,x_i},\theta)) \tag{12}$$

where the parameter θ is specified by the decision maker. The three objective values $\{E(\boldsymbol{x}), O(\boldsymbol{x},\theta), P(\boldsymbol{x},\theta)\}$ constitute a comprehensive evaluation of fitness of each solution $\boldsymbol{x}$ to the fuzzy optimization problem.

If a solution violates the constraints, we calculates the violation degree as:

$$v(x)=\max(\sum_{i=1}^{m}a_i(\overline{u}_{i,x_i}+\overline{g}_{i,x_i}\overline{v}_{i,x_i})-B,0)+\max(\sum_{i=1}^{m}l'(i,x_i)-L,0) \tag{13}$$

And then the objective function of the solution is added by a penalty of $Mv(\boldsymbol{x})$, where M is a large positive number.

4 Water wave optimization for the problem

4.1 Basic water wave optimization

WWO is a relatively new evolutionary algorithm inspired by the shallow water wave theory (Zheng, 2015), where the solution space is analogous to a seabed area, each solution $\boldsymbol{x}$ is analogous to a water wave associated with a wavelength bda_x, and the fitness of a solution is measured inversely according to its seabed depth. According to the shallow water wave theory, the shorter the distance between the seabed and the wave, the higher the wave height is and the smaller the wave length is, as illustrated in Figure 1.

At each iteration of the WWO algorithm, each wave $\boldsymbol{x}$ propagates in a range proportional to its wavelength, such that better solutions exploit smaller areas and worse solutions explore larger areas to balance between the local and global search to generate new solutions. In a high-dimensional continuous solution space, the propagation operation is executed by shifting each dimension i of the $\boldsymbol{x}$ as follows:

$$x_i=x_i+rand(-1,1)\cdot\lambda_x L_i \tag{14}$$

where *rand* is a function that generates a uniformly distributed random number within the specified range, and L_i is the length of the ith dimension of the solution space.

All wavelengths are initialized to 0.5 and then updated based on solution fitness at each iteration as follows:

$$\lambda_x=\lambda_x\alpha^{-(f(x)-f_{\min}+\epsilon)/(f_{\max}-f_{\min}+\epsilon)} \tag{15}$$

where $f_{\max}$ and $f_{\min}$ are the maximum and minimum fitness values among the population, respectively, ϵ is a very small value to avoid the zero-division error, and α is a parameter for wavelength reduction.

In addition to propagation, the basic WWO have two other operators: refraction and breaking. The refraction operator performs on any wave $\boldsymbol{x}$ that has not been improved after a certain number of generations by learning from the current best solution $\boldsymbol{x}^*$ at each dimension i as follows:

$$x_i'=N(\frac{x_i^*+x_i}{2},\frac{|x_i^*-x_i|}{2}) \tag{16}$$

where $N(\mu,\sigma)$ generates a Gaussian random number with mean μ and standard deviation σ. After refraction, its wavelength is updated according to the ratio between the new and original fitness values as:

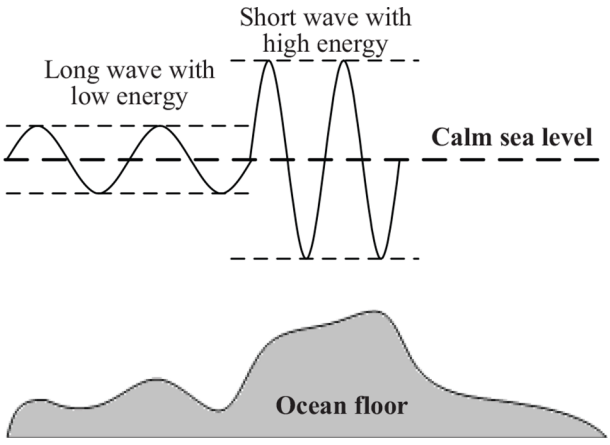


FIGURE 1
Illustration of wave heights and lengths in shallow water.

$$\lambda_{x'} = \lambda \frac{f(x)}{f(x')} \tag{17}$$

The breaking operator is used to split a newly found best wave $\boldsymbol{x}^*$ into a series of solitary waves, each of which is obtained by randomly selecting k dimensions (where k is a random number between 1 and a predefined upper limit $k_{\max}$) and at each dimension i updating the component as:

$$x_i' = x_i^* + rand(0,1) \cdot \beta L_i \tag{18}$$

where β is a parameter of breaking range. If the fittest one among the solitary waves is better than $\boldsymbol{x}^*$, it will replace $\boldsymbol{x}^*$ in the population.

In brief, propagation is the basic search mechanism for balancing global exploration and local exploitation, refraction helps stagnant waves to escape from local optima and improves the diversity of the population, while breaking further enhances the local search ability. The combination of these three operators makes WWO efficient in search in a high-dimensional solution space. Algorithm 18 presents the basic WWO algorithm framework.

The basic WWO algorithm is proposed for continuous optimization problems. Zheng et al. (2019) presented a systematic approach for adapting WWO to various combinatorial optimization problems. The key idea is to define a neighborhood search operation based on a neighborhood structure of the problem, and conduct the propagation on an solutions as a series of steps of neighborhood search, while the number of steps depends on the fitness or wavelength of the solution.

Since its proposal, WWO has attracted considerable attention in both academic and industrial communities. There have been a lot of work on modified WWO algorithms (Zheng and Zhang, 2015; Wu et al., 2017; Zhang et al., 2018; Zhang et al., 2019) and their applications to a variety of engineering optimization problems (Zheng et al., 2017c; Fard and Hajaghaei-Keshteli, 2018; Shao et al., 2018; Shao et al., 2019; Zhao et al., 2019; Zhou et al., 2019; Yan et al., 2021; Su et al., 2022; Zhang et al., 2022).

4.2 Adapted water wave optimization for fuzzy optimization

The basic WWO algorithm is for crisp optimization problems. To handle the presented fuzzy crop planning problem, we adapt the WWO in the following aspects.

- Instead of the single current best $\boldsymbol{x}^*$ in WWO, the adapted WWO keeps three current bests x_E^*, x_O^*, and x_P^* that have the best expected, optimistic, and pessimistic objective function values found so far, respectively.

```
1 Randomly initialize a population of NP solutions (waves);
2 Calculate the fitness of each solution, and let x* be the fittest one in the population;
3 while the termination condition is not satisfied do
4     foreach wave x in the population do
5         foreach dimension i do
6             Update xi according to Eq. (14);
7         Let x' be the propagated wave;
8         if f(x') > f(x) then
9             Replace x with x';
10            if f(x) > f(x*) then
11                Set x* to x;
12                for k = 1 to rand(1, k_max) do
13                    Select a random dimension i and create a solitary wave x' according to Eq. (18);
14                    if f(x') > f(x*) then
15                        Set x* to x';
16        else
17            if x has not been improved for h^max iterations then
18                Refract x to a new x' according to Eq. (16);
19            Update λ_x, according to Eq. (17);
20    Update the wavelengths of the solutions according to Eq. (15);
21 return the best wave found so far.
```

ALGORITHM 1
Basic WWO algorithm.

- At each iteration, let $E_{\max}$, $O_{\max}$, and $P_{\max}$ be the maximum expected optimistic, and pessimistic objective function values in the population, respectively, and $E_{\min}$, $O_{\min}$, and $P_{\min}$ be the corresponding minimum objective function values; for each solution $\boldsymbol{x}$ in the population, we select the maximum value among $(E(\boldsymbol{x})-E_{\min}+\epsilon)/(E_{\max}-E_{\min}+\epsilon)$, $(O(\boldsymbol{x},\boldsymbol{\theta})-O_{\min}+\epsilon)/(O_{\max}-O_{\min}+\epsilon)$, and $(P(\boldsymbol{x},\boldsymbol{\theta})-P_{\min}+\epsilon)/(P_{\max}-P_{\min}+\epsilon)$ as the exponent r, and update its wavelength as $\lambda_x=\lambda_x\alpha^{-r}$.
- When performing a propagation operation on a solution $\boldsymbol{x}$, each component x_i has a probability of λ_x of being changed to a new value x_i, which is determined by randomly selecting two other solutions, and then set to the corresponding component of the better one.
- A propagated solution $\boldsymbol{x}'$ will replace its original solution $\boldsymbol{x}$ if any of the following conditions is satisfied:

1. $\boldsymbol{x}'$ is a feasible solution, while $\boldsymbol{x}$ is an infeasible solution;
2. Both $\boldsymbol{x}'$ and $\boldsymbol{x}$ are feasible; $\boldsymbol{x}'$ is better than $\boldsymbol{x}$ in one objective functions and is not worse than $\boldsymbol{x}$ in either of the other two objective functions;
3. Both $\boldsymbol{x}'$ and $\boldsymbol{x}$ are feasible; $\boldsymbol{x}'$ is better than $\boldsymbol{x}$ in two or three objective functions.

- When performing a breaking operation on a solution $\boldsymbol{x}$, each solitary wave is obtained by selecting a random dimension i and setting x_i to a value, which, among all values in $[1,n]$, leads to the best improvement (in any of the three objective functions).
- The refraction operator is removed, and the removal of stagnant solutions is done by iteratively reducing the population size from an upper limit $N_P^{\max}$ to a lower limit $N_P^{\min}$, as suggested by Zheng and Zhang (2015).

In this way, the population evolves the solutions to improve the fitness in terms of the expected, optimistic, and pessimistic objective function values simultaneously. Finally, the three best solutions x_E^*, x_O^*, and x_P^* are returned to the decision maker for selection.

```
Randomly initialize a population of NP solutions (waves);
Calculate the fitness of each solution, and let x*_E, x*_O and x*_P be the solutions with the best expected, optimistic, and pessimistic objective function values, respectively;
while the termination condition is not satisfied do
    foreach wave x in the population do
        foreach dimension i do
            Update x_i according to Eq. (14);
        Let x' be the propagated wave;
        if x' is better than x in terms of the comprehensive comparison of the three objective functions then
            Replace x with x';
            if E(x) > E(x*) or O(x, θ) > O(x*, θ) or P(x, θ) > P(x*, θ) then
                Update the corresponding best solution;
                for k = 1 to rand(1, k_max) do
                    Select a random dimension i and create a solitary wave x' according to Eq. (18);
                    if x' leads to a new best solution then
                        Update the corresponding best solution;
            else
                if x has not been improved in any objective function for h^max iterations then
                    Refract x to a new x' according to Eq. (16);
                    Updateλx'accordingtoEq.(17);
    Update the wavelengths of the solutions;
return the best wave found so far.
```

ALGORITHM 2
WWO algorithm adapted for the fuzzy crop planning problem.

Compared to the basic WWO, the fuzzy WWO algorithm increases the time complexity in two aspects: (1) each solution is evaluated based on the three (related) objective functions; (2) the comparison of each pair of solutions is based on the three objective functions, at least once and at most three times. Consequently, the time complexity of the fuzzy WWO algorithm is at most triple that of the basic WWO.

5 Results

We applied the proposed algorithm to six selected agricultural regions in Zhejiang Province, East China. These regions were with different numbers of plots, crops, budgets, and allowable losses, as summarized in Table 1. The planning horizon was three months. The investment and cost are measured in RMB yuan.

After solving each problem instance, we presented the results to the decision-maker for selection, and obtained the actual total profit of the crop cultivation after the planning horizon. For comparison, we also implemented three evolutionary algorithms, including differential evolution (DE) (Omran and Engelbrecht, 2007), biogeography-based optimization (BBO) (Simon, 2008; Wang and Wu, 2014), and the basic WWO, to solve the crop planning problem by only maximizing the expected objective function (7). We executed each algorithm for 20 runs, and take the best solution among the 20 runs. The profit of each solution is evaluated based on the expected yields of the solution and actual costs and prices at the end of the planning horizon.

Table 2 presents the profits of the solutions obtained by the different algorithms on the instances, which are also compared in Figure 2. On instance 1, WWO solution obtained the maximum

TABLE 1 Basic information of the six agricultural regions for the applications of the proposed algorithm.

#Region	Area (hectares)	m (plots)	n (crops)	B (RMB Yuan)	L (RMB Yuan)
1	471	27	36	550,000	100,000
2	665	39	33	720,000	180,000
3	729	49	39	800,000	160,000
4	1202	76	36	1,200,000	300,000
5	1530	93	41	1,800,000	450,000
6	2808	121	39	2,400,000	500,000

TABLE 2 Profits of the solutions obtained by the three non-fuzzy evolutionary algorithms and the proposed WWO with fuzzy optimization (WWO-F).

#Region	DE	BBO	WWO	WWO-F
1	191275	184320	196009	207355
2	231060	208040	229750	250700
3	244960	231000	242580	252600
4	v408100	379100	391600	415200
5	357750	339120	362590	378100
6	560800	538800	557900	590200
Total	1993945	1880380	1980429	2094155

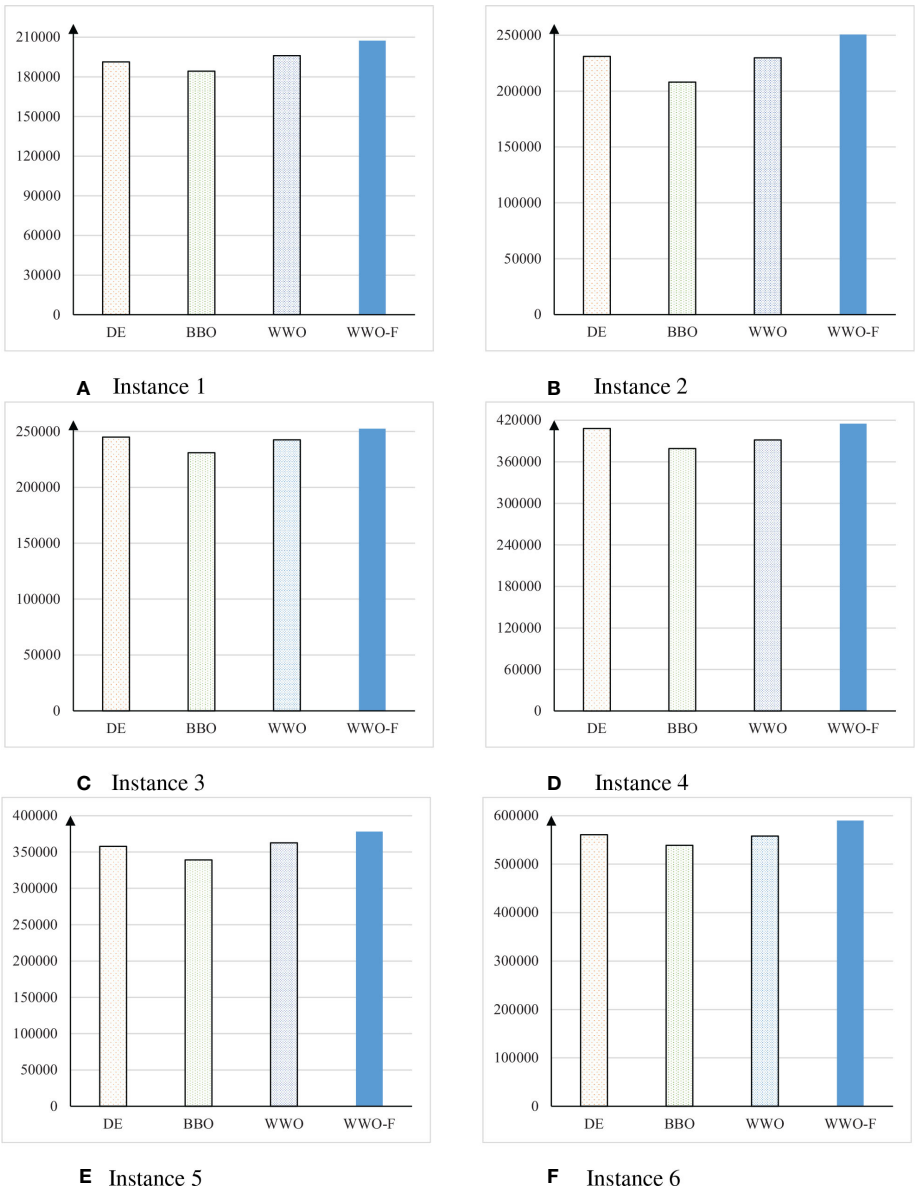


FIGURE 2
Comparison of the profits of the solutions obtained by the three non-fuzzy evolutionary algorithms and the proposed WWO with fuzzy optimization (WWO-F) on the six test instances **(A–F)**.

profit of 196,009 among the three non-fuzzy evolutionary algorithms, while WWO-F solution obtained a profit of 207,355, which was 11,346 more than maximum profit of the non-fuzzy algorithms. On instance 2, DE solution obtained the maximum profit of 231,060 among the three non-fuzzy evolutionary algorithms, while WWO-F solution obtained a profit of 250,700, which was 19,640 more than that of DE solution. On instance 3, DE solution obtained the maximum profit of 244,960 among the three non-fuzzy evolutionary algorithms, which was 7,640 less than the profit of 252,600 obtained by the WWO-F solution. On instance 4, DE solution obtained the maximum profit of 408,100 among the three non-fuzzy evolutionary algorithms, which was 7,100 less than the profit of 415,200 obtained by the WWO-F solution. On instance5, WWO solution obtained the maximum profit of 362,590 among the three non-fuzzy evolutionary algorithms, which was 15,510 less than the profit of 378,100 obtained by the WWO-F solution. On instance 6, DE solution obtained the maximum profit of 560,800 among the three non-fuzzy evolutionary algorithms, which was 29,400 less than the profit of 590,200 obtained by the WWO-F solution. The results show that, on all six instances, WWO-F always obtained a better profit than the non-fuzzy evolutionary algorithms. This is because the non-fuzzy evolutionary algorithms use only the expected objective function to evaluate the solution fitness; however, the estimated cost and selling price could deviate from the actual values, and hence a solution for maximizing the expected objective function often failed to fully utilize the budget to pursue the maximum profit. By simultaneously using the three criteria including expected, optimistic, and pessimistic values, WWO-F utilized the information contained in the fuzzy parameters much better than the non-fuzzy algorithms, evolved the solutions to keep a good trade-off between the overestimation of the profit and underestimation of the costs, and hence obtained solutions that are more robust and credible.

The last row of Table 2 summarizes the total profits of the algorithms on the six instances. DE obtained the maximum total profit of 1,993,945 among the three non-fuzzy evolutionary algorithms, while WWO-F obtained a total profit of 2,094,155, which was 102,210 more than the DE solution. In summary, the proposed fuzzy optimization approach obtained an over five percent increase over the best non-fuzzy algorithm. This result demonstrated the significant economic benefits brought by the application of our fuzzy optimization approach for crop planning.

6 Conclusion

This paper presents a crop cultivation planning problem with fuzzy parameters (including cost, yield, and selling price) for maximizing the expected total profit under the constraint of investment budget and potential loss. To fully utilize the information contained in fuzzy parameters, we evaluate the objective function with three metrics based on the expected, optimistic and pessimistic value models, and propose an adapted WWO algorithm that evolves the solutions to simultaneously improve the fitness in terms of the three related but different metrics. Results on a variety of test instances constructed on agricultural regions in East China validated that the solution of the proposed WWO algorithm with fuzzy optimization obtained an over five percent increase on the total profit over the best non-fuzzy algorithm.

The current work studies crop planning in a particular season. Currently, we are extending the fuzzy optimization problem and algorithm for annual crop planning, which involves cultivating and harvesting multiple crops with different seasonal lengths in a plot. Moreover, in the current study, the fuzzy parameters are mainly estimated based on experience or simple regression on historical data; in future work, we will estimate the parameters from big data, using fuzzy deep learning to discover highly nonlinear relationship with complex factors (Song et al., 2019; Elavarasan and Durai Raj Vincent, 2021) and employing transfer learning to utilize knowledge in similar domains to cope with the insufficiency of labeled data (Song et al., 2021; Song et al., 2022; Zheng et al., 2022) in a more comprehensive manner.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: http://compintell.cn/en/dataAndCode.html.

Author contributions

L-CL: Investigation, Software, Writing – Original draft preparation. K-CL: Data Curation, Validation. Y-JZ: Conceptualization, Methodology, Funding acquisition. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by National Natural Science Foundation of China under Grant 61872123 and Zhejiang Provincial Natural Science Foundation under Grant LR20F030002.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Adeyemo, J., Bux, F., and Otieno, F. (2010). Differential evolution algorithm for crop planning:single and multi-objective optimization model. *Int. J. Phys. Sci.* 5, 1592–1599. doi: 10.5897/IJPS.9000175

Alemany, M., Esteso, A., Ortiz, Á., and del Pino, M. (2021). Centralized and distributed optimization models for the multi-farmer crop planning problem under uncertainty: Application to a fresh tomato argentinean supply chain case study. *Comput. Ind. Eng.* 153, 107048. doi: 10.1016/j.cie.2020.107048

Chetty, S., and Adewumi, A. O. (2014). Comparison study of swarm intelligence techniques for the annual crop planning problem. *IEEE Trans. Evol. Comput.* 18, 258–268. doi: 10.1109/TEVC.2013.2256427

Detlefsen, N. K., and Jensen, A. L. (2004). A stochastic model for crop variety selection. *Agric. Syst.* 81, 55–72. doi: 10.1016/j.agsy.2003.08.004

Ekel, P., Pedrycz, W., and Schinzinger, R. (1998). A general approach to solving a wide class of fuzzy optimization problems. *Fuzzy Sets Syst.* 97, 49–66. doi: 10.1016/S0165-0114(96)00334-X

Elavarasan, D., and Durai Raj Vincent, P. M. (2021). Fuzzy deep learning-based crop yield prediction model for sustainable agronomical frameworks. *Neural Comput. Appl.* 33, 13205–13224. doi: 10.1007/s00521-021-05950-7

Esteso, A., Alemany, M. M. E., Ortiz, A., and Liu, S. (2022). Optimization model to support sustainable crop planning for reducing unfairness among farmers. *Centr. Euro. J. Oper. Res.* 30, 1101–1127. doi: 10.1007/s10100-021-00751-8

Fard, A. M. F., and Hajaghaei-Keshteli, M. (2018). A tri-level location-allocation model for forward/reverse supply chain. *Appl. Soft. Comput.* 62, 328–346. doi: 10.1016/j.asoc.2017.11.004

Fereidoon, M., and Koch, M. (2018). SWAT-MODSIM-PSO optimization of multi-crop planning in the karkheh river basin, iran, under the impacts of climate change. *Sci. Total Environ.* 630, 502–516. doi: 10.1016/j.scitotenv.2018.02.234

Fu, H., Chen, H., Ma, Q., Han, K., Wu, S., and Wu, L. (2023). Effect of planting and mowing cover crops as livestock feed on soil quality and pear production. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.1105308

Gavahi, K., Abbaszadeh, P., and Moradkhani, H. (2021). DeepYield: A combined convolutional neural network with long short-term memory for crop yield forecasting. *Expert Syst. Appl.* 184, 115511. doi: 10.1016/j.eswa.2021.115511

Hernández, S., and López, J. L. (2020). Uncertainty quantification for plant disease detection using Bayesian deep learning. *Appl. Soft. Comput.* 96, 106597. doi: 10.1016/j.asoc.2020.106597

Janová, J. (2012). Crop planning optimization model: the validation and verification processes. *Centr. Euro. J. Oper. Res.* 20, 451–462. doi: 10.1007/s10100-011-0205-8

Lin, C.-C., Deng, D.-J., Kang, J.-R., and Liu, W.-Y. (2021). A dynamical simplified swarm optimization algorithm for the multiobjective annual crop planning problem conserving groundwater for sustainability. *IEEE Trans. Ind. Informat.* 17, 4401–4410. doi: 10.1109/TII.2020.3029258

Liu, B. (2002). Toward fuzzy optimization without mathematical ambiguity. *Fuzzy Optim. Decision Making* 1, 43–63. doi: 10.1023/A:1013771608623

Liu, B. (2007). *Uncertainty theory. 2nd edn* (Berlin: Springer-Verlag).

López-Mata, E., Orengo-Valverde, J., Tarjuelo, J., Martínez-Romero, A., and Domínguez, A. (2016). Development of a direct-solution algorithm for determining the optimal crop planning of farms using deficit irrigation. *Agricultural. Water Manage.* 171, 173–187. doi: 10.1016/j.agwat.2016.03.015

Luhandjula, M. (2015). Fuzzy optimization: Milestones and perspectives. *Fuzzy Sets Syst.* 274, 4–11. doi: 10.1016/j.fss.2014.01.004

Márquez, A. L., Baños, R., Gil, C., Montoya, M. G., Manzano-Agugliaro, F., and Montoya, F. G. (2011). Multi-objective crop planning using pareto-based evolutionary algorithms. *Agric. Econom.* 42, 649–656. doi: 10.1111/j.1574-0862.2011.00546.x

Niu, G., Li, Y., Huang, G., Liu, J., and Fan, Y. (2016). Crop planning and water resource allocation for sustainable development of an irrigation region in china under multiple uncertainties. *Agric. Water Manage.* 166, 53–69. doi: 10.1016/j.agwat.2015.12.011

Omran, M., and Engelbrecht, A. (2007). Differential evolution for integer programming problems. *IEEE Congress Evol. Computation*, 2237–2242. doi: 10.1109/CEC.2007.4424749

Rabia, A. H., Neupane, J., Lin, Z., Lewis, K., Cao, G., and Guo, W. (2022). "Principles and applications of topography in precision agriculture," in *Advances in agronomy, chap. 4*, vol. 171 . Ed. D. L. Sparks (Cambridge, MA: Academic Press) 143–189. doi: 10.1016/bs.agron.2021.08.005

Sarker, R., and Ray, T. (2009). An improved evolutionary algorithm for solving multi-objective crop planning models. *Comput. Electr. Agricul.* 68, 191–199. doi: 10.1016/j.compag.2009.06.002

Sarker, R. A., Talukdar, S., and Haque, A. (1997). Determination of optimum crop mix for crop cultivation in bangladesh. *Appl. Math. Model.* 21, 621–632. doi: 10.1016/S0307-904X(97)00083-8

Shao, Z., Pi, D., and Shao, W. (2018). A novel discrete water wave optimization algorithm for blocking flow-shop scheduling problem with sequence-dependent setup times. *Swarm Evol. Comput.* 40, 53–75. doi: 10.1016/j.swevo.2017.12.005

Shao, Z., Pi, D., and Shao, W. (2019). A novel multi-objective discrete water wave optimization for solving multi-objective blocking flow-shop scheduling problem. *Knowl. Based Syst.* 165, 110–131. doi: 10.1016/j.knosys.2018.11.021

Simon, D. (2008). Biogeography-based optimization. *IEEE Trans. Evol. Comput.* 12, 702–713. doi: 10.1109/TEVC.2008.919004

Song, Q., Zheng, Y., Huang, Y., Xu, Z., Sheng, W., and Yang, J. (2019). Emergency drug procurement planning based on big-data driven morbidity prediction. *IEEE Trans. Ind. Informat.* 15, 6379–6388. doi: 10.1109/TII.2018.2870879

Song, Q., Zheng, Y.-J., Sheng, W.-G., and Yang, J. (2021). Tridirectional transfer learning for predicting gastric cancer morbidity. *IEEE Trans. Neural Netw. Learn. Syst.* 32, 561–574. doi: 10.1109/TNNLS.2020.2979486

Song, Q., Zheng, Y.-J., Yang, J., Huang, Y.-J., Sheng, W.-G., and Chen, S.-Y. (2022). Predicting demands of COVID-19 prevention and control materials *via* co-evolutionary transfer learning. *IEEE Trans. Cybern.*, 1–14. doi: 10.1109/TCYB.2022.3164412

Storn, R., and Price, K. (1997). Differential evolution - a simple and efficient heuristic for global optimization over continuous spaces. *J. Global Optim.* 11, 341–359. doi: 10.1023/A:1008202821328

Su, Z.-L., Jiang, X.-L., Li, N., Ling, H.-F., and Zheng, Y.-J. (2022).). optimization of false target jamming against UAV detection. *Drones* 6, 114. doi: 10.3390/drones6050114

Thilakarathne, N. N., Bakar, M. S. A., Abas, P. E., and Yassin, H. (2023). Towards making the fields talks: A real-time cloud enabled iot crop management platform for smart agriculture. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.1030168

Wang, Z.-C., and Wu, X.-B. (2014). Hybrid biogeography-based optimization for integer programming. *Sci. World J.* 2014, 9. doi: 10.1155/2014/672983

Wang, Y., Wu, P., Zhao, X., and Jin, J. (2012). Water-saving crop planning using multiple objective chaos particle swarm optimization for sustainable agricultural and soil resources development. *CLEAN – Soil Air Water* 40, 1376–1384. doi: 10.1002/clen.201100310

Wu, X., Zhou, Y., and Lu, Y. (2017). Elite opposition-based water wave optimization algorithm for global optimization. *Math. Prob. Eng.* 2017, 25. doi: 10.1155/2017/3498363

Yan, Z., Zhang, J., Zeng, J., and Tang, J. (2021). Water wave optimization algorithm for autonomous underwater vehicle path planning problem. *J. Intell. Fuzzy Syst.* 40, 9127–9141. doi: 10.3233/JIFS-201544

Zhang, M.-X., Wu, J.-Y., Wu, X., and Zheng, Y.-J. (2022). Hybrid evolutionary optimization for takeaway order selection and delivery path planning utilizing habit data. *Complex Intell. Syst.* 8, 4425–4440. doi: 10.1007/s40747-021-00410-0

Zhang, J., Zhou, Y., and Luo, Q. (2018). An improved sine cosine water wave optimization algorithm for global optimization. *J. Intell. Fuzzy Syst.* 34, 2129–2141. doi: 10.3233/JIFS-171001

Zhang, J., Zhou, Y., and Luo, Q. (2019). Nature-inspired approach: a wind-driven water wave optimization algorithm. *Appl. Intell.* 49, 233–252. doi: 10.1007/s10489-018-1265-4

Zhao, F., Zhang, L., Liu, H., Zhang, Y., Ma, W., Zhang, C., et al. (2019). An improved water wave optimization algorithm with the single wave mechanism for the no-wait flow-shop scheduling problem. *Eng. Optim.* 51, 1727–1742. doi: 10.1080/0305215X.2018.1542693

Zheng, Y.-J. (2015). Water wave optimization: A new nature-inspired metaheuristic. *Comput. Oper. Res.* 55, 1–11. doi: 10.1016/j.cor.2014.10.008

Zheng, Y.-J., Chen, S.-Y., Xue, Y., and Xue, J.-Y. (2017a). A Pythagorean-type fuzzy deep denoising autoencoder for industrial accident early warning. *IEEE Trans. Fuzzy Syst.* 25, 1561–1575. doi: 10.1109/TFUZZ.2017.2738605

Zheng, Y.-J., and Ling, H.-F. (2013). Emergency transportation planning in disaster relief supply chain management: A cooperative fuzzy optimization approach. *Soft. Comput.* 17, 1301–1314. doi: 10.1007/s00500-012-0968-4

Zheng, Y.-J., Lu, X.-Q., Du, Y.-C., Xue, Y., and Sheng, W.-G. (2019). Water wave optimization for combinatorial optimization: Design strategies and applications. *Appl. Soft. Comput.* 83, 105611. doi: 10.1016/j.asoc.2019.105611

Zheng, Y. J., Sheng, W. G., Sun, X. M., and Chen, S. Y. (2017b). Airline passenger profiling based on fuzzy deep machine learning. *IEEE Trans. Neural Netw. Learn. Syst.* 28, 2911–2923. doi: 10.1109/TNNLS.2016.2609437

Zheng, Y.-J., Song, Q., and Chen, S.-Y. (2013). Multiobjective fireworks optimization for variable-rate fertilization in oil crop production. *Appl. Soft. Comput.* 13, 4253–4263. doi: 10.1016/j.asoc.2013.07.004

Zheng, Y.-J., Wang, Y., Ling, H.-F., Xue, Y., and Chen, S.-Y. (2017c). Integrated civilian-military pre-positioning of emergency supplies: A multiobjective optimization approach. *Appl. Soft. Comput.* 58, 732–741. doi: 10.1016/j.asoc.2017.05.016

Zheng, Y.-J., Yu, S.-L., Song, Q., Huang, Y.-J., Sheng, W.-G., and Chen, S. (2022). Co-Evolutionary fuzzy deep transfer learning for disaster relief demand forecasting. *IEEE Trans. Emerg. Topics Comput.* 10, 1361–1373. doi: 10.1109/TETC.2021.3085337

Zheng, Y.-J., and Zhang, B. (2015). A simplified water wave optimization algorithm. *IEEE Congress Evol. Computation*, 807–813. doi: 10.1109/CEC.2015.7256974

Zhou, X.-H., Zhang, M.-X., Xu, Z.-G., Cai, C.-Y., Huang, Y.-J., and Zheng, Y.-J. (2019). Shallow and deep neural network training by water wave optimization. *Swarm Evol. Comput.* 50, 1–13. doi: 10.1016/j.swevo.2019.100561

Zuo, M., Kuo, W., and McRoberts, K. L. (1991). Application of mathematical programming to a large-scale agricultural production and distribution system. *J. Oper. Res. Soc.* 42, 639–648. doi: 10.1057/jors.1991.130

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 28 March 2023
DOI 10.3389/fpls.2023.1154176

OPEN ACCESS

EDITED BY
Yu Xue,
Nanjing University of Information Science and Technology, China

REVIEWED BY
Romany Mansour,
The New Valley University, Egypt
Peisen Yuan,
Nanjing Agricultural University, China

*CORRESPONDENCE
Fuquan Zhang
zfq@mju.edu.cn

SPECIALTY SECTION
This article was submitted to Sustainable and Intelligent Phytoprotection, a section of the journal Frontiers in Plant Science

RECEIVED 30 January 2023
ACCEPTED 28 February 2023
PUBLISHED 28 March 2023

CITATION
Chen Z, Wang C, Zhang F, Zhang L, Grau A and Guerra E (2023) All-in-one aerial image enhancement network for forest scenes. *Front. Plant Sci.* 14:1154176. doi: 10.3389/fpls.2023.1154176

All-in-one aerial image enhancement network for forest scenes

Zhaoqi Chen[1,2], Chuansheng Wang[3], Fuquan Zhang[2,4,5,6,7]*, Ling Zhang[4], Antoni Grau[3] and Edmundo Guerra[3]

[1]College of Computer and Big Data, Fuzhou University, Fuzhou, China, [2]Fujian Provincial Key Laboratory of Information Processing and Intelligent Control, Minjiang University, Fuzhou, China, [3]Department of Automatic Control, Polytechnic University of Catalonia, Barcelona, Spain, [4]College of Computer and Control Engineering, Minjiang University, Fuzhou, China, [5]Digital Media Art, Key Laboratory of Sichuan Province, Sichuan Conservatory of Music, Chengdu, China, [6]Fuzhou Technology Innovation Center of Intelligent Manufacturing information System, Minjiang University, Fuzhou, China, [7]Engineering Research Center for Intangible Cultural Heritage (ICH) Digitalization and Multi-source Information Fusion (Fujian Polytechnic Normal University), Fujian Province University, Fuzhou, China

Drone monitoring plays an irreplaceable and significant role in forest firefighting due to its characteristics of wide-range observation and real-time messaging. However, aerial images are often susceptible to different degradation problems before performing high-level visual tasks including but not limited to smoke detection, fire classification, and regional localization. Recently, the majority of image enhancement methods are centered around particular types of degradation, necessitating the memory unit to accommodate different models for distinct scenarios in practical applications. Furthermore, such a paradigm requires wasted computational and storage resources to determine the type of degradation, making it difficult to meet the real-time and lightweight requirements of real-world scenarios. In this paper, we propose an All-in-one Image Enhancement Network (AIENet) that can restore various degraded images in one network. Specifically, we design a new multi-scale receptive field image enhancement block, which can better reconstruct high-resolution details of target regions of different sizes. In particular, this plug-and-play module enables it to be embedded in any learning-based model. And it has better flexibility and generalization in practical applications. This paper takes three challenging image enhancement tasks encountered in drone monitoring as examples, whereby we conduct task-specific and all-in-one image enhancement experiments on a synthetic forest dataset. The results show that the proposed AIENet outperforms the state-of-the-art image enhancement algorithms quantitatively and qualitatively. Furthermore, extra experiments on high-level vision detection also show the promising performance of our method compared with some recent baselines.

KEYWORDS

image enhancement, all-in-one network, multi-receptive fields, drone image monitoring, forest protection, smoke detection

1 Introduction

Drone aerial image technology plays an indispensable role in forest fire monitoring. However, the images captured by drones are severely damaged because of the uncertainty and instability of aerial photography. Typical examples of aerial image degradation include atmospheric interference and motion blur caused by the vibration of the drone. Moreover, the aerial images could further suffer from the visual impact of compression when the images are transmitted back through the network. Therefore, how to restore degraded aerial images is particularly significant under the limitation of existing hardware. Recently, with the development of deep learning, data-driven methods designed for task-specific image enhancement have achieved great success, such as image dehazing (Ren et al., 2018; Qu et al., 2019; Wang et al., 2020; Song et al., 2022), image denoising (Zhang et al., 2017b; Ct et al., 2020), and image deblurring (Nah et al., 2017; Gao et al., 2019). However, an all-in-one image enhancement model seems more effective than its specific-task counterpart in practical application scenarios as real-world images usually suffer various degradations. For example, images of forest scenes collected by drones could be affected by adverse weather or blurred by remote sensor shaking. In contrast, integrating multiple image enhancement tasks in an all-in-one framework is a promising choice.

Recently, Li et al. proposed an all-in-one method, which uses a multi-encoder and single-decoder architecture to address various weather corruptions (Li et al., 2017). It also utilizes the neural architecture search to optimize the features extracted by the encoder, which performs better than previous task-specific image enhancement algorithms. But, designing such an architecture usually comes at the expense of computational costs. Due to its success in high-level tasks such as image classification, segmentation, and detection, the transformer has been used in low-level vision tasks. Valanarasu et al. proposed Transweather, an end-to-end multi-weather image restoration model, as an alternative solution to multi-encoders for the same application scenario (Valanarasu et al., 2022). Li et al. also proposed a unified framework capable of recovering images with unknown degradation types, which has demonstrated its effectiveness in image enhancement affected by natural weather (Li et al., 2022). Although the generalization performance of the network has been verified on multiple datasets, it has low practical application value due to its large number of parameters and computational delays. Moreover, nearly all of the representative models for aerial image enhancement are based on single-task design (Wang and Liu, 2022). Therefore, research on an all-in-one framework is still very necessary in this field.

We believe that the future development of aerial image enhancement research lies in all-in-one models, which is also a critical step toward general technology research. The motivation for this paper is two-fold: on the one hand, we wish to conduct an in-depth study on preserving the high-dimensional detail features of multi-scale objects, thus pushing the aerial image reconstruction methods to a new level. On the other hand, the all-in-one network can be utilized to study general strategies for a seamless transition between different tasks and domains. As shown in Figure 1, to this end, we propose an All-in-one Image Enhancement Network (AIENet) based on a Multi-Receptive Field (MRF) enhancement block. Specifically, the model only performs one downsampling operation on the original image. And the global skip connection is used to introduce the low-level feature information of the corresponding scale into the deconvolution process so that the model can obtain more high-resolution details during upsampling. In addition, with the multi-receptive field enhancement module, the model can fully use the prior hierarchical features on the same scale to explore different regions and then obtain the global context by aggregating the context information collected from different areas.

The main contributions of this work include the following:

- By comprehensively analyzing the characteristics of aerial imagery, we identify the importance of all-in-one models for forest scenarios. Furthermore, we accurately reconstruct local textures and microstructures in degraded images by maximizing the feature representation and learning capabilities of neural networks, thereby improving the accuracy of subsequent high-level computer vision tasks.
- We propose a lightweight image enhancement model AIENet, which can quickly solve the degradation problem in an all-in-one framework when collecting images. The proposed method utilizes the global and local skip connections to introduce high-resolution details into the output image. And the model designed in this paper cleverly uses the multi-receptive field fusion technique to perceive the same feature map from multiple scales, thus making up for the insufficient ability to capture global image features.
- We demonstrate that our method can achieve better visual performance and high scores based on the quantitative

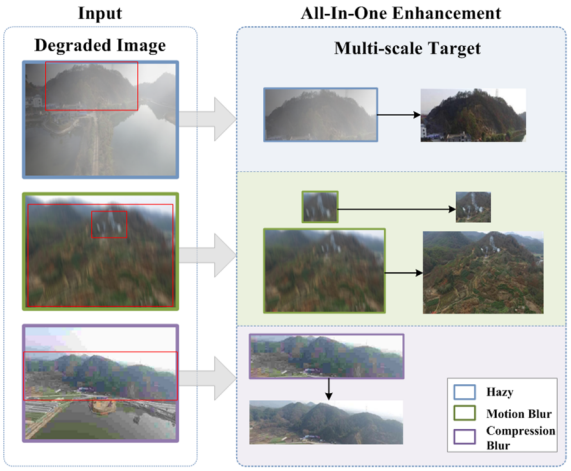


FIGURE 1
An illustration of our motivation. As shown, the forest scene images captured by drones could inevitably suffer from different degradation problems. The target size in the images captured at different locations is variable. Therefore, we explored the restoration of multi-scale target details in aerial images and proposed an all-in-one image enhancement method.

metrics in task-specific and all-in-one aerial image restoration on the forest wildfire dataset. The model proposed in this paper also provides an idea of integrated processing for restoring the visual quality of images with complex scenes.

The remainder of this paper is organized as follows. Section 2 discusses related work on image enhancement and multi-receptive field technology. The proposed method is introduced in Section 3. Section 4 analyzes the comparative experimental results, and our work is concluded in Section 5.

2 Related works

With the popularity of graphics processing units, the deep learning approaches (Alsubai et al., 2022; Farghaly et al., 2022) have developed the most advanced model in the computer vision field, and numerous elegant solutions (Xue et al., 2019; Xue et al., 2021) have been proposed for visual tasks in the last few years. In the field of image enhancement, most researchers work on task-specific image restoration. In this paper, we innovatively propose an all-in-one architecture to solve the image degradation problems encountered in various stages of aerial image acquisition, such as haze weather interference, the vibration of the remote sensing platform during shooting, and image compression during transmission distortion. Therefore, we first describe representative methods for each task. Then, we introduce related work on multi-scale receptive fields in low-level vision.

2.1 Image enhancement

Image Dehazing: Since Mccartney et al. proposed the atmospheric scattering model to approximate the haze effect which is shown as: $\hat{x} = x \odot t + A \odot (1 - t)$, where $\hat{x}$ and x mean the degradative images and restored images, respectively. t is the transmission map, which can express as: $t = e^{-\beta d}$, where β and d are the scattering coefficient and depth map (McCartney, 1976). $\boldsymbol{A}$ is the global atmospheric light and $\odot$ represented as pixel-wise multiplication. Recent image dehazing methods could be classified into two families, *i.e.*, prior-based methods and learning-based methods. In traditional prior-based methods, many image statistical priors are used as additional constraints to compensate for information loss during image degradation. He et al. proposed a classic image dehazing method that depends on the statistical results called Dark Channel Prior (DCP), which generates at least one low-intensity pixel in the color channel of each pixel local neighborhood (He et al., 2010). Then the learned transmission map is used to calculate the haze-free image through the physical model. Wang et al. found that the blurred areas are mainly concentrated on the brightness channel of the YCrCb color space (Wang et al., 2018) Therefore, it is possible to enhance the visual contrast of foggy scenes by recovering the missing texture information in the luminance channel. As for learning-based methods, the techniques such as attention (Liu et al., 2019; Zhang et al., 2020), feature fusion (Dong et al., 2020; Qin et al., 2020) and contrastive learning (Wu et al., 2021; Chen et al., 2022) are widely used to improve single-image dehazing performance. Moreover, they outperform the traditional prior-based image dehazing methods.

Motion Deblurring: Since large-scale real-world blur data is challenging to obtain, most traditional deblur methods are generally tested on synthetic images from $\hat{x}$ to x, which can be expressed as $\hat{x} = x \otimes k + n$, where $\hat{x}$ is the blurred image generated from clean image x, k is the blur kernel or convolution kernel, $\otimes$ denotes the convolution operator and n is additive noise. However, handcrafted methods are not good at capturing complex blur variations in authentic images. In contrast, CNN-based methods can handle real-world blurry images well if we have a dataset of paired images. Tao et al. proposed a multi-scale approach based on encoder-decoder recurrent networks (SRN), which is the first method to integrate recurrent neural networks (RNN) into deblurring models (Tao et al., 2018). Some methods (Kupyn et al., 2018; Kupyn et al., 2019) based on Generative Adversarial Networks (GAN) have also achieved competitive results on real-world deblur. Recently, multi-stage architecture networks (Chen et al., 2021; Zamir et al., 2021) have achieved state-of-the-art results in deblurring restoration tasks.

Compression Deblurring: Early image compression restoration methods use deblocking filters to reduce discontinuities between pixel blocks. To reduce blocking artifacts in compressed images, Lee et al. adaptively use various block predictors based on frequency components in the Discrete Cosine Transform (DCT) domain (Lee et al., 2004). Yoo et al. classifies blocks as flat or edge blocks and applies different deblocking filters depending on the classification result (Yoo et al., 2014). However, these methods employing deblocking filters only target blocking artifacts. But also other artifacts in compressed images, such as ringing artifacts. Therefore, most scholars have conducted extensive research on CNN-based compression deblur. Dong et al. introduce a super-resolution convolutional neural network for reducing compressed image artifacts (Dong et al., 2015). Zhang et al. use auto-encoders in both DCT and pixel domains, considering the output of auto-encoders and input images to reduce visual artifacts in compressed images (Zhang et al., 2018). Lee et al. utilize parallel atrous convolution residual blocks to extract a variety of features with large receptive fields, then use attention mechanism for the output of atrous convolution to obtain representations of the global region (Lee et al., 2021).

All-in-One Image Enhancement: Although the above image enhancement methods all perform well on specific tasks, real-world images are often easily corrupted by different degradation types, making task-specific image enhancement lack flexibility and generalization in practical applications. Recently, some work has focused on all-in-one visual enhancement networks. To deal with image degradation under severe weather conditions (such as rain, haze, and snow), Li. et al. present an ensemble model based on neural architecture search, whose generator has a multi-encoder and a typical decoder architecture (Li et al., 2020). In other words, the network must train different models for different degradation problems, which is unsuitable for an all-in-one solution in practical applications. Most recently, Valanarasu. et al. propose an alternative

state-of-the-art solution to this problem with TransWeather (Valanarasu et al., 2022). As an end-to-end vision transformer (Dosovitskiy et al., 2021) based multi-weather image restoration model, it exhibits more powerful versatility. Notably, these two all-in-one image enhancement methods focus on recovering the same combination of degradation types (i.e., weather disturbances). However, solving the image degradation problem under multiple conditions (such as weather, physical factors, etc.) in an all-in-one framework can better meet the most practical scenes. In addition, the real-time requirement of the remote sensing platform also means that the model design should be simplified. In this paper, we only use simple tricks to capture the global degradation representation of blurred images, building an all-in-one framework for handling different degradation types.

In summary, most image enhancement methods are designed for specific types of degradation, making it difficult to generalize to other image enhancement tasks. For example, state-of-the-art image dehazing methods typically rely on the atmospheric scattering physical model to recover images by estimating unknown parameters in the physical model. Similarly, image motion deblurring methods are usually designed based on linear motion blur physical models. Research on image compression deblur typically focuses on removing artifact blocks. Compared to these task-specific image enhancement methods, our model can restore images of different degraded types, which effectively alleviates the shortage of storage resources in complex application scenarios. Moreover, existing all-in-one image enhancement approaches focus on the study of image degradation caused by severe weather (*e.g.* haze, rain or snow), while our work committed on image degradation caused by different factors (*e.g.* haze, motion blur or compression blur). Notably, they usually insert modules such as transformers or attention mechanisms into the network, which can easily introduce a large number of parameters that make it difficult to meet practical requirements. Therefore, we devise an all-in-one model with characteristics of lightweight and obtain blur-free aerial images characterized by good visibility that is more responsive to practical scenarios.

2.2 Receptive field in low-level vision

The receptive field in the deep neural network represents the size of the area mapped on the original image by the pixels on the output feature map of each convolutional layer. Since the network generally uses convolutional and pooling layers which are locally connected, neurons cannot perceive all the characteristics of the original image. Therefore, Zhang et al. employ dilated filters to expand the receptive field (Zhang et al., 2017a). However, dilated filter inherently suffers from grid effects, where the receptive field only considers a sparse sampling of the input image with a checkerboard pattern. To avoid the increased computational burden and potential sacrifice of performance improvement, Liu et al. expand the receptive field by applying a wavelet transform to the U-Net architecture and propose a multi-layer wavelet CNN (MWCNN) model with reduced computational complexity (Liu et al., 2018). Fu et al. propose deep convolutional sparse coding architecture with atrous convolution to obtain a high-level receptive field (Fu et al., 2019). Although these methods are able to ensure that the neurons cover the image area entirely. However, an excessively large receptive field easily introduces redundant information to the small target area, which reduces the performance of the model. To solve the problem of differences in the distribution of target regions in aerial images, this paper uses parallel convolution with different convolution kernels to extract multi-scale target region features, so as to obtain a more effective global degradation representation.

3 Method

In this section, we elaborate on the architecture of the proposed all-in-one image enhancement network AIENet. The overall architecture of the model is shown in Figure 2. The model can strike a balance between speed and accuracy. Given a degraded image, AIENet first performs a unique downsampling operation. Subsequently, to yield a more effective and comprehensive degraded representation, we adopt multiple receptive fields, catering to a wide range of target region sizes. Lastly, the global skip connection is used to fill in the blank content of the deconvolution process to get purer high-resolution information. To showcase the competence of the proposed model, we present three typical image degradation problems encountered by drones when monitoring forest landscapes, namely haze, motion blur, and compression blur, as targeted examples in this paper. In the following sections, we first illustrate the multi-receptive field image enhancement block, which forms the fundamental component of AIENet, and then elaborate on the overall model architecture featuring a skip structure. Finally, the objective function of the model is discussed.

3.1 MRF enhancement block

The MRF enhancement block is a versatile module with a plug-and-play design, enabling its integration into any part of an existing network. Notably, this block offers multi-scale area perception, guaranteeing the inclusion of various scale feature details in the final outcome. It can be decomposed into two fundamental components: 1) a multi-scale perception module, responsible for extracting distinct scale representations; 2) a feature merging operator, which merges intermediate feature maps. Specifically, the features of the last layer are initially fed into two distinct branches, each engaging in feature extraction *via* diverse dimensions. The multi-scale perception refers to the lower-dimensional branch within the block, which employs convolution kernels of varying sizes to facilitate multi-scale feature perception. The enhancement block concludes by utilizing channel-wise concatenation, which enables the learning of comprehensive contextual information. We elaborate on these processes in detail below. The pipeline of the MRF enhancement block is shown in Figure 3.

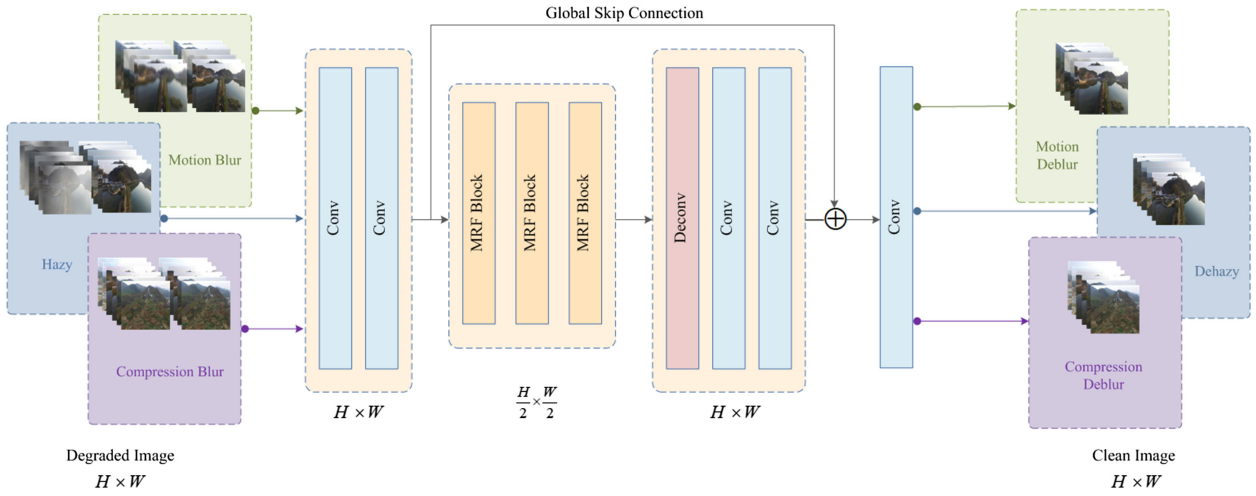


FIGURE 2
The overall structure of our All-in-one Image Enhancement Network (AIENet) which addresses image hazy, motion blur, and compression blur problems in an all-in-one model.

3.1.1 Multi-scale perception

The receptive field in a convolutional neural network represents the visual range of the network with respect to the input image. As only the input pixels within the receptive field contribute to the calculation, the size of the receptive field can be used to measure the ability of the model to leverage spatial information. However, it is not always optimal for the receptive field size to be maximized. In the case of larger targets, a larger receptive field can better integrate contextual information about the target area and restore its high-resolution details. For smaller objects, a larger receptive field can easily introduce excessive artifacts that may hinder the image restoration process. Especially for aerial images that are captured from multiple angles, the scale of the target area can constantly expand or shrink depending on the position of the drones. It is worth noting that a single receptive field may not always yield optimal results in learning the complex scale structures of aerial images. As such, Ren et al. and Liu et al. proposed a solution for multi-scale feature extraction Ren et al. (2016); Liu et al. (2019). While the design of multi-scale stacking allows the network to have a larger expression space in the receptive field, the network's receptive field is fixed in the inference stage when the model parameters are not updated. This is the statistical receptive field calculated by the model based on the data distribution of the training set, which may be suboptimal for each specific image. Additionally, by extracting intermediate features through this concatenated approach, gradient vanishing may occur, and the signal generated in earlier iterations may be disrupted.

To effectively address the aforementioned issues, this paper proposes the generation of intermediate feature maps through distinct branches. The aim of multi-scale perception is to utilize diverse receptive fields to enhance the comprehension of various regions and acquire multiple global degraded representations.

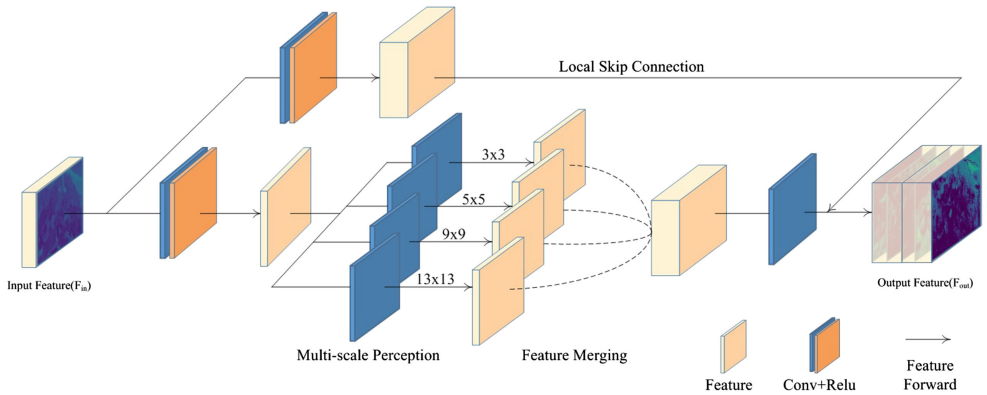


FIGURE 3
The MRF enhancement block comprises two branches: the high-dimensional extract branch and the multi-scale perception branch. The high-dimensional extract branch is responsible for preserving the high-resolution details of the input features, while the multi-scale perception branch employs convolution kernels of varying sizes to capture multi-scale features.

Indeed, it is possible to design the multi-scale perception module to be highly complex, maximizing the reasoning ability of the model. However, even with a simple parallel usage of several convolutional layers with different kernels and the use of skip connections to concatenate the shallow features with these multi-scale perception feature maps, we can already intuitively observe the efficiency of feature extraction under multi-scale perception. Specifically, in order to better preserve the high-resolution information of the original image, the resolution of all intermediate feature maps in the enhancement block is kept consistent with that of the input feature map. Formally, let's define $F_{in} \in R^{H \times W \times C_{in}}$ as the input feature map generated by the last layer. The enhancement block initially feeds the input feature map F_{in} into a dual-branch structure for feature extraction. The high-dimensional extract branch aims to learn more comprehensive original image features by expanding the channel dimension. The operation on the input feature map F_{in} can be defined as follows:

$$F_s = \text{Relu}\left(\text{Conv}_{c_1}(F_{in})\right) \tag{1}$$

Where $Conv_{c_1} \cdot (\cdot)$ is the convolution layer using c_1 convolution kernels and $F_s \in R^{H \times W \times C}{}_1$. Nevertheless, the local information captured by each pixel in F_s is limited. To address this issue, the multi-scale perception branch adopts a multi-kernel strategy, consisting of four convolutional layers with different kernels to capture more diverse features. This trick enables the network to conduct intricate feature extraction operations in low-dimensional feature spaces and reduce the number of model parameters. The resulting feature map generated by the parallel convolutional layers can be represented as follows:

$$F_d = \text{Relu}\left(\text{Conv}_{c_2}(F_{in})\right) \tag{2}$$

$$f_i = \omega_i(F_d)_{H \times W \times C_2}, \quad i \in \{1, 2, 3, 4\} \tag{3}$$

where $\omega_i(\cdot)$ denotes the parallel convolutional operation that generates the i^{th} scale receptive field and $F_d \in R^{H \times W \times C_2}$. This dual-branch structure maintains a relatively independent calculation scheme. The combination of a "deep network + multi-scale feature extraction" and a "shallow network + high-dimensional feature extraction" takes into account both the rich details of shallow features and the abstract semantics of multi-scale features. This combination also effectively manages the computational overhead to ensure the real-time performance.

3.1.2 Feature merging

Once the input feature map passes through multiple receptive fields of feature perception, as much contextual information as possible has been constructed between these different scales of receptive fields. We observe that for spatial tasks such as image enhancement, parallel multi-scale capabilities are required to handle perceptually large and small objects. In traditional processing methods, there are primarily two approaches for merging different feature maps: concatenation and element-wise addition. The latter requires the feature maps to have identical channels, necessitating the conversion of feature maps to uniform channels. However, this requirement restricts the flexibility of merging feature maps and direct summation of all feature maps may remove the generated image details. Instead, we adopt the concatenation approach to aggregate features from different receptive fields. At this stage, the feature channels already contain various local contextual information adapted to the target area size. The merged features can be further abstracted through convolution, allowing for soft transitions between receptive fields of varying scales and facilitating the construction of a holistic global context. To better convey the semantic details of the input features during training, skip connections are employed to ensure the effectiveness of the enhancement block, resulting in an output feature map with high-dimensional details and multi-scale perception. Let's define F_i as the feature map generated by the i^{th} MRF enhancement block. The output feature map can be expressed as:

$$F_m = \text{Cat}\left(\text{Conv}_{c_1}(Cat(f_i)), F_s\right), \quad i \in \{1, 2, 3, 4\} \tag{4}$$

where $Cat(\cdot)$ indicates channel-wise concatenation operator. In this way, feature maps of arbitrary numbers can be merged, which gives the MRF the potential to capture more details of multi-scale target areas.

3.2 Global-local skip connection

The residual network architecture has exhibited outstanding performance in computer vision tasks spanning from low-level to high-level tasks(Kim et al., 2016; Dong et al., 2020). This architecture was initially proposed by He et al. for image recognition (He et al., 2016). The purpose of the skip connection is to merge low-level features and high-level convolutional features with more intricate semantics. In spatial feature reconstruction tasks like image enhancement, the rich details preserved by high-level convolutions are extremely valuable. Nonetheless, the increase in the receptive field with network depth may result in the loss of high-dimensional details. To maintain fine details from the input image to the output image, we incorporate local skip connections, which significantly enhance the performance. To be more precise, the features extracted from the previous layer are first processed through convolution for high-dimensional feature extraction, and then combined with the multi-receptive field features before being passed on to the next module. This approach reduces the susceptibility of the model to loss of high-frequency information that may occur due to repeated convolution operations.

Although Liu et al. and Gao et al. successfully applied skip connections to image enhancement problems (Liu and Yang, 2018; Gao et al., 2019), it should be noted that deconvolution or upsampling often requires filling in a significant amount of missing content. It is important to acknowledge that generating high-quality results from scratch requires sufficient auxiliary information. To address the issue, we devise a novel skip connection that can take into account both global and local contextual information interaction. By employing a cross-layer global skip connection, the corresponding scale features are introduced into the deconvolution or upsampling process, which

can effectively preserve high-resolution details contained in the input images. This results in an enhanced ability of the network to recover image details, as illustrated in Figure 2. To exploit the merits of both designs, the model contains n MRF enhancement blocks and a global skip connection. Each enhancement block comprises local skip connections that fuse high-dimensional detail and multi-receptive field features. Such a residual structure allows the network to train deep models without sacrificing shallow information features.

3.3 Model architecture and loss function

Following the similar network design principle in Cai et al., we also design the overall network as a simple auto-encoder, where three residual blocks are inserted between the encoder and decoder to enhance its understanding capacity of different target regions (Cai et al., 2016; Ren et al., 2016; Li et al., 2017). Specifically, two convolutional layers are first used to encode the input blurred image into the feature map. This feature map is used as the encoder part, where only the last convolutional layer downsamples the feature map by a factor of 1/2. Correspondingly, a deconvolutional layer with a stride of 1/2 is used in the decoder part to upsample the feature maps to the original resolution. The feature maps are subsequently transformed back to image space using three convolutional layers to obtain the final blurred residual. For the middle residual block, we call it a "multi-receptive field enhancement block", because it uses four convolution kernels of different sizes to extract the details of varying target areas adaptively. The sizes of the four convolution kernels are set as 3x3, 5x5, 9x9, and 13x13, respectively. To obtain a good trade-off between performance and running time, we set the number of channels of all intermediate convolutional layers in the enhancement block to 32 or 128. Then an instance normalization (Ulyanov et al., 2016) and a Relu layer are placed after each convolutional layer. Each layer setting for the network is given in Table 1. Fan et al. has proved that in addition to the input image, pre-calculating the edges of the input image and feeding it into the network as auxiliary information is beneficial for network learning. By default, we also adopt this idea and concatenate the pre-calculated edges with the input blurred image along the channel dimension as the final input of the network (Fan et al., 2017; Fan et al., 2018; Ren et al., 2018).

Most learning-based image enhancement methods (Cai et al., 2016; Ren et al., 2016; Li et al., 2022) use Mean Square Error (MSE) loss to train the models. Following the same strategy, we also use this simple loss. Specifically, we adopt the strategy of (Ren et al., 2018) and set the learning objective of the model as the residual between the clear image and the input degraded one. In summary, the total loss can then be written as follows:

$$L = \|\hat{r} - r\|^2 \tag{5}$$

where r^ is the predicted residual, r is the residual of the degraded image and clear image at location (i,j), which can be calculated as follows:

$$r = \sum_{i=1}^{H}\sum_{j=1}^{W}(h(i,j) - g(i,j)) \tag{6}$$

Even with the only simple loss mentioned above, our method can still achieve state-of-the-art performance on aerial image enhancement. Further, this kind of loss function also enables efficient training due to the smaller number of parameters to update.

4 Experiments

This section provides qualitative and quantitative comparisons with state-of-the-art methods for three challenging aerial image enhancement tasks, *i.e.*, image dehaze, image motion deblur, and image compression deblur. We first introduce the dataset source and experimental settings. Then, we present the results of comparing our proposed algorithm with 15 state-of-the-art methods. Finally, the effectiveness of the proposed module is demonstrated through ablation experiments.

4.1 Dataset

We constructed datasets of degraded aerial images to evaluate the effectiveness of the proposed method in handling various types of image degradation. The raw images were sourced from a publicly available dataset, which comprised video frames captured by drones

TABLE 1 Network Setting.

	Layer Description	Output Size
Encoder		
#1	Conv (3, 32, 3, 1)	640x640x32
#2	Conv (32, 64, 3, 2)	320x320x64
3x MRF Enhancement Block(c)		
#1_1	Conv (c, 32, 3, 1)	320x320x32
#1_2	Conv (c, 128, 3, 1)	320x320x128
#2_1	Conv (32, 32, 3, 1)	320x320x32
#2_2	Conv (32, 32, 5, 1)	320x320x32
#2_3	Conv (32, 32, 9, 1)	320x320x32
#2_4	Conv (32, 32, 13, 1)	320x320x32
#3	Cat + Conv (128, 128, 3, 1)	320x320x128
Decoder		
#1	Deconv (256, 128, 4, 2)	640x640x128
#2	Conv (128, 64, 3, 1)	640x640x64
#3	Conv (64, 32, 3, 1)	640x640x32
#4	Conv (64, 32, 3, 1)	640x640x3

Where c indicates the number of feature channels entering the MRF enhancement blocks.

equipped with video surveillance cameras. To obtain degraded images of drone-monitored forest scenes under various conditions, we employ Python library Imgaug (Jung et al., 2020) to synthesize paired degraded images. For the 2007 original images, we generated two degraded images with different levels of degradation for each image by adjusting different parameters depending on the degradation type. In this paper, we synthesize haze images with different concentrations by setting the scattering coefficient to 2 or 3, generate motion blur images using blur kernels ranging from 25 to 34 and angles ranging from -150 to 360 degrees, and produce compression blur images by randomly selecting compression ratios between 89 and 93. Regarding the scattering coefficient used in generating hazy images, we based them on previous research about the atmospheric scattering physical model. The motion blur kernel and angle parameters were selected based on the linear motion blur physical model. The compression rate parameters are based on the pixel count of the image and the compression algorithm parameters. Finally, 4014 degraded images were generated for each of the three tasks, 70% of which are used for fully supervised training and the rest for testing. Since we use a mixture of three degraded types in the all-in-one image enhancement task, our all-in-one framework can effectively generate close-to-ground truth images for any degradation type.

4.2 Training specifications

All experiments are conducted using PyTorch on an Ubuntu 20.04 system, with NVIDIA RTX 3080Ti GPU to optimize the training speed. For each task, we compare the proposed method with the state-of-the-art methods separately. Then the generality of AIENet is demonstrated by further comprehensive training. We use almost the same training strategy for these models. For a fair comparison, all models are trained for 60 epochs. By default, we train our model with batch size 2 using the Adam optimizer (Kingma and Ba, 2014). The default initial learning rate is set to 0.001, decaying by 0.1 every 40 epochs. The changing trend of target loss is shown in Figure 4. In the early stages of training, the loss value is relatively high, indicating a large discrepancy between the predicted and ground truth images. However, the loss value of the model drops very quickly after training several epochs and plateaus at epoch 40. This suggests that our model may have converged earlier, but we still follow the default training strategy for comparison with previous work.

4.3 Comparisons with state-of-the-art methods

Our model is individually compared with five state-of-the-art methods on specific tasks for a comprehensive comparison. Specifically, we compare with (Li et al., 2017; Ren et al., 2018; Dong et al., 2020; Qin et al., 2020; Li et al., 2021) on the dehazing task. The methods to remove motion blur include (Nah et al., 2017; Tao et al., 2018; Kupyn et al., 2019; Zhang et al., 2019; Cho et al., 2021). The baselines for compression deblur are (Dong et al., 2015; Chang et al., 2020; Chen et al., 2021; Jiang et al., 2021; Zamir et al., 2021). To demonstrate the superiority of our all-in-one framework, we also compare models trained in an all-in-one manner on three tasks. In other words, we train the proposed model on an ensemble of all datasets consisting of degraded images with three different degradation types (i.e., haze, motion blur, and compression blur). And then test on a single type.

Quantitative evaluations between ground truth x and restored images y were performed *via* the conventional Peak Signal-to-Noise Ratio (PSNR) (Huynh-Thu and Ghanbari, 2008) and Structural Similarity (SSIM) (Wang et al., 2004) metrics. PSNR is a very important indicator in the field of image enhancement, which can be expressed as:

$$PSNR = 10\log_{10}\left(\frac{L^2}{MSE}\right) \tag{7}$$

where L is the possible maximal pixel value. The mean square error MSE between x and y is calculated as follows where H and W are the height and width of the images:

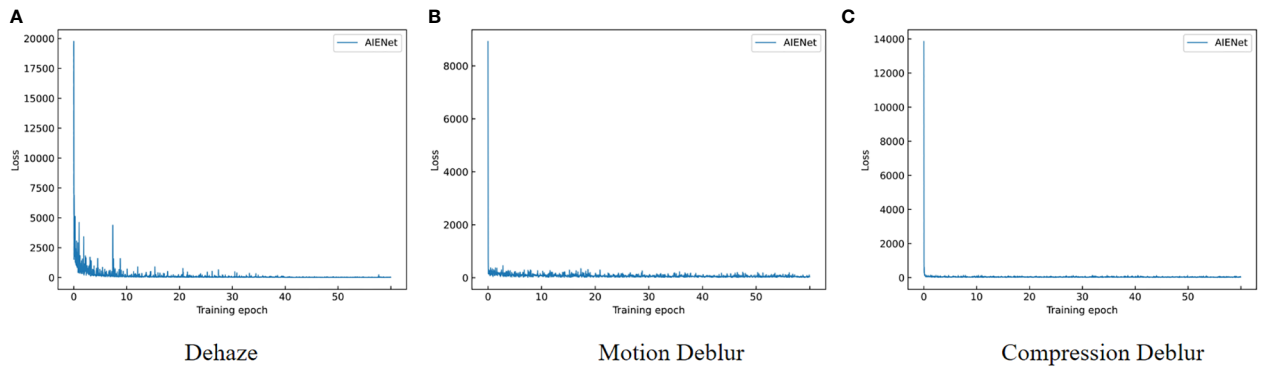


FIGURE 4
Curves of the changing trend of target loss. **(A)** shows the trend of the dehazing loss function, **(B)** shows the trend of the compression deblurring loss function, and **(C)** shows the trend of the motion deblurring loss function.

$$MSE(x,y) = \frac{1}{H \times W}\sum_{i=1}^{H}\sum_{j=1}^{W}(x(i,j) - y(i,j))^2 \quad (8)$$

In comparison to PSNR, the structural similarity indicator is more in line with human subjective system judgment on image quality. SSIM is designed to compute the luminance, contrast, and structural similarity between the x and y, which can be represented by:

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + c_1)\ (2\sigma_{xy} + c_2)}{(\mu_x^2+\mu_y^2 + c_1)\ (\sigma_x^2+\sigma_y^2 + c_2)} \quad (9)$$

where μ_x and μ_y are the mean of x and y, respectively. σ_x and σ_y are the variance of x and y, respectively. σ_{xy} is the covariance of x and y. By default, c_1=·$(0.01L)$ and c_2=·$(0.03L)$ are the constants used to avoid divisors by zero. We evaluated PSNR and SSIM based on the luminance channel Y of the YCbCr color space in accordance with the previous convention (Zamir et al., 2021; Valanarasu et al., 2022).

4.3.1 Task-specific image enhancement results

Quantitative Evaluation for Image Enhancement. Table 2 presents our quantitative evaluations. The top half of the tables contain results from task-specific image restoration. Our models achieve performances superior to all compared existing methods in PSNR on all tasks. For the image dehazing task, the proposed method yields the best PSNR of 35.69 dB, which also outperforms all dedicated to dehazing models. Notably, in our experiments, we found GCANet (Ren et al., 2018) to be the best-performing network for dehazing in SOTAs. And the method in this paper also achieves a breakthrough of 5.37%. Furthermore, we also get small victories in objectively evaluating SSIM close to the Human Visual System (HVS). In the motion deblurring task, our model exceeds all compared deblurring networks in terms of PSNR. It is worth noting that our model is the second best in the comparison of structural feature recovery. But compared to MIMO-UNet (Cho et al., 2021), the best network for motion blur removal in this experiment, our model parameters are only 10.62MB, while the MIMO-UNet network has a parameter amount of 25.97MB.

Qualitative Results for Image Dehazing. To illustrate that our model can better remove the visual effects of haze and restore more image details than other dehazing methods, Figure 5 depicts some visualizations of image dehazing reconstructions for aerial images of forests, comparing our method with FFANet (Qin et al., 2020) and GCANet (Ren et al., 2018). As illustrated, the FFANet does not completely remove the influence of haze, and its restored image has some artifacts. While GCANet seems to have the comparable visual quality to our model in image dehazing, our AIENet achieves visually pleasing results in detail enhancement (enlarged in red and blue bounding boxes).

Qualitative Results for Image Deblurring: To demonstrate the images restored by our model are sharper and produce fewer artifacts, Figure 6 visualizes motion deblurring examples, demonstrating the superiority of our model AIENet over MIMO-UNet (Cho et al., 2021) and DMPHN (Zhang et al., 2019).In particular, the state-of-the-art methods still retain obvious streak artifacts when restoring images, while our model can preserve the structural and textural image details. (*e.g.*, second example in Figure 6, forest enlarged in the red bounding boxes). Although in quantitative experiments, the proposed model does not show competitive performance on image compression deblurring. But in visual analysis, as shown in Figure 7, the proposed model can produce excellent visual quality on par with state-of-the-art methods.

Generality to Different Image Enhancement Tasks: To demonstrate the superior generalization of the proposed method, we compare it with different task-specific image enhancement methods on three challenging tasks including dehaze, motion deblur and compression deblur. As shown in Table 3, although HINet performs well in removing compression artifacts, it has uncompetitive results on image dehazing and motion deblurring. Similarly, FFANet exhibits significant performance discrepancies across different tasks, with a PSNR of 32.17 dB for compression deblurring but only 20.97 dB for dehazing. Evidently, these methods excel only in specific tasks while performing poorly in others. Although generalization seems to be visible in MIMO-UNet, our AIENet exhibits a more competitive restoration performance than it

TABLE 2 Quantitative comparisons in terms of PSNR and SSIM (the symbol "↑" means that higher value is better) with state-of-the-art image dehazing, motion deblurring, and compression deblurring methods.

		Dehaze			MotionDeblur			Compression Deblur	
		PSNR↑	SSIM↑		PSNR↑	SSIM↑		PSNR↑	SSIM↑
Task-specific	AODNet	13.99	0.7592	MS-CNN	28.47	0.8107	ARCNN	27.62	0.7986
	GCANet	<u>33.87</u>	<u>0.9632</u>	SRN	28.35	0.7968	FBCNN	30,99	0.9001
	MSBDN	14.96	0.8864	DeblurGAN-v2	27.54	0.7752	HINet	**33.11**	**0.9104**
	FFANet	20.97	0.9325	DMPHN	29.59	0.8248	SADNet	32.81	<u>0.9095</u>
	YOLY	10.22	0.5284	MIMO-UNet	<u>31.33</u>	**0.9317**	MPRNet	32.47	0.8626
	AIENet	**35.69**	**0.9642**	**AIENet**	**31.87**	<u>0.8648</u>	**AIENet**	<u>32.98</u>	0.8764
All-in-one	**AIENet**	32.50	0.9501	**AIENet**	28.88	0.7909	**AIENet**	31.70	0.8587

The best and second-best results are highlighted in bold and underlined, respectively. The above half of the table shows comparisons of our task-specific models individually evaluated for each task. The last row of the table show evaluations of our all-in-one model AIENet on all three test sets.

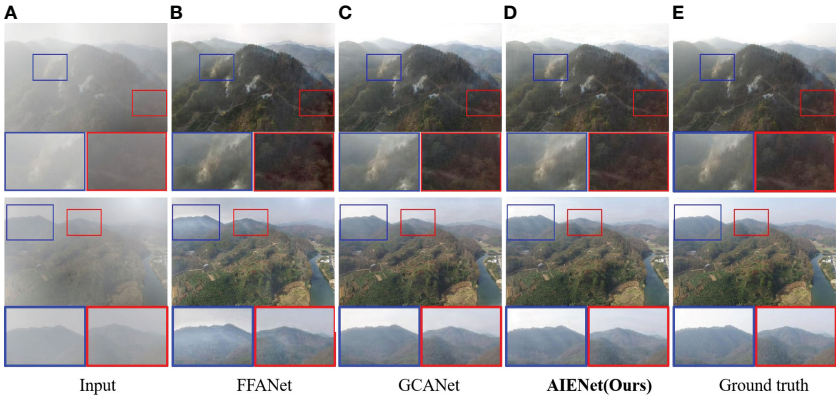


FIGURE 5
Qualitative enhancement comparisons of our model on synthetic hazy samples with FFANet (Qin et al., 2020) and GCANet (Ren et al., 2018). **(A)** is the input hazy images, **(B, C)** are the enhancement results of the state-of-the-art algorithms, **(D)** is the enhancement results of the proposed AIENet, and **(E)** is the ground truth images. Blue and red boxes correspond to the zoomed-in patch for better comparison.

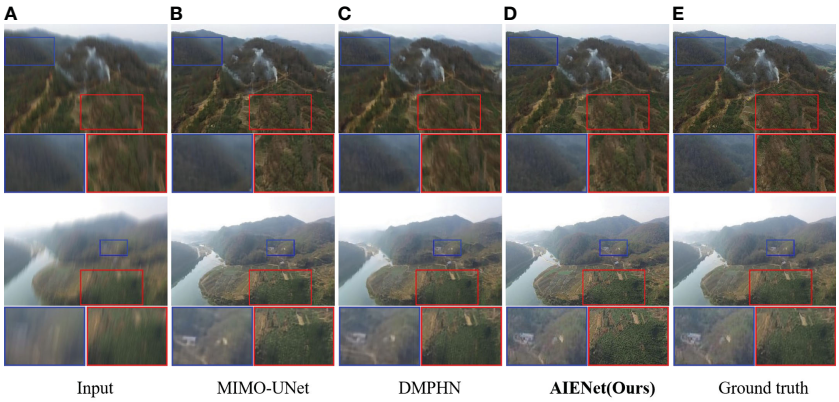


FIGURE 6
Qualitative enhancement comparisons of our model on synthetic motion blur samples with MIMO-UNet (Cho et al., 2021) and DMPHN (Zhang et al., 2019). **(A)** is the input motion blur images, **(B, C)** are the enhancement results of the state-of-the-art algorithms, **(D)** is the enhancement results of the proposed AIENet, and **(E)** is the ground truth images. Blue and red boxes correspond to the zoomed-in patch for better comparison.

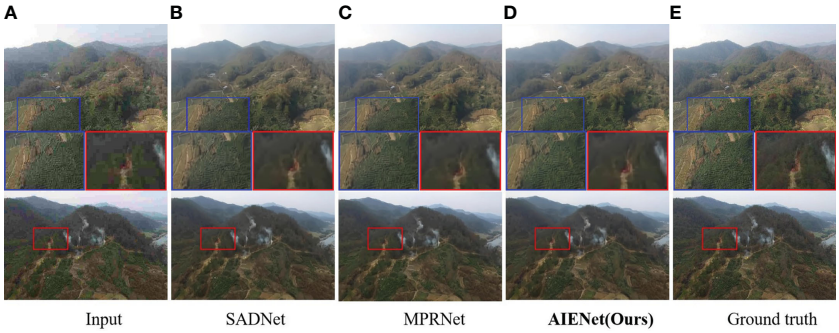


FIGURE 7
Qualitative enhancement comparisons of our model on synthetic compression blur samples with SADNet (Chang et al., 2020) and MPRNet (Zamir et al., 2021). **(A)** is the input compression blur images, **(B, C)** are the enhancement results of the state-of-the-art algorithms, **(D)** is the enhancement results of the proposed AIENet, and **(E)** is the ground truth images. Blue and red boxes correspond to the zoomed-in patch for better comparison.

TABLE 3 Quantitative comparison results (PSNR/SSIM) of some excellent methods for dehazy, motion deblurring and compression deblurring tasks.

Methods	Haze	Haze	Compression Blur
FFANet (Dehaze)	20.97/0.9325	29.60/0.8366	32.17/0.8785
FFANet (Dehaze)	30.12/**0.9842**	31.33/**0.9317**	32.76/**0.9408**
HINet (Compression Deblur)	21.94/0.9563	31.20/0.9035	**33.11**/0.9104
Ours (All-in-one)	32.50/0.9501	28.88/0.7909	31.70/0.8587
Ours	**35.69**/0.9642	**31.87**/0.864	32.98/0.8764

The best and second-best evaluation results are highlighted in bold and underlined, respectively.

in each task. The experimental results on datasets for hazy, motion blurring, and compression blurring show that our model excels at generalizing to diverse image domains.

4.3.2 All-in-one image enhancement results

The last row of the Table 2 presents quantitative evaluations for all-in-one image restoration. Generally, our method yields exceptional image quality and is faithful to the ground truth on all three test sets. Notably, for the image dehazing task, our trained all-in-one image enhancement network is second only to GCANet, the state-of-the-art model trained on the specific task, with PSNR/SSIM metrics reaching 32.50 dB/0.9501. Generally, the difference in image quality is less noticeable when the PSNR value reaches above 28 dB. Therefore, our model shows its outstanding performance and application value in environments that are sensitive to computational cost and running time.

4.4 Ablation study

In this section, we present ablation experiments to analyze the contribution of each component of our model. Specifically, we focus on two major components: with/without the skip connection and with the different number of enhancement blocks. Task-specific evaluation is performed on the synthetic haze dataset with the proposed models trained on the image size of 640Ã—640, and the results are shown in Table 4. To further validate the importance of each specific component in the all-in-one task, we also conduct analysis on the union of three datasets. Generally, we evaluate four different network configurations and follow the same training setup as the above experiments.

The influence of global-local skip connection: As mentioned in Section 3.2, skip connection can provide more high-resolution details of the original image for deconvolution or upsampling processes. Therefore, we demonstrate the influence of the design by removing them from our final model. Table 4 shows a substantial drop in PSNR of the image dehazing results from 35.69 dB to 32.36 dB when the global-local skip connection is removed. Correspondingly, the absence of the skip connection leads to poor performance as compared to employing it for all-in-one image enhancement. A similar trend is observed for the method without a local skip connection, where gains of the original model over it are 1.73 dB/0.0089 on PSNR/SSIM. We also provide two representative dehazing examples in Figure 8 for visual comparison. It can be seen that the images restored by removing skip connections contain either overly smooth contents or artifacts with grid textures. In contrast, the complete model is able to remove real noise, while preserving the structural and textural image details.

The effectiveness of the proposed MRF enhancement block: Since our model could employ different enhancement block number, we test different options. The results on image dehazing and all-in-one image enhancement tasks corresponding to different n are given in Table 4. This ablation study reveals that MRF block effectively increases the PSNR by 13.56% from n=1 to n=2, owning to the diverse receptive fields and the multi-scale perception mechanism. It is worth noticing that the model yields better performance in PSNR and SSIM respectively as the number of the MRF enhancement block increases, but the gains show a clear

TABLE 4 Ablation study on individual components of the proposed AIENet.

Method		Task-specific				All-in-one			
		PSNR↑		SSIM↑		PSNR↑		SSIM↑	
w/o global-local skip connection		32.36	(10.3%)	0.9290	(3.8%)	30.11	(7.9%)	0.8664	(9.7%)
w/o local skip connection		33.96	(5.1%)	0.9553	(0.9%)	29.61	(4.8%)	0.8681	(2.2%)
MRF	n = 1	30.16	(18.3%)	0.9147	(5.4%)	30.85	(5.3%)	0.8839	(7.5%)
	n = 2	34.25	(4.2%)	0.9541	(1.1%)	31.02	(4.8%)	0.8840	(7.5%)
	n = 4	36.18	(-1.4%)	0.9651	(-0.1%)	31.69	(-2.1%)	0.8890	(-0.2%)
Ours		**35.69**	(0.0%)	**0.9642**	(0.0%)	**31.05**	(0.0%)	**0.8875**	(0.0%)

The PSNR and SSIM of the proposed method are highlighted in bold. The symbol "↑" means that higher value is better.

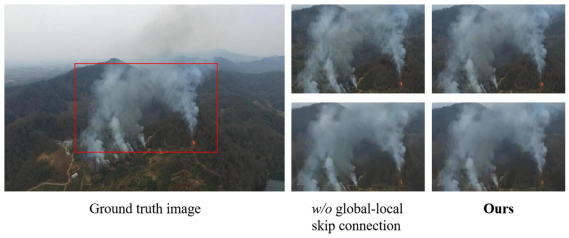


FIGURE 8
Examples on image dehazing(first row) and all-in-one image enhancement(second row) to show the superiority of global-local skip connection. Obviously, our model with glob-local skip connection improves the gridding artifacts and produces much better details.

downward trend. As such, the results also indicate that the model performance is not from the deeper layers but from a more efficient architecture, since more modules do not improve the performance much and our model has a smaller size.

4.5 Smoke detection results

As discussed in the introduction, aerial image enhancement could be helpful in improving the performance of fire detection approaches in forest fire prevention based on drone imagery monitoring. Therefore, we train a smoke detection algorithm (Wang et al., 2022) on the raw dataset. To verify the effectiveness of our method in boosting image detection, we use the results of image dehazing, image motion deblurring, and image compression deblurring as input exemplars for the detection algorithm, respectively. As a comparison, we also train models that perform better in each task, then test their enhanced results in the detection algorithm. As shown in Figure 9, the confidence below these images demonstrates the quantitative comparisons between the proposed model and the state-of-the-art methods. The results show that whether it is restoring blurred images or removing weather disturbances, our image enhancement method can effectively improve the confidence of the detection algorithm.

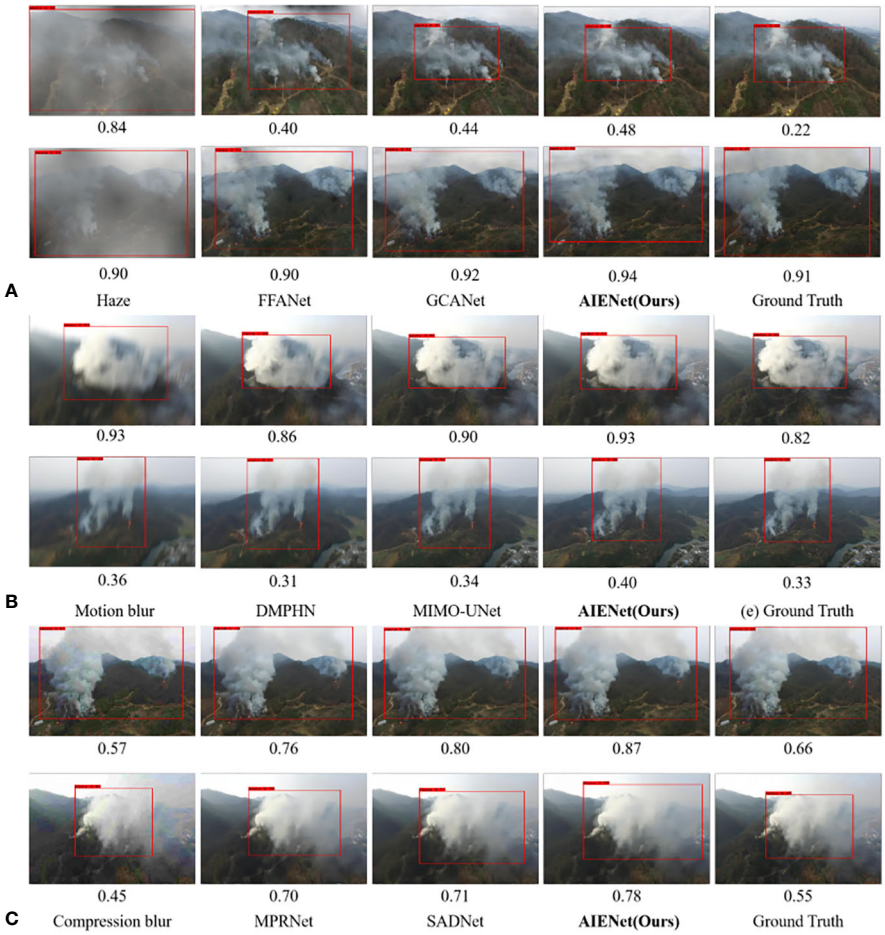


FIGURE 9
Qualitative comparison of our model with other works in improving fire detection performance. **(A)** is the visual comparison of image dehazing results. **(B)** is the visual comparison of image motion deblurring results. **(C)** is the visual comparison of image compression deblurring results.

5 Conclusion

This paper presents an image enhancement method based on multiple receptive fields to improve the visual effect of aerial images in forest scenes. We focused on building an all-in-one framework that eliminates any degradation in aerial imagery. Based on this, we also devise a novel multi-receptive field enhancement block, which can adapt to the distribution differences of object regions in aerial images. It also benefits the network by recovering high-resolution details of images more efficiently. Extensive experiments have validated the merits of our method over other state-of-the-art enhancement methods on benchmark datasets. Specifically, our AIENet has achieved considerable gains in both dehazing and motion deblurring tasks, *i.e.*, 5.3% improvement in PSNR on the haze dataset, and a 1.7% increase on the motion blur dataset. The results of all-in-one image enhancement also show that our model has the ability to obtain performance close to SOTAs, which avoids the lack of resources associated with storing models separately to handle individual enhancement tasks. And we have also experimentally demonstrated that AIENet generalizes well to other image domains. Moreover, we further conduct ablation experiments to demonstrate the influence of the proposed MRF enhancement block. We show that using three enhancement blocks leads to optimal performance (35.69 dB on a specific task, and 31.05 dB on an all-in-one task) as compared to employing other quantities of enhancement blocks. Notably, the proposed method introduces lightweight image enhancement capability since the architecture can be based on a simpler backbone network for image restoration with less running time, which is of great interest for devices with limited resources.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/supplementary material.

Author contributions

ZC designed the experiments and wrote the first draft of the manuscript. CW made substantial contributions to the design of the study and the revision of the manuscript. FZ received financial support for this project and performed the analysis of the manuscript. LZ performed the experimental data preparation. AG and EG contributed to the revision of the manuscript, and read and approved the submitted version. All authors contributed to the article and approved the submitted version.

Funding

This work was jointly supported by the project of Digital Media Art, Key Laboratory of Sichuan Province (Sichuan Conservatory of Music, Project No. 21DMAKL01), the first batch of industry-university cooperation collaborative education project funded by the Ministry of Education of the People's Republic of China (Minjiang University, Project No. 202101071001), Minjiang University 2021 school-level scientific research project (Minjiang University, Project No. MYK21011), Open Fund Project of Fuzhou Technology Innovation Center of Intelligent Manufacturing Information System (Minjiang University, Grant No. MJUKF-FTICIMIS2022), Open Fund Project of Engineering Research Center for ICH Digitalization and Multi-source Information Fusion (Fujian Polytechnic Normal University, Grant No. G3-KF2204), Guiding Project of Fujian Province (Minjiang University, Project No. 2020H0046). Key Technology Research and Industrialization Project for Software Industry Innovation in Fujian Province (MinjiangUniversity and Fujian Guotong Information Technology Co., Ltd., Project No. 36).

Acknowledgments

The authors would thank to Minjiang University for providing GPU device and the valuable comments from editors and reviewers.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Alsubai, S., Hamdi, M., Abdel-Khalek, S., Alqahtani, A., Binbusayyis, A., and Mansour, R. F. (2022). Bald eagle search optimization with deep transfer learning enabled age-invariant face recognition model. *Image Vision Computing* 126, 104545. doi: 10.1016/j.imavis.2022.104545

Cai, B., Xu, X., Jia, K., Qing, C., and Tao, D. (2016). Dehazenet: An end-to-end system for single image haze removal. *IEEE Trans. Image Process.* 25, 5187–5198. doi: 10.1109/TIP.2016.2598681

Chang, M., Li, Q., Feng, H., and Xu, Z. (2020). "Spatial-adaptive network for single image denoising," in *European Conference on Computer Vision*. (Glasgow, United: Springer), 171–187.

Chen, X., Fan, Z., Li, P., Dai, L., Kong, C., Zheng, Z., et al. (2022). "Unpaired deep image dehazing using contrastive disentanglement learning," in *European Conference on Computer Vision*. (Tel Aviv, Israel: Springer), 632–648.

Chen, L., Lu, X., Zhang, J., Chu, X., and Chen, C. (2021). "Hinet: Half instance normalization network for image restoration," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. (Nashville, TN, USA), 182–192.

Cho, S.-J., Ji, S.-W., Hong, J.-P., Jung, S.-W., and Ko, S.-J. (2021). "Rethinking coarse-to-fine approach in single image deblurring," in *Proceedings of the IEEE/CVF international conference on computer vision*. (Montreal, QC, Canada), 4641–4650.

Ct, A., Yong, X., and Wz, C. (2020). Image denoising using deep cnn with batch renormalization - sciencedirect. *Neural Networks* 121, 461–473. doi: 10.1016/j.neunet.2019.08.022

Dong, C., Deng, Y., Loy, C. C., and Tang, X. (2015). "Compression artifacts reduction by a deep convolutional network," in *Proceedings of the IEEE international conference on computer vision*. (Santiago, Chile), 576–584.

Dong, H., Pan, J., Xiang, L., Hu, Z., Zhang, X., Wang, F., et al. (2020). "Multi-scale boosted dehazing network with dense feature fusion," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. (Seattle, WA, USA), 2157–2167.

Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., et al. (2021). "An image is worth 16x16 words: Transformers for image recognition at scale," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 3156–3164. doi: 10.1109/CVPR.2021.00533

Fan, Q., Chen, D., Yuan, L., Hua, G., Yu, N., and Chen, B. (2018). "Decouple learning for parameterized image operators," in *Proceedings of the European Conference on Computer Vision (ECCV)*. (Munich, Germany), 442–458.

Fan, Q., Yang, J., Hua, G., Chen, B., and Wipf, D. (2017). "A generic deep architecture for single image reflection removal and image smoothing," in *Proceedings of the IEEE International Conference on Computer Vision*. (Venice, Italy), 3238–3247.

Farghaly, M., Mansour, R. F., and Sewisy, A. A. (2022). Two-stage deep learning framework for srgb image white balance. *Signal Image Video Process* 17, 277–284. doi: 10.1007/s11760-022-02230-2

Fu, X., Zha, Z.-J., Wu, F., Ding, X., and Paisley, J. (2019). "Jpeg artifacts reduction *via* deep convolutional sparse coding," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*. (Seoul, Korea (South)), 2501–2510.

Gao, H., Tao, X., Shen, X., and Jia, J. (2019). "Dynamic scene deblurring with parameter selective sharing and nested skip connections," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. (Long Beach, CA, USA), 3848–3856.

He, K., Sun, J., and Tang, X. (2010). Single image haze removal using dark channel prior. *IEEE Trans. Pattern Anal. Mach. Intell.* 33, 2341–2353. doi: 10.1109/CVPR.2009.5206515

He, K., Zhang, X., Ren, S., and Sun, J. (2016). Deep residual learning for image recognition. *IEEE*, 770–778. doi: 10.1109/CVPR.2016.90

Huynh-Thu, Q., and Ghanbari, M. (2008). Scope of validity of psnr in image/video quality assessment. *Electron. Lett.* 44, 800–801. doi: 10.1049/el:20080522

Jiang, J., Zhang, K., and Timofte, R. (2021). "Towards flexible blind jpeg artifacts removal," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*. (Montreal, QC, Canada), 4997–5006.

Jung, A. B., Wada, K., Crall, J., Tanaka, S., Graving, J., Reinders, C., et al. (2020) *Imgaug*. Available at: https://github.com/aleju/imgaug (Accessed 01-Feb-2020).

Kim, J., Lee, J. K., and Lee, K. M. (2016). "Accurate image super-resolution using very deep convolutional networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*. (Las Vegas, NV, USA), 1646–1654.

Kingma, D. P., and Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*. doi: 10.5555/3042817.3043064

Kupyn, O., Budzan, V., Mykhailych, M., Mishkin, D., and Matas, J. (2018). "Deblurgan: Blind motion deblurring using conditional adversarial networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*. (Salt Lake City, UT, USA), 8183–8192.

Kupyn, O., Martyniuk, T., Wu, J., and Wang, Z. (2019). "Deblurgan-v2: Deblurring (orders-of-magnitude) faster and better," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*. (Seoul, Korea (South)), 8878–8887.

Lee, K., Kim, D. S., and Kim, T. (2004). Regression-based prediction for blocking artifact reduction in jpeg-compressed images. *IEEE Trans. Image Process.* 14, 36–48. doi: 10.1109/TIP.2004.838699

Lee, D., Lee, C., and Kim, T. (2021). "Wide receptive field and channel attention network for jpeg compressed image deblurring," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. (Nashville, TN, USA), 304–313.

Li, B., Gou, Y., Gu, S., Liu, J. Z., Zhou, J. T., and Peng, X. (2021). You only look yourself: Unsupervised and untrained single image dehazing neural network. *Int. J. Comput. Vision* 129, 1754–1767. doi: 10.1007/s11263-021-01431-5

Li, B., Liu, X., Hu, P., Wu, Z., Lv, J., and Peng, X. (2022). "All-in-one image restoration for unknown corruption," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. (New Orleans, LA, USA), 17452–17462.

Li, B., Peng, X., Wang, Z., Xu, J., and Feng, D. (2017). "Aod-net: All-in-one dehazing network," in *Proceedings of the IEEE international conference on computer vision*. (Venice, Italy), 4770–4778.

Li, R., Tan, R. T., and Cheong, L.-F. (2020). "All in one bad weather removal using architectural search," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. (Seattle, WA, USA), 3175–3185.

Liu, X., Ma, Y., Shi, Z., and Chen, J. (2019). "Griddehazenet: Attention-based multi-scale network for image dehazing," in *Proceedings of the IEEE/CVF international conference on computer vision*. (Seoul, Korea (South)), 7314–7323.

Liu, J.-Y., and Yang, Y.-H. (2018). "Denoising auto-encoder with recurrent skip connections and residual regression for music source separation," in *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA) (IEEE)*. (Orlando, FL, USA), 773–778.

Liu, P., Zhang, H., Zhang, K., Lin, L., and Zuo, W. (2018). "Multi-level wavelet-cnn for image restoration," in *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*. 773–782.

McCartney, E. J. (1976). *Optics of the atmosphere: scattering by molecules and particles* (New York: IEEE).

Nah, S., Hyun Kim, T., and Mu Lee, K. (2017). "Deep multi-scale convolutional neural network for dynamic scene deblurring," in *Proceedings of the IEEE conference on computer vision and pattern recognition*. (Honolulu, HI, USA), 3883–3891.

Qin, X., Wang, Z., Bai, Y., Xie, X., and Jia, H. (2020). "Ffa-net: Feature fusion attention network for single image dehazing," in *Proceedings of the AAAI Conference on Artificial Intelligence*. (New York, USA), 11908–11915.

Qu, Y., Chen, Y., Huang, J., and Xie, Y. (2019). "Enhanced pix2pix dehazing network," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. (Long Beach, CA, USA), 8160–8168.

Ren, W., Liu, S., Zhang, H., Pan, J., Cao, X., and Yang, M.-H. (2016). "Single image dehazing *via* multi-scale convolutional neural networks," in *European conference on computer vision*. (Amsterdam, The Netherlands), 154–169.

Ren, W., Ma, L., Zhang, J., Pan, J., Cao, X., Liu, W., et al. (2018). "Gated fusion network for single image dehazing," in *Proceedings of the IEEE conference on computer vision and pattern recognition*. (Salt Lake City, UT, USA), 3253–3261.

Song, Y., He, Z., Qian, H., and Du, X. (2022). Vision transformers for single image dehazing. *arXiv e-prints*. doi: 10.48550/arXiv.2204.03883

Tao, X., Gao, H., Shen, X., Wang, J., and Jia, J. (2018). "Scale-recurrent network for deep image deblurring," in *Proceedings of the IEEE conference on computer vision and pattern recognition*. (Salt Lake City, UT, USA), 8174–8182.

Ulyanov, D., Vedaldi, A., and Lempitsky, V. (2016). Instance normalization: The missing ingredient for fast stylization. *arXiv preprint arXiv:1607.08022*. doi: 10.48550/arXiv.1607.08022

Valanarasu, J. M. J., Yasarla, R., and Patel, V. M. (2022). "Transweather: Transformer-based restoration of images degraded by adverse weather conditions," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. (New Orleans, LA, USA), 2353–2363.

Wang, Z., Bovik, A. C., Sheikh, H. R., and Simoncelli, E. P. (2004). Image quality assessment: from error visibility to structural similarity. *IEEE Trans. image Process.* 13, 600–612. doi: 10.1109/TIP.2003.819861

Wang, C., Grau, A., Guerra, E., Shen, Z., Hu, J., and Fan, H. (2022). Semi-supervised wildfire smoke detection based on smoke-aware consistency. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.980425

Wang, C., Li, Z., Wu, J., Fan, H., Xiao, G., and Zhang, H. (2020). Deep residual haze network for image dehazing and deraining. *IEEE Access* 8, 9488–9500. doi: 10.1109/ACCESS.2020.2964271

Wang, S., and Liu, B. (2022). "Deep attention-based lightweight network for aerial image deblurring," in *2022 26th International Conference on Pattern Recognition (ICPR) (IEEE)*. (Montreal, QC, Canada), 111–118.

Wang, A., Wang, W., Liu, J., and Gu, N. (2018). Aipnet: Image-to-image single image dehazing with atmospheric illumination prior. *IEEE Trans. Image Process.* 28, 381–393. doi: 10.1109/TIP.2018.2868567

Wu, H., Qu, Y., Lin, S., Zhou, J., Qiao, R., Zhang, Z., et al. (2021). "Contrastive learning for compact single image dehazing," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. (Nashville, TN, USA), 10551–10560.

Xue, Y., Tang, Y., Xu, X., Liang, J., and Neri, F. (2021). Multi-objective feature selection with missing data in classification. *IEEE Trans. Emerging Topics Comput. Intell.* 6, 355–364. doi: 10.1109/TETCI.2021.3074147

Xue, Y., Xue, B., and Zhang, M. (2019). Self-adaptive particle swarm optimization for large-scale feature selection in classification. *ACM Trans. Knowledge Discovery Data (TKDD)* 13, 1–27. doi: 10.1145/3340848

Yoo, S. B., Choi, K., and Ra, J. B. (2014). Post-processing for blocking artifact reduction based on inter-block correlation. *IEEE Trans. Multimedia* 16, 1536–1548. doi: 10.1109/TMM.2014.2327563

Zamir, S. W., Arora, A., Khan, S., Hayat, M., Khan, F. S., Yang, M.-H., et al. (2021). "Multi-stage progressive image restoration," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. (Nashville, TN, USA), 14821–14831.

Zhang, H., Dai, Y., Li, H., and Koniusz, P. (2019). "Deep stacked hierarchical multi-patch network for image deblurring," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. (Long Beach, CA, USA), 5978–5986.

Zhang, X., Wang, T., Wang, J., Tang, G., and Zhao, L. (2020). Pyramid channel-based feature attention network for image dehazing. *Comput. Vision Image Understanding* 197, 103003. doi: 10.1016/j.cviu.2020.103003

Zhang, X., Yang, W., Hu, Y., and Liu, J. (2018). "Dmcnn: Dual-domain multi-scale convolutional neural network for compression artifacts removal," in *2018 25th IEEE International Conference on Image Processing (ICIP)*. (Athens, Greece: IEEE), 390–394.

Zhang, K., Zuo, W., Gu, S., and Zhang, L. (2017a). "Learning deep cnn denoiser prior for image restoration," in *Proceedings of the IEEE conference on computer vision and pattern recognition*. (Honolulu, HI, USA), 3929–3938.

Zhang, K., Zuo, W., and Zhang, L. (2017b). FFDNet: Toward a fast and flexible solution for CNN-based image denoising. *IEEE Trans. Image Process.* 27(9), 4608–4622. doi: 10.1109/TIP.2018.2837275

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 14 April 2023
DOI 10.3389/fpls.2023.1158940

OPEN ACCESS

EDITED BY
Yu Xue,
Nanjing University of Information Science and Technology, China

REVIEWED BY
Huiling Chen,
Wenzhou University, China
Ronnie Concepcion II,
De La Salle University, Philippines

*CORRESPONDENCE
Zhenghong Yu
honger1983@gmail.com

†These authors have contributed equally to this work and share first authorship

SPECIALTY SECTION
This article was submitted to Sustainable and Intelligent Phytoprotection, a section of the journal Frontiers in Plant Science

RECEIVED 04 February 2023
ACCEPTED 27 March 2023
PUBLISHED 14 April 2023

CITATION
Yu Z, Ye J, Li C, Zhou H and Li X (2023) TasselLFANet: a novel lightweight multi-branch feature aggregation neural network for high-throughput image-based maize tassels detection and counting.
Front. Plant Sci. 14:1158940.
doi: 10.3389/fpls.2023.1158940

TasselLFANet: a novel lightweight multi-branch feature aggregation neural network for high-throughput image-based maize tassels detection and counting

Zhenghong Yu[1*†], Jianxiong Ye[1†], Cuina Li[2], Huabing Zhou[3] and Xun Li[3]

[1]College of Robotics, Guangdong Polytechnic of Science and Technology, Zhuhai, Guangdong, China, [2]Meteorological Observation Centre, China Meteorological Administration, Beijing, China, [3]Department of Computer Science and Engineering, Wuhan Institute of Technology, Wuhan, China

Accurately and rapidly counting the number of maize tassels is critical for maize breeding, management, and monitoring the growth stage of maize plants. With the advent of high-throughput phenotyping platforms and the availability of large-scale datasets, there is a pressing need to automate this task for genotype and phenotype analysis. Computer vision technology has been increasingly applied in plant science, offering a promising solution for automated monitoring of a large number of plants. However, the current state-of-the-art image algorithms are hindered by hardware limitations, which compromise the balance between algorithmic capacity, running speed, and overall performance, making it difficult to apply them in real-time sensing field environments. Thus, we propose a novel lightweight neural network, named TasselLFANet, with an efficient and powerful structure for accurately and efficiently detecting and counting maize tassels in high spatiotemporal image sequences. Our proposed approach improves the feature-learning ability of TasselLFANet by adopting a cross-stage fusion strategy that balances the variability of different layers. Additionally, TasselLFANet utilizes multiple receptive fields to capture diverse feature representations, and incorporates an innovative visual channel attention module to detect and capture features more flexibly and precisely. We conducted a series of comparative experiments on a new, highly informative dataset called MrMT, which demonstrate that TasselLFANet outperforms the latest batch of lightweight networks in terms of performance, flexibility, and adaptability, achieving an F1 measure value of 94.4%, a mAP.@5 value of 96.8%, and having only 6.0M parameters. Moreover, compared with the regression-based TasselNetV3-Seg† model, our proposed model achieves superior counting performance, with a mean absolute error (MAE) of 1.80, a root mean square error (RMSE) of 2.68, and a R^2 of 0.99. The proposed model meets the

accuracy and speed requirements of the vision system in maize tassel detection. Furthermore, our proposed method is reliable and unaffected by geographical changes, providing essential technical support for computerized counting in the field.

KEYWORDS

maize tassels, detection and counting, lightweight, feature aggregation, neural network

1 Introduction

Maize has been a crucial agricultural crop for several decades, serving as a primary source of food, feed, fuel (ethanol), and other industrial feedstocks (Andorf et al., 2019). However, the yield of maize kernels is prone to fluctuation due to planting density issues caused by poor germination and planting errors (Gonzalez et al., 2018). Counting maize tassels is therefore an essential task in crop management and breeding, as it provides valuable information on germination rates and population densities for farmers and breeders (Wang et al., 2021). Additionally, monitoring the development of maize tassels offers a reliable foundation for observing the growth cycle of maize (Yu et al., 2013), resulting in improved production and more effective farming practices. This technique can offer practical guidance and early warning to farmers (Bai et al., 2018), making it a powerful tool for crop management.

Traditional manual observation and counting rely on labor-intensive tasks, which are not only time-consuming but also limited in sample size. Fortunately, with the rapid development of computer vision and deep learning, more advanced research methods have emerged for extracting and processing visual information from image data, providing us with better options for efficiently detecting and counting plants (Yu et al., 2017). Building on breakthroughs from the past few decades (He et al., 2015; Russakovsky et al., 2015; Vaswani et al., 2017), the field of visual recognition has entered a new era of large-scale visual representation learning. The performance of visual representation learning systems is largely influenced by three main factors: the chosen neural network architecture, the method used to train the network, and the data used for training. In the field of visual recognition, every advancement in each of these areas contributes to the overall performance improvement. Innovations in neural network architecture design have played an important role in representation learning. Convolutional neural networks (CNNs, See Table 1 for all abbreviations in the paper.) (Chollet, 2017; Shafiq and Gu, 2022) have demonstrated high capability for learning discriminative visual representations and convincingly extended to a range of computer vision tasks such as image recognition, object detection, and semantic segmentation. Currently, related applications in object detection have gradually matured. The field is mainly divided into two-stage algorithms represented by Faster Region-based Convolutional Neural Network (Faster RCNN) (Chen and Gupta, 2017) and one-stage algorithms represented by You Only Look Once (YOLOv3) (Redmon and Farhadi, 2018). Among them, CenterNet (Duan et al., 2019), EfficientDet (Tan et al., 2020), RetinaNet (Lin et al., 2017), Yolov7-tiny (Wang et al., 2022) have achieved SOTA performance in many different fields.

In the field of plant science, various methods have been developed for applications such as counting trees (Li et al., 2016), fruits (Rahnemoonfar and Sheppard, 2017), wheat ears (Madec et al., 2019), and rice panicles (Wang et al., 2022a), with remarkable success. However, despite these advances, identifying maize tassels using visual techniques remains a challenging task due to the following factors:

TABLE 1 Summary of abbreviations.

Abb.	Full name
BN	Batch Normalization
MP	Max Pooling
REP	Re-Parameterization
GMP	Global Max Pooling
GAP	Global Average Pooling
MP-C	Max Pooling-Conv
ECA	Efficient Channel Attention
Mlt-ECA	Multi-Efficient Channel Attention
ELAN	Efficient Layer Aggregation Networks
ELAN-H	Efficient Layer Aggregation Networks-Higher
SPPCSPC	Spatial Pyramid Pooling Cross Stage Partial Connect
PANet	Path Aggregation Network
YOLO	You Only Look Once
CNN	Convolutional Neural Network
R-CNN	Region-based Convolutional Neural Network
UAV	Unmanned Aerial Vehicle
MAE	Mean Absolute Error
RMSE	Root-Mean Square Error
MAPE	Mean Absolute Percentage Error
FPS	Frames Per Second

1) The size, color, and texture of tassels vary depending on the maize cultivar and growth stage. The tassel margins are irregularly shaped, and their color may resemble that of leaves at a particular stage.
2) The illumination changes considerably under different weather conditions, leading to image degradation, wind-induced motion, camera shooting angle variations, and perspective distortion, causing differences in tassel poses.
3) Occlusion is commonly present, making it difficult to count even for human experts. The presence of a cluttered background diversifies and misleads the visual pattern of the maize tassel.

Therefore, there is an urgent need for machine learning models that can detect maize tassels in a wide variety of situations with reliable generalizability. Currently, vision-based research on grain ear numbers is mainly categorized into three groups as follows:

Segmentation-based methods: Segmentation-based methods are commonly used to segment plant or organs in image sequences using phenotypic features such as color and texture. The counting methods are then applied to calculate the number of ears. Several studies have been conducted to improve the accuracy and efficiency of this method for different plant species. For instance, Yu et al. (2016) utilized the spatio-temporal saliency properties of maize tassels and employed a low-rank matrix decomposition method to identify the pixel domain of the tassel. This was followed by image segmentation to extract the tassels. However, this method is only applicable to scenes where the tassels are relatively prominent against the background. Hayat et al. (2020) proposed a rice ear segmentation algorithm based on unsupervised Bayesian learning, which employs a multivariate mixture Gaussian model to represent the probability distribution of pixels. The algorithm achieved an average F1 measurement of 82.10%. Nonetheless, it may not be suitable for maize tassels due to their rich textural characteristics and different colors of different cultivars. Ma et al. (2020) developed an EarSegNet model based on semantic segmentation that integrates an encoder-decoder structure and dilated convolution. The model aims to improve the accuracy and efficiency of winter wheat ear segmentation and achieves an F1 measurement of 87.25%. Yu et al. (2022) proposed an Unmanned Aerial Vehicle (UAV) tassel image recognition algorithm based on the U-Net model. The algorithm combines lightweight and heavy extraction networks, striking a balance between accuracy and speed with a relative mean squared error RMSE of 4.4. Nevertheless, low-level noise can severely disrupt counting after phenotypic segmentation. This may result in errors accumulating and a decrease in accuracy.

Regression-based method: The direct counting by regression network is an alternative method for object counting. Lu et al. (2017) proposed TasselNet, a regression network that directly computes maize tassels. By modeling the local visual characteristics of field images and regressing the local counts of maize tassels, TasselNet can achieve good adaptability to in-field variations. However, this approach may not be as robust as object detectors in later growth stages. To improve counting accuracy and efficiency, the same research group proposed TasselNetV2 and TasselNetV2+ (Xiong et al., 2019; Lu and Cao, 2020). Similarly, Khaki et al. (2022) proposed WheatNet for counting wheat ears, achieving an overall prediction error of 8.7%. Liu et al. (2022) introduced IntegrateNet, a new network for maize stand counting that supervises the learning of density map and local count simultaneously. This approach balances the tradeoff between their errors, resulting in improved model performance. One disadvantage of direct counting by regression networks is that this method only provides ear counts that are as reliable as possible, making it difficult to analyze the ears phenotype accurately after counting.

Object detection-based method: Object detection is a popular approach for counting that involves detecting and drawing bounding boxes. This method not only provides the number of objects but also their size and location. For instance, Liu et al. (2020) used the Faster R-CNN to detect maize tassels in UAV images and achieved a detection accuracy of 89.96%. Ji et al. (2021) proposed an in-field maize tassels detection method that combines light saturation correction and Itti saliency-based systems to detect candidate regions, and false positives are removed using the LS-SVM classifier, resulting in an F1 score of 88.36%. However, object detection methods without deep learning models have relatively poor learning capabilities, which may limit their direct use in other applications. Yang S, et al. (2021) proposed an improved CenterNet that embeds location information in the feature extraction module and increases the detection accuracy to 92.4%. While the above detection methods are designed for specific cultivars in particular environments, they may not accurately distinguish tassels at early tasseling stages, making it difficult to continuously monitor maize breeding requirements. Miao et al. (2021) found that the convolutional neural network-based regression counting method had poor accuracy and high bias for plants with extreme leaf counts, while the count-by-detection method based on the Faster R-CNN object detection model achieved near-human performance for plants where all leaf tips are visible. However, the two-stage detection network used in the count-by-detection method ignores the real-time requirements of field applications. Notably, the YOLO series, another commonly used object detection method, is faster and more efficient than other methods and can meet the practical needs of plant detection and counting problems (Yang B, et al., 2021; Lyu et al., 2022; Zang et al., 2023). However, the accuracy of tassel detection still needs to be improved (Zou et al., 2020).

In recent years, embedded systems have gained significant attention as a promising approach for efficient target detection in crop growth monitoring systems, such as the identification of specific crop features. The real-time processing capabilities of embedded systems can significantly reduce the time delay between image acquisition and detection. To this end, several studies (Gajjar et al., 2021; Saddik et al., 2022) have presented intelligent plant disease diagnosis systems that integrate computer vision and machine learning techniques. However, the accuracy of such systems may be affected by image quality and lighting conditions, and they may require substantial computational resources and training time. In a similar vein, a fruit detection and counting system that leverages embedded systems and convolutional neural networks (CNNs) has been proposed

(Mazzia et al., 2020; Zhang et al., 2022). Although the system can detect and count fruits in real-time, its accuracy may be influenced by factors such as fruit color, size, and shape, and it requires significant computational resources. Therefore, there is a need for vision technologies that can ensure high accuracy while being easy to implement. Specifically, there is a need to address the impact of external environmental conditions and internal growth features on object detection while reducing the computational requirements of embedded systems.

To address these challenges, this paper proposes a new lightweight neural network called TasselLFANet, which is based on the deep learning framework. The network has a faster and stronger structure and a more efficient feature integration method. We enhance the feature-learning ability of the network by using a cross-stage fusion strategy that balances the variability of different layers. Moreover, our method makes use of diverse feature representations with multiple receptive fields and introduces an innovative visual channel attention module to detect and capture features more flexibly and accurately.

Overall, the main contributions of this paper are four-fold:

1) Through a series of comparative experiments, we demonstrate that detection-based methods hold great promise for plant counting applications. Compared to regression-based methods, detection-based methods provide more comprehensive object information, such as position and size, which can inform pre- and post-processing steps.
2) We propose a novel global regression framework for object detection, named TasselLFANet, which leverages an improved attention module to address scale changes of tassels and environmental variations in complex wild-field situations. Our framework achieves high efficiency while maintaining strong performance. In addition, our model has a small number of parameters, making it suitable for real-time detection in embedded systems.
3) We introduce MrMT (Multi-regional Maize Tassels), a highly informative, spatially and temporally continuous dataset containing 1968 images and 96434 corresponding bounding box annotations. This large-scale, high-cost dataset is designed to provide researchers with a more convenient and detailed resource for agricultural research, including crop growth stage detection.
4) Based on the evaluation of the MrMT dataset, we demonstrate that our proposed method outperforms the latest batch of high-performance lightweight networks for object detection. Furthermore, our method surpasses the state-of-the-art TasselNetV3-Seg† model in plant counting performance when compared to regression network-based counting methods.

2 Materials and methods

2.1 Image collection and annotation

We developed an automated ground-based observation system that captures farmland images continuously every day. The digital camera captures eight images per day, one every hour from 9:00 to 16:00. The system was installed in experimental fields located in Tai'an of Shandong province, Zhengzhou of Henan province, and Gucheng of Hebei province, China, as described in Yu et al. (2013). The experimental fields grew various cultivars of maize. The MrMT (Multi-regional Maize Tassels) dataset was collected, containing 12 independent image sequences from the tasseling stage to the flowering stage, totaling 1968 field images, as listed in Table 2. To capture subtle changes, the dataset covers various scales and

TABLE 2 Multi-regional Maize Tassels (MrMT dataset).

Image Sequence	Total Size	Train Size	Valid Size	Test Size
Shandong2010_1	116	71	30	15
Shandong2010_2	117	69	30	18
Shandong2011_1	140	79	30	31
Shandong2011_2	126	61	30	35
Shandong2012_2013	55	–	–	55
Henan2010	221	140	60	21
Henan2011	237	140	60	37
Henan2012	227	140	60	27
Henan2014	240	140	60	40
Hebei2010	225	140	60	25
Hebei2012	220	140	60	20
Hebei2014	44	–	–	44

environments through continuous sampling, which helps to learn network representations based on quantitative data. The MrMT dataset enriches the existing dataset of maize tassels images with more timepoints, as shown in Figure 1. The dataset includes images of maize tassels in different scenarios, such as tasseling stage to flowering stage in multiple time sequences (Figures 1A–F), and example images from different locations (Figures 1G–J), including Hebei2012, Shandong2011_1, Shandong2010_2, and Henan2012. These images highlight the complexity of the maize growth environment, making it challenging for automatic operations.

In accordance with standard annotation paradigms, box-level labeling was manually performed for each maize tassel using the open-source tool Labelimg (Tzutalin, 2022). An example of such labeling is shown in Figure 2. While this work was both expensive and time-consuming, it proved to be meaningful and valuable, resulting in the annotation of a total of 96,434 maize tassels. It is important to note that these labels may contain some level of noise, which can increase with the amount of data and make training the model more challenging. However, our proposed method has demonstrated excellent noise suppression capabilities, as we will show in later sections, and is also able to adapt to the domain of the MrMT dataset.

2.2 Lightweight feature aggregation network (TasslLFANet)

Our proposed method is a one-step global regression framework that directly maps image pixels to bounding boxes, coordinates, and classification scores. The network architecture is designed to be simpler and more efficient, allowing for real-time performance. Our method employs binary cross-entropy loss as the supervisory signal and includes a box regression branch that predicts four coordinates for each box, along with an objectness

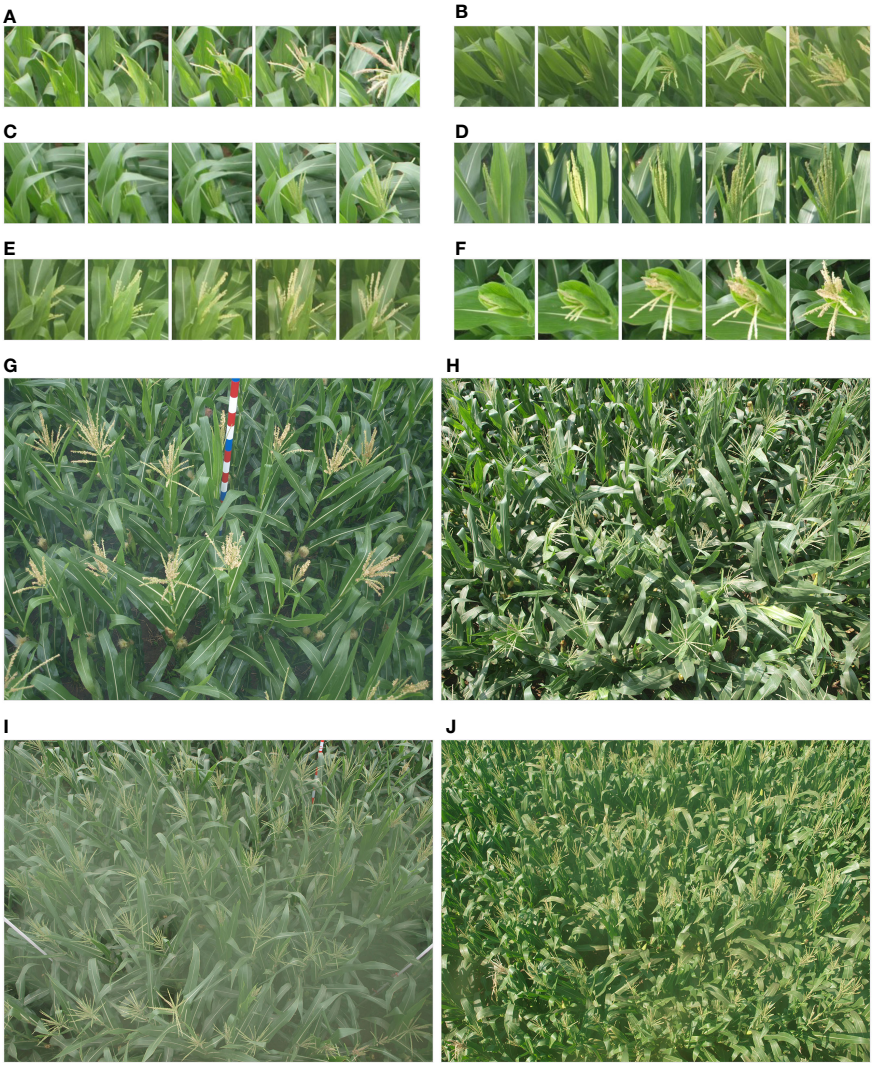


FIGURE 1
Example of field imaging of maize tassels in the MrMT dataset. **(A–F)** present images of tassels from tasseling stage to flowering stage in multiple time sequences; **(G–J)** are example images in different scenarios, respectively from Hebei2012, Shandong2011_1, Shandong2010_2, Henan2012, it can be observed that the growth environment of maize is complex and is challenging for automatic operations.

FIGURE 2
Box-level labeling for MrMT dataset.

score. The objectness score is equal to 1 if the anchor box overlaps with the ground-truth box more than any other anchor box. The overall architecture is depicted in Figure 3 and consists of Encoder, Decoder, and Dense Inference components, which we will explain in detail in the following sections.

2.2.1 Efficiency aggregation encoder

We use the convolution, batch normalization, and SiLU activation functions as basic components of the model. Efficient Layer Aggregation Network (ELAN) (Wang et al., 2022b) and Max Pooling-Conv (MP-C) modules constitute an Encoder for feature extraction. As shown in Figure 4, an image of size of H ×W ×3 is taken as input, the feature maps are performed by multi-dimensional aggregation, and the feature maps are output in two-fold down sampling manner. The encoder can then be divided into four stages, each stage outputting 32, 64, 128, and 256 channels respectively, and the mapping information is used as the encoding set for the input image. ELAN is our important means of feature encoding. Through cross-channel information interaction, it aggregates different feature layers into the output feature map. Such a jump-level structure integrates multi-layer outputs, which enhances the expressiveness and adaptability of the network. By the way, the MP-C down sampling method that combines pooling and convolution is adopted to reduce the image information lost after pooling. Combined with convolution, superficial information from shallow layers and semantic information from deep layers are aggregated to reduce feature dimensionality while retaining useful information, avoiding overfitting to some extent. Finally, we also add an attention module Multi-Efficient Channel Attention (Mlt-ECA), which improves the model's performance. Mlt-ECA is our innovation based on ECA (Wang Q, et al., 2020), which will be described next.

2.2.2 Multi-branch decoder

Multi-Branch Decoder, as a concise and effective decoder, utilizes the correlation between feature maps to enhance feature reuse through multi-dimensional interaction and aggregate rich information flows to enhance feature hierarchy. In order to better separate context information and obtain multi-level receptive fields, we embed the Spatial Pyramid Pooling Cross Stage Partial Connect (SPPCSPC) module at the connection between Encoder and Decoder. The SPPCSPC module uses group convolution, which is efficient for the model, where cross-stage feature fusion strategy and truncated gradient flow have been adopted to improve the

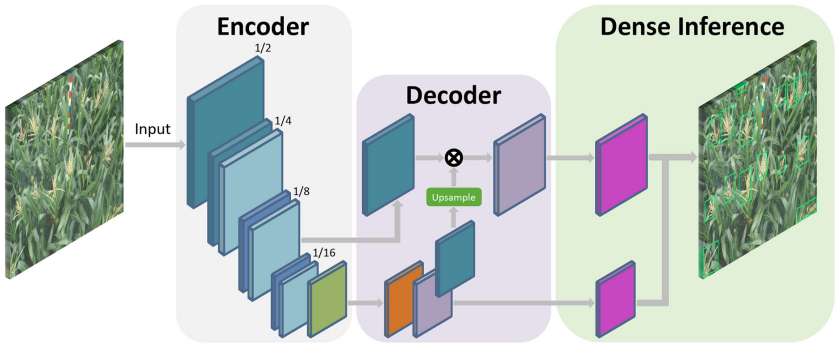


FIGURE 3
TasselLFANet global regression framework (⊗indicates Concat operation).

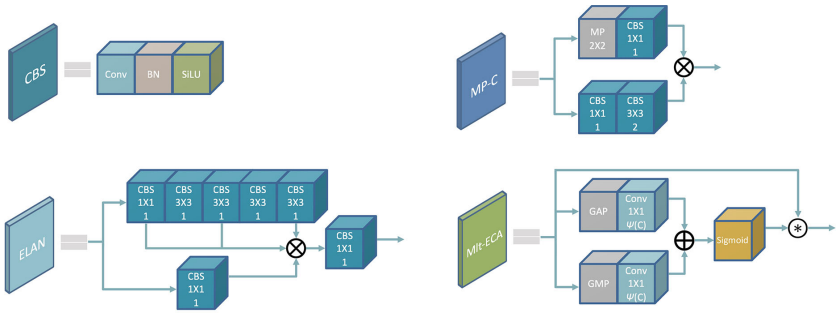


FIGURE 4
Encoder composition module: (CBS/Conv: K ×K, S; MP: S ×S. K ×K represents the convolution kernel size, S represents the step size).

variability of learned features within different layers (Wang et al., 2020), thereby obtaining aggregated information at different scales and enriching the expressive power of feature maps. To cope with the scale change and perspective transformation in the image, we merge the output feature map of the third stage in the Encoder, and increase the spatial dimension between the cascades by using the nearest neighbor upsampling on the Decoder output layer, so as to obtain feature maps that have the same size to the Encoder output. The 1 ×1 convolution is used to map the feature map to the same channel dimension to achieve concat splicing, and the spliced feature map is used as the second branch of the Decoder. The feature information is output after being remapped by the module Efficient Layer Aggregation Networks-Higher (ELAN-H). ELAN-H is an extension of the ELAN module that further enhances network learning capabilities without destroying the original gradient path. Two branches are used for detection on images with different scale transformations. They augment each other and adjust dynamically to improve the detection accuracy for object with multi-scales. Figure 5 shows the components of the Multi-Branch Decoder.

2.2.3 Dense inference

The Dense Inference module aims to merge the sub-images detection results from the information obtained at different stages, and further overlay the encoded information onto the original feature image. In real-world scenarios, training resources are generally relatively abundant, and people are more concerned about the inference overhead and performance, which is why we choose to use the Re-Parameterization (REP) module (Ding et al., 2021). The structure of REP is shown in Figure 5, which is based on a structural reparameterization implementation. Recently, the new concept of reparameterization has become an important topic in network training and object detection (Ding et al., 2022; Hu et al., 2022). During training, the network retains different receptive fields through multi-gradient flow path mapping features to enrich encoding information for better performance gain. During inference, all network layers are transformed into 3×3 convolutions using an Op-fusion strategy to achieve the goal of not increasing computational load, which is useful for network deployment and acceleration.

2.2.4 Mlt-ECA attention module

In recent years, attention mechanisms have been proven to have broad prospects for improving the performance of CNNs (Niu et al., 2021; Guo et al., 2022). They tend to be less computationally expensive than CNNs due to their lower structural complexity and number of parameters. Convolution operations are now an important means of extracting features of objects, and attention mechanism can modify the extracted features while preserving valuable information. This allows the model to ignore the background and pay more attention to the information it needs.

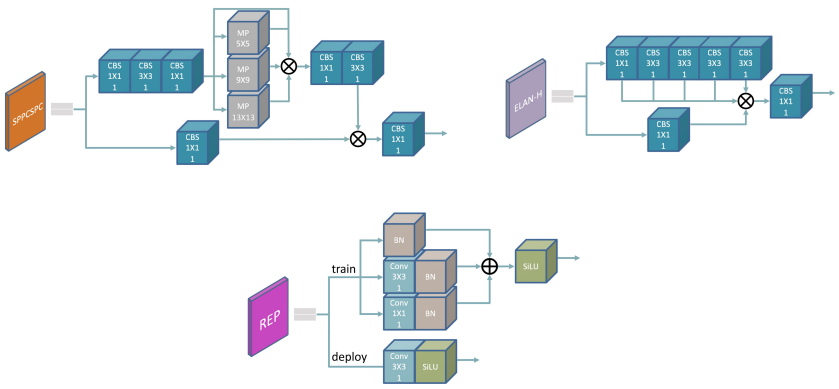


FIGURE 5
Decoder components and REP module (⊗means add operation).

The ECA method (Wang CY, et al., 2020) proposes a local cross-channel interaction strategy without dimensionality reduction. This strategy uses an adaptive one-dimensional convolutional kernel size method to determine the coverage of local cross-channel interactions. Based on the ECA step, we add another branch that considers building the dependencies between channel information in the module's internal feature map. As shown in Figure 4, this branch works similarly to the original ECA branch map, replacing average pooling with max pooling. Channel weights for both the average pooling branch and the global pooling branch are generated by performing a 1D convolution with a kernel size of k, where k is adaptively determined by mapping the channel dimension C and can be expressed as Eq.1. The features of the two channels are then aggregated, and finally output after the Sigmoid function feature recalibration.

$$k = \Psi(C) + \left|\frac{\log_2 C}{\gamma} + \frac{\beta}{\gamma}\right|_{odd} \tag{1}$$

Among them, k represents the size of the convolution kernel, C is the number of channels, odd indicates that k is odd, and γ and β are set to 2 and 1 respectively in the experiment to change the ratio between C and the convolution kernel. To easily distinguish the improved attention mechanism from the original ECA attention module, our proposed module is called Mlt-ECA.

Its working principle can be described by the following formula:

$$Ce(E) = \sigma(CL_k(E_{AP}^C) + CL_k(E_{MP}^C)) \tag{2}$$

$$ME = Ce(E) \circledast E \tag{3}$$

where E_{AP}^C and E_{MP}^C represent the features of average pooling and maximum pooling respectively, CL_k refers to the one-dimensional convolution of each pixel with its nearest neighbor k pixels, σrepresents the Sigmoid function, and $\circledast$ represents element-wise multiplication. ME denotes the output feature map obtained by Mlt-ECA attention.

2.2.5 Loss function

In deep learning, the loss function plays a crucial role in adjusting the weight of the neural network during backpropagation. In our work, we propose two loss functions to guide the bounding box regression. These functions are designed to calculate the error between the predicted values and the ground-truth values during the forward propagation stage, and are used to optimize the network parameters in the training process.

2.2.5.1 Localization loss

We use Loc to compute the localization difference between the prediction box and the ground-truth box for each image. Now let b_{pd}, b_{gt} be the center of the prediction bounding box and the ground-truth bounding box, respectively, then

$$Loc = IoU - \frac{\rho^2(b_{pd}, b_{gt})}{c^2} - \alpha v \tag{4}$$

where ρ is the Euclidean distance between the two centers, c is the diagonal length of the minimum enclosing rectangle of the prediction box and the ground-truth box, and IoU represents the intersection ratio between the ground-truth box and the prediction box. Among them, v is used to measure aspect ratio similarity, and α is the influencing factor of v, which are defined as:

$$v = \frac{4}{\pi^2}\left(arctan\frac{w_{pd}}{h_{pd}} - arctan\frac{w_{gt}}{h_{gt}}\right)^2 \tag{5}$$

$$\alpha = \frac{v}{1 - IoU + v} \tag{6}$$

where w_{pd} and w_{gt} are the widths of prediction boxes and ground-truth box, respectively; h_{pd} and h_{gt} are the heights of prediction boxes and ground-truth box, respectively.

2.2.5.2 Confidence loss and classification loss

Confidence Loss L_{xj}^{obj} and Classification Loss L_{xj}^{cls} use the binary cross-entropy function BCEWithLogitsLoss as supervision to measure the cross-entropy between the target and the output. As for a two-category task, for a sample, it is assumed that the predicted probability of one class is p, and the other class is $1 - p$.

$$P_i = \begin{cases} p, & if \quad y = 1 \\ 1 - p, & otherwise \end{cases} \tag{7}$$

P_i represents the probability that sample i is predicted to be a positive sample, then the loss can be defined as:

$$Lxj = -\frac{1}{N}\sum_i [y_i \ln P_i + (1 - y_i)\ln(1 - P_i)] \tag{8}$$

y_i represents the label of sample i, the y_i is set to 1 when i belongs to the positive class, the negative class is 0, and N represents the total number of samples.

Then, we combine the three losses to get the final loss of TasselLFANet, namely

$$L_{fan} = Loc + Lxj \tag{9}$$

where $Lxj = L_{xj}^{obj} + L_{xj}^{cls}$

3 Experiments and discussions

3.1 Implementation details

In this study, we used a deep learning approach to detect and count maize tassels. Our training dataset consisted of 1120 images randomly selected from the publicly available MrMT dataset. We reserved 480 and 368 images as validation and test sets, respectively. All experiments were conducted on a deep learning framework implemented with PyTorch 1.8 and CUDA 9.0, and executed on an Nvidia Quadro P5000 GPU with 16G of memory. We did not rely on pre-trained model weights during transfer learning to ensure that our model's performance reflected its true potential (Mesnil et al., 2012; He et al., 2019). Given the high resolution of the images in our dataset, we resized them to 640 x 640 pixels. We used the cosine function to schedule the learning rate, which started at 0.01.

The training was performed with stochastic gradient descent (SGD) optimizer with a momentum of 0.937, and lasted for 100 epochs. To prevent overfitting and enhance the robustness of our model, we applied various data augmentation techniques, including color distortion, random translation, random flipping, random scaling, and random stitching.

3.2 Comparison with object detection models

3.2.1 Evaluation metrics

The evaluation metrics of the TasselLFANet model compared to other models are mainly based on precision (P), recall (R), mean precision (mAP) and $F1-measure$, where TP, FP and FN are the number of true positives, false positives and false negatives, respectively. n is 1 because there is only one category (maize tassel) in the data. $mAP@0.5$ represents the average mAP at an IOU threshold of 0.5. $mAP@0.5:0.95$ represents the average of mAP at different IOU thresholds (from 0.5 to 0.95 in step of 0.05).

$$P = \frac{TP}{TP + FP} \tag{10}$$

$$R = \frac{TP}{TP + FN} \tag{11}$$

$$mAP = \frac{\sum_1^n \int_0^1 P(R)d(R)}{n} \tag{12}$$

$$IOU = \frac{Area\ \ of\ \ Overlap}{Area\ \ of\ \ Union} \tag{13}$$

$$F1 - measure = 2\frac{PR}{P + R} \in [0, 1] \tag{14}$$

3.2.2 Ablation study

Since attention is a plug-and-play module, we examine the impact of adding attention at various locations within the model architecture, conducting three different ablation implementations of ECA and Mlt-ECA. These implementations allow us to assess the benefits of integrating attention modules within the Encoder and after the Decoder has decoded the output layer. Our experimental results, as summarized in Table 3, demonstrate that Mlt-ECA consistently outperforms ECA when added in different positions. Moreover, we observe that the effect of adding after Decoder decoding the output layer is better than adding at two positions, and the highest performance is achieved when adding Mlt-ECA after the Encoder. Notably, our modified TasselLFANet with Mlt-ECA achieves an F1 score that is 0.4% higher than the original TasselLFANet, with P increasing by 0.6% to 0.946 and R increasing by 0.4% to 0.942. These findings indicate that Mlt-ECA is more robust and effective in suppressing background information, enabling the model to focus on foreground features and enhance high-level semantic understanding.

3.2.3 Comparing experimental results

Based on the MrMT dataset, we compare TasselLFANet with six state-of-the-art lightweight object detection methods according to the actual application needs, including SSD (Liu et al., 2016), RetinaNet (Lin et al., 2017), CenterNet (Duan et al., 2019), EfficientDet (Tan et al., 2020), Yolox-nano (Ge et al., 2021b) and Yolov7-tiny (Wang et al., 2022a). We train and test these networks using the same training and test sets. A comparison of each method is shown in the Table 4.

Obviously, our proposed model outperforms the other methods in different dimensions, and the model has fewer parameters. In general, as R increases, P decreases accordingly, and our model has a more robust trade-off. SSD, as a traditional one-stage object detection model without bounding boxes generation, uses smaller convolutional filters for dense sampling, which enables simple end-to-end training even on low-resolution input images. However, low feature layers and low number of convolutions will lead to insufficient extraction of shallow feature map information, making it difficult to meet the detection of small-scale tassels. RetinaNet and EfficientDet are good anchor-based object detection models that require extraction of candidate anchor points before making predictions. But they are also likely to miss small and dense objects due to background interference, large overlap of maize tassels and complex growth environment. The YOLO series of object detection algorithms are famous and well-balanced in speed and accuracy. Yolox-nano and Yolov7-tiny are

TABLE 3 Ablation experiments of ECA attention and Mlt-ECA attention.

Model	Encoder	Decoder	P	R	F1	mAP@.5	mAP@.5:.95	Param
TasselLFANet-original	–	–	0.940	0.938	0.938	**0.968**	0.543	6.0M
ECA	✓	✓	0.942	0.934	0.938	0.963	0.524	6.0M
MltECA			0.945	0.940	0.942	**0.968**	0.54	6.0M
ECA	–	✓	0.941	0.936	0.938	0.964	0.526	6.0M
MltECA			**0.948**	0.940	**0.944**	**0.968**	0.541	6.0M
ECA	✓	–	**0.948**	0.938	0.943	**0.968**	0.545	6.0M
MltECA			0.946	**0.942**	**0.944**	**0.968**	**0.546**	6.0M

"✓" means joining the corresponding module and the best performance is in boldface.

TABLE 4 Comparison of evaluation indicators of different models.

Model	Test Size	Backbone	P	R	F1	mAP@.5	mAP@.5:.95	Param
SSD	300	VGG-16	0.785	0.517	0.623	0.711	0.241	90.6M
RetinaNet	600	ResNet-50	0.914	0.719	0.805	0.828	0.376	138.0M
CenterNet	640	ResNet-50	0.874	0.928	0.900	0.902	0.452	124.0M
EfficientDet	512	EfficientNet-B0	0.914	0.712	0.800	0.825	0.444	15.1M
Yolox-nano	640	CSPDarknet-53	0.852	0.872	0.862	0.894	0.437	**3.7M**
Yolov7-tiny	640	E-ELAN	0.921	0.878	0.899	0.938	0.434	11.7M
TasselLFANet	640	ELAN	**0.946**	**0.942**	**0.944**	**0.968**	**0.546**	6.0M

The best performance is in boldface.

state-of-the-art detection models that use multi-scale information in combination with the path aggregation network PANet (Liu S, et al., 2018) to enhance the feature hierarchy, which significantly improves the detection accuracy, while at the same time it is clear that the parameter size is small. CenterNet employs keypoint estimation to find center points and regress other target attributes. This detection method is novel and specific, and has a higher recall rate. It should be noted that when the *IOU* increases above 0.5, TasselLFANet has a higher *mAP* compared to the other methods, 9.4% higher than CenterNet, the second-ranked model, indicating that the model has more accurate positioning performance. Therefore, our model can achieve the best results.

3.2.4 Maize tassels detection at different resolutions

Next, we explore the relationship between image resolution, F1-measure and frames per second (FPS) in image scaling. Speeds are measured on a desktop computer with a single Nvidia Quadro P5000 GPU (16G) at different resolutions and FPS averaged over 100 forward passes. Taking the image resolution of 640 ×640 as the benchmark, the results are shown in Figure 6. It can be seen that the F1-measure value shows a certain attenuation tendency as the image resolution increases or decreases. In general, increasing the image resolution of the input layer may potentially result in a higher fine-grained feature representation. The higher the resolution, the better the accuracy below a certain threshold. Beyond this threshold, the improvement in accuracy begins to decline. We believe that the reason for the attenuation is that the scale distribution of samples in our dataset tends towards medium-resolution and the gain from increasing resolution tends towards saturation. At the same time, high-resolution images increase the amount of computation, and image resolution is negatively correlated with FPS.

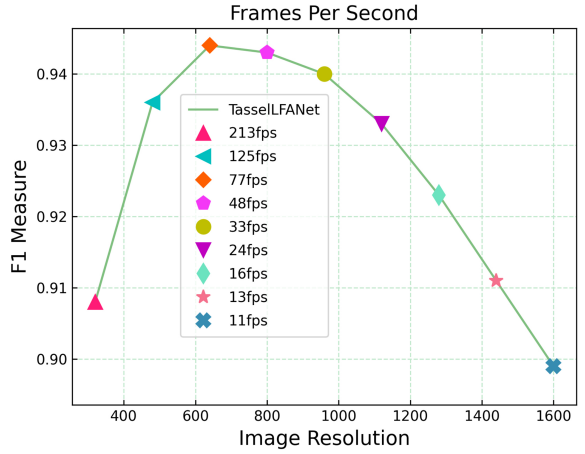


FIGURE 6
Relationship between image resolution, F1-measure and frames per second (FPS) in image scaling.

Please note that different devices and methods have been used to capture field images, so there may be differences in resolution. Therefore, it is important that the proposed method is robust to images with different resolutions. As shown in Figure 6, when the image resolution is reduced to 480 and 320, the F1-measure obtains high-efficiency fps of 125 and 213 respectively, showing high performance of 0.936 and 0.908. Based on this performance, we introduce a high-efficiency model TasselLFANet-HE with a resolution of 480 ×480 pixels. Its overall performance is shown in Table 5.

3.2.5 Overall performance evaluation

To further evaluate the effectiveness of the model's overall performance, we compared TasselLFANet and TasselLFANet-HE with other detection methods. Figure 7 shows a comparison of each model with respect to F1-measure and running speed. In brief, the TasselLFANet series model outperforms all other methods. Yolov7-tiny is faster than TasselLFANet. Although TasselLFANet has a lower parameter volume than Yolov7-tiny, TasselLFANet is less efficient than Yolov7-tiny due to feature aggregation reusability, which is to some extent expected. In addition, TasselLFANet-HE far surpasses other methods in speed, and achieves high performance that only TasselLFANet surpasses. In the future, even more speeds can be achieved with more advanced hardware devices.

3.2.6 Domain adaptation comparison

From the MrMT dataset, we selected a number of representative plant sample images such as sparse tassel(a), blurred image (b), strong illumination (c), severe occlusion (d), dense object (e), to test the domain adaptation of each model. As shown in Figure 8, the

TABLE 5 Performance comparison between TasselLFANet-HE and TasselLFANet.

Method	Test Size	P	R	F1	mAP@.5	mAP@.5:.95	FPS
TasselLFANet	640	0.946	**0.942**	**0.944**	**0.968**	**0.546**	77
TasselLFANet-HE	480	**0.947**	0.926	0.936	0.962	0.518	**125**

The best performance is in boldface.

visualization results further confirm the strong adversariality of our method in the face of scale transformation and environmental disturbance. The green box in the figure is the detection result, and the white numbers in the upper left corner of each image are the counting results. The first row in the Figure is the ground-truth result, and the rest rows are the prediction results of the eight models.

Figure 8A shows that all models can produce reasonable approximations to ground-truth counts when individual tassels are clearly visible. However, under the complex environment shown in Figures 8B, C, especially in Figure 8C, the image noise increases with the increasing external illumination, and the prediction errors of other models increases. In Figure 8D, the tassels are occluded by the background, and although our algorithm still performs relatively well, it is difficult to provide a targeted solution for detection. In this regard, we propose to improve the data sampling method, guide the model to learn the internal characteristics contained in the dataset with a tendency, and use a more effective strategy for positive and negative sample distribution (Zhu et al., 2020; Ge et al., 2021a; Li et al., 2022). As shown in Figure 8E, the SSD, RetinaNet and EfficientDet algorithms are weak in representation for shallow feature maps, and these algorithms are not robust when tassels are dense and small in size.

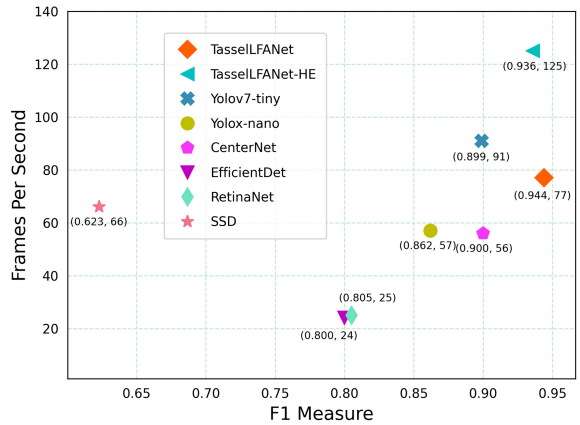


FIGURE 7
Performance and speed comparison of different models.

Significantly, our model performs well in this scene, with improvements that allow for more subtle image capture of tassels.

3.3 Comparing against the state-of-the-art counting method

3.3.1 Evaluation metrics

Four metrics including mean absolute error (MAE), root-mean square error (RMSE), mean absolute percentage error (MAPE), and the coefficient of determination (R^2) were used to evaluate the agreement between the predicted and ground-truth values. To be more specific, these metrics are computed by Eq. (15-18).

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|\hat{y}_i - y_i| \tag{15}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{16}$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{\hat{y}_i - y_i}{y_i}\right| \tag{17}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2} \tag{18}$$

To further demonstrate the superiority of our model, we compared all above-mentioned object detection methods with the state-of-the-art counting method, local density regression-based TasselNetV3-Seg† (Lu et al., 2021). The test images are 181 pictures randomly selected from the MrMT test set. We also report results on the number of model parameters and frames per second (FPS), as indicators of computational complexity. Table 6 shows the counting performance of all models related to the evaluation metrics.

3.3.2 Performance comparison

Numerical comparisons in Table 6 show that our proposed TasselLFANet series method has the best overall counting performance and leads TasselNetV3-Seg† in an all-round way. As the most advanced object detection algorithm, Yolov7-tiny is second

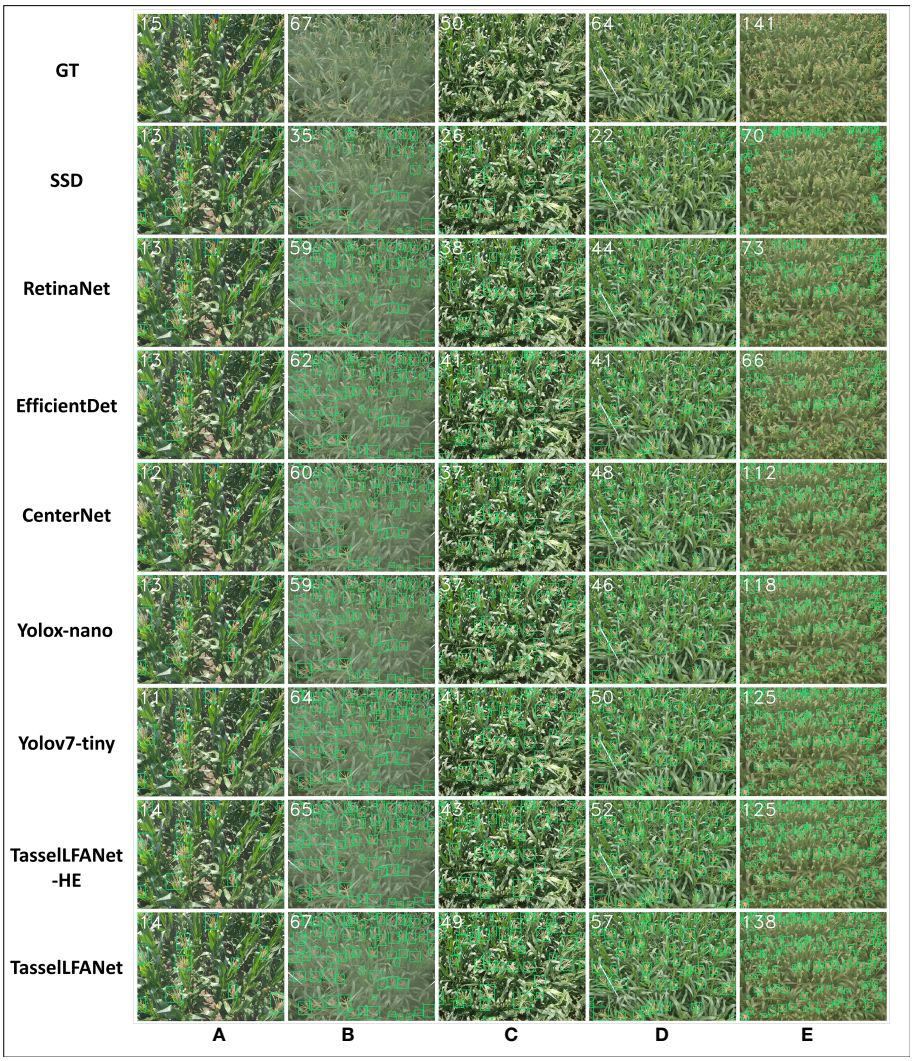


FIGURE 8
MrMT data set visualization results. The representative plant sample images for testing the domain adaptation of each model. GT means ground-truth count. **(A)** sparse tassel; **(B)** blurred image; **(C)** strong illumination; **(D)** severe occlusion; **(E)** dense object..

TABLE 6 Counting performance comparison of different models.

Model	Param	FPS	MAE	RMSE	MAPE	R^2
SSD	90.6M	66	9.04	11.99	34.3%	0.7477
EfficientDet	15.1M	24	4.22	6.65	19.8%	0.9225
RetinaNet	138.0M	25	4.04	6.42	19.7%	0.9278
TasselNetV3-Seg†	7.5M	36	3.75	4.68	27.5%	0.9632
Yolox-nano	3.7M	57	4.02	5.15	23.2%	0.9534
CenterNet	124.0M	56	2.87	3.85	17.8%	0.9740
Yolov7-tiny	11.7M	91	3.10	4.14	17.3%	0.9699
TasselLFANet-HE	6.0M	**125**	2.70	3.76	14.3%	0.9751
TasselLFANet	6.0M	77	**1.80**	**2.68**	**9.2%**	**0.9903**

The best performance is in boldface.

only to TasselLFANet in terms of comprehensive performance. TasselNetV3-Seg† is less efficient than Yolox-nano, and Yolox-nano has smaller parameters, because deep convolutions make poorer use of computational units than standard convolutions.

3.3.3 Coefficients of determination between different models and manual counting

We plot the linear regression relationship between manual counts and experimental algorithms on the MrMT dataset as shown in Figure 9. Presented by scatter diagram distribution, it shows that our model is able to capture changes with better robustness and generalization. The key lies in the dynamic interaction and adaptive adjustment of multiple branches of aggregate information flow, improving the performance of the model in sparse and dense scenes. Yet, the issue of occlusion has been a hot area of research in detection methods. If there are a lot of maize tassels in the image, many predicted bounding boxes will be filtered and underestimated by the non-maximum suppression of the detector. In contrast, TasselNetV3-Seg† can capture additional global information. Even though, our model achieves superior performance compared to TasselNetV3-Seg† which has large errors around the counting baseline.

4 Discussion

Computer vision technology has promoted the development of agricultural systems. However, plant counting is not a completely solved problem. We list the advantages and disadvantages of each comparative algorithm as well as their inference time and other parameters in Table 7. Detecting objects in natural canopy images with a large number of maize tassels is complex and challenging, as the size, shape, and color of maize tassel at different growth stages vary, and the degree of interference from external information is different. As shown in Figure 8D, when the object to be measured is occluded by background information, many underestimations can be observed, which is a popular problem in detection methods that has been continuously improved. Similarly, the scatter plot presented in Figure 9 illustrates more intuitively that the seemingly good image level mean absolute error (MAE) is usually dominated by underestimations in the detection performance. This is also related to the fact that too little contextual information was considered when constructing the model. The design of lightweight convolutional neural networks aims to achieve faster inference for mobile device applications. However, convolution operations can only capture local information in the window area, which hinders performance improvement in certain scenarios. Introducing self-attention or guided attention into convolution can capture global information well. Guided learning methods can be considered. The principle is to preprocess the incoming image in a way such as training stage to guide the neural network to learn important information more effectively. For example, in maize tassel detection, most of the time, green leaves in the field are the main background information. When weeds inevitably interfere with tassel detection, because the neural network has learned a lot about how to counter information dominated by green leaves and

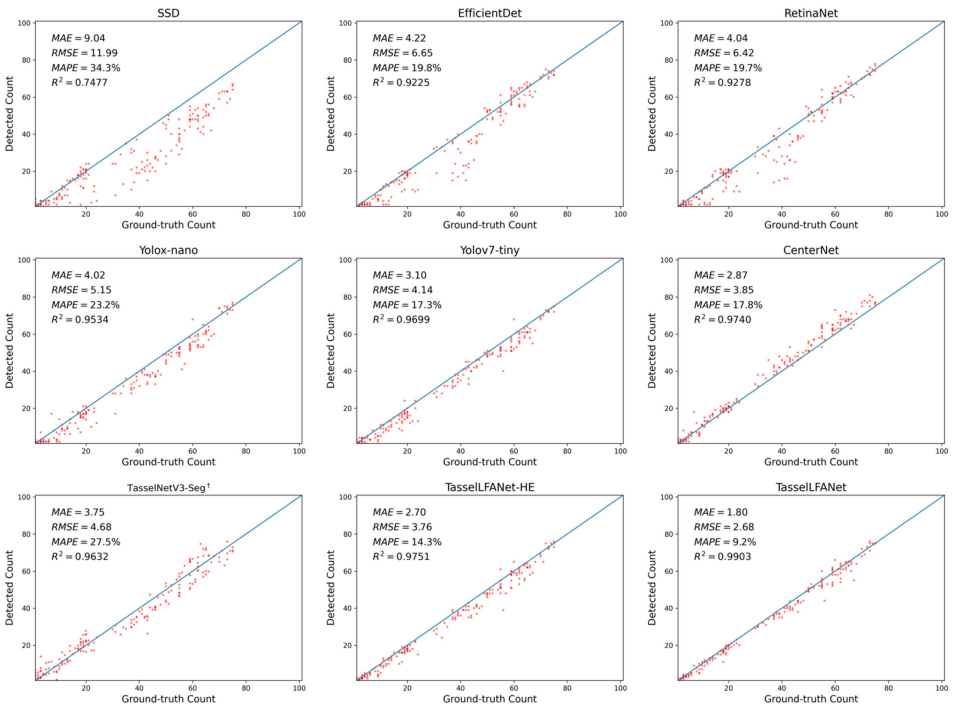


FIGURE 9
Scatter plots of the detected count by different models versus their ground-truth count.

TABLE 7 Characteristics of different models.

Model	Param (M)	Inference Time (s)	Advantage	Disadvantage
SSD	90.6	0.0152	High real-time performance, easy to train	Many parameters, poor generalization ability
EfficientDet	15.1	0.0417	Fewer parameters, easier to train	Lower real-time performance, lower detection accuracy for dense and small targets
RetinaNet	138.0	0.0401	Has multi-scale detection capability	Lower real-time performance, lower detection accuracy for dense and small targets
TasselNetV3-Seg†	7.5	0.0278	Has global modeling capability	Lower real-time performance
Yolox-nano	3.7	0.0172	High real-time performance, easy to train	Lower multi-scale detection accuracy
CenterNet	124.0	0.0179	Has global modeling capability	Weak domain adaptation ability
Yolov7-tiny	11.7	0.0111	High real-time performance	Weak domain adaptation ability
LFANet-HE	6.0	0.0081	High real-time performance	Domain adaptability is relatively weak
LFANet	6.0	0.0129	Balanced real-time performance and accuracy, strong domain adaptation ability	Parameters need to be improved

ignored learning about weeds and other interfering information, this will lead to a significant increase in the competitiveness of weeds and filaments in interfering with tassel detection during inference, as shown in Figure 10. In the early visual applications in agriculture, Yu et al. (2013) found that the color distribution of monochrome objects changes with brightness on the chromaticity saturation plane, and proposed a crop segmentation method AP-HI, which is insensitive to outdoor brightness and complex environmental elements and can adaptively extract crops. If this adaptive method is applied to suppress some green leaf information in the training images, the learning level of the neural network on green leaves can be reduced, thereby guiding the neural network to learn more about how to counter weeds and other interfering information, enhance the feature competition of the original model, and capture more discriminative feature information. To some extent, this is similar to the visual attention mechanism, which guides the neural network to focus more on certain features by learning the importance of information between features (Hu et al., 2018). Masking is direct method as another option. Pei et al. (2022) used an improved Yolov4 to detect and mask maize rows, and then detected weeds in the masked image. This preprocessing method reduces the imbalance between positive and negative samples and effectively improves detection accuracy.

In addition, while previous works mainly relied on larger backbone networks or higher-resolution input images to achieve higher accuracy, balancing all dimensions of network width/depth/resolution is crucial when considering both accuracy and efficiency (Tan and Le, 2019; Dollár et al., 2021; Tan and Le, 2021; Wang et al., 2022). Recent studies have shown that carefully designed lightweight networks can achieve comparable performance to their heavy counterparts with much less computational cost (Howard et al., 2017; Tan et al., 2020). Moreover, efficient network design can be further improved by leveraging knowledge distillation, which involves transferring the knowledge learned by a complex model to a simpler one (Hinton et al., 2015; Zagoruyko and Komodakis, 2017; Liu H, et al., 2018; Chen et al., 2020). In agriculture applications, this approach can enable the deployment of more lightweight models on edge devices with limited computational resources, while still maintaining satisfactory accuracy.

Overall, computer vision techniques have greatly facilitated the development of agricultural systems, but challenges remain in

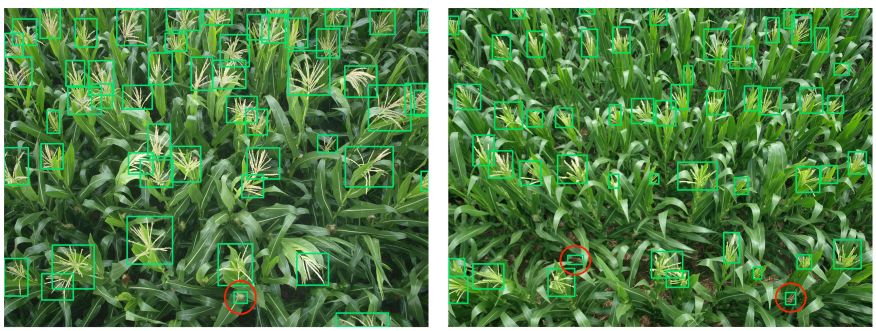


FIGURE 10
Interference from weeds and maize silks in the accurate detection of maize tassels (red circles indicate false detection).

accurately counting and detecting objects in complex and cluttered natural scenes. Various strategies such as introducing self-attention or guided attention into convolution operations, incorporating adaptive preprocessing methods, and leveraging efficient network design and knowledge distillation can be explored to improve detection performance and enhance the robustness of agricultural computer vision systems.

5 Conclusion

At this stage, designing a lightweight, effective, and easily implementable deep neural network for agricultural application scenarios is both challenging and important. In this study, we propose a novel neural network, TasselLFANet, for accurate and efficient detection and counting of maize tassels in high spatiotemporal image sequences. Our experiments confirm the promising performance of our proposed method compared to existing lightweight models, demonstrating superior performance, flexibility, and adaptability. In practical applications, we offer the following recommendations:

1) For optimal performance, speed, and robustness, we recommend using TasselLFANet.
2) For resource-constrained hardware devices, TasselLFANet-HE is recommended due to its optimal efficiency and reasonable accuracy in most cases.
3) For scenes with strong features during the early stages of plant growth, we recommend using TasselLFANet, which has higher sensitivity to the tasseling stage.
4) In training, enriching the size and environmental diversity of plant imaging can result in a more robust model that adapts to scale and environment transformations.

For future research, we plan to further deploy TasselLFANet in combination with machinery and equipment in agricultural and natural environments, and make targeted improvements based on feedback to meet higher standards and more demanding scenarios. The integration of computer vision technology with agriculture has promoted the advancement of smart agriculture and improved agricultural ecology. We expect that the dataset we release will attract the attention of researchers and facilitate further research on automated monitoring of plant growth. We believe that by collaborating across different fields, we can take the combination of computer vision technology and agriculture to the next level.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding author.

Author contributions

ZY and JY contributed to conception and design of the study. CL and XL provided the fields and collected the images. ZY and JY performed the data processing and modelling. JY wrote the first draft of the manuscript. ZY wrote sections of the manuscript. HZ edited the manuscript extensively. ZY and HZ participated in project management and obtained the funding for this study. All authors contributed to the article and approved the submitted version.

Funding

This work was supported in part by 2022 key scientific research project of ordinary universities in Guangdong Province under Grant 2022ZDZX4075, in part by 2022 Guangdong province ordinary universities characteristic innovation project under Grant 2022KTSCX251, in part by the Collaborative Intelligent Robot Production & Education Integrates Innovative Application Platform Based on the Industrial Internet under Grant 2020CJPT004, in part by 2020 Guangdong Rural Science and Technology Mission Project under Grant KTP20200153, in part by the Engineering Research Centre for Intelligent equipment manufacturing under Grant 2021GCZX018 and in part by the Guangke & Sany Marine Industry Collaborative Innovation Center in part by the National Natural Science Foundation of China under Grant 62171327.

Acknowledgments

The authors would like to thank the Wuxi Institute of Radio Science and Technology for developing and providing facilities and equipment, Rujun Wu, Yueming Wu, Kangqing Pan for assistance in annotating the MrMT dataset, and Yangxu Wang for assistance in editing and updating the code.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2023.1158940/full#supplementary-material

References

Andorf, C., Beavis, W. D., Hufford, M., Smith, S., Suza, W. P., Wang, K., et al. (2019). Technological advances in maize breeding: past, present and future. *Theor. Appl. Genet.* 132, 817–849. doi: 10.1007/s00122-019-03306-3

Bai, X., Cao, Z., Zhao, L., Zhang, J., Lv, C., Li, C., et al. (2018). Rice heading stage automatic observation by multi-classifier cascade-based rice spike detection method. *Agric. For. Meteorology* 259, 260–270. doi: 10.1016/j.agrformet.2018.05.001

Chen, X., and Gupta, A. (2017). An implementation of faster rcnn with study for region sampling. *arXiv preprint arXiv* 1702, 2138. doi: 10.48550/arXiv.1702.02138

Chen, T., Liu, L., Yang, M., Zhang, M., and Yang, J. (2020). DetNAS: Neural architecture search on object detection. *arXiv preprint arXiv* 2003, 09950. doi: 10.48550/arXiv.1903.10979

Chollet, F. (2017). "Xception: Deep learning with depthwise separable convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 1251–1258. doi: 10.48550/arXiv.1610.02357

Ding, X., Zhang, X., Han, J., and Ding, G. (2022). "Scaling up your kernels to 31x31: Revisiting large kernel design in cnns," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 11963–11975. doi: 10.48550/arXiv.2203.06717

Ding, X., Zhang, X., Ma, N., Han, J., Ding, G., and Sun, J. (2021). "Repvgg: Making vgg-style convnets great again," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 13733–13742. doi: 10.48550/arXiv.2101.03697

Dollár, P., Singh, M., and Girshick, R. (2021). "Fast and accurate model scaling," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 924–932. doi: 10.48550/arXiv.2103.06877

Duan, K., Bai, S., Xie, L., Qi, H., Huang, Q., and Tian, Q. (2019). "Centernet: Keypoint triplets for object detection," in *Proceedings of the IEEE/CVF international conference on computer vision*. 6569–6578. doi: 10.48550/arXiv.1904.08189

Gajjar, R., Gajjar, N., Thakor, V. J., Patel, N. P., and Ruparelia, S. (2021). Real-time detection and identification of plant leaf diseases using convolutional neural networks on an embedded platform. *Visual Comput.*, 1–16. doi: 10.1007/s00371-021-02164-9

Ge, Z., Liu, S., Li, Z., Yoshie, O., and Sun, J. (2021a). "Ota: Optimal transport assignment for object detection," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 303–312. doi: 10.48550/arXiv.2103.14259

Ge, Z., Liu, S., Wang, F., Li, Z., and Sun, J. (2021b). Yolox: Exceeding yolo series in 2021. *arXiv preprint arXiv* 2107, 8430. doi: 10.48550/arXiv.2107.08430

Gonzalez, V. H., Tollenaar, M., Bowman, A., Good, B., and Lee, E. A. (2018). Maize yield potential and density tolerance. *Crop Sci.* 58 (2), 472–485. doi: 10.2135/cropsci2016.06.0547

Guo, M. H., Xu, T. X., Liu, J. J., Liu, Z. N., Jiang, P. T., Mu, T. J., et al. (2022). Attention mechanisms in computer vision: A survey. *Comput. Visual Media* 8 (3), 331–368. doi: 10.1007/s41095-022-0271-y

Hayat, M. A., Wu, J., and Cao, Y. (2020). Unsupervised Bayesian learning for rice panicle segmentation with UAV images. *Plant Methods* 16 (1), 1–13. doi: 10.1186/s13007-020-00567-8

He, K., Girshick, R., and Dollár, P. (2019). "Rethinking imagenet pre-training," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 4918–4927. doi: 10.48550/arXiv.1811.08883

He, K., Zhang, X., Ren, S., and Sun, J. (2015). "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification," in *Proceedings of the IEEE international conference on computer vision*. 1026–1034. doi: 10.48550/arXiv.1502.01852

Hinton, G., Vinyals, O., and Dean, J. (2015). Distilling the knowledge in a neural network. *arXiv preprint arXiv* 1503, 2531. doi: 10.48550/arXiv.1503.02531

Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., et al. (2017). MobileNets: Efficient convolutional neural networks for mobile vision applications. *arXiv preprint arXiv* 1704, 4861. doi: 10.48550/arXiv.1704.04861

Hu, M., Feng, J., Hua, J., Lai, B., Huang, J., Gong, X., et al. (2022). "Online convolutional re-parameterization," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 568–577. doi: 10.48550/arXiv.2204.00826

Hu, J., Shen, L., and Sun, G. (2018). "Squeeze-and-excitation networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 7132–7141. doi: 10.48550/arXiv.1709.01507

Ji, M., Yang, Y., Zheng, Y., Zhu, Q., Huang, M., and Guo, Y. (2021). In-field automatic detection of maize tassels using computer vision. *Inf. Process. Agric.* 8 (1), 87–95. doi: 10.1016/j.inpa.2020.03.002

Khaki, S., Safaei, N., Pham, H., and Wang, L. (2022). Wheatnet: A lightweight convolutional neural network for high-throughput image-based wheat head detection and counting. *Neurocomputing* 489, 78–89. doi: 10.48550/arXiv.2103.09408

Li, W., Fu, H., Yu, L., and Cracknell, A. (2016). Deep learning-based oil palm tree detection and counting for high-resolution remote sensing images. *Remote Sens.* 9 (1), 22. doi: 10.3390/rs9010022

Li, S., He, C., Li, R., and Zhang, L. (2022). "A dual weighting label assignment scheme for object detection," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 9387–9396. doi: 10.48550/arXiv.2203.09730

Lin, T. Y., Goyal, P., Girshick, R., He, K., and Dollár, P. (2017). "Focal loss for dense object detection," in *Proceedings of the IEEE international conference on computer vision*. 2980–2988. doi: 10.48550/arXiv.1708.02002

Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., et al. (2016). "Ssd: Single shot multibox detector," in *Computer Vision–ECCV 2016: 14th European Conference*, Amsterdam, The Netherlands, October 11–14, 2016. 21–37 (Springer International Publishing). doi: 10.48550/arXiv.1512.02325

Liu, Y., Cen, C., Che, Y., Ke, R., Ma, Y., and Ma, Y. (2020). Detection of maize tassels from UAV RGB imagery with faster r-CNN. *Remote Sens.* 12 (2), 338. doi: 10.3390/rs12020338

Liu, S., Qi, L., Qin, H., Shi, J., and Jia, J. (2018). "Path aggregation network for instance segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 8759–8768. doi: 10.48550/arXiv.1803.01534

Liu, H., Simonyan, K., and Yang, Y. (2018). DARTS: Differentiable architecture search. *arXiv preprint arXiv* 1806, 9055. doi: 10.48550/arXiv.1806.09055

Liu, W., Zhou, J., Wang, B., Costa, M., Kaeppler, S. M., and Zhang, Z. (2022). "IntegrateNet: A deep learning network for maize stand counting from UAV imagery by integrating density and local count maps," in *IEEE Geoscience and Remote Sensing Letters*, Vol. 19. 1–5. doi: 10.1109/LGRS.2022.3186544

Lu, H., and Cao, Z. (2020). TasselNetV2+: A fast implementation for high-throughput plant counting from high-resolution RGB imagery. *Front. Plant Sci.* 11. doi: 10.3389/fpls.2020.541960

Lu, H., Cao, Z., Xiao, Y., Zhuang, B., and Shen, C. (2017). TasselNet: counting maize tassels in the wild *via* local counts regression network. *Plant Methods* 13 (1), 1–17. doi: 10.48550/arXiv.1707.02290

Lu, H., Liu, L., Li, Y. N., Zhao, X. M., Wang, X. Q., and Cao, Z. G. (2021). "TasselNetV3: Explainable plant counting with guided upsampling and background suppression," in *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 60. 1–15. doi: 10.1109/TGRS.2021.3058962

Lyu, S., Li, R., Zhao, Y., Li, Z., Fan, R., and Liu, S. (2022). Green citrus detection and counting in orchards based on YOLOv5-CS and AI edge system. *Sensors* 22 (2), 576. doi: 10.3390/s22020576

Ma, J., Li, Y., Liu, H., Du, K., Zheng, F., Wu, Y., et al. (2020). Improving segmentation accuracy for ears of winter wheat at flowering stage by semantic segmentation. *Comput. Electron. Agric.* 176, 105662. doi: 10.1016/j.compag.2020.105662

Madec, S., Jin, X., Lu, H., De Solan, B., Liu, S., Duyme, F., et al. (2019). Ear density estimation from high resolution RGB imagery using deep learning technique. *Agric. For. meteorology* 264, 225–234. doi: 10.1016/j.agrformet.2018.10.013

Mazzia, V., Khaliq, A., Salvetti, F., and Chiaberge, M. (2020). "Real-time apple detection system using embedded systems with hardware accelerators: An edge AI application," in *IEEE Access*, Vol. 8. 9102–9114. doi: 10.1109/ACCESS.2020.2964608

Mesnil, G., Dauphin, Y., Glorot, X., Rifai, S., Bengio, Y., Goodfellow, I., et al. (2012). "Unsupervised and transfer learning challenge: a deep learning approach," in *Proceedings of ICML Workshop on Unsupervised and Transfer Learning*. 97–110 (JMLR Workshop and Conference Proceedings).

Miao, C., Guo, A., Thompson, A. M., Yang, J., Ge, Y., and Schnable, J. C. (2021). Automation of leaf counting in maize and sorghum using deep learning. *Plant Phenome J.* 4 (1), e20022. doi: 10.1002/ppj2.20022

Niu, Z., Zhong, G., and Yu, H. (2021). A review on the attention mechanism of deep learning. *Neurocomputing* 452, 48–62. doi: 10.1016/j.neucom.2021.03.091

Pei, H., Sun, Y., Huang, H., Zhang, W., Sheng, J., and Zhang, Z. (2022). Weed detection in maize fields by UAV images based on crop row preprocessing and improved YOLOv4. *Agriculture* 12 (7), 975. doi: 10.3390/agriculture12070975

Rahnemoonfar, M., and Sheppard, C. (2017). Deep count: fruit counting based on deep simulated learning. *Sensors* 17 (4), 905. doi: 10.3390/s17040905

Redmon, J., and Farhadi, A. (2018). Yolov3: An incremental improvement. *arXiv preprint arXiv* 1804, 2767. doi: 10.48550/arXiv.1804.02767

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., et al. (2015). Imagenet large scale visual recognition challenge. *Int. J. Comput. Vision* 115, 211–252. doi: 10.48550/arXiv.1409.0575

Saddik, A., Latif, R., El Ouardi, A., Elhoseny, M., and Khelifi, A. (2022). Computer development based embedded systems in precision agriculture: Tools and application. *Acta Agriculturae Scandinavica Section B—Soil Plant Sci.* 72 (1), 589–611. doi: 10.1080/09064710.2021.2024874

Shafiq, M., and Gu, Z. (2022). Deep residual learning for image recognition: a survey. *Appl. Sci.* 12 (18), 8972. doi: 10.3390/app12188972

Tan, M., Chen, B., Pang, R., Vasudevan, V., and Sandler, M. (2020a). "MnasNet: Platform-aware neural architecture search for mobile," in *IEEE Transactions on Neural Networks and Learning Systems*, Vol. 31. 3408–3422. doi: 10.48550/arXiv.1807.11626

Tan, M., and Le, Q. (2019). "Efficientnet: Rethinking model scaling for convolutional neural networks," in *International conference on machine learning*. 6105–6114. doi: 10.48550/arXiv.1905.11946

Tan, M., and Le, Q. (2021). "Efficientnetv2: Smaller models and faster training," in *International conference on machine learning*. 10096–10106. doi: 10.48550/arXiv.2104.00298

Tan, M., Pang, R., and Le, Q. V. (2020). "Efficientdet: Scalable and efficient object detection," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 10781–10790. doi: 10.48550/arXiv.1911.09070

Tzutalin, D. (2022). *LabelImg is a graphical image annotation tool and label object bounding boxes in images*. Available at: https://github.com/tzutalin/labelImg.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., et al. (2017). Attention is all you need. *Adv. Neural Inf. Process. Syst.* 30, 5998–6008 doi: 10.48550/arXiv.1706.03762

Wang, C. Y., Bochkovskiy, A., and Liao, H. Y. M. (2022a). YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. *arXiv preprint arXiv* 2207, 2696. doi: 10.48550/arXiv.2207.02696

Wang, C. Y., Liao, H. Y. M., Wu, Y. H., Chen, P. Y., Hsieh, J. W., and Yeh, I. H. (2020). "CSPNet: A new backbone that can enhance learning capability of CNN," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops*. 390–391. doi: 10.48550/arXiv.1911.11929

Wang, C. Y., Liao, H. Y. M., and Yeh, I. H. (2022b). Designing network design strategies through gradient path analysis. *arXiv preprint arXiv* 2211, 4800. doi: 10.48550/arXiv.2211.04800

Wang, Q., Wu, B., Zhu, P., Li, P., Zuo, W., and Hu, Q. (2020). "ECA-net: Efficient channel attention for deep convolutional neural networks," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 11534–11542. doi: 10.48550/arXiv.1910.03151

Wang, L., Xiang, L., Tang, L., and Jiang, H. (2021). A convolutional neural network-based method for corn stand counting in the field. *Sensors* 21 (2), 507. doi: 10.3390/s21020507

Wang, X., Yang, W., Lv, Q., Huang, C., Liang, X., Chen, G., et al. (2022). Field rice panicle detection and counting based on deep learning. *Front. Plant Sci.* 2921. doi: 10.3389/fpls.2022.966495

Xiong, H., Cao, Z., Lu, H., Madec, S., Liu, L., and Shen, C. (2019). TasselNetv2: in-field counting of wheat spikes with context-augmented local regression networks. *Plant Methods* 15 (1), 1–14. doi: 10.1186/s13007-019-0537-2

Yang, B., Gao, Z., Gao, Y., and Zhu, Y. (2021). Rapid detection and counting of wheat ears in the field using YOLOv4 with attention module. *Agronomy* 11 (6), 1202. doi: 10.3390/agronomy11061202

Yang, S., Liu, J., Xu, K., Sang, X., Ning, J., and Zhang, Z. (2021). Improved CenterNet based maize tassel recognition for UAV remote sensing image. *Trans. Chin. Soc. Agric. Machinery* 52, 206–212. doi: 10.6041/j.issn.100-1298.2021.09.024

Yu, Z., Cao, Z., Wu, X., Bai, X., Qin, Y., Zhuo, W., et al. (2013). Automatic image-based detection technology for two critical growth stages of maize: Emergence and three-leaf stage. *Agric. For. meteorology* 174, 65–84. doi: 10.1016/j.agrformet.2013.02.011

Yu, Z., Yin, H., Feng, H., Chen, M., Zhou, H., Lu, T., et al. (2016). "An image-based approach to automatic crop organ extraction *via* low-rank matrix recovery," in *2016 15th International Symposium on Parallel and Distributed Computing (ISPDC)*. 376–381. doi: 10.1109/ISPDC.2016.63

Yu, X., Yin, D., Nie, C., Ming, B., Xu, H., Liu, Y., et al. (2022). Maize tassel area dynamic monitoring based on near-ground and UAV RGB images by U-net model. *Comput. Electron. Agric.* 203, 107477. doi: 10.1016/j.compag.2022.107477

Yu, Z., Zhou, H., and Li, C. (2017). Fast non-rigid image feature matching for agricultural UAV *via* probabilistic inference with regularization techniques. *Comput. Electron. Agric.* 143, 79–89. doi: 10.1016/j.compag.2017.10.002

Zagoruyko, S., and Komodakis, N. (2017). Paying more attention to attention: Improving the performance of convolutional neural networks *via* attention transfer. *arXiv preprint arXiv* 1612, 3928. doi: 10.48550/arXiv.1612.03928

Zang, H., Wang, Y., Ru, L., Zhou, M., Chen, D., Zhao, Q., et al. (2023). Detection method of wheat spike improved YOLOv5s based on the attention mechanism. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.993244

Zhang, Y., Yu, J., Chen, Y., Yang, W., Zhang, W., and He, Y. (2022). Real-time strawberry detection using deep neural networks on embedded system (rtsd-net): An edge AI application. *Comput. Electron. Agric.* 192, 106586. doi: 10.1016/j.compag.2021.106586

Zhu, B., Wang, J., Jiang, Z., Zong, F., Liu, S., Li, Z., et al. (2020). Autoassign: Differentiable label assignment for dense object detection. *arXiv preprint arXiv* 2007, 3496. doi: 10.48550/arXiv.2007.03496

Zou, H., Lu, H., Li, Y., Liu, L., and Cao, Z. (2020). Maize tassels detection: a benchmark of the state of the art. *Plant Methods* 16 (1), 108. doi: 10.1186/s13007-020-00651-z

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 09 June 2023
DOI 10.3389/fpls.2023.1152036

OPEN ACCESS

EDITED BY
Lei Shu,
Nanjing Agricultural University, China

REVIEWED BY
Jatinder Sangha,
Agriculture and Agri-Food Canada (AAFC), Canada
Saurav Raj,
Institute of Chemical Technology, India
Thippa Reddy Gadekallu,
Zhongda Group, China
Prabhujit Mohapatra,
VIT University, India

*CORRESPONDENCE
Rammohan Mallipeddi
mallipeddi.ram@gmail.com

RECEIVED 27 January 2023
ACCEPTED 19 May 2023
PUBLISHED 09 June 2023

CITATION
Ajani OS, Aboyeji E, Mallipeddi R, Dooyum Uyeh D, Ha Y and Park T (2023) A genetic programming-based optimal sensor placement for greenhouse monitoring and control. *Front. Plant Sci.* 14:1152036. doi: 10.3389/fpls.2023.1152036

A genetic programming-based optimal sensor placement for greenhouse monitoring and control

Oladayo S. Ajani[1], Esther Aboyeji[1], Rammohan Mallipeddi[1]*, Daniel Dooyum Uyeh[2], Yushin Ha[3,4] and Tusan Park[4,5]

[1]Department of Artificial Intelligence, School of Convergence, Kyungpook National University, Daegu, Republic of Korea, [2]Department of Biosystems and Agricultural Engineering, Michigan State University, East Lansing, MI, United States, [3]Upland-Field Machinery Research Center, Kyungpook National University, Daegu, Republic of Korea, [4]Department of Bio-Industrial Machinery Engineering, Kyungpook National University, Daegu, Republic of Korea, [5]Smart Agriculture Innovation Center, Kyungpook National University, Daegu, Republic of Korea

Optimal sensor location methods are crucial to realize a sensor profile that achieves pre-defined performance criteria as well as minimum cost. In recent times, indoor cultivation systems have leveraged on optimal sensor location schemes for effective monitoring at minimum cost. Although the goal of monitoring in indoor cultivation system is to facilitate efficient control, most of the previously proposed methods are ill-posed as they do not approach optimal sensor location from a control perspective. Therefore in this work, a genetic programming-based optimal sensor placement for greenhouse monitoring and control is presented from a control perspective. Starting with a reference micro-climate condition (temperature and relative humidity) obtained by aggregating measurements from 56 dual sensors distributed within a greenhouse, we show that genetic programming can be used to select a minimum number of sensor locations as well as a symbolic representation of how to aggregate them to efficiently estimate the reference measurements from the 56 sensors. The results presented in terms of Pearson's correlation coefficient (r) and three error-related metrics demonstrate that the proposed model achieves an average r of 0.999 for both temperature and humidity and an average RMSE value of 0.0822 and 0.2534 for temperate and relative humidity respectively. Conclusively, the resulting models make use of only eight (8) sensors, indicating that only eight (8) are required to facilitate the efficient monitoring and control of the greenhouse facility.

KEYWORDS

sensor aggregation, optimal sensor location, genetic programming, greenhouse, control

1 Introduction

Optimal sensor placement is aimed at realizing a sensor profile or layout that achieves minimum cost as well as satisfies some pre-specified performance criteria has gained traction in a broad spectrum of applications areas such as health monitoring (Tan and Zhang, 2020), distribution of medicine in disaster areas (Parque et al., 2019), indoor cultivation systems (Uyeh et al., 2022b) and smart cities (Du et al., 2019; Jena et al., 2021). Specifically, in indoor cultivation systems, optimal sensor placement has become attractive to facilitate the efficient coordination of sensors for monitoring plant life as well as providing the necessary control of the internal environmental conditions (micro-climate). Indoor cultivation systems such as greenhouses are cultivation systems that are controlled in order to support all year-round growing of plants or crops (Nordey et al., 2017). Although these systems are economical compared with open field cultivation systems, they rely on effective monitoring and control of micro-climate such as temperature and humidity which have a direct impact on crop growth, quality (Takahata and Miura, 2017; Syed and Hachem, 2019) and consequently, crop yield (Nordey et al., 2017). In fact, experimental analysis has shown that while effective control of the temperature favors plant growth and reduces the overall energy consumption of the system, appropriate levels of relative humidity are necessary to prevent fungal infections and control transpiration (Vox et al., 2010). In other words, efficient monitoring and control of micro-climate are crucial to achieving the economic and sustainability goals of controlled cultivation systems.

Traditionally, monitoring of greenhouse micro-climate and consequently its control is facilitated through randomly distributed sensors (based on the available resources and size of the greenhouse) (Yeon Lee et al., 2019). However, under such settings, there is no guarantee that such randomly placed sensors would provide measurements that are representative of the true micro-climatic conditions of the greenhouse. Furthermore, the use of a large number of sensors results in a large amount of data that requires efficient data management. In other words, the quality of information and the accuracy of the resulting micro-climate heavily relies on the number of sensors and their locations/placements. Therefore, the non-trivial task of optimizing the number of sensors and their locations becomes eminent as it forms the basis for accurate measurement of micro-climate and consequently optimal control of the cultivation system. Additionally, it reduces the overall operating cost of controlled cultivation systems.

Although several techniques (Kubrusly and Malebranche, 1985; Alonso et al., 2004; Flynn and Todd, 2010; Yi et al., 2011) for optimal sensor placement have been proposed in the literature for different applications, some of the proposed methods are not directly applicable for highly non-linear setups (complex systems) such as controlled cultivation systems. In the context of controlled cultivation systems, optimization, and machine learning-based algorithms have been proposed (Yeon Lee et al., 2019; Wu et al., 2020; Uyeh et al., 2021; Uyeh et al., 2022b).

In Yeon Lee et al. (2019), a setup which relies on the fusion of an error-based and entropy-based method was proposed for optimal location of temperature sensors. In the setup, a reference temperature is generated by averaging the temperature data from all the measurement locations. Consequently, sensor locations with measurements that are statistically close to reference temperature were selected. In addition, entropy related information was used to select locations that are significantly influenced by external environmental conditions. Based on these two methods, optimal sensor locations that provide representative data of the entire greenhouse condition as well as understanding regions with high variations in temperature were realized. A hierarchical cooperative particle swarm algorithm was proposed in Wu et al. (2020) for sensor placement in a vegetable-cultivating greenhouse with the aim of maximizing the entire coverage area (i.e., a non-occlusion coverage scheme). In the scheme, the decision space was designed based on the global effective coverage of each sensor as well as the orientation angles of the respective sensors. Based on the results, the model was argued to demonstrate the capability to overcome issues of occlusion between covered objects and also improved sensor utilization in general. However, the aforementioned works are limited because they were investigated over a limited period of time which does not account for different planting seasons and weather conditions. To address these issues, (Uyeh et al., 2021) proposed a Reinforcement Learning (RL) based method to optimally place sensors in a greenhouses using a robust dataset which features different planting seasons. The dataset consists data from 56 dual temperature and humidity sensors distributed within a greenhouse. In the work, RL-based ranking of the sensor locations was performed in order of their importance in estimating the greenhouse micro-climate for temperature and relative humidity respectively. The results show that the rank of each sensor location for effective measurement of the greenhouse micro-climate varies from month to month. This is very intuitive because it is expected that different temperature and humidity profiles would occur in different months and/or planting seasons based on the changes in external weather conditions. Based on the same dataset and extracted psychrometric features (dew point temperature, enthalpy, humid ratio, and specific volume) (Uyeh et al., 2022a) proposed a machine learning-based sensors clustering system to find the optimal sensor locations. The results indicate that less than 10 percent of the sensors were required to facilitate effective monitoring of the greenhouse.

Although the aforementioned works have considered optimal sensor location in controlled cultivation system over different planting seasons and environmental conditions, it is important to realize that the ultimate goal of monitoring in controlled cultivation systems is to maintain or regulate the micro-climates to be within the desired range and this is facilitated through the associated control systems. However, these works have only considered the problem of optimal sensor placement from a monitoring or measurement perspective without any notion of control. Therefore deviating from the large body of previous works, this paper proposes a Genetic Programming (GP)-based optimal sensor placement from a control perspective for controlled cultivation systems. In the approach, firstly, we show that reference micro-climate obtained from the aggregation of all measurements from the 56 sensors is highly correlated to measurements from each of the sensors. This means that the reference temperature is a robust estimate of the overall micro-climatic condition of the greenhouse. This is important because, in terms of regulating the micro-climate within the greenhouse, only such

reference micro-climate which are representative of the entire environmental condition are required to serve as input to the dedicated control systems. Consequently, reference micro climate obtained based on the weighted averaging aggregation method are used as targets to fit GP models that can effectively model the reference micro-climate using only measurements from sensors that are most vital to the reference micro-climate. In other words, through an optimization process, GP selects only the crucial sensors and effectively fuses them to realize the reference micro-climate. Therefore, the locations of the sensor that are featured in the resulting GP model are the optimal sensor locations required to facilitate monitoring and control of the entire greenhouse. Consistent with the findings in Uyeh et al. (2021), the results show that different optimal sensor locations are representative of the entire environmental condition across different months and different micro-climate. Furthermore, the economic impact of the results is reflected in the observation that only eight (8) sensors are required to monitor and control the controlled cultivation system. This implies that the energy cost of running the greenhouse as well as the sensor procurement cost is reduced drastically.

The rest of the paper is structured as follows; Section II presents a description of the data and featured pre-processing. Furthermore, an overview of data aggregation and the methods employed in this work, as well as correlation analysis of the resulting reference micro-climate compared to the measurements from each of the 56 sensors is presented. In Section III, the background of Genetic Programming as well as the proposed modules are presented. Section IV presents the results in terms of the models obtained as well as their implications. In Section V, conclusions and future directions are highlighted.

2 Data description and aggregation

This work leverages on the same data used in (Uyeh et al., 2021; Uyeh et al., 2022a). The dataset contains temperature and relative humidity measurements collected remotely from a research cultivation-controlled system in Kyungpook National University, South Korea. The data was collected over a period of seven months (February, March, April, May, June, July, and October) using 56 dual temperature and relative humidity sensors carefully distributed within the greenhouse. Specifics about the site location, description, greenhouse layout and the data collection protocol are detailed in (Uyeh et al., 2021; Uyeh et al., 2022a). Figure 1 presents the layout that is representative of the location of each of the sensors within the greenhouse.

The resulting data includes measurements recorded per minute for the two micro-climate (temperature and relative humidity). In terms of pre-processing, rows with missing data points were removed as the number of rows with missing data points is very insignificant compared with the entire observations.

2.1 Data aggregation

In the context of achieving a controlled or regulated environment, aggregate micro-climate (relative humidity and

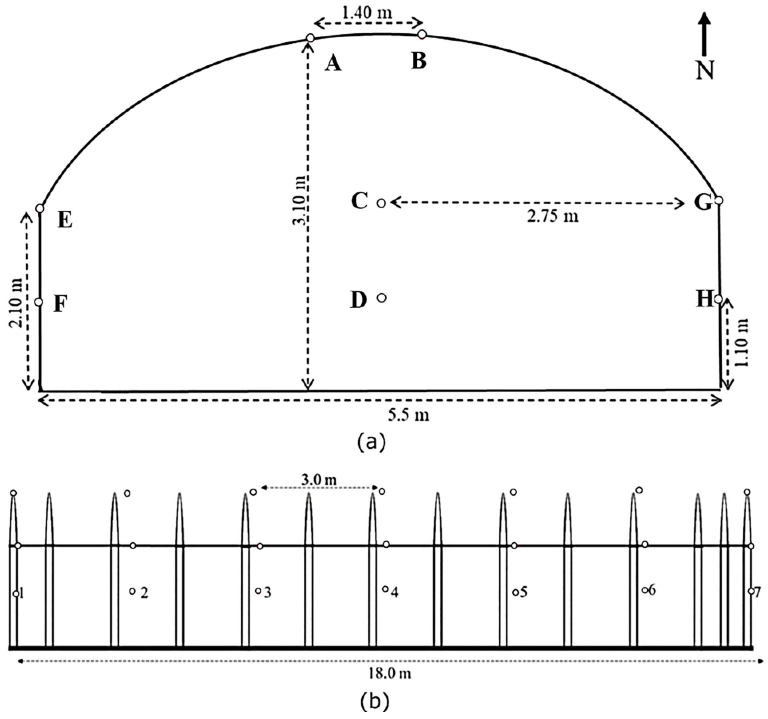


FIGURE 1
Layout of the 56 two-in-one temperature and relative humidity sensors within the greenhouse **(A)** Front view, **(B)** Side view. The small circles represent the positions of each of the sensors distributed over 8 rows (A–H) with each row containing 7 sensors.

temperature) are required as inputs to dedicated control systems for appropriate control actions within the cultivation systems (Yeon Lee et al., 2019). Data aggregation is the process of fusing information from different or multiple sensors together in order to derive a single reference measurement that is sent to a base station or controller depending on the intended application (Alkahtani and Karim, 2018; Kaur and Munjal, 2020; Yuan et al., 2021). Generally, in controlled cultivation systems the aim is to ensure that the micro-climate are controlled to support plant life and growth. To facilitate such control, there is need to have reference micro-climate that is representative of the environmental conditions of the cultivation systems and consequently take control actions based on the associated control laws. Several data aggregation methods such as weighted averaging (Hang et al., 2017; Yeon Lee et al., 2019), median (Cocco et al., 2015) and more complex fusing algorithms such as the unscented Kalman filter (Xia et al., 2022) and weighted least square method (Ren et al., 2017) etc. have been proposed in the literature for application in cultivating systems and other application domains.

In this work, we use the simple weighted averaging method given as

$$W = \frac{\sum_{i=1}^{N} w_i X_i}{\sum_{i=1}^{N} w_i} \quad (1)$$

where N is the total number of sensors to be averaged, w_i is the weights applied to each sensor value and X_i is the sensor values to be averaged. Similar to (Yeon Lee et al., 2019), we take the weight $w_i = 1$ for all the 56 sensors. This is to ensure that every variation or section of the greenhouse is given equal important. Furthermore, to ensure that the chosen aggregation method is representative of the response of each sensor we perform correlation analysis of the reference micro-climate with micro-climate from each of the sensors.

3 Genetic programming-based optimal sensor location

In this Section, a systemic overview of GP is presented and consequently, the protocols of the GP for the optimal sensor location based on the aforementioned data are presented.

3.1 Genetic programming

In artificial intelligence, Genetic programming (GP) is a class of bio-inspired algorithms generally known as evolutionary algorithms that are capable of generating solutions to problems that humans cannot solve or do not know how to solve directly. Formally, GP is a systematic method for getting computers to automatically solve a problem starting from a high-level statement of what needs to be done (Koza and Poli, 2005). Generally, based on different genetic operations (genetic events) such as crossover, mutation, reproduction, gene duplication, and gene deletion the idea is to randomly generate a large set of solutions and to evolve those solutions until the population converges to a global maxima/minima depending on the associated task and termination criteria. It is often used in the field of Machine Learning for hyper-parameter selection (Agrawal et al., 2021) or to determine relationships between features in data (Rodrigues et al., 2022). For example in the context of this work, the measurements from the 56 sensors are features and we intend to select the best features corresponding to the optimal sensor locations.

In terms of implementation, the typical evolution process of GP involves the following steps:

1. Define the problem objectives and randomly initialize or generate a population of solution candidates.
2. Repeat the following steps until a pre-defined termination criterion is reached:
   (a) Evaluate each of the solution candidates in the population based on the problem objective and assign it a function value.
   (b) Generate a new population of solution candidates by performing the following operations:
      1) Select a set of solution candidates for mating based on the assigned fitness value (selection).
      2) Include some of the selected solution candidates into the new population without modifying them (reproduction).
      3) Generate new solution candidates by genetically recombining randomly chosen parts of two selected individuals (crossover).
      4) Generate new solution candidates by replacing randomly chosen parts of some selected individuals with new randomly generated ones (mutation).
3. The resulting best solution candidates at any generation of the evolution process is chosen as the result of the GP process.

The aforementioned steps are summarized mathematically in Algorithm 1. In classical GP, solution candidates or programs are encoded as tree-based structures as shown in Figure 2 (Koza, 1993) because evaluating trees in a recursive manner is easy. Under this setting, mathematical expressions are evolved and evaluated with each tree nodes having an operator function and each terminal mode an operand. Furthermore, the crossover operation is achieved by swapping randomly selected sub-trees from two parent candidates while mutation is achieved by replacing a randomly chosen individual's sub-tree by a randomly generated one (Sotto et al., 2021). For example, in Figure 2A, the sub-tree of the parent solution is replaced to produce the offspring and in Figure 2B, two parents P1 and P2 are crossed to produced offspring 1 and 2 accordingly.

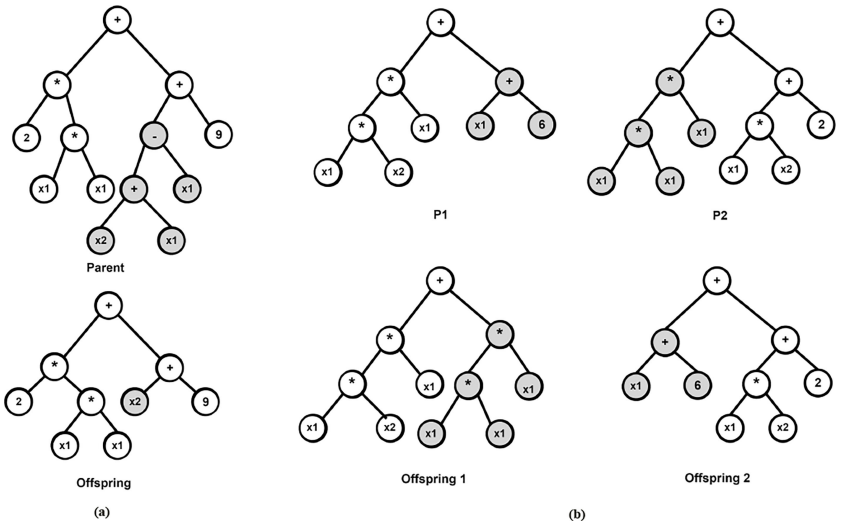


FIGURE 2
Examples of mutation and crossover operations in GP **(A)** Offspring generated by a single-parent mutation **(B)** Two offspring are generated by a bi-parent crossover.

3.2 GP-based optimal sensor location

In order to evolve a GP model that is representative of the optimal sensor locations, it's important to set the global task or objective. In terms of optimal sensor location, the goal is to realize an efficient combination of a limited number of sensors that can estimate the reference micro-climate obtained from the aggregation of the 56 sensors. Since the reference micro-climate are real continuous values, the problem at hand can be formalized as a classical symbolic regression task which is one of the most widely studied application of GP (Uy et al., 2010; He et al., 2022; Zojaji et al., 2022). Symbolic Regression (SR) is a class of machine learning approach that searches the space of mathematical expressions with the aim of identifying a model or expression that best describes the relations between a given dataset, both in terms of accuracy and simplicity. This can be summarized as a multi-objective framework where accuracy (error) is maximized (minimized) and the number of sensors is minimized (simplicity). Therefore, given the measurements of 56 sensors as input for each of the associated micro-climate and a set of operator functions, the GP builds a symbolic regression model and selects the minimum number of sensors sufficient to estimate the reference extcolor micro-climate from the 56 sensors. Consequently, the resulting locations of the chosen sensors are the optimal sensor locations for the associated micro-climate and month.

4 Computational experiments

Based on the data collected for each of the month featured in the aforementioned dataset, we construct GP models using variables (data from each of the sensors) as well as random numbers as terminals and arithmetic operators such as (addition, multiplication etc.) as operator functions. The choice of constructing GP models based on each month was motivated by intuition that different sensor profiles would be optimal for different months and seasons which was also validated in (Uyeh et al., 2021). For each month and the associated micro-climate, the data is divided randomly into training and testing set based on 70:30 ratio. Furthermore, the training set is further divided to obtain a validation set based on 80:20 ratio. The random division of the data is chosen to ensure that the opportunity to model the different time trends is not missed. Because the validation of the model would be affected if certain time trends are ignored in model development.

All the experiments were conducted in MATLAB installed on a 64-bit Windows 11 PC, with 3.00GHz Intel-i5-12500 CPU and 32GB RAM. The GP is initialized with a population size of 500 and is allowed to evolve for 100 generations. The best results obtained over 25 independent runs of the GP algorithm are reported. In terms of selection, tournament selection (Fang and Li, 2010) with size of 25 was used and an elite fraction of 0.3. For all the experiments, the set of function nodes used are basic arithmetic operators (+, -, ×) as well as minimum (min) and maximum (max) operators.

5 Results and discussion

To evaluate the resulting GP models, we employ a number of metrics namely; Pearson's Correlation Coefficient (R), Root Mean Squared Error (RMSE), Mean Average Error (MAE) and Maximum Absolute Error (Max.AE).

5.1 Correlation of sensor aggregation

Table 1 shows the average correlation of the reference micro-climate (temperature and relative humidity) with each of the

TABLE 1 Average correlation of the reference micro-climate with each of the measurements from the 56 sensors.

Months	Correlation Coefficients (*r*)	
	Temperature	Relative humidity
February	0.972	0.978
March	0.973	0.978
April	0.980	0.980
May	0.980	0.977
June	0.976	0.984
July	0.977	0.985
October	0.970	0.983

measurements from the 56 sensors over the even months. As seen in Table 1, the reference textcolormicro-climate are highly correlated with those measured from each of the sensors with the lowest being 97%. This demonstrates that the reference micro-climate based on the average aggregation method is satisfactorily representative of the global environmental conditions of the controlled cultivation systems and can be used to facilitate the control of the entire regions of the cultivation system.

5.2 Performance of GP-based model for temperature

In Table 2, the symbolic equations for the resulting model based on the associated sensors are presented. Specifically, the equations represent how to aggregate the information from each sensor as well as the bias term. Based on those models, Table 3 presents the performance of the model in terms of Pearson's Correlation Coefficient (*r*) with the reference temperatures, Root Mean Squared Error (RMSE), Mean Average Error (MAE) and the Maximum Absolute Error (Max.AE) of the predicted temperature compared to the reference temperature for each of the 7 months. In Figures 3, 4, comparisons between the actual and predicted values based on the test dataset are presented. The results in terms of the *r* values shows that the actual and predicted temperature based on the GP model are highly correlated with an average value of over 0.99 across the seven months. In terms of the error-related metrics, such as RMSE and MAE, it can be seen from Table 3 that the values are insignificant and within allowable limits. It is important to note that those error values are not from normalized samples but are based on the real magnitudes of the temperature measurements. The Max.AE metric presents the worst cases of error between the actual and real temperature values. These values are found to be in the region of the allowable measurement error from the device manufacturer which is ± 0.3°C (Uyeh et al., 2021). In terms of the qualitative analysis of the actual and predicted temperature values presented in Figures 3, 4, it can be clearly seen that the actual and predicted temperature measures are very similar.

5.3 Performance of GP-based model for relative humidity

In Table 4, the symbolic equations for the resulting model based on the associated sensors are presented. Based on those models, Table 5 presents the performance of the model in terms of Pearson's Correlation Coefficient (*r*) with the reference relative humidity, Root Mean Squared Error (RMSE), Mean Average Error (MAE) and the Maximum Absolute Error (Max.AE) of the predicted relative humidity compared to the reference relative humidity for each of the 7 months. In Figures 5, 6, comparison between the actual and predicted values based on the test dataset are presented. The results of the correlation analysis presented in Table 5 shows that the actual and predicted relative humidity based on the GP-model are highly correlated with an average value of over 0.99 across the seven months. In terms of the error related metrics, such as RMSE and MAE, it can be seen from Table 5 that the values are insignificant and within allowable limits. It is important to note that those error values are not from normalized samples but are based on the real magnitudes of the relative humidity measurements. The Max. AE metric presents the worst cases of the error between the actual and real relative humidity values. These values are found to be in the region of the allowable measurement error from the device manufacturer which is ± 2% C (Uyeh et al., 2021). In terms of the qualitative analysis of the actual and predicted relative humidity values presented in

TABLE 2 Resulting GP-based symbolic models (equations) for temperature.

Months	Symbolic Equations for Temperature
February	0.126A1 + 0.126A2 + 0.126B5 + 0.126B6 + 0.126D7 + 0.126E4 + 0.126F4 + 0.126H1 - 0.1880
March	0.126A1 + 0.126A2 + 0.126C5 + 0.126C7 + 0.126D5 + 0.126E2 + 0.126F5 + 0.126G3 - 0.0887
April	0.125A4 + 0.125B3 + 0.125B5 + 0.125C1 + 0.125D6 + 0.125E1 + 0.125E4 + 0.125F5 - 0.0672
May	0.125A5 + 0.125C1 + 0.125D4 + 0.125E4 + 0.125E6 + 0.125F3 + 0.125G5 + 0.125H1 - 0.0295
June	0.124B1 + 0.124B5 + 0.124B7 + 0.124D5 + 0.124E4 + 0.124E6 + 0.124F3 + 0.124H1 + 0.1330
July	0.125B3 + 0.125B4 + 0.125C7 + 0.125E4 + 0.125E6 + 0.125F5 + 0.125G4 + 0.125H1 + 0.0312
October	0.125B2 + 0.125B7 + 0.125D5 + 0.125E6 + 0.125F2 + 0.125G5 + 0.125H1 + 0.125H3 - 0.0088

10.3389/fpls.2023.1152036

TABLE 3 Performance of the GP-based models in terms of Pearson's Correlation Coefficient (*r*) with the reference temperature, Root Mean Squared Error (RMSE), Mean Average Error (MAE) and the Maximum Absolute Error of the predicted temperature (Max.AE).

Metrics	Months						
	February	March	April	May	June	July	October
r	0.9997	0.9998	0.9996	0.9999	0.9999	0.9999	0.9998
RMSE	0.0884	0.0862	0.1490	0.0761	0.0601	0.03178	0.0836
MAE	0.0634	0.0502	0.0644	0.0442	0.0383	0.0210	0.0527
Max.AE	0.4842	0.5808	1.1346	0.5672	0.4539	0.2613	0.6350

Figures 5, 6, it can be clearly seen that the actual and predicted relative humidity values are very similar.

5.4 Analysis of selected optimal sensor locations

In Table 6, the selected optimal sensor locations for each month are presented for temperature. From the Table, it can be observed that for the months of February, March and April an average of 5 out of the eight sensors selected are distributed along the center of the greenhouse (*A, B, C, D*), while other remaining three are either to the right or left side of the greenhouse. On the other hand for May to October, it can be seen that only an average of 3 of the sensors selected are distributed along the center of the greenhouse while the others are distributed to the left or right side of the greenhouse. This can be attributed to the different seasons of each month. Specifically, it can be inferred that the colder months have higher concentration of sensors along the center of the greenhouse while the more hotter months takes more advantage of sensor distributed along the facility.

In Table 7, the selected optimal sensor locations for each month are presented for relative humidity. It can be observed from the Table that the selected sensors were mostly distributed along the center of the greenhouse across the months. Specifically, each month had at least five (February, March, and July) or six (April, May, June, and October) of

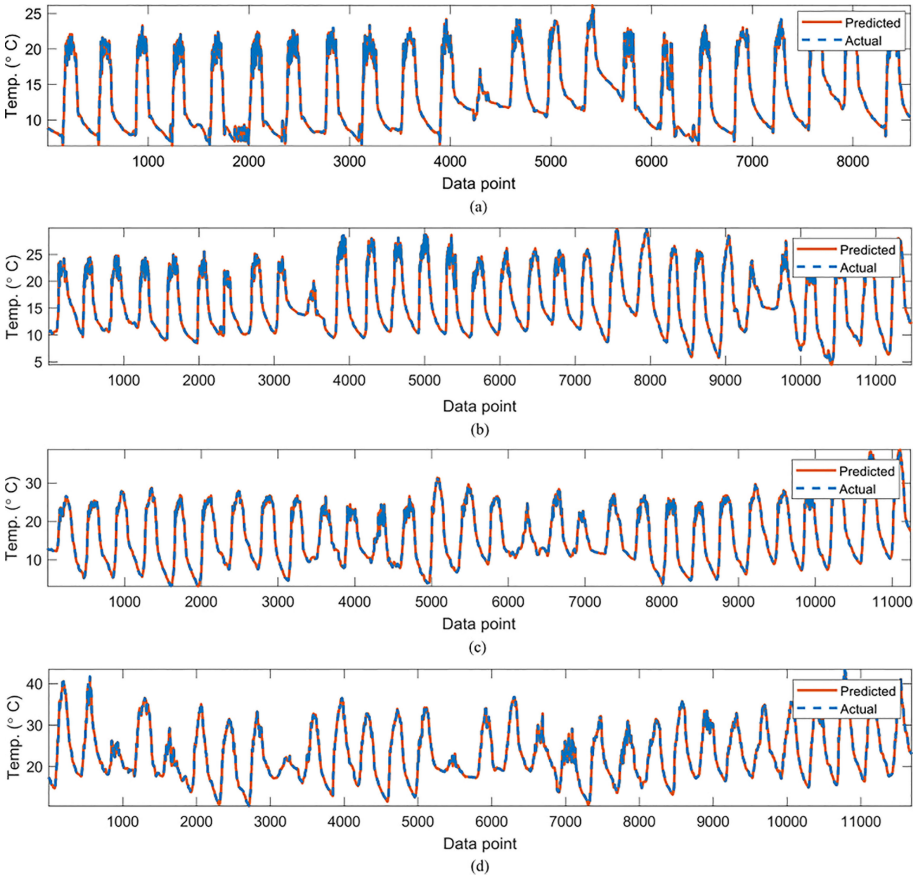


FIGURE 3
Comparisons of the actual reference temperature versus those predicted by the proposed GP-models for **(A)**February, **(B)** March, **(C)**, April **(D)** May.

10.3389/fpls.2023.1152036

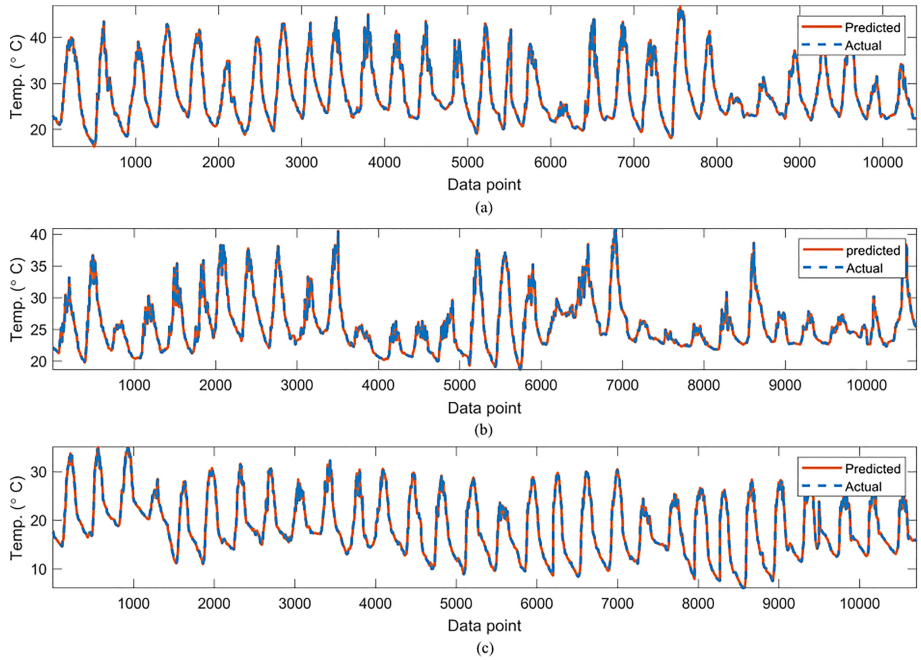


FIGURE 4
Comparisons of the actual reference temperature versus those predicted by the proposed GP-models for **(A)** June, **(B)** July **(C)** October.

the eight sensors distributed along the center of the greenhouse while the remaining three (3) or two (2) sensors respectively were distributed either to the right or left side of the greenhouse. This basically means that more sensors are selected from the center of the greenhouse during hotter months compared to colder ones with the exception of July which had the same number of sensors distributed in the middle as with February and March.

5.5 Implication of the GP-based model from a control and economic perspective

As mentioned earlier, the ultimate goal of monitoring in controlled cultivation systems is to achieve appropriate control. The advantage of the proposed framework from a control perspective is that it not only gives the optimal sensor locations for each month, but it also provides how to aggregate them efficiently to facilitate the needed control of the entire system. The symbolic representations presented in Tables 2, 4 for temperature and humidity respectively are the needed aggregation expressions required to obtain reference temperature and humidity that is representative of the micro-climate of the entire cultivation systems which can be fed into the control system and consequently provide control actions based on the associated control laws.

The results from the proposed model, indicate that only 8 optimally distributed sensors (less than 15% of the distributed sensors) are sufficient to facilitate efficient and effective monitoring and control of indoor environmental parameters. This reduces the entire operating cost in terms of energy use and most importantly, the cost of sensor procurement and installation can be reduced by about 75%.

TABLE 4 Resulting GP-based symbolic models (equations) for humidity.

Months	Symbolic Equations for Humidity
February	0.126A1 + 0.126A3 + 0.126A5 + 0.126B6 + 0.126D6 + 0.126E2 + 0.126F4 + 0.126G3 - 0.8330
March	0.125A1 + 0.125A4 + 0.125B7 + 0.125C5 + 0.125D3 + 0.125E1 + 0.125E6 + 0.125F3 + 0.0494
April	0.126A1 + 0.126A4 + 0.126A5 + 0.126B5 + 0.126B6 + 0.126C3 + 0.126F6 + 0.126H1 - 0.6660
May	0.125B4 + 0.125B6 + 0.125C1 + 0.125C2 + 0.125C3 + 0.125D7 + 0.125F6 + 0.125H1 - 0.1570
June	0.125B6 + 0.125B7 + 0.125C1 + 0.125C3 + 0.125D3 + 0.125D7 + 0.125F2 + 0.125G4 - 0.0069
July	0.124B2 + 0.124B7 + 0.124C1 + 0.124C4 + 0.124D6 + 0.124E2 + 0.124E6 + 0.124H3 + 0.4730
October	0.125A4 + 0.125B6 + 0.125B7 + 0.125C3 + 0.125E1 + 0.125E2 + 0.125F5 + 0.125H5 - 0.2610

10.3389/fpls.2023.1152036

TABLE 5 Performance of the GP-based models in terms of Pearson's Correlation Coefficient (*r*) with the reference relative humidity, Root Mean Squared Error (RMSE), Mean Average Error (MAE) and the Maximum Absolute Error (Max.AE) of the predicted relative humidity.

Metrics	Months						
	February	March	April	May	June	July	October
r	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
RMSE	0.3397	0.3152	0.2943	0.2196	0.2115	0.1536	0.2399
MAE	0.1786	0.1864	0.1722	0.1374	0.1462	0.1215	0.1659
Max.AE	4.1358	4.2093	4.1959	3.1712	2.2670	1.0564	3.4247

6 Conclusions and future works

In this work, an optimal sensor location for controlled cultivation system based on Genetic Programming (GP) is proposed. Using data collected from 56 dual temperature and humidity sensors distributed within a greenhouse, reference temperature and humidity values are obtained based on the weighted average aggregation of the data. Consequently, GP is used to build symbolic models which are representative of the optimal sensors as well as how to optimally aggregate the data from the sensors. The results based on the test data shows that the reference micro-climate from the GP-based model for each month is highly correlated to those obtained based on all the 56 sensors. Furthermore based on several error metrics, it was found that the resulting error from using only 8 sensors based on the GP model is within allowable measurement error as provided from the device manufacturer which is ±2%°C.

Although, this work has been limited to only Temperature and Relative Humidity, light or Photosynthetic active radiation is another important requirement in a greenhouse or any other controlled cultivation system. Therefore in the Future, we would be interested in considering the effect of light as well as other micro-climate within the greenhouse.

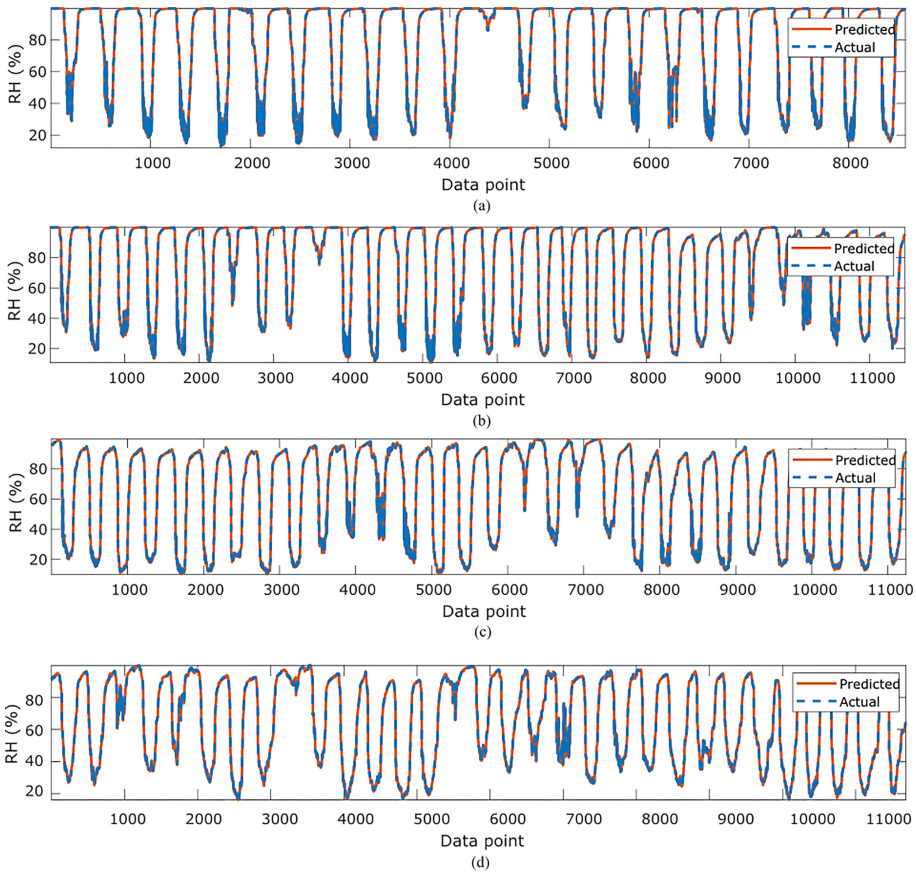


FIGURE 5
Comparisons of the actual relative humidity versus those predicted by the proposed GP-models for **(A)**February, **(B)** March, **(C)**, April **(D)** May.

10.3389/fpls.2023.1152036

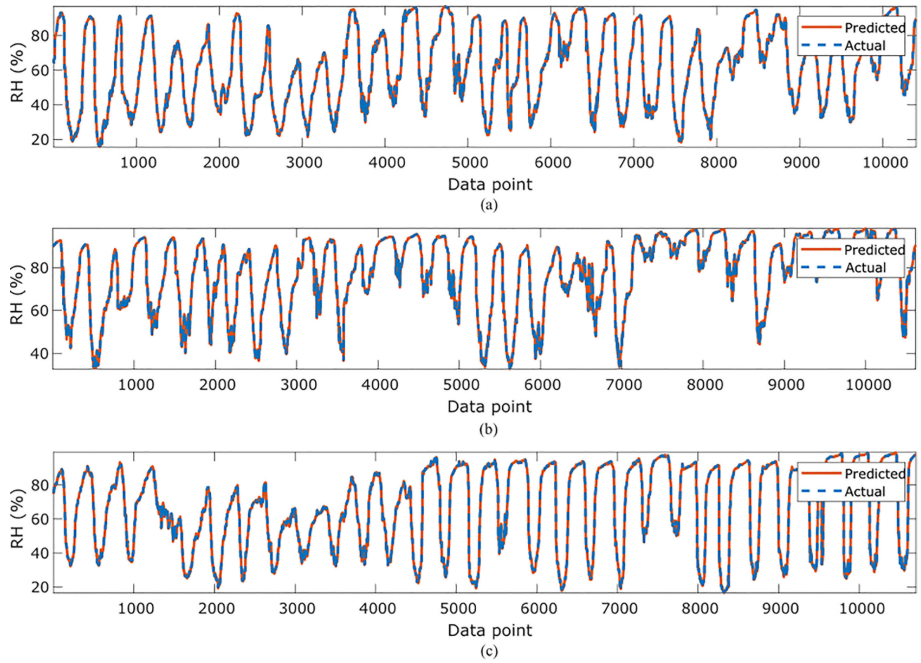


FIGURE 6
Comparisons of the actual relative humidity versus those predicted by the proposed GP-models for **(A)** June, **(B)** July **(C)** October.

TABLE 6 Optimal sensor locations for temperature.

Optimal sensor locations						
February	March	April	May	June	July	October
A1	A1	A4	A5	B1	B3	B2
A2	A2	B3	C1	B5	B4	B7
B5	C5	B5	D4	B7	C7	D5
B6	C7	C1	E4	D5	E4	E6
D7	D5	D6	E6	E4	E6	F2
E4	E2	E1	F3	E6	F5	G5
F4	F5	E4	G5	F3	G4	H1
H1	G3	F5	H1	H1	H1	H3

TABLE 7 Optimal sensor location for relative humidity.

Optimal sensor locations						
February	March	April	May	June	July	October
A1	A1	A1	B4	B6	B2	A4
A3	A4	A4	B6	B7	B7	B6
A5	B7	A5	C1	C1	C1	B7
B6	C5	B5	C2	C3	C4	C3
D6	D3	B6	C3	D3	D6	E1
E2	E1	C3	D7	D7	E2	E2
F4	E6	F6	F6	F2	E6	F5
G3	F3	H1	H1	G4	H3	H5

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

OA: Conceptualization, Methodology, Formal analysis, Software, Investigation, Writing - Original Draft; EA: Formal analysis, Writing - Original Draft, Writing - Review and Editing; RM: Supervision, Validation, Writing - Review and Editing; DD: Conceptualization, Data Curation; YH: Resources, Data Curation; TP: Resources, Data Curation, Supervision, Writing - Review and Editing. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2021R1I1A3049810) and the Korea Institute of Planning and Evaluation for Technology in Food, Agriculture, and Forestry (IPET) through Agriculture, Food and Rural Affairs Convergence Technologies Program for Educating Creative Global Leader, funded by Ministry of Agriculture, Food and Rural Affairs (MAFRA) (320001-4), Rep. Korea.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Agrawal, S., Sarkar, S., Alazab, M., Maddikunta, P. K. R., Gadekallu, T. R., and Pham, Q.-V. (2021). Genetic cfl: hyperparameter optimization in clustered federated learning. *Comput. Intell. Neurosci.* doi: 10.1155/2021/7156420

Al-kahtani, M. S., and Karim, L. (2018). Dynamic data aggregation approach for sensor-based big data. *Int. J. Of Advanced Comput. Sci. And Appl.* 9, 62–72. doi: 10.14569/IJACSA.2018.090710

Alonso, A. A., Kevrekidis, I. G., Banga, J. R., and Frouzakis, C. E. (2004). Optimal sensor location and reduced order observer design for distributed process systems. *Comput. Chem. Eng.* 28, 27–35. doi: 10.1016/S0098-1354(03)00175-3

Cocco, A., Serra, G., Lentini, A., Deliperi, S., and Delrio, G. (2015). Spatial distribution and sequential sampling plans for tuta absoluta (lepidoptera: gelechiidae) in greenhouse tomato crops. *Pest Manage. Sci.* 71 9, 1311–1323. doi: 10.1002/ps.3931

Du, R., Santi, P., Xiao, M., Vasilakos, A. V., and Fischione, C. (2019). The sensable city: a survey on the deployment and management for smart city monitoring. *IEEE Commun. Surveys Tutorials* 21, 1533–1560. doi: 10.1109/COMST.2018.2881008

Fang, Y., and Li, J. (2010). "A review of tournament selection in genetic programming," in *Advances in computation and intelligence*. Eds. Z. Cai, C. Hu, Z. Kang and Y. Liu (Berlin, Heidelberg: Springer Berlin Heidelberg), 181–192.

Flynn, E. B., and Todd, M. D. (2010). A bayesian approach to optimal sensor placement for structural health monitoring with application to active sensing. *Mechanical Syst. Signal Process.* 24, 891–903. doi: 10.1016/j.ymssp.2009.09.003

Hang, Z., Linda, S., Wangliang, L., Chuang, L., and Kaiyan, W. (2017). "Application of multi-sensor data fusion technique in greenhouse environmental monitoring," in *2017 International Conference on Smart Grid and Electrical Automation (ICSGEA)*, Changsha, China. 51–55. doi: 10.1109/ICSGEA.2017.47

He, B., Lu, Q., Yang, Q., Luo, J., and Wang, Z. (2022). "Taylor Genetic programming for symbolic regression," in *Proceedings of the genetic and evolutionary computation conference* (New York, NY, USA: Association for Computing Machinery), 946–954. doi: 10.1145/3512290.3528757

Jena, P. K., Ghosh, S., and Koley, E. (2021). Identification of optimal sensor location based on trade-off approach to improve resiliency of electricity market in smart grid. *IEEE SENSORS J.* 21, 17271–17281. doi: 10.1109/JSEN.2021.3078267

Kaur, M., and Munjal, A. (2020). Data aggregation algorithms for wireless sensor network: a review. *AD HOC Networks* 100, 1–20. doi: 10.1016/j.adhoc.2020.102083

Koza, J. R. (1993). Genetic Programming: On the Programming of Computers by Means of Natural Selection (Complex Adaptive Systems). *A Bradford Book* 1, 18.

Koza, J. R., and Poli, R. (2005). *Genetic programming* (Boston, MA: Springer US), 127–164. doi: 10.1007/0-387-28356-0

Kubrusly, C. S., and Malebranche, H. (1985). Sensors and controllers location in distributed systems - a survey. *Autom* 21, 117–128. doi: 10.1016/0005-1098(85)90107-4

Nordey, T., Basset-Mens, C., de Bon, H., Martin, T., Delétré, E., Simon, S., et al. (2017). Protected cultivation of vegetable crops in sub-saharan africa: limits and prospects for smallholders. A review. *Agron. Sustain. Dev.* 37, 1–20. doi: 10.1007/s13593-017-0460-8

Parque, V., Obasekore, H., Oladayo, S., and Miyashita, T. (2019). "On planning distributed minimal sensor networks," in *2019 IEEE 1st Global Conference on Life Sciences and Technologies (LifeTech)*, Osaka, Japan, pp. 26–28. doi: 10.1109/LifeTech.2019.8884022

Ren, M., Sun, L., Liu, M., Cheung, C., Yin, Y., and Cao, Y. (2017). A weighted least square based data fusion method for precision measurement of freeform surfaces. *Precis. Eng.* 48, 144–151. doi: 10.1016/j.precisioneng.2016.11.014

Rodrigues, N. M., Batista, J. E., La Cava, W., Vanneschi, L., and Silva, S. (2022). "Slug: feature selection using genetic algorithms and genetic programming," in *Genetic programming*. Eds. E. Medvet, G. Pappa and B. Xue (Cham: Springer International Publishing), 68–84.

Sotto, L. F. D. P., Kaufmann, P., Atkinson, T., Kalkreuth, R., and Basgalupp, M. P. (2021). Graph representations in genetic programming. *Genet. Programming Evolvable Machines* 22, 607–636. doi: 10.1007/s10710-021-09413-9

Syed, A. M., and Hachem, C. (2019). Review of construction; geometry; heating, ventilation, and air-conditioning; and indoor climate requirements of agricultural greenhouses. *J. Biosyst. Eng.* 23, 18–27 Syed2019ReviewOC. doi: 10.1007/s42853-019-00005-1

Takahata, K., and Miura, H. (2017). Effects of growth period and air temperature on the position of the inflorescence on the stem of tomato plants. *Horticulture J.* 86, 70–77. doi: 10.2503/hortj.MI-110

Tan, Y., and Zhang, L. (2020). Computational methodologies for optimal sensor placement in structural health monitoring: a review. *Struct. Health Monit.* 19, 1287–1308. doi: 10.1177/1475921719877579

Uy, N. Q., Hoai, N. X., O'Neill, M., McKay, R. I., and López, E. G. (2010). Semantically-based crossover in genetic programming: application to real-valued symbolic regression. *Genet. Programming Evolvable Machines* 12, 91–119. doi: 10.1007/s10710-010-9121-2

Uyeh, D. D., Akinsoji, A., Asem-Hiablie, S., Itoro Bassey, B., Osinuga, A., Mallipeddi, R., et al. (2022a). An online machine learning-based sensors clustering system for efficient and cost-effective environmental monitoring in controlled environment agriculture. *Comput. Electron. Agric.* 199, 107139. doi: 10.1016/j.compag.2022.107139

Uyeh, D. D., Bassey, B. I., Mallipeddi, R., Asem-Hiablie, S., Amaizu, M., Woo, S., et al. (2021). A reinforcement learning approach for optimal placement of sensors in protected cultivation systems. *IEEE Access* 9, 100781–100800. doi: 10.1109/ACCESS.2021.3096828

Uyeh, D. D., Iyiola, O., Mallipeddi, R., Asem-Hiablie, S., Amaizu, M., Ha, Y., et al. (2022b). Grid search for lowest root mean squared error in predicting optimal sensor location in protected cultivation systems. *Front. IN Plant Sci.* 13, 1–17. doi: 10.3389/fpls.2022.920284

Vox, G., Teitel, M., Pardossi, A., Minuto, A., Tinivella, F., and Schettini, E. (2010). Sustainable greenhouse systems. *Sustainable Agriculture: Technology, Planning and Management*. A. Salazar and I. Rios Eds. 1–78.

Wu, H., Li, Q., Zhu, H., Han, X., Li, Y., and Yang, B. (2020). Directional sensor placement in vegetable greenhouse for maximizing target coverage without occlusion. *Wireless Networks* 26, 4677–4687. doi: 10.1007/s11276-020-02370-8

Xia, S., Nan, X., Cai, X., and Lu, X. (2022). Data fusion based wireless temperature monitoring system applied to intelligent greenhouse. *Comput. Electron. Agric.* 192, 106576. doi: 10.1016/j.compag.2021.106576

Yeon Lee, S., Bok Lee, I., Hyeon Yeo, U., Woo Kim, R., and Gyu Kim, J. (2019). Optimal sensor placement for monitoring and controlling greenhouse internal environments. *Biosyst. Eng.* 188, 190–206. doi: 10.1016/j.biosystemseng.2019.10.005

Yi, T., Li, H., and Gu, M. (2011). Optimal sensor placement for health monitoring of high-rise structure based on genetic algorithm. *Math. Problems Eng.* 1–12. doi: 10.1155/2011/395101

Yuan, J., Liu, W., Wang, J., Shi, J., and Miao, L. (2021). An efficient framework for data aggregation in smart agriculture. *CONCURRENCY AND COMPUTATION-PRACTICE Exp.* 33, 1–14. doi: 10.1002/cpe.6160

Zojaji, Z., Ebadzadeh, M. M., and Nasiri, H. (2022). Semantic schema based genetic programming for symbolic regression. *Appl. Soft Comput.* 122, 1–25. doi: 10.1016/j.asoc.2022.108825

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 22 August 2023
DOI 10.3389/fpls.2023.1168900

OPEN ACCESS

EDITED BY
Lei Shu,
Nanjing Agricultural University, China

REVIEWED BY
Huiling Chen,
Wenzhou University, China
Changji Wen,
Jilin Agricultural University, China

*CORRESPONDENCE
Yanhua Zhang
zhangyanhua02@caas.cn
Ke Yang
82101201039@caas.cn

RECEIVED 18 February 2023
ACCEPTED 07 August 2023
PUBLISHED 22 August 2023

CITATION
Yu Z, Yang M, Hu Z, Gu F, Peng B, Zhang Y and Yang K (2023) Kinematic analysis and process optimization of root-cutting systems in field harvesting of garlic based on computer simulation technology. *Front. Plant Sci.* 14:1168900. doi: 10.3389/fpls.2023.1168900

Kinematic analysis and process optimization of root-cutting systems in field harvesting of garlic based on computer simulation technology

Zhaoyang Yu[1,2], Mingjin Yang[1], Zhichao Hu[2], Fengwei Gu[2], Baoliang Peng[2], Yanhua Zhang[2*] and Ke Yang[3*]

[1]College of Engineering and Technology, Southwest University, Chongqing, China, [2]Nanjing Institute of Agricultural Mechanization, Ministry of Agriculture and Rural Affairs, Nanjing, China, [3]Key Laboratory of Modern Agricultural Equipment, Ministry of Agriculture and Rural Affairs, Nanjing, China

Introduction: Root cutting is an important process in garlic field harvesting but is the weakest link in the full mechanization of garlic production. To improve the current situation of technological backwardness and poor operational quality of mechanized garlic root-cutting in the main garlic-producing regions of China, this study combined the physical characteristics and agronomic requirements of garlic plants, and proposed an innovative floating root-cutting technology for garlic combine harvesters that enables the top alignment of bulb, adaptive profiling floating of cutter, and embedded cutting of roots.

Methods: Through the kinematic analysis of the floating cutting process, the coordinate equations of the initial contact point of the bulb, the mathematical model of the floating displacement of the cutting component. Using computer simulation techniques, the dynamic simulation study of the floating cutting process was carried out in the rigid-flexible coupling numerical simulation model of root-cutting mechanism and garlic plant. The influence law of garlic conveying speed, extension spring preload force and stiffness on the floating displacement of the cutting component and the angular velocity of swing arm reset and its formation causes were analyzed by a single-factor simulation test. The key operating parameters of the root-cutting mechanism were optimized through the computerized virtual orthogonal test and fuzzy comprehensive evaluation.

Results and discussion: The significance of the factors affecting the floating cutting performance decreased in the following order: extension spring preload force, garlic conveying speed and extension spring stiffness. The optimal parameter combination of the root cutting mechanism obtained from the optimization were as follow: extension spring preload force was 16 N, garlic conveying speed was 0.8 m/s, and extension spring stiffness was 215 N/m. Tests

conducted with the optimal parameter combination yielded a root excision rate of 92.72%, which meets the requirements of Chinese garlic field harvesting quality. This study provides computer simulation optimization methods for the optimal design of the root-cutting mechanism, and also provides technical and equipment support for the full mechanization of garlic production in China.

KEYWORDS

garlic field harvesting, floating root-cutting, computer simulation, virtual orthogonal test, fuzzy comprehensive evaluation, root excision rate

1 Introduction

Garlic (*Allium sativum L.*) is an herb of the genus Allium in the lily family, whose bulbs have a high food value not only for direct consumption but also as a cooking ingredient or condiment for everyday dishes. According to the statistics of the Food and Agriculture Organization of the United Nations (FAO), China accounts for 80% of the global garlic production, while 90% of cultivated areas are located in Asia and Africa (FAOSTAT, 2020a; FAOSTAT, 2020b). China is the world's largest garlic producer. Its perennial garlic planting area, harvest area, and export volume rank first in worldwide (FAOSTAT, 2020a; FAOSTAT, 2020b). Figure 1 shows the garlic field production process, which has three major stages: sowing, field management and harvesting. In China, old machinery has been used for garlic sowing and field management, while harvesting is done manually or with small excavators. Highly efficient garlic combine harvesting equipment can be considered as "unavailable" (Yu et al., 2023). It is difficult to achieve combined harvesting of garlic in China due to the prevalence of mulching, narrow spaced dense planting and the tendency of garlic plants to fall over during harvest period. In 2022, the total mechanized harvesting area of garlic in China was less than 6%. The low level of harvest mechanization has hindered the development of China's garlic industry. The slow development of combine harvesting technology is the main factor affecting the promotion and application of mechanized garlic harvesting technology in China.

The garlic combine harvesting process includes garlic digging, soil removal, seedling cutting, root cutting, and bulb bagging (Zhao et al., 2020). Root-cutting is the most difficult process as well as the main factor that hiders the development of garlic combine harvesting technology (Yang et al., 2015; Quynh Anh et al., 2022). Owing to characteristics such as deep rooting, wide lateral growth of garlic roots, soft and tough roots, and the unsupported state of roots during cutting, it is difficult to efficiently remove the roots all at once. The size of the bulbs also varies, and traditional fixed cutters cannot be change according to the size of the bulb (Wang et al., 2018). Moreover, because the roots are attached to the root disc of the bulb, and when cutting the roots near the root disc, it is very easy to cause bulb cutting damage. It is a key technical challenge that must be addressed in the combine harvesting of garlic to achieve one-time and high-efficiency root removal without bulb damage.

To date, research on the root-cutting technology for garlic combine harvesting remains in the exploratory and experimental stages worldwide. Years of continuous and extensive literature searches have identified few academic reports on the mechanized garlic root-cutting technology. Yu et al. (2021b) developed a profiling, roller-type root-cutting mechanism for garlic combine harvesters. The rotational trajectory of the cutter was designed as a rotating hyperbolic surface that follows the shape of the lower surface of a garlic bulb. The operating parameters of the mechanism were determined, and the root fracture process was analyzed in

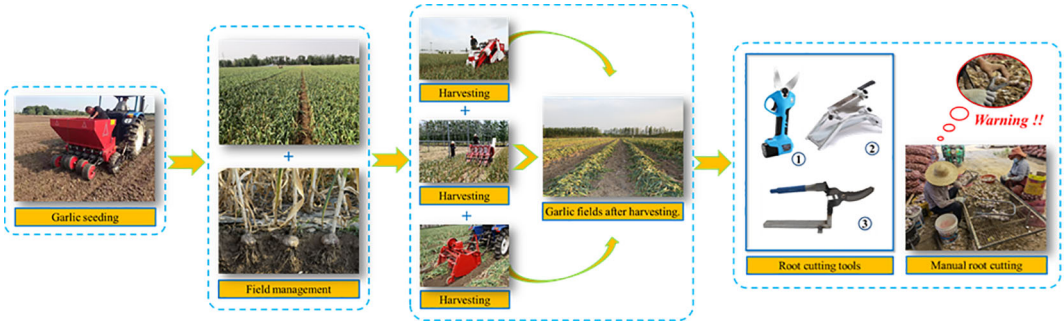


FIGURE 1
Garlic field production process (sowing, field management, harvesting).

detail by high-speed photographic tests. However, this mechanism was not highly adaptable to bulbs of different sizes. Cai (2019) developed a U-blade bite-type root-cutting mechanism for the primary processing of postharvest garlic, in which two symmetrically arranged special U-blades bite each other around their respective turning points to form a conical cutting track. By aligning the bottom of the bulb with the conical tip of the conical cutting track, both the garlic roots and root discs could be removed in one operation. However, this mechanism requires the manual orientation and placement of bulbs before cutting. Yang et al. (2022) proposed a noncontact positioning root-cutting method based on machine vision. They constructed a deep convolutional neural network with an improved YOLO v2 model to detect specific bulbs and roots, and predict the root-cutting line position. Two serrated disc cutters were automatically adjusted to the cutting line position for root-cutting. However, the complexity of the image acquisition environment requires improvements in image denoising technology. Using machine vision technology, Zhao et al. (2011) developed a special algorithm for the rapid positioning of the intersection of garlic bulb and roots. They used graphical user interface programming to develop an automatic and precise positioning system for the intersection of garlic bulb and roots, which provided a theoretical basis for the development and application of machine vision technology in an automated garlic root-cutting system. Thuyet et al. (2020) developed an automated grading and robotic-sorting system for root-trimmed garlic. The system used a deep convolutional neural network based image analysis technique to grade and sort root-trimmed garlic. This study provided a theoretical basis for the development and application of fully automated garlic root-cutting and bulb-sorting robots. The main garlic harvesting equipment being developed in China were garlic excavator and garlic seedling cutting combine harvester, while garlic combine harvesting equipment with root-cutting function is not yet available (Yu et al., 2021b). Erme (France), J.J. Broch (Spain) and Top Air (USA) are the main garlic harvesting equipment manufacturers in Europe and the United States. Their products include garlic excavators, garlic pickers, and stalk-cutting and baling-type garlic harvesters. However, these machines do not have a garlic root-cutting function. The HZ1 self-propelled garlic harvester by the Yanmar Company of Japan is the only garlic harvester with a root-cutting function from a developed country. The root-cutting mechanism consists of two superimposed serrated disc cutters. Because the serrated disc cutters are the fixed, the cutting height could not be adaptively adjusted according to the size of the bulb, resulting in poor root-cutting performance (Hou et al., 2021).

In recent years, the advances in computer simulation technology for agricultural machinery research have increased. The use of computer simulation technology has become an important step in the virtual simulation modeling and dynamic simulation analysis of crop mechanization harvesting to determine the interaction mechanism between harvesting machinery and crops, and to optimize the operating parameters. ADAMS is the most widely used software for the simulation and analysis of multibody system dynamics. Its rigid-flexible coupling modeling approach consists of three methods (Pu and Wu, 2009): the discrete beam method, modal neutral file (MNF) method, and AutoFlex method. The effects of the proposed method on crop-harvesting domains are discussed below. Using the discrete beam method, Wang et al. (2021) developed a flexible body model of the chrysanthemum stalk in ADAMS and fused it with a chrysanthemum picker to construct a rigid-flexible coupled numerical simulation model of the picker and chrysanthemum. The main factors affecting chrysanthemum picking were determined using dynamic simulation tests. Using the MNF method, Xie et al. (2020) completed the meshing of the sugarcane model in ANSYS, imported it into ADAMS, and fused it with the sugarcane top-breaking roller model to construct a rigid-flexible coupled numerical simulation model of the top-breaking rollers and sugarcane. The motion characteristics of sugarcane during top-breaking were studied using dynamic simulation tests. Shi et al. (2017) used the AutoFlex method to generate a flexible body model of the *Artemisia selengensis* (*A. selengensis*) stalk in ADAMS, which was fused with a cutter model of the *A. selengensis* harvester to construct a rigid-flexible coupled numerical simulation model of the cutter and *A. selengensis* stalk. Kinetic simulation analysis of the cutting process of the *A. selengensis* stalk was carried out, and the working parameters of the cutter were optimized by virtual simulation orthogonal tests. Yu et al. (2021a) used the AutoFlex method to transform a coffee stalk model into a flexible body in ADAMS, and constructed a simulation model of vibrating comb-type coffee threshing. The generalized force and sensor functions were used to monitor and control the shedding of coffee grains, and the dynamic simulation of the coffee threshing process was achieved. In addition, Chen et al. (2021) constructed finite element models of a garlic bulb and cutter based on ANSYS/LS-DYNA. Though the simulation analysis of root-disc cutting, they investigated the effects of the structural parameters of the cutter on the root-disc cutting force, and obtained the optimal structural parameters of the cutter. The abovementioned study serves as a reference of the present study in optimizing the floating root-cutting process of garlic using computer simulation technology.

To address the current problems of technological backwardness and poor operational quality of mechanized garlic root-cutting in China, this study combined the physical characteristics and agronomic requirements of garlic plants, and proposed an innovative floating root-cutting technology for garlic combine harvesters that enables the top alignment of bulb, adaptive profiling floating of cutter, and embedded cutting of roots. The kinematic and dynamic simulation analyses of the floating process of the root-cutting mechanism were conducted, and the optimal parameter combination of the root-cutting mechanism was obtained. The three main contributions of this study are as follows.

(1) The coordinate equations of the initial contact point of the bulb and a mathematical model of the floating displacement of the cutting component were established. The influence law and factors of the floating displacement of the cutting component in the two stages of floating were revealed. The theoretical basis for the optimal design of the root-cutting mechanism was provided.

(2) Using computer simulation techniques, the dynamic simulation analysis of the floating cutting process was conducted. The effects of each factor on the floating cutting performance were analyzed through single-factor simulation test. The key operating parameters of the root-cutting mechanism were optimized through virtual orthogonal test and fuzzy comprehensive evaluation. The computer simulation optimization method was provided for the optimal design of the root-cutting mechanism.

(3) Finally, the accuracy and reliability of the optimal parameter combination of the root-cutting mechanism were verified by simulation and field verification tests. The verification test results showed that the operating indexes of the root-cutting mechanism met the requirements of Chinese garlic field harvesting quality. The results of this study provide a reference for technological improvements and mechanism optimization of root-cutting for garlic combine harvesting in China.

2 Materials and methods

2.1 Structure of the root-cutting mechanism

The structure of the garlic root-cutting mechanism is shown in Figure 2 and the root-cutting mechanism mainly consists of a clamping chain, alignment chain, roller brush, front rotary cutter, extension spring, swing arm, double elastic guide plates, and horizontal disc cutting component (hereinafter referred to as "cutting component") (Yu et al., 2023). Among them, the clamping and alignment chains are in an inclined configuration. The double elastic guide plates were configured on both sides of the bulb running track and below the alignment chain. The double elastic guide plates with rubber pads were placed on the inside. The cutting component was configured below the double elastic guide plates, and is the main component for roots-bulb separation. It mainly consists of double protective fences, horizontal disc rotary cutter group, and DC motor. The double protective fences were arranged in parallel above the horizontal disc rotary cutter group. The exterior of the double protective fences could be fitted with rubber sleeves. One end of the swing arm was hinged to the frame, while the end was hinged to the cutting component. The two swing arms, cutting component and frame formed a parallel four-link structure. Therefore, during the floating cutting process, the cutting component moves in translation but did not rotate, thus ensuring that the spatial posture of the cutting component remained constant. The extension spring was connected to the frame at one end and to the swing arm at the other end; hence the cutting component produced elastic floating during the cutting process.

2.2 Working principle of the root-cutting mechanism

The garlic plant enters the clamping chain from the clamping chain feed inlet. The clamping chain holds the upper stalk of the plant and transports it backward. Then, the lower stalk of the plant enters the alignment chain and the plant is transported backward by the clamping action of the clamping chain and the pushing action of the alignment chain. As the plant is transport backward, the high-speed rotating roller brush removes the soil that adhered to the roots and neatly combs the disorganized roots. The high-speed rotating front rotary cutter, which acts on the lower part of the roots, completes the pre-cutting of the roots and makes them form a neat cross-section.

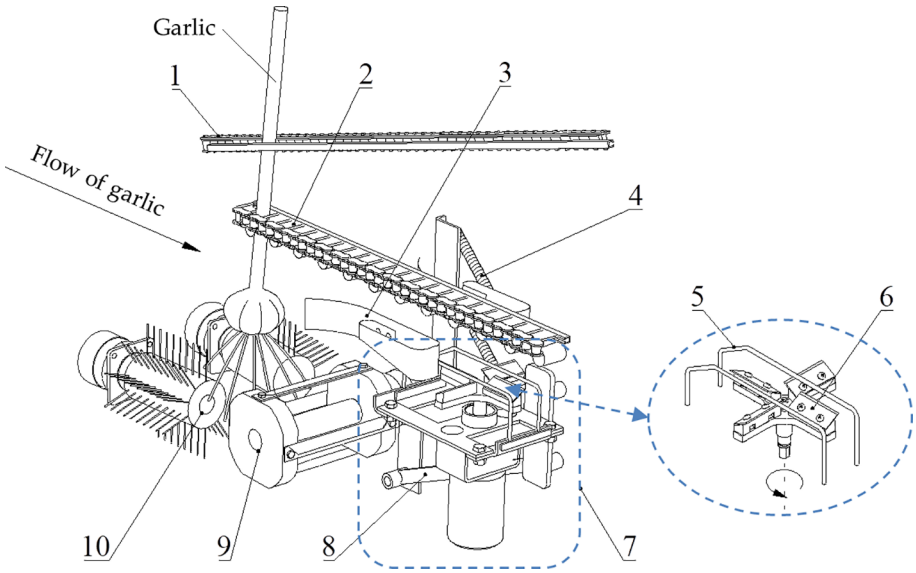


FIGURE 2
Structure of the root-cutting mechanism. 1. Clamping chain; 2. Alignment chain; 3. Double elastic guide plates; 4. Extension spring; 5. Double protective fences; 6. Horizontal disc rotary cutter group; 7. Horizontal disc cutting component; 8. Swing arm; 9. Front rotary cutter; 10. Roller brush.

Owing to the inclined configuration of the clamping and the alignment chains (i.e., the distance between the fronts of the clamping and alignment chains was small, while the distance between their ends was large) and given that the alignment chains restrict the movement of the plant in the forward and backward directions, the bulb gradually approached the alignment chain and completed the top alignment. Once the bulb was firmly attached to the alignment chain, it restricted the movement of the plant in both upward and downward directions. As the distance between the clamping and alignment chains increased, the bulb slid downward relative to the clamping chain; however, it always remained close to the alignment chain and was transported backward. Subsequently, the double elastic plates guided the bulb into the cutting channel to prevent its lateral deflection. The lower surface of the bulb was pressed against the double protective fences, which brought the cutting component close to it. The roots were simultaneously cut off by the horizontal disc rotary cutter group through the double protective fences, thus separating the roots from the bulb.

2.3 Kinematic analysis of the floating process of the root-cutting mechanism

The elastic floating of cutting component is an important guarantee for efficient and low damage cutting of the roots. Studying the kinematic characteristics of elastic floating of the cutting component provides a theoretical basis for optimizing the root-cutting mechanism.

From a kinematic perspective, the floating displacement of the cutting component and the position of the initial contact point between the bulb and cutting component are the key parameters reflecting the kinematic characteristics of the elastic floating of the cutting component (Yu et al., 2021c). Because the floating displacement of the cutting component determines the relative position of the bulb and the cutting component, it can be used to determine the distance between the rotary cutter group and root disc during the floating process. Simultaneously, the floating displacement of the cutting component determines the elongation of the extension spring, which affects the squeezing force applied to the lower surface of the bulb. The position of the initial contact point between the bulb and cutting component (hereinafter referred to as the initial contact point of the bulb) determines the floating displacement of the cutting component in two different floating phases and the point where the bulb is subjected to the maximum impact force.

In this section, the mathematical model of the floating displacement of the cutting component and the coordinate equations of the initial contact point of the bulb are established through theoretical analysis. The purpose is to clarify the influence law and factors of the floating displacement of the cutting component and the initial contact point of the bulb.

2.3.1 Determination of the initial contact point of the bulb

As shown in Figure 3, the hinge point of lower swing arm o is the origin of the coordinates, the conveying direction of the alignment chain is the positive direction along the x-axis, while the upward direction is the positive direction along y-axis. The bulb is considered an ellipsoid, with a long axis (bulb diameter) of $2a$ and short axis (bulb height) of $2b$. The equation of the bulb elliptical outline at the initial moment of contact between the bulb and cutting component is expressed as:

$$\frac{(x-x_{o'})^2}{a^2}+\frac{(y-y_{o'})^2}{b^2}=1 \tag{1}$$

where, $x_{o'}$ and $y_{o'}$ are the horizontal and vertical coordinates of the center point o' of the bulb (m).

As seen from Figure 2, the protective fence has two vertical edges at the front and back, a top beveled edge, and a top horizontal edge. The top beveled edge is the part that makes initial contact with the bulb and guides it to the top horizontal edge. As shown in Figure 3, a line mn, which coincides with the top beveled edge of the protective fence, was created. The slope k of the line mn remained constant during the floating process. At the initial moment of contact between the bulb and cutting component, the equation of line mn is expressed as:

$$y=kx+c \tag{2}$$

where, k is the slope of line mn and c is the value of the vertical coordinate of the intersection at the line mn and y-axis.

As seen from Figure 3, the tangent point between line mn and the bulb elliptical outline is the initial contact point i of the bulb. From Equations (1) and (2), the coordinates of the initial contact point i of the bulb are expressed as:

$$\begin{cases} x_i=\frac{a^2k}{\sqrt{a^2k^2+b^2}}+x_{o'} \\ y_i=-\frac{b^2}{\sqrt{a^2k^2+b^2}}+y_{o'} \end{cases} \tag{3}$$

where, x_i and y_i are horizontal and vertical coordinates of the initial contact point i of the bulb (m). As seen from Equation (3), the position of the initial contact point of the bulb is related to the size of the elliptical outline of the bulb, the coordinates of the center point of the bulb, and the slope of the top beveled edge of the protective fence. To reduce the bulb collision damage, the position of the initial contact point of the bulb should be shifted toward the bottom of the bulb during the cutting process.

2.3.2 Mathematical model of the floating displacement of the cutting component

After the bulb made contact with the top beveled edge of the protective fence, the plant continued to be transported backward, the bulb was pressed against the double protective fences to float the cutting component downward. In this study, the downward floating process of the cutting component was divided into two stages. In the first stage, the bulb makes contact with the top beveled edge of the protective fence, and the cutting component floats downward. In the second stage, the bulb makes contact with the top horizontal edge of the protective fence, and the cutting component floats downward.

In the first stage, relative sliding between the bulb and protective fence occurred. The points of action of the squeezing and frictional

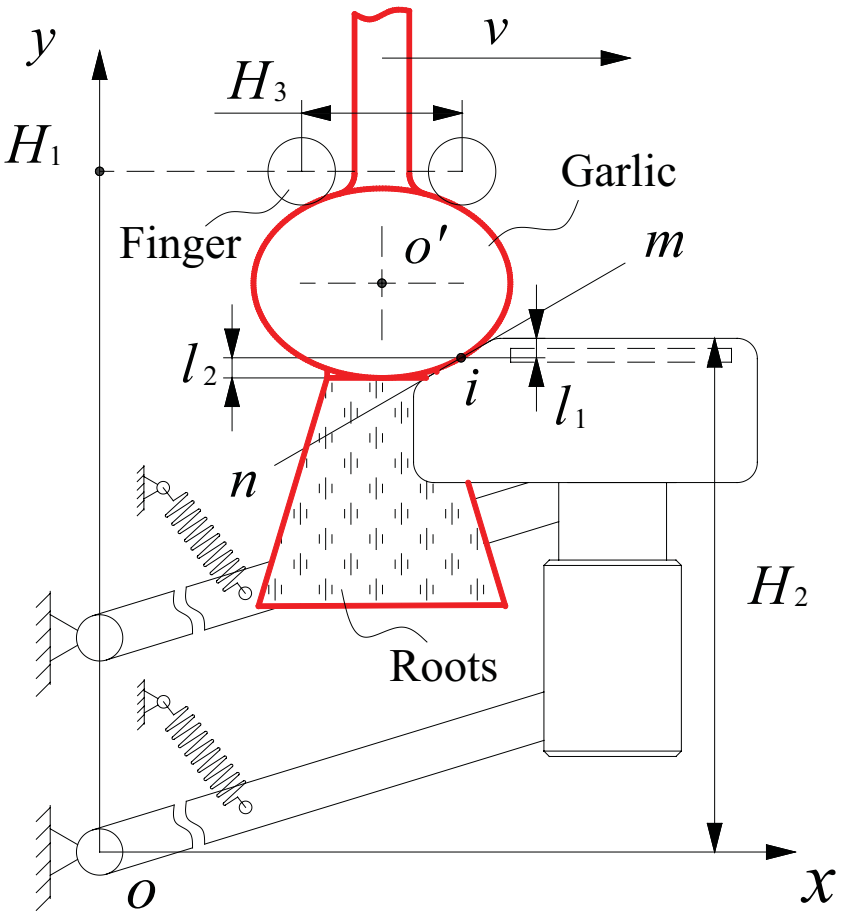


FIGURE 3
Schematic diagram of the relative position of the floating process of the cutting component.

forces exerted by the top beveled edge of the protective fence on the bulb remained the same. From the positional relationship shown in Figure 3, the floating displacement l_1 of the cutting component in the first stage of floating is expressed as:

$$l_1 = |H_2 - y_i| = \left| \frac{b^2}{\sqrt{a^2k^2 + b^2}} + (H_2 - y_{o'}) \right| \quad (4)$$

where, H_2 is the initial height of the top horizontal edge of the protective fence (m). In the mechanism design, when the slope of the top beveled edge of the protective fence was determined, the initial height H_2 should be greater than the height y_i of the initial contact point of the bulb to minimize the bulb collision damage. Otherwise, the bulb will collide with the bending area between the top beveled edge and top horizontal edge of the protective fence, which could damage the bulb. As seen from Equation (4), the first stage floating displacement l_1 was positively correlated with the bulb height $2b$ and initial height H_2, and negatively correlated with the bulb diameter $2a$ and the slope of the top beveled edge of the protective fence.

When the height of the top horizontal edge of the protective fence was gradually lowered and below the initial contact point i of the bulb, it meant the second stage of floating. The bulb disengaged from the top beveled edge and made contact with the top horizontal edge of the protective fence. The bulb was pressed against the top horizontal edge of the protective fence, which caused the cutting component to continue floating downward. The point of action of the squeezing and frictional forces exerted by the top horizontal edge of the protective fence on the bulb gradually lowered. When the lower surface of the bulb was above the horizontal edge of the protective fence, the squeezing and frictional forces on the bulb shifted toward the lowest point at the bottom of the bulb, and the cutting component floated downward to the lowest point. At this time, the floating displacement l_2 of the cutting component increased to the maximum value expressed as:

$$l_2 = |y_i - y_{o'} + b| = \left| b - \frac{b^2}{\sqrt{a^2k^2 + b^2}} \right| \quad (5)$$

As seen from Equation (5), the floating displacement l_2 of the cutting component in the second stage of floating is related to the bulb diameter $2a$, bulb height $2b$, and slope k of the top beveled edge of the protective fence. As the point of action of squeezing and frictional forces exerted by the top horizontal edge of the protective

fence on the bulb gradually lowered, the change of the point of action causes the change of the damaged part of the bulb and thus expanded the damaged area of the bulb. The larger the floating displacement l_2 of the cutting component, the wider the damaged area of the bulb would be, and the long the time of the squeezing and frictional forces on the bulb would be.

The above analysis not only obtained the coordinate equations of the initial contact point between the bulb and cutting component and the mathematical model of the floating displacement of the cutting component, but also revealed the influence law and factors of the floating displacement of the cutting component in the two stages of floating. These provide a theoretical reference for the analysis of bulb collision damage, extension spring elongation and preload force, and kinematic analysis of the elastic floating of the cutting component.

2.4 Numerical simulation modeling

ADAMS is a multi-body system dynamics analysis software that can simulate the floating cutting process of the root-cutting mechanism to significantly reduce the development cost and cycle time of the mechanism (Wang et al., 2022). This section used the ADAMS software to carry out the dynamic simulation of the floating cutting process of the root-cutting mechanism, with the aim of studying the effects of different operating parameters on the floating cutting performance to optimize the parameters of the root-cutting mechanism.

2.4.1 Models of garlic plant and root-cutting mechanism

Figure 4 shows the parts of a garlic plant, which include the stalk, bulb, and roots. As shown in Figure 5, the ANSYS and ADAMS coupling method was used to establish the flexible body model of the stalk and bulb (Yu et al., 2023). First, the 3D models of the stalk and bulb were imported into the Mechanical APDL module in ANSYS for meshing. Then, the modal neutral files of the stalk and bulb were imported into ADAMS using the coupling interface "ADAMS Flex: Create a Flexible Body". Finally, the model properties and modalities of each order were checked and verified. Jinxiang garlic is a typical garlic variety grown in China's main garlic producing areas. Flexible body model of stalk with a diameter of 13 mm and a length of 230 mm was constructed and a flexible body model of bulb with a diameter of 61 mm and a height of 45 mm was constructed. The densities of the stalk and bulb were 855.2 kg/m^3 and 1057 kg/m^3; the Poisson's ratios were 0.30 and 0.23; and the elastic modulus were 8.0×10^6 Pa and 2.38×10^7 Pa, respectively.

As shown in Figure 5, the ADAMS discrete beam method was used to construct a flexible body model for the garlic root group (Yu et al., 2023). First, a 3D model of the root group was created and then imported into ADAMS to discretize the single root one by one. The diameter and length of a single root were 1.6 mm and 63 mm; the density, Poisson's ratio, and elastic modulus of the root were 837.3 kg/m^3, 0.385, and 2.26×10^6 Pa, respectively. A single root was discretized into flexible beams.

A flexible body model of the garlic plant was constructed by fusing the stem, bulb and root group.

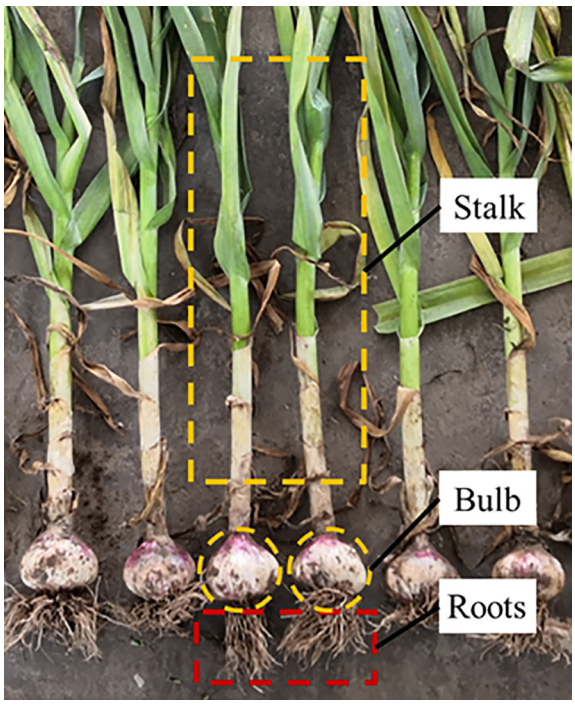


FIGURE 4
The composition of garlic plant.

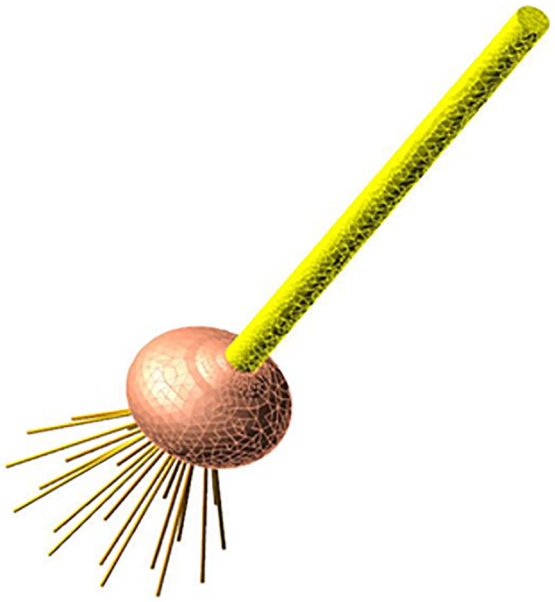


FIGURE 5
The flexible body model of the garlic plant.

Model of root-cutting mechanism: A 3D simplified model ofv the main components of the root-cutting mechanism was created by INVENTOR, and then imported into ADAMS along with the material properties of each component (Wang et al., 2013).

2.4.2 Rigid-flexible coupling model of the root-cutting mechanism and garlic plant

As shown in Figure 6, based on the kinematic analysis of the root-cutting mechanism, constraints, forces, drives, and dummy objects were added to build a rigid-flexible coupling model of the root-cutting mechanism and garlic plant in ADAMS (Yu et al., 2023). The stationary components were fixed to the reference system, while the revolute and translational pairs were used to define the constraints between components with rotational and linear relative motion relationships (Prastiyo and Fiebig, 2021). Simultaneously, dummy objects were constructed to assist the addition of motion constraints between the flexible body model of the garlic plant and interacting components. Contact constraints between the flexible body model of garlic plant and the interacting components were applied.

The impact function (IMPACT) was used to define the contact force between the bulb and protective fence, which has a contact type of Flex Body to Solid. The general expression of the impact function (Chen and Dun, 2012; Xie, 2019), which consists of the elastic force generated by the mutual extrusion of two objects and the damping force generated by their relative motion, is expressed as:

$$Q_{Impact}(x_i) = \begin{cases} 0 & x_i < x_0 \\ k_s(x_0 - x_i)^e - step(x_i, x_0 - d, 1, x_0, 0) \cdot c_{\max} & x_i \geq x_0 \end{cases} \quad (6)$$

where, Q_{Impact} is the contact force between two objects (N); x_i is the actual distance of the collision zone of two objects (mm); x_0 is the initial distance of the collision zone before the contact between two objects (mm); e is the contact force index; d is the penetration depth

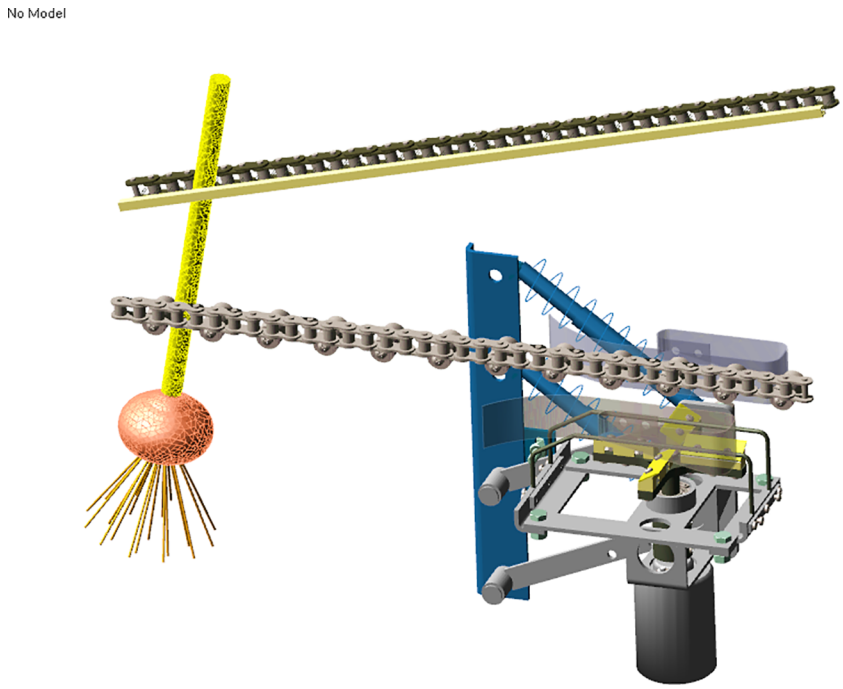


FIGURE 6
Rigid-flexible coupling model of the root-cutting mechanism and garlic plant.

(mm); c_{max} is the maximum damping coefficient (N·s/mm^{-1}); *step* is the step function that prevents the discontinuity of damping force during the collision; and k_s is the contact stiffness coefficient (N/mm). The Hertz collision contact model was used to calculate the contact stiffness coefficient k_s, which is expressed as:

$$k_s = \frac{16R^* E^{*2}}{9} \tag{7}$$

where, R^* is the equivalent relative radius of curvature (mm); $\frac{1}{R^*} = \frac{1}{R_1} + \frac{1}{R_2}$, where R_1 and R_2 are the radius of curvature of the bulb (17.25 mm) and protective fence (1.5 mm) in the collision zone; E^* is the equivalent elastic modulus (Mpa); $\frac{1}{E^*} = \frac{1-\mu_1^2}{E_1} + \frac{1-\mu_2^2}{E_2}$, where E_1 and E_2 are the elastic moduli of the bulb and protective fence, respectively. The elastic modulus of the protective fence is 2.1×10^5 MPa; μ_1 and μ_2 are the Poisson's ratios of the bulb and protective fence, respectively. The Poisson′s ratio of the protective fence is 0.3. According to Equation (7), the contact stiffness coefficient k_s between the bulb and protective fence is 1546 N/mm. The parameters of the impact function are listed in Table 1. (The parameters taken from reference (Chen and Dun, 2012) or the official recommended values of ADAMS software were tested and modified).

The drive was added to the clamping chain, alignment chain, and rotary cutter group. To reduce the residual root length after cutting, the root disc should be parallel to the rotary plane of the rotary cutter group when it cuts the roots. The above requirements could be achieved by the reasonable configuration of the speed ratio of the alignment and clamping chains. According to the research method of Yu et al. (2021b), the speed ratio between the alignment and the clamping chains was calculated as 1.015 based on the displacements of the alignment and the clamping chains when the garlic plant was jointly clamped and transported.

3 Dynamic simulation of the floating cutting process

The elastic expansion and contraction of the extension spring made the root-cutting mechanism drive the up and down floating of the cutting component. By optimizing the mechanical structure, motion parameters, and mechanical parameters of the extension spring, it would be possible to achieve the protection fence always closed to the lower bulb surface during the floating cutting process. After the basic structural form and parameters of the root-cutting mechanism were determined, the optimal floating cutting performance of the root-cutting mechanism can be achieved by optimizing the motion parameters of the root-cutting mechanism and the mechanical parameters of the extension spring.

TABLE 1 The parameters of the impact function.

Parameters	Values
Contact Stiffness Coefficient k_s (N·mm^{-1})	1546
Contact Force Index e	2.2
Damping Coefficient c_{max} (N·s·mm^{-1})	10
Penetration Depth d (mm)	0.1

In this section, using computer simulation techniques, the dynamic simulation study of the floating cutting process would be carried out in the rigid-flexible coupling numerical simulation model of root-cutting mechanism and garlic plant. Firstly, through single-factor simulation test, the influence law of the motion parameters of the root-cutting mechanism and the mechanical parameters of the extension spring on the floating cutting performance and its formation causes would be analyzed. Then, through virtual orthogonal test and fuzzy comprehensive evaluation, the optimal parameter combination of the root-cutting mechanism would be determined and verification test would be carried out.

3.1 Single-factor simulation test

3.1.1 Test index

The root excision rate is the main index used to evaluate the quality of garlic root-cutting. The key to the efficient cutting of garlic roots is to ensure that the protective fence is always close to the lower surface of the bulb during the floating cutting process. Because the rotary cutter group was positioned beneath the protective fence, and the cutting edge was close to the protective fence, the protective fence was close to the lower surface of the bulb, which ensured that the cutting edge was close to the root disc when cutting the root, and only a short length of the root remained after cutting.

The above theoretical analysis and preliminary experimental study showed that two main technical problems must be solved to ensure that the protective fence is always close to the lower surface of the bulb. First, when the bulb collides with the cutting component, the cutting component is ejected downward by the contact force and temporarily moves away from the lower surface of the bulb, making it impossible for the protective fence to be close to the lower surface of the bulb in the early stage of cutting. Second, after the completion of floating cutting process (i.e., the garlic is disengaged from the cutting component), it takes time for the cutting component to rise from the lowest point to the highest point of floating. Thus, when the next garlic enters the protective fence, the protective fence has yet to return to the highest point. This will affect the subsequent garlic to repeat the floating cutting operation successively. To quantitatively assess the floating cutting performance of the root-cutting mechanism, the floating displacement of the cutting component and angular velocity of swing arm reset (i.e., the angular velocity of the swing arm when the bulb disengaged from the cutting component and the cutting component floated upward) were used as the single-factor simulation test indexes.

3.1.2 Test factor

Based on kinematic analysis results of the floating process of the root-cutting mechanism and the pre-test results of the simulation, this study selected the key factors affecting the floating cutting performance, such as the extension spring preload force, extension

spring stiffness, and garlic conveying speed (i.e., clamping chain conveying speed) as the test factors to carry out the simulation analysis, and devoted to investigate the influence law of each factor on the floating cutting performance and its formation causes. The pre-test results of the extension spring preload force showed that to reduce the effects of collisions on the downward floating of the cutting component, the minimum preload force of 12 N was taken. The maximum preload force of 16 N was chosen to reduce the bulb collision damage caused by excessive preload force. Therefore, the extension spring preload forces were set as 12, 13, 14, 15 and 16 N. If the extension spring stiffness or preload force was different, the tension required to stretch the spring was different, then there may be an effect on the floating effect of the cutting component. To study the effects of the spring stiffness on the floating process, the test began with the minimum value of 15 N/m. Based on the results of the preliminary bench test, the selected extensile spring stiffness values were 15, 65, 115, 165 and 215 N/m. If garlic conveying speed was different, impact force between the bulb and the protective fence was different, then there may be an effect on the floating effect of the cutting component. The selected garlic conveying speeds were 0.8, 0.9, 1.0, 1.1 and 1.2 m/s.

3.1.3 Single-factor test results and analysis

(1) Effect of extension spring preload force on the floating cutting performance

In the single-factor test of the extension spring preload force, the garlic conveying speed and extension spring stiffness were set as 0.8 m/s and 215 N/m, respectively. The effect of the extension spring preload force on the floating displacement of the cutting component and angular velocity of swing arm reset was obtained through the simulation test, as shown in Figure 7. First, the changing process of the floating displacement of the cutting component during the simulation was analyzed, as shown in Figure 7A. When the simulation started, the displacement of the cutting component first remained constant. When the bulb collided with the top beveled edge of the protective fence, the cutting component was ejected and displaced downward. Then, the ejected cutting component gradually rose under the action of the extension spring. After the top horizontal edge of the protective fence made contact with the lower surface of the bulb and was pressed against it, and the displacement of the cutting component remained constant. When the root-cutting operation was completed, the bulb disengaged from the cutting component and the cutting component floated upward again. Finally, the cutting component returned to the height before the collision. The displacement of the cutting component remained constant while waiting for the subsequent garlic to repeat the operation.

In Figure 7A, the red circle marked the downward displacement caused by the cutting component being ejected downward. As seen from Figure 7A, the greater the preload force, the smaller the downward displacement caused by the cutting component being ejected downward would be, and the smaller the floating displacement of the cutting component would be. The possible reasons are as follows. The greater the preload force, the greater the resistance of the downward floating of the cutting component would be, the smaller the initial speed of the cutting component caused by the bulb collision would be, thus decreasing the floating displacement of the cutting component.

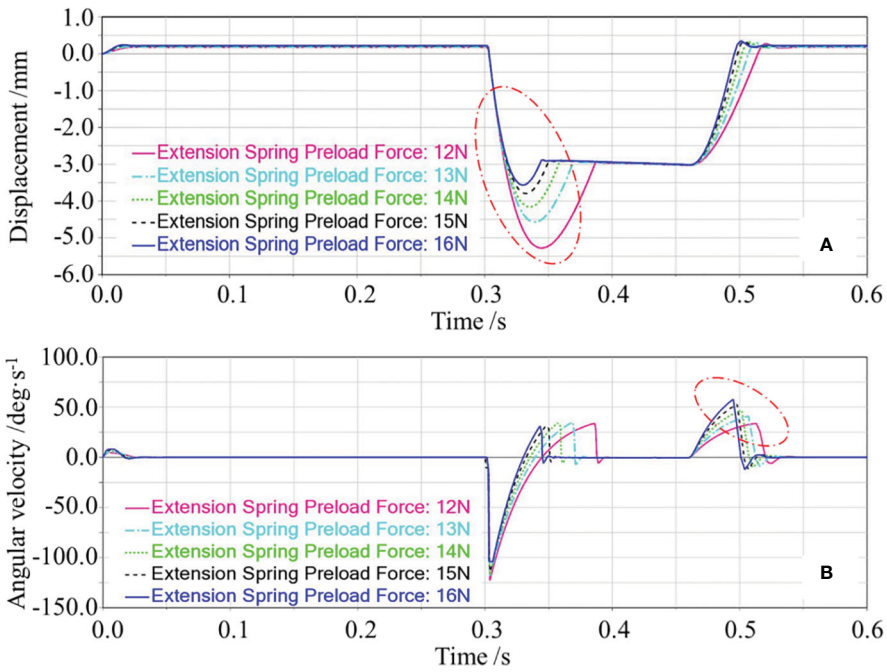


FIGURE 7
Effects of the extension spring preload force on the floating displacement of the cutting component and the angular velocity of swing arm reset. In **(A)**, the red circle marked the downward displacement caused by the cutting component being ejected downward; In **(B)**, the red circle marked the change in the angular velocity of swing arm during the reset phase of the cutting component.

When the root-cutting operation was completed, the bulb gradually disengaged from the cutting component; under the action of the extension spring, the cutting component gradually floated upward with a certain reset angular velocity. At this time, the angular velocity of swing arm reset determined the time of the cutting component reset, which affected the repeated cutting operation of the subsequent garlic. In Figure 7B, the red circle marked the change in the angular velocity of swing arm during the reset phase of the cutting component, i.e., the change in the angular velocity of swing arm reset. As shown in Figure 7B, the angular velocity of swing arm reset increased with the increase of the extension spring preload force. The possible reasons are as follows. The greater the preload force, the greater the tension of the extension spring on swing arm would be, and the higher the angular acceleration of swing arm reset would be, thus increasing the angular velocity of swing arm reset.

(2) Effect of extension spring stiffness on the floating cutting performance

In the single-factor test of the extension spring stiffness, the extension spring preload force and garlic conveying speed were set as 16 N and 0.8 m/s, respectively.

The effects of the extension spring stiffness on the floating displacement of the cutting component are shown in Figure 8A. The floating displacement of the cutting component decreased with the increase of the extension spring stiffness, but this trend was not obvious. The possible reasons are as follows. The greater the stiffness, the greater the tension required for the extension spring to elongate the same length would be, the greater the resistance of the floating downward of the cutting component would be, and the smaller the initial speed of the cutting component caused by the bulb collision would be, thus decreasing the floating displacement of the cutting component. When the stiffness increased to a certain value, the floating displacement of the cutting component was smaller, and the variation of the floating displacement of the cutting component caused by different stiffness was not obvious.

The effects of the extension spring stiffness on the angular velocity of swing arm reset are shown in Figure 8B. The angular velocity of swing arm reset was basically constant with the increase of the extension spring stiffness. The possible reasons are as follows. The five extension spring stiffnesses tested in the single-factor test corresponded to the five tensions of the extension spring. However, when the bulb disengaged from the cutting component and the cutting component gradually floated upward, compared with the gravity of the cutting component, the five tensions of the extension spring did not differ much. So the five angular accelerations of swing arm reset caused by five tensions of the extension spring did not differ much, thus leading to little change in the angular velocity of swing arm reset.

(3) Effect of garlic conveying speed on the floating cutting performance

In the single-factor test of the garlic conveying speed, the extension spring stiffness and extension spring preload force were set as 215 N/m and 16 N, respectively.

The effects of the garlic conveying speed on the floating displacement of the cutting component are shown in Figure 9A. The floating displacement of the cutting component increased with the

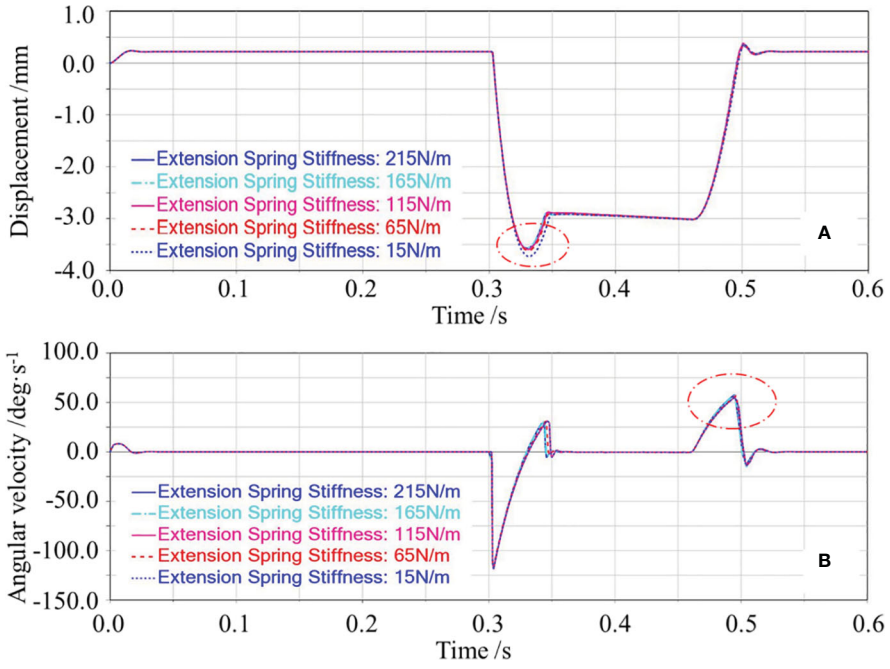


FIGURE 8
Effects of the extension spring stiffness on the floating displacement of the cutting component and the angular velocity of swing arm reset. In **(A)**, the red circle marked the downward displacement caused by the cutting component being ejected downward; In **(B)**, the red circle marked the change in the angular velocity of swing arm during the reset phase of the cutting component".

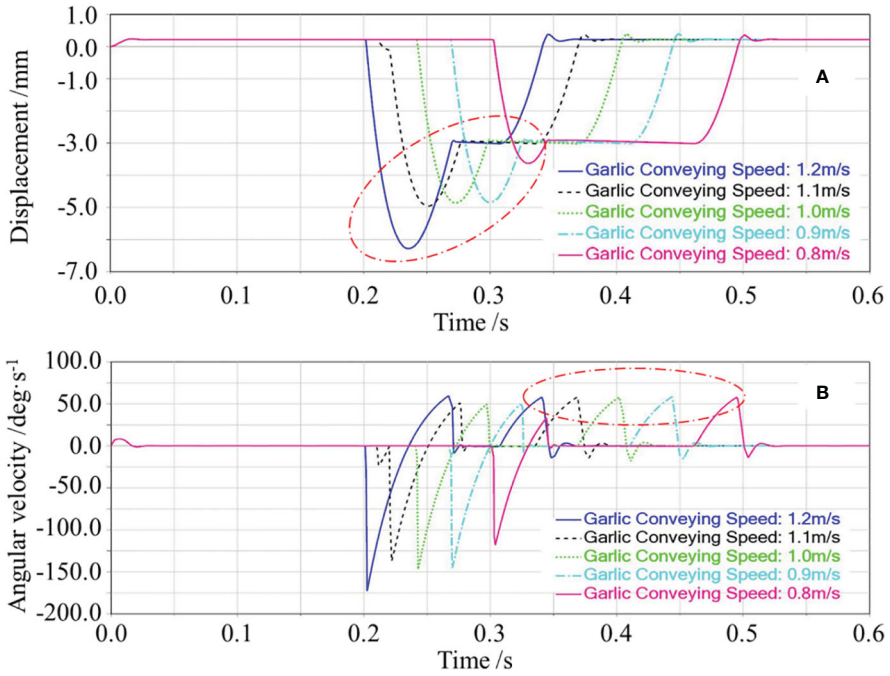


FIGURE 9
Effects of the garlic conveying speed on the floating displacement of the cutting component and the angular velocity of swing arm reset. In **(A)**, the red circle marked the downward displacement caused by the cutting component being ejected downward; In **(B)**, the red circle marked the change in the angular velocity of swing arm during the reset phase of the cutting component".

increase of the garlic conveying speed. The possible reasons are as follows. Because the cutting component was connected to the frame by an extension spring, at the moment of collision between the bulb and the cutting component, it is approximated as an elastic collision. According to the laws of conservation of momentum and conservation of energy, the higher the garlic conveying speed, the greater the speed of the cutting component after the collision would be, thus increasing the floating displacement of the cutting component.

The effects of the garlic conveying speed on the angular velocity of swing arm reset are shown in Figure 9B. The angular velocity of swing arm reset was basically constant with the increase of the garlic conveying speed. The possible reasons are as follows. As shown in Figure 9A, the displacement curve of the cutting component was already in the horizontal state before the bulb disengaged from the cutting component, which means that the protective fence was already close to the lower surface of the bulb. Therefore, for the five states of garlic conveying speed, the cutting component was at the same displacement and was subject to the same tension of the extension spring when the bulb disengaged from the cutting component. Thus, the angular velocity of swing arm reset was the same.

3.2 Virtual orthogonal test

Through above-mentioned single-factor simulation test, the influence law of the motion parameters of the root-cutting mechanism and the mechanical parameters of the extension spring on the floating cutting performance and its formation causes were investigated. In order to select the parameter combination that were as optimal as possible for the floating displacement of the cutting component and the angular velocity of swing arm reset, and further investigate the combined effects of extension spring preload force A, extension spring stiffness B, and garlic conveying speed C on the floating cutting performance, a three-factor, three-level virtual orthogonal test would be conducted in this section.

3.2.1 Design and scheme of virtual orthogonal test

To facilitate the quantitative assessment, the maximum floating displacement l of the cutting component at the moment of collision between the bulb and cutting component, and the maximum angular velocity ω of swing arm reset at the moment the bulb disengaged from the cutting component were selected as the quantitative assessment indexes of the virtual orthogonal test. The maximum floating displacement l of the cutting component is the increase of the displacement of the cutting component before the collision and the maximum displacement of the cutting component after the collision. Both indexes were automatically obtained by the ADAMS/postprocessor module. According to the above single-factor test results, the extensile spring preload force was selected as 14, 15, and 16 N, the extensile spring stiffness was 115, 165, and 215 N/m, and the conveying speed was 0.8, 0.9, and 1.0 m/s. The test was designed using the $L_9(3^4)$ orthogonal test table. The test scheme is presented in Table 2.

10.3389/fpls.2023.1168900

TABLE 2 Test schemes and results of virtual orthogonal tests.

Test Number	Extension Spring Preload Force *A*		Extension Spring Stiffness *B*	Garlic Conveying Speed *C*	Maximum Floating Displacement of The Cutting Component *l* (mm)	Maximum Angular Velocity of Swing Arm Reset ω (deg·s^{-1})
1	1(14)		1(115)	1(0.8)	4.45	45.50
2	1(14)		2(165)	2(0.9)	5.16	46.49
3	1(14)		3(215)	3(1.0)	5.90	47.13
4	2(15)		1(115)	2(0.9)	4.78	52.01
5	2(15)		2(165)	3(1.0)	5.45	52.06
6	2(15)		3(215)	1(0.8)	4.03	52.94
7	3(16)		1(115)	3(1.0)	5.15	56.43
8	3(16)		2(165)	1(0.8)	3.81	57.29
9	3(16)		3(215)	2(0.9)	4.39	58.58
Maximum Floating Displacement of The Cutting Component *l*	K_{11}	15.51	14.38	12.29		
	K_{12}	14.26	14.42	14.33		
	K_{13}	13.35	14.32	16.5		
	Range	2.16	0.10	4.21		
	Significance Sequence of Factors *C*>*A*>*B* Optimal Parameter Combination $C_1A_3B_3$					
Maximum Angular Velocity of Swing Arm Reset ω	K_{21}	139.12	153.94	155.73		
	K_{22}	157.01	155.84	157.08		
	K_{23}	172.3	158.65	155.62		
	Range	33.18	4.71	1.46		
	Significance Sequence of Factors *A*>*B*>*C* Optimal Parameter Combination $A_3B_3C_2$					

K_{11}~K_{13} were the sum of the maximum floating displacement of the cutting component measured for a factor at level 1, level 2 and level 3, respectively; K_{21}~K_{23} were the sum of the maximum angular velocity of swing arm reset measured for a factor at level 1, level 2 and level 3, respectively.

3.2.2 Orthogonal test results and analysis

Based on the experimental scheme above, virtual orthogonal simulation tests were conducted, and the test results are listed in Table 2. Using IBM SPSS Statistics 22 software, the data processing and statistical analysis of the test results were performed (Gong et al., 2012; Li and Zhang, 2015).

First, range analysis was conducted on the test results. The range analysis results are listed in Table 2. The significance of the factors affecting the maximum floating displacement of the cutting component decreased in the following order: garlic conveying speed, extension spring preload force, and extension spring stiffness. The optimal parameter combination was $C_1A_3B_3$; The significance of the factors affecting the maximum angular velocity of swing arm reset decreased in the following order: extension spring preload force, extension spring stiffness, and garlic conveying speed. The optimal parameter combination was $A_3B_3C_2$.

Second, ANOVA was performed on the test results. The ANOVA results are summarized in Table 3. The degree of influence of factors on the maximum floating displacement of the cutting component and the maximum angular velocity of swing arm reset varied. At the 95% confidence level, the effects of the garlic conveying speed and extension spring preload force on the maximum floating displacement were highly significant ($P< 0.01$), and the effects of extension spring stiffness on the maximum floating displacement were not significant ($P > 0.05$). At the 95% confidence level, the effects of the extension spring preload force on the maximum angular velocity were highly significant ($P< 0.01$), the effects of extension spring stiffness on the maximum angular velocity was significant ($0.01< P< 0.05$) and the effects of garlic conveying speed on the maximum angular velocity were not significant ($P > 0.05$).

3.2.3 Fuzzy comprehensive evaluation and comprehensive optimization

The above analysis revealed that the three factors (i.e., such as extension spring preload force, extension spring stiffness, and garlic

TABLE 3 ANOVA of virtual orthogonal tests.

Item	Source	Sum of Squares	Df	Mean Squares	*F* Value	*P* Value
Maximum Floating Displacement of The Cutting Component *l*	*A*	0.784	2	0.392	108.557	0.0091
	B	0.002	2	0.001	0.234	0.8105
	C	2.955	2	1.477	409.148	0.0024
	Error	0.007	2	0.004		
Maximum Angular Velocity of Swing Arm Reset *ω*	*A*	183.861	2	91.930	3395.053	0.0003
	B	3.743	2	1.872	69.122	0.0143
	C	0.441	2	0.220	8.137	0.1094
	Error	0.054	2	0.027		

P<0.01 (Highly significant), 0.01<P<0.05 (Significant), P>0.05 (Not significant).

conveying speed) had different orders of significance, different significance, and different optimal parameter combinations for the maximum floating displacement of the cutting component and maximum angular velocity of swing arm reset. In view of this, it is necessary to conduct a comprehensive evaluation of the two indexes obtained from each group of tests. Based on the comprehensive evaluation results of floating cutting performance for each group test, the comprehensive optimization of factors was carried out to obtain the parameter combination that were as optimal as possible for the floating displacement of the cutting component and the angular velocity of swing arm reset.

Owing to orders of magnitude and dimensions of the two indexes were different, the fuzzy comprehensive evaluation method (Xu et al., 2021) was used to establish the membership model of the two indexes and obtain the same orders of magnitude and dimensionless membership values. The maximum floating displacement of the cutting component was a small offset index, i.e., the smaller the maximum floating displacement, the better. Its membership model is shown in Equation (8). The maximum angular velocity of swing arm reset was a large offset index, i.e., the larger the maximum angular velocity, the better. Its membership model is shown in Equation (9).

$$t_{1n} = \frac{l_{\max} - l_n}{l_{\max} - l_{\min}} (n = 1, 2, \ldots 9) \tag{8}$$

$$t_{2n} = \frac{\omega_n - \omega_{\min}}{\omega_{\max} - \omega_{\min}} (n = 1, 2, \ldots 9) \tag{9}$$

where t_{1n} and t_{2n} are the membership values of the maximum floating displacement l of the cutting component and the maximum angular velocity ω of swing arm reset for the nth test, respectively; $l_{\max}$ and $l_{\min}$ are the maximum and minimum values of index l; l_n is the value of index l for the nth test; $\omega_{\max}$ and $\omega_{\min}$ are the maximum and minimum values of index ω; ω_n is the value of index ω for the nth test; n is the test number. The membership values t_{1n} and t_{2n} were obtained from Equations (8) and (9), respectively, and are listed in Table 4.

The fuzzy relationship matrix $\boldsymbol{T_n}$ was constructed from the membership values of the two indexes. The fuzzy relationship matrix $\boldsymbol{T_n}$ is expressed as:

$$T_n = \begin{pmatrix} t_{11} & \cdots & t_{19} \\ \vdots & \ddots & \vdots \\ t_{21} & \cdots & t_{29} \end{pmatrix} \tag{10}$$

TABLE 4 Comprehensive scores of the two indexes.

Test Number	Membership values of the maximum floating displacement of the cutting component t_{1n}	Membership values of the maximum angular velocity of swing arm reset t_{2n}	Comprehensive scores *U*
1	0.694	0	0.416
2	0.354	0.076	0.243
3	0	0.125	0.050
4	0.536	0.498	0.521
5	0.215	0.502	0.330
6	0.895	0.569	0.765
7	0.359	0.836	0.550
8	1	0.901	0.960
9	0.722	1	0.833

This test was dedicated to reducing the maximum floating displacement of the cutting component and the maximum angular velocity of swing arm reset. According to the importance of the two indexes, the weight of the maximum floating displacement of the cutting component was set as 0.6 and the weight of the maximum angular velocity of swing arm reset was set as 0.4. Thus, the weight assignment set $\boldsymbol{W}$ was constructed as $\boldsymbol{W}$ = [0.6 0.4].

According to the fuzzy relationship matrix $\boldsymbol{T_n}$ and the weight assignment set $\boldsymbol{W}$, the comprehensive score set $\boldsymbol{U}$ was obtained by fuzzy transformation, where $\boldsymbol{U}$=W·T_n. The comprehensive scores of floating cutting performance for each test were obtained, as listed in Table 4. The higher the comprehensive score, the better the performance of the test scheme would be.

In this study, the comprehensive scores of floating cutting performance were used as the comprehensive optimization basis of the root-cutting mechanism. The Shapiro-Wilk test was conducted using the IBM SPSS Statistics 22 software for comprehensive scores of floating cutting performance of the root-cutting mechanism in Table 4. The P value of the Shapiro-Wilk test was 0.957 ($P > 0.05$). The results indicate that the comprehensive scores of floating cutting performance of the root-cutting mechanism obtained from nine groups of simulation tests conformed to a normal distribution. Range analysis was performed on the comprehensive scores. From the range analysis results in Table 5, it could be seen that the significance of the factors affecting the floating cutting performance decreased in the following order: extension spring preload force, garlic conveying speed, and extension spring stiffness. The optimal parameter combination was $A_3C_1B_3$ (i.e., extension spring preload force of 16 N, garlic conveying speed of 0.8 m/s, and extension spring stiffness of 215 N/m). Based on the range analysis results, a radar diagram that was drawn to visually describe the effects of factors on the comprehensive scores is shown in Figure 10. As seen from Figure 10, the comprehensive score is positively correlated with the extension spring preload force and stiffness and negatively correlated with the garlic conveying speed. ANOVA was performed on the comprehensive scores. The ANOVA results are shown in Table 6, which revealed that the effects of the extension spring preload force and garlic conveying speed on the comprehensive score were highly significant ($P< 0.01$), while the effect of extension spring stiffness on the comprehensive score was not significant ($P > 0.05$) at the 95% confidence level.

3.2.4 Simulation verification of optimization results

In order to verify the accuracy and reliability of the above optimization results, we needed to measure the performance of the optimal parameter combination through simulation test. The simulation verification test was conducted in the rigid-flexible coupling model of the root-cutting mechanism and garlic plant constructed in section 2.4. Same as the above virtual orthogonal test, the simulation model used the impact function method (IMPACT) to define the contact force between the bulb and protective fence, its contact type was Flex Body to Solid, the contact stiffness coefficient was 1546 N/mm, the contact force index was 2.2, the damping coefficient was 10 N·s/mm, the penetration depth was 0.1 mm, and the speed ratio between the alignment and the clamping chains was calculated as 1.015. In the simulation verification test, the optimal parameter combination was used, and the optimal parameter combination was a garlic conveying speed of 0.8 m/s, an extension spring preload force of 16 N, and an extension spring stiffness of 215 N/m. After the above parameters were set, the simulation verification test was carried out.

Through simulation verification test, it was determined that the maximum floating displacement of the cutting component was 3.74 mm and the maximum angular velocity of swing arm reset was 57.93 deg/s under the condition of the optimal parameter combination. The orthogonal test results of groups 1-9 in Table 2 and the optimal parameter combination test results were selected to form 10 groups of test results, as shown in Table 7. Then, a fuzzy comprehensive evaluation of the 10 groups of test results was performed. Membership models of the maximum floating displacement of the cutting component and the maximum angular velocity of swing arm reset used for the 10 groups of test results in Table 7 are shown in Equations (11) and (12), respectively. The membership values s_{1n} and s_{2n} were obtained from Equations (11) and (12), respectively, and are listed in Table 7. The fuzzy relationship matrix $\boldsymbol{S_n}$ was constructed from the membership values s_{1n} and s_{2n} is shown in Equation (13).

$$s_{1n} = \frac{l'_{\max} - l'_n}{l'_{\max} - l'_{\min}} (n = 1, 2, \ldots . 10) \tag{11}$$

TABLE 5 Range analysis of comprehensive scores.

Item	Extension Spring Preload Force *A*	Extension Spring Stiffness *B*	Garlic Conveying Speed *C*
K_{31}	0.709	1.487	2.141
K_{32}	1.616	1.533	1.597
K_{33}	2.343	1.648	0.930
Range	1.634	0.161	1.211
Significance Sequence of Factors	*A*>*C*>B		
Optimal Parameter Combination	$A_3C_1B_3$		

K_{31}~K_{33} were the sum of the comprehensive scores for a factor at level 1, level 2 and level 3, respectively.

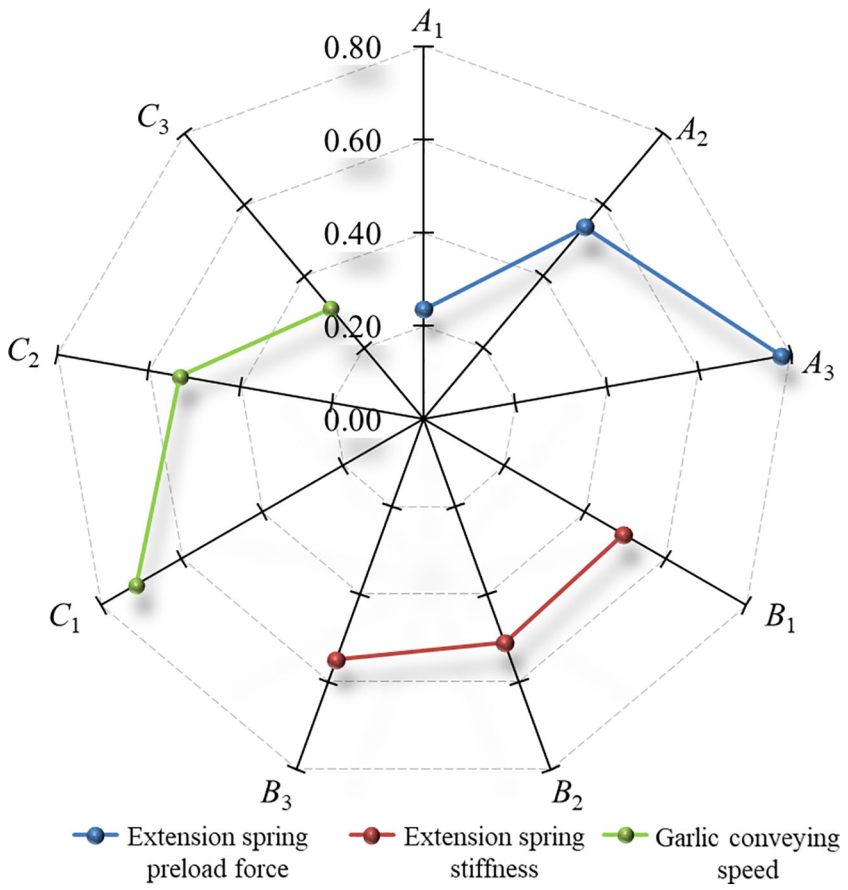


FIGURE 10
Effects of factors on the comprehensive scores for the floating cutting performance. Note: A_1~A_3 are the level 1, level 2 and level 3 of the extension spring preload force, respectively; B_1~B_3 are the level 1, level 2 and level 3 of the extension spring stiffness, respectively; C_1~C_3 are the level 1, level 2 and level 3 of the garlic conveying speed, respectively; The nine radial axes in the figure are the comprehensive scores for the floating cutting performance.

$$s_{2n} = \frac{\omega'_n - \omega'_{\min}}{\omega'_{\max} - \omega'_{\min}} (n = 1, 2, \ldots . 10) \tag{12}$$

$$S_n = \begin{pmatrix} s_{11} s_{12} s_{13} s_{14} s_{15} s_{16} s_{17} s_{18} s_{19} s_{110} \\ s_{21} s_{22} s_{23} s_{24} s_{25} s_{26} s_{27} s_{28} s_{29} s_{210} \end{pmatrix} \tag{13}$$

Where, s_{1n} and s_{2n} are the membership values of the maximum floating displacement l of the cutting component and the maximum angular velocity ω of swing arm reset for the nth test in Table 7, respectively; $l'_{\max}$ and $l'_{\min}$ are the maximum and minimum values of index l in Table 7; l'_n is the value of index l for the nth test in Table 7; $\omega'_{\max}$ and $\omega'_{\min}$ are the maximum and minimum values of index ω in Table 7; ω'_n is the value of index ω for the nth test in Table 7; n is the test number; s_{11}-s_{110} are the membership values of the maximum floating displacement of the cutting component for tests of groups 1-10 in Table 7, respectively; s_{21}-s_{210} are the membership values of the maximum angular velocity of swing arm reset for tests of groups 1-10 in Table 7, respectively.

Same as the above virtual orthogonal test, the weight of the maximum floating displacement of the cutting component was set as 0.6 and the weight of the maximum angular velocity of swing arm reset was set as 0.4. Thus, the weight assignment set $\boldsymbol{M}$ was constructed as $\boldsymbol{M}$ = [0.6 0.4]. By fuzzy transformation, we get the comprehensive score set $\boldsymbol{V}$, where $\boldsymbol{V}$=$\boldsymbol{M}$·$\boldsymbol{S}_n$. Through the above

TABLE 6 ANOVA of comprehensive scores.

Source	Sum of Squares	Df	Mean Squares	*F* Value	*P* Value
Extension Spring Preload Force *A*	0.447	2	0.223	1301.815	0.0009
Extension Spring Stiffness *B*	0.005	2	0.002	12.873	0.0777
Garlic Conveying Speed *C*	0.245	2	0.123	716.903	0.0016
Error	0.0004	2	0.0002		

TABLE 7 Comprehensive scores of floating cutting performance based on 10 groups of test results.

Parameter combination		Maximum Floating Displacement of The Cutting Component l (mm)	Membership values of the maximum floating displacement of the cutting component based on 10 groups of test results s_{1n}	Weight of the maximum floating displacement of the cutting component	Maximum Angular Velocity of Swing Arm Reset ω (deg·s^{-1})	Membership values of the maximum angular velocity of swing arm reset based on 10 groups of test results s_{2n}	Weight of the maximum angular velocity of swing arm reset	Comprehensive scores based on 10 groups of test results V
Orthogonal test parameter combinations in Table 2	1	4.45	0.671	0.6	45.50	0.000	0.4	0.403
	2	5.16	0.343		46.49	0.076		0.236
	3	5.90	0.000		47.13	0.125		0.050
	4	4.78	0.519		52.01	0.498		0.510
	5	5.45	0.208		52.06	0.502		0.326
	6	4.03	0.866		52.94	0.569		0.747
	7	5.15	0.347		56.43	0.836		0.543
	8	3.81	0.968		57.29	0.901		0.941
	9	4.39	0.699		58.58	1.000		0.819
Optimal parameter combination	10	3.74	1.000		57.93	0.950		0.980

The 10 groups of test results, i.e., the test results of the optimal parameter combination and the test results of 9 groups of orthogonal tests in Table 2; Parameter combinations 1-9 corresponded to orthogonal test parameter combinations 1-9 in Table 2, respectively; Parameter combination 10 was the optimal parameter combination.

calculations, the comprehensive scores of floating cutting performance for 10 groups of test were obtained, as listed in Table 7. As can be seen from Table 7, the comprehensive score of the optimal parameter combination was 0.980, which was higher than the comprehensive scores of the other 9 groups of orthogonal tests. The floating cutting performance of the optimized root-cutting mechanism was better than other parameter combinations. Therefore, the comprehensive optimization results are reliable.

3.3 Field verification test

3.3.1 Test condition

A test bench of the garlic root-cutting mechanism was constructed to conduct a field verification test, with the purpose of verifying the accuracy of the numerical simulation model and reliability of the optimal parameter combination. The test bench is shown in Figure 11, in which each component is individually driven by speed-controlled motors. Parameters, such as the tilt angle and height of the test stand could be adjusted as required. Garlic plants were obtained from an experimental field in Jinxiang County, Shandong Province, China.

3.3.2 Design and scheme of field verification test

The two indexes assessed in the simulation tests, such as the maximum floating displacement of the cutting component and maximum angular velocity of swing arm reset, correspond to the root excision rate, which is the actual index of the root-cutting quality. The smaller the maximum floating displacement of the cutting component and the greater the angular velocity of swing arm reset, the easier it is to ensure that the protective fence was closely attached to the lower bulb surface for the floating root-cutting process. In this way, the cutting edge is close to the root disc of the garlic, the remaining length of the root after cutting is shorter, and the root excision rate is higher. Therefore, the field verification tests were conducted for nine parameter combinations of the virtual orthogonal test listed in Table 2 and the optimal parameter combination in Section 3.2.3, using the root excision rate as the root-cutting quality index. The field verification tests were conducted sequentially for the nine parameter combinations listed in Table 2.

The garlic plants selected for the test were characterized by good uprightness, well-developed roots, and uniform bulb maturity. Before the tests, the garlic plants were tidied up and cleaned up to remove residual film, large pieces of soil, and debris on the roots. Each test was fed 90 garlic plants, and repeated three times for each parameter combination. The root excision rate was measured in the three tests, and the average value of the three tests was taken as the test result.

The root excision rate is the ratio of the total mass of the roots removed to the total mass of all roots. The root excision rate equation is expressed as:

$$P = Q_1/(Q_1 + Q_2) \tag{14}$$

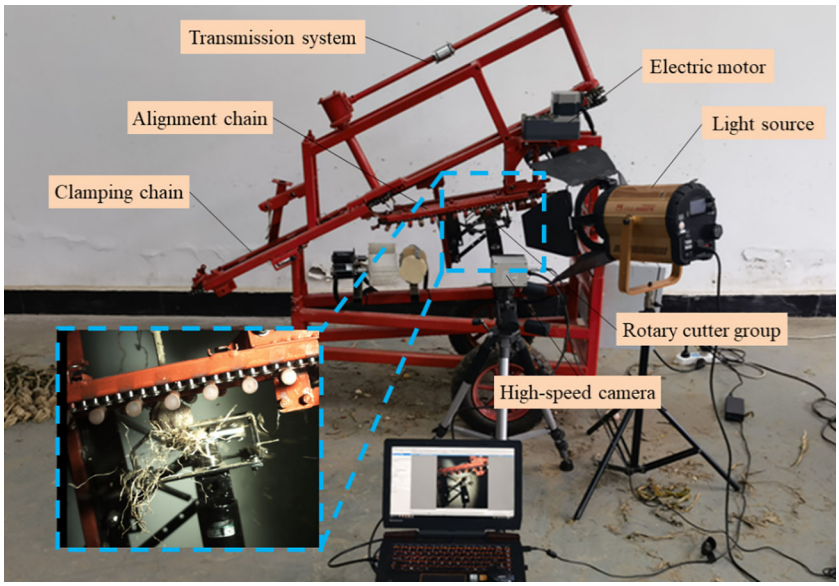


FIGURE 11
Test bench of garlic root-cutting mechanism.

where, P is the root excision rate (%); Q_1 is the total mass of removed roots (g); Q_2 is the total mass of remaining roots (g). Before weighing, the soil and debris on the root were also removed.

3.3.3 Results and analysis of field verification test

The results of the verification test are summarized in Table 8. Using IBM SPSS Statistics 22 software, the data processing and statistical analysis of the test results were performed.

The Shapiro-Wilk test was performed for the root excision rate in Table 8. The P value of the Shapiro-Wilk test was 0.788 ($P > 0.05$). The results indicate that the root excision rate obtained from 9 groups of field verification tests conformed to a normal distribution. Range analysis was performed on the test results, and the results are listed in Table 8. The significance of the factors affecting the root excision rate decreased in the following order: extension spring preload force, garlic conveying speed, and extension spring stiffness. The optimal

TABLE 8 Test schemes and results of field verification test.

Test number		Extension Spring Preload Force *A*	Extension Spring Stiffness *B*	Garlic Conveying Speed *C*	Root excision rate *P* (%)
1		1(14)	1(115)	1(0.8)	85.03
2		1(14)	2(165)	2(0.9)	83.77
3		1(14)	3(215)	3(1.0)	81.96
4		2(15)	1(115)	2(0.9)	87.61
5		2(15)	2(165)	3(1.0)	85.32
6		2(15)	3(215)	1(0.8)	90.56
7		3(16)	1(115)	3(1.0)	88.23
8		3(16)	2(165)	1(0.8)	92.37
9		3(16)	3(215)	2(0.9)	91.05
Root excision rate *P*	K_1	250.76	260.87	267.96	
	K_2	263.49	261.46	262.43	
	K_3	271.65	263.57	255.51	
	Range	20.89	2.7	12.45	
	Significance Sequence of Factors $A>C>B$ Optimal Parameter Combination $A_3C_1B_3$				

K_{31}~K_{33} were the sum of the root excision rate for a factor at level 1, level 2 and level 3, respectively.

parameter combination was $A_3C_1B_3$. Based on the range analysis results, a radar diagram that was drawn to visually describe the effects of factors on the root excision rate is shown in Figure 12. As seen from Figure 12, the root excision rate was positively correlated with the extension spring preload force and stiffness and negatively correlated with the garlic conveying speed. ANOVA was performed on the test results. The ANOVA results are shown in Table 9. At the 95% confidence level, the effect of the extension spring preload force and garlic conveying speed on the root excision rate was highly significant ($P< 0.01$), and the effects of extension spring stiffness on the root excision rate were not significant ($P > 0.05$).

The comparison of the results of the virtual orthogonal and field verification tests showed that the optimal parameter combinations obtained from the two tests were the same. The influence law and significance of each factor on the root excision rate and comprehensive score for floating cutting performance were also the same. Thus, the accuracy of the numerical simulation model and the reliability of the simulation results were verified.

In order to verify the accuracy and reliability of the optimization results, the optimal parameter combination of $A_3C_1B_3$, which was optimized by the field verification tests, was used as the test condition to determine the root excision rate by the field test. To eliminate random errors, the test was repeated three times and the average value was taken as the test result. When the extension spring preload force was 16 N, garlic conveying speed was 0.8 m/s, and extension spring stiffness was 215 N/m, the root excision rate was 92.72% as measured in the field test.The test result (i.e., root cutting rate of 92.72%) was better than those of the other nine orthogonal test schemes listed in Table 8, which meets the requirements of Chinese garlic field harvesting quality. The bulbs

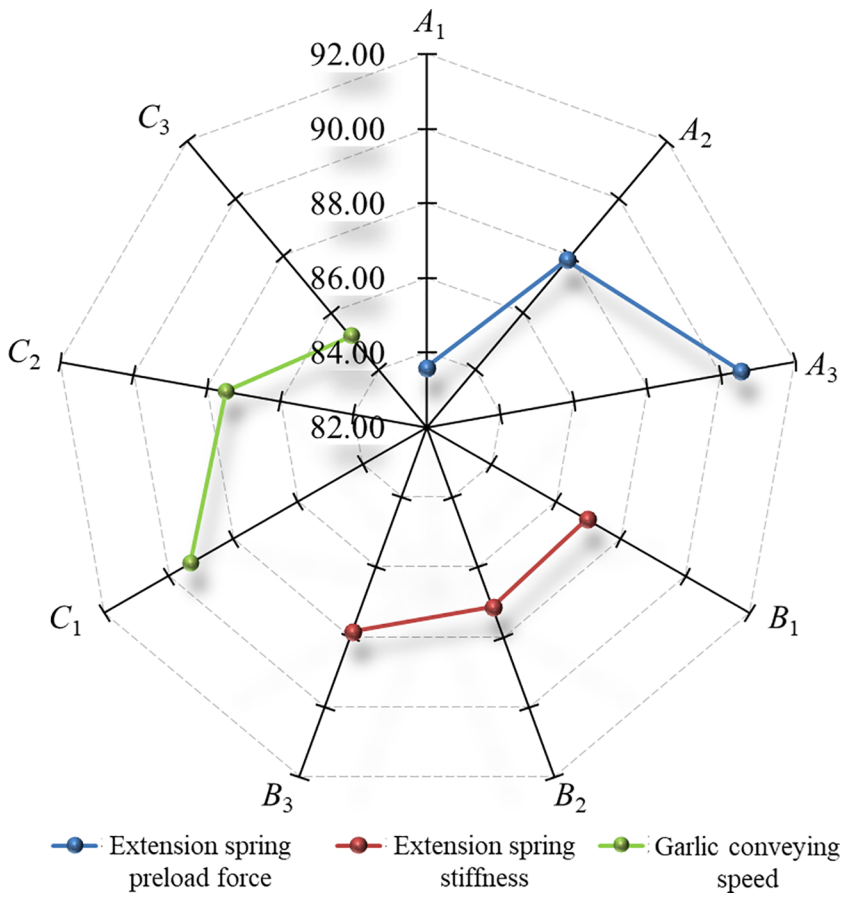


FIGURE 12
Effects of factors on the root excision rate. The nine radial axes in the figure were the root excision rate.

TABLE 9 ANOVA of field verification test.

Source	Sum of Squares	Df	Mean Squares	*F* Value	*P* Value
Extension Spring Preload Force *A*	73.892	2	36.946	328.963	0.0030
Extension Spring Stiffness *B*	1.343	2	0.672	5.981	0.1433
Garlic Conveying Speed *C*	25.941	2	12.971	115.488	0.0086
Error	0.225	2	0.112		

after root-cutting are shown in Figure 13. Therefore, the optimal parameter combination $C_1A_3B_3$ is reliable.

In addition, the bulb damage caused by the cutting operation was studied by theoretical analysis in Section 2.3. To comprehensively assess the applicability and reliability of the optimal parameter combination for the root-cutting mechanism, the bulb damage rate during the cutting operation was tested. The test was repeated three times and the average value was taken as the test result. Under the optimal parameter combination, the average value of bulb damage rate was less than 6.0%, which meets the requirements of Chinese garlic field harvesting quality.

4 Discussion

In this study, the kinematic characteristics of the floating cutting process of the root-cutting mechanism, and technical methods for improving the root excision rate were investigated in detail. The optimal parameter combination for root-cutting mechanism was obtained through simulation analysis and verified by simulation and field tests. Based on kinematic analysis results of the floating process of the root-cutting mechanism and the pre-test results of the simulation, this study selected the extension spring preload force, extension spring stiffness, and garlic conveying speed as the test factors to carry out the simulation analysis, and devoted to investigate the influence law of each factor on the floating cutting performance and its formation causes. The effects of the different garlic varieties, different bulb sizes, and garlic plant collapse on the floating cutting performance of the root-cutting mechanism were not further investigated because of the short garlic harvesting period in China. Therefore, all test results only apply to Jinxiang garlic with good uprightness and uniform bulb maturity. Factors such as garlic variety, plant growth, plant collapse, and the assembly with the combine harvester need to be further studied. In the next step, the

FIGURE 13
The bulbs after root-cutting.

root-cutting mechanism will be configured in the garlic combine harvester to perform field tests. The adaptability of the root-cutting mechanism to different garlic varieties, plant collapse and bulb maturity will be examined, and the root-cutting mechanism will be further optimized.

The tests showed that because there were only two protective fences between the bulb and rotary cutter group, the garlic roots easily passed through the protective fences and were fully cut by the rotary cutter group, resulting in a high root excision rate. At the same time, it was observed in the tests that the risk of cutting the bulb during the root-cutting process was effectively reduced by using protective fences that formed a physical barrier between the bulb and the rotary cutter group, which separates the bulb from the rotary cutter group. However, we found that the reliability of the rotary cutter group and protective fences still need to be further improved. After a long period of operation, the rotary cutter group was worn out, while the protective fences were deformed. In future product development, it is recommended that the rotary cutter group and the protective fences be made of high-quality high-speed steel and high quality spring steel respectively with strict heat treatment process to further improve the reliability of the root-cutting mechanism.

5 Conclusions

In response to the problems of technological backwardness and poor operational quality of mechanized root-cutting in garlic field harvesting in China, we combined the physical characteristics and agronomic requirements of garlic plants to develop a new floating root-cutting technology for garlic combine harvesters. The coordinate equations of the initial contact point of the bulb and a mathematical model of the floating displacement of the cutting component were established. Using computer simulation techniques, the influence law of the garlic conveying speed, extension spring preload force and stiffness on the floating cutting performance and its formation causes were analyzed, and the optimal parameter combination of the root-cutting mechanism was determined. The test results showed that the floating cutting performance was positively correlated with the extension spring preload force and stiffness and negatively correlated with the garlic conveying speed. Among them, the effects of extension spring preload force and garlic conveying speed on the floating cutting performance were highly significant. The simulation and field validation tests showed that the garlic floating root-cutting technology enabled the top alignment of bulb, adaptive profiling floating of cutter, and embedded cutting of roots and its technology scheme was feasible and effective and achieved the best operating performance. The floating cutting performance of the root-cutting mechanism met the requirements of Chinese garlic field harvesting quality.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material. Further inquiries can be directed to the corresponding authors.

Author contributions

Conceptualization, MY and ZH. Methodology, FG. Formal analysis, BP. Data curation, KY. Writing—original draft preparation, ZY and KY. Writing—review and editing, ZY and YZ. All authors contributed to the article and approved the submitted version.

Funding

This research was funded by the Modern Agricultural Machinery Equipment and Technology Demonstration and Promotion of Jiangsu Province (NJ2021-22), National Key R&D Program of China (2017YFD0701305-02).

Acknowledgments

We are grateful to the teachers for their academic guidance. We also sincerely appreciate the anonymous reviewers for their helpful suggestions.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Cai, S. L. (2019). Design and experimental study on automatic cutting root production line of garlic. (Hangzhou, China: Zhejiang Sci-Tech University). 1–77. doi: 10.27786/d.cnki.gzjlg.2020.000752

Chen, H. T., and Dun, G. Q. (2012). Optimization of parameters for soybean lifter based on dynamic simulation of virtual prototype. *Trans. Chin. Soc Agric. Eng.* 28, 23–39. doi: 10.3969/j.issn.1002-6819.2012.18.004

Chen, J. N., Zhou, B. S., Jia, J. M., Chen, Z. W., Yu, C. N., and Cai, S. L. (2021). Design and parameters optimization of root cutting tool based on garlic numerical simulation model. *J. Food Process Eng.* 44, 1–11. doi: 10.1111/jfpe.13753

FAOSTAT (2020a) *Crops and livestock products: Garlic Area Harvested*. Available at: https://www.fao.org/faostat/en/#data/QCL.

FAOSTAT (2020b) *Crops and livestock products: Garlic Export Quantity*. Available at: https://www.fao.org/faostat/zh/#data/TCL.

Gong, J., Shi, P. C., and Li, C. Y. (2012). Application of SPSS software in multivariate analysis of variance. *Agric. Network Info.* 4, 31–33.

Hou, J. L., Li, C., Zhang, Z. L., Li, T. H., Li, Y. H., and Wu, Y. Q. (2021). Design and test of double-row walking garlic combine harvester. *Trans. Chin. Soc Agric. Eng.* 36, 1–11. doi: 10.11975/j.issn.1002-6819.2021.12.001

Li, X., and Zhang, M. M. (2015). *SPSS 22.0 Statistical Analysis from Introduction to Mastery* (Beijing, China: Electronic Industry Press).

Prastiyo, W., and Fiebig, W. (2021). Multibody simulation and statistical comparison of the linear and progressive rate double wishbone suspension dynamical behavior. *Simul. Model. Pract. Th.* 108, 102273. doi: 10.1016/j.simpat.2021.102273

Pu, M. H., and Wu, J. (2009). Study on flexible sugarcane modeling based on ADAMS software. *J. System Simulation* 21, 1930–1932. doi: 10.16182/j.cnki.joss.2009.07.046

Quynh Anh, P. T., Thuyet, D. Q., and Kobayashi, Y. (2022). Image classification of root-trimmed garlic using multi-label and multi-class classification with deep convolutional neural network. *Postharvest Biol. Tec.* 190, 111956. doi: 10.1016/j.postharvbio.2022.111956

Shi, Y. Y., Chen, M., Wang, X. C., Zeng, Y. N., and Odhiambo, M. O. (2017). Dynamc smuaon and experiments on *artemisia selengensis* orderly harvester cutter. *Trans. Chin. Soc Agric. Mach.* 48, 110–116. doi: 10.6041/j.issn.1000-1298.2017.02.015

Thuyet, D. Q., Kobayashia, Y., and Matsuo, M. (2020). A robot system equipped with deep convolutional neural network for autonomous grading and sorting of root-trimmed garlics. *Comput. Electron. Agric.* 178, 105727. doi: 10.1016/j.compag.2020.105727

Wang, S. Y., Hu, Z. C., Peng, B. L., Wu, H. C., Gu, F. W., and Wang, H. O. (2022). Simulation and parameter optimisation of pickup device for full-feed peanut combine harvester. *Comput. Electron. Agric.* 192, 106602. doi: 10.1016/j.compag.2021.106602

Wang, S. Y., Hu, Z. C., Yao, L. J., Peng, B. L., Wang, B., and Wang, Y. W. (2013). Simulation of auto-follow row detection mechanism in beet harvester based on ADAMS. *Trans. Chin. Soc Agric. Mach.* 44, 62–67. doi: 10.6041/j.issn.1000-1298.2013.12.011

Wang, H. X., Li, T. H., Wu, Y. Q., and Geng, A. J. (2018). Research status and prospects of garlic harvesting machinery. *J. Chin. Agric. Mech.* 39, 102–107. doi: 10.13733/j.jcam.issn.2095-5553.2018.06.021

Wang, R. Y., Zheng, Z. A., Lu, X. F., Gao, L., Jiang, D. L., and Zhang, Z. M. (2021). Design, simulation and test of roller comb type Chrysanthemum (*Dendranthema morifolium Ramat*) picking machine. *Comput. Electron. Agric.* 187, 106295. doi: 10.1016/j.compag.2021.106295

Xie, L. X. (2019). *Design and parameter optimization of stalk chopping, leaf stripping and top breaking devices for sugarcane harvester* (Hangzhou, China: Zhejiang University).

Xie, L. X., Wang, J., Cheng, S. M., Zeng, B. S., and Yang, Z. Z. (2020). Optimisation and dynamic simulation of a conveying and top breaking system for whole-stalk sugarcane harvesters. *Biosyst. Eng.* 197, 156–169. doi: 10.1016/j.biosystemseng.2020.06.017

Xu, X. B., Nie, C. W., Jin, X. L., Li, Z. H., Zhu, H. C., Xu, H. G., et al. (2021). A comprehensive yield evaluation indicator based on an improved fuzzy comprehensive evaluation method and hyperspectral data. *Field Crop Res.* 270, 108204. doi: 10.1016/j.fcr.2021.108204

Yang, K., Hu, Z. C., Peng, B. L., Yu, Z. Y., Wang, S. Y., You, Z. Y., et al. (2015). Research and test of garlic mechanical cutting roots technology. *J. Chin. Agric. Mech.* 36, 153–159. doi: 10.13733/j.jcam.issn.2095-5553.2015.03.038

Yang, K., Hu, Z. C., Yu, Z. Y., Peng, B. L., Zhang, Y. H., and Gu, F. W. (2022). Design and experiment of garlic harvesting and root cutting device based on deep learning target determination. *Trans. Chin. Soc Agric. Mach.* 53, 123–132. doi: 10.6041/j.issn.1000-1298.2022.01.013

Yu, Z. Y., Hu, Z. C., Yang, K., Peng, B. L., Zhang, Y. H., and Yang, M. J. (2021b). Operation mechanism analysis and parameter optimization of garlic root floating cutting device. *Trans. Chin. Soc Agric. Mach.* 52, 111–119. doi: 10.6041/j.issn.1000-1298.2021.05.012

Yu, Z. Y., Hu, Z. C., Yang, M. J., Yang, K., Peng, B. L., and Zhang, Y. H. (2021c). Kinematical characteristics analysis and simulation of garlic root profiling floating cutting mechanism. *J. Chin. Agric. Mech.* 38, 17–23. doi: 10.13733/j.jcam.issn.2095-5553.2021.12.01

Yu, Y. J., Wang, J., Lai, Q. H., Jia, G. X., Yu, F., and Cao, Y. (2021a). Design and experiment of hand-held vibrating comb-type *coffea arabica* L. harvester. *Trans. Chin. Soc Agric. Eng.* 38, 24–32. doi: 10.6041/j.issn.1000-1298.2021.09.014

Yu, Z. Y., Yang, K., Hu, Z. C., Peng, B. L., Gu, F. W., Yang, L., et al. (2023). Parameter optimization and simulation analysis of floating root cutting mechanism for garlic harvester. *Comput. Electron. Agric.* 204, 107521. doi: 10.1016/j.compag.2022.107521

Zhao, D., Cai, D. M., Qin, L. X., Gao, X., Huang, W. T., and Liu, C. (2020). Design and experiment of modularized garlic combine harvester. *Trans. Chin. Soc Agric. Mach.* 51 (4), 95–102. doi: 10.6041/j.issn.1000-1298.2020.04.011

Zhao, L. Q., Chen, Z. W., Liu, X. Z., and Guo, D. (2011). Based on image processing and GUI garlic cut the umbilical. *J. Qingdao Agric. Univ. (Natural Science).* 28, 305–308. doi: 10.3969/j.issn.1674-148X.2011.04.013

frontiers | Frontiers in Plant Science

TYPE Methods
PUBLISHED 05 September 2023
DOI 10.3389/fpls.2023.1153226

OPEN ACCESS

EDITED BY
Yunchao Tang,
Guangxi University, China

REVIEWED BY
Nisha Pillai,
Mississippi State University, United States
Syahrul Fithry Senin,
Universiti Teknologi Teknologi MARA, Cawangan Pulau Pinang, Malaysia
Fengyun Wu,
South China Agricultural University, China

*CORRESPONDENCE
Yonghui Zhao
hero9968@nefu.edu.cn

RECEIVED 29 January 2023
ACCEPTED 15 August 2023
PUBLISHED 05 September 2023

CITATION
Li C, Chen Z, Jing W, Wu X and Zhao Y (2023) A lightweight method for maize seed defects identification based on Convolutional Block Attention Module. *Front. Plant Sci.* 14:1153226. doi: 10.3389/fpls.2023.1153226

A lightweight method for maize seed defects identification based on Convolutional Block Attention Module

Chao Li[1], Zhenyu Chen[1], Weipeng Jing[1], Xiaoqiang Wu[2] and Yonghui Zhao[1]*

[1]College of Computer and Control Engineering, Northeast Forestry University, Harbin, China, [2]School of Mechanical Engineering, Inner Mongolia University for Nationalities, Tongliao, Inner Mongolia Autonomous Region, China

Maize is widely cultivated and planted all over the world, which is one of the main food resources. Accurately identifying the defect of maize seeds is of great significance in both food safety and agricultural production. In recent years, methods based on deep learning have performed well in image processing, but their potential in the identification of maize seed defects has not been fully realized. Therefore, in this paper, a lightweight and effective network for maize seed defect identification is proposed. In the proposed network, the Convolutional Block Attention Module (CBAM) was integrated into the pretrained MobileNetv3 network for extracting important features in the channel and spatial domain. In this way, the network can be focused on useful feature information, and making it easier to converge. To verify the effectiveness of the proposed network, a total of 12784 images was collected, and 7 defect types were defined. Compared with other popular pretrained models, the proposed network converges with the least number of iterations and achieves the true positive rate is 93.14% and the false positive rate is 1.14%.

KEYWORDS

MobileNetv3-Large, image classification, transfer learning, CBAM, lightweight network

Introduction

Maize is an important feed crop in animal husbandry and aquaculture industry, as well as one of the essential raw materials in food industry. Rapid and accurate identification of maize seed varieties not only has important application value in maize planting, but also can effectively avoid the occurrence of adverse phenomena such as mixed seed and mixed processing. In maize automatic sorting system, the industrial camera needs to collect images of the seeds, so that we can screen and compare whether the appearance of the seeds is complete and normal, and the final classification accuracy of the sorting system is highly dependent on the performance of maize image classification algorithm.

Traditional methods (Hinton et al., 2006; Kiratiratanapruk and Sinthupinyo, 2011; Huang et al., 2019) for the classification of maize seed images generally pre-process the images first, then extract the shape, color and other feature information of each maize image by artificial means, screen and integrate the extracted features. Using professional industrial machines to extract maize seeds can not only avoid the influence of other light sources, but also achieve automatic acquisition to improve the collection efficiency. Previous study has proposed several classification algorithm by deep learning. (Krizhevsky et al., 2017) proposed the convolutional neural network AlexNet, which won the first place in the ImageNet competition, and reduced the Top-5 classification error rate of 1000 types of images to 10%. (Xu et al., 2021) established a variety classification model by using multi-layer perceptron, the overall classification accuracy reached more than 95%. (Alotaibi and Alotaibi, 2020) used the residual network ResNet to classify hyperspectral images. The accuracy was 95.33% on the Salinas dataset and 90.57% on the Indian Pines dataset. (Cui et al., 2019) uses the improved VGG16 to classify the four types of images, and replaces the SoftMax classifier in the VGG16 network with the 4-label SoftMax classifier. The final test accuracy of the model is 95%. EfficientNet (Soleimanipour et al., 2022) was used to classify the four pistachio varieties, based on Python programming the Kears API and TensorFlow machine learning framework. The results show that, the average accuracy rate and recall rate of this model are 96.73% and 96.70%, respectively. Compared with traditional machine learning, using deep learning to solve the problem of image classification can not only avoid the influence of human factors, but also greatly improve the recognition accuracy and have stronger observability.

The application of deep learning in agricultural production has a significant impact as it can improve crop productivity and quality. Taking bananas and apples as examples, (Wu et al., 2021; Fan et al., 2022b; Wu et al., 2022) proposed deep learning-based methods for defect detection in bananas and apples, respectively. These methods can quickly and accurately detect defects in fruits and achieve real-time detection, providing farmers with more efficient means of fruit screening and quality control.

For the overall recognition and localization of banana clusters, (Wu et al., 2023) proposed a deep learning based method with an accuracy rate of 93.2%. Using the YOLOv5-B model for banana localization and detection, the average processing time per image is only 9 ms. These research results can help farmers automatically identify and count the number of banana clusters during the growth and harvesting process, improving production management efficiency and accuracy.

In apple production, (Fan et al., 2022b) proposed a real-time defect detection method based on the YOLOv4 network. The average detection accuracy in online testing reached 93.9%, and it can evaluate 4 fruits per second. This method can assist farmers in quickly and accurately screening out defective apples from a large quantity, improving apple quality and market competitiveness.

In addition, (Wang et al., 2022) proposed a maize defect detection method based on the watershed algorithm, providing a new solution for maize production. This method can effectively identify good seeds and bad seeds, with an average accuracy rate of 95.63% and an average recall rate of 95.29%. It can help farmers quickly select high-quality seeds during the planting process, improving planting success rate and yield.

In conclusion, the application of deep learning methods in agricultural production plays a crucial role. Through automated image recognition and defect detection techniques, it can enhance the quality and productivity of agricultural products, provide farmers with more accurate and efficient production management tools, and make a significant contribution to promoting agricultural development.

In this paper, our main contributions are as follows: Firstly, we propose an improved MobileNetv3 network, which can reduce model parameters and reduce real-time computation. The requirements for equipment resources are low and can be applied to mobile devices. Secondly, we introduce the Convolutional Block Attention Module to the MobileNetv3 network to improve the fitting ability and generalization ability of the network. Finally, we use our own maize seed dataset to test the new method, which is derived from the maize image in GrainSpace (Fan et al., 2022a), and has achieved more than 90% recognition accuracy. Experimental verification has demonstrated that the proposed algorithm in this paper is capable of classifying maize based on its appearance, and further detecting the types of defects present, thus achieving intelligent detection of maize defects and effectively reducing labor costs. This algorithm holds great potential for wide-ranging applications and promising future development.

Related work

Transfer learning

The transfer learning method can greatly save the time of training model and significantly reduce the hardware resources required by deep learning. For this reason, more and more scholars adopt transfer learning method to train their own data sets. (Xie et al., 2022) used Inceptionv3 and DenseNet201 as feature extractors at the same time, and used the dual transfer learning framework to reach the accuracy rate, recall rate and F1 values of 95.11%, 95.33%, and 95.15% on 10 types of seabird data sets. (Fraiwan et al., 2022) applied the transfer learning technology to the convolution neural network model to classify three maize diseases. The average accuracy rate reached 98.6%, which proved that transfer learning can greatly improve the accuracy of classification. (Xiang et al., 2019) applied lightweight network MobilleNetv2 to the task of fruit classification, and used Softmax as the classifier for feature classification. Finally, the accuracy rate was obtained 85.12%. (Das et al., 2023) applied the transfer learning to the DenseNet169 network to classify 11 kinds of garbage that frequently occur in outdoor environment, and finally achieved the accuracy of 93.10%. (Huang et al., 2017) put forward the method of transfer learning to solve the obstacles in SAR target classification. (Tammina, 2019) proposed to use the VGG16 model of transfer learning to classify cat and dog pictures, and achieved an accuracy rate of 95.40%, which is 16.2% higher than that without transfer learning. (Kaya et al., 2019) used DNN

network of transfer learning to verify the influence of deep neural network on plant classification in four different transfer learning models. The experimental results show that transfer learning can optimize automatic plant identification and improve low-performance plant classification models.

MobileNet

MobileNet network model is a lightweight network model proposed by Google. Compared with the traditional convolutional neural network, MobileNet has the advantages of fewer parameters and lower delay. At present, MobileNet series networks include MobileNetv1, MobileNetv2 (Huu et al., 2022; Młodzianowski, 2022) and MobileNetv3 (Howard et al., 2019; Hussain et al., 2021; Zhao and Wang, 2022; Zhao et al., 2022). MobileNetv3 is Google's new invention after MobileNetv2, and its main improvement is to add SE-net after the deep separable convolution in MobileNetv2, which automatically obtains the importance of each feature channel by learning, and suppresses some feature information that is not useful for the current task. In addition, MobileNetv3 integrates the four characteristics of MobileNetv2 and MobileNetv1, which first uses the convolution of 1×1 for dimensionality upgrading, and introduces the inverse residual structure of MobileNetv2 linear bottleneck. Then, 3x3 depth separable convolution is performed to reduce the computational amount of the network. Then, through the lightweight SE-net attention model, the net-work pays attention to more useful channel information to ad-just the weight of each channel. Finally, the h-swish activation function is used instead of the swish function to reduce the amount of operation and improve performance. After changing the Bneck structure, compared with MobileNetv2, MobileNetv3 achieves the same accuracy on COCO, the speed is 25% faster, and the segmentation algorithm is also improved.

Convolutional block attention module

The application of Attention mechanisms (Zhang et al., 2022) enables the neural network to focus on important features and suppress unnecessary features. This paper uses the Convolutional Block Attention Module (Woo et al., 2018). (Liang et al., 2022) used CNN to mine the deep features of gold price data and improved the feature extraction ability of the network through CBAM. The experiment proved that this method could improve the prediction accuracy and was superior to other models. (Ma et al., 2019) proposed CBAM-GAN generative adversarial network, which can significantly improve the quality of generated images. (Li et al., 2022) adopted MobileNetv3 to detect floating objects on the water surface and CBAM to enhance feature fusion. The experiment showed that the detection accuracy of the improved model increased by 2.9% and the detection speed increased by 55%. (Pei et al., 2022) used the YOLOv4 model with CBAM to detect weeds in the field, which reduced the total number of tags in 1000 images by half, and the average accuracy reached 86.89%, thus improving the efficiency of weed detection. (Chen et al., 2023) used the MobileNetv3-large network combined with CBAM for crack detection. The experiment showed that the model had better performance and the crack identification accuracy reached 99.69%. CBAM can improve the training efficiency and prediction accuracy of CNN, which is a combination of channel attention mechanism and spatial attention mechanism, which can be embedded in the module of CNN and conduct end-to-end training together with CNN, with only a small amount of computation added.

Methods

The method of MobileNetv3-Large

We used MobileNetv3-Large as the image classification model (Qian et al., 2021). MobileNetv3 introduces the depthwise convolution. The convolution kernel of the depthwise separable convolution is equal to the number of input channels, that is, a convolution kernel performs a convolution operation on a feature map of the previous layer separately to obtain the number of output channels equal to the number of input channels. Compared with conventional convolution, the number of parameters of the depthwise convolution can be saved by 1/3, and the number of layers of the neural network can be deeper under the premise of the same number of parameters, so that the model can be lightweight.

MobileNetv3 uses an inverted residual structure. It can effectively avoid the problem of gradient disappearance or gradient explosion. At the same time, the lightweight attention structure SE block is used. The idea of SE block is to start from the weight of spatial information, and use the feature map to obtain the weight of each layer after batch normalization optimization through a series of convolution operations. Firstly, the feature map with input of $H \times W \times C$ is averaged and pooled into a vector of $1 \times 1 \times C$, and then the weight of $1 \times 1 \times C$ is obtained by two convolutions of 1×1. Finally, the feature map of each input channel is multiplied by the vector to obtain the final output.

The activation function used by MobileNetv3 is h-swish, which is obtained by h-sigmoid. It has the characteristics of easier calculation and faster learning, as shown in Equation 1:

$$h - swish = x\frac{ReLU6(x+3)}{6} \quad (1)$$

In the Mobilenetv3 model, the bottleneck layer is the middle layer between the image feature input and the last fully connected layer, which could retain the feature input. Since the model parameters of MobileNetv3 model have good generalization ability after extensive training of ImageNet data set, the training parameters of this network in ImageNet were transferred.

The input image outputs a feature vector of $1 \times 1 \times 1280$ through the bneck layer and then enters the full connection layer through the dropout layer to reduce the dimension of the feature vector to $1 \times 1 \times c$, where c represents the number of categories in the data set. When c=7, there are seven types of maize. Finally, the output of the full connection layer is processed by Softmax regression to obtain the probability distribution of each class. The

calculation formula is:

$$P_j = \frac{n_j}{\sum_{k=1}^{c} n_k} \tag{2}$$

In Equation 2, p_j represents the classification probability of category j, n_j represents the output of the full connection layer, and c represents the total number of categories in the data set.

We combined the feature extraction of the MobileNetv3 model with the ImageNet data set in the source domain with the fully connected modules, so as to adapt to the classification task and perform the training on the double-side maize image data set to obtain the maize seed network model.

Improvement of network structure

Based on the above analysis, we know that the MobileNetv3 network model has the characteristics of few parameters and low delay. To solve the problem of low accuracy in current maize classification due to similar morphology, transfer learning can be introduced on the basis of MobileNetv3 network. Based on the characteristics of MobileNetv3 network model, the requirements for device performance can be greatly reduced. In the process of transfer learning, the large-scale shared parameters in the neural network of the pre-training were first migrated, and the initial training weight of the source domain model was transferred to the new network for initialization, so as to obtain the prior knowledge on the large data set. Then, the training was conducted according to the self-built dataset, and the parameters in the model were adjusted through subsequent learning, so as to obtain the classification model. The maize seed classification algorithm combined with transfer learning can not only improve the training efficiency and accuracy of the network model, but also improve the algorithm performance based on the pre-training basis of large data sets. The final model in training will have better generalization ability and robustness.

In addition, the improved MobileNetv3 network introduced CBAM attention mechanism. The structure of CBAM is shown in Figure 1. The application of attention mechanisms enabled the neural network to focus on important features and suppress unnecessary features. CBAM module successively induced a one-dimensional channel attention mechanism feature graph M_c and a two-dimensional space attention mechanism feature graph M_s, and the calculation process is shown as follows:

$$P' = M_c(P) \otimes P, P'' = M_s(P') \otimes P' \tag{3}$$

In Equation 3, P is the input image, $\otimes$ is elements multiplication. In the process of multiplying, the attention mechanism replicates: channel attention mechanisms propagate along spatial dimensions and vice versa. P' is the weighted result of the channel attention mechanism, and P' is the final output obtained. CBAM includes the channel attention mechanism CAM and the spatial attention mechanism SAM. Input the feature map P and apply max-pooling and avg-pooling to each spatial position, then it can obtain two $C \times 1 \times 1$ vectors, which are respectively sent to shared MLP. Finally, the two were combined, and the channel attention M_c was obtained after activation function. The input feature map P and apply max-pooling and avg-pooling to each integration to obtain two $1 \times H \times W$ feature vectors, which were then sent to the standard convolution layer according to the channel cost, and M_s was obtained after function activation. Suppose the size of the input image is $C \times H \times W$, and C, H and W correspond to the channel, height and weight respectively. Firstly, the spatial information in the feature map was extracted by means of average pooling and maximum pooling, and two different operation results were generated: P_{avg} and P_{max}, which represent the feature results generated by average pooling and maximum pooling of input feature map P, respectively. Then, the two description results were extracted into the shared network to generate channel attention feature graph M_c. After the shared network was applied to the description of each feature, the output feature vector was obtained by adding elements by elements. The calculation method is as follows:

$$\begin{aligned} M_c &= \sigma(p_1(p_2(P)) + p_1(p_3(P))) \\ &= \sigma(W_1(W_o(P_{avg})) + W_1(W_o(P_{max}))) \end{aligned} \tag{4}$$

In Equation 4, P is the input image, σ is the nonlinear activation function Sigmoid, p_1 is the forward calculation function, p_2 and p_3 are the average-pooling function and max-pooing function respectively. W_0 and W_1 are the weights of two linear layers. In order to calculate spatial attention, average pooling and maximum pooling were carried out along the channel first, and they were connected together to generate effective feature descriptions, which were represented by $M_s(P)$. The channel information of the feature map was extracted through two kinds of operations to generate two two-dimensional feature maps P_{avg} and P_{max}, which represent the average pooling feature and the maximum pooling feature respectively. Then the two-dimensional spatial attention feature map was generated by concatenation operation and convolution operation, whose expression is as follows:

$$M_s(P) = \sigma(p(p_c(P_{avg}, P_{max}))) \tag{5}$$

In Equation 5, σ represents sigmoid activation function, $p_c$is splicing operation, and p is 7×7 convolution operation. The improved MobileNetv3 is shown in Figure 2. The structure introduces CBAM attention mechanism to replace the original SE module. The size of each image sample will be resized to the size of 224 × 224 before input.

The Loss Function used in the experiment is the cross entropy, and its expression is shown as follows:

$$L = \frac{1}{N}\sum_i L_i = -\frac{1}{N}\sum_i \sum_{c=1}^{M} y_{ic} \log p_{ic} \tag{6}$$

In Equation 6, M is the number of classes, and y_{ic} is the sign function (0 or 1), if the true class of sample i is equal to c takes 1, otherwise takes 0. p_{ic} is the probability that the observed sample i belongs to class c.

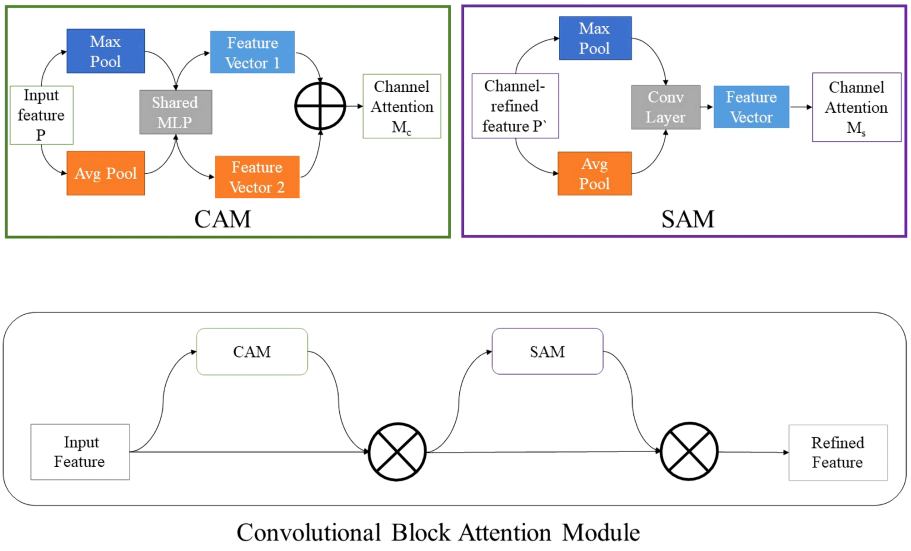


FIGURE 1
The structure of CBAM. CAM and SAM are included, representing the channel attention mechanism and the spatial attention mechanism respectively.

Experiments

Datasets

In this experiment, we used the GrainSpace (Fan et al., 2022a) as the dataset for the test model. GrainSpace is a large-scale data set for fine-grained and domain adaptation, which is mainly used for grain recognition. GrainSpace contains three types of grains, namely maize, rice and wheat. We selected the maize image as the data set of this experiment, which is divided into seven categories, and divided into double-sided images and single-sided images.

The maize image is shown in Table 1. According to the classification standard formulated by ISO55270 Cereals, maize seeds can be divided into normal and six types of damaged and unsound seeds: fusarium (FM) seeds, sprouted (SD) seeds, mouldy (MY) seeds, broken (BN) seeds, attacked by pests (AP) seeds and heated (HD) seeds. Among these maize seeds, FM, MY, BP maize seeds represent the proportion of maize seeds polluted by Fusarium or fungi; SD, AP and HD maize seeds correspond to the nutritional

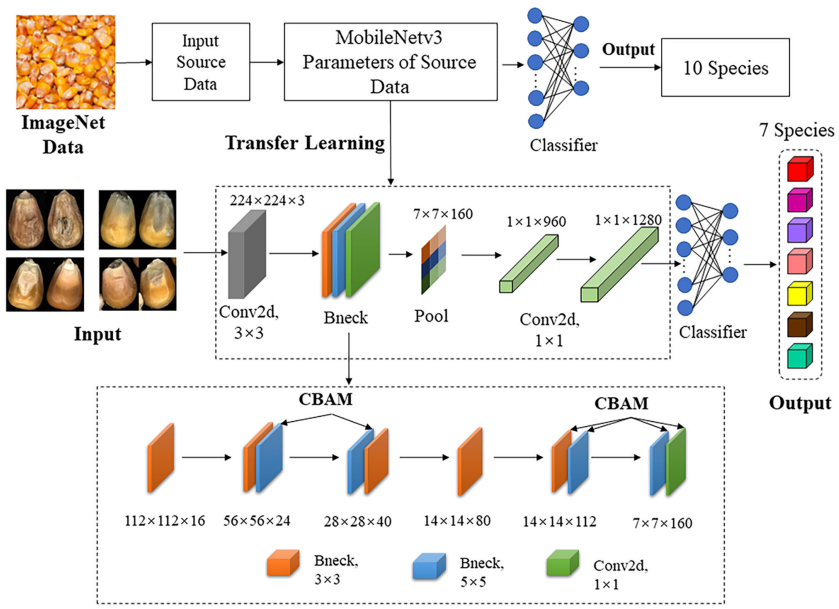
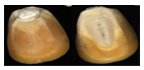
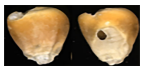
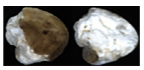
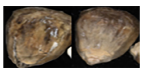
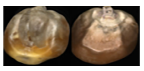


FIGURE 2
The overview of MobileNetv3-large using CBAM attention mechanism and transfer learning.

TABLE 1 Part of maize image display.

Class No.	Name	Description	Image(double-side)	Image(single-side)	Train	Validation
1	NOR	Normal			1671	186
2	AP	Attacked by Pests			1634	181
3	BN	Broken			1647	183
4	FM	Fusarium			1651	183
5	HD	Heated			1638	182
6	MY	Mouldy			1624	180
7	SD	Sprouted			1639	182

components of maize. All images are in PNG format. Due to the interference of various factors during the shooting of various maize images, maize grain images with complete shape and standard features are preferred as the data set of the experiment. All the acquired maize images were obtained through cropping. Due to the varying sizes of the maize, the dimensions of each image are not fixed, but the maximum value will not exceed 300 × 400. A total of 12784 training data were obtained. The ratio of train set to validation set was 9:1.

Implementation details

The server platform configuration of the experiment is as follows: CPU is Intel Core I5-10400, 6 core and 12 threads, GPU is NVIDIA GeForce RTX2080ti. The operating system is Windows10 and the Python version is 3.7. The torch version in the deep learning framework is 1.10.0, and the torchvision version is 0.11.0; CUDA version is 11.0; CUDnn version is 8.0. In the experiment, the pre-training weights of MobileNetv3-large models trained on ImageNet were used to initialize the model, and the modified model was used to train the maize seed data set. Adam is used by the model designer, and the Loss Function is cross entropy. Considering the memory and model generalization ability, the learning rate is set to 0.0001. Adam optimization algorithm and h-swish activation function were adopted. Set batchsize to 32 and epochs to 40. Using the Adam optimizer can automatically modify the learning rate with the training process, which can speed up the training speed and improve the performance of the model. We use the h-swish function because it can alleviate the disappearance of the gradient, prevent over-fitting, and accelerate the convergence rate of the gradient descent. Moreover, the computational cost is low, and the computer runs faster.

Result analysis and visualization

At present, there are many convolutional networks applied in deep learning. In order to reflect the effectiveness of the models, popular models such as MobileNetv2, ResNet50 and VGG were selected to compare with the MobileNetv3 network, and the above models were respectively used for training in the self-built dataset of this study. During the training process, the test set and training loss value were recorded for each training cycle completed by the model. The variation of the accuracy rate of each network with the number of epochs is shown in Figure 3, and the variation of the loss function with the number of epochs is shown in Figure 4.

The accuracy of the improved MobileNetv3 algorithm is more than 10% higher than that of other computational majority methods. The

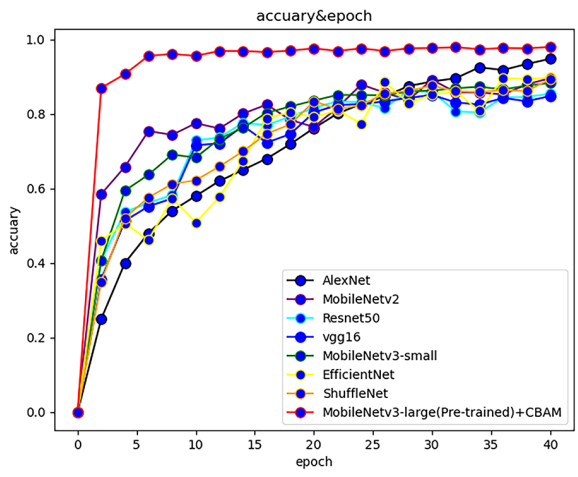


FIGURE 3
The accuracy of various algorithms varies with the number of epochs.

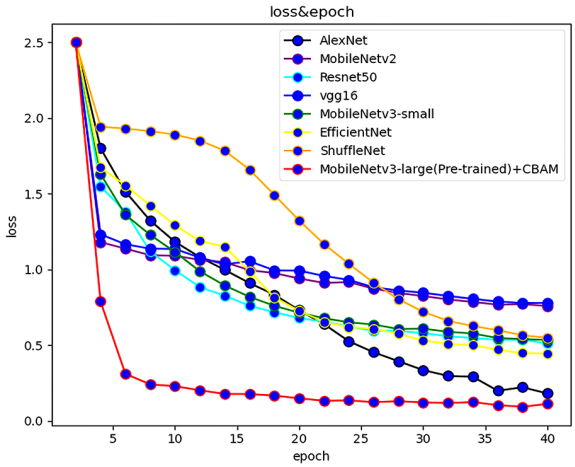


FIGURE 4
The loss of various algorithms varies with the number of epochs.

training speed is the fastest, reaching 8.23bit/s, and the test time is only 35.25ms. The reasons are as follows: (1) Depthwise separable convolution was used to reduce the amount of computation and parameters of the model, and the reversed residual network structure constructed was helpful for feature extraction; (2) The CBAM was introduced to integrate the attention mechanism in channel and space, so that the important features are enhanced, while the unimportant features were suppressed; (3) The weight of large data sets in the model can be used to pool network parameters in similar tasks, and it improved the performance of the model. Table 2 shows the summary of experimental results of each model on the maize seed data set. Based on the aforementioned results, in addition to our method, AlexNet, EfficientNet, and MobileNetv3 - large also demonstrated excellent training outcomes with validation accuracies exceeding 90%. However, ResNet101 and VGG16 exhibited slower training speeds due to their larger model parameters and greater computational requirements, resulting in longer processing times.

A group of maize seed images were selected from other data sets and introduced into the model for classification. The average accuracy of maize seed classification is more than 97%, and the accuracy of some categories of maize seeds classification is more than 99% (such as BN, FM), but the classification accuracy of HD is low. Through analysis, there are two reasons for the decline of HD maize recognition accuracy: Firstly, some maize seeds have complex features, which are highly similar to other maize seeds; Secondly, when the images in the data set were screened, the number of maize seed samples with certain features was small. When CBAM attention mechanism was introduced, the recognition accuracy of image classification by MobileNetv3 algorithm has been improved slightly, and the number of references and the time required for image recognition were reduced.

We selected 50 images for each type of maize seed in other datasets, then observed the prediction results of each type of maize by this method, and drew the confusion matrix. The confusion matrix has five evaluation indexes: *P*, *R*, *F*1, *TPR*, *FPR*. Their calculation is as follows:

$$P = \frac{TP}{TP + FP} \tag{7}$$

$$R = \frac{TP}{TP + FN} \tag{8}$$

$$F1 = \frac{2 \times P \times R}{P + R} \tag{9}$$

$$TPR = \frac{TP}{TP + FN} \tag{10}$$

$$FPR = \frac{FP}{FP + TN} \tag{11}$$

Where *P* represents precision, *R* represents recall, and *F*1 represents F1-score. *TP* (True Positives) represents the number of true samples correctly classified as positive. *TN* (True Negative)

TABLE 2 Experimental results of various algorithms.

Model	Prediction (%)							Accuracy (%)	Training speed(bit/s)	Test time (s)
	NOR	AP	BN	FM	HD	MY	SD			
ResNet50	92.80	82.53	97.65	98.14	64.71	73.02	92.24	81.77	2.71	126.15
ResNet101	90.64	81.65	96.51	96.84	63.37	72.53	91.55	78.73	1.43	129.45
VGG16	97.14	91.15	98.33	94.71	66.57	83.52	95.63	86.44	1.33	195.41
AlexNet	82.81	88.14	95.62	96.65	82.20	92.31	93.43	94.80	2.63	128.33
ShuffleNet	80.16	75.67	75.94	96.63	68.14	76.37	79.43	77.24	3.67	50.58
EfficientNet	89.73	96.54	96.83	99.26	70.38	85.59	95.64	89.41	2.64	62.67
EfficientNetv2	90.65	97.65	98.32	99.12	72.23	86.54	96.75	90.90	2.20	46.89
MobileNetv2	96.80	94.96	87.57	97.32	81.36	86.81	97.82	85.26	6.71	43.52
MobileNetv3-small	96.81	94.34	98.96	97.76	84.32	87.93	98.74	87.40	7.82	40.63
MobileNetv3-large	97.24	98.00	97.54	97.74	90.21	96.84	98.85	96.31	8.20	37.64
MobileNetv3-large+CBAM	97.40	98.51	99.03	99.06	91.67	98.49	99.10	97.95	8.23	35.25

10.3389/fpls.2023.1153226

represents the number of true samples incorrectly classified as negative. *FP* (False Positive) represents the number of false samples incorrectly classified as positive. *FN* (False Negative) represents the number of false samples correctly classified as negative. *TPR* (True Positice Rate) represents the proportion of the positive samples of the predicted pair to the total positive samples. The *FPR* (False Positive Rate) represents the proportion of positive samples with wrong prediction to all positive samples.

The confusion matrix after completion is shown in Figure 5. Calculated from the confusion matrix, the accuracy true positive rate of the maize seed image classification model reached 93.14% and the false positive rate reached 1.14%. In addition, the precision, recall rate and F1 values of all kinds of maize seed images were caused by Table 3. Compared with previous methods for maize seed detection, the method proposed in this paper not only accurately detects maize affected by defects, but also identifies the types of defects, with a shorter detection time. In conclusion, the method in this paper has high feasibility in the task of maize classification. However, there are certain limitations in our experiment. Firstly, we only conducted detection on 7 types of maize, and when faced with multiple types of defects, we cannot determine the effectiveness of this detection method. Secondly, when detecting maize with highly similar defect types, our method still has issues with category detection errors. Lastly, during the experimental process, we only selected 50 images per category of maize and created a confusion matrix based on the classification results. We then calculated metrics such as accuracy, recall, F1 score for each category of maize, but these test results may contain some degree of error.

Ablation experiment

In the ablation experiment, we use the control variable method to prove the necessity of a module by removing or adding the function of a module. If the performance results obtained after the ablation experiment change significantly, it shows that the module plays a decisive role.

The ablation experiment was divided into four phases. The first step was tested and validated using only the MobileNetv3 algorithm without the use of transfer learning and attention mechanisms. The second step was to use the attention mechanism, but not transfer learning. The third step was to use transfer learning without the attention mechanism. The fourth step was to use both the attention mechanism and transfer learning. In order to verify the influence of attention mechanism, the attention mechanism module was frozen so that it cannot be used. Similarly, this method was also applicable to transfer learning.

After experimental verification, the accuracy and loss values of each method change with the number of iterations, as shown in Figures 6, 7.

After the use of transfer learning, the recognition accuracy of the model was greatly increased, which increases by 62% when epoch is 1. With the increase of epoch, the gap between the two gradually decreases. When epoch is 20, the gap between the two Narrows to less than 5%; when epoch is 40, the accuracy of the two is almost the same. After the introduction of CBAM attention mechanism, the recognition accuracy

TABLE 3 the precision, recall, F1 value of various maize classifications in double-sided images.

Class No.	Class Name	P (%)	R (%)	F1(%)
1	NOR	94.23	98.00	96.08
2	AP	96.00	96.00	96.00
3	BN	100.00	94.00	96.91
4	FM	94.23	98.00	96.08
5	HD	88.88	80.00	84.21
6	MY	85.18	92.00	88.46
7	SD	94.11	96.00	95.05

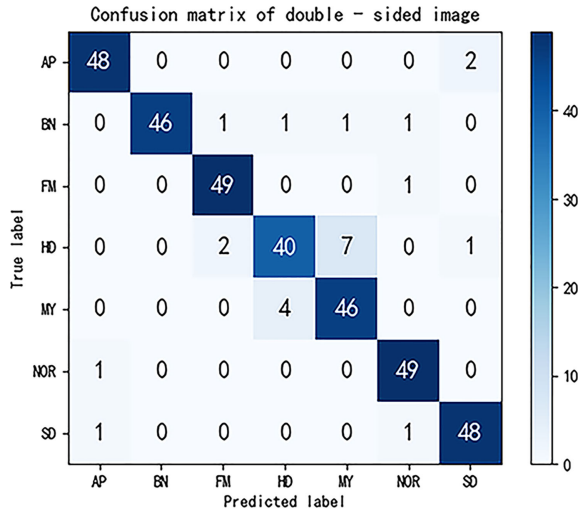


FIGURE 5
Confusion matrix diagram for double-sided maize seeds forecasts.

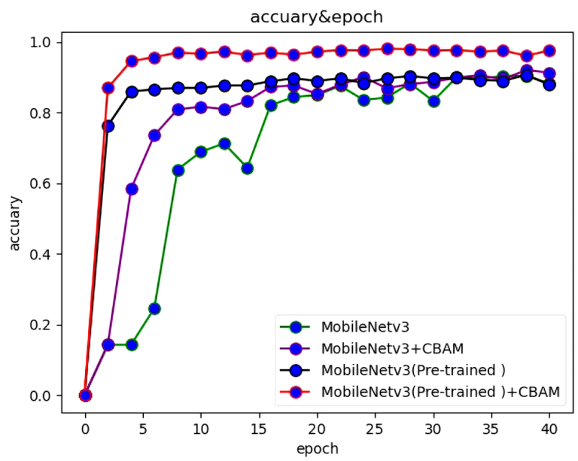


FIGURE 6
The accuracy of various ablation methods varies with the number of epochs.

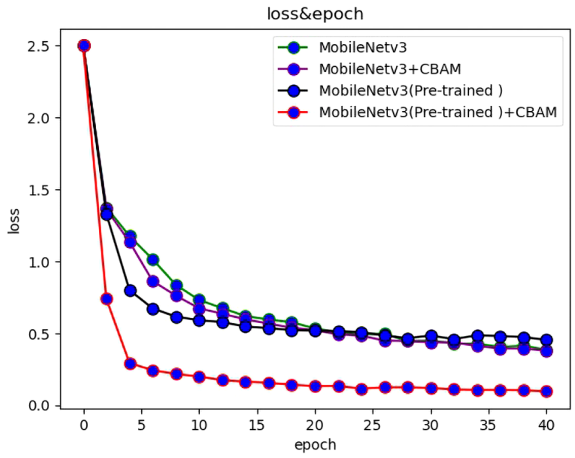


FIGURE 7
The loss of various ablation methods varies with the number of epochs.

of the model can be improved by 7%. In conclusion, for the MobileNetv3 network for double-sided maize seed classification and recognition, the primary factor affecting the accuracy is transfer learning, and the secondary factor is attention mechanism.

At the same time, in order to verify the improvement of the classification accuracy of double-sided maize seed images, a dataset of single-sided maize seed images was used as the control experiment. According to the confusion matrix shown in Figure 8, the true positive rate was 81.14% when using single-sided maize seed images, which was 12% lower than that using double-sided maize seed images. The false positive rate was 3.14%, which was 2% higher than that of double-sided maize seed images. The accuracy, recall and F1 values of various types of maize classification calculated according to the confusion matrix are shown in Table 4. From this table, it can be analyzed that the overall classification accuracy under one-sided maize seeds is about

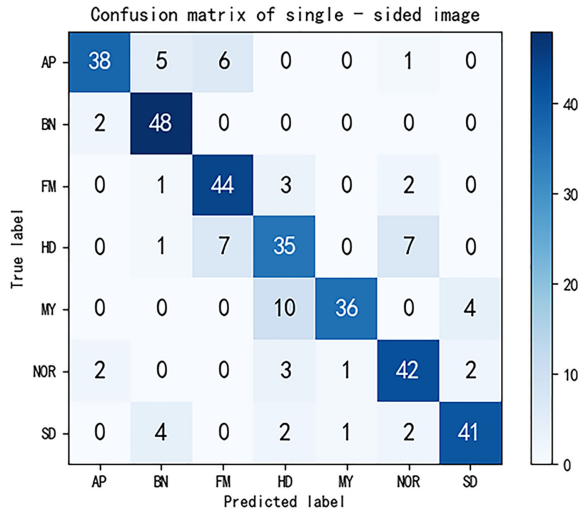


FIGURE 8
Confusion matrix diagram for single-sided maize seeds forecasts.

TABLE 4 The precision, recall, F1 value of various maize classifications in single-sided images.

Class No.	Class Name	P (%)	R (%)	F1 (%)
1	NOR	77.78	84.00	80.77
2	AP	90.48	76.00	82.61
3	BN	81.36	96.00	88.07
4	FM	77.19	88.00	82.24
5	HD	66.04	70.00	67.96
6	MY	94.74	72.00	81.82
7	SD	87.23	82.00	84.54

15% lower than the overall classification accuracy under double-sided maize seeds.

Through the comparative analysis of the above ablation experiments, it is found that transfer learning has the greatest influence on the accuracy, which can increase the prediction rate by 62% during the first training, and it is the most important factor affecting the accuracy of image recognition. Secondly, the accuracy of double-sided image can be improved by 12% than that of single-sided image. The effect of attention mechanism on image recognition accuracy is 8% higher than without attention mechanism. In summary, the influencing factors were ranked as transfer learning greater than double-sided maize seed images and greater than attention mechanisms.

Conclusion

The main contribution of this paper is to introduce deep learning into the field of maize seed image classification and recognition, and propose an improved MobileNetv3 algorithm model, achieving a comprehensive classification true positive rate of 93.14%. The average prediction, recall, and F1 score of maize classification reached 93.23%, 93.43%, 93.26% and the average test time is only 35.25 ms. In future research, we will attempt to conduct defect detection studies on other types of grains such as rice and wheat, and explore different methods to achieve superior detection results. Our approach enables automatic classification and detection of maize, thereby identifying higher quality maize and contributing to the efficient allocation of human and material resources, while promoting the advancement of intelligent agricultural production.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

CL was responsible for the conception and design of research, the design of experimental methods, and the review and revision of

papers. ZC was responsible for software development and design, experimental design verification and verification, experimental data analysis, and writing the first draft of the paper. WJ was responsible for supervision and guidance of research projects, review and revision of papers. YZ was responsible for the conception and design of research, the acquisition of research funds, the collection of research resources, the supervision and guidance of research topics, and the review and revision of papers. XW was responsible for data collation and management, research project management, paper review and revision. All authors contributed to the article and approved the submitted version.

Funding

This research was funded by Science and Technology Program of Inner Mongolia Autonomous Region, grant number 2022YFSJO037.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Alotaibi, B., and Alotaibi, M. (2020). A hybrid deep resnet and inception model for hyperspectral image classification. *PFG–Journal Photogrammetry Remote Sens. Geoinformation Sci.* 88, 463–476. doi: 10.1007/s41064-020-00124-x

Chen, L., Yao, H., Fu, J., and Ng, C. T. (2023). The classification and localization of crack using lightweight convolutional neural network with cbam. *Eng. Structures* 275, 115291. doi: 10.1016/j.engstruct.2022.115291

Cui, H., Liu, X., Han, L., and Wei, Z. (2019). "Road crack classification based on improved vgg convolutional neural network," in *Fuzzy systems and data mining V* (Lausanne, Switzerland: IOS Press), 542–547.

Das, D., Sen, A., Hossain, S. M., and Deb, K. (2023). "Trash image classification using transfer learning based deep neural network," in *International conference on intelligent computing & Optimization* (Cham: Springer), 561–571.

Fan, L., Ding, Y., Fan, D., Di, D., Pagnucco, M., and Song, Y. (2022a). "Grainspace: A large-scale dataset for fine-grained and domain-adaptive recognition of cereal grains," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. (EEE Computer Society, New Orleans Louisiana, United States), 21116–21125.

Fan, S., Liang, X., Huang, W., Zhang, V. J., Pang, Q., He, X., et al. (2022b). Real-time defects detection for apple sorting using nir cameras with pruning-based yolov4 network. *Comput. Electron. Agric.* 193, 106715. doi: 10.1016/j.compag.2022.106715

Fraiwan, M., Faouri, E., and Khasawneh, N. (2022). Classification of corn diseases from leaf images using deep transfer learning. *Plants* 11, 2668. doi: 10.3390/plants11202668

Hinton, G. E., Osindero, S., and Teh, Y.-W. (2006). A fast learning algorithm for deep belief nets. *Neural Comput.* 18, 1527–1554. doi: 10.1162/neco.2006.18.7.1527

Howard, A., Sandler, M., Chu, G., Chen, L.-C., Chen, B., Tan, M., et al. (2019). "Searching for mobilenetv3," in *Proceedings of the IEEE/CVF international conference on computer vision*. (IEEE Computer Society, Seoul, Korea), 1314–1324.

Huang, S., Fan, X., Sun, L., Shen, Y., and Suo, X. (2019). Research on classification method of maize seed defect based on machine vision. *J. Sensors* 1–9. doi: 10.1155/2019/2716975

Huang, Z., Pan, Z., and Lei, B. (2017). Transfer learning with deep convolutional neural network for sar target classification with limited labeled data. *Remote Sens.* 9, 907. doi: 10.3390/rs9090907

Hussain, A., Barua, B., Osman, A., Abozariba, R., and Asyhari, A. T. (2021). "Performance of mobilenetv3 transfer learning on handheld device-based real-time tree species identification," in *2021 26th International Conference on Automation and Computing (ICAC)*. (IEEE Robotics and Automation Society, Portsmouth, United Kingdom), 1–6.

Huu, P. N., Quang, V. T., Le Bao, C. N., and Minh, Q. T. (2022). Proposed detection face model by mobilenetv2 using asian data set. *J. Electrical Comput. Eng.* 1–19. doi: 10.1155/2022/9984275

Kaya, A., Keceli, A. S., Catal, C., Yalic, H. Y., Temucin, H., and Tekinerdogan, B. (2019). Analysis of transfer learning for deep neural network based plant classification models. *Comput. Electron. Agric.* 158, 20–29. doi: 10.1016/j.compag.2019.01.041

Kiratiratanapruk, K., and Sinthupinyo, W. (2011). "Color and texture for corn seed classification by machine vision," in *2011 international symposium on intelligent signal processing and communications systems (ISPACS)*. (IEEE, Chiang Mai, Thaila), 1–5.

Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2017). Imagenet classification with deep convolutional neural networks. *Commun. ACM* 60, 84–90. doi: 10.1145/3065386

Li, H., Yang, S., Liu, J., Yang, Y., Kadoch, M., and Liu, T. (2022). A framework and method for surface floating object detection based on 6g networks. *Electronics* 11, 2939. doi: 10.3390/electronics11182939

Liang, Y., Lin, Y., and Lu, Q. (2022). Forecasting gold price using a novel hybrid model with iceemdan and lstm-cnn-cbam. *Expert Syst. Appl.* 206, 117847. doi: 10.1016/j.eswa.2022.117847

Ma, B., Wang, X., Zhang, H., Li, F., and Dan, J. (2019). "Cbam-gan: generative adversarial networks based on convolutional block attention module," in *International conference on artificial intelligence and security* (Cham: Springer), 227–236.

Młodzianowski, P. (2022). "Weather classification with transfer learning-inceptionv3, mobilenetv2 and resnet50," in *Conference on multimedia, interaction, design and innovation* (Cham: Springer), 3–11.

Pei, H., Sun, Y., Huang, H., Zhang, W., Sheng, J., and Zhang, Z. (2022). Weed detection in maize fields by uav images based on crop row preprocessing and improved yolov4. *Agriculture* 12, 975. doi: 10.3390/agriculture12070975

Qian, S., Ning, C., and Hu, Y. (2021). "Mobilenetv3 for image classification," in *2021 IEEE 2nd International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering (ICBAIE)*. (IEEE, Nanchang, China), 490–497. doi: 10.1109/ICBAIE52039.2021.9389905

Soleimanipour, A., Azadbakht, M., and Rezaei Asl, A. (2022). Cultivar identification of pistachio nuts in bulk mode through efficientnet deep learning model. *J. Food Measurement Characterization* 16 (4), 1–11. doi: 10.1007/s11694-022-01367-5

Tammina, S. (2019). Transfer learning using vgg-16 with deep convolutional neural network for classifying images. *Int. J. Sci. Res. Publications (IJSRP)* 9, 143–150. doi: 10.29322/IJSRP.9.10.2019.p9420

Wang, L., Liu, J., Zhang, J., Wang, J., and Fan, X. (2022). Corn seed defect detection based on watershed algorithm and two-pathway convolutional neural networks. *Front. Plant Sci.* 13, 730190. doi: 10.3389/fpls.2022.730190

Woo, S., Park, J., Lee, J.-Y., and Kweon, I. S. (2018). "Cbam: Convolutional block attention module," in *Proceedings of the European conference on computer vision (ECCV)*. (Springer Link, Munich, Germany), 3–19.

Wu, F., Duan, J., Ai, P., Chen, Z., Yang, Z., and Zou, X. (2022). Rachis detection and three-dimensional localization of cut off point for vision-based banana robot. *Comput. Electron. Agric.* 198, 107079. doi: 10.1016/j.compag.2022.107079

Wu, F., Duan, J., Chen, S., Ye, Y., Ai, P., and Yang, Z. (2021). Multi-target recognition of bananas and automatic positioning for the inflorescence axis cutting point. *Front. Plant Sci.* 12, 705021. doi: 10.3389/fpls.2021.705021

Wu, F., Yang, Z., Mo, X., Wu, Z., Tang, W., Duan, J., et al. (2023). Detection and counting of banana bunches by integrating deep learning and classic image-processing algorithms. *Comput. Electron. Agric.* 209, 107827. doi: 10.1016/j.compag.2023.107827

Xiang, Q., Wang, X., Li, R., Zhang, G., Lai, J., and Hu, Q. (2019). "Fruit image classification based on mobilenetv2 with transfer learning technique," in *Proceedings of the 3rd International Conference on Computer Science and Application Engineering*. (Association for Computing Machinery, New York, NY, United States), 1–7.

Xie, J., Zhu, M., and Hu, K. (2022). Improved seabird image classification based on dual transfer learning framework and spatial pyramid pooling. *Ecol. Inf.* 72, 101832. doi: 10.1016/j.ecoinf.2022.101832

Xu, P., Yang, R., Zeng, T., Zhang, J., Zhang, Y., and Tan, Q. (2021). Varietal classification of maize seeds using computer vision and machine learning techniques. *J. Food Process Eng.* 44, e13846. doi: 10.1111/jfpe.13846

Zhang, Y., Zhou, A., Zhao, F., and Wu, H. (2022). A lightweight vehicle-pedestrian detection algorithm based on attention mechanism in traffic scenarios. *Sensors* 22, 8480. doi: 10.3390/s22218480

Zhao, Y., Huang, H., Li, Z., Yiwang, H., and Lu, M. (2022). Intelligent garbage classification system based on improve mobilenetv3-large. *Connection Sci.* 34, 1299–1321. doi: 10.1080/09540091.2022.2067127

Zhao, L., and Wang, L. (2022). A new lightweight network based on mobilenetv3. *KSII Trans. Internet Inf. Syst. (TIIS)* 16, 1–15. doi: 10.3837/tiis.2022.01.001

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 06 November 2023
DOI 10.3389/fpls.2023.1169091

OPEN ACCESS

EDITED BY
Lei Shu,
Nanjing Agricultural University, China

REVIEWED BY
Bimlesh Kumar,
Indian Institute of Technology Guwahati, India
Huiling Chen,
Wenzhou University, China

*CORRESPONDENCE
Hongping Zhou
hpzhou@njfu.edu.cn
Zhong Xue
xxyyxz006@163.com

RECEIVED 18 February 2023
ACCEPTED 16 October 2023
PUBLISHED 06 November 2023

CITATION
Zhao L, Zhou H, Xu L, Yuan W, Shi M, Zhang J and Xue Z (2023) Parameter optimization of the spiral fertiliser discharger for mango orchards based on the discrete element method and genetic algorithm. *Front. Plant Sci.* 14:1169091. doi: 10.3389/fpls.2023.1169091

Parameter optimization of the spiral fertiliser discharger for mango orchards based on the discrete element method and genetic algorithm

Liang Zhao[1], Hongping Zhou[1]*, Linyun Xu[1], Weidong Yuan[1], Minghong Shi[1], Jian Zhang[1] and Zhong Xue[2]*

[1]College of Mechatronics Engineering, Nanjing Forestry University, Nanjing, China, [2]South Subtropical Crops Research Institute, Chinese Academy of Tropical Agricultural Sciences, Zhanjiang, China

Introduction: In order to solve the problems of inaccurate fertilization, unstable fertilization and low fertiliserutilization rate in mango orchard.

Methods: A small spiral fertiliser discharger was designed based on the agronomic characteristics of fertilization in mango orchard. The fertilizing performance test and parameter optimization of thespiral fertiliser discharger were carried out by combining bench and simulation test. Firstly, the main influencing factors of the fertilizing performance of the spiral fertiliser discharger were analyzed by theoretical calculation formula, and the range of its value was preliminarily determined. At the same time, the digital and discrete element models of the spiral fertiliser discharger were established. Then,the discrete element model of granular fertiliser was established on the basis of the physical and related mechanical simulation parameters of granular fertiliser obtained by experimental statistics.Taking the variable coefficient of fertilizing stability as the response value, the method of singlefactor simulation fertilizing test was used to explore the parameters that have a significant influence on the variable coefficient of fertilizing stability. The response surface method (RSM) was used tosimulate the fertilizing performance of three significant parameters. Based on the quadraticregression orthogonal rotation combination design test, a second-order regression mathematicalmodel between the variable coefficient of fertilizing stability and the significant parameters wasestablished. The variable coefficient of fertilizing stability was as small as possible. The geneticalgorithm (GA) was used to optimize the regression model. Finally, the verification test of thefluidity and applicability of different fertilisers was carried out.

Results: The results of single factor test showed that the diameter of spiral blade, pitch and rotationalspeed of fertilizing shaft have significant influence on the variable coefficient of fertilizing stability.The optimal parameter combination of the spiral fertiliser discharger was obtained: 98.44 mm for thediameter of spiral blade, 54.8 mm for the pitch, and 24.43 r/min for the rotational speed of

fertilizingshaft. The verification results showed that the average relative error of the test was small, and themass flow rate of different fertilisers and the variable coefficient of fertilizing stability could meetthe agronomic requirements of fertilization in mango orchards. The reliability of the discrete elementsimulation test results and research methods of the spiral fertiliser discharger was verified.

Conclusion: The results and methods of this study can provide reference for the development of mangoorchard fertilization machinery and related fertilizing performance test.

KEYWORDS

mango orchard, spiral fertiliser discharger, parameter optimisation, genetic algorithm, discrete element method, response surface methodology

Highlights

1. DEM-GA is used to optimize the parameters of spiral fertilizer discharger.
2. The optimal combination of parameters is obtained by optimizing regression model.
3. Fertilizing stability of spiral fertiliser discharger for mango orchard is improved.
4. The fertilizer utilization rate and environmental protection effect were improved.
5. Providing references for optimised design of mango orchard fertilization machinery.
6. Providing references for related study on fertilizing performance in other orchards.

1 Introduction

Mango is one of the world's five major tropical fruits (banana, mango, pineapple, longan, and sugar apple), which has a rich nutrition and unique taste, and it has been favored by people, known as the 'king of tropical fruits' (Wei and Aiping, 2021). Mango is native to the Himalayas, India, Bangladesh, Indochina Peninsula, and Malay Peninsula, and in China, it is mainly distributed in Yunnan, Guangxi, Guangdong, Sichuan, Fujian, Guizhou, Taiwan, and Hainan (He et al., 2018). According to statistics, by 2020, the national mango planting area has reached 5.241 million mu. The total yield in China ranks third in the global mango-producing countries (Wei, 2021), and the planting area and yield maintain a steady growth trend.

Fertilization is a crucial operation in mango orchard management, which plays a key role in mango quality and yield (Liu et al., 2021). At present, there are some problems with fertilization in mango orchards in China, such as high intensity, low efficiency, and poor quality of artificial fertilization. Unreasonable fertilization can easily lead to soil compaction, unbalanced nutrition of fruit trees, low fertilizer utilization, and environmental pollution (Huang et al., 2018; Qu et al., 2021). Mechanical fertilization can effectively reduce the labor intensity of fruit farmers, improve fertilization efficiency and quality, and reduce the cost of agricultural production. The fertiliser discharger is a key component of fertilization machinery, and its working performance directly affects fertilization quality (Zhao et al., 2022). Currently, the fertiliser dischargers at home and abroad can be mainly divided into the outer groove wheel type, the spiral type, the centrifugal type, and the rotary disc type. These fertiliser dischargers have their advantages and disadvantages and different applicable scenarios, but there is a certain pulsation degree in the fertiliser discharge operation (Zhu et al., 2018; Liu et al., 2020; Dun et al., 2022), resulting in poor stability and uniformity of fertilization. Therefore, it is necessary to perform optimization and experimental study of fertiliser dischargers to improve fertilization performance.

The traditional test method has been adopted to optimize the fertilizing performance of the fertiliser discharger with low efficiency, high cost, and long cycle, and the simulation test method can effectively solve the above problems. The discrete Element Method (DEM) is a numerical simulation method based on the discontinuity assumption proposed by Professor CUNDALL (Wei and Gao, 2021) in 1971. It can be used to simulate and analyse the interaction between agricultural granular materials and mechanical equipment. It provides a new approach for the digital design of modern agricultural equipment, greatly improves the research and development efficiency of agricultural equipment, and has a good application prospect in agricultural engineering (Liao et al., 2023; Tan et al., 2022; Liu et al., 2022).

In recent years, domestic and foreign scholars have mainly studied the performance of large-scale fertilization for fruits and crops such as apples, bananas, corn, and soybeans based on the discrete element method (DEM). For example, Yuan et al. (2021) investigated the effect of the blade configuration of the auger on the mixing performance of orchard soil-fertiliser particles by the discrete element method and experimental measurement method, and the optimal parameters of blade rotation speed, lateral angle, pitch angle, and blade number were obtained. Liu et al. (Liu et al., 2020).

designed a fertilization shunt plate according to the fertilization characteristics of alfalfa, optimised its structural parameters by discrete element simulation test, determined the optimal knob width, tilt angle, and horizontal distance, and verified the accuracy of the simulation results through field experiments. Sun et al. (2020) designed and developed a groove wheel fertilization device by using 3D printing technology. Then, the EDEM simulation test was used to analyse the influence of various factors on the fertilization performance, and the optimal parameters of groove wheel radius, helix angle, rotation speed, and inclination angle were determined. Thaper R. (Thaper, 2014) applied the discrete element method to analyse the relationship between different fertiliser types and leaf shapes and the uniformity of the double disc spreader. By establishing a discrete element simulation model of an orchard centrifugal fertiliser applicator, COETZEE (Coetzee and Lombard, 2011) investigated the influence of structural parameters such as orifice flow and blade inclination angle on the consistency of the orchard centrifugal fertiliser spreader. The experimental results indicated that the discrete element simulation model has a good prediction effect. By using the discrete element method, van Liedekerke et al. (Van Liedekerke et al., 2009). studied the effect of different disc rotating speeds on the performance of a centrifugal seeder. Zhu et al. (2022) designed a self-propelled orchard organic fertiliser strip paver according to the planting status of dwarf and dense apple orchards in China and the agronomic requirements of organic fertiliser strip furrow fertilization. The discrete element method was employed to optimize the structure of the fertiliser discharge port of the strip paver. The optimal structure of the fertiliser discharge port was determined to be an oblique mouth shape, and its fertilization performance was verified by a strip paving test and field test. Song et al. (2020) designed a rotary variable fertiliser discharger according to the distribution pattern of banana roots, and they optimized the parameters of the fertiliser discharger through discrete element simulation analysis. The optimal parameters of the forward speed, the rotation period, the central angle, and the opening size of the curved groove were obtained, and the performance of the fertiliser discharger was verified by prototype and field experiments. Jia et al. (2022) designed a pneumatic centralized precision mixed fertilization device according to the requirement of topdressing in the middle and late stages of maize growth. Based on fluid dynamics and the discrete element coupling method, the optimal inclination angle of the fertiliser outlet, the conveying gas velocity, and the bellows length were obtained, and the fertilization performance was verified by field experiment. Dun et al (Dun et al., 2018). developed a ratio control and position stratified fertilization device that is suitable for the soybean planting pattern in Northeast China, and the ditching fertilization performance of the device was investigated by combining discrete element simulation and field experiment. Liu et al. (Liu et al., 2021). designed a layered quantitative fertilization device according to the rape compartment noodle planting mode and the rape root growth law. The discrete element simulation test was conducted to determine the optimal parameters of the diameter of the retaining rod, the number of the retaining rod groups, and the spacing of the retaining rod groups. Then, the operating performance was verified by bench and field experiments. For the research of optimization methods, Thawkar et al. (Thawkar et al., 2021). proposed a hybrid feature selection method based on butterfly optimization algorithm (BOA) and ant lion optimization algorithm (ALO). The results show that BOAALO is superior to the original BOA and ALO in terms of accuracy, sensitivity, specificity, kappa value, type I and type II errors, and the receiver operating characteristic curve, which can be better applied to the diagnosis of breast cancer. Sayed et al (Sayed et al., 2021). proposed a new hybrid version of convolutional neural network architecture and bald eagle search (BES) optimization. The BES algorithm is used to find the optimal value of the hyperparameters of the SqueezeNet architecture. The overall accuracy of the proposed melanoma skin cancer prediction model is 98.37%, indicating that the method can be better applied to skin lesion classification. Xing et al. (Xing et al., 2023). proposed an improved whale optimization algorithm (QGBWOA). A comparative experiment with dimensions of 10, 30, 50, and 100 was conducted on CEC 2014, and a comparative experiment with dimensions of 30 was conducted on CEC 2020. The experimental results show that the convergence accuracy and convergence speed are significantly improved. Pir et al. (Piri and Mohapatra, 2021). used a multi-objective quadratic binary HHO (MOQBHHO) technique with K-Nearest Neighbor (KNN) method as the wrapper classifier to extract the optimal feature subset, and carried out relevant comparative experiments. The results show that the MOQBHHO proposed in this paper effectively finds a set of non-dominated feature subsets, which has advantages in obtaining the best trade-off between the two fitness evaluation criteria. Santana et al. (2018) used genetic algorithm to optimize the parameters of vacuum cooling treatment conditions for broccoli after harvest. The results showed that GA had good performance for the optimization of broccoli cooling process, and the best conditions for vacuum cooling process of broccoli were obtained. The above research on optimization methods has certain guiding significance for this study.

Though many studies have been conducted on the fertilizing performance of fruits and crops at home and abroad, there are few reports on the fertilizing performance of fertilization machinery in mango orchards. The application of discrete element simulation analysis to apple, banana, rape, and other crop fertilization machinery provides a basis for studying fertilization machinery design and fertilizing performance in mango orchards. In this study, a small spiral fertiliser apparatus was designed according to the agronomic characteristics of fertilization in mango orchards. The fertilizing performance test and parameter optimization of the spiral fertiliser discharger were carried out by adopting the bench test and simulation test. Firstly, the main influencing factors of the fertilizing performance of the fertiliser discharger were analysed by theoretical calculation, and the ranges of values were preliminarily determined. Meanwhile, the digital and discrete element models of the spiral fertiliser discharger were established. Then, by taking the variable coefficient of fertilizing stability as the evaluation index, the single factor and quadratic regression orthogonal rotation combination design test method were exploited to simulate the fertilizing performance, and the parameters significantly affecting the variable coefficient of fertilizing stability were explored.

Subsequently, the second-order regression mathematical model of the variable coefficient of fertilizing stability and the significant parameters was established, and the genetic algorithm was adopted to optimize the regression model to obtain the optimal parameter combination. Finally, the accuracy of discrete element simulation test results of the spiral fertiliser discharger was verified by different fertiliser flowability and applicability tests, which provide a reference for the development of mango orchard fertilization machinery and related experimental research. So as to reduce the labor intensity of fruit farmers, improve fertilization efficiency and fertilization quality, as well as mango quality and yield.

2 Design and parameterisation of spiral fertiliser discharger

2.1 Structure design and working principle

In this study, a small spiral fertiliser discharger was designed according to the agronomic characteristics of mango orchard fertilization. It is mainly composed of a fertiliser box, a fertilizing box, a spiral blade, a fertilizing shaft, and a fertilizing pipe. The three-dimensional model is shown in Figure 1. Specifically, the three-dimensional size of the fertiliser box is 280 mm × 260 mm × 270 mm, the inclined plate angle is set to 45°, the length of the fertiliser box is 260 mm, the diameter of the fertilizing box is 105 mm, the diameter of the fertilizing shaft is 25 mm, and the diameter of the fertilizing pipe is 65 mm. Additionally, the parameters such as the diameter of the spiral blade, the pitch, and the rotational speed of the fertilizing shaft will be calculated according to the agronomic requirements of the mango orchard and the relevant theoretical formulas. The main working process of the spiral fertiliser discharger is introduced as follows. Firstly, an appropriate amount of fertiliser is poured into the fertiliser box, and then the speed-regulating motor is started to drive the fertilizing shaft and the spiral blade to rotate. Under the drive of the two, the fertiliser is orderly transported to the right and falls into the fertiliser pipe and soil, thus completing the fertilization operation.

2.2 Determination of the main parameters of the fertiliser discharger

2.2.1 The fertilizing amount

The fertilizing amount is a key index to evaluate the performance of the spiral fertiliser discharger, and it is related to the structure and motion parameters of the fertiliser discharger. In the process of fertiliser application, the cross-sectional area of the fertilizing shaft affects the performance of the fertiliser discharger. However, compared with the whole fertiliser discharger, the shaft diameter is small, and its axial blocking effect is often ignored. Therefore, the fertilizing amount discharged by the spiral fertiliser discharger can be approximately calculated as:

$$Q = 47D^2 \cdot n \cdot S \cdot \lambda \cdot \epsilon \cdot \phi \tag{1}$$

where, Q is the fertilizing amount (unit: t/h), D is the diameter of the spiral blade (unit: mm), n is the rotational speed of the fertilizing shaft (unit: r/min), S is the pitch (unit: mm), λ is the fertiliser density (unit: t/m^3), ϵ is the coefficient of inclined conveying, and ϕ is the filling factor.

From the formula (1), it can be seen that the fertilizing amount Q f the spiral fertiliser discharger is related to D, n, S, λ, ϵ, ϕ, etc. When the fertilizing amount Q is determined, parameters such as the diameter of the spiral blade D, the rotational speed of the fertilizing shaft n, and the pitch S can be adjusted appropriately to meet the demand for the fertilizing amount per mango tree. When the fertiliser spreader travels continuously forward one plant distance, the fertilizing amount of the spiral fertiliser discharger can be calculated as:

$$Q = \frac{v \cdot f}{R} \tag{2}$$

where, v is the forward speed of the fertiliser applicator (unit: km/h); f is the fertilizing amount per mango tree (unit: kg/plant), and R is the average spacing length of mango trees (unit: m).

In this study, the forward speed v of the fertiliser discharger was set to 2 km/h. The national standard 'Mango Cultivation Technical Regulations' indicates that the fertilizing amount f during the

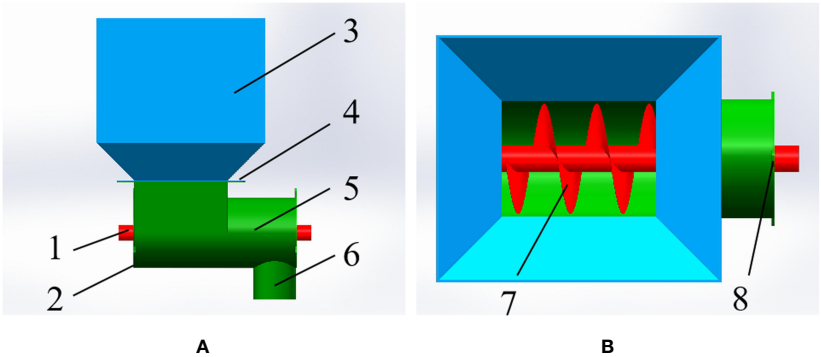


FIGURE 1
The three-dimensional diagram of spiral fertiliser discharger for the mango orchard: **(A)** the main view, **(B)** the top view, fertilizing shaft 2. left cover plate 3. fertiliser box 4. connection plate 5. fertilizing box 6. fertilizing pipe 7. spiral blade 8. right cover plate.

FIGURE 2
Mango orchard planting situation.

topdressing period of mango trees is generally 0.3-0.5 kg/plant. Meanwhile, the average plant spacing and row spacing of mango trees were measured through on-the-spot investigation. As shown in Figure 2, the average plant spacing R was finally measured to be 1.82 m, and the average row spacing was 3.16 m. Then, the above data are substituted into Formula (2) to calculate the fertilizing amount required for mango orchards.

$$Q = \frac{v \cdot f}{R} = 0.329 \sim 0.549 t/h \tag{3}$$

2.2.2 The diameter of the spiral blade and pitch

The diameter of the spiral blade is an important structural parameter of the spiral fertiliser discharger, and it directly affects the fertilizing performance. The diameter of the spiral blade is usually determined according to the structure of the fertiliser discharge device, the physical and mechanical properties of the fertiliser, and the fertilizing amount required of fruit trees. In this study, the diameter of the spiral blade of the fertiliser discharger is calculated by Formula (4).

$$Q = 47K_1 \cdot J \cdot \lambda \cdot \epsilon \cdot \phi \cdot D^{\frac{5}{2}} \tag{4}$$

Let

$$K = \left(\frac{1}{47K_1 J}\right)^{\frac{2}{5}} \tag{5}$$

Therefore,

$$D = K\left(\frac{Q}{\lambda \cdot \epsilon \cdot \phi}\right)^{\frac{2}{5}} \tag{6}$$

where, K_1 is the ratio coefficient between the diameter of the spiral blade and pitch, and is usually 0.5-0.9 for a horizontally arranged fertiliser discharger; J is the fertiliser composite characteristic coefficient, and it is set to 28 with reference to Supplementary Table 1; λ is the fertiliser density, and it is measured to be 0.913 t/m^3; ϵ is the inclined conveying coefficient, and it usually takes the value of 1 for the spiral fertiliser discharger designed in this study, which is installed horizontally; ϕ is the filling factor, and it is set to 0.2 with reference to Supplementary Table 1; K is the fertiliser synthesis factor, and it is set to 0.0632 with reference to Supplementary Table 1.

Then, the above data are substituted into Formula (4) to calculate as 80.08-98.2 mm. Since the diameter D of the spiral blade is usually a standard integer, so the diameter of the spiral blade is initially set to 80-100 mm.

Pitch is a key structural parameter of the spiral fertiliser discharger. It determines the size of the spiral lift angle and changes the movement speed and direction of the granular fertiliser, so it is closely related to the fertilizing performance of the spiral fertiliser discharger. The pitch is usually calculated according to the following empirical formula:

$$S = K_1 \cdot D \tag{7}$$

2.2.3 The rotational speed of the fertilizing shaft

The rotational speed of the fertilizing shaft significantly affects the fertilizing performance of the spiral fertiliser discharger. Usually, the greater the speed of the fertilizing shaft, the larger the fertilizing amount of the fertiliser discharger. However, the speed of the fertilizing shaft should not be too large because when the speed of the fertilizing shaft exceeds a critical value, excessive centrifugal force will be generated to throw the fertiliser outwards, resulting in the inability to discharge fertiliser normally, so the speed of the fertilizing shaft needs to be limited not greater than the critical value. The speed of the fertilizing shaft is usually determined according to the physical and mechanical properties of the fertiliser, the diameter of the spiral blade, and other parameters. When there is no radial movement of the granular fertiliser, the relationship between the maximum centrifugal force on the outside of the spiral and its gravity is represented below.

$$m\omega_{max}^2 r \le mg \tag{8}$$

That is,

$$2\pi n_{max}/60 \le \sqrt{gr} \tag{9}$$

Considering the influence of fertiliser discharger by different physical properties of fertilisers, we have:

$$\pi n_{max}/30 \le K\sqrt{gr} \tag{10}$$

$$n_{max} = \frac{30K}{\pi}\sqrt{\frac{g}{r}} = \frac{30K}{\pi}\sqrt{\frac{2g}{D}} \tag{11}$$

Let $J = 30K\sqrt{2g}/\pi$. Then, eq. (11) can be translated into the common empirical formula:

$$n_{max} = J/\sqrt{D} \tag{12}$$

where, m is the mass of fertiliser (unit: kg), ω_{max} is the critical angular velocity of the fertilizing shaft (unit: rad/s), r is the radius of the spiral (unit: m), n_{max} is the critical speed of fertilizing shaft (unit: r/min), and g is the gravitational acceleration (unit: m/s^2).

Through calculation, the maximum rotational speed of the fertilizing shaft is 89 r/min. To satisfy the amount of fertiliser required for the mango garden, the rotational speed of the fertilizing shaft should not be too high, and it is not allowed to exceed the critical value, i.e., the following condition should be met.

$$n \leq n_{max} \tag{13}$$

where, n is the actual speed of the fertilizing shaft (unit: r/min).

To sum up, the diameter of the spiral blade, the pitch, and the rotational speed of the fertilizing shaft are the key parameters affecting the fertilizing performance of the spiral fertiliser discharger. According to the agronomic characteristics of fertilization in mango orchards and the related theoretical calculation formulas, the adjustment ranges of the relevant parameters of the small spiral fertiliser discharger in mango orchards are preliminarily determined, i.e., the diameter of the spiral blade is 80-100 mm, the pitch is 50-70 mm, and the rotational speed of fertilizing shaft is 15-55 r/min. Bassed on this, the three-dimensional solid model of all parts of the spiral fertiliser discharger is established.

3 Materials and methods

3.1 Discrete element modelling of the fertiliser discharger and the granular fertiliser

Before using the discrete element simulation software EDEM to simulate the fertilizing test of the spiral fertiliser discharger, it is necessary to understand the physical and mechanical properties of the granular fertiliser. The shape and density of the granular fertiliser directly affect its movement in the fertilizing box and the fertilizing performance. To make the discrete element model of the granular fertiliser closer to the actual fertiliser, by taking the compound fertiliser commonly used in mango orchards as the object, 100 compound fertiliser particles were randomly selected, the length (L), width (W), and thickness (T) of the compound fertiliser particles were measured by an electronic digital vernier calliper with an accuracy of 0.01 mm, and then the measured data were collated to calculate the equivalent diameter (D_1), sphericity (Φ), and standard deviation. After statistical analysis, the relevant parameters of the triaxial size of the compound fertiliser particles are shown in Supplementary Table 2, and the relevant calculation formulas are as follows.

$$D_1 = \sqrt[3]{LWT} \tag{14}$$

$$\Phi = \frac{D_1}{L} \tag{15}$$

where, D_1 is the equivalent diameter (unit: mm), L is the length (unit: mm), W is the width (unit: mm), T is the thickness (unit: mm), and Φ is the sphericity.

From Supplementary Table 2, it can be seen that the average length of the compound fertiliser particles is 4.168 mm, the average width is 3.819 mm, the average thickness is 3.535 mm, the average equivalent diameter is 3.829 mm, the average sphericity is 92.1%, and the shape of its profile is approximately ellipsoidal. Therefore, the EDEM software was exploited to build a discrete element model of the compound fertiliser particles, and the ellipsoid was set to have a long axis of 4.2 mm and a short axis of 3.6 mm. As the EDEM software cannot directly establish the ellipsoid, five spheres were used for filling to establish the discrete element model of the compound fertiliser particles, and the radius of the five spheres was set to 1.7 mm, 1.75 mm, 1.8 mm, 1.75 mm, and 1.7 mm respectively, and the discrete element model of the compound fertiliser particles is shown in Supplementary Figure 1.

Then, the 3D design software SolidWorks was used to build the digital model of the spiral fertiliser discharger and imported into the solving environment of the EDEM software in the igs format. The relevant physical and mechanical parameters of the compound fertiliser particles and carbon structural steel Q235 were obtained by actual measurements with instruments such as the WD-E micro-controlled electronic universal testing machine, the MV-VS078FM high-speed photographic instrument, and the MXD-01 friction coefficient instrument, as listed in Supplementary Table 3. Meanwhile, the material of all components of the spiral fertiliser discharger was set to be carbon structural steel Q235, and the relevant simulation parameters of the compound fertiliser particles were set.

3.2 Contact model selection

The contact model is usually selected according to the characteristics of the study object, and the main contact models include the Hertz-Mindlin, Hertz-Mindlin with JKR, Hertz-Mindlin with bonding, etc. Different contact models have different application scenarios. According to the physical characteristics of granular fertilisers and their low surface adhesion, and assuming that the force, velocity, and displacement changes of granular fertilisers in the fertiliser discharge process are determined by the small elastic deformation between granular fertilisers or between granular fertilisers and geometry, and combined with Newton's second law, each granular fertiliser moves under the action of torque and force. Therefore, this study selected the Hertz-Mindlin no-slip contact mechanics model for the simulation tests of fertilization. The model provides an accurate representation of the physical situation, and the mechanical calculations are accurate and efficient, which can shorten the simulation time and improve the simulation efficiency (Zhao et al., 2022). The working principle of the model is shown in Figure 3.

Based on the Hertz-Mindlin no-slip contact mechanics model, during discharge application, the granular fertiliser is mainly subject to gravity, granular fertiliser normal force F_n, tangential force F_τ, normal damping force F_d^n, and tangential damping force F_d^τ, which is represented as follows:

$$F_n = \frac{4}{3} E^\star \sqrt{R^\star} \delta_n^{\frac{3}{2}} \tag{16}$$

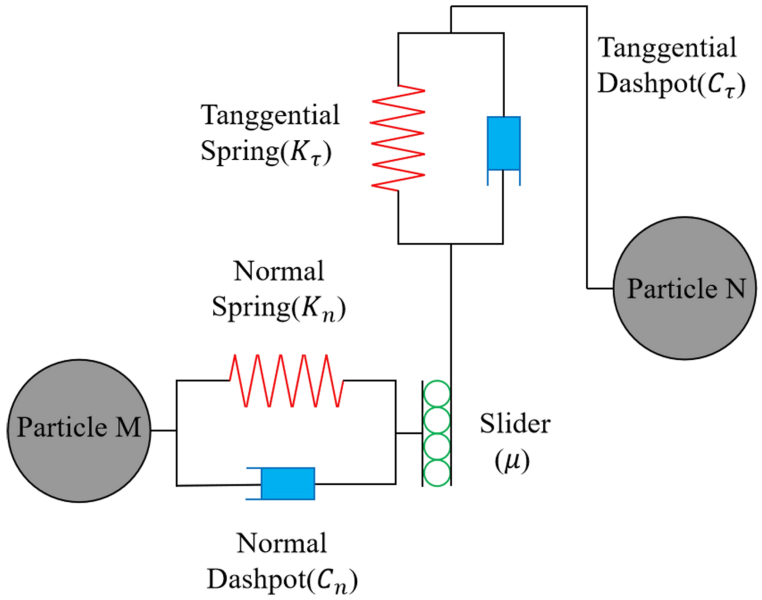


FIGURE 3
The working principle of the Hertz-Mindlin no-slip contact mechanics model.

$$F_\tau = -S_\tau \delta_\tau \tag{17}$$

$$F_d^n = -2\sqrt{\frac{5}{6}}\beta\sqrt{S_n m^*}\,v_n^{\overline{rel}} \tag{18}$$

$$F_d^\tau = -2\sqrt{\frac{5}{6}}\beta\sqrt{S_\tau m^*}\,v_\tau^{\overline{rel}} \tag{19}$$

with

$$\frac{1}{E^*} = \frac{1-v_i^2}{E_i} + \frac{v1-_j^2}{E_j} \tag{20}$$

$$\frac{1}{R^*} = \frac{1}{R_i} + \frac{1}{R_j} \tag{21}$$

$$S_\tau = 8_G^*\sqrt{R^*\delta_n} \tag{22}$$

$$\beta = \frac{lne}{\sqrt{ln^2 e + \pi^2}} \tag{23}$$

$$S_n = 2_E^*\sqrt{R^*\delta_n} \tag{24}$$

where, E^* is the equivalent modulus of elasticity; R^* is the equivalent radius; δ_n is the normal overlap; E_i and E_j are the modulus of elasticity; v_i and v_j are Poisson's ratio; R_i and R_j are the radius of contact particles; S_τ is the tangential stiffness; δ_τ is the tangential overlap; G^* is the equivalent shear modulus, S_n is the normal stiffness; m^* is the equivalent mass; e is the recovery factor; β is the damping ratio; $v_n^{\overline{rel}}$ is the normal component of relative velocity; $v_\tau^{\overline{rel}}$ is the tangential component of relative velocity.

The tangential force F_τ between granular fertilisers and between granular fertilisers and geometry is limited by Coulomb friction $\mu_s F_n$; meanwhile, granular fertilisers are susceptible to rolling friction during discharge application, which can be expressed in terms of moments T_i on the contact surfaces of granular fertilisers as follows.

$$T_i = \mu_r F_n R_i \omega_i \tag{25}$$

where, μ_s is the static friction factor, μ_r is the rolling friction factor, and ω_i is the unit angular velocity vector of the granular fertiliser at the point of contact.

The granular fertiliser moves under torque and force. When the tangential force is greater than the Coulomb friction, the granular fertiliser slides; therefore, the tangential moment and the rolling friction moment together determine the motion of the granular fertiliser.

3.3 Simulation of the fertiliser discharge test method

Before the simulation fertilizing test, the relevant simulation parameters of the granular fertiliser and the spiral fertiliser discharger were first set up, and a rectangular granular plant was established at the position directly above the fertiliser box for granular fertiliser generation, and granular fertiliser was generated dynamically in a fixed form to improve the efficiency of simulating fertiliser discharge, with 80,000 and 40,000 fertiliser granules being generated per second and the total simulation time of 6 s. In the fertiliser discharge process, the fertiliser granules are free to fall in the -Z-axis direction at a speed of 2 m/s. When all the fertiliser granules are generated and fall into the fertiliser box and the fertilizing box, the spiral blade and the fertilizing shaft are set to

rotate at the corresponding speed. Then, under the drive of the spiral blade and the fertilizing shaft, the granular fertiliser is transported to the right in an orderly manner and falls into the fertiliser discharge tube, and a fertiliser discharge monitoring area is set up at the outlet of the fertiliser discharge tube. The mass flow rate of the granular fertiliser can be obtained through the EDEM post-processing module after the end of the simulation of fertiliser discharge, thus completing the whole simulation of fertiliser discharge test of the spiral fertiliser discharger. The process of simulating fertiliser discharge is shown in Figure 4.

(1) Single-factor test. According to the previous theoretical analysis of the spiral fertiliser discharger, the main factors affecting the performance of the spiral fertiliser discharger are the diameter of the spiral blade, the pitch, the rotational speed of the fertilizing shaft, etc. Further single-factor tests were conducted to verify the performance of the spiral fertiliser discharger. In this study, the diameter of and spiral blade A, the pitch B, and the rotational speed of the fertilizing shaft C were used as test factors, and the variable coefficient V of fertilizing stability was used as the response value to perform a single-factor simulation test of fertiliser discharge performance. Each group of trials was repeated three times, the variable coefficient of fertilizing stability was obtained by the following formulas, and the factor levels for the single-factor simulation of fertiliser discharge trials are shown in Table 1.

$$\delta = \frac{\sum_{x=1}^{y} Q_n}{x} \tag{26}$$

$$\sigma = \sqrt{\sum_{x=1}^{y} \frac{(Q_n - \delta)^2}{x}} \tag{27}$$

$$V = \frac{\sigma}{\delta} \times 100\,\% \tag{28}$$

where, δ is the average fertilizing amount for all data in the fertiliser monitoring area (unit: g), Q_n is the fertilizing amount from a particular trial collected in the fertiliser monitoring area (unit: g); x is the number of trials; y is the total number of data points collected in the fertiliser monitoring area; σ is the standard deviation of the fertilizing amount for all data in the fertiliser monitoring area, and V is the variable coefficient of fertilizing stability (unit: %).

(2) Quadratic regression orthogonal rotational combination design test. To reduce the fertiliser discharge pulsation degree of the spiral fertiliser discharger and improve the fertilizing stability, based on the results of the single-factor simulation fertilizing test, the quadratic regression orthogonal rotational combination design test was carried out by using the response surface method (RSM) to further investigate the influence of the diameter of the spiral blade A, the pitch B, and the rotational speed of the fertilizing shaft C on the fertilizing performance of the spiral fertiliser discharger, with the variable coefficient V of fertilizing stability as the evaluation index. The test parameters were taken at five levels, namely, high, upper, middle, lower, and low, expressed in the form of codes +1.682, 1, 0, -1, and -1.682, respectively, and the factor level codes of the multi-factor simulation fertilizing test are shown in Table 2. Further, a second-order regression mathematical model between the variable coefficient of fertilizing stability and the significance parameter was established based on a quadratic regression orthogonal rotational combination design test, which provides a theoretical basis for optimising the parameters of the spiral fertiliser discharger and improving its fertilizing performance.

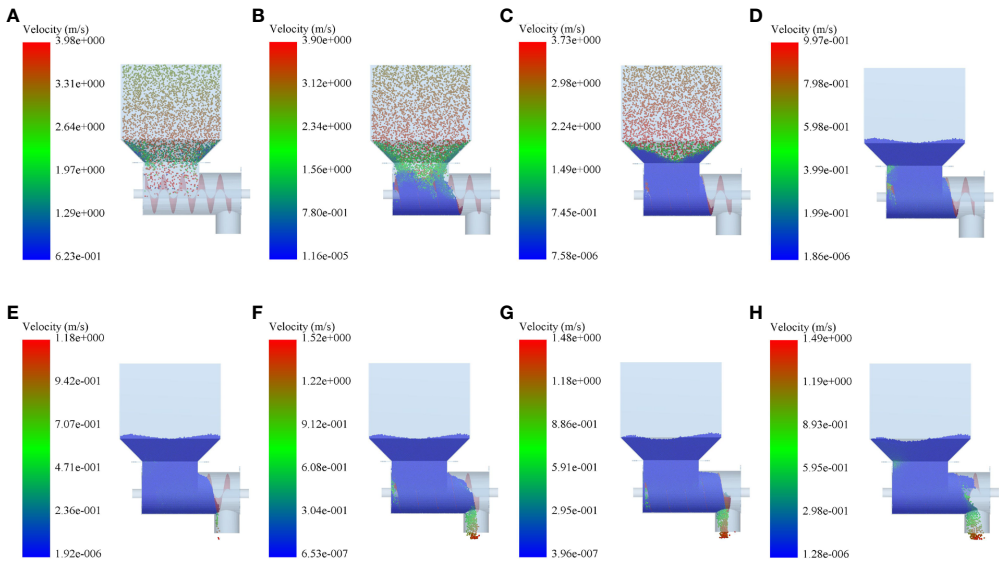


FIGURE 4
The process of simulating fertiliser discharge: **(A)** 0.1 s, **(B)** 0.6 s, **(C)** 1.1 s, **(D)** 2.2 s, **(E)** 2.8 s, **(F)** 3.7 s, **(G)** 5.3 s, **(H)** 6.0 s.

TABLE 1 The factor levels for the single-factor simulation fertilizing trial.

Horizontal	The diameter of the spiral blade A (mm)	The pitch B (mm)	The rotational speed of fertilizing shaft C (r/min)
1	80	50	15
2	85	55	25
3	90	60	35
4	95	65	45
5	100	70	55

4 Results and discussion

4.1 Single-factor test

4.1.1 The diameter of the spiral blade

The variable coefficient of fertilizing stability decreased with the increase in the diameter of the spiral blade when the pitch was 60 mm and the rotational speed of the fertilizing shaft was 35 r/min. The results of the simulated fertiliser discharge test are shown in Figure 5. A polynomial function was used to fit the diameter of the spiral blade and the variable coefficient of fertilizing stability, and a regression curve between the two was obtained. The analysis of the variance and significance test results on the regression mathematical model and coefficients showed that the diameter of the spiral blade had a significant effect on the variable coefficient of fertilizing stability ($P< 0.05$), the fit of the regression model $R^2 = 0.963$ indicated a good curve fit, and the diameter of the spiral blade was negatively correlated with the variable coefficient of fertilizing stability. The fitting equation is shown in Equation (29), and the standard errors of -5.01, 0.545, and 0.004 are 2.37, 0.53, and 0.003 respectively. It can be seen that the standard errors of each parameter in the model are small, indicating that the established mathematical model is accurate and reliable.

$$V = -5.01 + 0.545A - 0.004A^2 \tag{29}$$

4.1.2 The pitch

The variable coefficient of fertilizing stability increased with the pitch when the diameter of the spiral blade was 90 mm and the rotational speed of the fertilizing shaft was 35 r/min. The results of the fertiliser discharge simulation are shown in Figure 6. An exponential function was used to fit the pitch and the variable coefficient of fertilizing stability, and the regression curve between them was obtained. The analysis of variance and significance test results on the regression model and coefficients indicated that the effect of pitch on the variable coefficient of fertilizing stability was very significant ($P< 0.01$), and the fit of the regression mathematical model $R^2 = 0.956$ demonstrated a good curve fit and a positive correlation between the pitch and the variable coefficient of fertilizing stability, The fitting equation is shown in Equation (30), and the standard errors of 0.572, 0.029, and 0.00003 are 2.45, 0.08, and 0.0006 respectively. It can be seen that the standard errors of each parameter in the model are small, indicating that the mathematical model is accurate and reliable.

$$V = \exp 0.572 + 0.029B + 0.00003B^2 \tag{30}$$

4.1.3 The rotational speed of the fertilizing shaft

When the diameter of the spiral blade was 90 mm and the pitch was 60 mm, the variable coefficient of fertilizing stability increased with the rotational speed of the fertilizing shaft, and the results of the fertiliser discharge simulation are shown in Figure 7. The regression curve was obtained by fitting a power model to the rotational speed of the fertilizing shaft and the variable coefficient of fertilizing stability. Then, the regression mathematical model and coefficients were analysed by ANOVA (analysis of variance) and significance test. It can be seen that the influence of the rotational speed of the fertilizing shaft on the variable coefficient of fertilizing

TABLE 2 The factor level codes for multi-factor simulation fertilizing trials.

Normative variables	Natural variables		
	The diameter of the spiral blade A (mm)	The pitch B (mm)	The rotational speed of the fertilizing shaft C (r/min)
Upper asterisk arm γ (+1.682)	100	70	55
Upper level 1	96	66	47
Zero level 0	90	60	35
Lower level -1	84	54	23
Lower asterisk arm $-\gamma$ (-1.682)	80	50	15

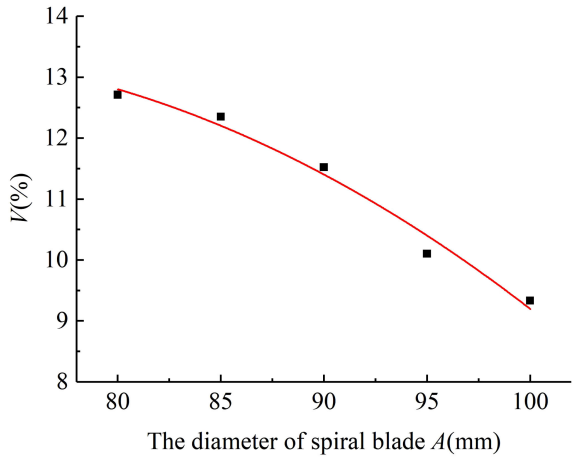


FIGURE 5
The regression fitting curve between the diameter of the spiral blade and the variable coefficient of fertilizing stability.

stability was very significant (P< 0.01), the fit of the regression model R^2 = 0.952 indicated a good fitting, and the rotational speed of the fertilizing shaft was positively correlated with the variable coefficient of fertilizing stability. The fitting equation is shown in Equation (31), and the standard errors of 7.463, 0.005, and 2.032 are 2.15, 0.01, and 0.73 respectively. It can be seen that the standard errors of each parameter in the model are small, indicating that the mathematical model has high accuracy and reliability.

$$V = 7.463 + 0.005C^{2.032} \tag{31}$$

To sum up, the results of the single-factor simulated fertilizing performance tests show that the diameter of the spiral blade, the pitch, and the rotational speed of the fertilizing shaft significantly affect the variable coefficient of fertilizing stability of the spiral fertiliser discharger. Especially, the diameter of the spiral blade is negatively correlated with the variable coefficient of fertilizing stability, and the pitch and the rotational speed of fertilizing shaft are positively correlated with the variable coefficient of fertilizing stability. These results verify the reliability of the theoretical analysis of the spiral fertiliser discharger in the previous study.

4.2 Quadratic regression orthogonal rotational combination design trial

According to factor level codes in Table 2, the multi-factor simulated fertilizing test scheme was determined by using the Design-Expert software. Specifically, Nine central points were used for error estimation in the simulated fertilizing test, a total of 23 sets of tests were conducted, and each set of tests was repeated three times. The corresponding variable coefficient of fertilizing stability was obtained through the calculation of formulas (26) - (28), and then a series of data analyses was conducted. The results of the multifactorial simulation fertilizing test are shown in Table 3.

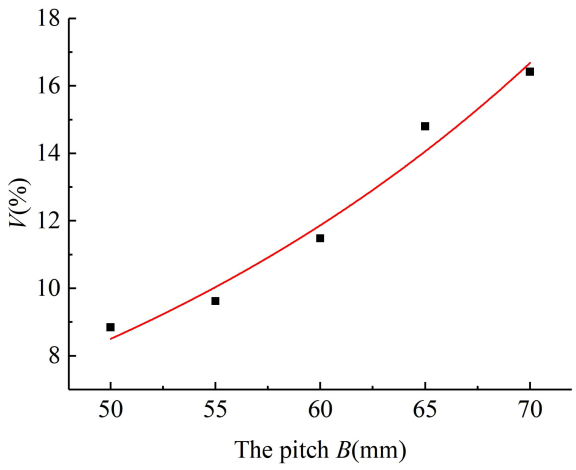


FIGURE 6
The regression fitting curve between the pitch and the variable coefficient of fertilizing stability.

10.3389/fpls.2023.1169091

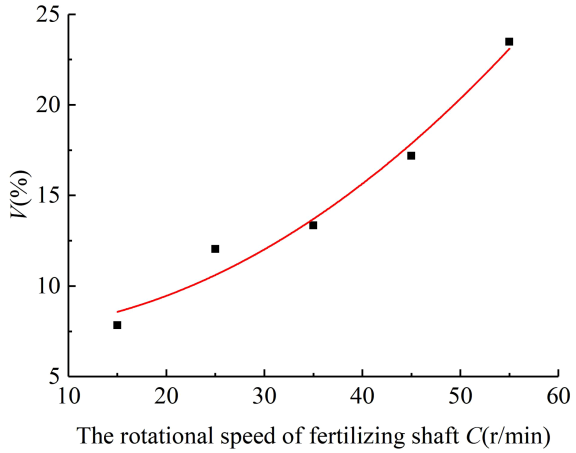


FIGURE 7
The regression fitting curve between the fertiliser discharge shaft speed and the variable coefficient of fertilizing stability.

TABLE 3 The scheme and results of the multi-factor simulation fertilizing trial.

Serial number	The diameter of the spiral blade A (mm)	The pitch B (mm)	The rotational speed of the fertilizing shaft C (r/min)	The variable coefficient of fertilizing stability V (%)
1	-1	-1	-1	8.76
2	1	-1	-1	7.31
3	-1	1	-1	9.59
4	1	1	-1	8.62
5	-1	-1	1	14.13
6	1	-1	1	12.65
7	-1	1	1	18.15
8	1	1	1	16.32
9	-1.682	0	0	12.26
10	1.682	0	0	10.09
11	0	-1.682	0	9.21
12	0	1.682	0	15.86
13	0	0	-1.682	7.12
14	0	0	1.682	21.87
15	0	0	0	11.09
16	0	0	0	11.24
17	0	0	0	12.27
18	0	0	0	11.41
19	0	0	0	12.67
20	0	0	0	13.15
21	0	0	0	11.28
22	0	0	0	12.63
23	0	0	0	11.45

4.2.1 Regression model, analysis of variance and significance test

Through multiple regression analysis of the above experimental results by using the Design-Expert software, a mathematical model can be obtained for the regression between the variable coefficient of fertilizing stability and the diameter of spiral blade, the pitch, and the rotational speed of the fertilizing shaft.

$$V = -74.77 + 2.13A - 0.28B - 0.49C + 0.0005AB - 0.0016AC + 0.009BC - 0.01A^2 + 0.0013B^2 + 0.005C^2 \quad (32)$$

Then, the ANOVA and significance test was conducted on the regression model, and the ANOVA results of the multi-factor simulation fertilizing test in Table 4 show that the diameter of the spiral blade A, the pitch B, the rotational speed of the fertilizing shaft C, and the quadratic term C^2 of the rotational speed of the fertilizing shaft have a significant effect on the variable coefficient of fertilizing stability V. The factors influencing the variable coefficient of fertilizing stability V in the order of priority are the rotational speed of the fertilizing shaft C, the pitch B, and the diameter of the spiral blade A. The p-value of this mathematical model is< 0.0001, indicating that the regression model of the variable coefficient of fertilizing stability is highly significant, and the misfit is not significant because the p-value of the misfit term is equal to 0.1125>0.05, indicating that the regression equation is well fitted, with the coefficient of determination R^2 = 0.9529 and the corrected coefficient of determination R^2_{adj} = 0.9204. Both values are close to 1, and the standard error is 0.33, indicating that the regression equation is highly reliable; meanwhile, the accuracy is 20.735, indicating that the regression mathematical model is highly accurate.

In the case that the regression model is significant and the misfit term is not significant, the above regression mathematical model is optimised to obtain a new equation by excluding the very insignificant term.

$$V = -77.02 + 2.09A - 0.08B - 0.64C + 0.0098BC - 0.01A^2 + 0.0039C^2 \quad (33)$$

The ANOVA results of the optimised regression mathematical model are shown in Table 5, which shows that the performance of the optimised regression model is improved. The coefficient of determination R^2 = 0.9524 and the corrected coefficient of determination R^2_{adj} = 0.9346 are close to 1, and the standard error decreased to 0.26, indicating that the reliability of the regression equation is enhanced; meanwhile, the accuracy is improved to 27.634, indicating the accuracy of the regression model is higher than that before the optimisation.

4.2.2 Response surface analysis of interaction effects

The interactions between the test factors were analysed to investigate their effects on the fertilizing performance of the spiral fertiliser discharger. Also, the optimal parameter region was further narrowed down. Figure 8 shows the response surface plot of the interaction effect on the variable coefficient of fertilizing stability.

Figure 8A shows the response surface of the interaction effect between the diameter of the spiral blade and the pitch on the variable coefficient of fertilizing stability when the rotational speed of the fertilizing shaft was 35 r/min. When the diameter of the spiral blade was 80-100 mm and the pitch was 50-70 mm, the variable coefficient of fertilizing stability decreased with the increase in the

TABLE 4 The analysis of the variance results of the multi-factor simulation fertilizing test.

Source of variance	Sum of squares	Degrees of freedom	F-value	p-value	Significance
Model	250.81	9	29.25	< 0.0001	**
A	6.44	1	6.76	0.022	*
B	32.33	1	33.94	< 0.0001	**
C	196.3	1	206.07	< 0.0001	**
AB	2.11E-03	1	2.22E-03	0.9632	
AC	0.099	1	0.1	0.7523	
BC	3.85	1	4.04	0.0656	
A^2	3	1	3.15	0.0991	
B^2	0.034	1	0.035	0.8539	
C^2	8.68	1	9.11	0.0099	**
Residuals	12.38	13			
Misfit	7.64	5	2.58	0.1125	
Error	4.74	8			
Total	263.19	22			

** indicates that the term is highly significant (p<0.01), * indicates that the term is significant (p<0.05). The same notation are used below.

TABLE 5 The ANOVA results of the optimised model for the multi-factor simulation fertilizing test.

Source of variance	Sum of squares	Degrees of freedom	F-value	p-value	Significance
Model	250.67	6	53.4	< 0.0001	**
A	6.44	1	8.23	0.0111	*
B	32.33	1	41.33	< 0.0001	**
C	196.3	1	250.89	< 0.0001	**
BC	3.85	1	4.92	0.0413	*
A^2	3.01	1	3.85	0.0675	
C^2	8.67	1	11.08	0.0043	**
Residuals	12.52	16			
Misfit	7.78	8	1.64	0.2501	
Error	4.74	8			
Total	263.19	22			

diameter of the spiral blade and increased with the increase in the pitch. The variable coefficient of fertilizing stability can be reduced by increasing the diameter of the spiral blade and decreasing the pitch. Figure 8B shows the response surface of the variable coefficient of fertilizing stability due to the interaction between the diameter of the spiral blade and the rotational speed of the fertilizing shaft when the pitch was 60 mm. It can be seen that the variable coefficient of fertilizing stability decreased with the increase in the diameter of the spiral blade and increased with the increase in the rotational speed of the fertilizing shaft when the diameter of the spiral blade was 80-100mm and the rotational speed of fertilizing shaft was 15-55 r/min. Therefore, the variable coefficient of fertilizing stability can be reduced by appropriately increasing the diameter of the spiral blade and decreasing the rotational speed of the fertilizing shaft. Figure 8C shows the response surface of the interaction between the pitch and the rotational speed of the fertilizing shaft on the variable coefficient of fertilizing stability when the diameter of the spiral blade was 90 mm, the pitch was 50-70 mm, and the rotational speed of the fertilizing shaft was 15-55 r/min. It can be seen that the variable coefficient of fertilizing stability increased with the increase in the pitch and the rotational speed of the fertilizing shaft. Thus, reducing the pitch and the rotational speed of the fertilizing shaft can reduce the variable coefficient of fertilizing stability. The results of the response surface analysis indicate that the variable coefficient of fertilizing stability increases with the pitch and the rotational speed of the fertilizing shaft.

Therefore, when the diameter of the spiral blade is 90-100 mm, the pitch is 50-60 mm, and the rotational speed of the fertilizing shaft is 15-35 r/min, the variable coefficient of fertilizing stability can be reduced by increasing the diameter of the spiral blade and decreasing the pitch and the rotational speed of the fertilizing shaft.

4.3 Parameters optimisation of the spiral fertiliser discharger

The genetic algorithm (GA) is a global search algorithm based on the principle of biological genetic evolution. It simulates the phenomenon of reproduction, mating, and mutation in natural selection and genetic processes. This algorithm can overcome the shortcomings of traditional nonlinear programming algorithms

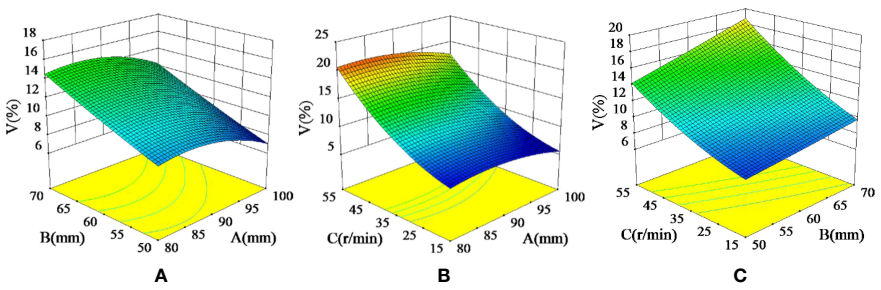


FIGURE 8
Interaction effects on the variable coefficient of fertilizing stability (*V*): **(A)** The response surface plot of the interaction between the diameter of the spiral blade and the pitch (*AB*) on the variable coefficient of fertilizing stability, **(B)** The response surface plot of the interaction between the diameter of the spiral blade and the rotational speed of the fertilizing shaft (*AC*) on the variable coefficient of fertilizing stability, **(C)** The response surface plot of the interaction between the pitch and the rotational speed of the fertilizing shaft (*BC*) on the variable coefficient of fertilizing stability.

that are easy to fall into local optima. It has the advantages of strong global optimization ability, robustness, and efficiency, and it has been widely used and developed (Yang et al., 2019; Bahiraei et al., 2020; Chen et al., 2021). Therefore, this study adopted the genetic algorithm to find the optimal parameter combination of the spiral fertiliser discharger. The specific process of applying the genetic algorithm is shown in Figure 9.

To improve the fertilizing stability of the spiral fertiliser discharger, under the premise of ensuring the safe and reliable operation of the fertiliser discharger, the variable coefficient of fertilizing stability should be as small as possible. Therefore, the objective optimization function of the variable coefficient of fertilizing stability is established, as shown in formula (34). According to the variables in the objective optimization function, the parameters to be optimized include the diameter of the spiral blade A, the pitch B, and the rotational speed of the fertiliser shaft C. Therefore, the parameter optimization problem of the spiral fertiliser discharger can be defined as finding an optimal vector X, as shown in formula (35). Meanwhile, according to the analysis results of the response surface, the variation ranges of the parameters to be optimized of the spiral fertiliser discharger are determined, and the corresponding constraint conditions are established, as shown in formula (36).

$$V(x) = \min V(A, B, C) \tag{34}$$

$$X = [A \ B \ C]^T \tag{35}$$

$$s.t \begin{cases} 90 \le A \le 100 \\ 50 \le B \le 60 \\ 15 \le C \le 35 \end{cases} \tag{36}$$

In this study, the parameter optimization of the spiral fertiliser discharger is a nonlinear constrained optimization problem, which can be effectively solved by the genetic algorithm. The specific operation steps are as follows:

(1) Parameter coding. The genetic algorithm uses binary coding. There are three independent variables, namely, the spiral blade diameter A, the pitch B, and the fertilizer shaft speed C. Each independent variable is set to 6 genes, a total of 18 genes.

(2) The initial population is generated, and the population number is set to 50.

(3) Fitness evaluation function. After weighting the objective optimization function, the fitness evaluation function is defined as $F(x) = 1/V(x)$.

(4) Genetic operation, including selection, crossover, and mutation. Copying is the basic operator of the genetic algorithm. It reproduces excellent chromosomes in the next generation of new groups, consistent with the natural selection principle of ‘ survival of the fittest’. Whether the chromosomes are copied is determined according to the size of their fitness. The larger the fitness is, the more the chromosomes are copied. Meanwhile, the chromosomes with a smaller fitness is eliminated so that the total number of chromosomes in the new group is the same as that of the original group. In this paper, the selection probability is set to 0.9, and the population is set to 50, so a total of 50 * 0.9 = 40 chromosomes are saved at each cycle. In addition, 50 * (1-0.9) = 10 optimal individuals are copied to ensure the stability of the population in the iterative process. The number of chromosomes is still 40 + 10 = 50 so that the next round of iteration is conducted. The

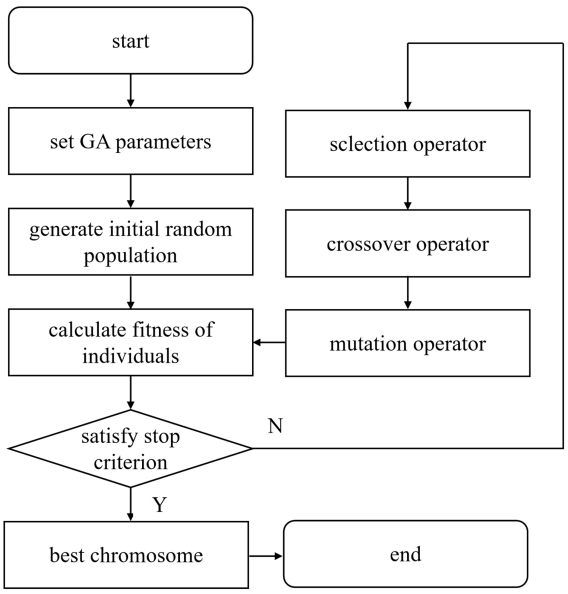


FIGURE 9
The flow chart of applying the genetic algorithm,.

chromosomes are selected by the roulette method. In the genetic algorithm, exchange is the main way to generate new chromosomes. It imitates the principle of hybridization in biology and exchanges some genes of the two chromosomes. The chromosomes that perform the exchange are randomly selected. First, the exchange probability is determined to be 0.75. Then, the above-mentioned roulette selection method is adopted to select the exchanged chromosomes according to the fitness size, and pairwise exchanges are performed in turn. Mutation is another method for generating new chromosomes in genetic algorithms. It mutates a character of a chromosome, such as changing the original gene 0 into 1 or changing the original gene 1 into 0. The selection of mutant chromosomes and the determination of mutation location are all generated by random methods. First, the number of chromosomes that need to be mutated is determined, e.g., the mutation probability is 0.02, and the number of populations is 50. Therefore, 50 * 0.02 = 1 chromosome is randomly mutated in each iteration, and then a position is randomly selected in the gene fragment of the chromosome to change the original gene value. The group P (t) is promoted to the next generation group P (t + 1) by three operations with guessing properties.

(5) Termination condition judgment. The maximum evolutionary algebra method is taken as the stop rule of the program. In this study, the iterative algebra 100 is set as the termination condition, i.e., the genetic operation process is terminated when 100 consecutive generations do not reproduce a new generation, and the maximum fitness individual obtained in the previous calculation process is taken as the optimal solution. The corresponding iterative convergence curve is presented in Supplementary Figure 2.

Through multiple iterative calculations, the optimal parameter combination of the spiral fertiliser discharger was finally obtained, i.e., 98.44 mm for the diameter of the spiral blade, 54.8 mm for the pitch, and 24.43 r/min for the rotational speed of the fertilizing shaft. In this case, the variable coefficient of fertilizing stability was 6.51%, which could meet the fertilization agronomic requirements in the national standard "Technical Specification for Quality Evaluation of Fertilization Machinery" (the total variable coefficient of fertilizing stability).

4.4 Experimental validation

4.4.1 The flowability verification test of the granular fertiliser

To further validate the reliability of the results of the discrete element simulation fertilizing test and the optimal parameters of the spiral fertiliser discharger, the flowability verification test of the granular fertiliser was conducted by using a combination of bench and simulation tests, the mass flow rate was taken as the evaluation index, and the measured and simulated values of the mass flow rate were compared as shown in Figure 10. In the fertilizing test, an appropriate amount of compound fertiliser granules are first added to the fertiliser box, and the optimal parameters of the fertiliser discharger are set; then, the power is activated, the speed-controlled motor drives the fertilizing screw to discharge the fertiliser, and the granular fertiliser drops into the collection box below the fertilizing pip. The entire process of the granular fertiliser was recorded by a camera, the flow time and the mass of the granular fertiliser in the collection box were recorded, and the mass flow rate of the granular fertiliser was calculated. The test process was repeated five times. Subsequently, the digital model of the spiral fertiliser discharger and the discrete element model of the granular fertiliser were imported into the EDEM software, and the simulated fertilizing test was

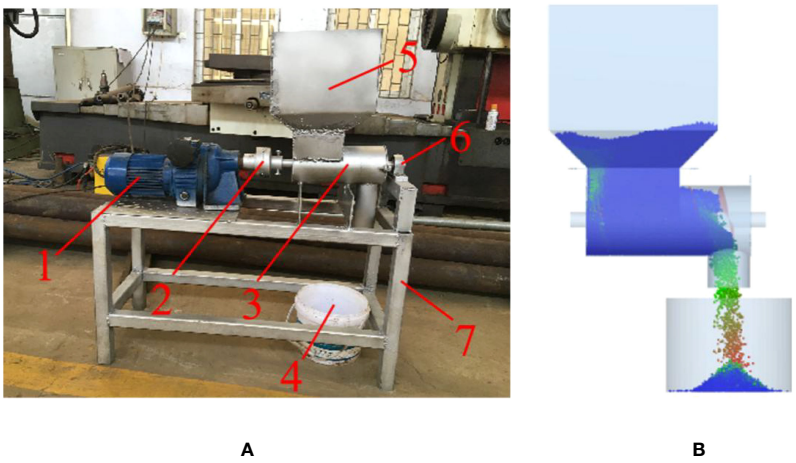


FIGURE 10
The process of granular fertiliser flow verification test: **(A)** the bench test, **(B)** the simulation test. 1. JWB-X0.37-8D type electrodeless transmission 2. coupling 3. fertilising box 4. gathering barrel 5. fertiliser box 6. bearing seat, 7. frame.

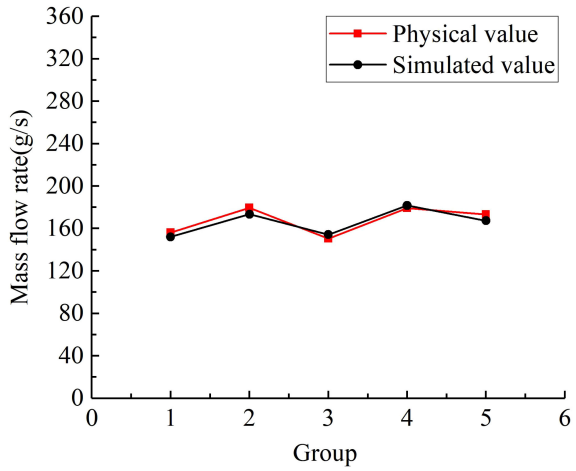


FIGURE 11
The comparison results of flowability validation test on the granular fertiliser.

carried out under the same conditions as the bench test. After this test, the mass flow rate of the granular fertiliser was derived, the average relative error between the physical value and the simulated value of the mass flow rate was calculated, and the test results were compared and analysed.

Figure 11 shows the comparison results of the flowability verification test of the granular fertiliser. The average relative error of the fertiliser flow rate between the bench and the simulated fertilization was calculated to be 2.64%. The small error indicates that there is no significant difference between the physical and simulated values of the granular fertiliser flow rate. Meanwhile, the results demonstrate the high accuracy of the regression mathematical model developed in this study and the optimal parameter combination determined for the fertiliser discharger.

4.4.2 The applicability verification test of different fertilisers

To verify the applicability of different fertilisers to the spiral fertiliser discharger optimized in this study, three types of fertilisers (urea, potassium chloride, and compound fertiliser) commonly used in mango orchards are taken as the research objects, as shown in Figure 12. The physical characteristic parameters of the three fertilisers are measured, and the results are presented in Table 6. Taking the mass flow rate and the variable coefficient of fertilizing stability as response indicators, the optimal parameter combination of the fertiliser discharger determined above was exploited to conduct different fertiliser applicability verification tests. Five groups of tests were carried out for each fertiliser, and each group of tests was repeated three times. The mass flow rate and the variable coefficient of fertilizing stability of the three fertilisers were calculated by formulas (26)-(28). Finally, the average relative error was calculated, and the test results were compared and analysed.

Figure 13 shows the comparison results of the fertiliser mass flow rate. According to statistical analysis, the average mass flow rate of urea particles, potassium chloride particles, and compound fertiliser particles is 172.35 g/s, 158.42 g/s, and 173.11 g/s, respectively. When the forward speed of the fertiliser applicator is 2 km/h and the plant spacing of mango trees is 1.82 m, the three types of fertilisers under the optimal parameter combination of the spiral fertiliser discharger can meet the fertilization requirements of 0.3-0.5 kg/plant for mango trees. Figure 14 shows the comparison results of different fertilizing stability tests. The average variable coefficient of fertilizing stability of urea granules, potassium chloride

FIGURE 12
Three types of fertilisers commonly used in mango orchards: **(A)** urea, **(B)** potassium chloride, **(C)** compound fertiliser.

TABLE 6 The physical properties parameters of the three types of fertilisers.

Type	Average equivalent diameter (mm)	Spherical rate (%)	Density (g/cm^3)	Water content (%)	Angle of repose (°)
Urea	3.581	91.08%	0.938	2.53%	33.7
Potassium chloride	3.463	87.41%	0.965	2.71%	35.9
Compound fertiliser	3.829	92.13%	0.913	3.28%	32.5

granules, and compound fertiliser granules is 6.85%, 7.16%, and 6.73%, respectively. It can be seen that the average variable coefficient of fertilizing stability of the three types of fertilisers is small, which meets the agronomic fertilization requirements of the national standard "Technical Specification for Quality Evaluation of Fertilization Machinery" (the total variable coefficient of fertilizing stability $\leq 7.8\,\%$). The applicability test results of different types of fertilisers further verified the reliability and accuracy of the discrete element simulation test results and the optimal parameter combination determined for the spiral fertiliser discharger. The designed small spiral fertiliser discharger can be used in the fertilization operation of mango orchards.

Through theoretical analysis, the error in the fluidity verification test of granular fertiliser may be due to the difference between the established discrete element model of granular fertiliser and the actual fertiliser. Through research, it is found that the average mass flow rate of potassium chloride particles in the applicability verification test of different fertilisers is smaller than that of urea and compound fertiliser particles, and the average variable coefficient of fertilizing stability of potassium chloride particles is larger than that of urea and compound fertiliser particles. It may be because the spherical rate of potassium chloride particles (87.41%) is smaller than that of urea (91.08%) and compound fertiliser (92.13%) particles, which leads to a slightly smaller mass flow rate of potassium chloride particles and a slightly larger coefficient of variation of fertiliser discharge stability. In addition, there are still some shortcomings in this study, such as considering only one moisture content, granular fertilisers with different moisture content may have different physical and mechanical properties and fertilizing performance, which will be further studied in the future. At the same time, the uniformity of fertiliser discharge is also an important index of the fertilizing performance of the spiral fertiliser discharger. The next step will be to build a relevant test platform and develop a mango orchard fertilization machine to conduct an experimental study on the uniformity of fertiliser discharge.

5 Conclusion

Based on the agronomic fertilization characteristics of mango orchards, a small spiral fertiliser discharger was designed in this study. First, the fertilizing performance test and parameter optimization of the spiral fertiliser discharger were conducted by combining the bench test and simulation test. Then, by taking the variable coefficient of fertilizing stability as the response value, the single-factor test and the quadratic regression orthogonal rotation combination design test were designed by the Design Expert software. Subsequently, the parameters significantly affecting the variable coefficient of fertilizing stability were explored, and the second-order regression mathematical model between the variable coefficient of fertilizing stability and the significant parameters was

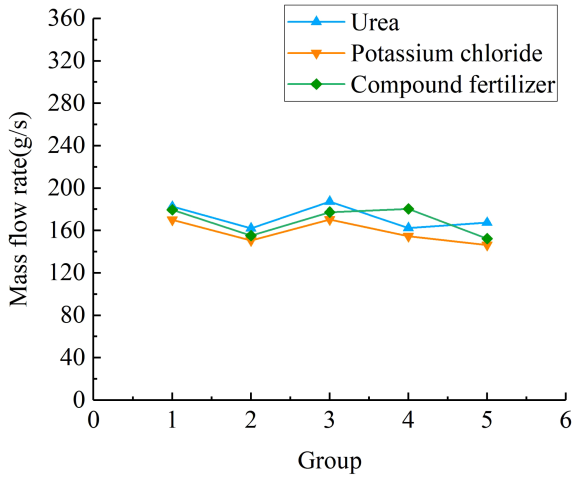


FIGURE 13
The comparison of mass flow rate results for different fertilisers.

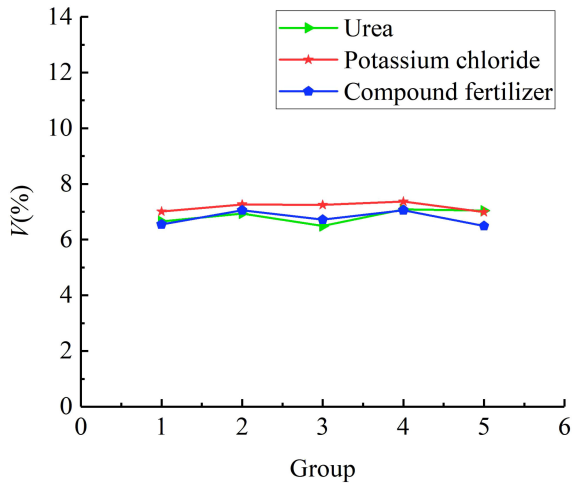


FIGURE 14
The comparison of fertilizing stability test results for different fertilisers.

established. Finally, the optimal parameter combination of the fertiliser discharger was determined by the genetic algorithm. The accuracy of the simulation test results and the optimal parameter combination of the spiral fertiliser discharger was further verified by experiments. Based on the experimental results, the following conclusions were drawn:

(1) The single-factor test results showed that the diameter of the spiral blade, the pitch, and the rotational speed of the fertilizing shaft have significant effects on the variable coefficient of fertilizing stability, and the variable coefficient of fertilizing stability decreases with the increase in the diameter of the spiral blade and increases with the pitch and the rotational speed of the fertilizing shaft.

(2) The quadratic regression orthogonal rotation combination design test results showed that the established regression model of the variation coefficient of the fertiliser discharge stability has good reliability and precision. Meanwhile, the optimal parameter combination of the spiral fertiliser discharger optimized by the genetic algorithm is 98.44 mm for the diameter of the spiral blade, 54.8 mm for the pitch, and 24.43 r/min for the rotational speed of the fertilizing shaft. The variable coefficient of fertilizing stability is 6.51%, which meets the fertilization agronomic requirements of the national standard 'Technical Specification for Quality Evaluation of Fertilization Machinery'.

(3) Under the optimal parameter combination of the fertiliser discharger, the verification test results on flowability and applicability of granular fertiliser show that the average relative error of the mass flow rate of the bench test and the simulated fertiliser discharge test is 2.64%, and the error is small. Meanwhile, the average mass flow rate of three types of fertilisers commonly used in mango orchards can meet the fertilization requirements of mango trees; the variable coefficient of fertilizing stability of the three types of fertilisers is small, which meets the agronomic fertilization requirements of mango orchards.

The above test results verify the reliability and accuracy of the discrete element simulation fertilizer test results and the optimal parameter combination of the determined fertilizer discharger. Therefore, the small spiral fertilizer discharger designed in this paper can be used for fertilization in mango orchards, which can effectively improve the accuracy of fertilization in mango orchards, the stability of fertilizer discharge and the utilization rate of fertilizer, thereby improving the yield and quality of mango, and playing a certain role in environmental protection. The results of this study provide basic data and research methods for the development of mango orchard fertilization machinery and other orchard related fertilizer performance tests.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding authors.

Author contributions

LZ: Methodology, Data curation, Software, Validation, Writing – original draft. HZ: Conceptualization, Supervision, Funding

acquisition. LX: Writing – review & editing. WY: Formal analysis. MS: Visualization. JZ: Investigation. ZX: Resources. All authors contributed to the article and approved the submitted version.

Funding

This project has received funding from the National Key Research and Development Plan of China (Grant No. 2017YFD0202102).

Acknowledgments

'The authors would like to thank Dr. Weidong Yuan, Nanjing Forestry University, for linguistic assistance during the preparation of this manuscript. Meanwhile we warmly appreciate associate professor Zhong Xue, South Subtropical Crops Research Institute, Chinese Academy of Tropical Agricultural Sciences, many thanks for providing the testing equipment and site.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2023.1169091/full#supplementary-material

References

Bahiraei, M., Nazari, S., Moayedi, H., and Safarzadeh, H. (2020). Using neural network optimized by imperialist competition method and genetic algorithm to predict water productivity of a nanofluid-based solar still equipped with thermoelectric modules. *Powder Technol.* 366, 571–586. doi: 10.1016/j.powtec.2020.02.055

Chen, Z., Wang, G., Xue, D., and Cui, D. (2021). Simulation and optimization of crushing chamber of gyratory crusher based on the DEM and GA. *Powder Technol.* 384, 36–50. doi: 10.1016/j.powtec.2021.02.003

Coetzee, C. J., and Lombard, S. G. (2011). Discrete element method modelling of a centrifugal fertiliser spreader. *Biosyst. Eng.* 109 (4), 308–325. doi: 10.1016/j.biosystemseng.2011.04.011

Dun, G. Q., Liu, W. H., Du, J. X., Zhou, C., Mao, N., and Ji, W. Y. (2022). Optimization design and experiment of double spiral fertiliser discharger with arc groove. *J. Agric. Machinery* 53 (10), 118–125+174.

Dun, G. Q., Liu Yang, Y. Z., Chen, H. T., Du, J. X., and Zhang, J. T. (2018). Simulation and experiment of EDEM-based controlled-position layered fertiliser opener operation process. *J. Hunan Agric. Univ. (Natural Sci. Edition)* 44 (01), 95–100. doi: 10.13331/j.cnki.jhau.2018.01.018

He, C. C., Feng, H. D., Wei, Z. Y., Hou, X. W., and Chen, Y. Y. (2018). Investigation of current situation of fertilization and analysis of soil nutrients of mango orchard in HainanIsland. *Chin. J. Trop. Crop* 12, 2336–2348. doi: 10.3969/j.issn.1000-2561.2018.12.002

Huang, S., Long, D., and He, Y. (2018). Research progress and prospect of mango fertilization in china. *Agric. Technol.* 38 (16), 7–8. doi: 10.11974/nyyjs.20180833004

Jia, H. L., Tan, H. W., Wen, X. Y., Wang, G., Yuan, H. F., and Huang, D. Y. (2022). Design and experiment of a pneumatic collection and discharge type precision blending and fertiliser application device. *J. Agric. Machinery* 53 (S2), 109–119+203. doi: 10.6041/j.issn.1000-1298.2022.S2.013

Liao, Q. X., Chen, Y., Zhang, Q. S., Wang, L., Ling, J. X., and Du, W. B. (2023). Design and experiment of a side deep cavity fertiliser application device for oilseed rape. *J. Agric. Machinery* 54 (2), 1–13. [2022-12-06]. doi: 10.604/j.issn.1000-1298.2023.02.004

Liu, J. S., Gao, C. Q., Nie, Y. J., Yang, B., Ge, R. Y., and Xu, Z. H. (2020). Numerical simulation of fertilizer shunt-plate with uniformity based on EDEM software. *Comput. Electron. Agric.* 178, 105737. doi: 10.1016/j.compag.2020.105737

Liu, X. D., Ding, Y. C., Shu, C. X., Liu, W. P., Wang, K. Y., Du, C. Q., et al. (2020). Design and experiment of spiral disturbance cone centrifugal fertilizer apparatus. *Trans. CSAE* 36 (2), 40–49. doi: 10.11975/j.issn.1002-6819.2020.02.006

Liu, W., Si, R., Fan, J., and Lin, D. (2021). Effect of different fertilization patterns on yield and quality of mango. *J. Trop. Crops* 42 (03), 761–768. doi: 10.3969/j.issn.1000-2561.2021.03.022

Liu, X., Wang, X., Chen, L., Zhang, C., Liu, W., and Ding, Y. (2021). Design and experiments of layered and quantitative fertilization device for rapeseed seeder [J]. *Trans. CSAE* 37 (5), 1–10. doi: 10.11975/j.issn.1002-6819.2021.05.001

Liu, M. C., Zhao, Q. J., Han, S. Q., Song, Z. H., Li, F. D., and Yan, Y. F. (2022). Design and experiment of self-propelled variable ratio fertiliser directional spreader for mulberry gardens. *J. Agric. Machinery* 53 (S2), 120–130+140. doi: 10.6041/j.issn.1000-1298.2022.S2.014

Piri, J., and Mohapatra, P. (2021). An analytical study of modified multi-objective harris hawk optimizer towards medical data feature selection. *Comput. Biol. Med.* 135, 104558. doi: 10.1016/j.compbiomed.2021.104558

Qu, H., Peng, J., Zhou, L., Huang, J., Zhu, X., and Zeng., Y. (2021). Analysis of fertiliser application status and reduction potential of orchards in baise mango main production area. *J. South. Agric.* 52 (12), 3375–3381. doi: 10.3969/j.issn.2095-1191.2021.12.021

Santana, J. C. C., Araújo, S. A., Alves, W. A., Belan, P. A., Jiangang, L., Jianchu, C., et al. (2018). Optimization of vacuum cooling treatment of postharvest broccoli using response surface methodology combined with genetic algorithm technique. *Comput. Electron. Agric.* 144, 209–215. doi: 10.1016/j.compag.2017.12.010

Sayed, G. I., Soliman, M. M., and Hassanien, A. E. (2021). A novel melanoma prediction model for imbalanced data using optimized SqueezeNet by bald eagle search optimization. *Comput. Biol. Med.* 136, 104712. doi: 10.1016/j.compbiomed.2021.104712

Song, S., Duan, J., Zou, X., Yang, Z., Ou, Z., and Wang, B. (2020). Parameter optimization and test of variable fertilizer apparatus based on root distribution pattern of bananas. *Trans. Chin. Soc Agric. Eng.* 36, 11–18. doi: 10.11975/j.issn.1002-6819.2020.06.002

Sun, J., Chen, H., Duan, J., Liu, Z., and Zhu, Q. (2020). Mechanical properties of the grooved-wheel drilling particles under multivariate interaction influenced based on 3D printing and EDEM simulation. *Comput. Electron. Agric.* 172, 105329. doi: 10.1016/j.compag.2020.105329

Tan, H., Xu, L., Ma, S., Niu, C., Yan, C., and Shen, C. (2022). Design and test of a scraper-type organic fertiliser strip spreading and rotary tillage hybrid fertiliser applicator. *J. Agric. Machinery* 1–24, 2022– 12-06. doi: 10.6041/j.issn.1000-1298.2022.11.016

Thaper, R. K. (2014) *Effect of vane shape and fertilizer product on spread uniformity using a dual-disc spinner spreader (Doctoral dissertation, auburn university)*. Available at: https://www.proquest.com/dissertations-theses/effect-vane-shape-fertilizer-product-on-spread/docview/2778644784/se-2?accountid=44047.

Thawkar, S., Sharma, S., and Khanna, M. (2021). Breast cancer prediction using a hybrid method based on butterfly optimization algorithm and ant lion optimizer. *Comput. Biol. Med.* 139, 104968. doi: 10.1016/j.compbiomed.2021.104968

Van Liedekerke, P., Tijskens, E., Dintwa, E., Rioual, F., Vangeyte, J., and Ramon, H. (2009). DEM simulations of the particle flow on a centrifugal fertilizer spreader. *Powder Technol.* 190 (3), 348–360. doi: 10.1016/j.powtec.2008.08.018

Wei, P. (2021). How the largest mango base in china was built–baisai's observation of developing special industrial clusters based on resource advantages. *Farmers' Friend* 08), 2–7.

Wei, Z., and Aiping, G. (2021). Research progress of mango nutrition and fertilization technology. *China Forestry Specialties* 02), 98–99+104. doi: 10.1016/j.powtec.2008.08.018

Wei, Z. Y., and Gao, A. P. (2021). Current status and outlook of the application of discrete element method in agricultural engineering research. *J. Agric. Machinery* 52 (04), 1–20. doi: 10.13268/j.cnki.fbsic.2021.02.040

Xing, J., Zhao, H., Chen, H., Deng, R., and Xiao, L. (2023). Boosting whale optimizer with quasi-oppositional learning and gaussian barebone for feature selection and COVID-19 image segmentation. *J. Bionic Eng.* 20 (2), 797–818. doi: 10.1007/s42235-022-00297-8

Yang, L., Wang, J., Sun, X., and Xu, M. (2019). Multi-objective optimization design of spiral demister with punched holes by combining response surface method and genetic algorithm. *Powder Technol.* 355, 106–118. doi: 10.1016/j.powtec.2019.07.030

Yuan, Q., Xu, L., Ma, S., Niu, C., Yan, C., and Zhao, S. (2021). The effect of paddle configurations on particle mixing in a soil-fertilizer continuous mixing device. *Powder Technol.* 391, 292–300. doi: 10.1016/j.powtec.2021.06.022

Zhao, L., Zhou, H. P., Xu, L. Y., Song, S. Y., Zhang, C., and Yu, Q. X. (2022). Design and experiment of a powdered organic fertiliser strip application and discharge device. *J. Agric. Machinery* 53 (10), 98–107. doi: 10.1016/j.powtec.2021.09.065

Zhao, L., Zhou, H., Xu, L., Song, S., Zhang, C., and Yu, Q. (2022). Parameter calibration of coconut bran substrate simulation model based on discrete element and response surface methodology. *Powder Technol.* 395, 183–194. doi: 10.1016/j.powtec.2021.09.065

Zhu, X. H., Li, X. D., Gao, X., Tan, C., and Deng, H. T. (2022). Design and test of a self-propelled organic fertiliser spreader for orchards. *J. Agric. Machinery* 53 (05), 136–146. doi: 10.6041/j.issn.1000-1298.2022.05.014

Zhu, Q., Wu, G., Chen, L., Zhao, C., and Meng, Z. (2018). Influences of structure parameters of straight flute wheel on fertilizing performance of fertilizer apparatus. *Trans. CSAE* 34 (18), 12–20. doi: 10.11975/j.issn.1002-6819.2018.18.002

10.3389/fpls.2023.1169091

Glossary

DEM	discrete element method
GA	genetic algorithm
Q	the fertilizing amount
D	the diameter of the spiral blade
n	the rotational speed of the fertilizing shaft
S	the pitch
v	the forward speed of the fertilizer applicator
f	the fertilizer amount per mango tree
R	the average spacing length of mango trees
K_1	the ratio coefficient between the diameter of the spiral blade and pitch
J	the fertiliser composite characteristic coefficient
K	the fertiliser synthesis factor
m	the mass of fertiliser
r	the radius of the spiral
n_{max}	the critical speed of fertilizer shaft
g	the gravitational acceleration
D_1	the equivalent diameter
L	the length
W	the width
T	the thickness
Φ	the spherical rate
F_n	the normal force
F_τ	the tangential force
F_d^n	the normal damping force
F_d^τ	the tangential damping force
E^*	the equivalent modulus of elasticity
R^*	the equivalent radius
E_i, E_j	the modulus of elasticity
v_i, v_j	the poisson's ratio
R_i, R_j	the radius of contact particle
S_τ	the tangential stiffness
S_n	the normal stiffness
G^*	the equivalent shear modulus
m^*	the equivalent mass
E	the recovery factor
$v_n^{\overline{rel}}$	the normal component of relative velocity
$v_\tau^{\overline{rel}}$	the tangential component relative velocity
T_i	the moments

(Continued)

Continued

Q_n	the fertilizing amount from a particular trial collected in the fertiliser monitoring area
x	the number of trials
Y	the total number of data points collected in the fertilizer monitoring area
V	the variable coefficient of fertilizing stability
A	the diameter of the spiral blade
B	the pitch
C	the rotational speed of fertilizing shaft
AB	the interaction of the diameter of the spiral blades and the pitch
AC	the interaction of the diameter of the spiral blades and the rotational speed of fertilizing shaft
BC	the interaction of the pitch and the rotational speed of fertilizing shaft
A^2	the quadratic term of the diameter of the spiral blades
B^2	the quadratic term of the pitch
C^2	the quadratic term of the rotational speed of fertilizing shaft
λ	the fertilizer density
ε	the coefficient of inclined conveying
ϕ	the filling factor
ω_{max}	the critical angular velocity of the fertilizer shaft
δ_τ	the tangential overlap
δ_n	the normal overlap
β	the damping ratio
μ_s	the static friction factor
μ_r	the rolling friction factor
ω_i	the unit angular velocity vector of the granular fertiliser at the point of contact
δ	the average fertilizing amount for all data in the fertiliser monitoring area
σ	the standard deviation of the fertilizing amount for all data in the fertiliser monitoring area

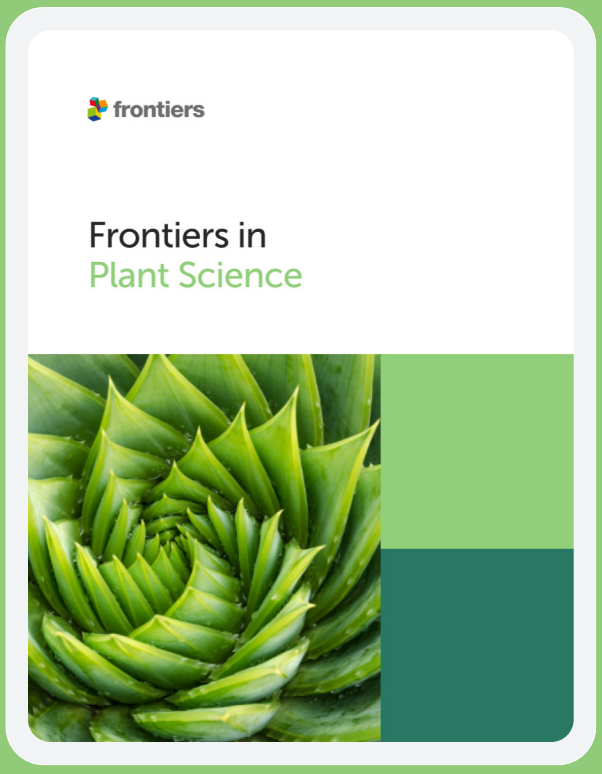
Frontiers in
Plant Science

Cultivates the science of plant biology and its applications

The most cited plant science journal, which advances our understanding of plant biology for sustainable food security, functional ecosystems and human health.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact